INSTITUTO DE SISTEMAS E ROBÓTICA
UNIVERSIDADE DE COIMBRA

Multi-Camera Wireless Capsule Endoscopy:

Segmentation of Motion and Local Panoramic Views for Capsule Localization and Lesion Detection

Marina Oliveira, Supervised by Professor Helder Araujo

Introduction

Wireless Capsule Endoscopy (WCE) allows full inspection of the gastrointestinal (GI) tract in regions that are difficult to reach [1]. Its imitations include extensive reading time, limited field of view, poor image quality and lack of position data [2].

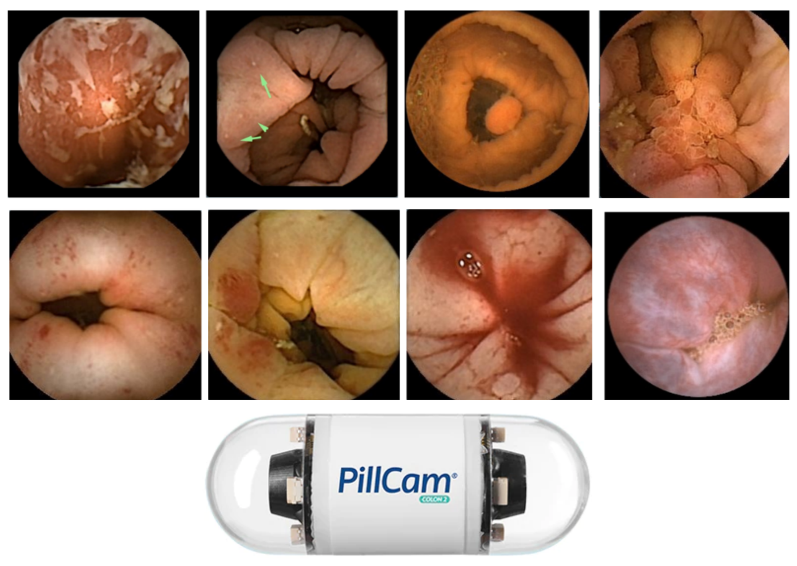

Figure 1: Sample of WCE frames of the most frequent lesions in the GI tract obtained with the multi-camera capsule PillCam Colon2. (Adapted from PillCam Rapid Reader Atlas)

A panoramic image ideally enables the simultaneous viewing of multiple frames with a broader field of view without information loss for localization and lesion detection purposes [3].

Objectives

Local panorama to map the location of the capsule along the GI tract to the timeline when an abnormality is detected.

Pipe Representation

Forward motion can be handled with the generalized pipe representation [4] with the 3D position of point Q expressed as:

$$Q = k\hat{s} + Rcos(\alpha)\hat{d} + Rsin(\alpha)\hat{r} \quad (1)$$

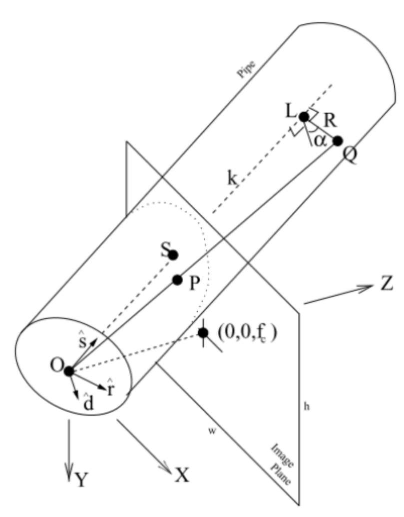

Figure 2: Particular case where the camera's forward motion is along the optical axis that coincides with the z axis. (Adapted from [4])

Pipe Mosaicing

A flow vector (u,v) between corresponding points P_k and P_{k+1} in sequential images is a function of the position. The scanning broom is the curve $F(x, y) = 0$ perpendicular to the optical flow. Assuming uniform radial optical flow, $F(x, y)$ is a circle around the center of FOE. [4].

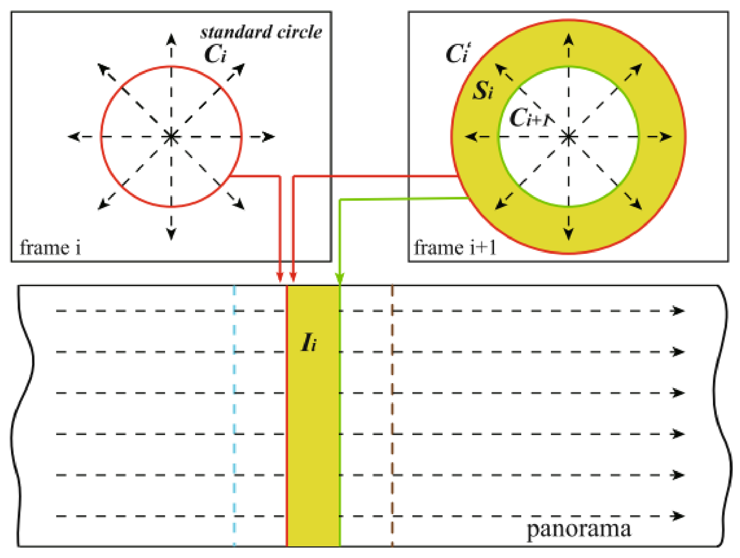

Figure 3: Representation of the scanning broom with uniform optical flow for the panorama. (Adapted from [4])

