



**HAL**  
open science

# Preliminary results on 3D reconstruction of small bowel by Non-Rigid Structure-From-Motion

Daubresse-Chasle Marie, Parashar Shaifali, Andreff Nicolas

► **To cite this version:**

Daubresse-Chasle Marie, Parashar Shaifali, Andreff Nicolas. Preliminary results on 3D reconstruction of small bowel by Non-Rigid Structure-From-Motion. CRAS, 2023, Paris, France. hal-04393919

**HAL Id: hal-04393919**

**<https://hal.science/hal-04393919v1>**

Submitted on 25 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Preliminary results on 3D reconstruction of small bowel by Non-Rigid Structure-From-Motion

Marie Daubresse–Chasle<sup>1</sup>, Shaifali Parashar<sup>2</sup>, and Nicolas Andreff<sup>1</sup>

<sup>1</sup>Femto-ST Institute, <sup>2</sup>LIRIS

## INTRODUCTION

The observation of lesions on the surface of the small intestine is important for the diagnosis and monitoring of inflammatory diseases. This part of the gut is not accessible by endoscopy. Since the 2000s, wireless capsule endoscopy (WCE) has made it possible. However, capsules are subject to irregular peristaltic motion. Thus it is sometimes difficult to interpret capsule movements from the images alone and to correctly count the lesions. Reconstruction of the small bowel from WCE video will help the physician to have a better spatial understanding of the lesion localization [1], [2].

In the context of gut reconstruction, the most commonly used methods belong to the structure-from-motion (SfM) paradigm, with the assumption of visualizing a rigid scene [3], [4]. It is relevant with an endoscope because two consecutive images are very close in time. However, for WCE, the acquisition frequency is too low. Non-Rigid Structure-From-Motion (NRSfM) seems to be suitable for our case. Nonetheless, existing algorithms [5], [6] assume to observe a surface with fronto-parallel deformations (Fig. 1) which are not representative of the topology of the gut, where one rather observes radial deformations.

A first approach to take into account the tubular topology of the gut [1] creates a mosaic of the small bowel by applying a circular mesh at the center of each image (Fig. 2a). Each image is unfolded together with the mesh and thus forms a vertical slice of the mosaic (Fig. 2b).

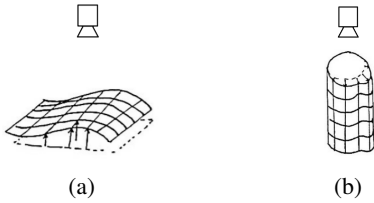


Fig. 1: (a) Fronto-parallel deformations and (b) radial deformations

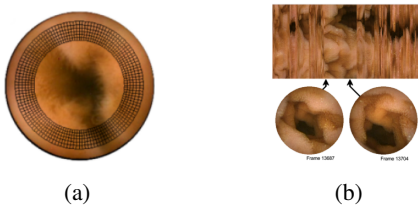


Fig. 2: (a) The mesh applied on a WCE video frame and (b) an example of map with peristalsis motion [1]

Nevertheless, the mesh does not take the camera position into account to correctly fit real data because it is an inherent 2D mesh rather than the image projection of a 3D shape.

A recent algorithm [2] solves the tubular NRSfM in two steps: first, solve a conventional N-view NRSfM, then apply the tubular constraints on the reconstructed 3D point cloud to get the gut geometry. In this paper, we present a 2-view single-step NRSfM algorithm adapted to the tubular topology of the small intestine by using a cylindrical template. Unfolding the images along this mesh allows us to transform the Tubular NRSfM problem into a fronto-parallel NRSfM problem.

## MATERIALS AND METHODS

Our method is an update of the fronto-parallel NRSfM algorithm [6]. The latter works as follows. First,  $N$  points are matched from 2 images. Assuming the observed surface to be smooth, it can be approximated locally by a planar patch around each detected point. Thereby, a local homography (aka a warp) can be computed between a neighborhood around each point in the first image and another one around the matching point in the second image. Similarly, the deformation undergone by the 3D surface between the two views corresponds to the displacement and rotation of the 3D planar patch around the 3D point to be reconstructed. This transformation can not be fully recovered, but, using isometric constraints and differential geometry, the change of orientation of the 3D patch normal vector can be estimated. It can express as a function of the Jacobian of the reconstruction function  $\Gamma$  (inverse of the image projection) and of the local homography computed from the pairs of image points  $\eta$ . Then, fitting a Thin Plate Spline (TPS) over the reconstructed normal vector field provides a smooth estimate of the surface. Yet, as stated above, this method is not robust when applied to tubular topology. Therefore, inspired by the cylindrical

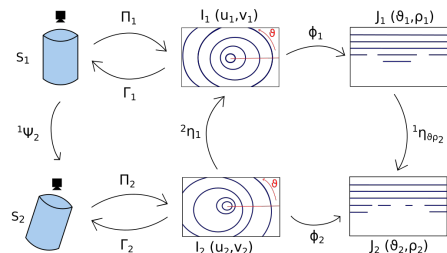


Fig. 3: Outline of the proposed tubular NRSfM.

mesh mosaicking method [1], we transform each actual image point in Cartesian  $(u,v)$  coordinates into a virtual point in polar  $(\theta, \rho)$  coordinates (Fig. 3), thus transforming the actual image of a tubular surface into a virtual fronto-parallel image of a nearly planar surface. To do so, the mesh is centered on the vanishing point rather than the image center. In practice, we thus create a pair of transformations  $(\Phi_1, \Phi_2)$  from the pair of actual images  $I_1, I_2$  to the pair of virtual images  $J_1, J_2$ . We can thereby model the warps between  $I_1$  and  $I_2$  as  ${}^2\eta_1 = \Phi_1^{-1} \circ {}^2\eta_{\theta\rho_1} \circ \Phi_2$ , where  ${}^2\eta_{\theta\rho_1}$  are the warps between virtually fronto-parallel images. Then, using this model and the warps  ${}^2\eta_1$  computed from actual images allows us to plug the tubular topology into the fronto-parallel NRSfM algorithm described above.

