Inter-Operative Biopsy Site Relocalization in Endoluminal Surgery

Anant S. Vemuri, Stéphane Nicolau, Adrien Sportes, Jacques Marescaux, Luc Soler, Nicholas Ayache

Inter-Operative Biopsy Site Relocalization in Endoluminal Surgery

Anant Suraj Vemuri, Stephane Nicolau, Adrien Sportes, Jacques Marescaux, Luc Soler, and Nicholas Ayache

***Abstract*—Barrett's oesophagus, a pre-malignant condition of the oesophagus has been on a rise in the recent years. The standard diagnostic protocol for Barrett's involves obtaining biopsies at suspicious regions along the oesophagus. The localization and tracking of these biopsy sites "inter-operatively" poses a significant challenge for providing targeted treatments and tracking disease progression. This paper proposes an approach to provide guided navigation and relocalization of the biopsy sites using an electromagnetic tracking system. The characteristic of our approach over existing ones is the integration of an EM sensor at the flexible endoscope tip, so that the endoscopic camera depth inside the oesophagus can be computed in real-time, allowing to retrieve and display an image from a previous exploration at the same depth. We firstly describe our system set-up and methodology for inter-operative registration. We then propose three incremental experiments of our approach. First, on synthetic data with realistic noise model to analyze the error bounds of our system. The second on *in-vivo* pig data using an optical tracking system to provide a pseudo ground-truth. Accuracy results obtained were consistent with the synthetic experiments despite uncertainty introduced due to breathing motion, and remain inside acceptable error margin according to medical experts. Finally, a third experiment designed using data from pigs to simulate a real task of biopsy site relocalization, and evaluated by 10 GI experts. It clearly demonstrated the benefit of our system towards assisted guidance by improving the biopsy site retrieval rate from 47.5% to 94%.**

***Index Terms*—Computer assisted intervention, gastro-intestinal (GI) endoscopy, Barrett's Oesophagus, biopsy relocalization, electromagnetic tracking, video synchronization.**

I. Introduction

BARRETT'S oesophagus is the pre-malignant dysplastic condition that is considered to be the cause for oesophageal adenocarcinoma [1]. The sequence of events from Barrett's oesophagus to adenocarcinoma has several steps, encompassing low grade intra-epithelial neoplasia, and high grade intra-epithelial neoplasia. Since its evolution spans many years, endoscopic surveillance for patients with Barrett's oesophagus has been advocated to detect neoplasia at early and curable stages. In recent years endoscopic imaging techniques have improved greatly. However, according to the Seattle protocol, the gold standard approach of comprehensive endoscopic biopsy protocol involves direct sampling from suspicious areas in combination with systematic random four quadrant biopsies (4QBs) every 2 cm along the length of the Barrett's segment [2], [3]. In a typical surveillance endoscopy the primary objective is the inter-operative relocalization of these biopsy sites to guide the treatment. Often relocalization is performed using the markings on the endoscope which is not accurate for three reasons. 1) The markings on the endoscope are every 5 cm; 2) Depending on the position of the patient the length inside the oesophagus can change between two procedures; 3) The orientation of the endoscope is not fixed inside the oesophagus (particularly the roll angle), so the accurate location of biopsy is hard to determine. The aim of this paper is to provide a novel solution to such inter-operative relocalization in endoluminal procedures involving surveillance endoscopies.

Anant S. Vemuri is with the IHU, Strasbourg at 1, Place d'Hopital, 67000 Strasbourg, France. He is also affiliated with the INRIA, Sophia Antipolis, France (e-mail:anant.vemuri@ihu-strasbourg.eu).

Stephane Nicolau, Jacques Marescaux and Luc Soler are with the IRCAD, 67000 Strasbourg, France (e-mail: stephane.nicolau@ircad.fr).

Adrien Sportes is with the CHRU, 67000 Strasbourg, France (e-mail: adrien.sportes@me.com).

Nicholas Ayache is with INRIA, Sophia Antipolis 2004 Route des Lucioles, 06902 Valbonne, France (e-mail: nicholas.ayache@ircad.fr).

A. Related Work

"*Inter-operative*" localization in surgery is a challenging problem. In gastroenterological procedures, the surgeon relies on the anatomical landmarks (such as vasculature) and the distance from the gastro-oesophageal junction as recorded using the markings on the endoscope to determine his or her previous location. This approach heavily relies on the surgeon's experience to track a disease. One of the most significant difficulties arise from application of localized diagnostic tools such as confocal microscopy [4], [5] or optical coherence tomography [6]. These tools, image at a surface area of about 0.4 mm^2 and the surface to be explored is more than 1200 mm^2 (note that a typical oesophagus diameter is about 15$\sim$20 mm). The lack of an approach to locate the previous biopsy sites precisely has limited the localized tracking and treatment of Barrett's oesophagus.

Several approaches to track the biopsy sites "*intra-operatively*" exist [7]–[11]; each of them relying on the recovery of the 3D structure of the anatomy and/or the repeatability of image features, 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. [10] employ epipolar geometry, Mountney et al. [7] propose a simultaneous localization and mapping (SLAM) based method. Atasoy et al. [8], propose a probabilistic region matching approach in narrow-band images by using feature matches obtained from affine invariant anisotropic feature detector. More recently in [9] propose to formulate the relocalization as an 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 to relocalize

the biopsy sites. However, it is unclear how they provide spatial relations of the extracted segments inter-operatively. Hence, the application of their result in the clinical context for relocalization seems limited. All the previously described approaches have relied entirely on image based tracking or image feature cues to provide tracking in GI procedures. However, inter-operatively, temporal changes in tissue texture and large deformation of the endoscopic scene are observed. Hence repeatability of feature extraction, matching and tracking poses a significant challenge.

To overcome some of these difficulties we propose here a method utilizing an Electromagnetic tracking system (EMTS) for inter-operative re-localization. EMTS have been used previously for medical applications [12]–[14]. EMTS offers a distinct advantages; firstly it does not suffer from line-of-sight, secondly, it can be easily integrated into flexible instruments and thus employed for tracking under deformable conditions. Several studies have been published demonstrating the accuracy of EMTS. However depending on the surgical environment, it is susceptible to distortions, but methods have been proposed in literature [15], [16] to overcome these challenges.

In this paper propose a solution for two aspects of the re-localization problem, firstly *w.r.t* position in the oesophagus and secondly the orientation at that position. We present here a significant extension on an earlier work [17], where we had originally introduced this approach. The rest of the paper is organised as follows. In II, we highlight the limitations of our previous approach and propose usage of two supplementary sensors in our set-up, along with a dedicated registration method, to overcome the problem. Our system being only qualitatively evaluated with a questionnaire in [17]; we propose here two sets of quantitative experiments on synthetic and real-data from pigs to rigorously evaluate our guidance platform. Finally, in IV, we provide a realistic qualitative assessment of our system on pigs by clinical experts using biopsy sites simulated by coagulation device.

II. System Setup and Approach

The system setup consists of an NDI Aurora® tracker. It comprises of a box-type electromagnetic field emitter (EMFE), three EM sensors (two Aurora 6-DOF Cable Tools and an Aurora 6-DOF Catheter, Type 2), an EM control interface, a titanium arm to mount the EMFE and a dual channel Karl Storz flexible endoscope. The EMFE is a box which generates an EM field to track the sensors. The range of the EMFE is a cubic region of volume 50 x 50 x 50 cm^3 in front of it. An EM sensor is a copper wire shielded in a rubber sheath of 2 mm (diameter). The EM tracker is connected to the EMFE and EM sensors and provides an interface with the computer. The EM sensors provide 6-DOF pose information which includes 3D position and 4D quaternion. The EMFE is mounted onto the titanium arm and is positioned $\sim$ 6 cm above the desired working volume.

Our navigation system aims to provide inter-operative video synchronization for a live intervention with an intervention performed a few months prior. The goal of our video synchronization approach is; given a video recording from a previously conducted intervention on a patient, we display a video frame from the previous intervention, that corresponds to the location of the endoscope in the oesophagus during the live intervention.

In this section, II-A discusses the limitations of [17] and proposes a method to overcome them. II-B describes the data acquisition and details the notation used in the rest of the paper. Section II-C introduces the inter-operative registration, II-D presents the estimation of orientation difference between matched locations in the trajectories and II-E briefly describes the interface presented to the clinician.

A. Limitations of our previous approach

In [17], we used a single sensor, inserted into the endoscope channel, which posed difficulties in clinical settings. Inter-operative registration was performed using the complete oesophagus trajectory from the exploration with the gastro-oesophageal junction (GOJ) as the anatomical landmark. This required the expert to necessarily reach the GOJ first before receiving any guidance. This however is not always possible; in case of oesophagus stenosis due to a tumour, for instance. It is also a constraint since the clinician may come across lesions before reaching the sphincter, and may prefer to treat them immediately. Additionally, the combined variation in head and neck orientation introduces an uncertainty that is difficult to model. In fact, the patient is not always lying in supine position and can be set in decubitus lateral, depending on the surgical conditions making the orientation compensation (II-D) from the 3D endoscope trajectory alone unreliable.

