On Simultaneous Localization and Mapping inside the Human Body (Body-SLAM)

by

Guanqun Bao

A Dissertation

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Doctor of Philosophy

in

Electrical and Computer Engineering

by

———————————————

April 2014

APPROVED:

————————————————   ————————————————————
Professor Kaveh Pahlavan, Major Thesis Advisor   Professor Yehia Massoud, Head of Department

————————————————   ————————————————————
Professor Lifeng Lai, ECE Dept., WPI   Professor Emmanuel Agu, CS Dept., WPI

————————————————   ————————————————————
Dr. Allen Levesque, ECE Dept., WPI   Dr. Kamran Sayrafian, NIST

"To my family"

Guanqun Bao

Worcester Polytechnic Institute

Abstract

Doctor of Philosophy

Body-SLAM: Simultaneous Localization and Mapping inside Human Body

by Guanqun BAO

Wireless capsule endoscopy (WCE) offers a patient-friendly, non-invasive and painless investigation of the entire small intestine, where other conventional wired endoscopic instruments can barely reach. As a critical component of the capsule endoscopic examination, physicians need to know the precise position of the endoscopic capsule in order to identify the position of intestinal disease after it is detected by the video source. To define the position of the endoscopic capsule, we need to have a map of inside the human body. However, since the shape of the small intestine is extremely complex and the RF signal propagates differently in the non-homogeneous body tissues, accurate mapping and localization inside small intestine is very challenging. In this dissertation, we present an in-body simultaneous localization and mapping technique (Body-SLAM) to enhance the positioning accuracy of the WCE inside the small intestine and reconstruct the trajectory the capsule has traveled. In this way, the positions of the intestinal diseases can be accurately located on the map of inside human body, therefore, facilitates the following up therapeutic operations. The proposed approach takes advantage of data fusion from two sources that come with the WCE: image sequences captured by the WCE's embedded camera and the RF signal emitted by the capsule. This approach estimates the speed and orientation of the endoscopic capsule by analyzing displacements of feature points between consecutive images. Then, it integrates this motion information with the RF measurements by employing a Kalman filter to smooth the localization results and generate the route that the WCE has traveled. The performance of the proposed motion tracking algorithm is validated using empirical data from the patients and this motion model is later imported into a virtual testbed to test the performance of the alternative Body-SLAM algorithms. Experimental results show that the proposed Body-SLAM technique is able to provide accurate tracking of the WCE with average error of less than 2.3cm.

Acknowledgements

Contents

List of Figures

List of Tables

Abbreviations

ASIFT   Affine Scale Invariant Feature Transform

BAN   Body Area Network

CDF   Cumulative Distribution Function

CT   Computer Tomography

CRLB   Cramer Rao Lower Bound

FDA   Food and Drug and Administration

FCC   Federal Communications Commission

FP   Feature Points

GI   Gastro Intestinal

GPS   Global Positioning System

HOG   Histogram of Oriented Gradients

IPS   Indoor Positioning System

ISM   Industrial Scientific and Medical

IMU   Inertial Measurement Unit

LBP   Local Binary Pattern

LS   Least Square

KF   Kalman Filter

MLE   Maximum Llikelihood Estimation

MSE   Mean Square Error

MRI   Magnetic Resonance and Imaging

NIST   National Institute of Science and Technology

PDF   Probability Density Function

RSS   Received Signal Strength

SNR   Signal to Noise Ratio

SLAM   Simultaneous Localization And Mapping

SRM   Statistical Region Merging

SIFT   Scale Invariant Feature Transform

ToA   Time of Arrival

WCE   Wireless Capsule Endoscopy

Symbols

$\boldsymbol{I_n}$ $n \times n$ identity matrix

$(.)^T$ matrix transpose

$(.)^H$ $n \times n$ matrix conjugate transpose

$|\boldsymbol{S}|_i$ ith column of matrix $\boldsymbol{S}$

$|\boldsymbol{S}|_{i,j}$ i, jth column of matrix $\boldsymbol{S}$

$|\boldsymbol{S}|_{i,j}$ signum function

$sign(.)$ i, jth column of matrix $\boldsymbol{S}$

$diag(\boldsymbol{A})$ vector resulting from extraction of diagonal elements of $\boldsymbol{A}$

$tr(.)$ matrix trace

$\bigtriangledown(f)$ gradient with respect to f

$E(.)$ expectation

$\delta(.)$ discrete Kronecker delta function

$Bel^-(.)$ prior belief

$Bel^+(.)$ posterior belief

Chapter 1

Introduction

Many of the profound innovations in science and engineering start with metaphors presented in the science fictions. The wireless information networking industry was motivated by the Captain Kirk's communicator in the 1960s science fiction series "Star Trek". The idea was formed in the early 1980s; the Federal Communications Commission (FCC) released the Industrial, Scientific and Medical (ISM) bands; the IEEE 802.11 standardization committee created the WLAN standard in 1997 [1]. After almost half a century, modern smart phones are what the evolution of the "Star Trek" communicator fantasy brought to us. Recently, another 1960s science fiction, the "Fantastic Voyage", in which a space craft with its crew were shrunken to become a micro-device capable of traveling inside human body to remove a brain clot, has stimulated a new wave of innovative science and engineering for the Body Area Network (BAN) [2–5]. That space craft lost its navigation capabilities and went through an unguided dramatic traveling experience within the human body before it exits through tears from the eye of the human subject. Today, wireless endoscopic capsules are traveling inside the digestive system in the same way as the space craft in the fantastic voyage traveled and one can

envision emergence of a number of other similar applications for micro-robots inside the human body.

1.1 Evolution of Wireless Capsule Endoscopy (WCE)

Endoscopy [6] is a medical procedure used to examine the interior wall of the digestive system. According to a study conducted in 2002 [7], approximately 19 million people in the United States were estimated to be affected by disorders of the small intestine. This statistic indicates that effective advancements in endoscopy technology are extremely worthy of investigation. When using the conventional endoscopic instrument, a long flexible tube with a miniature camera needs to be inserted through the mouth or the anus in order to get into the gastrointestinal (GI) tract. Owing to its rigidity and large size, it causes much discomfort to whoever undergoes this procedure. This generally limits the willingness of patients to have their GI tract examined regularly. Furthermore, the lack of capability to reach the entire small intestine is also a significant shortcoming of the current wired endoscope.

Wireless Capsule Endoscopy (WCE) [8–11], a significant step in the efforts of developing a more effective endoscopy technique, was invented to overcome the above limitations. The first WCE prototype for the small intestine was approved by the Food and Drug Administration (FDA) in 2001. Over subsequent years, this technology has been evolving into one of the most popular non-invasive imaging tools of the intestinal disease diagnosis. WCE is a pill-shaped device which consists of a short focal length CMOS camera, light source, battery and radio transmitter [12, 13]. After the endoscopic capsule is swallowed by a patient , this miniature device propelled by peristalsis of GI tract begins to work and record images at least 2 frames per second (permitting the acquisition of over 50

000 images) while moving along the GI tract. At the same time, images are sent out wirelessly in Ultra High Frequency (UHF) at 432 MHz to a small portable recorder attached to the waist [14]. The images are subsequently downloaded from the portable recorder to a workstation for analysis off line. The whole examine process takes about 8 h, during this period, the patient do not need to be confined to a hospital or clinic environment during the examination and is free to continue their daily routine. Up to now, WCE has been used to detect the following diseases [15–17] small intestinal blooding, Crohn disease, ulcer, tumors, vascular lesions and colon cancers.

1.2   Motivation

Although WCE provides a non-invasive wireless imaging technology for observing the entire GI tract, one significant drawback of this technology is that it cannot localize itself during its several hours journey. Therefore, when an abnormality is detected by the video source, the physicians have limited idea where the abnormality is located which prevents the following up therapeutic operations being executed immediately. Therefore, having a precise localization system for the endoscopic capsule would greatly enhance the benefits of WCE.

However, localization of the WCE inside the human body is not trivial. There are some fundamental technical challenges which make accurate localization inside human body a difficult task.

- First, we don't have a clear map of inside human body. A map of the digestive system plays a very important role in refining the localization results [18–20] since everything goes through the GI tract follows the same route. However, existing

computed tomography (CT) and magnetic resonance imaging (MRI) imaging tools are not able to provide enough resolution to extract the path of the small intestine.

- Second, conventional single source localization techniques, for example RF localization techniques, cannot provide satisfactory localization results due to the non-homogeneity and severe attenuation of body tissues [21]. We need to design more complicated hybrid localization algorithms that integrate all possible data sources to enhance the localization accuracy. To do this, we need researchers with multidisciplinary background including wireless localization, robotics and image processing etc.

- Third, validation of existing localization algorithms are challenging [22]. After the capsule is swallowed by the patient, we have limited control of the endoscopic capsule. Exploratory clinical procedures such as planar X-ray imaging and Ultrasound cannot be easily used for verifying the position and motion status of the capsule due to their high cost and potential risk to the patient's health.

- Last but most importantly, operating experiments inside human body is extremely difficult. As we mentioned previously, there are practical challenges to verify the performance of any localization algorithm. Moreover, human subjects are different from one and another, we need a uniform platform to do comparative performance evaluation for different algorithms.

These challenges make deign of an accurate localization system for the WCE inside human body a unsolvable engineering problem for more than 13 years. And this became the motivation of my research : To design a localization system that is able to precisely localize the endoscopic capsule as it travels along the digestive system and meanwhile reconstruct the map inside of the small intestine.

1.3   Contributions

To meet the challenges introduced above, in this dissertation, we present an in-body simultaneous localization and mapping technique (Body-SLAM) to enhance the positioning accuracy of WCE inside the small intestine and meanwhile reconstruct the trajectory the capsule has traveled. The contributions of this multi-disciplinary and inter-disciplinary dissertation are:

- Design and performance evaluation of a Body-SLAM algorithm to accurate localize the position of WCE and reconstruct the 3D map the capsule has traveled. The proposed Body-SLAM technique estimates the speed and orientation of the endoscopic capsule by analyzing displacements of feature points between consecutive images and this motion information is integrated with the RF measurements by employing a Kalman filter to smooth the localization results and the generated 3D map.

- To achieve this objective, we modeled the motion of the endoscopic capsule using empirical data obtained from a actual patients. This motion model is further imported into a emulation testbed for performance evaluation.

- We designed a tested for performance evaluation of hybrid localization algorithms that benefits from content of the endoscopic images as well as the features of the RF signal emitted from the video capsule. We used this testbed to demonstrate the effectiveness of hybrid localization algorithms for Body-SLAM inside small intestine.

The specific contributions of mine are reflected in the following publications:

1. K. Pahlavan, G. Bao, Y. Ye, S. Makarov, U. Khan ... K. Sayrafian, "Rf localization for wireless video capsule endoscopy". International Journal of Wireless Information Networks, Vol.19 (4), pp.326-340, 2012.

2. G. Bao, Y. Ye, U. Khan, X. Zheng and K. Pahlavan, "Modeling of the Movement of the Endoscopy Capsule inside GI Tract based on the Captured Endoscopic Images", The 2012 International Conference on Modeling, Simulation and Visualization Methods (MSV), Las Vegas, USA, July, 2012.

3. G. Bao and K. Phalavan, "Motion Estimation of the Endoscopy Capsule using Region-based Kernel SVM Classifier", 2013 IEEE International Conference on Electro/Information Technology (EIT), Rapid City, SD, May 9-11, 2013.

4. G. Bao, L. Mi and K. Phalavan, "Emulation on Motion Tracking of Endoscopic Capsule inside Small Intestine", 2013 World Congress in Computer Science,Computer Engineering, and Applied Computing (WORLDCOMP'13), Las Vegas, USA, 2013.

5. G. Bao, L. Mi and K. Phalavan, "A Video Aided RF Localization Technique for the Wireless Capsule Endoscope (WCE) inside Small Intestine", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

6. L. Mi, G. Bao and K. Phalavan, "Design and Validation of a Virtual Environment for Experimentation inside the Small Intestine", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

7. R. Fu, G. Bao and K. Pahlavan, "Activity Classification with Empirical RF Propagation Modeling", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

8. L. Mi, G. Bao and K. Pahlavan, "Geometric Estimation of Intestinal Contraction for Motion Tracking of Video Capsule Endoscope", SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, San Diego, California, February 15-20, 2014.

9. G. Bao, L. Mi, Y. Geng and K. Pahlavan, "A Computer Vision based Speed Estimation Technique for Localizing the Wireless Capsule Endoscope inside Small Intestine," submitted to Signal Processing Letters, IEEE, April, 2014.

10. G. Bao, L. Mi, Y. Geng, M. Zhou and K. Pahlavan, "A Video-based Speed Estimation Technique for Localizing the Wireless Capsule Endoscope inside Gastrointestinal Tract, " submitted to IEEE Engineering in Medicine and Biology Society (EMBC 14), March, 2014.

11. M. Zhou, G. Bao and K. Pahlavan, "Mutual Information based Motion Tracking Technique for the WCE inside Large Intestine", submitted to IEEE Engineering in Medicine and Biology Society (EMBC 14), March, 2014.

12. G. Bao, L. Mi and K. Pahlavan, "Hybrid Localization of Micro-robotic Endoscopic Capsule inside Small Intestine by Data Fusion of Vision and RF Sensors", submitted to Sensor Journal, IEEE, March, 2014.

A full publication list can be found in Appendix A.

1.4   Outline of the Dissertation

This dissertation focuses on the hybrid localization which we called "Body-SLAM" for the wireless capsule endoscopy and testbed design for comparative performance evaluation of various localization algorithms inside human body. The rest of the dissertation

is organized as follows: in chapter 2, we give a overview of the existing localization technologies of WCE and addressed the technical challenges in this field. In Chapter 3, we present a hybrid localization technique, which we called "Body-SLAM", that uses endoscopic images for motion tracking and combines the motion information with the RF signal radiated from the capsule to enhance the localization accuracy, and meanwhile reconstruct the trajectory the capsule has traveled. In chapter 4, performance evaluation of the proposed localization algorithm are given by using empirical data and design of emulation testbed. Finally, we conclude the dissertation in chapter 5 and give the suggested direction of future work.

Chapter 2

Challenges in WCE Localization

While physicians can receive clear images of the interior of the entire digestive system using WCE, they have little idea of the exact location of the capsule when an abnormality is found by the video source. To localize intestinal abnormalities, physicians have to administrate successive radiological, endoscopic or surgical operations, which are invasive and potentially harmful to patient's health. If we could develop a wireless localization system to localize these devices, not only can physicians diagnose the medical diseases, but they can also learn where the diseases are located. However, designing such a localization system is a very challenging task. In this chapter, we review the existing localization techniques, especially the RF localization techniques, discuss their limitations and address the challenges in designing localization system for inside human body.

2.1 Overview of Wireless Capsule Endoscopy (WCE)

Wireless Capsule Endoscopy (WCE) is a pill-shaped device which consists of a short focal length CMOS camera, light source, battery and radio transmitter [12, 13] as shown in Figure 2.1. After the endoscopic capsule is swallowed by a patient, this miniature device begins to work and record images at least 2 frames per second while moving along the GI tract. At the same time, images are sent out wirelessly to a data recorder attached to the patient's waist. The whole process takes about 8 h, then all the image data are downloaded into a work station and physicians could inspect the whole video and diagnosis diseases in the GI tract. Being such an innovative technique without cable connection, WCE offers a patient-friendly, non-invasive and painless investigation of the entire GI tract, especially the small intestine, where other conventional endoscopic instrument can barely reach. Up to now, WCE has been used to detect the following diseases: small intestinal blooding, Crohn disease, ulcer, tumors, vascular lesions and colon cancers [15–17].

A typical capsule endoscopy system consists of 3 components shown in Figure 2.2 [12, 13]:

1. A wireless capsule endoscope

   All capsule endoscopes have similar components: a disposable plastic capsule, a complementary metal oxide semiconductor or high-resolution charge-coupled device image capture system, a compact lens, white-light emitting diode illumination sources, and an internal battery source.

2. A sensing system with sensing pads or a sensing belt to attach to the patient, a data recorder, and a battery pack.

The mode of data transmission is either via ultra-high frequency band radio teleme-try (PillCam, EndoCapsule) or human body communications (MiroCam). The latter technology uses the capsule itself to generate an electrical field that uses human tissue as the conductor for data transmission. Currently PillCam SB2 and Miro-Cam are available with extended battery life, which may be beneficial in patients with delayed small-bowel transit.

3. A personal computer workstation with proprietary software for image review and interpretation.

   Major visualization systems are RAPID Reader from Given Imaging, WS-1 En-doCapsule from Olympus America and MiroView from IntroMedic.

Specifications of each individual WCE system are outlined in Table 2.1.

1	Optical dome
2	Lens holder
3	Lens
4	Illuminating LEDs
5	CMOS imager
6	Battery
7	RF transmission model
8	Antenna

FIGURE 2.1: The architecture of WCE

(a) (b) (c)
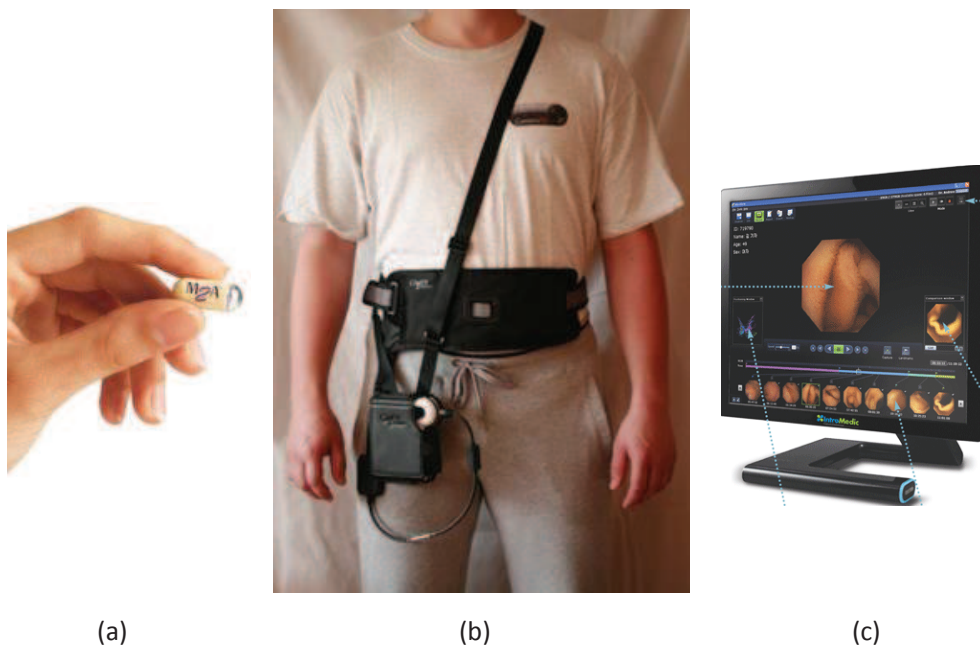
FIGURE 2.2: Wireless Capsule Endoscopy

TABLE 2.1: FDA-approved wireless capsule systems and specifications

WCE company	Size, mm	Weight	View angle	Frame rate	Battery life	Resolution
EndoCapsule						
Olympus America	11×26	3.5 g	145°	2 /sec	8 hours	512×512
PillCam SB2						
Given Imaging	11×26	2.8 g	156°	2 /sec	8 hours	256×256
PillCam SB3						
Given Imaging	11×26	2.8 g	156°	2-6/sec	12 hours	320×320
PillCam SB2EX						
Given Imaging	11×26	3.3 g	156°	2 /sec	12 hours	256×256
MiroCam						
Intromedic Co Ltd	11×26	3.3 g	170°	3 /sec	11 hours	320×320

2.2 Literature Review

WCE provides a noninvasive way to inspect the entire small intestine. As a critical component of capsule endoscopic examination, physicians need to know the precise position of the endoscopic capsule in order to identify the position of intestinal disease after it is found by the video source [23–25]. The follow up therapeutic operations and effect of drug administration are heavily dependent on the accuracy of capsule's position information [26]. Therefore, having a precise and reliable localization system plays an important role in enhancing the benefits of WCE. During the past few years, many attempts have been made to develop accurate and reliable localization systems for the WCE. A good review of existing localization techniques is given in [27]. These technologies can be divided into those using magnetic field [28–32] or inertial systems [33], using image processing techniques [34] and techniques using RF signals [4, 35–37].

In magnetic sensing based techniques, a magnet is inserted into the WCE and the WCE is located by measuring the magnetic field [38, 39]. This technique increases the weight and size of the WCE and the magnetic field of the WCE used for localization will be interfered by the external magnetic fields used for other applications such as the Magnetic Resonance Imaging (MRI) systems. One can also insert radiation opaque material into the WCE and trace the location of the WCE using X-ray or Computed Tomography (CT) scan. Continuous imaging using X-ray or CT scan is very expensive and it bears the health risks for the patient [27, 40].

