

Inter-operative Trajectory Registration for Endoluminal Video Synchronization: Application to Biopsy Site Re-localization

Anant Suraj Vemuri^{1,3}, Stephane A. Nicolau², Nicholas Ayache³,
Jacques Marescaux², and Luc Soler²

¹ IHU, Strasbourg, France

² IRCAD, Virtual-Surg, Strasbourg, France

³ INRIA, Sophia Antipolis, France

anant.vemuri@inria.fr

Abstract. The screening of oesophageal adenocarcinoma involves obtaining biopsies at different regions along the oesophagus. The localization and tracking of these biopsy sites inter-operatively poses a significant challenge for providing targeted treatments. This paper presents a novel framework for providing a guided navigation to the gastro-intestinal specialist for accurate re-positioning of the endoscope at previously targeted sites. Firstly, we explain our approach for the application of electromagnetic tracking in achieving this objective. Then, we show on three *in-vivo* porcine interventions that our system can provide accurate guidance information, which was qualitatively evaluated by five experts.

1 Introduction

Oesophageal adenocarcinoma (OAC) is rapidly increasing in frequency in the United States and other western countries. Gastroesophageal reflux disease, a benign complication caused by the stomach acid coming into the esophagus, as a chronic condition, leads to Barrett's esophagus (BE). It refers to the metaplasia in the cells of the lower esophagus and in most cases is a precursor to OAC. The evolution of BE to an adenocarcinoma is observed to progress from low-grade to a high-grade dysplasia. The guidelines [1] prescribe different levels of surveillance intervals depending on the degree of dysplasia with a minimum of two biopsies per year. A typical surveillance procedure involves taking four quadrant biopsies every 2cms towards the distal end of the esophagus and in suspicious regions. The biopsied tissue is sent to the pathology for evaluation. With the introduction of devices such as the probe-based confocal laser endomicroscopy real-time visualization and diagnosis of suspected regions can be performed intera-operatively. High resolution narrow band imaging has also been used for diagnosis and surveillance by visual inspection of the mucosa and the subepithelium. In each of these cases, during a follow-up inspection, the gastro-intestinal (GI) specialist is required to locate the previously biopsied or surveyed location. This problem in the literature has been termed as the re-localization

issue. Typically, the GI specialist uses the markings on the endoscope, which can be highly unreliable and which limit his or her ability to accurately re-position the endoscope and the optical biopsy probe and hence to effectively track the disease. Due to the lack of deterministic tools for providing such re-localization inter-operatively, the GI specialist has to survey or biopsy the entire affected oesophagus region, which prevents targeted treatments.

2 Related Works

To our knowledge, there is no previous work which tackles this issue of re-localization of the flexible endoscope inter-operatively. However, several approaches to track the biopsy points intraoperatively exist [2,3,4]; each of them relying on the recovery of the 3D structure of the anatomy, to map and track the biopsy sites as they move in and out of the field-of-view of the endoscope frame. Allain *et al.* [2,4] employ epipolar geometry, Mountney *et al.* [3] propose a simultaneous localization and mapping (SLAM) based method. Atasoy *et al.* [5], propose a probabilistic region matching approach in narrow-band images by using feature matches obtained from affine invariant anisotropic feature detector.

More recently Atasoy *et al.* [6] propose to formulate the re-localization as image-manifold learning process. By projecting endoscopic images on low dimensional space they propose to classify and cluster the images collected during multiple interventions into manageable segments, which they claim would aid in re-localization of the biopsy sites. However, they do not provide any spatial relations of the extracted segments inter-operatively, and so have not sufficiently clarified the application of their result in a clinical context for re-localization. We believe that, relying only on image based information for information extraction, that has to be mapped across multiple interventions can be highly unreliable; especially, due to temporal changes in tissue texture over multiple procedures, coupled with a highly deformable endoscopic scene, where repeatability of feature extraction, matching and tracking poses a significant challenge.