Here we propose an alternative using two additional sensors. One sensor is attached at the Suprasternal (or Jugular) notch. It is the anatomical landmark located at the superior border of the manubrium of the sternum, between the clavicular notches. The second sensor is attached along the sternum about 10-15 cms below the Jugular notch. The third sensor is inserted in one of the channels of the endoscope, as in our previous work. Fig. 1 shows the placement of the sensors on the patient.

B. Data acquisition

The process of acquiring data during an intervention involves a synchronized capture of video from the endoscope and tracking data from the EMTS. For each corresponding position of the endoscope in the oesophagus, the corresponding position of (all three) sensors is recorded into a database. It should be noted that the accuracy of the EMTS falls when tracking fast moving objects [18]. Hence, any tracking data above a certain velocity threshold (10 mm/sec, as was suggested in [18]) was discarded. We define a recording of an intervention ($\mathcal{I}$) as a collection of the following four items;

1) The trajectory of the endoscopic sensor in the oesophagus $\mathcal{T} = \{P_i^E \in \mathbb{R}^3; R_i^E \in SO(3),\ i \in [1, N]\}$ for N points recorded on the oesophagus trajectory, where; P_i^E = position and R_i^E = orientation as generated by the EM sensor.
2) A collection of sensor positions for the external landmarks on the Jugular notch ($\mathcal{L}_A$) and the sternum ($\mathcal{L}_B$) for each corresponding position (P_i^E) of the endoscopic

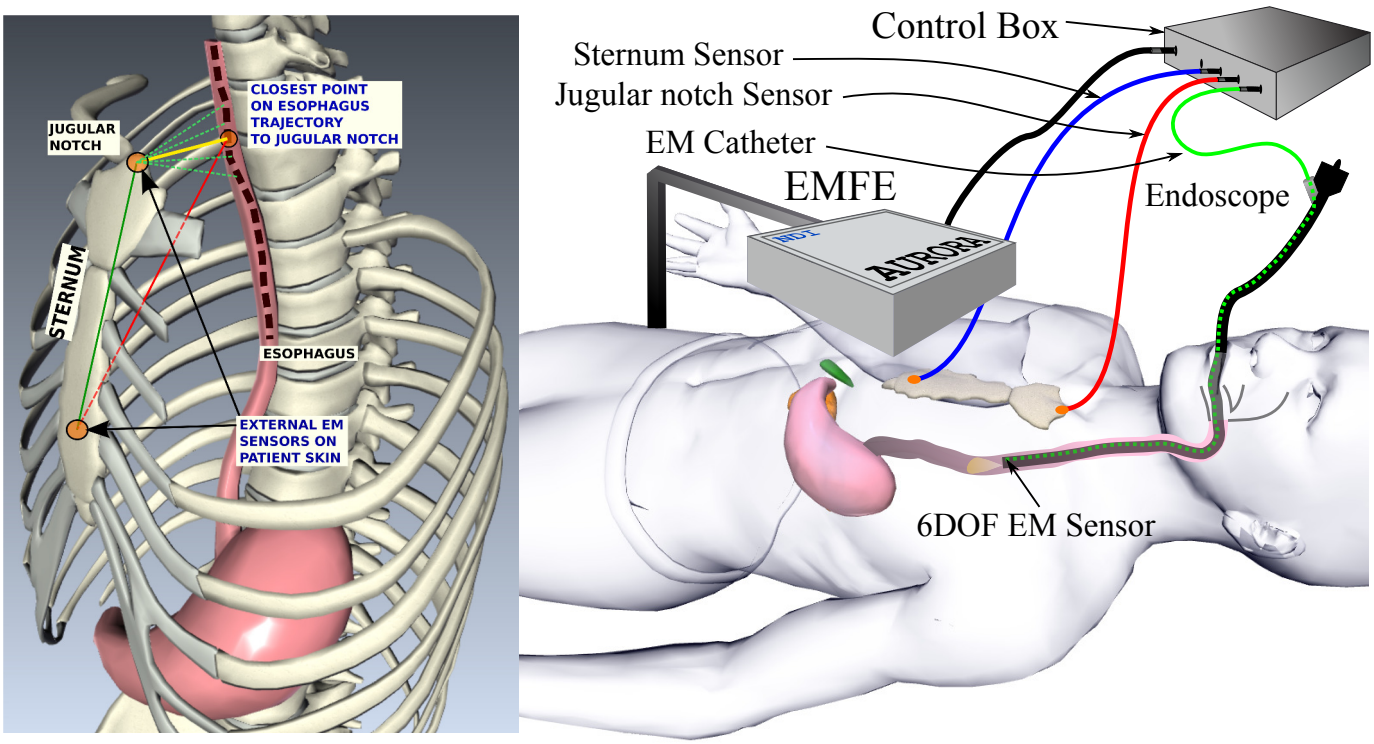

Fig. 1. Placement of sensors (left) and System setup (right).

sensor; $\mathcal{L}_{A,B} = \{P_i^A, P_i^B \in \mathbb{R}^3; R_i^A, R_i^B \in SO(3),\ i \in [1, N]\}$.

3) Set of images recorded $\boldsymbol{I} = [I_1, ...I_N]$ in the oesophagus at each position of the endoscopic sensor.
4) Tag for the entry, $k \in [1, N]$ in $\mathcal{T}$ and $\mathcal{L}_{A,B}$, which corresponds to the shortest distance from the endoscopic sensor to the sensor at the jugular notch. As we move along the oesophagus, the distance of the point P_i^E from P_i^A first decreases monotonically and then increases. The index k marks the position at which this distance is a minimum. k identifies a reference plane ($\mathcal{D}$) formed by the three sensors $\mathcal{D} = \{P_k^E, P_k^A, P_k^B; \in \mathbb{R}^3\}$ as show in in Fig. 1. The entry k is computed at the beginning of the procedure. For notational simplicity we drop the suffix k and represent reference plane as $\mathcal{D} = \{\mathbf{P}^E, \mathbf{P}^A, \mathbf{P}^B; \in \mathbb{R}^3\}$. $\mathcal{D}$ is used for inter-operative registration as will be explained in II-C.

The first intervention acquisition usually contains many back and forth movements inside the oesophagus, and the entire recording thus contains plenty of quasi-superimposed trajectories (within a distance of 1-2 mm). Here we assume that the video has been sorted and only single smooth continuous trajectory has been extracted and annotated. This preliminary work of sorting the first exploration is not the focus of this paper and will be tackled in future work. Further details will be discussed in V. In our experiments smooth motion was performed to avoid such a situation.

C. Inter-operative synchronization

To synchronize (live) intervention $\mathcal{I}_1$, which is updated in real-time, with $\mathcal{I}_2$ (intervention performed 3 months prior), the reference frames of the EMTS in these two interventions have to be registered. We accomplish this by performing a rigid registration between $\mathcal{D}_1$ and $\mathcal{D}_2$. However, there is an uncertainty in the positioning of $\mathbf{P}_l^B$ and $\mathbf{P}_l^E$ for plane $\mathcal{D}_l, l = [1, 2]$. Hence instead of performing a point based registration, we represent the plane $\mathcal{D}_l$ using a reference point and two vectors as,

$$\mathcal{D}_l \equiv [\mathbf{P}_l^A; \mathbf{n}_1^l = \widehat{\mathbf{P}_l^A\mathbf{P}_l^B}; \mathbf{n}_2^l = \mathbf{n}_1^l \times \widehat{\mathbf{P}_l^A\mathbf{P}_l^E}] \quad (1)$$

We use this representation to perform the inter-operative registration ($\mathcal{I}_1$ with $\mathcal{I}_2$; $T \in SE(3)$), which is performed on-line, once at the beginning of the procedure. So, there are no additional steps to initialize the video synchronization. If there is a significant movement of the patient as noted from change in $\mathcal{D}_l$, k, in item (4), is recomputed and registration is performed again.

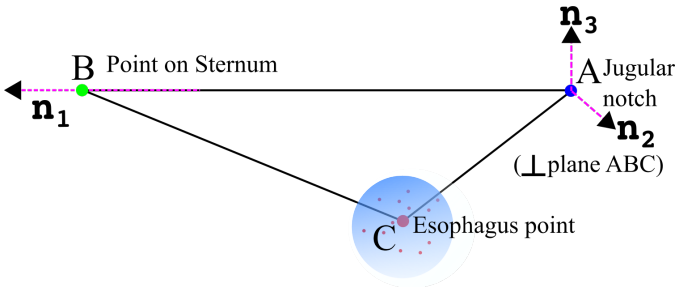

Fig. 2. Formation of the reference frame for an intervention $\mathcal{I}_l$. The point C is computed as barycentre of points in a sphere of radius 5mm (chosen empirically) centred at the closest point found.