In [33], Ciuti and his colleagues magnetic inertial sensing based localization system. They inserted a three-axis accelerometer LIS331DL into the capsule. the This inertial sensing not only provides the approximate location and orientation of the capsule in digestive tract, but also provides feedback to the actuation system to preserve a reliable

magnetic link between the external permanent magnet and the capsule. However, it would be difficult to make a compact capsular mechanism to be swallowable with the integration of such a inertial sensing subsystem and four cylindrical magnets. Also, this localization technique only offers rough spatial information (an average error of 3 cm) without data in a vertical direction.

Besides the magnetic field based and inertial sensing based techniques, using computer vision based technique for localization the WCE is being investigated [34, 41–44]. Because the capsule endoscope changes position and direction very slow, some identical areas exist in the successive two endoscopic images, so we can find the correspondent point pairs in these two images. Using the image correspondences, we can determine the motion (rotation and translation) parameters of capsule endoscope with an appropriate algorithm. This approach can be a complementary method for improving the magnetic localization and orientation method.

Using the RF signal used for image transmissions for the WCE to also locate the capsule offers a natural and low cost solution that does not add to the capsule extra complexity and payload [4, 45–49]. Therefore, it has been chosen for use with the smartpill capsule in USA and the M2A capsule in Israel. RF signal has been widely used for locating an object in both outdoor and indoor environments with the accuracy achieved up to hundreds of millimeters [19, 50]. Nevertheless, applying radio frequency in the task of tracking an object when it moves inside a special environment, such as the GI tract, is a challenge. This is because high-frequency signals suffer significant attenuation at different levels when they pass through different living tissues, whereas low-frequency signals due to their long wavelengths are not able to deliver the desired precision of several millimeters. The most commonly used RF techniques are Received Signal Strength (RSS) and Time of Arrival (ToA). In the following sections, we will explain the principle

of using RF signal for localization and address their limitations and challenges when applied inside human body.

2.3   RF Localization Techniques

The wireless localization industry was initiated by Global Positioning System (GPS) for outdoor navigation in early 1970s and later evolved into Indoor Positioning System (IPS) in 1990s [51–57]. Soon after, with the release of Body Area Network (BAN) IEEE 802.15.6 standard and arising of implantable micro-robots, the future trend of this localization technique is moving inside the human body [4, 58–60]. The first major application for this localization technology is the wireless capsule endoscopy [7, 29, 61–63].

A commonly used RF localization infrastructure is to attach many calibrated external RF sensors to the anterior abdominal wall of the human body to detect the RF signal emitted by the wireless capsule as shown in Figure 2.3. By interpreting the character of the received signal (RSS or ToA) into distance between the capsule and body mounted sensor array, position of the capsule can be estimated by pattern matching algorithms such as least square algorithm and maximum likelihood algorithm [28, 36]. However, RF localization of micro-robots inside humans is not trivial. Compared to outdoor and indoor environments, the inside of the human body is a complex environment making engineering design and visualization a formidable task [64]. The inside of the human body is an extremely complex medium for RF propagation because it is a non-homogeneous liquid-like environment with irregularly shaped boundaries and severe path-loss. Things become more complex when it comes inside human body since the road map for the movements of the micro-robot is blurry and the body mounted sensors

used as references for localization are also in motion. More importantly, reliable designs need testing the hardware implementation, but we cannot easily test devices inside the human bodies. Therefore, existing RF localization systems sometimes end up providing discontinuous and scattered estimations with large errors.
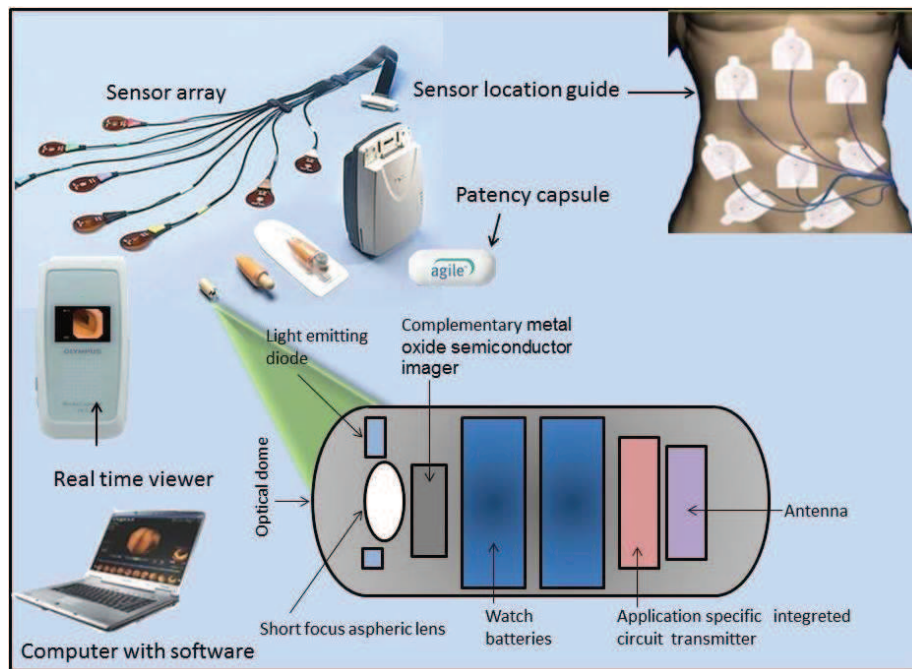
FIGURE 2.3: A typical RF localization system

2.3.1 RSS based techniques

The name of "wireless" capsule endoscope indicates its capability to transmit the images by RF signal. The transmitter embedded inside the capsule sends endoscopic images, which are captured during its travel along the GI tract, to several receivers placed uniformly on the exterior of the patient abdomen as shown in Figure 3.12. Taking advantage of this integrated function, people can measure the strength of the received RF signals at each sensor and use each sensor as a reference node to localize the capsule (mobile node). The tracking algorithm is based on the observation that the closer the receiver is to the transmitter, the stronger signal it catches. The relationship between

TABLE 2.2: Parameters for the statistical implant to body surface pathloss model

Implant to body surface	$LP(d_0)$ dB	α	σ_{dB}
Deep tissue	47.14	4.26	7.85
Near surface	49.81	4.22	6.81

the RSS reading and the distance from the transmitter to the receiver can be expressed by a pathloss model as given below [45, 46, 65, 66]:

$$RSS(d) = P_t - PL(d_0) - 10\alpha log_{10}\frac{d}{d_0} + S(d > d_0) \tag{2.1}$$

where d is the distance between transmitter and receiver, PT is the transmit power, $PL(d_0)$ is the path loss for a reference distance d_0 (i.e. 50 mm), α is the path loss gradient which is determined by the propagation environment. For example, in free space, α equals to 2. Since the human body tissue strongly absorbs RF signal, a much higher value for the path loss gradient is expected for inside human body. S is a Gaussian random variable caused by shadow fading. From Eq. 2.1, the distances between the capsule and each of the sensors can be roughly estimated by the RSS readings. Then, the capsule's location can be calculated using trilateration method.

A propagation attenuation model plays a vital role in the RSS technique. In order to reduce the positioning error, it is necessary to develop an appropriate implant to body surface path loss model. The parameters of one of the most cited signal attenuation model developed by National Institute of Standards and Technology (NIST) at MICS band are summarized in Table 2.2.

The empirical model mentioned previously is not accurate enough for the complex environment of the GI tract. The model was developed by National Institute of Standards and Technology (NIST) at MICS band.

Instead of using a signal propagation model, another RSS based localization scheme is called "finger printing" technique [47]. The way of finger printing technique works is to create a lookup table for position estimation first. Offline measurement survey needs to be done in advance, in which at each position of the capsule, both the corresponding signal strength measured by each of the sensors and its position data were recorded into the table. During the experiment, online data were compared with the data stored in the lookup table to find the closest match and thus to select the most appropriate position. However, since we don't have a map of inside the body to do the survey and people are different in term of body shape, this method doesn't have too much practical value.

2.3.2 ToA based techniques

For RF based localization, a widely known benefit of ToA based techniques is their high accuracy compared to RSS based techniques [60, 67]. The ToA based technique relies on measurements of travel time of signals between the known reference nodes and unknown mobile node. Ranging distance is calculated by multiplying the propagation velocity of RF signal and the measured ToA value [60, 68].

$$d_i = c \times \tau_i \tag{2.2}$$

However, since the human body is formed of tissues with different characteristics of conductivity and relative permittivity, the RF signal propagates with various speed through different organs [69]. These variations in the speed are the dominant source of error for the ToA-based RF localization inside the human body. Also, in near-field application, time-based methods are difficult because radio waves travel with a very high speed ($3 \times 10^8 m/s$); thus, an extremely strict time synchronization of less than 1 ns is

required in order to obtain the position resolution of 0.3 m. Another geometric location method, time difference of arrival (TDoA), does not have these disadvantages. All it needs is a transmission that has a recognizable unambiguous starting point. The data used in the location calculations is the time difference in the reception of that starting point at the several reference nodes, and not the actual time of flight of the signal from the target to the fixed sensors. But in order to have sufficient data to find the mobile's coordinates, TDoA requires one more reference node than ToA.

2.3.3 Localization Algorithms

In every ranging based localization, the position of the mobile node is determined as the intersection of the spheres [70], of which centers are the coordinates of the reference nodes and radius are the ranging distance m_i between the reference nodes $\begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^T$ and the target node $\begin{bmatrix} x & y & z \end{bmatrix}^T$, where

$$m_i^2 = (x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 \tag{2.3}$$

Since inside the human body is an non-homogeneous environment, there is difference between the true distance and the ranging distance using ToA. Therefore, the spheres do not always intersect at one single point. The goal of the localization algorithm is to find out the best estimation of the target's actual position based on the noisy measurements. Two most commonly used optimal estimation algorithms are least square algorithm and maximum likelihood algorithm. In the following subsections, we explained how these two algorithms works.

2.3.3.1   least square algorithm

In least square (LS) algorithm [71, 72], at least three reference nodes are needed to solve the least square problem. Substituting

$$x' = x - x_1 \quad y' = y - y_1 \quad z' = z - z_1 \tag{2.4}$$

and

$$x'_i = x_i - x_1 \quad (i = 2, 3) \tag{2.5}$$

into Eq. 2.3 and subtracting the first one $(i = 1)$ successively from it for $i = 2, 3$ results in an equation set in the matrix form as

$$
\begin{bmatrix}
x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\
x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\
& \cdots & \\
x_n - x_1 & y_n - y_1 & z_n - z_1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= \frac{1}{2}
\begin{bmatrix}
m_1^2 - m_2^2 + k_2 - k_1 \\
m_1^2 - m_3^2 + k_3 - k_1 \\
\cdots \\
m_1^2 - m_n^2 + k_n - k_1
\end{bmatrix}
\tag{2.6}
$$

where

$$k_i = x_i^2 + y_i^2 + z_i^2 \tag{2.7}$$

it can be denoted as

$$2\boldsymbol{A}\boldsymbol{t} = \boldsymbol{b} \tag{2.8}$$

where

$$\boldsymbol{t} = \begin{bmatrix} x & y & z \end{bmatrix}^T \tag{2.9}$$

$$\boldsymbol{A} = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \\ & ... & \\ x_n - x_1 & y_n - y_1 & z_n - z_1 \end{bmatrix} \tag{2.10}$$

$$\boldsymbol{b} = \begin{bmatrix} m_1^2 - m_2^2 + k_2 - k_1 \\ m_1^2 - m_3^2 + k_3 - k_1 \\ ... \\ m_1^2 - m_n^2 + k_n - k_1 \end{bmatrix} \tag{2.11}$$

The solution can be obtained by using the least square method [72, 73]:

$$\boldsymbol{t} = \frac{1}{2}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b} \tag{2.12}$$

2.3.3.2 maximum likelihood algorithm

This section talks about how to using maximum likelihood (ML) algorithm [74, 75] to do the localization. Assume the RSS measurements intensity for each sensor is

$$R_i = \gamma_i \sum_{k=1}^{K} \frac{C_k}{|\rho_k - r_i|^\alpha} + \omega_i \tag{2.13}$$

where R_i is the t-th sample: γ_i is gain factor, C_k is intensity of the k-th contaminant source, ρ_k is the position of the k-th source, r_i is the position of the mobile node, ω_i is the background noise.

Eq. 2.14 can be also expressed as

$$R_i = \gamma_i \frac{C}{m_i^2} + \omega_i \tag{2.14}$$

where m_i is shown in Eq. 2.3, which is the Euclidean distance between the mobile node and sensor nodes.

Setting $\xi_i = (\omega_i - \mu_i)/\sigma_i \sim N(0,1)$, $\frac{(R_i - \mu_i)}{sigma_i} \sim N(\frac{\gamma_i}{\sigma_i} \frac{C}{m_i^2}, 1)$, we can define the following matrix notation:

$$\boldsymbol{Z} = \left[\frac{(R_1 - \mu_1)}{\sigma_1}, \frac{(R_2 - \mu_2)}{\sigma_2} ... \frac{(R_N - \mu_N)}{\sigma_N} \right]^T \tag{2.15}$$

$$\boldsymbol{G} = diag \left[\frac{\gamma_1}{\sigma_1}, \frac{\gamma_2}{\sigma_2} ... \frac{\gamma_N}{\sigma_N} \right] \tag{2.16}$$

$$\boldsymbol{D} = \left[\frac{1}{m_1^2}, \frac{1}{m_2^2} ... \frac{1}{m_N^2} \right]^T \tag{2.17}$$

$$\boldsymbol{\xi} = \left[\xi_1, \xi_2 ... \xi_N \right]^T \tag{2.18}$$

We use MLE method to estimate the location. The joint probability density function can be expressed as:

$$f(Z|\theta) = (2\pi)^{N/2} exp\left\{-\tfrac{1}{2}(Z - GDC)^T(Z - GDC)\right\} \qquad (2.19)$$

its log likelihood function is:

$$L(\theta) \sim -\frac{1}{2}\sum_{i=1}^{N}\left\|Z_i - \gamma_i\frac{C}{m_i^2}\right\| = -\frac{1}{2}\sum_{i=1}^{N}\left\|\frac{R_i - \mu_i}{\sigma_i} - \gamma_i\frac{C}{m_i^2}\right\| \qquad (2.20)$$

where the θ is the estimated mobile position. Thus, we can get the maximum likely mobile position by minimizing this function [76].

2.3.4 Cramer-Rao Lower Bound (CRLB)

Cramer-Rao lower bound (CRLB), named in honor of Harald Cramer [77] and Calyam-pudi Radhakrishna Rao [78] who were among the first to derive it, expresses a lower bound on the variance of estimators of a deterministic parameter. In the localization literature [79, 80], CRLB defines the lower bound on the precision of a localization that one algorithm can reach. To calculate the CRLB for localization inside human body, we define a performance evaluation scenario and models for the behavior of the localization metrics mentioned above, the RSS and ToA, for RF signaling in between the GI tract and the body-mounted sensors used for localization. In this section, we introduce a general scenario for comparative performance evaluation of RSS and ToA based localization for capsule endoscopy application. The scenario is designed to reflect the performance in different organs, the path of movement of the WCE inside the small intestine, and the number and pattern of installation of body mounted sensors on the torso. Since the received signal on the body-mounted sensors is distorted with the multipath receptions

caused by the refraction at the boundary of organs and tissues inside the human body, models for behavior of the RSS and ToA are fairly complicated.
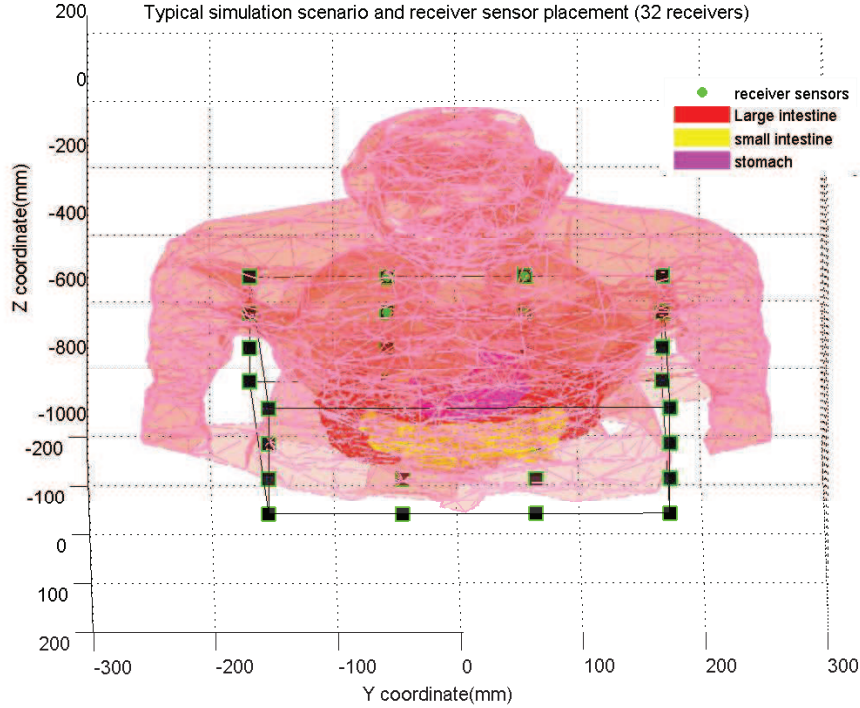
FIGURE 2.4: A typical 3D pattern of body mounted sensors used as reference points of the performance evaluation scenario for localization of the WCE

Consider the WCE whose location is being indexed as 1 and m body mounted receiver sensors denoted with indexes $2...m+1$ as shown in Figure 2.4. Each receiver sensor i is capable of measuring the ToA τ_i or RSS r_i from the WCE. The observation vector is $X = |\tau_2...\tau_{m+1}|$ for the ToA case or $X = |r_2...r_{m+1}|$ for the RSS. Assume the localization coordinate of the WCE is $\theta_1 = [x_1, y_1, z_1]$, then our objective here is to estimate the location of the WCE $\hat{\theta}_1$. The τ_i observation are modeled as normal random variables $f_{\tau_i|\theta_1,\theta_i} \ N(d_{i,1}|\bar{v}, \sigma^2_T)$, where $d_{i,1}$ is the distance between the WCE and receiver sensor i. $\bar{v}$ is the average propagation speed of the RF signal inside the human GI tract, and σ_T is the parameter describing the ToA ranging error caused by human tissue non-homogeneity. The r_i measurements are log-normally distributed $f_{r_i dB|\theta_1,\theta_i} \ N(P_r(dB), \sigma^2_{sh})$, with $P_r(dB) = P_0(dB) - 10\alpha log10(d_{1,i})$. $P_0(dB)$ is the RSS

at the reference distance from the WCE. α is the pathloss gradient and $\sigma^2{}_{sh}$ is the variance of the log normal shadowing.

The CRLB of $\hat{\theta}_1$ is $cov(\hat{\theta}_1) \geqslant I(\theta_1)^{-1}$ is the Fisher information matrix (FIM)

$$I_{\theta_1} = -E\nabla_{\theta_1}(\nabla_{\theta_1} ln\iota(X|\theta_1,\theta)) = \begin{bmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{bmatrix} \tag{2.21}$$

where $\iota(X|\theta_1,\theta)$ is the logarithm of the joint conditional probability density function:

$$\iota(X|\theta_1,\theta) = \sum_{i=2}^{m+1} log f_{\tau_1|\theta_1,\theta_i} \quad (for\ ToA) \tag{2.22}$$

$$\iota(X|\theta_1,\theta) = \sum_{i=2}^{m+1} log f_{r_1|\theta_1,\theta_i} \quad (for\ RSS) \tag{2.23}$$

and

$$I_{xx} = -\sum_{m+1}^{i=2} E[\frac{\partial^2 log f_{\tau_i|\theta_1,\theta_i}}{\partial^2 x_1^2}] \quad (for\ ToA) \tag{2.24}$$

$$I_{xx} = -\sum_{m+1}^{i=2} E[\frac{\partial^2 log f_{r_i|\theta_1,\theta_i}}{\partial^2 x_1^2}] \quad (for\ RSS) \tag{2.25}$$

Similar expressions can be extend to $I_{yy}, I_{zz}, I_{xy}, I_{xz}$ and I_{yz}. The CRLB on the variance of the ToA/RSS location estimation is

$$\sigma_1^2 = tr\left\{cov_\theta(\hat{x}_1, \hat{y}_1, \hat{z}_1)\right\}$$

$$= min \ tr(cov(\hat{\theta}_1)) = tr(I(\theta_1)^{-1})$$

$$= (-I_{xx}(I_{yy}+I_{zz})+I_{xy}I_{xy}+I_{xz}I_{xz}.../(-I_{xx}I_{yy}I_{zz}+I_{xx}I_{yz}I_{yz}...+I_{xz}I_{yy}I_{zz})) \quad (2.26)$$

2.4 Challenges of Localization inside Small Intestine

There are a number of fundamental multi-disciplinary scientific and technological challenges facing the RF localization of the WCEs inside the human body. To design an accurate localization system for inside human body we need to consider the following [4]:

- Modeling of the Movements of WCE inside the GI Tract

  The first challenge for meaningful analysis of RF localization inside the human body is to use clinical databases and clinical procedures performed by GI specialist, to model the movements of the endoscopy capsule inside the GI tract [81]. Previously acquired and stored databases of patients with approximately 55,000 images per patient could be examined for detection of landmarks or fixed points such as the pylorus and the ileocecal valve [29, 82]. Using the location of these landmarks, the number of images that observes the landmark, and the fact that the images are taken at a rate of two frames /sec (recently released WCE can take up to six frames / sec), we should design a model for the movements of the capsule in the GI tract to be mapped into the hardware and visualization platform. In the future, inertial sensing units that are small enough to be embedded in a pill size device could be used to provide real time information about pitch and roll angles of

the endoscopic capsule. This information could be used to enhance the movement model provided by examining the images reported by the capsule. The improved model for the movements of the WCE using inertial sensors would enhance the RF localization result. The feedback controlled inertial sensors have been already used to monitor the robotic end luminal system using magnetic field to efficiently perform diagnostic and surgical medical procedures [33].

