

Uniscale Multi-view Registration Using Double Dog-Leg Method

Chao-I Chen^a, Dusty Sargent^b, Chang-Ming Tsai^a, Yuan-Fang Wang^a and Dan Koppel^b

^aDept. of Computer Science, University of California, Santa Barbara, CA, USA 93106

^bSTI Medical Systems, 733 Bishop Street, Honolulu, HI, USA 96813

ABSTRACT

3D computer models of body anatomy can have many uses in medical research and clinical practices. This paper describes a robust method that uses videos of body anatomy to construct multiple, partial 3D structures and then fuse them to form a larger, more complete computer model using the structure-from-motion framework. We employ the Double Dog-Leg (DDL) method, a trust-region based nonlinear optimization method, to jointly optimize the camera motion parameters (rotation and translation) and determine a global scale that all partial 3D structures should agree upon. These optimized motion parameters are used for constructing local structures, and the global scale is essential for multi-view registration after all these partial structures are built. In order to provide a good initial guess of the camera movement parameters and outlier free 2D point correspondences for DDL, we also propose a two-stage scheme where multi-RANSAC with a normalized eight-point algorithm is first performed and then a few iterations of an over-determined five-point algorithm is used to polish the results. Our experimental results using colonoscopy video show that the proposed scheme always produces more accurate outputs than the standard RANSAC scheme. Furthermore, since we have obtained many reliable point correspondences, time-consuming and error-prone registration methods like the iterative closest points (ICP) based algorithms can be replaced by a simple rigid-body transformation solver when merging partial structures into a larger model.

Keywords: Modeling, registration, visualization

1. INTRODUCTION

3D computer models of body anatomy have been widely used for medical research and clinical practices. While sensor modalities, such as CT and MR, provide 3D data, they are not part of the standard protocols of many clinic procedures where video cameras are used instead. In such medical procedures, using computer vision structure-from-motion framework to construct 3D anatomical models from videos is a promising alternative, as the model provides both the appearance and structure information of the body anatomy.¹

In the model building process, many partial, local 3D structures are first constructed. A larger, more complete model can then be built by fusing all these partial structures. To merge structures using the above scheme, two main problems need to be solved: global scale determination and partial structure registration. Theoretically speaking, based on video data alone, a partial 3D structure can only be faithfully constructed up to an undetermined scale. This is because the exact physical distance of the camera movement is unknown, and the same video sequence can be generated by, for example, an object that is twice as large, moving twice as fast, and at twice the distance to the camera. Therefore, it is important to establish one uniform global scale that every local, partial 3D structure can agree upon before these partial models are registered together.

We propose to use a nonlinear optimization method, the Double Dog-Leg (DDL)² method, to determine this consistent global scale and optimize all the camera motion parameters jointly at the same time. Once a consistent global scale is determined, instead of applying the commonly used iterative closest points (ICP) algorithm³ for model registration which can be time-consuming and error-prone, we solve a simple rigid-body transformation problem with the help of reliable ground control points.

Further author information: (Send correspondence to Chao-I Chen)
Chao-I Chen: E-mail: chaoichen@gmail.com, Telephone: 1 805 280 9655