This paper proposes to solve the above problem, by introducing an Electromagnetic tracking system (EMTS) into the loop and providing a framework for utilizing it for re-localization inter-operatively. Firstly, we define the re-localization problem using a two-step approach; *a) Gross localization*: referring to an approximate positioning of the endoscope close to a reference point (typically a biopsy site) in a previously conducted procedure; and *b) Fine positioning*: referring to a mapping of the biopsy points taken during a previous procedure onto the images in the current intervention. Here, we address the gross localization issue. By attaching an electromagnetic (EM) sensor to the tip of a flexible endoscope, its movement within the body can be tracked with respect to a fixed external reference frame. By computing the correspondence between the EM sensor positions in two successive interventions, we propose to provide a guided-view to the GI specialist. This guided-view consists of a matching image extracted from the previously recorded intervention that best matches the live view (Fig. 3). This can, as will be shown, provide more information during the intervention

to the GI specialist. The rest of the paper is organized as follows: Sec. 3 presents the system set-up and the work-flow of our approach. Sec. 4 describes the evaluation strategy and results.

3 System Setup and Methodology

The system setup consists of an EM field generator with a working volume of $50 \times 50 \times 50\text{cm}^3$, which is placed roughly above the chest of the patient and fixed in position using a titanium arm, as shown in Fig. 1. A 6-*dof* EM sensor is inserted into the working channel of the endoscope and fixed at the tip of the endoscope (Fig. 1b). The EMTS and the endoscope are connected to a laptop, that synchronously records the data. A recording of an intervention consists of a list of EM sensor poses (trajectory of the endoscope tip), with the corresponding image frames captured from the endoscope. During a live procedure, given the recording of a previously conducted intervention, we find a corresponding image in the recording, that spatially matches the endoscope's current location in the oesophagus. Fig. 2 provides an overview of the work-flow of our method. The process is divided into three parts *a) Acquisition phase*: in which the recorded data is tagged and stored for further processing; *b) Registration phase*: to perform registration of EMTS reference frames of the live procedure and a previously recorded intervention chosen for providing a guided view.; and *c) Synchronization phase*: to perform spatial synchronization between the trajectories of the live intervention and the recording that was previously registered.

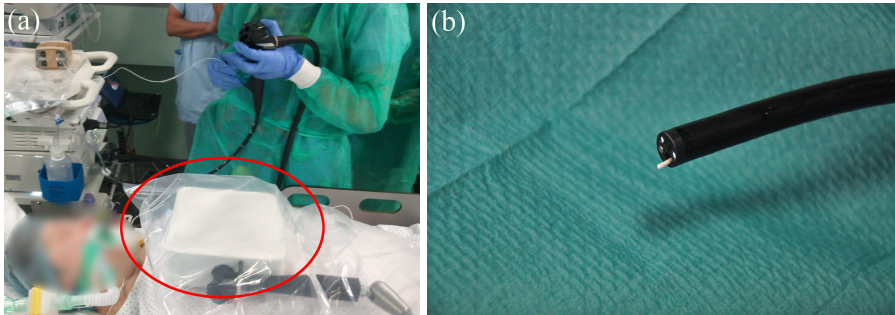


Fig. 1. (a) The apparatus setup in the operating room; (b) Placement of the sensor in the endoscope working channel

3.1 Acquisition Phase

During this phase, the GI specialist performs the recording of an intervention and tagging of relevant images. The flexible endoscope is slowly guided through the oesophagus, while the EM sensor pose and the corresponding image acquired

from the endoscope are recorded to the database. The recording contains many uninformative frames; with bubbles, motion-blur, specular highlights, and out-of-focus images. Firstly, these uninformative frames are detected and left out from further processing using the method described in [7]. The GI specialist tags the images containing the sphincter as it is used as a landmark in the registration phase (Sec. 3.2). We use the sphincter as the anatomical landmark because it is stable and can be reached with good repeatable accuracy ($\sim 3\text{mm}$). The endoscopic frames that contain the biopsy procedure are tagged and in an offline step, the expert reviews the tagged biopsy images and selects those most relevant for the procedure. At this stage the expert can choose to add supplementary information to the images of the recordings, which will be available during the synchronization phase.