- In every localization technique, map always plays a very important role in terms of refining the localization results [18, 19]. Existing literature [20] reported that a clear street map is able to reduce the GPS localization error from tens of meters to several meters in the urban area. In case of the localization inside human body, "map" is even more important since everything goes through the GI tract follows the same route. Knowing a clear pattern of the intestinal tract will greatly enhance the localization accuracy. Therefore, tracing the path of intestinal tract is essential to the accurate capsule localization.

- Modeling of the Wideband RF Propagation from Inside the Human Body

  The second challenge is to model the wideband characterization of the RF propagation channel between an endoscopy capsule and body-mounted sensors [83, 83, 84]. We could use measurements inside phantoms and on the human subject's body surface to calibrate existing software simulation tools for direct solution of Maxwell's equations inside the human body. We then could use the software to determine the waveforms observed by a body-mounted sensor used as a reference point for localization or another endoscopy capsule inside the tract that could be used for cooperative localization purposes. Finally, it should be possible to design models for the temporal and spatial features of these waveforms (that are extracted for

localization techniques) as capsules travel along the GI tract, to be used by the CPS for performance evaluation and visualization .

- Design of Complicated Algorithms for Localization inside the GI Tract

  The third challenge would be the design and comparative performance evaluation of alternative localization algorithms and discovery of methods for visualization of the results. For this part one needs to consider the use of channel models for spatial and temporal variation of the signal, the model for the track of physical movement of the capsule inside digestive system, and landmarks detected from video frames of the endoscopy capsule camera [41, 63, 85, 86]. In addition to the RF localization features, we may expect that these algorithms could exploit the knowledge of pattern of movements and the visual data observed by the camera inside the tract. The Cramer-Rao lower bound (CRLB) for the performance of basic RSS and ToA based localization algorithms for capsule endoscopy are already available in the literature [87]. We can use these bounds as a guideline for the expected performance of the designed algorithms [88].

- Security and Reliability Issues

  One last challenge in RF localization for WCEs would be to examine and where possible quantify the security, reliability, and privacy of implantable WCEs in human bodies. Here, there is an impending need to understand and analyze radio propagation of signals from WCEs outside the human body at larger distances where they may (a) cause interference (accidental or malicious) to the localization of WCEs and or devices inside a human body (b) recovered by more powerful devices towards identifying existence of such WCEs in specific patients. The former impacts the reliability of localization of the WCEs inside the human body while

the latter impacts the privacy of patients and the medical procedures that may be conducted on the patients.

In the following chapters, we are going to elaborate how we meet these challenges.

Chapter 3

Design of Algorithms for

Body-SLAM

Since the RF signal suffers from the noisy characteristics of wireless channel and multi-path distortions, it is natural to resort to other techniques to improve the overall performance of the localization system. One way to enhance the performance of RF localization is to combine the motion information of the capsule by employing a data fusion algorithm such as Kalman filter [89] or particle filter [43, 90, 91]. In our previous work [53], we have used both filters to integrate the RSS-based Wi-Fi localization and the movement models from inertial sensors including accelerometers and gyroscopes for cooperative robotic localization in indoor areas. The results were promising since this method shown the potential to enhance the localization accuracy by combining data from various sensor sources. However, as we mentioned previously, inertial sensors that meet the accuracy requirement for the WCE applications are too large to be embedded inside a video capsule and even if they can be embedded as the assembly technology

improves, the cost of the capsule will be increased dramatically. In the localization literature, there has been a trend to extract motion parameters from consecutive image sequence to improve the accuracy of RF localization, which is known as visual based Simultaneous Localization and Mapping (V-SLAM) algorithm [92–94]. In the WCE application, since the endoscopic capsule continuously takes pictures with very short time interval (up to 6 frames / sec), it is possible to extract the motion information of the capsule by processing the video stream captured by the embedded vision sensor. This motion information can be used as an alternative of inertial sensors to smooth the RF localization results and meanwhile to reconstruct the trajectory that the capsule has traveled in the same manner of V-SLAM for the indoor geo-location.

3.1   Formulation of Body-SLAM

For every location aware application, higher positioning accuracy can be achieved by employing hybrid techniques which take advantage of data fusion of different sensors [89]. Since the only two data sources come with the endoscopic capsule are video stream captured by the embedded vision sensor and wireless signal received by the body mounted RF sensors, an intuitive idea to enhance the localization accuracy of the capsule is through combination of the two [95]. As we mentioned before, the endoscopic capsule continuously takes pictures at short time interval (2 - 6 frames / sec) as it travels. Thus, it is possible to obtain information such as how quick the capsule moves and the direction of moving to track the position of the capsule. In this chapter, we present a novel motion tracking algorithm for the endoscopic capsule by analyzing the displacements of unique portion of the scene, which referred as feature points (FPs), between consecutive image frames. The proposed motion tracking algorithm consists of 3 steps: feature

points matching, image unrolling and quantitative calculation of motion parameters.

Detailed procedures of each step are explained in the upcoming subsections.
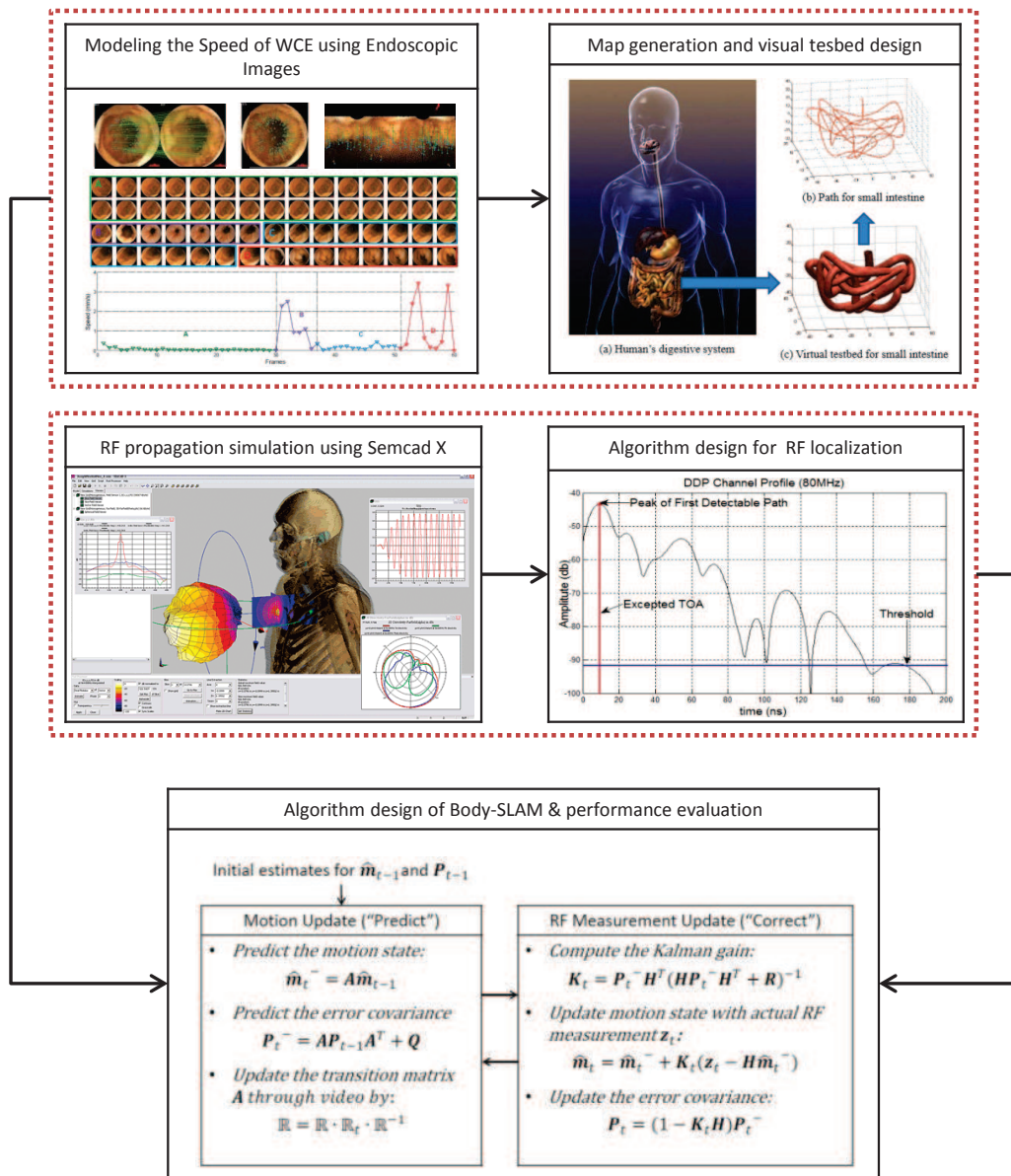
FIGURE 3.1: Overall flow chart of Body-SLAM

3.2 Motion Tracking using Endoscopic Images

The movement of the endoscopic capsule is highly unpredictable. It may move fast, slow, rotate and with any combination of the movements stated above. This complicated movement of the wireless capsule creates great errors to the localization accuracy since the Received Signal Strength (RSS) various a lots due to fast fading and sudden change of antenna gain caused by flipping and rotating. Thus, knowing how the capsule moves will help us to better understand the radio propagation channel inside human body and therefore enhance the accuracy of the existing localization methods.

3.2.1 Analyzing the Content of Endoscopic Images

3.2.1.1 Image segmentation using SRM

To model the pattern of movements of the endoscopic capsule, we need to categorize the endoscopic images first to get a conceptual idea how the capsule moves [85]. Based on our observation, the received endoscopic images can be briefly categorized into two basic categories: "facing the tunnel" (FT) and "facing the lumen" (FL). Two sets of typical FT and FL images are shown in Figure 3.2. FT is the case when the focal axis of the camera is parallel to the center of the intestinal tube. The major feature of this set of images is always there would be a black hole (we call it "tunnel" here) somewhere in the picture representing the vanishing line of the intestinal tube. Through a sequence of consecutive FT images, we can clearly see the capsule either moves propelled by the intestinal motility. On the contrast, FL is the case where the capsule tends to stop or moves not as fast as in the FT. The reason why we do such classification is we are trying to develop a geometric model for the FT images to quantitatively calculate the speed of the capsule.
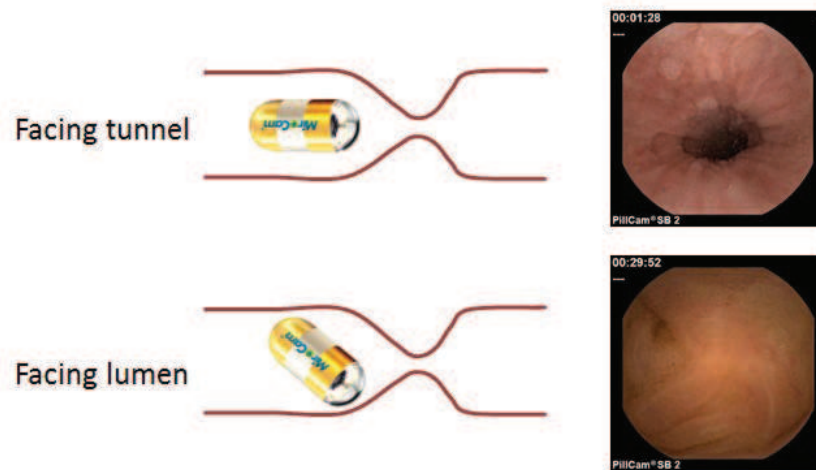
FIGURE 3.2: Two basic categories of endoscopic images

To distinguish the FT images from the FL images seems to be a fairly easy task for the human eye, however, it has been proved to be extremely difficult for the machines. Major sources of difficulties include highly complicated shape of the scene, various lightning conditions and uncontrolled noise include liquid and bubbles inside the GI tract [85]. Given a set of labeled images, finding what is in common among each set and what is difference between different set can provide inductive clues for classifier design. Some normally used feature descriptors such like Histogram of Oriented Gradients (HOG) and Local Binary Pattern (LBP) [96] doesn't work well for our application since no distinguish difference can be found between the two image sets. Thus, in terms of image representation, our approach is a region-based method. We used a Statistical Region Merging (SRM) techniques described in [97] to segment the original image into several sub-regions with each region represent an object. The basic idea of this technique is to grow the major regions iteratively by combining smaller regions or pixels with homogeneous properties. A typical example of segmented FT image is shown in Figure 3.3 with segmented regions shown in their representative colors. From Figure 3.3 we can clearly see that after the segmentation, the image preserved the tunnel shape (the darkest component in the center) while the complex textures around the tunnel are get rid of. This

will effectively reduce the variance in the feature space.
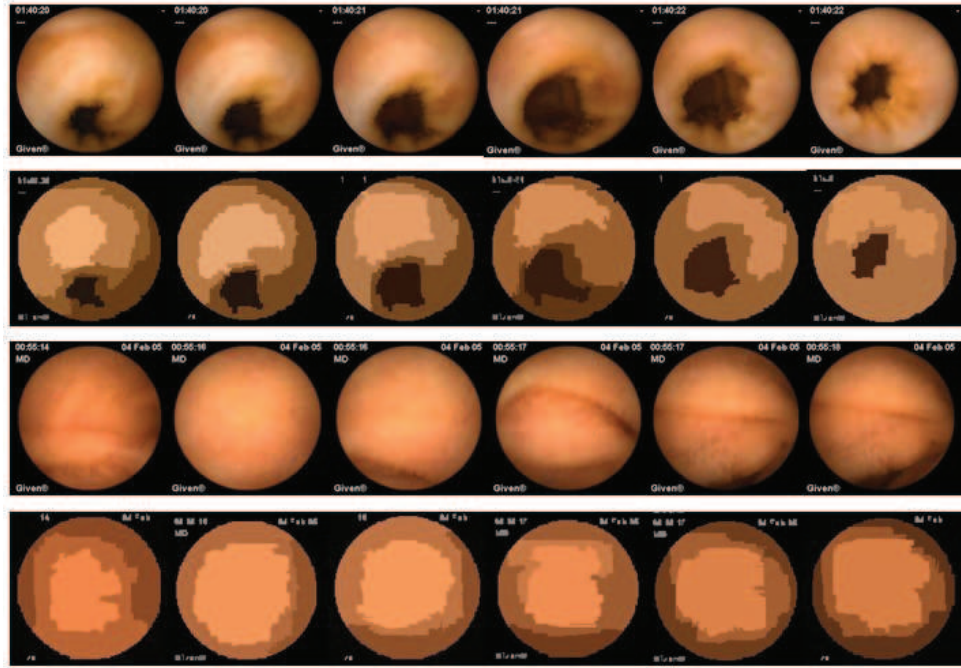
FIGURE 3.3: Two sequences of segmented image with Q = 16 Top 2 rows: FT, Bottom 2 rows: FL

The region merging rule is following:

$$P(R, R') = \begin{cases} Merge & if\, |\bar{R} - \bar{R}'| \leqslant \sqrt{s^2(R,Q) + s^2(R',Q)} \\ Not\ merge & Otherwise \end{cases} \tag{3.1}$$

where $\bar{R}$ is the average value of a certain color channel inside region R, $s(R,Q)$ is a threshold function whose value is controlled by Q. Detailed expression of $s(.)$ can be found in [97]. A good threshold is to find balance between preserving the major components of the scene and the risk of over merging. The choice Q control the coarseness of the segmentation: a large Q will keep more detailed regions while a small Q tends to merge the small regions. From the experimental point of view, we set the value of Q to be 16.

After merging, pixels inside each isolated region share a common color expectation while the expectations between adjacent regions are different for at least one color channel. Then, we can extract features out of the segmented images. To reach a good classification performance, the feature selection must obey the following rule: choose the feature that is more likely to appear in one set other than in the other set. This can be measured by calculating the co-occurrence of similar instances from different sets with the same label. Features that are more distinguishable may increase the precision of classification. Nine features are selected to classify the images. They are the size of the darkest region of the segmented image, length of the darkest region, length of the darkest region, RGB value for the darkest region and RGB value for the remaining regions.

3.2.1.2 Image classification using SVM classifier

After feature extraction, the segmented images are classified using a Kernel Support Vector Machine (K-SVM). The training data set are labeled in the following format $\left\{ \boldsymbol{x}_i, y_i \right\}$, where $\boldsymbol{x}_i$ is a $n \times 1$ feature vector, each element of the feature vector is composed by the feature we extracted from the previous section, since we extracted 9 features to represent an image, here $n = 9$, and $y_i \in \left\{ +1, -1 \right\}$ is the label of the image. If the endoscopic image is FT, $y_i = +1$, otherwise, $y_i = -1$. Suppose we have some hyperplanes which separates the positive from the negative examples, the points $\boldsymbol{x}$ that lie on the hyperplane must satisfy

$$\boldsymbol{w}^T \boldsymbol{x} + b = 0 \tag{3.2}$$

where $\boldsymbol{w}$ is a weight vector with the same dimension of $\boldsymbol{x}$ and b is a bias term, which is a real number. The distance between a training sample $\boldsymbol{x}_i$ and the boundary, usually

called "geometric margin", can be expressed as follows:

$$\frac{\boldsymbol{w}_i^T + b}{\|\boldsymbol{w}\|} \tag{3.3}$$

Since the hyperplane expressed by Eq. 3.2 are identical after $\boldsymbol{w}$ and b are scaled by a common constant, we can add a normalized restriction to this expression:

$$min|\boldsymbol{w}_i^T + b| = 1 \tag{3.4}$$

Then, the optimal solution is the boundary that maximize the minimum distance which expressed by Eq. 3.3. By restriction of Eq. 3.4, this can be reduced to maximization of $\frac{1}{\|\boldsymbol{w}\|}$.

The above equations are only applicable for the linear separable case. However, for our application, since the content of endoscopic images from different sets sometimes share similar features, the two set of images are not always linear separable. In another word, a hyperplane that can perfectly classify the two image sets does not exist. Thus, we need an approach that able to achieve nonlinear boundaries. Kernel mapping [98] is a technique which is used to solve nonlinear separation data set. The basic concept of the Kernel method is to map the vector $\boldsymbol{x}_i$ to a higher dimensional space (possibly infinite dimensional) and do the SVM in this higher dimensional space. Figure 3.4 shows a one dimensional example of kernel mapping. The transformed space should satisfy that the distance is defined in the transformed space and the distance has a relationship to the distance in the original space.
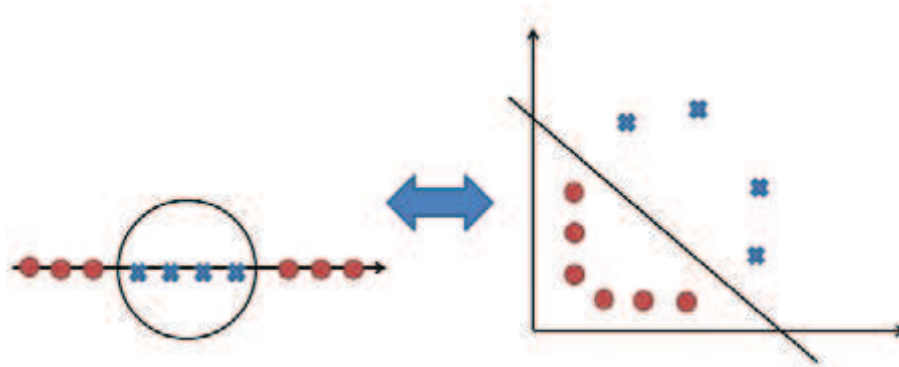
FIGURE 3.4: Illustration of feature mapping using Kernel function

$$x \in \vec{\boldsymbol{R}}^n \overset{mapping}{\to} \boldsymbol{\phi}(\boldsymbol{x}) \in \vec{\boldsymbol{R}}^m \qquad (3.5)$$

where ϕ is the mapping function. Since $\boldsymbol{\phi}(\boldsymbol{x})$ is very high dimensional, it would not be very easy to work with $\boldsymbol{\phi}(\boldsymbol{x})$ explicitly. If we measure the margin by the kernel function and perform the optimization, a nonlinear boundary can be obtained.