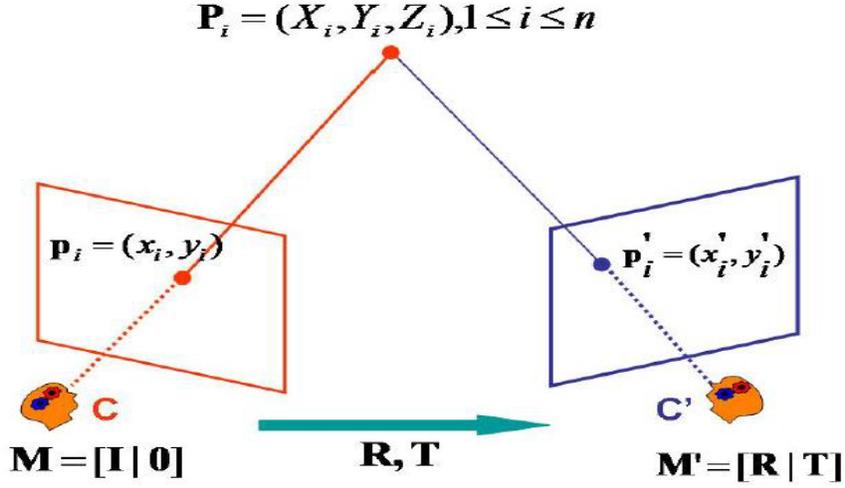


Figure 1. Two-view imaging configuration

2. METHOD

2.1 Bundle Adjustment

In this section, we first review how the structure and motion parameters are inferred from a two-view imaging configuration. Figure 1 shows one example of such a configuration where a number of prominent surface feature points $P_i, 1 \leq i \leq n$ are observed. The pixel locations observed in the first image frame are unprimed lower case p_i while those observed in the second image frame are primed lower case p'_i . The movement of the camera between two shots are denoted as R (rotation) and T (translation). We use a shorthand P (without a subscript) to denote the collection of P_i , or $P = [P_1, P_2, \dots, P_n]$. Similar shorthand for other symbols will be used as well.

In the analysis of structure and motion, prominent 2D feature points (p and p') and their correspondences are first established by using a 2D feature detection, description and matching scheme such as.⁴ Based on these data, camera motion (R and T) and 3D structure (P) can be computed. We will discuss in detail how we eliminate outliers and compute reliable camera motion parameters in section 2.3. The key observation is that *if the values of these unknown variables are inferred correctly, one would reasonably expect that the structure and motion parameters together should explain the image observation.* That is, if we position the camera pair according to the correct motion parameters (R and T) and place 3D features in space according to the correct P , image projections should conform with p_i in the first image and p'_i in the second image (assuming that camera has been calibrated and image distortion has been corrected beforehand⁵). Equation 1 explains the above relation. The best P , R and T are the choices that provides the smallest projection error in the two images.

$$P, R, T = \underset{P, R, T}{\operatorname{argmin}} \sum_{i=1}^n (d(p_i, P_{rj}(P_i)) + d(p'_i, P'_{rj}(P_i)))^2 \quad (1)$$

Function $d(\bullet, \bullet)$ denotes the 2D Euclidean distance between the two arguments. P_{rj} and P'_{rj} are the projection operators for the first and the second cameras respectively. They can be shown to be equation 2 and equation 3 respectively.⁶

$$P_{rj}(P_i) = M \begin{bmatrix} P_i \\ 1 \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} P_i \\ 1 \end{bmatrix} \quad (2)$$

$$P'_{rj}(P_i) = M' \begin{bmatrix} P_i \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} P_i \\ 1 \end{bmatrix} \quad (3)$$

Since equation 1 is a nonlinear equation with many variables, an advanced nonlinear optimizer is required to solve the equation. We employ the truth-regions methods to optimize camera motion parameters and determine a consistent global scale for all partial structures.

2.2 Nonlinear Optimization Methods

Four trust-region methods are evaluated in this work. Double Dog-Leg (DDL), Dog-Leg (DL), Hook-Step (HS) and the better-known Levenberg-Marquardt (LM) methods are all heuristic procedures that combine the gradient-descent method and the Gauss-Newton method for nonlinear optimization.^{2,6-8} The gradient-decent (or steepest-decent) method^{9,10} guarantees progress toward the goal of a local minimum. However, if the solution landscape is not well conditioned (there are elongated ridge and valley), it will take a long time for the program to reach the final goal. The Newton’s method,^{9,10} on the other hand, converges quickly in the vicinity of a local minimum, but can diverge or cycle when starting far from a local minimum. Since the strength of one option is the weakness of the other, it is logical to combine the two methods. This combined approach forms a basis of the optimization algorithms called the trust-region methods that guarantee convergence to a local minimum from any starting search point.

The first such algorithm, the Levenberg-Marquardt (LM) method, searches in a direction that interpolates between steepest descent and Newton’s method, favoring steepest descent when progress is slow and moving to Newton’s method as progress improves. This allows LM to combine steepest descent’s advantage of guaranteed convergence with the fast local convergence of Newton’s method, while avoiding the pitfalls of both individual methods. Other trust region methods, called Powell’s Dog Leg (DL) method and the Double Dog Leg (DDL) method proposed by Dennis and Schnabel,^{2,8} and the Hook-Step (HS) method, are more sophisticated variants of LM. They use approximations to improve the performance of the algorithm without sacrificing its attractive convergence properties.

