Sparse Reconstruction of Liver Cirrhosis from Monocular Mini-Laparoscopic Sequences

Jan Marek Marcinczak, Sven Painer and Rolf-Rainer Grigat
TU Hamburg-Harburg, Harburger Schlossstrasse 20, 21079 Hamburg, Germany

ABSTRACT

Mini-laparoscopy is a technique which is used by clinicians to inspect the liver surface with ultra-thin laparoscopes. However, so far no quantitative measures based on mini-laparoscopic sequences are possible. This paper presents a Structure from Motion (SfM) based methodology to do 3D reconstruction of liver cirrhosis from mini-laparoscopic videos. The approach combines state-of-the-art tracking, pose estimation, outlier rejection and global optimization to obtain a sparse reconstruction of the cirrhotic liver surface. Specular reflection segmentation is included into the reconstruction framework to increase the robustness of the reconstruction. The presented approach is evaluated on 15 endoscopic sequences using three cirrhotic liver phantoms. The median reconstruction accuracy ranges from 0.3 mm to 1 mm.

Keywords: Structure from Motion, 3D Reconstruction, Liver Cirrhosis, Laparoscopy

1. DESCRIPTION OF PURPOSE

Mini-laparoscopy is a technique used by clinicians to inspect the cirrhotic liver surface with ultra-thin endoscopes. Due to the mini-laparoscope, this technique is less invasive than classical laparoscopy and has been proven to increase the reliability of diagnosis. However, so far no quantitative analysis of the liver surface based on mini-laparoscopy exists. To obtain accurate measures of the 3D structure, this paper evaluates the accuracy of Structure from Motion (SfM) based reconstructions in the context of mini-laparoscopy.

For 3D reconstruction in mini-laparoscopy, a passive method is desirable as the hardware which is already present in the clinic can be used. One of the most evolved passive Shape from Shading based approaches to 3D reconstruction in laparoscopy has been published by Collins et al. However, it requires a constant albedo which must be known in advance. Furthermore, ambiguities of the Shape from Shading solution lead to false reconstructions. Their final conclusion is that 3D reconstruction in laparoscopy cannot be performed by shading cues alone. Due to the limited practicability of Shape from Shading, the focus of this article is set to develop a SfM based reconstruction for mini-laparoscopy which is robust to specular reflections and illumination change.

Well-known solutions to the Structure from Motion problem have been published by M. Pollefeys and Sturm and Triggs. Early attempts to SfM using monocular endoscopes have been proposed by Thormhlen et al. while more recent approaches can be found in. Most of these SfM strategies can be divided into the following steps: compute corresponding features, compute an initial reconstruction and global optimization/bundle adjustment (BA). However, after an initial reconstruction is obtained there are many possible ways to proceed and the optimal strategy depends on the application. Feature selection, tracking strategy and the arrangement of triangulation and optimization steps depend on the image data and performance requirements; in offline reconstruction other strategies can be used as in real-time applications where the entire sequence is not known in advance.

2. METHOD

The presented SfM approach is based on a subset of frames from the sequence which are called keyframes. Using these keyframes a 3D reconstruction of the scene and the camera motion is computed. The classical SfM approach uses a rigid motion model. Therefore, non-rigid motion or deformations will lead to significant inaccuracies or even failure of the reconstruction. In the case of mini-laparoscopic videos of the liver surface, the liver and the abdomen wall are deformed by respiratory motion. Due to the increased liver stiffness caused by fibrosis, there is no notable deformation of the liver surface by respiratory motion in the case of liver cirrhosis. However, the abdomen wall undergoes a non-rigid deformation caused by respiration. This causes problems when it comes to motion estimation from 2D correspondences. The
relative pose estimation is based on the epipolar constraint which is violated if a non-rigid deformation occurs. One solution to this problem is, to perform a segmentation of the cirrhotic liver surface and to use only corresponding points which are located on the liver surface. Another option is to rely on robust estimators.\textsuperscript{17,18} As the points located on the abdomen wall move non-rigidly they do not fulfill the rigid motion hypothesis and are detected as outliers by the robust estimators. The latter approach has the advantage that points which are not located on the liver surface and fulfill the rigid motion hypothesis are not automatically discarded and consequently more corresponding points are kept. This is especially important as there are often few corresponding points if the liver surface is sparsely textured – this depends on the type of cirrhosis.