3.2.2 Feature Points Matching

For the FT images, the translation of the endoscopic capsule inside the small intestine can be modeled as a tiny camera passing through a elastic cylindrical tube as shown in Fig. 3.5. Since the WCE continuously takes pictures at a rate up to 6 frames/sec, common portions of the scene may present between consecutive images [34]. These portions of the images are called "feature points" (FP). The pattern and magnitudes of the displacements of these feature points can be used as a hint to reveal the speed of the endoscopic capsule.

To make an accurate estimation of the capsule's speed, it is very important that the FPs extracted from the reference (first) frame can be accurately located in the following frames.

The Affine Scale-invariant Feature Transform (ASIFT) [99–101] defined by the affine camera model in Eq. 3.7, is a perfect matching tool for the WCE images due to its immune property to viewpoint changes, blur, noise and spatial deformations.

$$A = H_\lambda R_1(\Psi) T_t R_2(\Phi)$$

$$= \lambda \begin{bmatrix} cos\Psi & -sin\Psi \\ sin\Psi & cos\Psi \end{bmatrix} \begin{bmatrix} t & 0 \\ t & 1 \end{bmatrix} \begin{bmatrix} cos\Phi & -sin\Phi \\ sin\Phi & cos\Phi \end{bmatrix} \tag{3.6}$$

where R represents rotation and T represents tilt. Ψ is rotation angle of camera around optical axis. Φ is longitude angle between optical axis and a fixed vertical plane. λ is zoom parameter. Detailed procedure of FPs matching using ASIFT can be found in [100]. An example of feature points matching is given in Fig. 3.6 (a), in which blue "o" represents the coordinates of detected FPs in the reference frame, red "o" represent the coordinates of matched FPs on the second frame. If we connect the corresponded FP pairs on the same frame (as shown in Fig. 3.6 (b)), a bunch of motion vectors will be generated representing the displacements of FPs between frames.
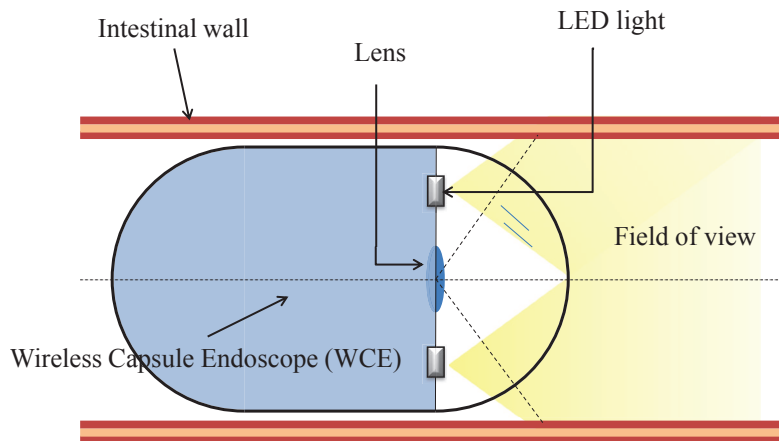
FIGURE 3.5: A WCE moving inside the small intestine

(a) Corresponding feature points between
two consecutive frames

(b) Formation of
motion vectors
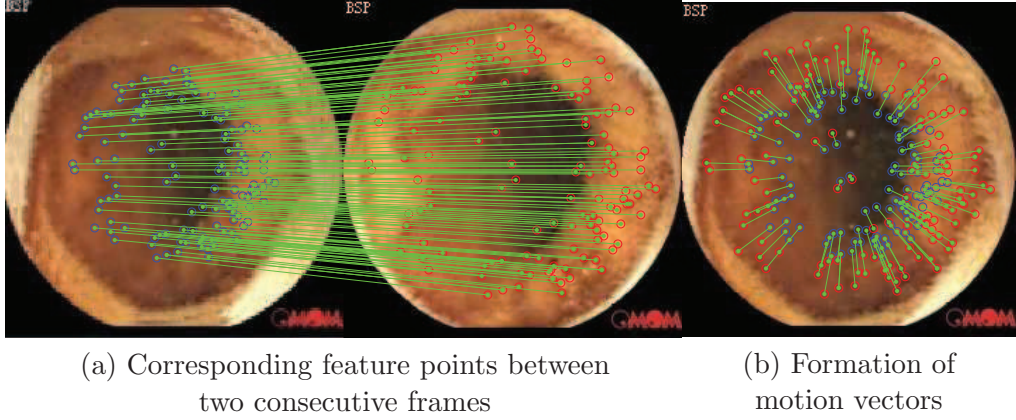
FIGURE 3.6: Feature matching between two consecutive images using A-SIFT

3.2.3 Image Unrolling

To standardize the displacement of each FP pair and facilitate the quantitive calculations of motion parameters that are useful for localization, we need to perform an inverse cylindrical projection [102] (also referred as "image unrolling" in [63, 103]) to project the original cylindrical image onto an flatten view coordinate system, which we called "unrolled" image domain. As shown in Fig. 3.7, given a point P at distance d away from the camera, the angler depth of P is defined as:
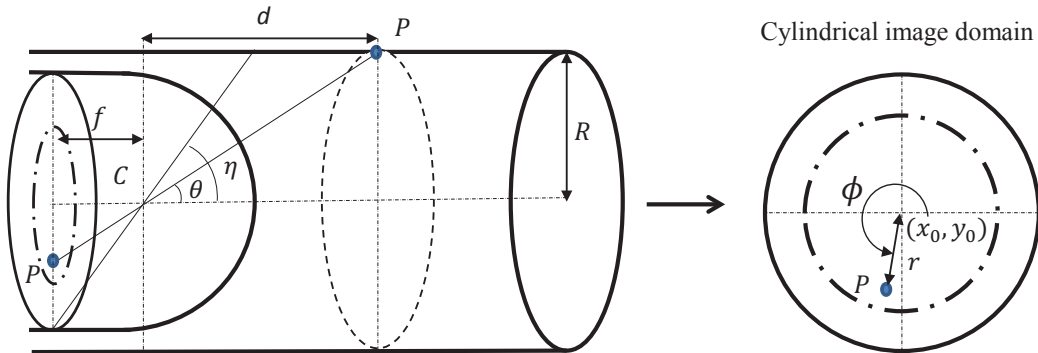
FIGURE 3.7: Image acquisition system of WCE.

$$\theta = tan^{-1}\left(\frac{R}{d}\right) \qquad (3.7)$$

where R represents the radius of the intestinal tube. It can be seen from Eq. 3.7 that a smaller angler depth indicates a larger distance away from the camera. To facilitate the derivation of angler depth, we map the coordinate (x, y) of any point on the cylindrical image plane to the unrolled image plane (x', y') by:

$$x' = \frac{L\phi}{2\pi} \qquad y' = r \qquad\qquad (3.8)$$

where ϕ is the angle between point P and the horizontal axis in the cylindrical image plane (shown in Fig. 3.8 (a)).

$$\phi = tan^{-1}\left(\frac{y - y_0}{x - x_0}\right) \qquad\qquad (3.9)$$

r is the radius of the circular ring associated with point P that can be calculated by:

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2}. \qquad\qquad (3.10)$$

L and H are length and height of the unrolled image plane respectively. Fig. 3.8 illustrates the procedure of image unrolling.

In this unrolled image plane, x' axis represents the radian angle ϕ whose value ranges from 0 (when $x' = 0$) to 2π (when $x' = L$). y' axis represents angular depth which reflect the distance away from the camera. $y' = 0$ represents a 0 angular depth and $y' = H$ gives the maximum field of view η of the camera. As can be seen in Fig. 3.8(a), after the mapping, the circular rings in the cylindrical image plane are stacked up vertically in the unrolled image plane. Under this new coordinate system, the angular depth of any point P can be calculated directly through its y' value by:

$$\theta \cong \left(\frac{y'}{H}\right)\eta \tag{3.11}$$

The angular depth obtained from Eq. 3.11 would facilitate the calculation on the speed of the capsule and values changes in x' direction would facilitate the calculation of rotation of the capsule. Detailed calculation based on this new coordinate will be presented in the upcoming subsection.
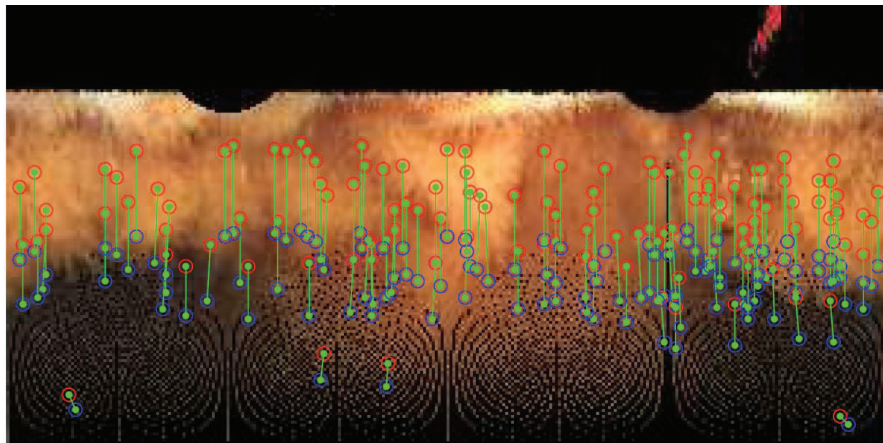
FIGURE 3.8: The process of "unrolling" the cylindrical image

3.2.4 Speed Estimation

As mentioned in previous sections, motions of a video capsule can be detected by measuring the displacements of the FPs. To explain better, we use Fig. 3.9 to illustrate the procedure of calculating the transition speed of a capsule traveling through the intestinal tube.

In Fig. 3.9 (a), point P is a FP detected at a distance D from the initial position of the camera C with its angular depth equals to θ_1. After the camera has moved forward by

a distance d to a new position C', the angular depth of P changes to θ_2. The changes in angular depth can be used to calculate the transition speed of the capsule.

$$\theta_1 = tan^{-1}\frac{R}{D} \implies D = \frac{R}{tan\theta_1} \tag{3.12}$$

$$\theta_2 = tan^{-1}\frac{R}{D-d} \tag{3.13}$$

Replacing D in Eq. 3.13 with Eq. 3.12, we get:

$$d = \frac{R}{tan\theta_2}\left(1 - \frac{tan\theta_2}{tan\theta_1}\right) \tag{3.14}$$

since the time interval for this distance d is half a second, the speed of the capsule can be calculated by:

$$v = \frac{\frac{1}{N}\sum_{i=0}^{N} d_i}{0.5} = \frac{2}{N}\sum_{i=0}^{N}\frac{R}{tan\theta_{2i}}\left(1 - \frac{tan\theta_{2i}}{tan\theta_{1i}}\right) \tag{3.15}$$

where N equals to the total number of all detected FPs.

From Eq. 3.15 it can be seen that information on depth of FP is factored into the final expression of distance moved by the capsule. In this way, the actual displacement d of the camera is independent of the location of the FP chosen in the image. To reemphasize, the unrolling process facilitates the deriving of θ_1 and θ_2 in Eq. 3.11 and therefore facilitates the deduction of v. Similarly, if the capsule moves backward, the speed can be calculated in the same manner as well.
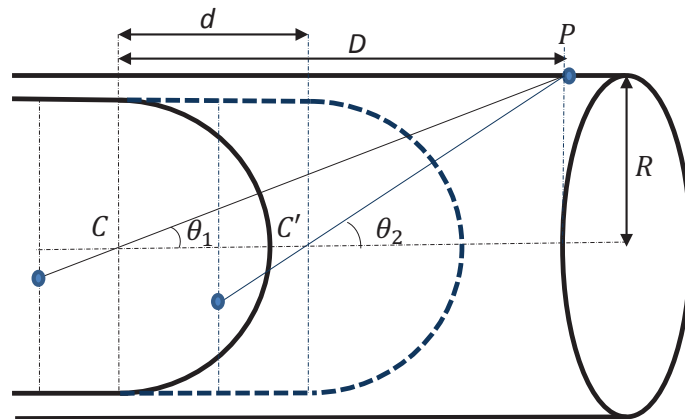
FIGURE 3.9: Speed estimation
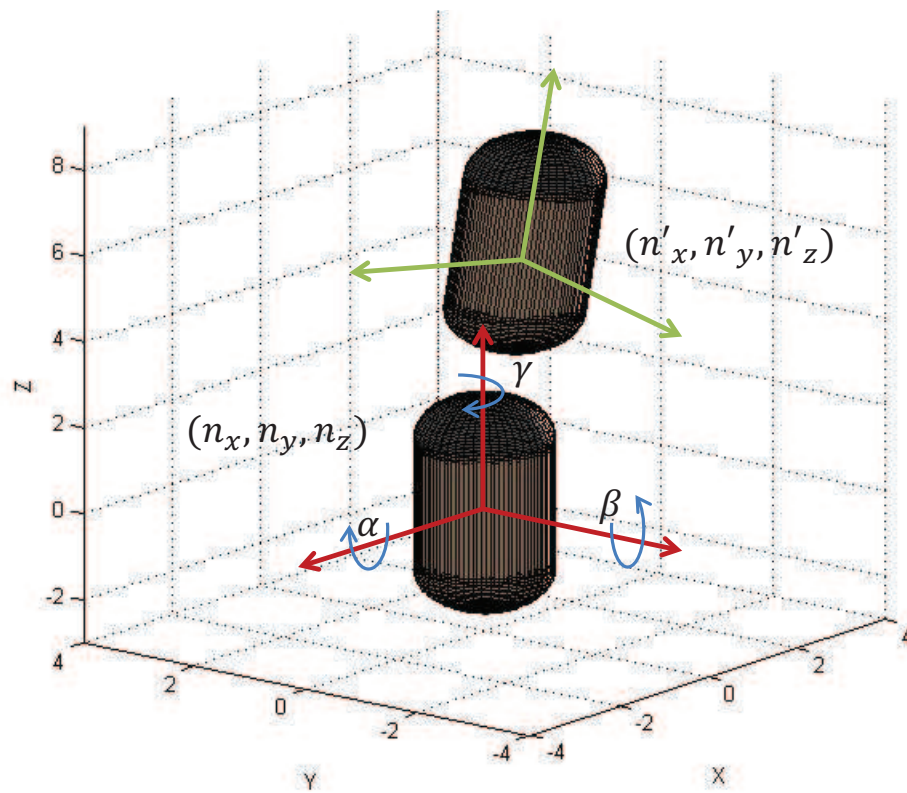
FIGURE 3.10: Direction of moving of the capsule

3.2.5 Direction of Moving Estimation

Another important aspect for motion tracking is estimating the direction of moving of the capsule. As illustrated in Fig.3.10. If we define the world coordinate as (X, Y, Z)

and capsule's coordinate as (X', Y', Z'). The moving direction of the capsule is given by a norm vector $(n_x, n_y, n_z)^T$ in the world coordinate. After the capsule rotated with angle α around its X' axis (pitch), angle β around its Y' axis (yaw) and angle γ around its Z' axis (roll), the new direction of the capsule $(n_x', n_y', n_z')^T$ can be calculated by:

$$
\begin{bmatrix} n_x' \\ n_y' \\ n_z' \end{bmatrix} = \mathbb{R} \cdot \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}
\tag{3.16}
$$

where $\mathbb{R}$ is an accumulative rotation matrix which relates the camera's coordinate system (X', Y', Z') to the world coordinate system (X, Y, Z). If we assume the camera's coordinate system was initially aligned with the world coordinate system with it is focal axis pointed to the Z axis, then, the initial value of $\mathbb{R}$ equals to a 3×3 identical matrix. As the capsule moves away from the original position, $\mathbb{R}$ is updated at each time step by:
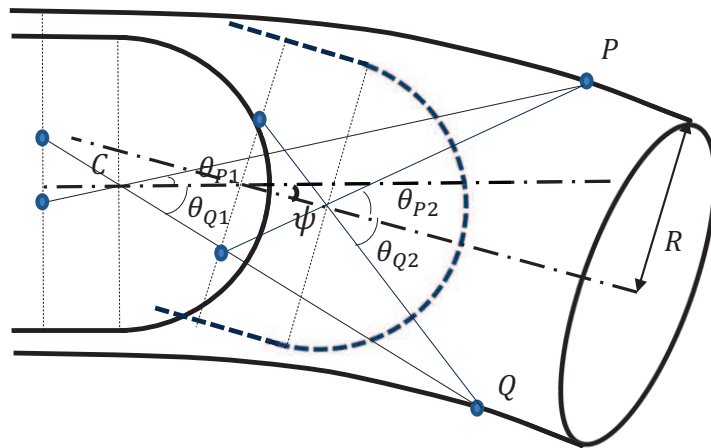
FIGURE 3.11: Direction of moving estimation

$$
\mathbb{R} = \mathbb{R} \cdot \mathbb{R}_t \cdot \mathbb{R}^{-1}
\tag{3.17}
$$

where $\mathbb{R}_t$ is an direction updating matrix that has a following expression:

$$\mathbb{R}_t = \begin{bmatrix} cos\alpha cos\gamma & cos\gamma sin\alpha sin\beta - cos\alpha sin\gamma & cos\alpha cos\gamma sin\beta - sin\alpha sin\gamma \\ cos\beta sin\gamma & cos\alpha cos\gamma + sin\alpha sin\beta sin\gamma & -cos\gamma sin\alpha + cos\alpha sin\beta sin\gamma \\ -sin\beta & cos\beta sin\alpha & cos\alpha cos\beta \end{bmatrix} \quad (3.18)$$

where α, β and γ are the pitch, yaw and roll angles about the capsule's X', Y' and Z' axises, respectively, during the elapsed time interval. Again, these angles can be obtained without complicated computation in the unrolled image domain.

- pitch (α) and yaw (β) estimation

During the transition of the capsule, the capsule will tilt toward the direction of the curly intestinal tube. As illustrated in Figure 3.10, point P and point Q are of the same distance from the initial position of the camera C. After the camera moves to C' and tilted with angle φ towards Q, the angular depths of the two FPs changes with different amount of magnitudes.

$$\triangle P = \theta_{P2} - \theta_{P1} \quad \triangle Q = \theta_{Q2} - \theta_{Q1} \quad (3.19)$$

Figure 3.10 shows that angular displacement $\triangle P$ is obviously larger than the angular displacement $\triangle Q$. The magnitude of tilting can be roughly estimated by:

$$\varphi \cong \frac{\triangle Q - \triangle P}{max(\triangle P, \triangle Q)} \quad (3.20)$$

The direction of tilting can be obtained by finding the group of with smallest displacement in y' in the unrolled image domain. Therefore, this tilting angle ϕ can be further decomposed into pitch angle α and yaw angle β by:

$$\alpha = \varphi \cdot cos\phi \quad \beta = \varphi \cdot sin\phi \tag{3.21}$$

- roll (γ) estimation

The calculation of roll angle γ is even easier in the unrolled image domain by measuring the horizontal displacements of FPs in the x' axis:

$$\gamma = \frac{1}{N} \sum_{i=0}^{N} \frac{\triangle x_i'}{L} 2\pi \tag{3.22}$$

where $\triangle x'$ denotes the horizontal displacement of a FP in the unrolled domain. L is the length of the unrolled image. It can be seen from Eq. 3.22, a greater $\triangle x'$ reflects a greater rolling angle γ and vice versa.

3.3 Data Fusion of Visual and RF Information

The two data sources that come with the endoscopic capsule provide complementary characteristics: visual motion tracking is very accurate at low-velocities but suffers from accumulative estimation errors which leads to drifting away, while RF localization provides absolute localization results that is independent from the previous measurements but with certain amount of error for estimation. In this section, we talk about how to fuse the data from both sensors to improve the reliability and accuracy of WCE localization inside small intestine. The proposed hybrid solution utilizes a Kalman filter to

predict the position of the capsule based on the motion model extracted from images and obtains feedback from the RF measurements to correct the position estimations [41].