The way we employ DDL to solve our problem is as follows: given a set of camera movement parameters, from a sequence of N image pairs, we parameterize rotation (R) and translation (T) by 3 variables each in between two adjacent views. Because there is one global scale that should be shared by all partial structures, we take out one degree of freedom from one of the T to ensure this constraint will be satisfied. That is, the direction of the particular T can change but not its magnitude. The fixed magnitude enforces a fixed global scale. The resulting number of parameters is therefore $6N - 1$ total. Given a set of reliable 2D feature points (p and p') and their correspondences, we can employ DDL to solve these $6N - 1$ parameters in equation 1.

The initial locations of all 3D points P can be computed through light-path triangulation as long as 2D point correspondences are known and camera motion parameters (R and T) are determined. During the optimization procedure, only 2D feature points remain unchanged while other parameters including 3D points, camera rotation and translation will be updated to establish an jointly optimal configuration.

When employing a nonlinear iterative optimization procedure, the convergence speed and the quality of the final solution depends upon if a good initial guess of the camera movement parameters and correct 2D point correspondences have been obtained. To obtain a good initial guess with high quality 2D point input for the nonlinear optimizer, we use the two-stage method described in the next section.

2.3 Two-stage Outlier-filtering Scheme

Medical images in general are much more challenging than indoor or outdoor data due to repetitive patterns and similar colors of tissue and organs, which result in a small number of detected 2D features with a relatively high percentage of outliers (mismatches of point correspondences.) Although the classic RANSAC algorithm,¹¹ probably the most widely used outlier-filtering technique in the field of computer vision, has been successfully applied to many applications, we find the standard scheme does not always produce satisfactory results when applying on medical images because of the inherited nature of the challenging data. Inspired by Chum’s work¹² where iterative methods and over-determined schemes are experimented. We propose a more sophisticated two-stage outlier-filtering scheme to help tackle difficult data for our project.

2.3.1 Multi-RANSAC algorithm

In the first stage, the normalized eight-point algorithm, together with a multi-RANSAC scheme, is performed to detect outliers. The normalization scheme proposed by Hartley¹³ first translates 2D feature points so that their centroid is at the origin and then rescales the point locations to ensure the average distance from the origin equals $\sqrt{2}$. This simple transformation is applied to each of the two images independently and has been shown to provide a better numerical condition for the eight-point algorithm in computing the fundamental matrix.

At the end of each RANSAC run, a candidate fundamental matrix F is generated and all feature correspondences that are considered to be correct (or compatible with this F) should satisfy the equation $p'^T F p = 0$. A more intuitive error measure is to compute the square sum of Euclidian distances of the points from their corresponding epipolar lines. The point-line distances of both images are taken into account and the relation can be described by equation 4.

$$d(p', Fp)^2 + d(p, F^T p')^2 \quad (4)$$

Instead of blindly setting a fixed threshold to detect outliers, we apply the box-plot method, a simple statistic method used to identify outliers without any assumption of the underlying data distribution, to automatically determine the threshold. The box-plot method computes the lower quartile (the 25th percentile Q_L) and the upper quartile (the 75th percentile Q_U) of the data set. The difference between these two values ($Q_U - Q_L$) is called the inter-quartile range or IQ . Any 2D feature with distance larger than $Q_U + 1.5 * IQ$ is considered as an outlier and will be excluded from the further computation.

The algorithm of the first stage is summarized in table 1.

Table 1. Summary of the multi-RANSAC algorithm

Repeat until no outliers are detected	
Step	Description
1.	Run classic RANSAC algorithm with normalized eight-point algorithm to generate the best fundamental matrix F
2.	Given the F matrix, compute point-line residue error of each 2D feature (the Euclidean distance between a feature and its corresponding epipolar line)
3.	Use box-plot method to identify and eliminate outliers

This multi-RANSAC scheme is motivated by one important observation - *with the present of many outliers, the final model (the fundamental matrix in our application) suggested by the classic RANSAC algorithm can not best describe the underlying data, but this semi-optimal result is good enough to help identify some outliers.* After eliminating these outliers, more accurate model can be expected in the next run of the RANSAC. Our experiments show that this multi-RANSAC scheme greatly stabilizes the fundamental matrix computation.