After the registration of the reference frames of the patient in $\mathcal{I}_1$ with $\mathcal{I}_2$, a spatial correspondence between the sensor position in live from $\mathcal{T}_1 \in \mathcal{I}_1$ is made with $\mathcal{T}_2 \in \mathcal{I}_2$, by a search for the closest neighbour (in euclidean space). Due to small movement of the endoscope tip (yaw and pitch motion) along the cross-section of the oesophagus, the trajectories are not completely smooth. If we use only the euclidean distance to compute the nearest neighbour, it can lead to false matches along the oesophagus as depicted in Fig. 3. Globally, the trajectory of the endoscope in the oesophagus is sufficiently linear along $\mathbf{n}_1^l$, computed in Eq. (1). Hence to alleviate the false matches during synchronization, the search space is constrained along $\mathbf{n}_1^l$ ($= 4mm$ chosen empirically) using $\mathcal{T}_l^{projected}$ as in Eq. (2) and as shown in Fig. 3.

$$\mathcal{T}_l^{projected} = \langle \mathcal{T}_l, \mathbf{n}_1^l \rangle \qquad l \in [1, 2] \tag{2}$$

The closest neighbour then gives the corresponding best matching image of the region in the oesophagus. Let the function for finding the closest neighbour be denoted by f. In particular, the matched images that were tagged (performed offline by the GI expert) to contain locations of biopsy sites, provide the GI specialist a more localized region for review. Fig. 4 presents an example result of the synchronization.

D. Orientation Correction

Depending on the orientation of the patient (decubitus lateral or supine), the EMTS setup, and the orientation of the endoscope in the oesophagus; the view inside the oesophagus could be rotated when matching the live and a recorded interventions. After the inter-operative synchronization, considering a match for the u^{th} point $\{P_{u,1}^E R_{u,1}^E\} \in \mathcal{I}_1$ and v^{th} point $\{P_{v,2}^E R_{v,2}^E\} \in \mathcal{I}_2$ was obtained; we propose to estimate the correction in orientation of image $I_v^2 \in \boldsymbol{I}_2$ *w.r.t* image $I_u^1 \in \boldsymbol{I}_1$. Now for $R_{u,1}^E$ and $\tilde{R}_{v,2}^E$ (corrected with T from II-C), the residual orientation can be computed as,

$$R = ({R_{u,1}^E}^{-1})\, \tilde{R}_{v,2}^E \tag{3}$$

The image orientation correction can then be estimated from the roll angle (θ_c) by decomposing R into Euler angle formulation. However, this decomposition is not stable; since the estimation of θ_c depends on the complete decomposition into yaw-pitch-roll. Since, during an intervention the endoscope can perform yaw and pitch motion which would affect the estimated roll angle in case of a direct estimation from R. An ideal case would be to dissociate the effect of yaw-pitch combination when estimating the roll angle (θ_c). To achieve this we propose the following approach.

Considering the component vectors of the orientation, $R_{u,1}^E = (\mathbf{n}_x^u\ ,\ \mathbf{n}_y^u\ ,\ \mathbf{n}_z^u)$ and $\tilde{R}_{v,2}^E = (\mathbf{n}_x^v\ ,\ \mathbf{n}_y^v\ ,\ \mathbf{n}_z^v)$. To effectively estimate the rotation about $\mathbf{n}_z^u$ by reducing the effects due to angle between $\mathbf{n}_z^u$ and $\mathbf{n}_z^v$, we compute firstly the projections $\mathbf{n}_x^{v,proj}$ and $\mathbf{n}_y^{v,proj}$ on the $\mathbf{n}_x^u\mathbf{n}_y^u$ plane, as shown in Fig. 5 and compute the roll angle; $\theta_c = (\alpha + \beta)/2$, about $\mathbf{n}_z^u$ required to align vectors $(\mathbf{n}_x^{v,proj},\ \mathbf{n}_y^{v,proj})$ with $(\mathbf{n}_x^u,\ \mathbf{n}_y^u)$. This gives the estimate for orientation difference between the endoscope position in live intervention $\mathcal{I}_1$ and the corresponding match computed in II-C.

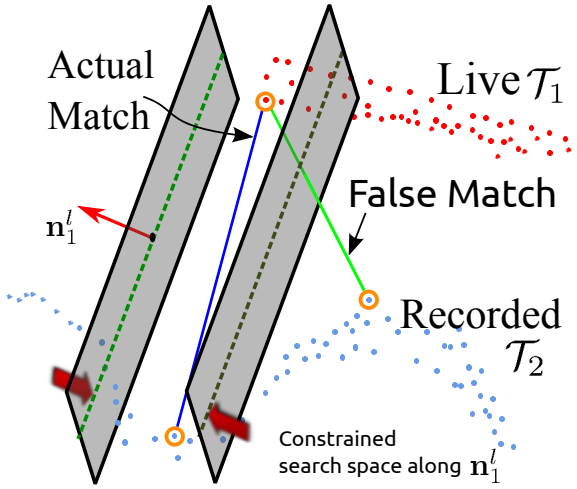

Fig. 3. False match using euclidean distance. The dotted line indicates the closest matched point using euclidean distance. However, the actual match is further away as shown above.

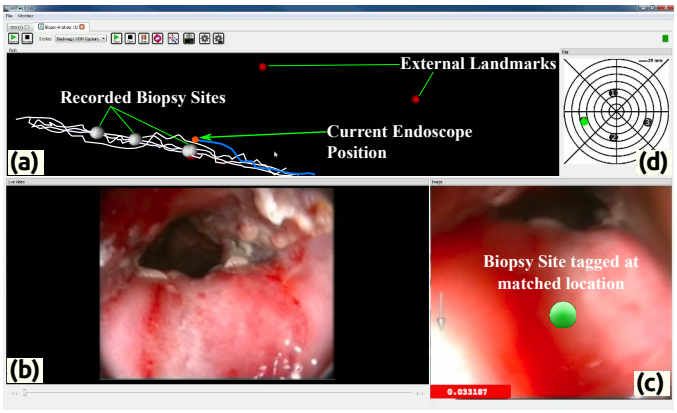

Fig. 4. Application view as presented to the GI expert. (a) Frame indicating the 3D view containing the trajectory of the sensor in the oesophagus. (b) Live video frame from the endoscope. (c) The corresponding matched image from a previous intervention. (d) Displays the map of biopsy sites from intervention $\mathcal{I}_2$ in dark black circles and the current position and orientation of the endoscope in the oesophagus by bright green circle.

E. Application interface

Fig. 4 presents a view of the application available to the clinician during the procedure. There are four parts to the interface; 1) The 3D-view, depicting the recorded 3D position of the EM sensors during the live intervention $\mathcal{I}_1$ and previously recorded intervention $\mathcal{I}_2$, 2) the live endoscopic image $\{I_u \in \boldsymbol{I}_1 : \boldsymbol{I}_1 \in \mathcal{I}_1\}$, 3) the corresponding matched image $\{I_v \in \boldsymbol{I}_2 : \boldsymbol{I}_2 \in \mathcal{I}_2\}$ and, 4) the map of the biopsies taken during $\mathcal{I}_2$. The biopsy map is an important aspect of the interface since it facilitates plotting the biopsy points in a context clearly understood by a clinician. We briefly explain here the generation of this biopsy map.

It consists of concentric circles, with each annuli depicting regions of depth of 5 cms inside the oesophagus. The circle in the centre depicts the z-line (gastro-oesophageal junction). During an intervention $\mathcal{I}_2$, each biopsy point is first mapped onto the corresponding annulus. The orientation angle θ_{map}, is estimated as the angle between the patient's reference frame

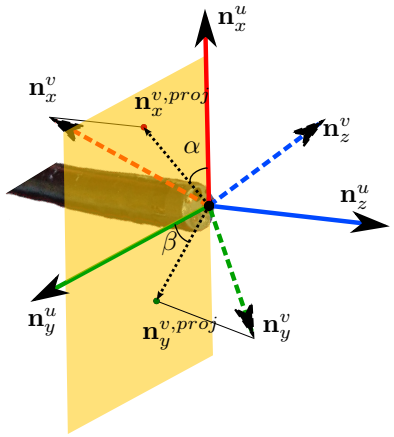

Fig. 5. Figure showing the approach to obtain roll angle about the endoscope axis by negating the effects of yaw and pitch motions. We use this for orientation correction between two interventions. Note that the orientation directions of $\mathbf{n}_z^u$ and $\mathbf{n}_z^v$ have been exaggerated to explain the point.

$\mathcal{D}_2$ and the reference frame of the EM sensor at the biopsy site. The approach is similar to one presented earlier in II-D, where the EM sensor reference frame is projected onto the patient's reference frame at $\mathcal{D}_2$. In Fig. 4 all the black smaller circles indicate biopsy sites performed during $\mathcal{I}_2$.

To incorporate sensor position and orientation from $\mathcal{I}_1$, firstly, the position of the endoscopic sensor from $\mathcal{I}_1$ is localized in the corresponding annulus. If a biopsy site exists in this region that is closest to the live endoscope position, it is highlighted by a change in colour. Secondly, to incorporate the orientation correction in the clinical work-flow, the absolute orientation of the endoscope *w.r.t* the patient reference frame is computed as $\theta = \theta_{map} + \theta_c$, where θ_{map} is the estimated orientation of the closest biopsy point to live endoscope position. As the gastroentrologist rotates the endoscope inside the lumen, $\theta_c \rightarrow 0$ and the live view I_u^1 is aligned to fit the matched view I_v^2. The orientation correction is incorporated in the map shown in Fig. 4(d). This enables the gastroentrologist to physically rotate the endoscope to align the live view; since rotating the endoscopic image itself would be visually difficult to interpret.