As it is necessary to use a robust estimator for the computation of the relative pose because of false correspondences,\textsuperscript{13} it can also be used to discard the points which perform a non-rigid motion. The drawback of this approach is that in the case where most corresponding points perform a non-rigid motion and only a few points are detected on the liver surface, many iterations can be necessary to find a well supported hypothesis or the robust estimator even fails. To improve the estimation it is highly recommended to use a pre-calibrated endoscope. This reduces the number of parameters that need to be estimated and thereby the number of iterations of the robust estimator.

To increase the quality of the corresponding points, several filters are used to detect false 2D correspondences before the motion estimation is performed. Points which are located in specular reflections are detected by the approach described in\textsuperscript{19} but are not discarded for motion estimation. Stable (not moving) reflections support the motion estimation, while moving reflections disturb the estimation. Again, the robust estimator does the job and decides if a reflection supports the motion estimation. Correspondences which fulfill the epipolar constraint are used for motion estimation and outliers are discarded. More important is specular reflection segmentation when it comes to back-projection. Corresponding points which are located in specular reflections in multiple frames usually lead to inaccurate points on the surface and are therefore not used for triangulation even if they are classified as inlier. In the proposed SfM strategy the calibrated five point algorithm\textsuperscript{15} is used for estimation of the relative pose while the robust triangulation algorithm\textsuperscript{13} is used to compute the 3D points from 2D correspondences. After triangulation, outliers in the point cloud are detected by a nearest neighbor analysis. In a next step, the poses of the laparoscope for the frames which are in between both keyframes are initialized by interpolating the estimated relative motion. The poses could also be estimated by using 2D-3D correspondences.\textsuperscript{20} However, after initialization, bundle adjustment\textsuperscript{8,21} is carried out and the initial poses are corrected by minimizing the re-projection error. Due to the short feature trajectories in laparoscopic videos, keyframes are close to each other and the interpolation works well. Naturally, this cannot be done in sequences where long trajectories are available and the linear interpolation does not reflect the motion between the keyframes. After the bundle adjustment, points with high geometric error are removed from the reconstruction and the bundle adjustment is repeated until all reconstructed points have a small geometric error. Afterwards, the next keyframe is processed. The entire SfM strategy is illustrated in Figure 1.

3. RESULTS

The accuracy of the reconstructed 3D point cloud was evaluated using 15 sequences of 3 different cirrhotic liver phantoms. The camera pose and the location of the liver phantoms were measured by an optical tracking system which allows to generate ground truth depth maps. During the evaluation process, the point cloud is registered to the reference point cloud. For every frame, the 2D correspondences of all visible 3D points are used to obtain a reference depth value which is used for comparison. Therefore, the accuracy stated in Table 1 reflects the accuracy of the structure and not of the absolute placement of the point cloud. After registration the distance in millimeters between corresponding surface points in the reconstructed and reference point cloud is calculated. The 25%, 50%, 75%, 90%, 95% and 99% percentiles of the error distribution are given in Table 1. As expected, the reconstructions based on the Sequences 1 and 3 have a lower reconstruction error. These sequences are recorded close to the liver surface which reduces the reconstruction error. As the structure is evaluated frame by frame, the reconstruction error reflects the accuracy of the local structure which might differ from an evaluation where the complete reconstruction is compared in a single step. The median reconstruction error is below 1 mm – apart from Sequence 5 of phantom 2 where the median error is 1.01 mm. Difficult is the comparison to other approaches which have been present in similar fields. To the best of the authors knowledge no work has been published with focus on the reconstruction accuracy of Structure from Motion in the context of liver cirrhosis. However, the work of Groch et al.\textsuperscript{22} used Structure from Motion to reconstruct the surface from two different liver phantoms (no nodules) among other organ surfaces using an HD endoscope. The resulting average accuracy was 4.3 mm for one phantom and 5.6 mm for the other phantom. As the resulting accuracy depends on the distance to the scene, illumination, motion
Figure 1: Overview of the SfM strategy for 3D reconstruction from mini-laparoscopic videos. The strategy includes a specular reflection segmentation which is used to exclude pixels located in specular reflections from triangulation.
Table 1: Reconstruction error in millimeters of the sparse point cloud. The $p_{25}$, $p_{50}$, $p_{75}$, $p_{90}$, $p_{95}$ and $p_{99}$ percentiles of the reconstruction error are shown for the 15 sequences of the three liver phantoms.