2.3.2 Iterative over-determined five-point algorithm

It is known that the more features we use to estimate a model the more likely the adverse effect of random Gaussian noise that causes the slight displacements of 2D features can be averaged out during the computation. This is one of the reasons we propose to add a second stage process after the multi-RANSAC scheme. After outliers are removed from the feature list in the previous stage, all the remaining features then serve as input to an over-determined five-point algorithm¹⁴ routine, the state-of-the-art algorithm for solving the relative camera pose problem. However, unlike the eight-point algorithm that can be applied to compute both the fundamental matrix and the essential matrix, the five-point algorithm requires the knowledge of camera calibration. The

five-point algorithm thus works only on the essential matrix E which later can be decomposed into the camera rotation (R) and translation (T) parameters. With these parameters in hand, we can locate feature points in 3D space through light-path triangulation. We use these 3D points as additional information to help identify outliers in this stage.

The traditional RANSAC method works well on many model-fitting problems, but for certain models, the fundamental matrix for example, the optimal model can be obtained from a set of features that contaminated by outliers that are consistent with the epipolar geometry by coincidence.

Theoretically speaking any 2D feature correspondences that satisfy the equation $p'^T F p = 0$ are considered compatible with the given fundamental matrix F . It is, therefore, impossible to identify mismatches that move roughly along the epipolar lines using only Euclidian distances between feature points and their corresponding epipolar lines as the error measure. To detect this type of mismatches using extra information is the second motivation for this proposed stage. Fortunately, during our simulations we observed that these mismatches often cause five-point algorithm to fail due to numerical instability. Even if the essential matrix E can be successfully computed, these outliers' 3D points reconstructed through light-path triangulation are usually located behind the cameras, which is not a valid configuration. Testing whether a reconstructed 3D point located in front of the camera is referred to as identifying the *cheirality* of this point with respect to the camera.¹⁵ Both facts, together with the point-line distances, become valuable tools for identifying outliers. The whole procedure again will be repeated a number of times until no more outliers are detected. The algorithm is summarized in table 2.

Table 2. Summary of the iterative over-determined five-point algorithm

Repeat until no outliers are detected	
Step	Description
1.	Use all remaining features to compute essential matrix E
2.	if five-point algorithm fail to computer Report error and terminate the program
3.	Use light-path triangulation to construct all features' 3D points
4.	Eliminate features that lie behind the cameras, go back to step 1
5.	Compute point-line residue error of each 2D feature
6.	Use box-plot method to identify and eliminate outliers

We employ many layers of feature filtering operations to ensure all the outliers are excluded in the later process. This is important because while nonlinear optimizers handle noisy data very well, they are sensitive to outliers and the traditional single RANSAC method often cannot provide satisfactory filtering results when used on medical data.

2.4 Model Registration

Multi-view registration is an essential step to generate a larger, more complete model by fusing all partial structures together. Many techniques have been proposed to solve this 3D model registration problem in which the rigid-body transformation relation between two 3D-point clouds will be established. These techniques can be roughly divided into two broad categories: the iterative approach and the feature-based approach.¹⁶ The former approach minimizes the registration error iteratively if an initial estimate of the rigid-body transformation parameters is given in advance. Among all the iterative methods, the iterative closest point (ICP) algorithm³ is probably the most widely used due to its easy-to-implement characteristic. Various improvements on the classic

ICP algorithm to enhance the reliability and reduce the running time have been explored in the past decade.¹⁷ However, there is one inherited drawback of the iterative methods - they require a good initial estimate of the transformation parameters, otherwise there is no guarantee of finding correct solution even with noiseless partial structures.

Feature-based approach, on the other hand, extracts prominent local features first and then finds the correspondences of these features to estimate the transformation parameters without any initial guess from the users. However, to extract salient local features by analyzing 3D geometric structures such as principle curvatures¹⁸ and structured-meshes¹⁹ is a very computation intensive task. Furthermore, if no prominent features can be detected from the 3D input data, this approach will fail.