III. Quantitative Evaluation

In this section we perform our quantitative evaluation of the above described system in two phases. First, by generating synthetic data using a gaussian noise model for the uncertainty in our system; followed by experimental evaluation on real-data. This section is organised as follows. III-A discusses the sources of error, and presents the experimental set-up to measure the uncertainty in $\mathcal{D}$, the generation of synthetic data, the characterization of the final error propagated to video synchronization and results. III-B describes the experiments on real-data collected from pigs.

A. Evaluation on synthetic data

To generate realistic synthetic data, it is essential to identify the various sources of error in our application domain. There are five main sources of uncertainty in our system;

(*a*) In the placement of landmarks (external and internal), which corresponds to the uncertainty in the estimation of $\mathcal{D} = \{\mathbf{P}^E, \mathbf{P}^A, \mathbf{P}^B \in \mathbb{R}^3\}$

(*b*) Deformation in the patient during intervention and between two interventions. This can occur due to change in the posture in which the patient is lying in on the table (decubitus lateral or supine) and also due to the change in the positioning of the neck.

(*c*) Deformation due to breathing. Typically under general anaesthesia, the breathing is steady, however under rare conditions of heavy breathing there can be extension of the oesophagus up to 1-2 cms.

(*d*) Tracking errors due to EMTS.

(*e*) Position of the endoscope tip in the cross section of the oesophagus.

In this section, we discuss how we empirically measure and model the errors (*a*) and (*e*) and evaluate their effects in the final matching of the trajectories. The effect of (*b*) would not be significant since the oesophagus does not exhibit longitudinal elongation. However the important effect would be due to the position of the neck, which would not affect our approach since it does not rely completely on the oesophagus trajectory. The effect due to breathing (*c*) is non-linear along the oesophagus length and is maximum closest to the GI junction. However, we have not modelled this error in our current work, but it can be easily incorporated as discussed in V. The error (*d*) due to EMTS tracking has been explored extensively in [14], [18]–[20] to characterize the distribution of error in the working volume. However, the effect on the tracking by EMTS is dependent on the environment and varies non-linearly in the working volume. In our case the working volume of the EMTS is very small (25 x 10 x 10 cm^3) and the effect of distortions in EM field is minimal and hence can be neglected. In synthetic evaluation we do not model the errors (*b*), (*c*) and (*d*), which, however, have been implicitly considered in real-data, hence their effect has been studied in combination with (*a*). Due to the way the ground-truth is defined (III-A3) the effect of uncertainty (*e*) has not be included in the final error computation. However, it does not significantly impact our analysis since it only affects the viewpoint in the oesophagus.

1) Measurement of uncertainty in $\mathcal{D}$: The registration of two interventions $\mathcal{I}_1$ and $\mathcal{I}_2$ relies on the registration of their corresponding reference planes $\mathcal{D}_1$ and $\mathcal{D}_2$. We can model the uncertainty in $\mathcal{D}$ by measuring the uncertainty in the positioning of the landmarks at, *a*) Jugular notch; $\mathbf{P}^A$ *b*) landmark on the sternum; $\mathbf{P}^B$ and *c*) the closest point on oesophageal trajectory to $\mathbf{P}^A$; $\mathbf{P}^E$. We model the uncertainty at each of these landmarks using a normal distribution $\mathcal{N}(0, \Sigma)$. The experiments for uncertainty measurement were sub-divided in two independent steps. First for the external landmarks we conduct empirical study on 18 test subjects. Each subject was positioned on the surgical table once, in the decubitus lateral position and the external sensors on the jugular notch and sternum were repositioned 12-15 times by our experts. To be able to average the uncertainty of all subjects, a third fixed sensor was attached to a platform so that it was perpendicular to the subject's chest at the jugular notch to provide a frame of reference for the patient (Fig. 6). All the sensor positions were then transformed to the reference frame formed at the jugular notch and the positional uncertainty

$(\Sigma_{\mathbf{P}^A}, \Sigma_{\mathbf{P}^B})$ on the landmarks was measured in this frame of reference, for all subjects. In the second step, we conducted the experiment on a phantom model of the oesophagus. After placing the external landmarks, we estimate the closest point in oesophagus trajectory in 108 trials. The process was repeated for several different positions of sensor placed at the jugular notch, sternum sensor and we observed a similar uncertainty matrix in each case. It led us to conclude that the uncertainty in the point $\mathbf{P}^E$ was independent of uncertainty in $\mathbf{P}^A$ and $\mathbf{P}^B$. Finally, the points were transformed to the reference frame at the jugular notch so that the computed covariance matrix $(\Sigma_{\mathbf{P}^E})$ was independent of the position of the EMFE *w.r.t* the patient.

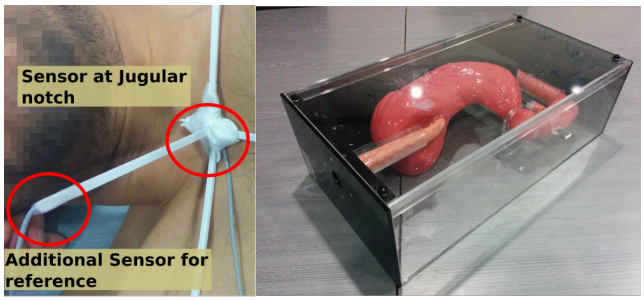

Fig. 6. (Left) Using the second sensor at the jugular notch for providing a reference frame when measuring uncertainty in $\mathcal{D}$. (Right) Oesophagus and stomach phantom, used for measuring uncertainty in the oesophagus cross-section.

2) Measuring uncertainty in endoscope tip: An additional source of uncertainty is the position of the tip of the endoscope inside the cross-section of the oesophagus. We recorded 100 back and forth trajectories and selected 10 regions of 2 mm thickness and computed the uncertainty of points in the selected cross-sections. We model the uncertainty of points in this region as a normal distribution $\mathcal{N}(0, \Sigma_{CS})$. The experiment was conducted on oesophagus phantom Fig. 6. Eq. (4) presents the empirically measured uncertainties.

$$\Sigma_{\mathbf{P}^A} = \begin{bmatrix} 8.5 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 12.5 \end{bmatrix} \quad \Sigma_{\mathbf{P}^B} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 13.759 & 0 \\ 0 & 0 & 94.5 \end{bmatrix}$$
$$\Sigma_{\mathbf{P}^E} = \begin{bmatrix} 26 & 0 & 0 \\ 0 & 25 & 0 \\ 0 & 0 & 109 \end{bmatrix} \quad \Sigma_{CS} = \begin{bmatrix} 25.7 & 0 & 0 \\ 0 & 22.3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{4}$$

3) Generation of synthetic data: Using a phantom model Fig. 6, we performed a recording $\mathcal{I}_1$ as was explained earlier in II-B. To generate synthetic data ($\mathcal{I}_2$) at a new position, we apply a known transform T_{GT}; generated using a random rotation vector and a random translation (± 1000) in x, y and z direction to $\mathcal{I}_1$. Then, using the estimated uncertainties $\mathcal{N}(0, \Sigma_{\mathbf{P}^A})$ and $\mathcal{N}(0, \Sigma_{\mathbf{P}^B})$ a noisy version $\tilde{\mathcal{I}}_2$ was created. Let,

$$\begin{aligned} [\ \tilde{\mathbf{P}}_2^A &= T_{GT}\mathbf{P}_2^A + \mathcal{N}(0, \Sigma_{\mathbf{P}^A}), \\ \tilde{\mathbf{P}}_2^B &= T_{GT}\mathbf{P}_2^B + \mathcal{N}(0, \Sigma_{\mathbf{P}^B}), \\ \tilde{\mathcal{T}}_2 &= T_{GT}\mathcal{T}_1 + \mathcal{N}(0, \Sigma_{CS})\ ] \in \tilde{\mathcal{I}}_2 \end{aligned} \tag{5}$$

be the external landmark positions. Now using $\tilde{\mathbf{P}}_2^A$, the new closest point $(\hat{\mathbf{P}}_2^E)$ on the oesophageal trajectory $\tilde{\mathcal{T}}_2$ was computed. $\tilde{\mathbf{P}}_2^E = \hat{\mathbf{P}}_2^E + \mathcal{N}(0, \Sigma_{\mathbf{P}^E})$, a noisy version of $\mathbf{P}_2^E$ was generated. Finally, for $\tilde{\mathcal{I}}_2$ the new reference plane $\tilde{\mathcal{D}}_2 = \{\tilde{\mathbf{P}}_2^E, \tilde{\mathbf{P}}_2^A, \tilde{\mathbf{P}}_2^B; \in \mathbb{R}^3\}$ is obtained. To eliminate any bias in the registration using $\mathcal{D}_1$ and $\tilde{\mathcal{D}}_2$; gaussian noise was applied to $\mathcal{D}_1$ in the same way as described above to obtain $\tilde{\mathcal{D}}_1$. Using the above data, $f : \mathcal{T}_1 \rightarrow \tilde{\mathcal{T}}_2$ is bijective (f defined in II-C), which provides the necessary ground-truth to estimate the error in synchronization. We used 7 recorded $\mathcal{I}_1$, which were transformed using 30 different T_{GT}. For each $\mathcal{I}_2$ generated by transformed $\mathcal{I}_1$, 1500 $\tilde{\mathcal{I}}_2$ samples were generated. Thus, for each base trajectory $\mathcal{I}_1$, 45000 synthetic *trajectory matches* were computed.