## RESULTS

Our algorithm is more suitable for tubular topologies than NRSfM [6], [5]. We compare our algorithm with Sengupta’s method [2], which is the only state-of-the-art method to successfully reconstruct a tubular scene using NRSfM. For that, we use their two synthetic datasets (Fig. 4). **NR-Synth-1** consists of 69 images of a cone deformed at each frame created on Blender. **NR-Synth-2** is similar to **NR-Synth-1** but the starting shape is a cylinder and the dataset is composed of 79 images. We compute the average error  $e_p$  between the computed 3D points and the ground truth and the associated standard deviation  $\sigma_p$ . The results of [2] for an image are obtained from the N-1 other images. In our case, since our method only uses 2 views, we average the results over the N-1 images for comparison. See table I for results. Our algorithm obtains better results than [2] on **NR-Synth-2** when the cylinder is little deformed or when the deformation between two images is limited (Fig. 5a). Moreover, [2] is a bit more sensitive to outliers than the proposed method (Fig. 5a, 5b). On the other hand, our algorithm does not succeed yet in reconstructing a cone as in frame 17 (Fig. 5c). Indeed, there is an ambiguity between height and radius in the projection of a cylinder onto an image (Fig. 6). So far, our approach was to fix the cylinder

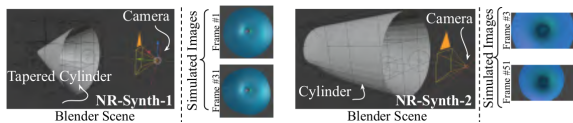
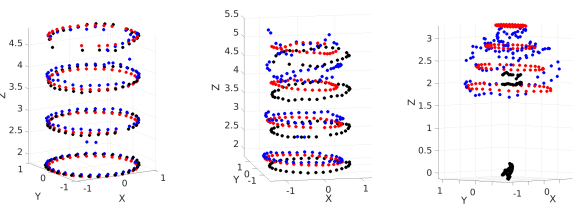


Fig. 4: Images taken from [2]. Simulation set-up from Blender and sample frames



(a) Frame 1 of NR-Synth-2 (b) Frame 79 of NR-Synth-2 (c) Results for frame 17 of NR-Synth-1

Fig. 5: Results for the proposed method (black points) and for Sengupta and Bartoli’s method [2] (blue points) in comparison with Ground truth (red points)

	NR-Synth-2				NR-Synth-1	
	Frame 1		Frame 79		Frame 17	
	$e_p$	$\sigma_p$	$e_p$	$\sigma_p$	$e_p$	$\sigma_p$
Proposed method	0.04	0.04	0.17	0.12	0.26	0.18
Sengupta & Bartoli [2]	0.09	0.20	0.15	0.21	0.10	0.08

TABLE I: Experimental results of our method compare to Sengupta and Bartoli’s method [2] for **NR-Synth-1** and **NR-Synth-2**

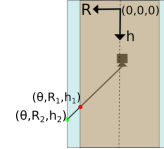


Fig. 6: Ambiguity between height and radius of a cylinder

radius to determine a unique height. Therefore, it is natural that the cone can only be reconstructed as an elongated cylinder.

## CONCLUSIONS AND DISCUSSION

Our NRSfM method adapted to the cylinder obtains good results for the reconstruction of a cylinder and competes with the results of Sengupta and Bartoli [2]. However, it creates distortions on the reconstruction of non purely cylindrical shape. This is due to the existence of an ambiguity between height and radius when projecting a point of the cylinder. The future works are to resolve this ambiguity by constraining the deformation between two points. Moreover, taking better account of the camera pose to create the 2D mesh could improve results and will be more relevant to the actual movement of the capsule in the intestine. Finally, this paper present preliminary results and our algorithm needs to be test on scene with more important deformations as observed during peristaltic movement.

Acknowledgment. This work was funded by Grand Prix Scientifique 2018, Fondation Ch. Defforey-Institut de France

## REFERENCES

- [1] P. M. Szczypiński, R. D. Sriram, P. V. Sriram, and D. N. Reddy, “A model of deformable rings for interpretation of wireless capsule endoscopic videos,” *Medical Image Analysis*, vol. 13, no. 2, pp. 312–324, 2009.
- [2] A. Sengupta and A. Bartoli, “Colonoscopic 3D Reconstruction by Tubular,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, pp. 1237–1241, 2021.
- [3] N. Mahmoud, “OrbSlam-based endoscope tracking and 3d reconstruction,” *International MICCAI Workshop on Computer Assisted and Robotic Endoscopy*, Oct 2016.
- [4] R. Miranda-Luna, C. Daul, W. C. Blondel, Y. Hernandez-Mier, D. Wolf, and F. Guillemin, “Mosaicing of bladder endoscopic image sequences: Distortion calibration and registration algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 541–553, 2008.
- [5] S. Parashar, D. Pizarro, and A. Bartoli, “Local Deformable 3D Reconstruction with Cartan’s Connections,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 12, pp. 3011–3026, 2020.
- [6] S. Parashar, Y. Long, M. Salzmann, and P. Fua, “A closed-form solution to local non-rigid structure-from-motion,” *arXiv e-prints*, pp. arXiv–2011, 2020.