Both iterative approach and feature-based approach are designed to solve a general case where only two partially overlapped 3D-point clouds are given but no correspondence information is available. Fortunately, we can narrow down the problem scope by making good use of existing information - the 2D feature correspondences which have been served as guides for camera motion inference, image rectification and stereo matching in the previous steps of our project. In the model registration step, we first compute 3D locations of all remaining point correspondences through light-path triangulation and then use these 3D points as ground control points. With the help of these ground control points, the whole registration procedure has become a simple rigid-body transformation problem which can be solved efficiently using SVD.²⁰

There are three main advantages of using these ground control points: (1). The proposed method does not require any initial estimate of the transformation parameters which is essential information for iterative methods. (2). Unlike feature-based methods, our method completely eliminates the time consuming step of analyzing 3D geometric structures for extracting prominent local features. (3). Using a few reliable anchor points to estimate the transformation parameters is not only more efficient but also more accurate than using all points in noisy point clouds.

The third point discussed above is especially true when dealing with input models generated through the structure-from-motion framework. Unlike range images acquired by hardware, in this framework a stereo matching program is first employed to establish point correspondences and then light-path triangulation is applied to compute 3D locations of these points. It is well known that establishing point correspondences through stereo matching techniques has been, and still is, a very challenging problem. Noticeable amount of incorrect point correspondences are expected from stereo matching programs. These incorrect matches can only lead to wrong 3D positions after light-path triangulation, resulting in less precise 3D models than range images. Using every point in the noisy partial models increases the model registration program's chances of being trapped in a local minimum. Therefore, reliable ground control points based on SURF features should be used for computing rigid-body transformation.

To further ensure the robustness in the computation of the ground control points, a RANSAC-based scheme is again applied. The algorithm for model registration step is summarized in table 3.

3. EXPERIMENTAL RESULTS

We apply the proposed method in our Colon-CAD project¹ to improve the accuracy of the model building process. The colon images we used were acquired by an endoscope. These images depict a diverticulum, an abnormal air-filled outpouching of the colonic wall, which may lead to infection and surgery may be needed. Using the structure-from-motion framework, we were able to successfully identify this type of structures.

To visualize the quality of the inferred camera movement parameters, we employ the polar rectification algorithm²¹ to generate rectified images. Image rectification, an important step in a 3D model building process, rearranges input image pixels so that the corresponding points (that result from the projection of the same 3D point) will lie on the same scan lines on the rectified images to enable efficient and robust stereo correspondence computation. Given the input images and the fundamental matrix F which encodes the camera movement information, a rectification program outputs the corresponding rectified images. Hence, by inspecting the rectified images, we can evaluate the quality of the inferred camera movement parameters.

Figure 2(a) depicts a pair of rectified images generated using the fundamental matrix F computed by the proposed method while figure 2(b) depicts the output using the alternative F matrix computed by the traditional

Table 3. Summary of the model registration step

Step	Description
1.	Use only ground control points and run classic RANSAC algorithm with rigid-body transformation solver to estimate best transformation parameters (R, T)
2.	Given the transformation parameters (R, T), perform rigid-body transform and compute 3D Euclidean distance between correspondent ground control points
3.	Use box-plot method to identify and eliminate outliers
4.	Use all remaining ground control points to estimate transformation parameters again
5.	Perform rigid-body transform on point clouds and fuse partial models

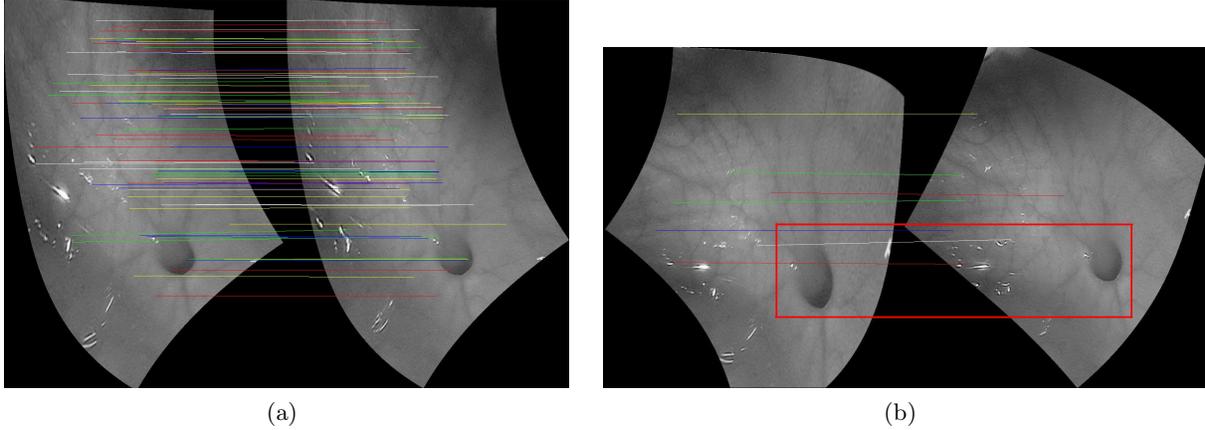


Figure 2. Rectification output comparison using (a)Proposed method (b)Traditional RANSAC