4) Error Measurement: The statistic that is important from a clinician's viewpoint is the error in the final synchronization of images. Computing a metric to measure similarity between matched images from $\mathcal{I}_1$ and $\tilde{\mathcal{I}}_2$ would not yield much information since the tissue in the oesophagus may undergo a drastic change due to natural oesophagus muscular contractions and tissue structural changes due to disease propagation. Here we characterize the metric as the error in depth along the oesophagus. As shown in Fig. 7, since $\mathcal{T}$ exhibits linearity along $\mathbf{n}_1$ (computed in Eq. 1), hence the 3D position error was projected on to $\mathbf{n}_1$ to obtain the depth error along the oesophagus; normalizing the position in the oesophagus in the range [0=Beginning of the oesophagus, 1=Close to the gastro-oesophageal junction].

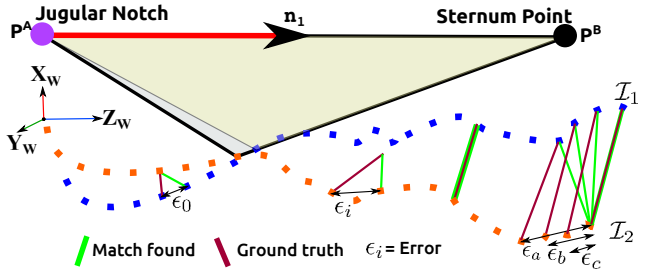

Fig. 7. This figure shows the approach to error estimation when the reference planes are aligned with the world reference frame. The error in-effect becomes the error in depth inside the oesophagus.

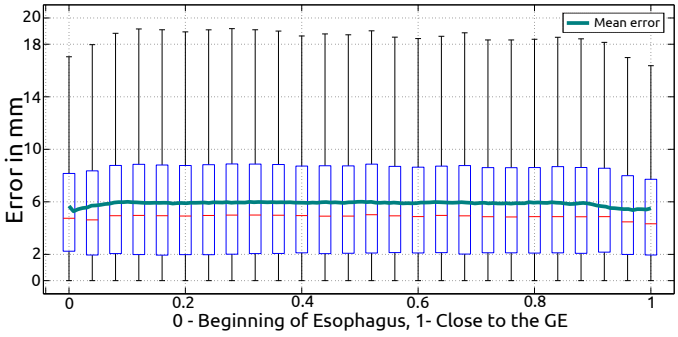

Fig. 8. Absolute value of error along the oesophagus length for the each of the 7 $\mathcal{I}_1$ trajectories matched with their corresponding 45000 synthetic $\tilde{\mathcal{I}}_2$ trajectories. We observe a constant offset due to the error in the translational component of registration.

5) Results: Each $\mathcal{I}_1$ is matched with the $\tilde{\mathcal{I}}_2$ synthetic samples generated. From the error measurement presented in Fig. 8, a consistent bias in the trajectory due to a mismatch along $\mathbf{n}_1$ can be observed. This bias remains fairly constant along the trajectory. On average, we observe a roughly constant offset in the range (5-6 mm). Practically, this is due to the translational part of the registration error. Indeed, rotational error is small and has a negligible influence on the depth estimation, whereas translational error along Z has a direct influence on depth estimation, and mainly depends on the (simulated) error in positioning the landmarks at the jugular notch $\mathbf{P}^A$, the variance of which along $\mathbf{n}_1^l$ is $12.5mm^2$. When adding variance on both $\tilde{\mathbf{P}}_1^A$ and $\tilde{\mathbf{P}}_2^A$, we obtain a standard deviation of $5mm$, which corresponds to the reported registration error. Hence, we can consider a single point along the trajectory to provide sufficient information about the distribution of error. We observe that from Fig. 9 that the error is within 16 mm in 96% of the cases. The synchronization error observed is within acceptable limits since in effect it finally corresponds to a manageable variation in image matches inside the oesophagus Fig. 10. In Fig. 7 we observed a decrease on either ends of the error curve. This is due to the fact the towards the end of the trajectory, the error due to mismatch decreases, as can be seen from ϵ_a, ϵ_b and ϵ_c in Fig. 7. The maximum error observed is about 4.9 cms, corresponds to the largest observed error in 315000 matches. It is an unlikely situation to occur in reality that is reported in Fig. 11. As it can be seen, the error in $\tilde{\mathbf{P}}_2^E$ and $\tilde{\mathbf{P}}_2^B$ are unrealistically large. This can happen only in situations where the patient is breathing heavily. However, in most procedures the patients are placed under general anaesthesia and the breathing is very steady. Hence we believe that this effect would not really be encountered in a real scenario. But, such a case can be tackled with using a more robust registration model which we discuss briefly in V.

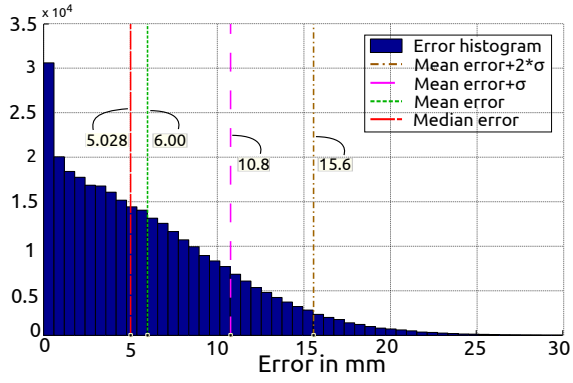

Fig. 9. Error at a single point on the oesophagus length for 315000 synthetic trajectories.

B. Evaluation on Real Data

We performed the quantitative evaluation on data collected on pigs. To our original set-up for data acquisition discussed in II, an optical tracking system (OTS) was included to provide the necessary ground-truth for the pose of the EMFE. An OT marker M_1 was fixed on the EMFE as shown in Fig. 12. This section firstly presents our approach to do the hand-eye calibration of OTS and EMTS. Secondly, using the additional information from OTS we provide a description of the methodology to establish ground-truth in the collected data to compute the error (as was described in III-A4), which is followed by the results.

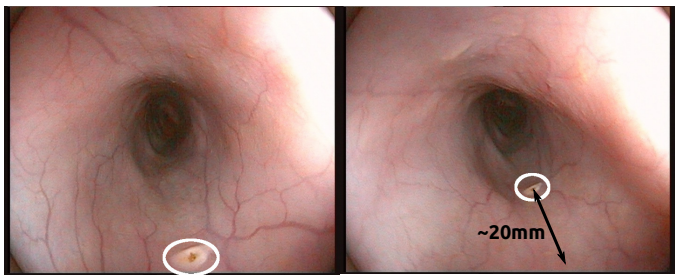

Fig. 10. Approximate distance measure inside the oesophagus translated to displacement in image space.

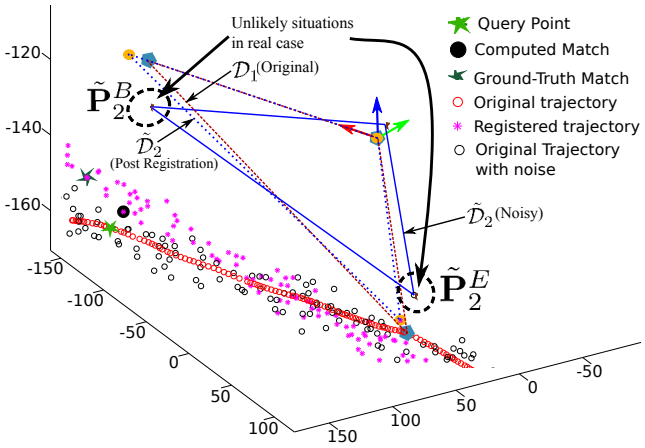

Fig. 11. Case of maximum error of $\approx$ 4.9 cm. It can be seen that $\tilde{\mathbf{P}}_2^E$ and $\tilde{\mathbf{P}}_2^B$ is larger than that encountered in a real scenario. In this particular simulated case the error in $\tilde{\mathbf{P}}_2^E$ causes it to be tending into the patient body and the noise in $\tilde{\mathbf{P}}_2^B$ causing it to be further away from the oesophageal trajectory. A registration using these points leads to an error in rotation along an axis perpendicular to the plane $\mathcal{D}_1$ through $\mathbf{P}^A$. The final registration causes the two trajectories to be crossed leading to larger errors as we move into the oesophagus.

1) Hand-Eye Calibration: Mathematically the hand-eye calibration involves solving the equation $AT_{EM}^{M_1} = T_{EM}^{M_1}B$. Where $A \in SE(3)$ is the relative motion of the EMFE and $B \in SE(3)$ represents the relative motion of the marker M_1 attached to EMFE. Several approaches have been presented in literature to solve this [21]–[24].