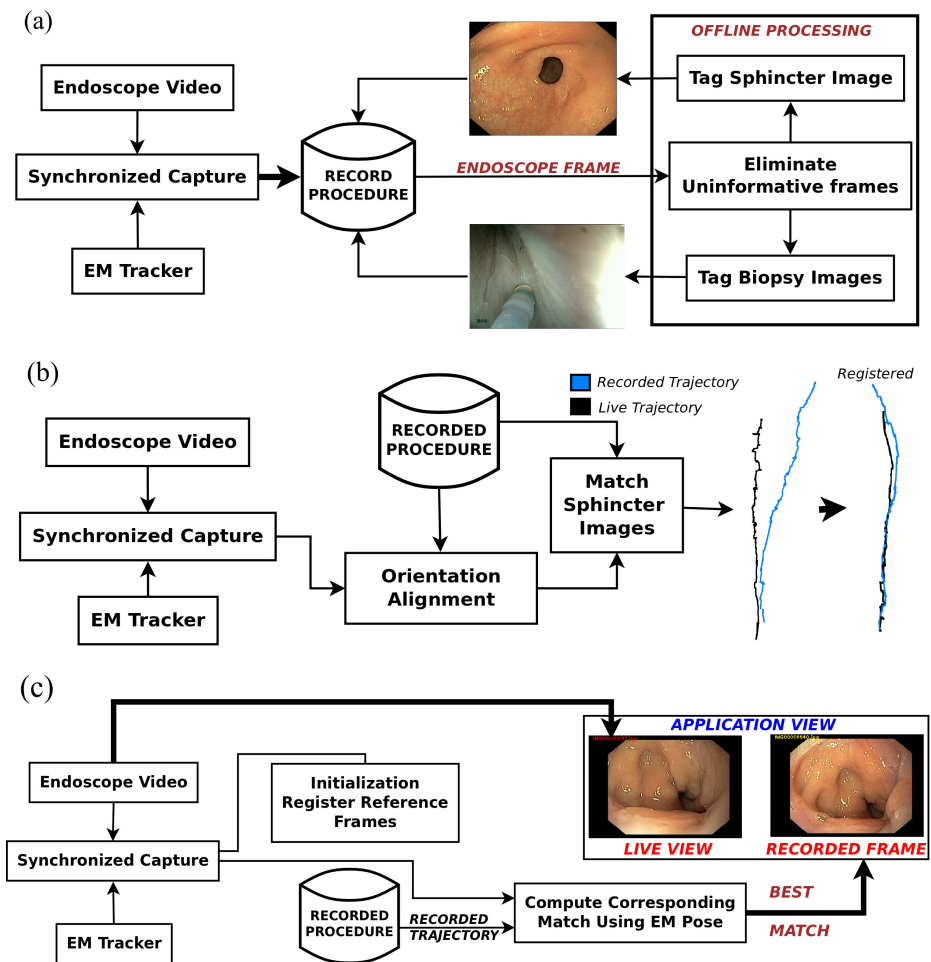


Fig. 2. Workflow, (a) Acquisition Phase; (b) Registration Phase; (c) Synchronization Phase

3.2 Registration Phase

Since the set-up of the EMTS would always change inter-operatively; registration must be performed between the EMTS reference frames of the live and a recording of a previous intervention chosen for providing a guided-view. We propose an on-line registration, without introducing additional constraints in the operating room. To achieve this, firstly, the EM sensor position is recorded while the GI specialist introduces the endoscope into the oesophagus and guides it until the sphincter. We use the contextual knowledge that the oesophagus is fairly linear and exhibits very minimal lateral movement. Hence, the largest principal components of the live trajectory and the trajectory of the previous intervention can be used to obtain 3D orientation alignment (R) along $\hat{z} = [0\ 0\ 1]^T$ for each trajectory. The tagged sphincter positions can then be used to obtain translation t , which along with R provides an initialization for the iterative closest point (ICP) for further registration refinement.