RANSAC method. As can be seen, figure 2(b) shows a large alignment error along the y axis especially within the highlighted rectangle region. Please note that, the image size used in this paper is about 10 percent of the original image size, and hence, any visible misalignments will be magnified many times in the original images and will be at least more than 30 pixels apart. With this large alignment error, it will be very difficult to match the corresponding points in the later process, resulting in an unsatisfactory 3D model.

Furthermore, a thorough analysis of four different non-linear optimization algorithms (DL, DDL, LM and HS) is performed using synthetic data. A set of 3D points and camera poses are randomly generated first and then the 2D points can be determined by projecting the 3D points onto the image planes. The input for these non-linear optimizers includes the camera movement information in between these images as well as the locations and the correspondence relations of 2D points on these images. In our simulation, different amounts of noise were added into the input 2D positions to test the robustness of these algorithms.

The accuracy results of using these methods for inferring the camera translation (T), rotation (R), 2D and 3D point locations are presented in Figures 3(a), 3(b), 3(c) and 3(d) respectively. There is no significant difference among these four algorithms in terms of accuracy since all of them converge to the ground truth with error less than 10^{-13} . Furthermore, Figure 3(e) shows that all these methods took about the same number of iterations to converge. However, because different algorithms perform slightly different computation in each

iteration, it is clear in Figure 3(f) that the DDL method is twice as fast as the LM method.²²

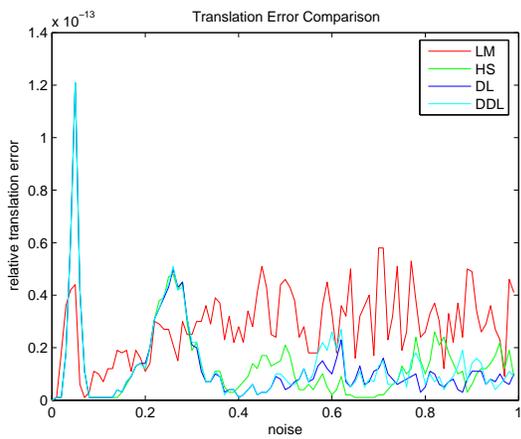
Figures 4(a) and 4(b) are 2D views of a preliminary reconstructed model with texture on top of it. A more clear view of the structure can be seen from a Matlab output in figure 4(c). At this perpendicular view angle, we can see a relatively flat structure with one clear cavity (the rectangle area.)

4. CONCLUSIONS

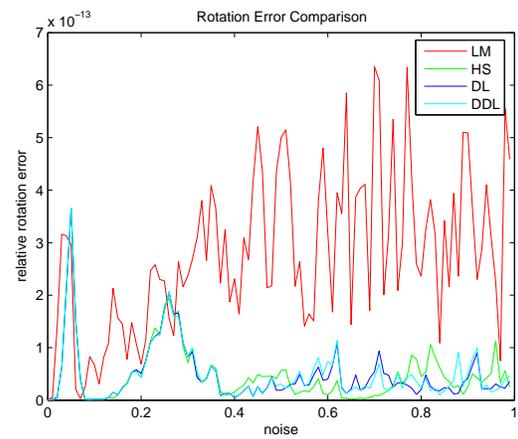
A technique for robust and efficient 3D structure inference and registration from video is presented. Extensive tests of four different nonlinear optimization algorithms show that there is no significant difference among these algorithms in terms of accuracy but the DDL method requires the least amount of running time. We apply DDL to solve the multi-view registration problem by simultaneously optimizing the camera movement parameters and determining a consistent global scale. With this global scale and reliable point correspondences, we can avoid using the time-consuming iterative closest point (ICP) based algorithms for multi-view registration. Instead, a simple rigid-body transformation with RANSAC can be applied.