As explained in [25], a simultaneous estimation (of rotation and translation) is usually error prone. Partially motivated by their approach, we propose an alternative by subdividing the estimation of the $T_{EM}^{M_1} \in SE(3)$ transform in two separate steps. First for estimating the $R \in SO(3)$ and the second data set for computing the translation $t \in \mathbb{R}^3$. Fig. 12 shows the set-up for calibrating OTS with EMTS. We fix an OT marker M_1 on the EMFE for which we want to estimate the hand-eye transform T_{EM}^{M1}. We use a non-ferromagnetic rod to fix an EM sensor on one end and an OT marker on the other end.

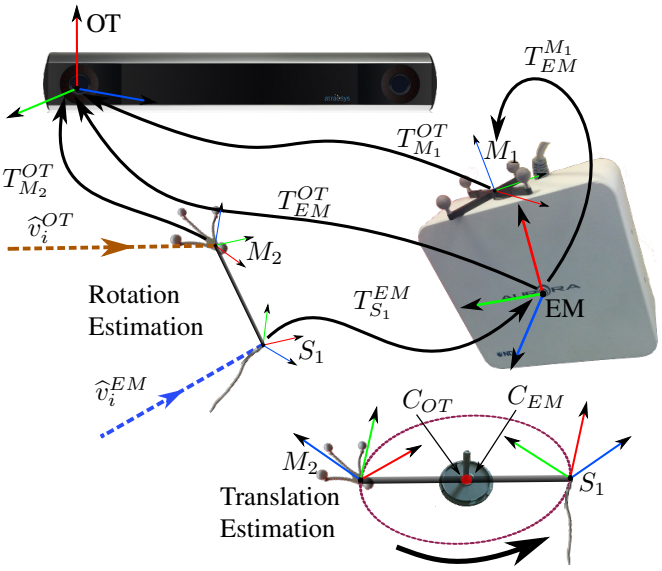

Fig. 12. Calibration set-up, M1 and M2 are optical markers. S1 is an electromagnetic sensor. The calibration is performed in two steps; line segments for estimating rotation and circle for estimation of translation.

2) Estimation of Rotation: This rod was translated along a straight line and the motion of the EM sensor and the OT marker were captured in their respective reference frames. For each of the lines we obtained the unit vectors $\hat{v}^{EM}, \hat{v}^{OT} \in \mathbb{R}^3$. Then using the approach presented in [26], where each point represents the coordinates of a unit vector in $\mathbb{R}^3$, we estimated R to align the two sets of vectors by minimizing the following criterion.

$$\min_{R} \sum_{i=1}^{N} \|\hat{v}_i^{OT} \times (R\hat{v}_i^{EM})\| \tag{6}$$

3) Estimation of translation: The planar platform with the sensor and marker, is mounted onto a turn-table as shown in Fig. 12. This allows both the EM sensor and the OT marker to rotate about a fixed axis and point in space. Using the arc formed by the EM sensor and OT marker, we estimate a circle $\in \mathbb{R}^3$ the centres (C_{EM}, C_{OT}) of which can be used to estimate the translation $t = C_{OT} - RC_{EM}$. We performed the calibration using 11 data-sets and averaged the final hand-eye calibration matrix.

4) Ground truth for Real Data: The following steps describe the approach to establish ground-truth on real data;

Step-1 Record trajectory $\mathcal{I}_1$ as described in II-B at position EM_1.

Step-2 Move the EMTS to a new position EM_2. Remove and reposition the sensors at landmarks $\mathcal{L}_A$ and $\mathcal{L}_B$.

Step-3 At EM_2 perform a new recording $\mathcal{I}_2^{tmp}$. Here only $\mathcal{D}_2$, the reference plane, was the only relevant information that was stored and the rest was discarded.

Step-4 Compute the ground-truth transform from the optical tracker as, $X = T_{EM}^{M_{11}} T_{M_{11}}^{M_{12}} T_{M_{12}}^{EM}$.

Step-5 Using the registration approach presented in II-C obtain $\hat{X}$ from registration $\mathcal{D}_1 \rightarrow \mathcal{D}_2$.

Step-6 Compute $\mathcal{I}_2 = X\mathcal{I}_1$. This represents the ground-truth.

To compute the matching error we use the trajectories $\mathcal{I}_2$ and $\hat{\mathcal{I}}_2 = \hat{X}\mathcal{I}_1$ and the matching, $f : \mathcal{T}_2 \rightarrow \hat{\mathcal{T}}_2$ is again bijective. The synchronization error along the oesophagus trajectory was obtained as in the synthetic case.

5) Results: In our experimental set-up, we used the OTS from Atracsys. The EMTS and OTS were both configured on a Linux machine (i7 3.2 GHz octa-core processor) for synchronized capture (Blackmagic Decklink Intensity pro acquisition card). 8 recordings were performed (with an average of 200 points in $\mathcal{T}$) for separate positions of the EMFE and the pig. Each recorded trajectory was registered and matched with the rest 7 recorded trajectories as described in III-B4. Fig. 13 and Fig. 14 show the final matching error along the oesophagus trajectory. Fig. 13 indicates the error results are comparable with what we observed in the synthetic case. It thus indicates that our model is a reasonable choice and the empirical values are relevant. We observe a similar error phenomenon at the beginning and end of the trajectory as was previously explained in III-A4. In 96% of the cases the error is in the range $[10.9mm, 13.47mm]$, which is within reasonable limits as was explained in III-A5.

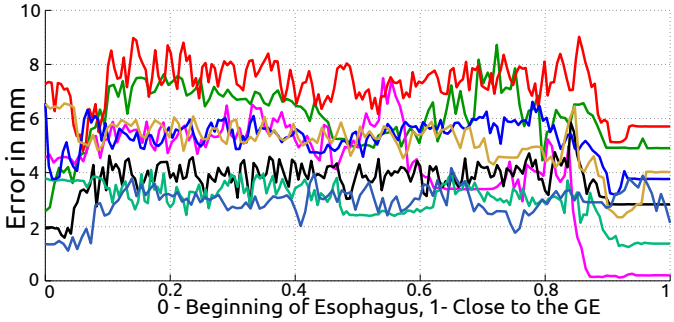

Fig. 13. Each curve represents the mean error along the oesophagus for a trajectory when synchronized with 7 other recorded trajectories. The above figure shows the result for 8 such trajectory synchronizations.

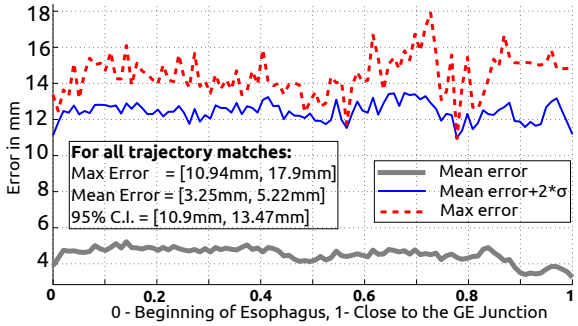

Fig. 14. Error distribution along the oesophagus trajectory for all 56 trajectory matches for interventions on pig. This figure consolidates the results of the errors in Fig. 13

IV. Qualitative Evaluation

The evaluation was conducted to compare the performance of our system against the classical approach currently being clinically followed. In the classical approach a gastroentrologist prints a picture of the locations where the biopsies were

performed and notes the corresponding approximate length of the endoscope inside the oesophagus. For a follow-up procedure, the biopsy site are approximately relocalized using the endoscope length.

We compared the performance of our system against the classical approach to guidance with 10 gastroenterologists. The evaluation was conducted a posteriori on data collected from pigs. In order to simulate real conditions, we performed two sets of recordings in the pig's oesophagus. First set of recordings ($\mathcal{S}_1$) were performed with no coagulation markings on the oesophagus tissue. Then the clinician used a coagulation device to make 12 markings (tags) along the oesophagus lining. The markings corresponded to a typical four quadrant biopsy performed in any real diagnostic intervention. A second set of recordings ($\mathcal{S}_2$) were performed on the oesophagus tagged with coagulation markings. Between each recording of $\mathcal{S}_1$ and $\mathcal{S}_2$ all the sensors were removed and replaced and the EMFE was repositioned at a new location. For the actual experimental evaluation, each of the recordings from the first set $\mathcal{S}_1$ were used as a simulated live interventions which would be matched with each of the recordings from $\mathcal{S}_2$. Using a test environment built specifically for qualitative assessment, we present the evaluation data in two phases. The test environment presented the evaluator with a the video sequence from an intervention in $\mathcal{S}_1$ which was navigated using a keyboard interface, to go back and forth on the recorded images.