3.3 Synchronization Phase

Once the reference frames have been aligned, a spatial correspondence between the sensor position from the current intervention and the previously recorded intervention is computed. By partitioning the trajectory of the previous intervention as a binary tree, a search for the closest neighbour (in euclidean sense) to the current sensor position is made. Due to localized deformations of the oesophagus during navigation, the trajectories are not smooth and exhibit deviations (Fig. 3a) from the central oesophagus line (connecting the throat to the sphincter, along $\hat{z}$), which can lead to a false match with marked depth difference. Since the trajectories have been aligned along $\hat{z}$, the search space is constrained to lie within Δz ($\approx 2\text{mm}$). The closest neighbour gives the corresponding best matching image of the region in the oesophagus taken during the previous procedure. In particular, the matched images that were tagged to contain locations of biopsy sites, provide the GI specialist a more localized region for review. Fig. 3 presents the result of the synchronization phase. ICP finds the transform up to a rotation along $\hat{z}$, however, since the search space is constrained along this direction, the determination of the closest neighbour is unaffected by it.

4 Results

For the purposes of the experiments we used a NDI ©Aurora EMTS (accuracy of 0.5mm in a working volume of $30 \times 30 \times 30\text{cm}^3$). The proposed approach was tested on three sets of *in-vivo* sequences on different porcines, with each set consisting of four recordings. Prior marking were made on the oesophagus using a dual knife at every 2cm to simulate biopsy locations. To replicate realistic surgical procedures conducted at different times, the EM field emitter was randomly repositioned before each recording. The steps as explained in Sec. 3

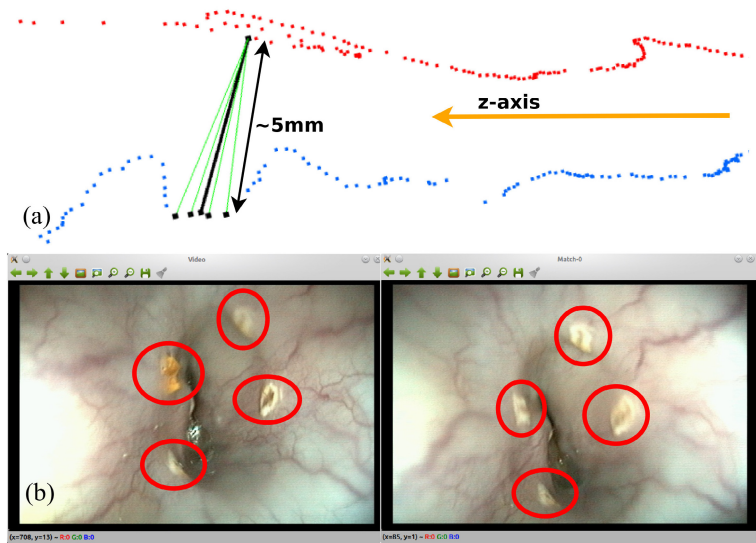


Fig. 3. View of the application window. (a) The magnified view of the live and previously recorded trajectories and the corresponding matches; (b) The image from the live view (left) and best matched image (right) from previously recorded intervention. The encircled regions indicate markings made using the dual knife.

were performed, with the first recording as the reference and three other recordings to mimic a follow-up procedure. A qualitative evaluation of the approach was performed by presenting the output of the synchronization phase to five experts. As shown in Fig. 3(b) two images were presented, one displaying frames to emulate the live stream from an intervention and the second, displaying the corresponding matching frame from a previously recorded intervention. Using the markings made by the dual knife and other visible landmarks as reference, the expert assessed the relevance and quality of the matched image presented

Table 1. Each question is rated on a scale of 1 (definitely not) to 5 (definitely yes). Five experts were asked to evaluate the results. All showed strong agreement for usefulness of the approach.