REFERENCES

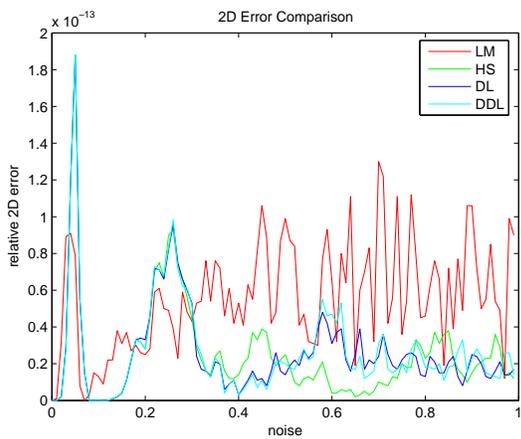
1. D. Koppel, C.-I. Chen, Y.-F. Wang, H. Lee, J. Gu, A. Poirson, and R. Wolters, "Toward automated model building from video in computer assisted diagnoses in colonoscopy," in *Proceedings of the SPIE Medical Imaging Conference*, 2007.
2. J. E. Dennis and H. H. W. Mei, "Two new unconstrained optimization algorithms which use function and gradient values," *Journal of Optimization Theory and Application* **28**, pp. 453–482, August 1979.
3. P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**, pp. 239–256, February 1992.
4. D. Sargent, C.-I. Chen, C.-M. Tsai, Y.-F. Wang, and D. Koppel, "Feature detector and descriptor for medical images," in *Proceedings of the SPIE Medical Imaging Conference*, 2009.
5. W. Li, S. Nie, M. Soto-Thompson, C.-I. Chen, and Y. I. A-Rahim, "Robust distortion correction of endoscope," in *Proceedings of the SPIE Medical Imaging Conference*, 2008.
6. R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second ed., 2004.
7. A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods (MPS-SIAM Series on Optimization)*, SIAM, Philadelphia, PA, 2000.
8. J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, PA, 1996.
9. J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997.
10. Y. Saad, *Iterative Methods for Sparse Linear System*, SIAM, Philadelphia, PA, second ed., 2003.
11. M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM* **24**, pp. 381–395, June 1981.
12. O. Chum, J. Matas, and J. Kittler, "Locally optimized ransac," *Lecture Notes in Computer Science* **2781**, pp. 236–243, 2003.
13. R. I. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**, pp. 580–593, June 1997.
14. D. Nistér, "An efficient solution to the five-point relative pose problem," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**, pp. 756–770, June 2004.
15. R. I. Hartley, "Cheirality invariants," in *In Proc. DARPA Image Understanding Workshop*, pp. 745–753, Morgan Kaufmann Publishers, Inc, 1993.
16. C.-S. Chen, Y.-P. Hung, and J.-B. Cheng, "Ransac-based darces: A new approach to fast automatic registration of partially overlapping range images," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**, pp. 1229–1234, 1999.
17. S. Rusinkiewicz and M. Levoy, "Efficient variants of the icp algorithm," in *Proceedings of the 3rd International Conference on 3-D Digital Imaging and Modeling*, pp. 145–152, 2001.



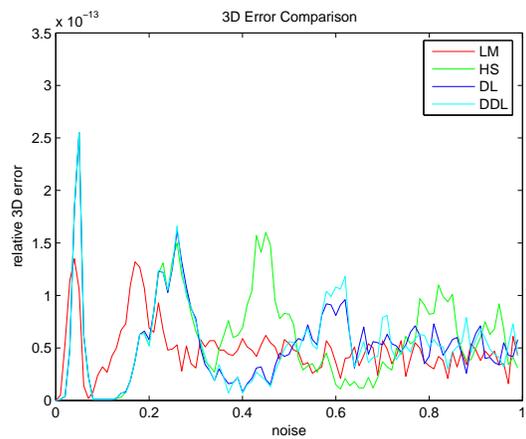
(a)



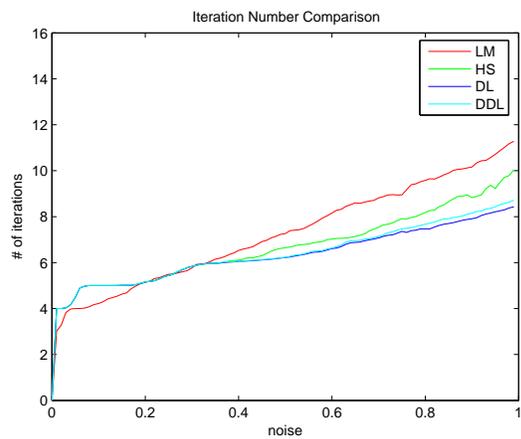
(b)