Experimental Dataset

- A tubular model with synthetic colon-like texture that was constructed with Blender;
- Crohn's disease patient videos, obtained with the multi-camera capsule PillCam Colon2;

Methodology

Since the motion of the capsule is the result of the peristaltic movements and it is not purely translational, the following steps were taken:

1. Optical flow (OF) estimation between sequential frames with the RAFT network [5];
2. Comparative analysis of the OF between sequential frames from both cameras to extract pure translational motion segments;
3. A panorama was constructed with purely translational sequential frames from the synthetic model with the bilinear interpolation of the images according to a global homography (H) computed given the correspondences obtained with the OF vectors:

$$P_{k+1} = HP_k \quad (2)$$

Preliminary Results

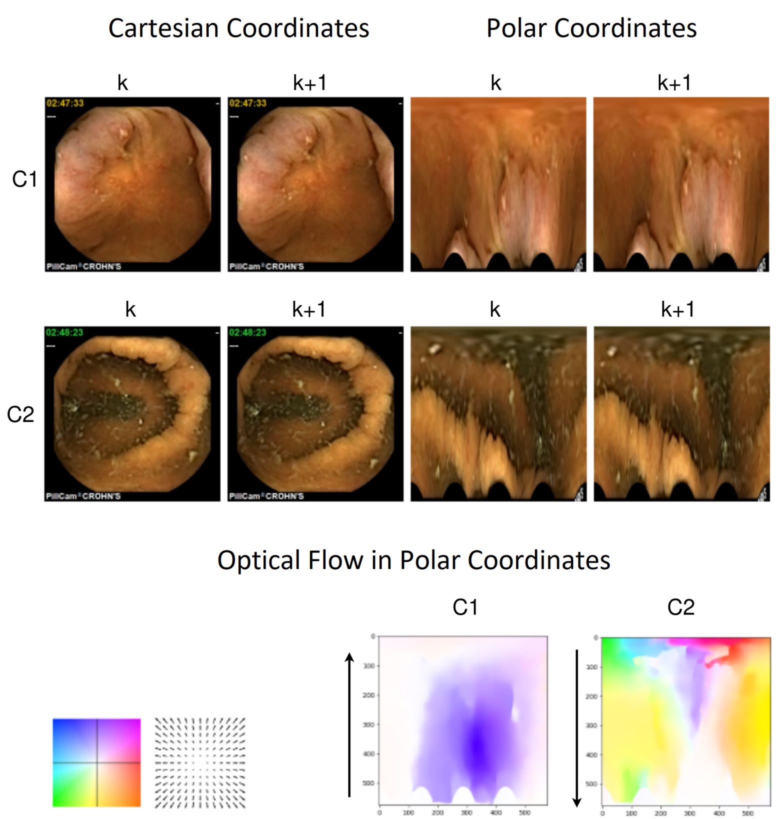

Figure 4: Example of a sequential pair of frames from the cameras on both ends of the capsule (C1 and C2) and the optical flow estimation given the polar representation with the RAFT network [5].

	C1	C2
α_{vert}^{meanOF} (°)	28.4388	178.7171

Table 1: Angle α between the mean Optical Flow vectors of each pair of consecutive frames (previously manually annotated as purely translational) in polar representation, from C1 and C2, and the vertical axis in degrees.