3.3.1 Kalman Filter

The state of the mobile robot (in this case the capsule) in the 3D space can be modeled as its x, y, z coordinates and orientation. These parameters can be combined into a state variable vector. As we introduced above, the capsule continuously takes pictures as it moves along, by processing the video stream, we are able to extract the motion information about how far it has moved and its orientation. However, due to the low resolution and low frame rate, these estimations include errors and these errors accumulate frame by frame. This will cause the capsule drifting away from the correct path. The Kalman Filter (KF), which has been widely used for mobile robot navigation, is a smarter way to optimally estimate the state by integrating all available data from various sensor sources. Since the only two data sources come with the endoscopic capsule are video stream captured by the embedded vision sensor and wireless signal received by the body mounted RF sensors, an intuitive idea to enhance the localization accuracy of the capsule is through combination of the two.

The main idea of KF is to estimate the conditional probability of being in state $\boldsymbol{m}_t$ given available measurements $\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_t$. We call the probability of being in state $\boldsymbol{m}_t$ given measurements $\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_t$ the belief,

$$Bel(\boldsymbol{m}_t) = P(\boldsymbol{m}_t|\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_t) \tag{3.23}$$

We can split this belief definition into the prior belief $Bel^-(\boldsymbol{m_t})$ and the posterior belief $Bel^+(\boldsymbol{m_t})$ using Bayesian rule and Markov assumption.

$$Bel^-(\boldsymbol{m}_t) = P(\boldsymbol{m}_t|\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_{t-1}) \tag{3.24}$$

$$= \int P(\boldsymbol{m}_t|\boldsymbol{m}_{t-1})Bel^+(\boldsymbol{m}_{t-1})d\boldsymbol{m}_{t-1}$$

$$Bel^+(\boldsymbol{m}_t) = P(\boldsymbol{m}_t|\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_t) \tag{3.25}$$

$$= \frac{P(\boldsymbol{z}_t|\boldsymbol{m}_t)Bel^-(\boldsymbol{m}_t)}{P(\boldsymbol{z}_t|\boldsymbol{z}_1, \boldsymbol{z}_2...\boldsymbol{z}_{t-1})}$$

The prior belief is the conditional probability of being at state $\boldsymbol{m}_t$ given all the measurements $\boldsymbol{z}$ up to step t. The posterior belief is the conditional probability of being at state $\boldsymbol{m}_t$ given all the measurements $\boldsymbol{z}$ up to and include step t. In order to compute the beliefs, we need to find expressions for the system model $P(\boldsymbol{m}_t|\boldsymbol{m}_{t-1})$ and the measurement model $P(\boldsymbol{z}_t|\boldsymbol{m}_t)$.

In order to predict and correct the belief, the KF needs a model for the system and a model for the measurements. The KF assumes a Linear Dynamic System description of the system of which it is estimating the state. The dynamic system may be corrupted by noise sources, which the KF assumes can adequately be modeled by independent, white, zero-mean, Gaussian distributions.

- Assumptions

    The KF assumes that the system state and measurements can be described by a linear dynamic system. This is a set of linear equations that models the evolution of

the state of the system over time and that describes how measurements are related to the state. The KF assumes a linear model, since it simplifies the computations and since often a linear approach is adequate for the problem to be modeled. When the problem is not linear, then we can use linearizing techniques to transform a non-linear problem into a linear. A linear dynamic system consists of a system and a measurement model.

- System Model

  The system model describes how the true state of the system evolves over time. The KF needs this model in order to make predictions about the state. The KF assumes that the state of the system evolves according to the linear equation.

$$\boldsymbol{m}_t = \boldsymbol{A}\boldsymbol{m}_{t-1} + \boldsymbol{\omega}_{t-1} \tag{3.26}$$

  The true state $\boldsymbol{m}_t \in \boldsymbol{R}^n$ of the system at time t depends on the state of the system one step earlier $\boldsymbol{m}_{t-1}$ and some noise. Matrix $\boldsymbol{A}$ is an $n \times n$ matrix that, without taking into account possible noise in the system, relates the state of the previous time step $t-1$ to the state at the current step t. The vector $\boldsymbol{\omega} \in \boldsymbol{R}^n$ represents the noise in the system.

- Measurement Model

  The measurement model describes how measurements are related to states. The KF needs a model of the measurements in order to correct the state prediction when a measurement is available. If it has a model that given the true state of the system describes what the measurement will be, then it can compare the real measurement with the measurement that the model gives to correct the state

prediction. The KF assumes that the measurements can be modeled by an equation
that linearly relates the state of the system to a measurement,

$$z_t = Hm_t + \psi_t \tag{3.27}$$

The true measurement $z_t \in R^m$ at time t depends linearly on the state of the
system m_t. The $m \times n$ matrix H relates the current state m_t to the measurement
z_t. The vector $\psi \in R^m$ represents the noise during the measurements.

- Markov process

  Notice in Eq. 3.26 that the state m_t at time t does not depend on all other
  states and measurements given m_{t-1}. Also notice in Eq. 3.27 that, given m_t, the
  measurement z_t does not depend on all other states and measurements. These
  properties make the system a Markov process.

3.3.2 Relative Position Predictions using Images

Given the motion model derived from the previous section, the priori motion state $\widehat{m_t^-}$
at time step t (without any knowledge of RF measurement) is given by:

$$\widehat{m_t^-} = A_{t-1} \cdot \widehat{m_{t-1}} + \omega_{t-1} \tag{3.28}$$

motion state vector m_t is defined as $[x, y, z, n_x, n_y, n_z]^T$, where (x, y, z) is the 3D position
of the capsule in the world coordinate system and $[n_x, n_y, n_z]$ is a norm vector that
indicates the direction of moving of the capsule. ω_t is a noise term caused by inaccurate
motion estimation which follows a normal probability distributions with covariance equal
to Q_t. A is a 6×6 transition matrix that relates the previous motion state at time

$t-1$ to the current motion state at time t. Given that the sum of two Gaussian random variables results in another Gaussian random variable, we derive the probabilistic system model as

$$P(\boldsymbol{m}_t|\boldsymbol{m}_{t-1}) = N(\boldsymbol{A}\boldsymbol{m}_{t-1}, \boldsymbol{Q}_t) \tag{3.29}$$

If we plug in all the parameters, Eq. 3.28 can be rewritten as:

$$
\begin{bmatrix} x_t \\ y_t \\ z_t \\ n_{x|t} \\ n_{y|t} \\ n_{z|t} \end{bmatrix}
=
\begin{bmatrix}
1 & 0 & 0 & v\triangle t & 0 & 0 \\
0 & 1 & 0 & 0 & v\triangle t & 0 \\
0 & 0 & 1 & 0 & 0 & v\triangle t \\
0 & 0 & 0 & & & \\
0 & 0 & 0 & & \begin{bmatrix} \mathbb{R} \end{bmatrix} & \\
0 & 0 & 0 & & &
\end{bmatrix}
\cdot
\begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \\ n_{x|t-1} \\ n_{y|t-1} \\ n_{z|t-1} \end{bmatrix}
\tag{3.30}
$$

where v is the transition speed of the capsule derived from Eq. 3.15. $\triangle t$ is the time interval between frames (half a second). $\mathbb{R}$ is the same rotation matrix introduced in Eq. 3.16.

In the same way we want to find the probabilistic characteristics of the RF measurement model from Eq. 3.27, since this is the distribution $P(\boldsymbol{z}_t|\boldsymbol{m}_t)$ needed to compute the posterior belief from equation Eq. 3.25. The RF measurements are modeled according to the measurement model

$$\widehat{\boldsymbol{z}}_t = \boldsymbol{H}_t \cdot \widehat{\boldsymbol{m}}_t^- + \boldsymbol{\nu}_t \tag{3.31}$$

where $\boldsymbol{\nu_t}$ is a measurements noise term. Similar to $\boldsymbol{\omega_t}$, $\boldsymbol{\nu_t}$ also followed a normal distribution with covariance equal to $\boldsymbol{R_t}$. In this, we again look at the term $boldsymbol H_t \cdot \widehat{\boldsymbol{m}}_t^-$ as a Gaussian distribution with mean $N(\boldsymbol{H_t} \cdot \widehat{\boldsymbol{m}}_t^-, 0)$. Given these two Gaussian distributions, we see that the conditional probability of observing $\boldsymbol{z}_t$ given $\boldsymbol{m}_t$ is Gaussian distributed as

$$P(\boldsymbol{z}_t|\boldsymbol{m}_t) = N(\boldsymbol{H}\boldsymbol{m}_t, \boldsymbol{R}_t) \tag{3.32}$$

$\boldsymbol{H}$ is a 3×6 matrix which predicts the RF localization based on the prior motion state at time t.

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{3.33}$$

Once the actual RF localization result $\boldsymbol{z_t}$ is available, we can use it to correct the predicted position of the capsule. The covariance gives an indication of how precise the KF thinks the state estimate elements are estimated. The smaller the variances are, the more precise the state estimate is.

3.3.3 Absolute Position Measurements by RF Localization

To obtain the actual RF measurement $\boldsymbol{z_t}$, several calibrated external sensors are attached to the anterior abdominal wall of the human body as shown in Figure 3.12 to detect the wireless signal emitted by the wireless capsule [104]. As we mentioned in Chapter 2, the power of received signal (RSS) is used to identify the distance between the capsule and body mounted sensors using statistical channel models.
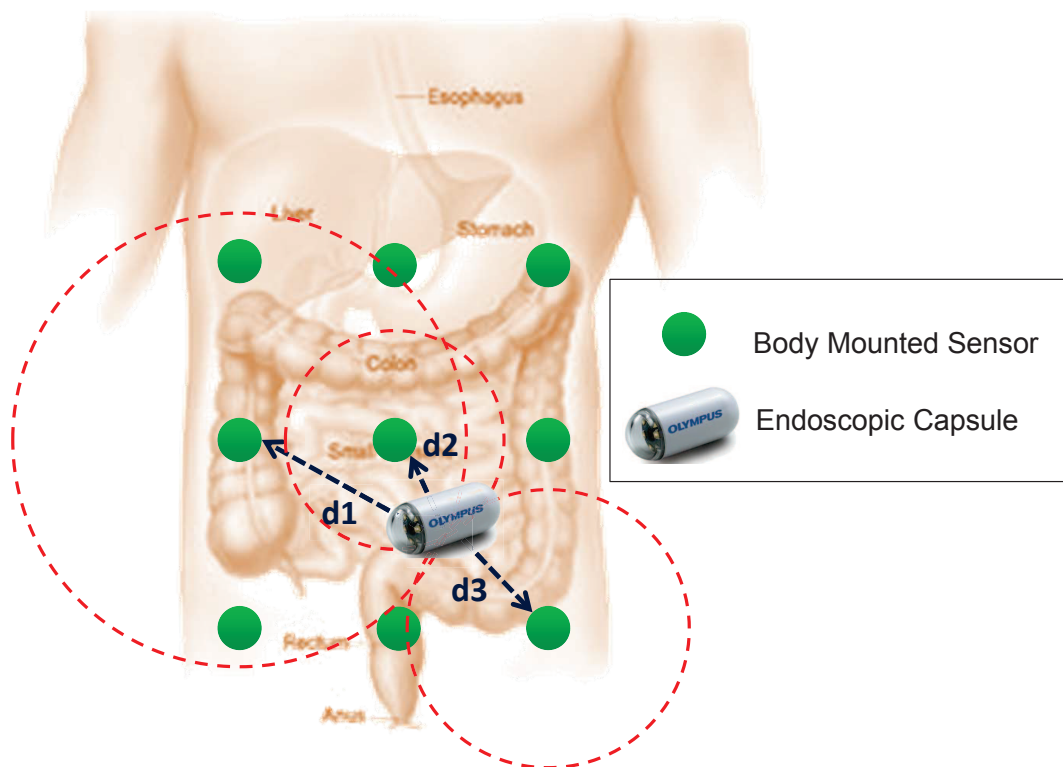
FIGURE 3.12: A typical RF localization system

$$L_p(d) = L_p(d_0) + 10\alpha log_{10}\left(\frac{d}{d_0}\right) + S(d > d_0) \tag{3.34}$$

where $L_p(d)$ represents the path loss in dB at some distance d between the transmitter and receiver, d_0 is a threshold distance and α is the path loss gradient which is determined by the propagation environment. The parameters of the path loss model developed by National Institute of Standards and Technology (NIST) at MICS band [66] was summarized in Table 2.2.

Let (x, y, z) be the potential position of the capsule and (x_i, y_i, z_i) be the position of body mounted sensor i. The ranging distance between the capsule and sensors can be expressed as:

$$d_i = 10\left(\frac{L_p(d) - L_p(d_0)}{10\alpha}\right)d_0 \tag{3.35}$$

Given 3 or more estimated distances between the capsule and body mounted sensors, the 3D position of the capsule $\boldsymbol{z_t}$ can be calculated using a least square algorithm introduced in Chapter 2 by minimizing the function below:

$$f(x, y, z) = \sum_{i=1}^{N}\left(\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} - d_i^2\right)^2 \tag{3.36}$$

3.3.4 Correction using RF Localization

After the actual RF localization $\boldsymbol{z_t}$ is obtained, we can use the priori motion estimate $\widehat{\boldsymbol{m}_t^-}$ and a weighted difference between the actual RF measurement $\boldsymbol{z_t}$ and the predicted RF measurement $\widehat{\boldsymbol{z_t}}$ to correct the localization results.

Initial estimates for $\widehat{m}_{t-1}$ and P_{t-1}

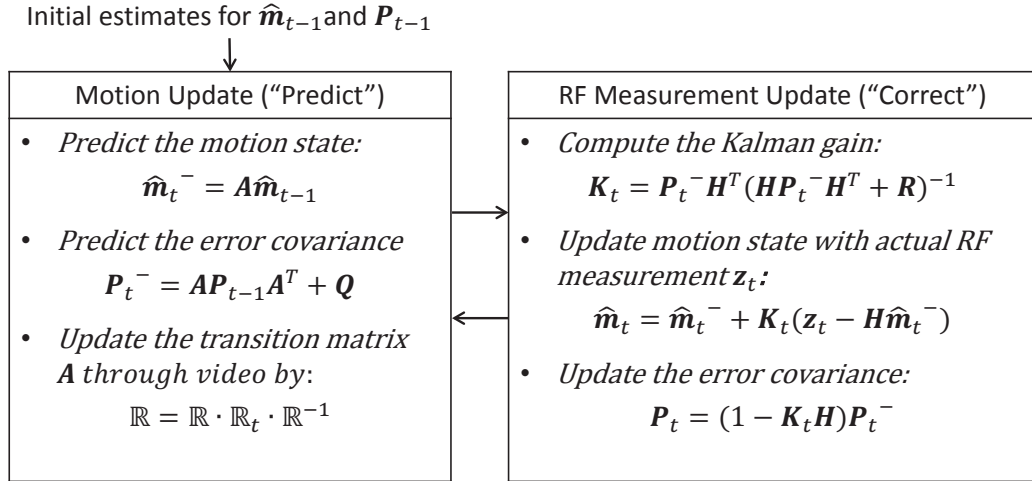| Motion Update ("Predict")	RF Measurement Update ("Correct")
• *Predict the motion state:* $$\widehat{m}_t{}^- = A\widehat{m}_{t-1}$$ • *Predict the error covariance* $$P_t{}^- = AP_{t-1}A^T + Q$$ • *Update the transition matrix* *A through video by*: $$\mathbb{R} = \mathbb{R} \cdot \mathbb{R}_t \cdot \mathbb{R}^{-1}$$	• *Compute the Kalman gain:* $$K_t = P_t{}^-H^T(HP_t{}^-H^T + R)^{-1}$$ • *Update motion state with actual RF measurement z_t:* $$\widehat{m}_t = \widehat{m}_t{}^- + K_t(z_t - H\widehat{m}_t{}^-)$$ • *Update the error covariance:* $$P_t = (1 - K_tH)P_t{}^-$$

FIGURE 3.13: A complete flowchart of data fusion of images and RF measurements using a Kalman filter

$$\widehat{m}_t = \widehat{m}_t^- + K_t(z_t - \widehat{z}_t) \tag{3.37}$$

where $\widehat{m}_t$ is defined as a posteriori motion state estimate given the RF measurement z_t. The 3×6 matrix K in Eq. 3.37 is called Kalman gain. If we define the priori estimate errors covariance as $P_t^- = E[(m_t - \widehat{m_t^-})(m_t - \widehat{m_t^-})^T]$ and a posteriori estimate errors covariance as $P_t = E[(m_t - \widehat{m_t})(m_t - \widehat{m_t})^T]$, the Kalman Gain can be expressed as:

$$K_t = P_t^- H^T \left(HP_t^- H^T + R\right)^{-1} \tag{3.38}$$

The Kalman Gain controls the weighs of both sensors on the final position estimation: if RF measurement noise is low, then the final estimation is more dependent on the RF measurement. Otherwise, the final estimation is more dependent on the motion model. The whole process of data fusion is illustrated in Fig. 3.13

Chapter 4

Performance Evaluation of

Body-SLAM

One of the major challenges in localization inside human body is performance evaluation. That's because any experiments inside human body is extremely difficult to operate. After the WCE is swallowed by the patient, we have limited control of the capsule and there is no effective way to verify their positions. Moreover, human subjects are different from one and another, we need a uniform platform for comparative performance evaluation. In this chapter, we talk about the validation of our proposed localization algorithm. Both the empirical results and emulation results are given to verify the performance of our algorithm.

4.1 Empirical Results of Motion Tracking

We need the motion model of the WCE for the simulation and the analysis of the temporal and spatial variation of the observed signals by body mounted sensors, design
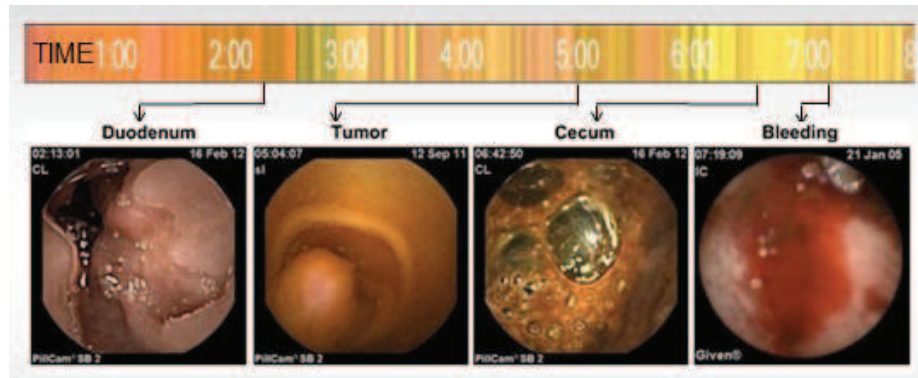
FIGURE 4.1: Some typical landmarks for the WCE

of algorithms for localization, and the emulation of the channel characteristics for performance evaluation and visualization of locations of WCEs. Unlike the movement of vehicles on roads or human beings in indoor areas, the movement of WCE inside the human body is very inconsistent and varies with the type of organs. While we cannot develop completely generalized models, we should be able to develop empirical movement models for these movements. Some pre-defined landmarks are detected by image processing techniques or identified by a GI specialist through the video source taken by WCE. These landmarks include entrance and exit of each of the four organs traversed by the endoscopy capsule: esophagus, stomach, small intestine and large intestine as well as tumors and bleeding identified in the tract. Figure 4.1 shows pictures of landmarks inside the GI tract associated with pictures of duodenum, bleeding, tumor and cecum.

As introduced in Chapter 3, if we know the motion information of the capsule by processing the endoscopic images, we can integrate this motion information with the RF measurements to enhance the localization accuracy. Conventionally, one simple approach of extracting the speed information of the capsule is to assume the capsule travels at a constant average speed and the approximate position of the capsule is calculated according to the time of travel away from some pre-defined land marks such like pylorus and ileocecal valve. Apparently, when using this approach, the further the capsule moves

away from the land marks, the greater the error is. Especially after the video capsule has entered a few centimeters of the small intestine, the localization error will increase dramatically. This is mainly due to the high complicity level of the shape of the small intestine. The distribution of small intestine is like a curled snake with its length varies from 5m to 9m (the average value for human being is 7m) and the tendency of loops is highly indistinguishable. Besides, the intestinal motility is not consistent. Peristalsis may make the wireless capsule sometimes move quickly, sometimes slow or sometimes even stop and then progress with any combination of the movements above. To enhance the localization accuracy, a precise knowledge of how fast the capsule moves is urgently needed. In this section, we apply the speed estimation algorithm introduced in chapter 3 to the clinical endoscopic image database to model the speed of the capsule traveling through the small intestine.