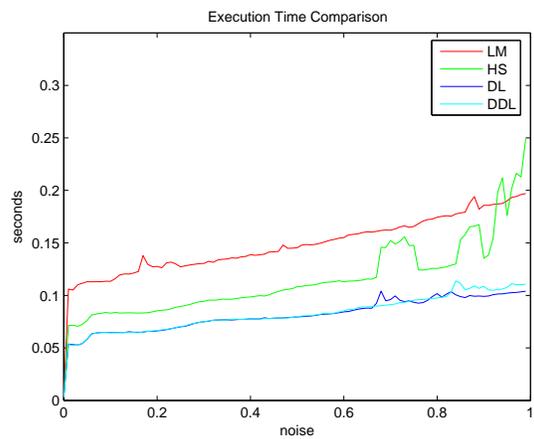
(c)



(d)

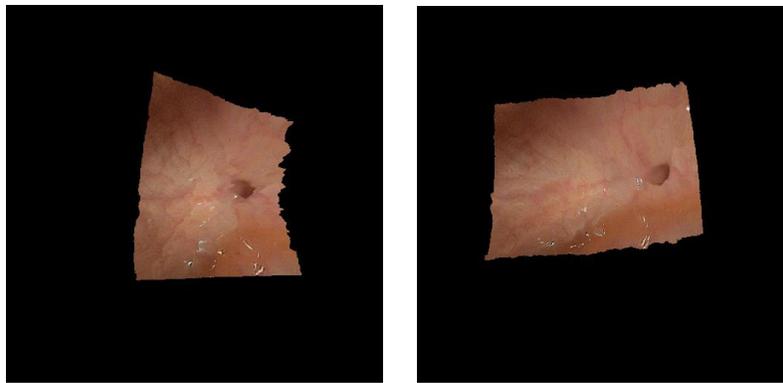


(e)



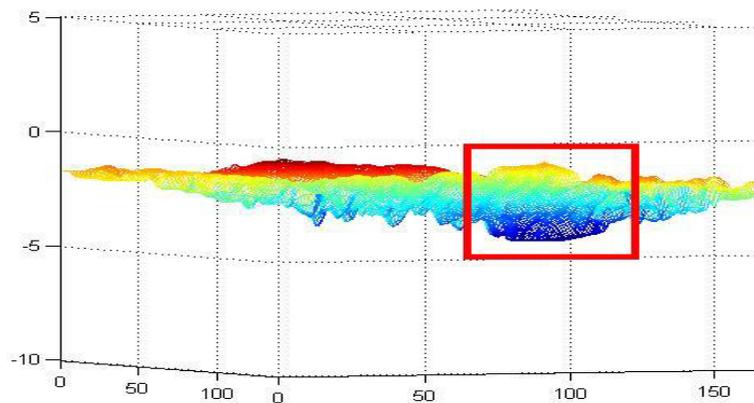
(f)

Figure 3. Analysis of four different non-linear optimization algorithms (a) Translation error (b) Rotation error (c) 2D location error (d) 3D location error (e) Number of iterations required (f) Execution time required



(a)

(b)



(c)

Figure 4. 3D model results

18. C. S. Chua and R. Jarvis, "3d free-from surface registration and object recognition," *International Journal of Computer Vision* **17**, pp. 77–99, 1996.
19. R. Bergevin, M. Soucy, and H. Gagnon, "Towards a general multi-view registration technique," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**, pp. 540–547, May 1996.
20. D. W. Eggert, A. Lorusso, and R. B. Fisher, "Estimating 3-d rigid body transformations: A comparison of four major algorithms," *Machine Vision and Applications* **9**, pp. 272–290, March 1997.
21. M. Pollefeys, R. Koch, and L. V. Gool, "A simple and efficient rectification method for general motion," in *Proceedings of International Conference on Computer Vision*, pp. 496–501, 1999.
22. D. Sargent, *Tracker-Endoscope Calibration for Colonoscopy*. PhD thesis, Department of Computer Science, University of California, Santa Barbara, 2008.