<table>
<thead>
<tr>
<th>Phantom</th>
<th>Sequence</th>
<th>$p_{25}$</th>
<th>$p_{50}$</th>
<th>$p_{75}$</th>
<th>$p_{90}$</th>
<th>$p_{95}$</th>
<th>$p_{99}$</th>
<th>[ mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phantom 1</td>
<td>1</td>
<td>0.24</td>
<td>0.41</td>
<td>0.68</td>
<td>1.03</td>
<td>1.35</td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td><strong>0.15</strong></td>
<td><strong>0.29</strong></td>
<td>0.53</td>
<td>0.86</td>
<td>1.15</td>
<td>1.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.18</td>
<td>0.32</td>
<td>0.53</td>
<td><strong>0.80</strong></td>
<td><strong>1.01</strong></td>
<td><strong>1.56</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.39</td>
<td>0.73</td>
<td>1.36</td>
<td>2.44</td>
<td>3.46</td>
<td>6.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.34</td>
<td>0.59</td>
<td>0.98</td>
<td>1.75</td>
<td>3.01</td>
<td>7.26</td>
<td></td>
</tr>
<tr>
<td>Phantom 2</td>
<td>1</td>
<td>0.22</td>
<td>0.41</td>
<td>0.69</td>
<td>1.03</td>
<td>1.29</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.19</td>
<td>0.35</td>
<td>0.63</td>
<td>1.00</td>
<td>1.33</td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.22</td>
<td>0.41</td>
<td>0.67</td>
<td>0.98</td>
<td>1.21</td>
<td>1.84</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.29</td>
<td>0.51</td>
<td>0.87</td>
<td>1.43</td>
<td>1.97</td>
<td>5.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.62</td>
<td>1.01</td>
<td>1.93</td>
<td>4.47</td>
<td>6.28</td>
<td>10.20</td>
<td></td>
</tr>
<tr>
<td>Phantom 3</td>
<td>1</td>
<td>0.16</td>
<td>0.30</td>
<td><strong>0.52</strong></td>
<td>0.81</td>
<td>1.06</td>
<td>1.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.47</td>
<td>0.82</td>
<td>1.30</td>
<td>2.25</td>
<td>3.22</td>
<td>6.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.17</td>
<td>0.32</td>
<td>0.56</td>
<td>0.92</td>
<td>1.27</td>
<td>2.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.39</td>
<td>0.70</td>
<td>1.27</td>
<td>2.29</td>
<td>3.30</td>
<td>5.91</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.47</td>
<td>0.82</td>
<td>1.30</td>
<td>2.25</td>
<td>3.22</td>
<td>6.66</td>
<td></td>
</tr>
</tbody>
</table>

and texture of the phantoms, it is not possible to compare both accuracies directly. One major difference is that the pose of the endoscope has been measured by an optical tracking system while in the presented approach no additional hardware is required and the camera pose is estimated from image correspondences. In a more recent publication by Maier-Hein et al., the precision of SfM based reconstruction in endoscopy is given as "from a few to a dozen millimeters". Three classical categories for macroscopic liver cirrhosis are: micronodular (< 3 mm), nodular (3–7 mm) and macronodular (> 7 mm). So far no evaluation of the detection rate of nodules based on 3D reconstructions has been performed. However, the median accuracy of less than 1 mm in the presented evaluations lets one expect that most of the nodular and macronodular structures of liver cirrhosis will be detectable in the reconstruction. This is illustrated by Figure 2, where a nodule in the image and its corresponding reconstruction is shown. The reconstruction is based on 200 frames of a real mini-laparoscopic sequence. The big nodule is clearly visible as local extremum in the reconstructed point cloud.

4. CONFIRMATION

This work has not been submitted for presentation or publication elsewhere.

REFERENCES


(a) Sparse 3D reconstruction of a mini-laparoscopic sequence.

(b) Projections of the 3D point cloud into frame v_{ref}.

Figure 2: Sparse 3D reconstruction of a real mini-laparoscopic sequence. a.) Reconstructed point cloud and camera trajectory of a real mini-laparoscopic sequence (view from above). The liver surface and the big nodule are clearly visible in the reconstructed point cloud. b.) Projection of the point cloud into frame. One can verify that the nodule in the point cloud corresponds exactly to a big nodule visible in the images.