4.1.1 Speed Estimation using PillCam COLON 2

PillCam COLON 2 is another product of Given Imaging, which specially aims at visualization of the colon mucosa and detecting polyps [105, 106]. After FDA Rejected PillCam Colon application in USA in 2008, Given Imaging developed the second-generation Pill-Cam COLON 2 and received a CE Mark in 2009 and was commercially available in Europe in 2010. The size of PillCam COLON 2 is 11×31 mm and equipped with two identical image cameras on both ends as shown in Figure 4.2. As we mentioned in the previous chapters, performance evaluation of the speed estimation of WCE is extremely difficult. The double camera feature of the PillCam COLON 2 is a perfect tool for statistical validation. Since the images taken from both cameras are almost the same but pointing to opposite direction, after applying the speed estimation algorithm, the speed extracted from both cameras should be the same.

FIGURE 4.2: PillCam COLON 2 with double cameras

The data set we used for this validation is from UMass Memorial Medical Center. The video clip is consist of 2600 continuous image pairs. After applying the proposed motion tracking algorithms to each image pair, the result of the speed estimation (magnitude) is shown in Figure 4.3, where blue line represents the speed estimated from camera 1 and red line represents the speed estimated from camera 2. As we can see, the trend of both line match pretty well which indicate our proposed speed estimation is accurate. The plot shows that the capsule doesn't travel at a constant speed, it sometimes travels fast propelled by the intestinal motility and sometimes moves slowly even stops. This result confirms the assumption that the capsule translates at various speed. We also plotted the PDF and CDF of the speed estimation results from both cameras in Figure 4.4, they also share almost identical distributions.

4.1.2 Statistical Speed Modeling

We also tested our algorithm with the clinical data from different individuals using OMOM [107] capsule. Figure 4.7 shows a sequence of 60 endoscopic frames. After applying our speed estimation algorithm, a plot of the corresponding speed is given underneath. Since we do not have precise control of the capsule, we are not able to

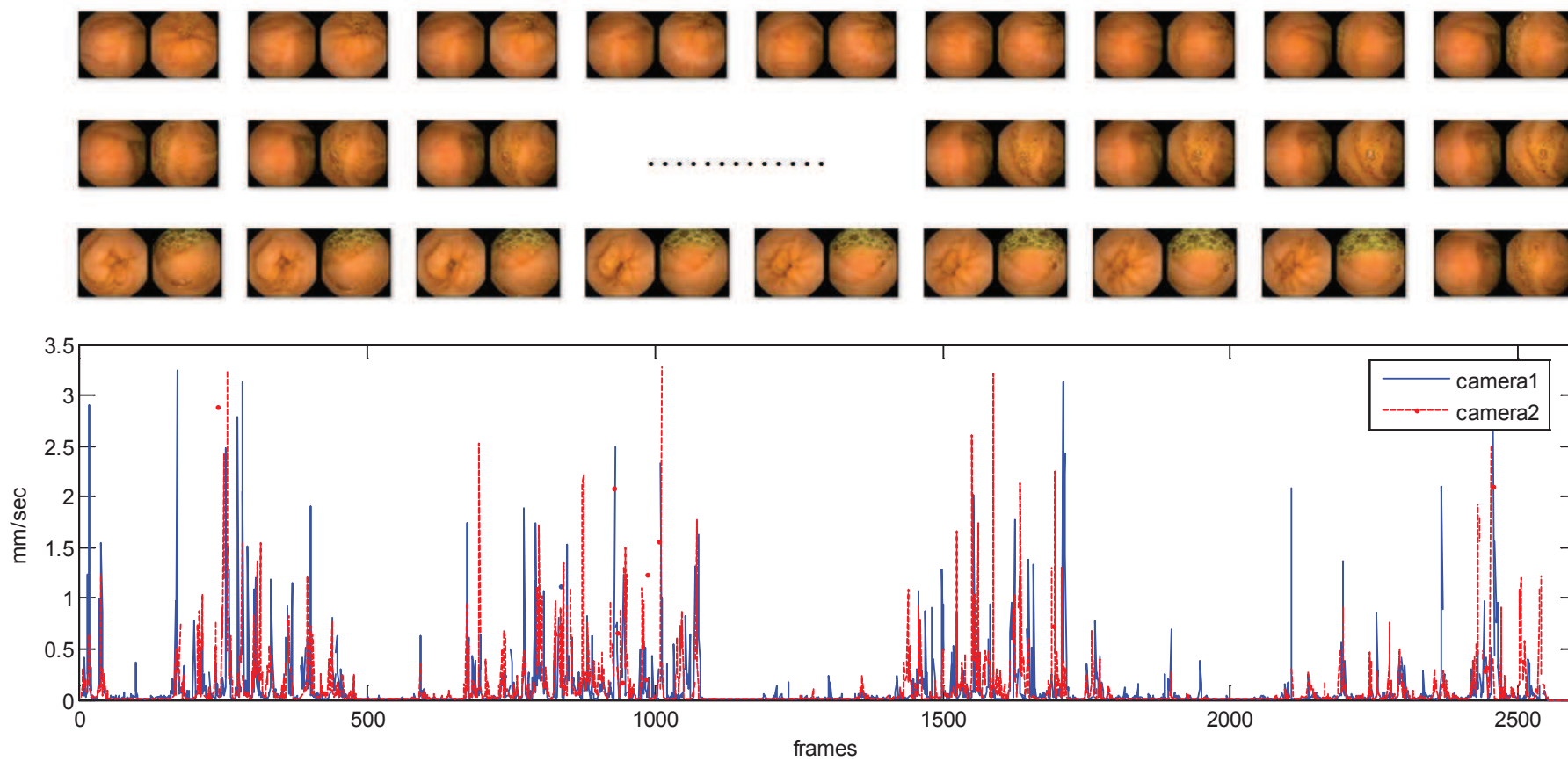FIGURE 4.3: Speed estimation results from PillCam COLON 2 double cameras
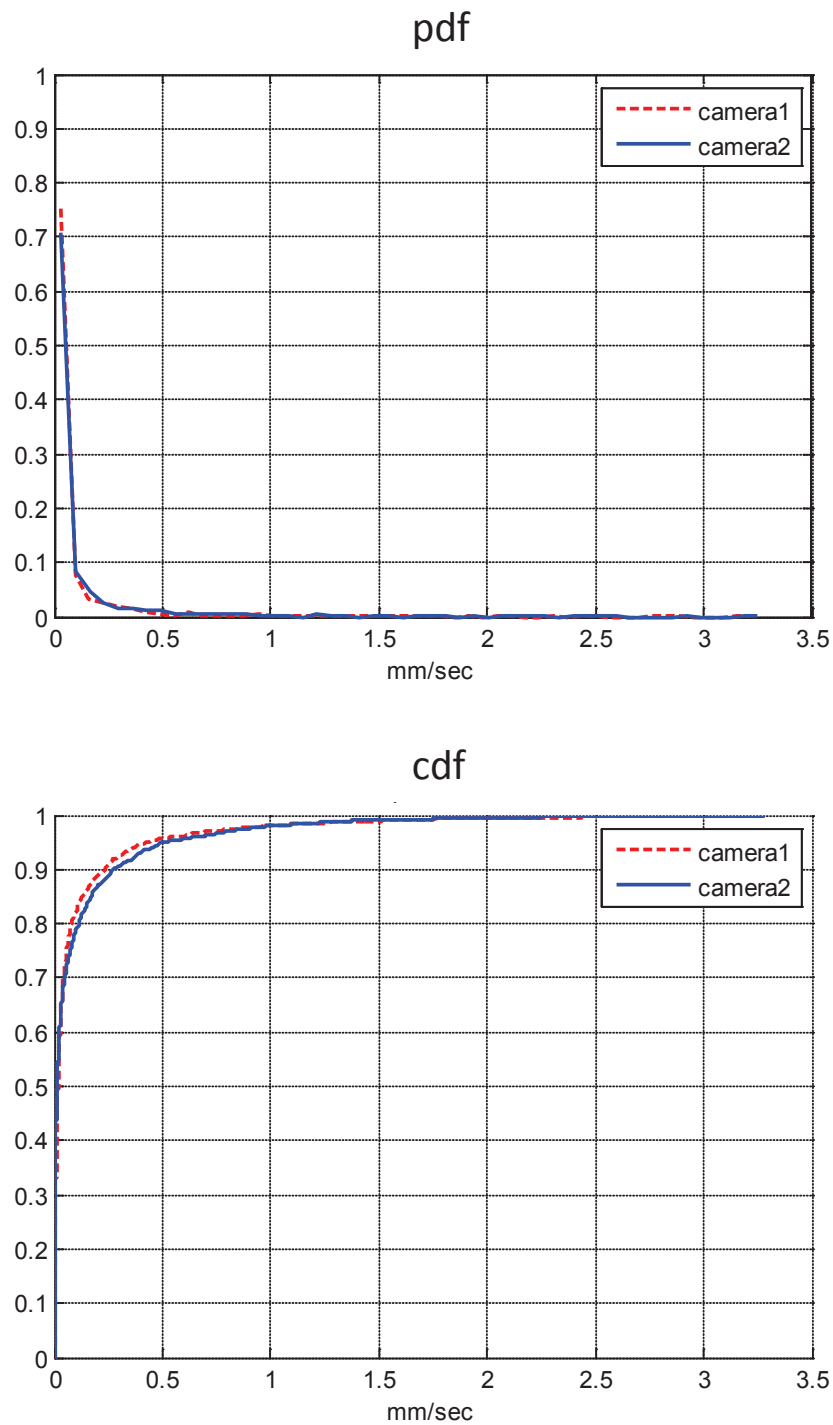
FIGURE 4.4: Statistics of speed estimation using PillCam COLON 2 double cameras

perform quantitative evaluation, but we can examine the images manually by naked eye. As shown in Figure 4.7, the whole image sequence can be divided into 4 sections marked by A, B, C, and D. It can be seen that during section A and C, the scene almost

stays still which indicates a slow motion of the capsule, while in section B and D, the capsule moves faster. The trend of the corresponding speed estimation plot matches this observation. To further validate our algorithm, we compared the statistical results
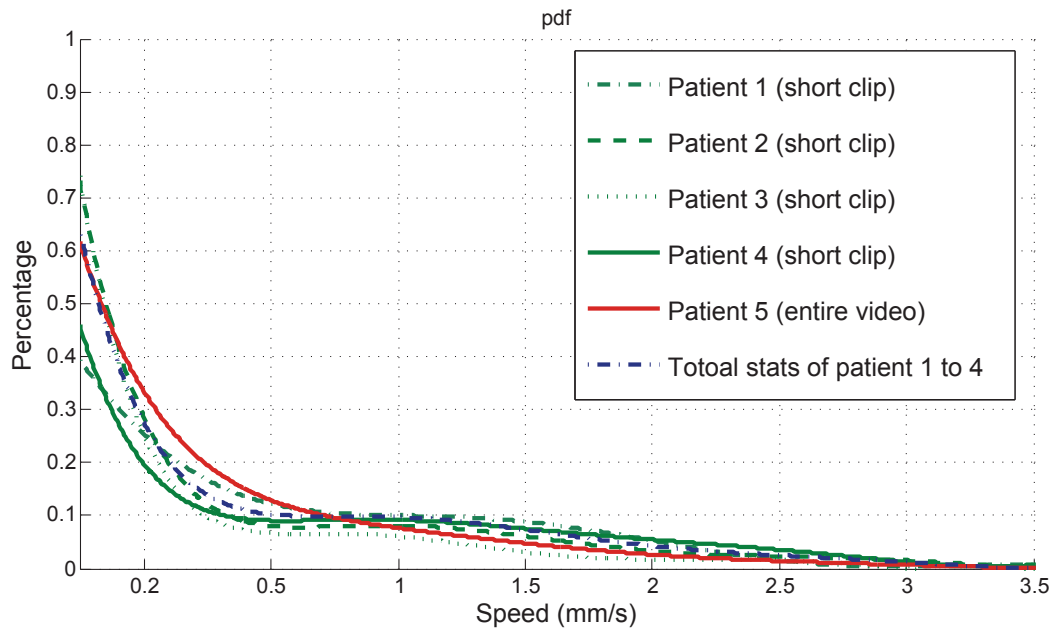
FIGURE 4.5: PDF of the speed estimation from different individuals
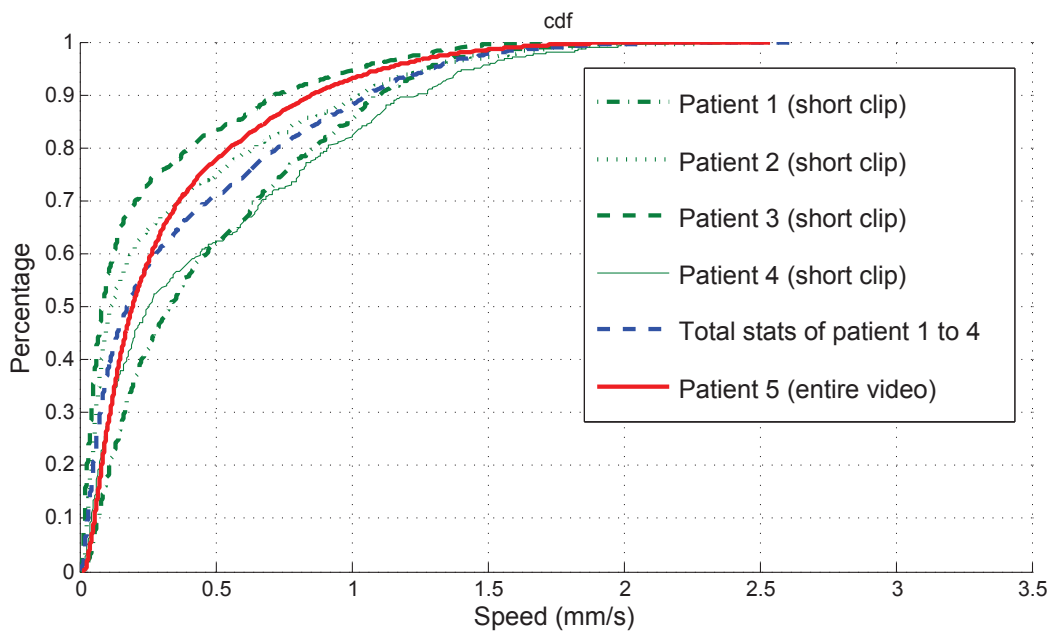
FIGURE 4.6: CDF of the speed estimation from different individuals

of clinical data (4 short video clips and 1 long video clip) from 5 different patients. The results are shown in Figure 4.5 and Figure 4.6. It can be seen although the video clips are from different individuals, after applying our speed estimation algorithm, the estimated speed shares very similar distributions in term of probability density function (PDF) and cumulative distribution function (CDF). Some typical speed estimation pattern are shown in Figure 4.8 and Figure 4.9.
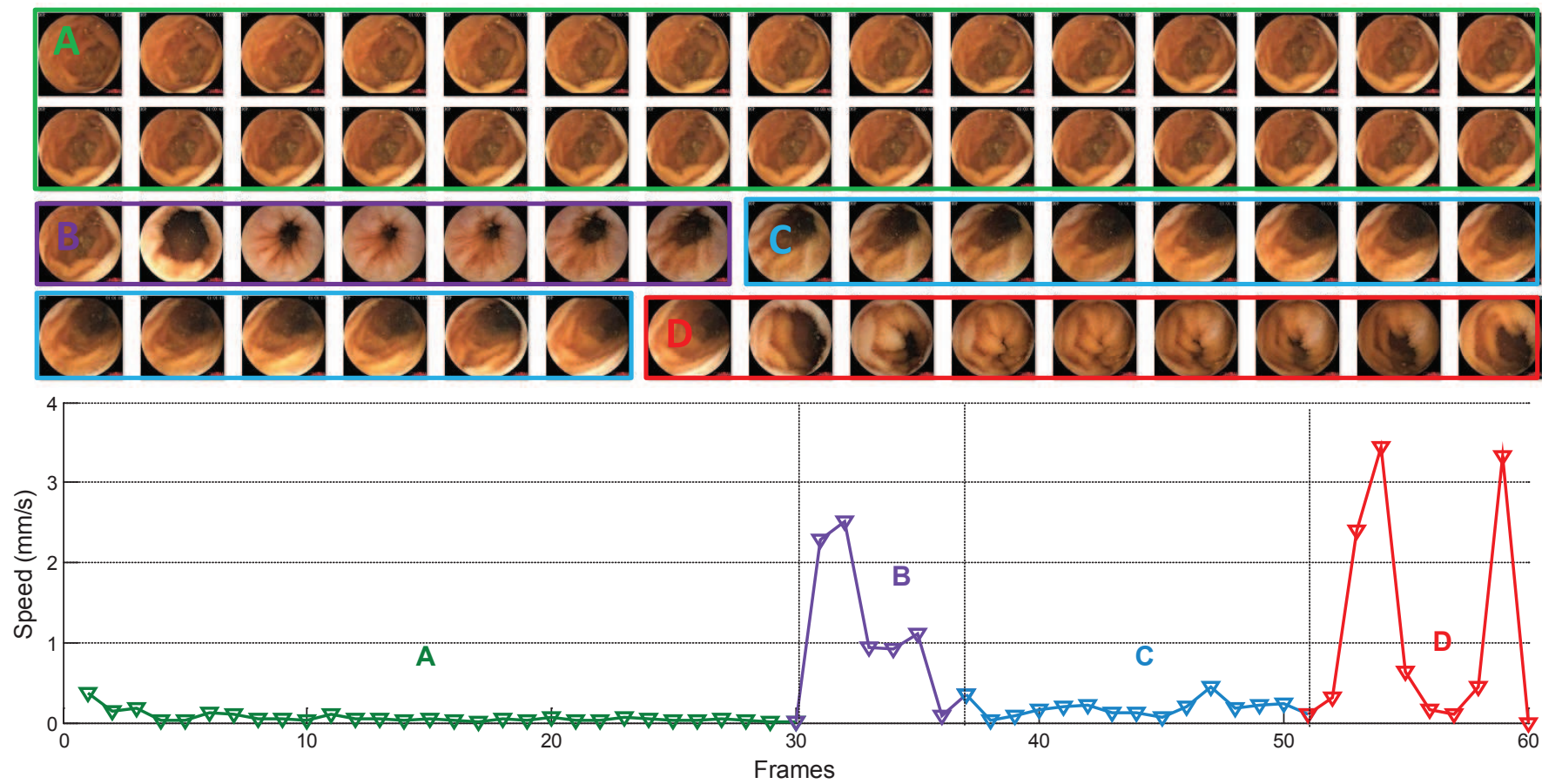
FIGURE 4.7: Speed estimation results of a sequence of real endoscopic images
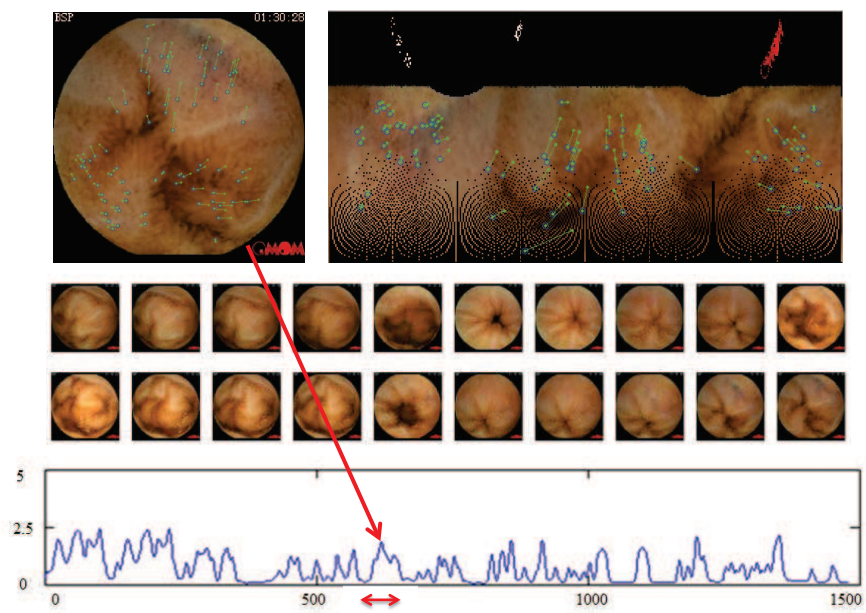
FIGURE 4.8: A typical speed pattern of moving fast
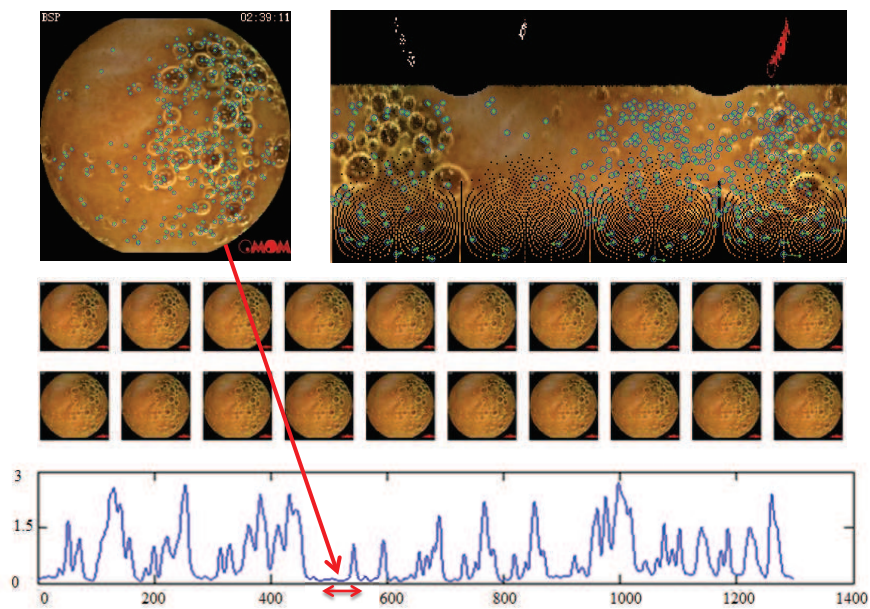
FIGURE 4.9: A typical speed pattern of moving slow

4.2 Design of Testbed for Performance Evaluation

As we mentioned before, one of the major challenges of implementing any capsule localization algorithms when it comes inside the human body is validation. That is because we have limited control of the capsule after it is swallowed by the patient so we could not verify the performance of the algorithms [4, 108]. Besides, carrying out experiments on the real human beings is extremely costly and restricted by law. Thus, the only way to verify the performance of our localization algorithm is to build up an emulation test bed. Similar idea was applied to wireless wide area network (WWAN) and wireless local area network (WLAN) deployments. In WWAN and WLAN, instead of doing real measurements, researchers model the characteristic of the RF signal under a certain scenario by creating emulation testbed. Algorithms and protocols can be imported into the testbed to test and compare their performances. In case of WCE localization, since experiments inside human body is impossible, we need to design a testbed for performance evaluation of the proposed Body-SLAM algorithm. In this section, we describe the details how to establish a testbed for localization inside human body.