Question	E1	E2	E3	E4	E5	Mean	Stdv
Does the proposed image contain the information that you expected?	4	4	5	4	5	4.4	0.55
Would the information be helpful for your procedure?	5	5	4	4	4	4.4	0.55
Would this information change the way you would explore the oesophagus?	3	4	4	3	3	3.4	0.55
Do you think the approach would reduce the duration of exploration?	4	4	3	3	4	3.4	0.55
Are you likely to be interested in using such an application in the future?	5	5	5	5	5	5	0
Would you recommend it to a fellow colleague?	4	5	5	5	5	4.8	0.45
Rate the quality of information?	5	5	5	4	5	4.8	0.45



Fig. 4. The 1st column shows the frames from the live video, the 2nd column shows the closest match. Columns 3,4 and 5 present frames randomly selected from the 20 best matches obtained from the synchronization phase.

using our approach. We quantified the feedback of their experience using a questionnaire presented in Table. 1. The evaluation experiment clearly shows that the experts found the current system very useful for re-localizing the simulated biopsy sites. However, it was suggested that the matched image did not necessarily provide the ideal viewpoint (Fig 4a,b and e). Upon added investigation we observed that by considering ($\sim$) 20 best neighbours, the selection of best image can be refined to accommodate the best viewpoint. Fig. 4 shows a sample of these frames. While our current approach already provides the GI specialist with a more advanced guidance, we can further improve the frame selection by performing image analysis on these best neighbours.

5 Conclusion

This paper introduces a novel framework for combining images with instrument tracking information for inter-operative biopsy site re-localization. An application of this approach using EMTS for Barrett's oesophagus has been presented. To our knowledge, this is the first time such a system has been proposed in the literature. The system set-up is quite straightforward, it works in real-time (trajectory matching $\sim 100\text{Hz}$), can be used with minimal change to the operating room protocol and most of the offline steps can easily be automated. Moreover, because the system is scalable in time, the recordings can be shared between GI specialists without loss of information. Finally, since our method does not rely on the quality of images, it is robust to typical endoscopic image artefacts for inter-operative comparison. In future, we will tackle the issue of best viewpoint selection and the mapping of recorded regions of interest to the live view. Currently the accuracy of the approach relies on the repeatability of accurately localizing the sphincter inter-operatively by the GI specialist, we would also address this issue in the future.

Acknowledgement. This work is part of the “ARFlex” project supported by IHU, Strasbourg. We would like to extend our thanks to Dr. Hyunsoo Chung, Dr. Jérôme Huppertz, Prof. Gérard Gay and Prof. Michel Delvaux for their invaluable support.

References

1. Wang, K.K., Sampliner, R.E.: Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *The American Journal of Gastroenterology* 103(3), 788–797 (2008)
2. Allain, B., Hu, M., Lovat, L.B., Cook, R., Ourselin, S., Hawkes, D.: Biopsy site re-localisation based on the computation of epipolar lines from two previous endoscopic images. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 491–498. Springer, Heidelberg (2009)
3. Mountney, P., Giannarou, S., Elson, D., Yang, G.-Z.: Optical biopsy mapping for minimally invasive cancer screening. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 483–490. Springer, Heidelberg (2009)
4. Allain, B., Hu, M., Lovat, L.B., Cook, R.J., Vercauteren, T., Ourselin, S., Hawkes, D.J.: A system for biopsy site re-targeting with uncertainty in gastroenterology and oropharyngeal examinations. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010, Part II. LNCS, vol. 6362, pp. 514–521. Springer, Heidelberg (2010)
5. Atasoy, S., Glocker, B., Giannarou, S., Mateus, D., Meining, A., Yang, G.-Z., Navab, N.: Probabilistic region matching in narrow-band endoscopy for targeted optical biopsy. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009, Part I. LNCS, vol. 5761, pp. 499–506. Springer, Heidelberg (2009)
6. Atasoy, S., et al.: Endoscopic video manifolds for targeted optical biopsy. *IEEE Transactions on Medical Imaging* 31(3), 637–653 (2012)
7. Bashar, M.K., et al.: et al. Automatic detection of informative frames from wireless capsule endoscopy images. *Medical Image Analysis* 14(3), 449–470 (2010)