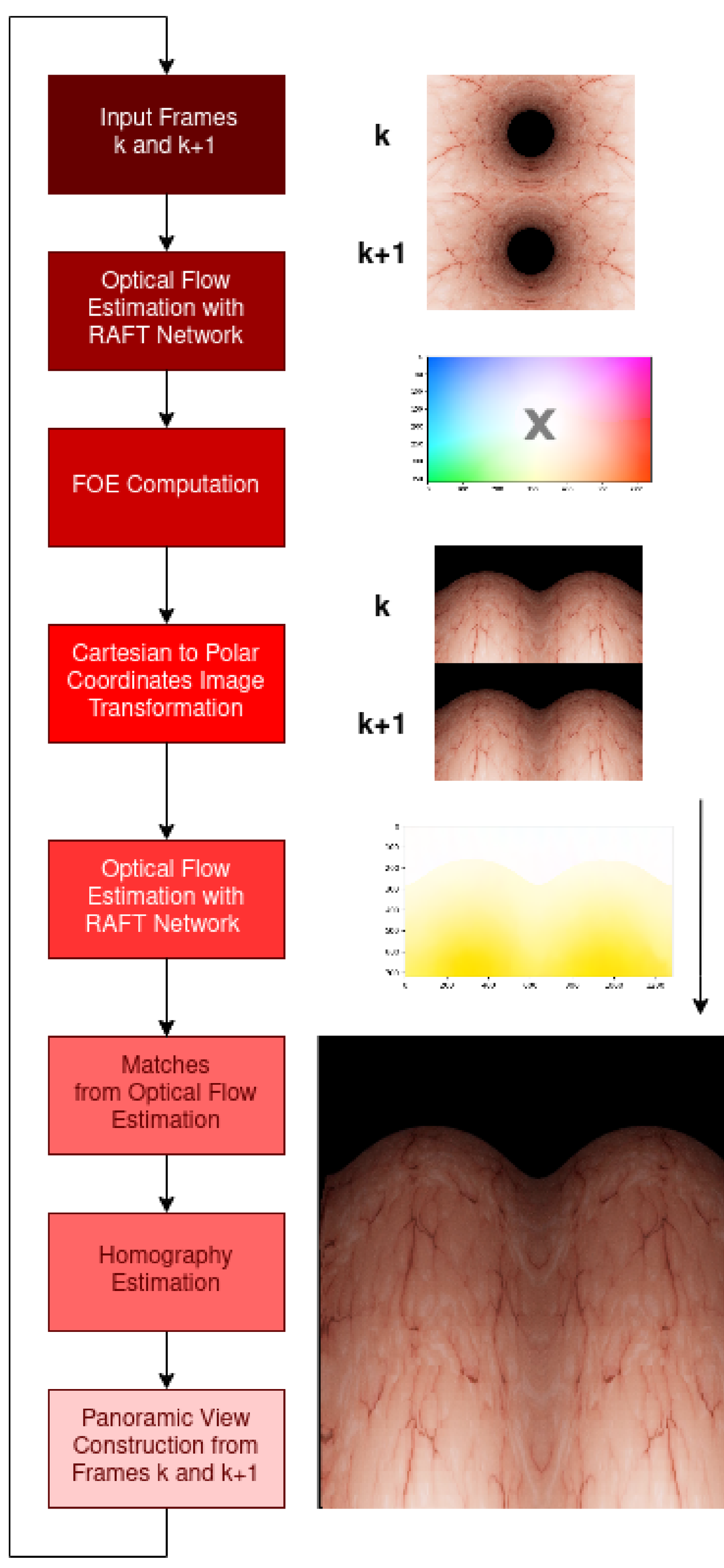

Figure 5: Summary diagram of the methodology for the construction of the preliminary result of the local panorama given the synthetic purely translational frames with central FOE from the Blender model.

Future Work

- A more robust metric for the comparative analysis of the OF vector field of consecutive frames from the opposing end cameras (C1 and C2).
- An approach to compute the FOE in the case that it is non-central.
- A more robust method for the mosaicing of the panorama, for example, the computation of several local homographies instead of a global one.
- The computation of the local panorama with the patient video given the frames extracted as purely translational.

References

[1] Muhammad Attique Khan, Muhammad Rashid, Muhammad Sharif, Kashif Javed, and Tallha Akram. Classification of gastrointestinal diseases of stomach from wce using improved saliency-based method and discriminant features selection. *Multimedia Tools and Applications*, 78(19):27743–27770, 2019.

[2] Ilknur Umay, Barış Fidan, and Billur Barshan. Localization and tracking of implantable biomedical sensors. *Sensors*, 17(3):583, 2017.

[3] Evaggelos Spyrou, Dimitris Diamantis, and Dimitris K. Iakovidis. Panoramic visual summaries for efficient reading of capsule endoscopy videos. pages 41–46, 2013.

[4] B. Rousso, S. Peleg, I. Finci, and A. Rav-Acha. Universal mosaicing using pipe projection. *Proceedings of the IEEE International Conference on Computer Vision*, pages 945–952, 1998.

[5] Zachary Teed and Jia Deng. Raft: Recurrent all-pairs field transforms for optical flow. In *European conference on computer vision*, pages 402–419. Springer, 2020.

Acknowledgements

- FCT funded research project PTDC/EMD-EMD/28960/2017, entitled "Multi-Cam Capsule Endoscopy Imagery: 3D Capsule Location and Detection of Abnormalities";
- FCT funded PhD Scholarship 2020.06592.BD;
- Institute of Systems and Robotics - University of Coimbra (ISR-Coimbra, UIDB/0048/2020)

Contact Information

Email: marina.oliveira@student.uc.pt