Since the only two data sources we get from the WCE are images and RF signals used for transmitting these images, the testbed should include two major components: visual component and RF component. The overall flow chart of designing this testbed is illustrated in Figure 4.10.

Visual component   RF component

FIGURE 4.10: Design of emulation testbed for quantitative performance evaluation
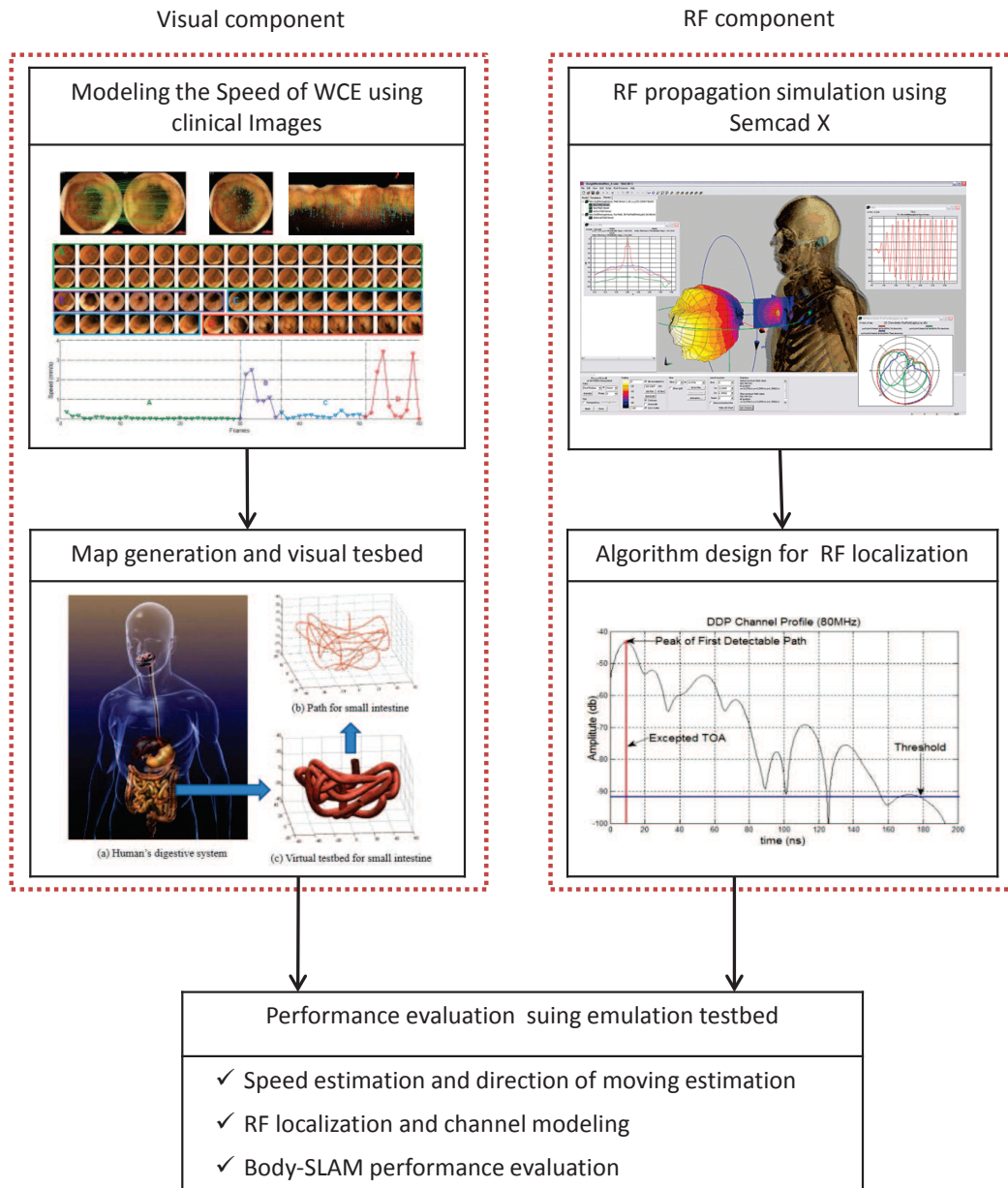
4.2.1   Visual Component

To establish the visual component of the testbed, we need to create a visual environment that is able to produce artificial images that look the same as the real endoscopic images taken by the WCE while we have fully control of the motion status of the camera.

4.2.1.1   Physical testbed

Our first attempt to emulates the digestive tract is to build a physical hardware [109]. This hardware was created by bending and twisting a 1.5 meter long 3 centimeter diametric Polyvinyl Chloride (PVC) tube. The outer surface of the tube was painted with flesh color to give it a more realistic interior look. A layer of tinfoil paper was covered around the tube to prevent outer light from transmitted into the tube and also preventing the light of camera from escaping outside.

To simulate the transition of the endoscopic capsule, we inserted a wired endoscopy camera equipped with four LED lights (as shown in Figure 4.11 (a)) into the tube with a constant step of 0.03 cm and took a picture after each step. In the endoscopic pictures, the tube surface that lied physically closer to the camera had a brighter intensity value. The brightness decreased as the distance increased and finally at the far end of the tube, which was corresponding to the center of the endoscopic pictures, a black hole would form. If the camera was about to tilt, the black hole would move toward to the edge of the endoscopic pictures. Figure 4.11 (c) indicates a test pictures take from inside the physical testbed. We can see that it looks similar with real pictures taken from the small intestine.

One big advantage of the physical testbed is that we can put it into water to simulate the liquid environment inside the small intestine. Also we could insert a antenna inside

(a)   (b)   (c)
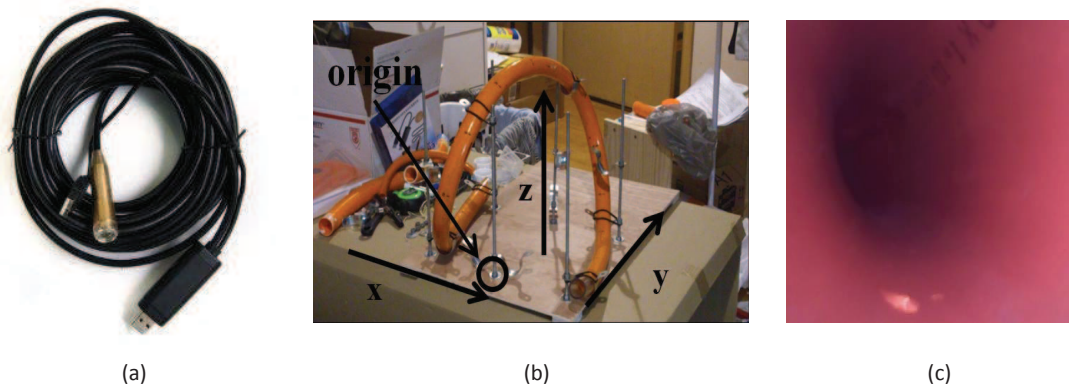
FIGURE 4.11: A physical visual model for the small intestine (a) wired endoscopic camera (b) appearance of the physical model (c) pictures taken from inside the physical model

the tube to emulate RF signal emitted by the endoscopic camera. However, there are some fundamental drawbacks with this physical model:

- The major drawback of this physical model was the restriction in camera control. After the camera was inserted into the tube, we can only control the speed of the camera by pulling the wire connected to of the camera. However, we don't have full control of other movements of the camera such as tilt and rotation.

- The PVC plastic tube is not soft enough to bend into the complicated shape (especially the sharp turn) of the small intestine. Besides, the endoscopic camera is longer than the actual size of the endoscopic capsule, this would make smooth transition of the camera challeging since it sometimes get stuck somewhere in the tube.

- Adding texture to the interior of the PVC tube is difficult. We meed map the actual texture of inside the digestive tract to the interior to the physical model to make it looks more similar to the real endoscopic images.

Due to the drawbacks stated above, an alternative way to emulate the inside environment of the small intestine was to build a virtual testbed which will be introduced in the upcoming subsection.

4.2.1.2 Virtual testbed

The major reason that we choose a virtual visual testbed is that it gives full control of the camera's motion. This is critical in validating the performance of our motion tracking algorithm. Besides, we can easily change the shape of the testbed and attach any color and texture to the testbed to make it looks more realistic. To verify the feasibility of the virtual testbed, we compared the images generated from the virtual testbed and physical testbed and make sure they share similar characteristics.

The first step is to map the physical model to the virtual 3D space [110]. To do so, we measured the x, y, and z coordinates of the pipe every 10cm and imported those coordinates into Matlab (as shown in Figure 4.12). After linking those 3D points, we got the path of the physical testbed. This path is used for generating a cylinder the same size as the physical pipe as shown in Figure 4.13. We attached the texture extracted from the real endoscopic images to the interior of this cylinder to give a realistic look as the small intestine. Using Matlab graphic toolbox, we placed a camera view point and moved it along this path and meanwhile taking pictures. An light source was placed behind the camera view point to simulate the illumination system. Some sample images generated by this virtual environment is shown in Figure 4.14. As we can see, the images generated from the virtual environment shares the same features of the images from the physical testbed, therefore, it can be used as an alternative for creating the visual component of the testbed. However, the cylinder shown in Figure 4.13 is too short and too simple

to represent the real small intestine. We need a more realistic and complicated virual

model of the digestive tract.
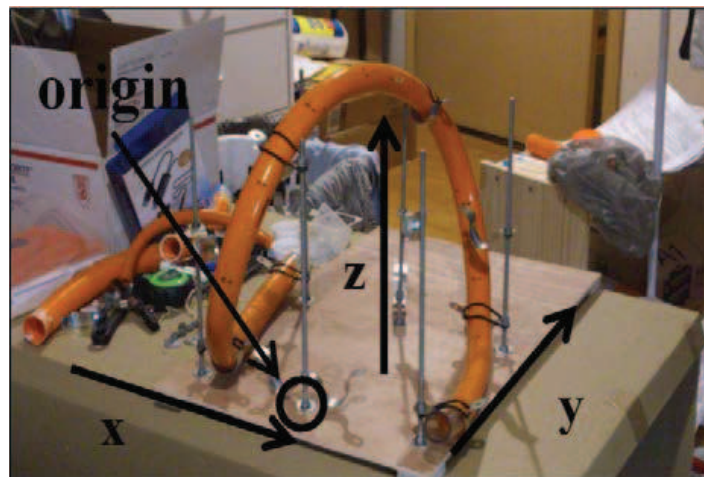
(a) Physical testbed
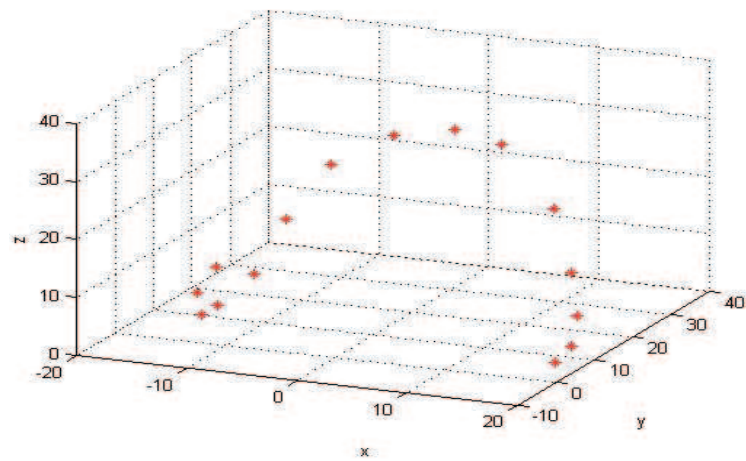
(b) Mapping the coordinates of physical testbed into Matlab
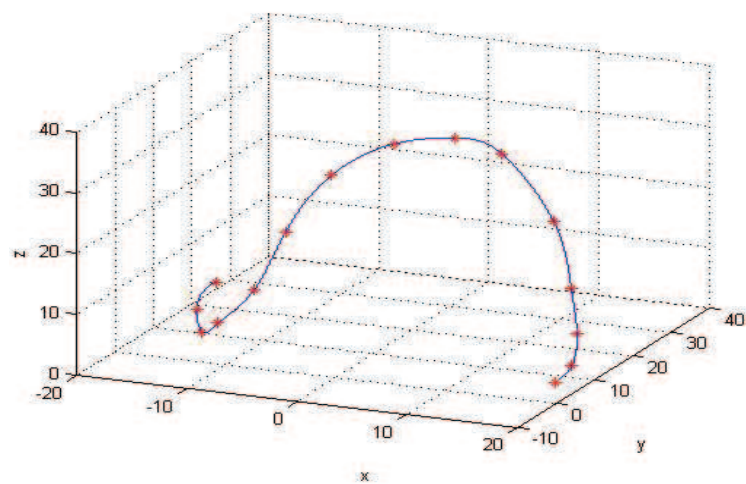
(c) Create the 3D path of the physical testbed by interpolation

FIGURE 4.12: Mapping the physical testbed into virtual 3D space
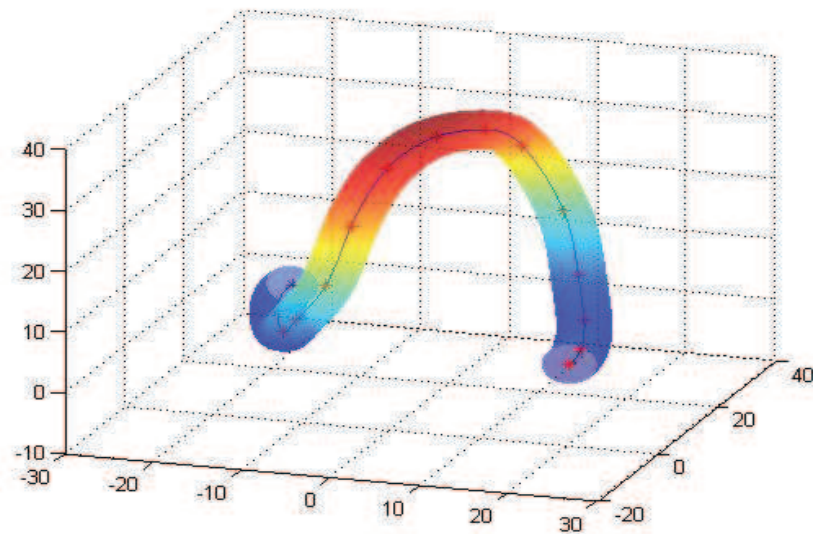
FIGURE 4.13: 3D testbed
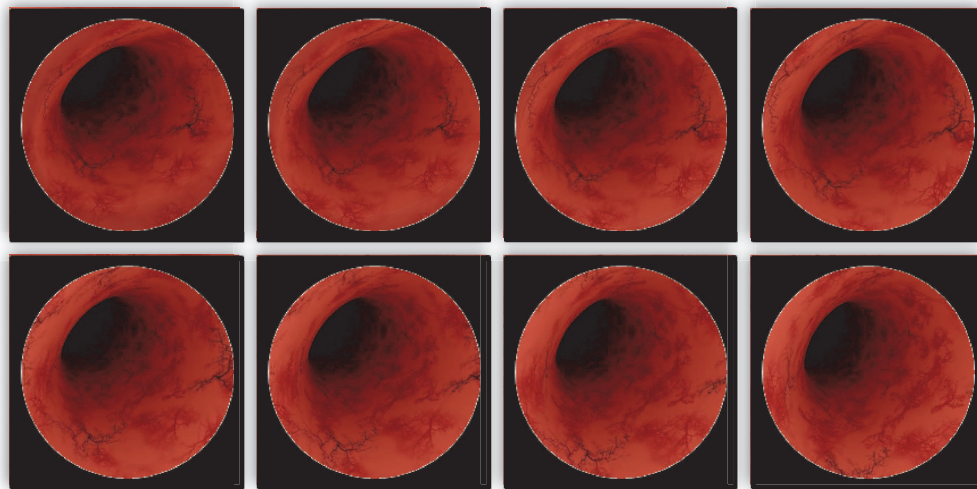
FIGURE 4.14: Emulated endoscopic images from virtual visual testbed

To create a more realistic scenario, we generated a cylindrical tube with the same size and shape of the small intestine. The actual path of this virtual test bed was extracted from a an anatomical 3D model of the small intestine shown in Figure 4.16 by using some 3D image processing techniques. For the large intestine, since it already has a very clear pattern which looks like a big hook, we applied 3D skeletonization technique [111] to extract the path of it. As for the small intestine, since the shape of the small intestine is much more complicated (the trend of the small intestine can be hardly recognized

by human eyes), we developed an element sliding technique to trace the path. The basic idea behind this technique is to define an element shape (ES) with its radius automatically adjustable to the radius of the small intestine [86]. This ES is propelled forward by a factor associated proportional to the average distance between the vertices within certain range and the physical center of the ES. As the ES goes along the small intestine, the position of its physical center is recorded and this will give us a clear path of the small intestine. The path extracted from the 3D model is shown on the bottom of Figure 4.15.