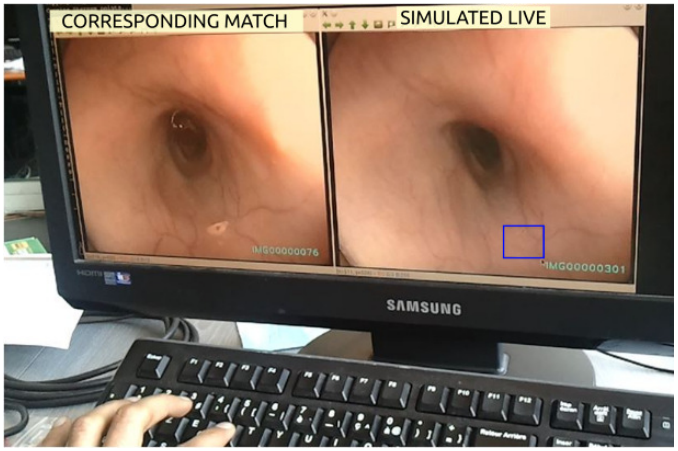

Fig. 15. Interface for qualitative evaluation. Right image corresponds to the simulated live view. Left image corresponds to the image matched using the EM position. The blue rectangle is the marking made by the evaluator on the estimated location of the biopsy point in the simulated live view. The navigation between each frame is done using a keyboard. The evaluator has to mark on only a single image where he/she expects the coagulation markings were made in the matched view.

In the first phase, the evaluator was provided with a printed picture of all the biopsy locations taken from $\mathcal{S}_2$ with their approximate length inside the oesophagus. As the sequence was being navigated the approximate distance inside the oesophagus was also presented on the screen. It corresponded to the length of the endoscope inside the oesophagus, which was shown as the approximated depth recorded using the EMTS. In the second evaluation phase, we used our system to provide a guided navigation. As the evaluator navigated the sequence of images from an intervention in $\mathcal{S}_1$, a corresponding matched

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10
	Classical Approach									
B1	≤ 1	≤ 1	≤ 1	≤ 1	NF	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
B2	≤ 1	≤ 1	≤ 1	3.5	NF	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
B3	≤ 1	≤ 1	NF	3	NF	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
B4	5	≤ 1	≤ 1	≤ 1	NF	≤ 1	≤ 1	4.2	≤ 1	4.2
B5	5.9	≤ 1	NF	≤ 1	4.5	NF	≤ 1	5.3	≤ 1	6.1
B6	NF	≤ 1	NF	≤ 1	NF	NF	≤ 1	4.1	≤ 1	NF
B7	≤ 1	≤ 1	NF	≤ 1	NF	NF	≤ 1	5.1	≤ 1	NF
B8	≤ 1	≤ 1	NF	4.3	NF	NF	≤ 1	5.5	≤ 1	NF
B9	≤ 1	≤ 1	NF	5.8	6.3	NF	NF	NF	≤ 1	≤ 1
B10	5.4	5.1	NF	≤ 1	3.2	≤ 1	3.5	NF	3.9	NF
B11	4.3	3.8	NF	NF	4.3	≤ 1	3.9	NF	4.1	≤ 1
B12	4.1	6.1	NF	NF	4.4	≤ 1	4.1	NF	≤ 1	NF
# Err ≤ 1cm	6	9	3	6	0	7	8	3	10	5
Accuracy	50%	75%	25%	50%	0%	58%	67%	25%	83%	42%
	Synchronization Approach									
B6	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	2.8	≤ 1	≤ 1	≤ 1	NF
B9	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	3	≤ 1	≤ 1	NF	≤ 1
B10	NF	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	NF
B11	3.3	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1	≤ 1
# Err ≤ 1cm	10	12	12	12	12	10	12	12	11	10
Accuracy	83%	100%	100%	100%	100%	83%	100%	100%	92%	83%

TABLE I

THE ABOVE TABLE PROVIDES THE DETAILS OF THE BIOPSY SITES AS MARKED BY 10 CLINICAL EXPERTS (COLUMNS). THE ROWS MARKED B1-B12 INDICATE THE BIOPSY SITES. EACH CELL (IN ROWS B1-B12) PROVIDES THE ERROR IN DEPTH (IN CMS) FOR IDENTIFYING A BIOPSY SITE BY THE EXPERT FOR THE CLASSICAL APPROACH. THE LABEL NF IMPLIES THE BIOPSY SITE WAS "NOT FOUND" BY THE EXPERT. FOR THE SYNCHRONIZATION APPROACH ONLY THE BIOPSY SITES INCORRECTLY IDENTIFIED OR MARKED NF HAVE BEEN PRESENTED HERE. THE % VALUES INDICATE THE ACCURACY.

image was presented from an intervention in $\mathcal{S}_2$ as shown in Fig. 15.

In each phase, the evaluator was asked to select an ROI in the images from $\mathcal{S}_1$ where he/she felt confident about the possible location of the biopsy sites by reviewing the images from $\mathcal{S}_2$. To establish the ground-truth, the expert who performed the original procedures, tagged the ROI in all the image sequences from $\mathcal{S}_1$; where the coagulation markings were made within few minutes after the completion of the original data collection, by closely reviewing the images from the sequences in $\mathcal{S}_2$. By comparing the tags made by the evaluators against ground-truth, the relocalization accuracy was computed as the number of tags successfully identified by the clinician in 3 separate trajectory matches between sets $\mathcal{S}_1$ and $\mathcal{S}_2$. To avoid any bias, we conducted both phases of the evaluation on different days.

A. Results

Table. I provides the result of qualitative evaluation of the proposed navigation system with 10 GI experts. It is evident that our approach definitely improves the localized targeting of biopsies over the classical approach by increasing the confidence in locating the target region during a surveillance endoscopy. Additionally, we noted that in presence of ambiguity our experts chose to acknowledge the uncertainty by marking the site as "Not Found" than identifying it as a wrong

location. Clinically, this would typically lead to sampling a larger portion of the oesophagus for biopsy. Using our proposed approach the number of sites "Not Found" reduced to 4 and the number of sites incorrectly identified to 3. Overall, the accuracy of re-localization improved from 47.5% to 94% and the recall from 63.3% to 97%.

V. Discussion

We have presented a complete description of our system and have rigorously evaluated the efficiency of the proposed approach using three different experiments. The first experiment, identified the sources of uncertainty in our system; and determined the extent of noise empirically from the data recorded on 18 test subjects, to generate synthetic (but realistic) data. Accuracy was measured as the discrepancy in depth (in millimetres) in the oesophagus between the matched frames. The second experiment was performed on real *in-vivo* data and ground truth established using an optical tracking system. The final experiment replicated an actual procedure offline by simultaneously presenting an image from $\mathcal{I}_1$ and the corresponding matched image from $\mathcal{I}_2$ using pig data. Here the biopsies where simulated with coagulation markings. This experiment allowed the evaluation of the true function of the system; i.e. to characterize the extent of the ability of the system to enable the GI expert to relocate biopsy sites between two procedures.

Overall, the evaluation experiments provided satisfactory results for GI experts. One could be disappointed by the results from the synthetic experiments, since errors of magnitude of up to 49 mm can be observed. However, such large errors correspond to the Gaussian tail, which would not be expected in practical scenarios (Fig. 9 and Fig. 11). Indeed, it is highly unlikely that the GI expert would position both the external EM sensors on the patient skin with an error above 10 mm. Obviously, the Gaussian should be truncated to reflect more precisely the true noise on the EM sensor positioning. We thus expect that the maximum error that will be encountered should remain less than 15 mm. This expectation is rather well confirmed by the second set of experiments involving *in-vivo* data on pigs and realistic sensor positioning error. We could indeed observe that maximum error average on 7 data-sets is less than 14 mm (Fig. 14), which corresponds to an acceptable visual discrepancy by our GI experts.

The third set of experiments most clearly indicated the benefit to the medical community in terms of an accuracy measurement that they could easily put into context. The experiments corresponded closely to their clinical work-flow and measured the extent that the system could help in the task of relocalizing biopsy sites from a first exploration in a simulated live video sequence. This experiment also showed that an approximate depth information (like the markings on the endoscope shaft) are clearly not sufficient for relocalization since the experts could not retrieve simulated biopsies with the same accuracy in comparison to our guided navigation.

We organized the synthetic experiments to be as close as possible to the realistic scenario, in terms of noise affecting the process. However, we did not model the possible deformation due to breathing or due to patient repositioning on the table. We believe the latter should have little influence on the system accuracy since an analysis of the thoracic CT of the same patient, indicated no variation in the oesophagus relative shape and position. Certainly, this should be rigorously assessed but a first comparison on several patients showed us this tendency.

In the case when the patient is taking long and deep breaths, which would result in a large shift (1 cm or more) in the relative position of the gastro-oesophageal junction [27] should not be neglected. In our context, the patient is always under general anaesthesia and the breathing is regulated. Thus, we believe that the elongation of the oesophagus at its distal end would be negligible. Further experiments on patients are nevertheless necessary to assess this phenomenon. However, since our approach uses external sensors on the sternum, we believe that by studying the relative displacement of the sensor at B ($\mathcal{L}_b$), we could propose a dynamic model of the oesophagus length (typically a deformation which quadratically decreases when the distance to the stomach increases) as a function of the relative position of the sensor P_i^B, as in [28]. The accuracy of the EM apparatus can degrade in presence of distortion due to ferromagnetic equipment. However in our application the equipment is compatible with the EMTS and we did not observe any errors during a procedure. Additionally, due to the limited working volume, we never really reach the extremities of the EMFE workspace, which means the EMFE accuracy should remain within 1 mm as specified by the manufacturer. However, in some cases when the patient size is large, it would be necessary to provide a robust approach to correcting errors in the EM tracker [15], [16].

Although our system is already very useful for the GI experts; displaying in the live view, an augmented reality based view of the previous biopsy positions can be considered as a final goal. However, a few issues still need to be addressed to reach such navigational help.

a) *Oesophagus model*: A typical surveillance procedure includes a few dense (corresponding to biopsy sites) and many sparsely populated points as recorded $\mathcal{T}$. A patient specific model that would include a smooth trajectory representing the central line in the oesophagus and corresponding frames with clear lumen visibility is necessary to eliminate redundant data and provide seamless video synchronization. This could be further used to generate patient specific virtual models for training using simulators.

b) *Improving registration*: In addition to using the three landmarks for registration as discussed in this paper; incorporating a patient specific oesophagus model could aide in reducing registration errors due to placement of sensor at $\mathcal{L}_A$ and $\mathcal{L}_B$ and in alleviating the error in depth matching.

c) *Scene matching*: The approach in this paper presents a matched image only in terms of the depth in the oesophagus. However, such an image may not contain the right view point corresponding to the live-view. Additionally the matched image could be uninformative. Choosing the image from $\mathbf{I}_2 \in \mathcal{I}_2$ in the vicinity of the match obtained using the EM tracker, such that the image best matches the live view in terms of the visual content; is an important aspect that needs to be tackled.

d) *Biopsy site mapping*: Once the optimal view-point has been obtained the next step involves mapping the regions of interest such as, biopsy sites from the matched image to the live view, as in [8]. However, to obtain such inter-operative mapping an understanding of the variation in tissue texture with time and the reliability of the key-point matching approaches is necessary. Indeed, the vasculature network along the oesophageal mucosa is visible on some patients only, and also depends on the imaging modality (NBI^{TM} by Olympus, FICE^{TM} by Fuji, SPIES^{TM} by Karl Storz and i-scan^{TM} by Pentax). The visibility of this vascular structure changes due to contractions of the oesophagus muscles. Also, it is yet unknown if the vascular network remains static or changes over a period of time. In case of a Barrett's oesophagus, it is well known that the length of the Barrett's evolves with time. Hence, patient data must be regularly collected to analyse the behaviour of the oesophagus mucosa, under different endoscopic modalities.

VI. Conclusion

In this paper, we have proposed a system to relocalize biopsy sites in a live intervention $\mathcal{I}_1$; that were performed during a previous intervention $\mathcal{I}_2$ several weeks or months prior. Our system synchronizes the video from $\mathcal{I}_2$ with the live video during $\mathcal{I}_1$, displaying in real-time an image from $\mathcal{I}_2$ which was recorded approximately at the same depth in the oesophagus. Our clinicians believe our current system can already provide a dramatic improvement for navigation in flexible endoscopy, we plan in the near future to validate our system directly on patients. The purpose would be to evaluate the extent of enhancement our system brings for better guidance, and the extent to which it modifies the clinical decision making and the eventual treatment outcome. We also plan to investigate how this work can be extended for colonoscopy. The general ideology proposed for providing a synchronized view of the video frames seems well suited for this application. However, we will need to acquire patient data to better understand and model the elastic behaviour of colon.

References

[1] B. R. Walker *et al.*, *Davidson's principles and practice of medicine*. Elsevier Health Sciences, 2013.

[2] D. S. Levine *et al.*, "An endoscopic biopsy protocol can differentiate high-grade dysplasia from early adenocarcinoma in barrett's esophagus," *Gastroenterology-Baltimore then Philadelphia*, vol. 105, pp. 40–40, 1993.

[3] K. K. Wang and R. E. Sampliner, "Diagnosis, surveillance and therapy of barrett's esophagus," *The American Journal of Gastroenterology*, vol. 103, no. 103, pp. 788–797, 2008.

[4] R. Kiesslich *et al.*, "In vivo histology of barrett's esophagus and associated neoplasia by confocal laser endomicroscopy," *Clinical Gastroenterology and Hepatology*, vol. 4, no. 8, pp. 979–987, 2006.

[5] A. L. Polglase *et al.*, "A fluorescence confocal endomicroscope for in vivo microscopy of the upper-and the lower-gi tract," *Gastrointestinal endoscopy*, vol. 62, no. 5, pp. 686–695, 2005.

[6] X. Li *et al.*, "Optical coherence tomography: advanced technology for the endoscopic imaging of barrett's esophagus." *Endoscopy*, vol. 32, no. 12, pp. 921–930, 2000.

[7] P. Mountney *et al.*, "Optical biopsy mapping for minimally invasive cancer screening," in *proceedings of MICCAI*. Springer, 2009, pp. 483–490.

[8] S. Atasoy *et al.*, "Probabilistic region matching in narrow-band endoscopy for targeted optical biopsy," in *proceedings of MICCAI*. Springer, 2009, pp. 499–506.

[9] S. Atasoy, D. Mateus *et al.*, "Endoscopic video manifolds for targeted optical biopsy," *Medical Imaging, IEEE Transactions on*, vol. 31, no. 3, pp. 637–653, 2012.

[10] B. Allain *et al.*, "Re-localisation of a biopsy site in endoscopic images and characterisation of its uncertainty," *Medical image analysis*, vol. 16, no. 2, pp. 482–496, 2012.

[11] M. Ye *et al.*, "Online scene association for endoscopic navigation," in *proceedings of MICCAI*. Springer, 2014, pp. 316–323.

[12] A. M. Franz, T. Haidegger, W. Birkfellner, K. Cleary, T. M. Peters, and L. Maier-Hein, "Electromagnetic tracking in medicine-a review of technology, validation, and applications," *Medical Imaging, IEEE Transactions on*, vol. 33, no. 8, pp. 1702–1725, 2014.

[13] K. Mori *et al.*, "Bronchoscope tracking without fiducial markers using ultra-tiny electromagnetic tracking system and its evaluation in different environments," in *proceedings of MICCAI*. Springer, 2007, pp. 644–651.

[14] Z. Yaniv *et al.*, "Electromagnetic tracking in the clinical environment," *Medical Physics*, vol. 36, no. 3, pp. 876–892, 2009.

[15] T. Reichl, J. Gardiazabal, and N. Navab, "Electromagnetic servoinga new tracking paradigm," *Medical Imaging, IEEE Transactions on*, vol. 32, no. 8, pp. 1526–1535, 2013.

[16] X. Luo and K. Mori, "Robust endoscope motion estimation via an animated particle filter for electromagnetically navigated endoscopy," *Biomedical Engineering, IEEE Transactions on*, vol. 61, no. 1, pp. 85–95, 2014.

[17] A. S. Vemuri *et al.*, "Inter-operative trajectory registration for endoluminal video synchronization: Application to biopsy site re-localization," in *proceedings of MICCAI*. Springer, 2013, pp. 372–379.

[18] D. D. Frantz *et al.*, "Accuracy assessment protocols for electromagnetic tracking systems," *Physics in medicine and biology*, vol. 48, no. 14, p. 2241, 2003.

[19] G. S. Fischer and R. H. Taylor, "Electromagnetic tracker measurement error simulation and tool design," in *proceedings of MICCAI*. Springer, 2005, pp. 73–80.

[20] Y. Wang *et al.*, "Positional accuracy and transmitter orientation of the 3d electromagnetic tracking system," *Measurement Science and Technology*, vol. 24, no. 10, p. 105105, 2013.

[21] R. Y. Tsai and R. K. Lenz, "Real time versatile robotics hand/eye calibration using 3d machine vision," in *Robotics and Automation, 1988. Proceedings., IEEE International Conference on*. IEEE, 1988, pp. 554–561.

[22] J. C. Chou and M. Kamel, "Finding the position and orientation of a sensor on a robot manipulator using quaternions," *The International Journal of Robotics Research*, vol. 10, no. 3, pp. 240–254, 1991.

[23] K. Daniilidis, "Hand-eye calibration using dual quaternions," *The International Journal of Robotics Research*, vol. 18, no. 3, pp. 286–298, 1999.

[24] J. Barreto *et al.*, "Automatic camera calibration applied to medical endoscopy," in *20th British Machine Vision Conference (BMVC)*. The British Machine Vision Association (BMVA), 2009, pp. 1–10.

[25] A. Malti and J. P. Barreto, "Robust hand-eye calibration for computer aided medical endoscopy." in *proceedings of IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2010, pp. 5543–5549.

[26] A. K. Somani, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-d point sets," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 5, pp. 698–700, 1987.

[27] B. P. Yaremko *et al.*, "Determination of respiratory motion for distal esophagus cancer using four-dimensional computed tomography," *International Journal of Radiation Oncology: Biology, Physics*, vol. 70, no. 1, pp. 145–153, 2008.

[28] A. Hostettler *et al.*, "Toward and accurate real-time simulation of internal organ motions during free breathing from skin motion tracking and an a priori knowledge of the diaphragm motion," in *Computer-Assisted Radiology and Surgery*, vol. 2, no. 1 SUPPL, 2007, pp. S100–S102.