As illustrated on the top of Fig. 4.16 (c), the virtual test bed shared the same topology with the real small intestine which is intertwined back and forth. To make the interior of the testbed look more realistic, we extracted color and texture from the real endoscopic images and mapped it onto the interior surface of the tube. Similar emulation set up can be found in [108, 112, 113].
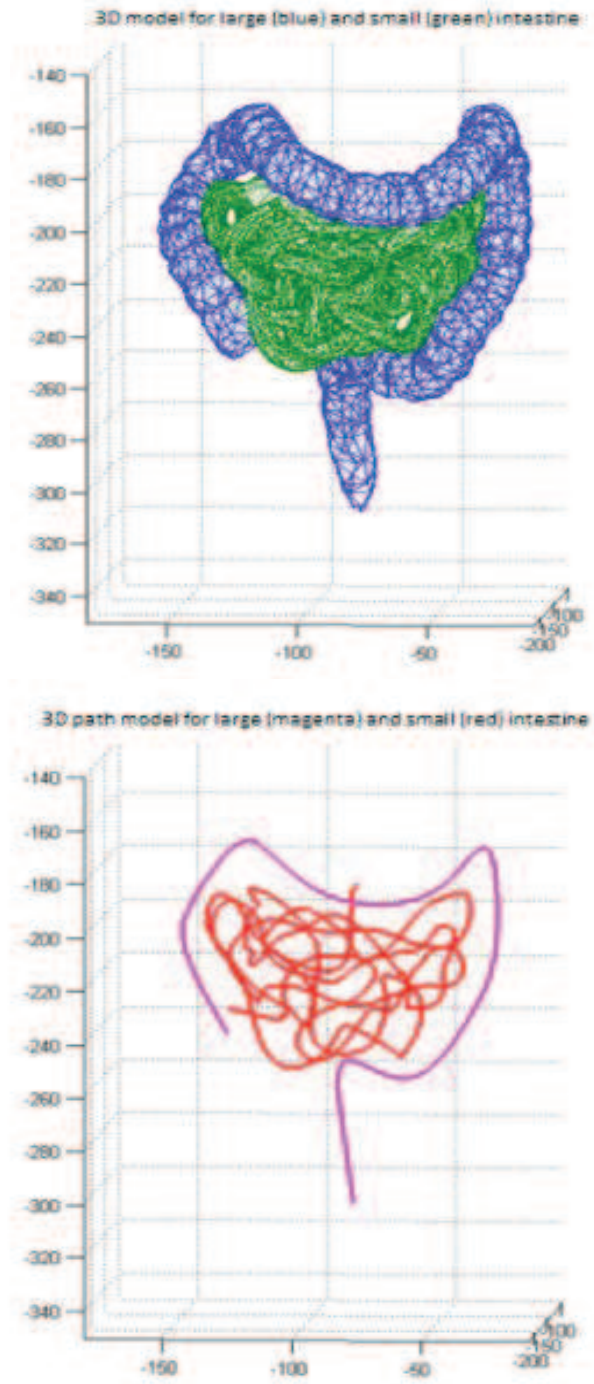
FIGURE 4.15: 3D path generation from a 3D GI tract model

(a) Human's digestive system

(b) Path for small intestine

(c) Virtual testbed for small intestine

(d) Real endoscopic images

(e) Emulated endoscopic images
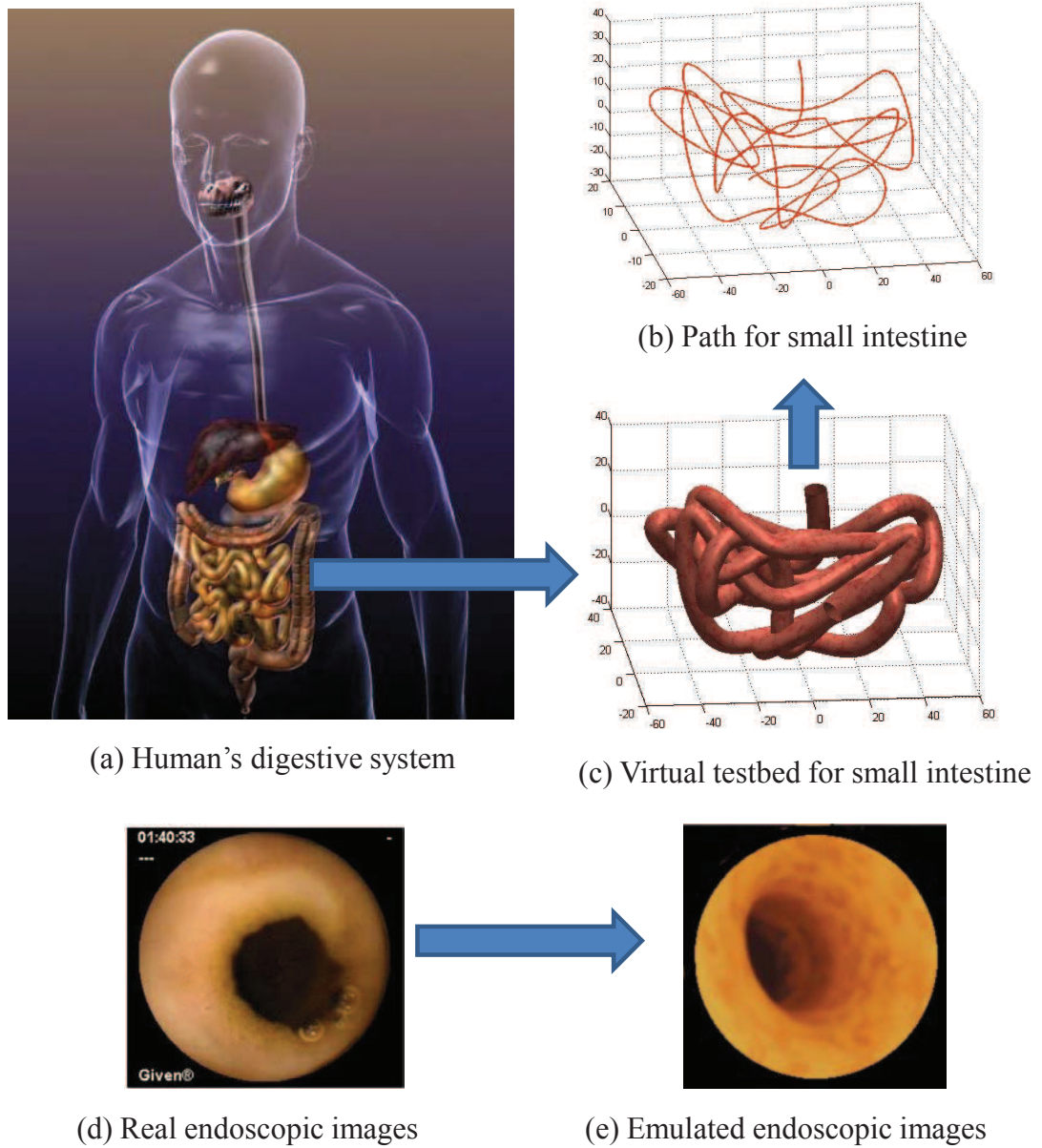
FIGURE 4.16: Emulation testbed set up

4.2.2 RF Component

Research in localization inside of the human body has reached a bottleneck due to the difficulty of conducting measurements inside the human body. Two major limitations causing difficulty are the existence of a nonhomogeneous environment and difficulties in antenna implantation inside the human body for experimental purposes. Previously, the efficiency of different simulations around a human body was assessed and theoretically analyzed [114, 115]. Phantoms with emulated tissues were used to validate surface measurements of a human body [116–118]. However, the simulation analysis of the small intestine remains to be done.

4.2.2.1 RF propagation emulation using FDTD

The RF propagating simulation was carried out using SEMCAD X [119]. SEMCAD X is a full-wave electromagnetic simulation platform based on the Finite Difference Time Domain (FDTD) method [120]. This software provides an abundant library of anatomical non-homogeneous human body models for waveform transmission problems. The models can be used to simulate wave propagation in and around the human body. Additionally, this software runs faster than other electromagnetic simulation platforms due to its algorithm optimization.

The FDTD method was first introduced by Yee in 1966 [121]. It solves Maxwell curl equations in the time domain. The FDTD method has been proven to be an effective simulation method in terms of the accuracy of obtaining electrical and magnetic field parameters. It has been widely used in indoor localization and microwave simulations [122, 123]. In this section, we present the simulation results of the waveform propagation in both homogeneous and non-homogeneous tissues. By comparing simulation results

and empirical measurements, we show that the SEMCAD X platform is a reliable tool for waveform transmission.

In SEMCAD X, we simulated the in-body waveform transmission. The bandwidth range is from 50MHz to 400MHz. We imported a model of a 34-year male and positioned two sensors internally around the human torso. We placed two edge sensors inside the human body model at the torso. One sensor acts as a transmitter (WCE) and the other sensor as a receiver. We fix the receiver at a single location and move the transmitter at increments of 1 cm as shown in Figure 4.17. Using this procedure we obtain the values of TOA and received signal strength at different distances. The results are shown in Figure 4.18.
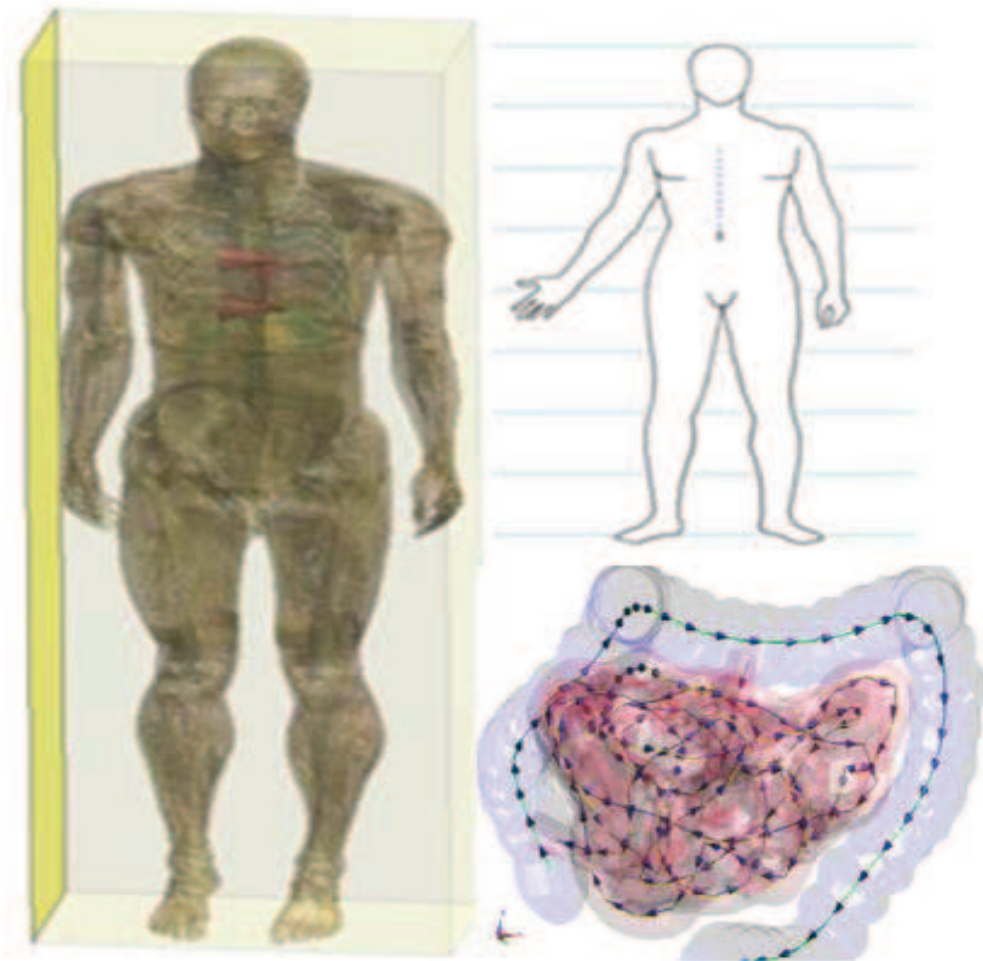
FIGURE 4.17: RF propagation setup using SEMCAD X

4.2.2.2 RSS vs ToA
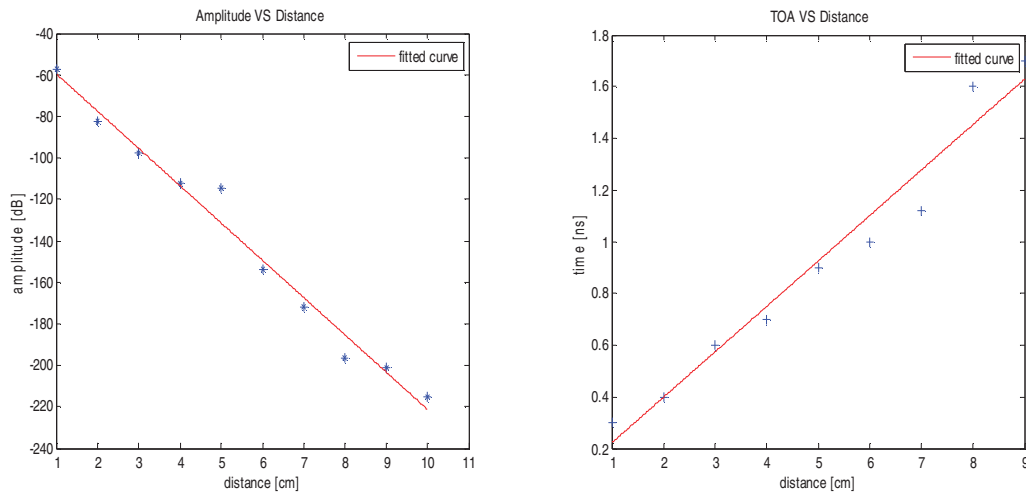
FIGURE 4.18: RSS versus distance (left) and Time-of-Arrival (TOA) versus distance (right) inside human body

To measure the statistics of the temporal and spatial behavior of the signal, we used computational techniques for direct solution of Maxwell's equations for extensive measurements of wide-band characteristics of RF signals inside the human body. We have used these techniques to find the wide-band received signal at body mounted sensors and other WCEs when a waveform is transmitted from a VCE in a specific location inside the GI tract. Then we extracted the RSS, TOA and DOA of the received wide-band signal by other capsules or by body-mounted sensors to model them for use in RF localization algorithm design.

For RSS based localization techniques, we need a path-loss model that relates the statistical behavior of the power to the distance to calculate the estimated distance of the capsule from the body-mounted sensors used as the reference point. For TOA-based localization algorithms, we need a model for the multipath arrival and the relationship between distance measurement error and the bandwidth of the system to account for the measurement noise and various biases in distance estimation from TOA measurements.

The current body of literature only provides a few path-loss models for implant communication applications. Modeling of the effects of multipath on TOA- and DOA-based localization is at its infancy and new models for these purposes are needed.

4.3   Performance Evaluation for Body-SLAM

We first test the performance of the motion tracking algorithm using the visual testbed. The artificial images were generated by moving the camera view point along a preknown path. During the transition of the capsule, it took pictures at resolution of 420×560 pixels. For the purpose of inverse cylindrical projection, we only used the square portion (1:420, 71:490) of the original image. The size of the unrolled domain was set to be $500(L) \times 300(H)$ pixels. The radius of the cylindrical tube was set to be 1.5 cm, which was close to the actual radius of the small intestine.

Fig. 4.19 shows some artificial images generated by the visual testbed and typical movements that were detected by the proposed motion tracking algorithm. Fig. 4.19 (a) shows a scenario that the capsule moves forward. It can be seen that the displacements of FPs are originally pointed to the outer ring in cylindrical image domain while stacked vertically with their orientation pointing upward when mapped into the unrolled image domain. The magnitude of displacements in the unrolled domain reflects the speed of moving of the capsule. Similarly, if the capsule moves backward (shown in Fig. 4.19 (b)), motion vectors in the cylindrical image domain are pointed to the center and correspondingly, when mapped into the unrolled image domain, they are vertically pointing downward. Fig. 4.19(c) shows a case when the capsule rotated, motion vectors in the cylindrical image domain formed a circle around the focal axis and the corresponding

motion vectors in the unrolled domain were horizontally pointing to the right indicating a clockwise rotation. The magnitude of $\triangle x'$ reveals the rotation angle. Finally in Fig. 4.19(d), the capsule tilted toward ϕ during the transition, thus the magnitudes of the motion vectors in this area were smaller than the others. The difference in magnitude indicates the degree that the capsule tilts.

Cylindrical image domain   Unrolled image domain

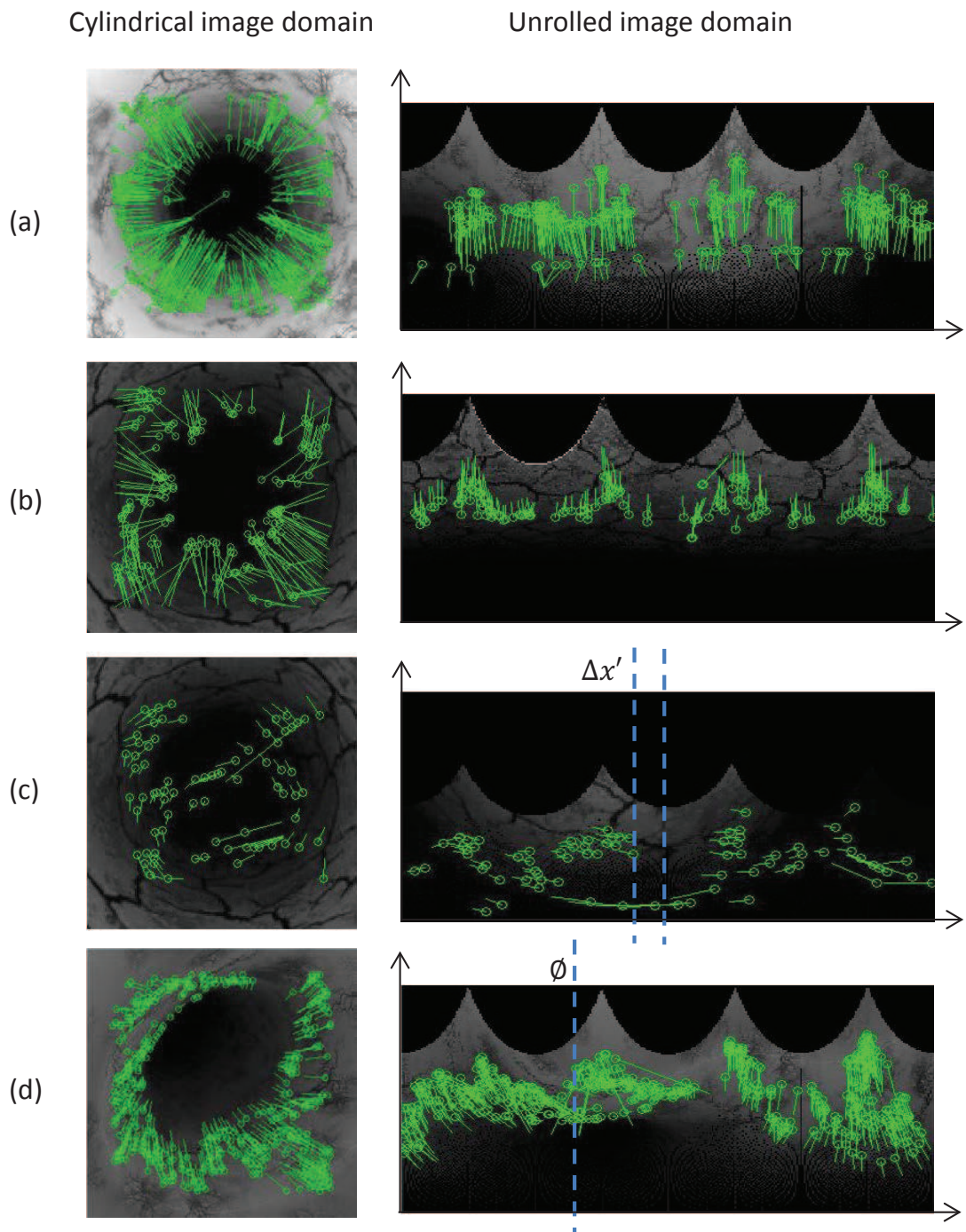(a)

(b)

(c)

$\Delta x'$

(d)

$\emptyset$

FIGURE 4.19: Typical movements detected by the proposed motion tracking algorithm

Experimental results show that the linear transitions and rotations can be accurately inferred from the vertical component and horizontal component of the motion vectors in the unrolled domain, respectively. Meanwhile, differences in the magnitude of motion vectors reflect tilting direction of the capsule. Based on the quantitative calculation, the tracking process is implemented as follows: given the initial position of the capsule, the subsequent positions of the capsule are estimated by multiply the current position with the transition matrix, which consists of distance, rotation, and tilt angle. The tracking results are illustrated in Figure 4.20 compared with the ground truth path. Also, the MSE of estimated position of the capsule for each step was plotted in Figure 4.21. It can be seen from the plotting that the motion estimation for the first 50 steps were very accurate, whose average MSEs were below 1 cm. However, the error increased as the step goes further. This was due to the accumulative characteristic of all motion tracking techniques. In every motion tracking technique, the next state is purely decided by the current state plus the current transition information, which was estimated from the displacements of FPs between consecutive image frames. If an error happened in the estimation of this transition information, even with very little magnitude, this error would accumulate and the overall error would keep increasing. This is what we call "drifting effect" in Robotics. Thus, we cannot judge the performance of a motion tracking algorithm based on the overall accuracy. Instead, we should evaluate the performance of an algorithm by measuring the accuracy of estimation within each step. Statistics in Table 4.1 show that our proposed algorithm worked accurately in calculating the proceeding distance and rotation angle for each step. The average distance error was 0.04 cm which was way below the unit step size, and the average rotation error was 1.8^o, which was also very small compared with average rotation angle of 7.8^o. The estimation of tilt was not as accurate because the differences in motion vectors were not

TABLE 4.1: Motion tracking performance for each step

	Average estimates	Average error
Distance	0.41 cm	0.04
Rotation	7.8^o	1.8^o
Tilt	4.3^o	3.0^o

always obvious and the range of smaller motion vectors may cover up to more than 45

degree of x' axis. Therefore, it was very difficult to quantize the tilt angle which leaded

to great calculation errors.
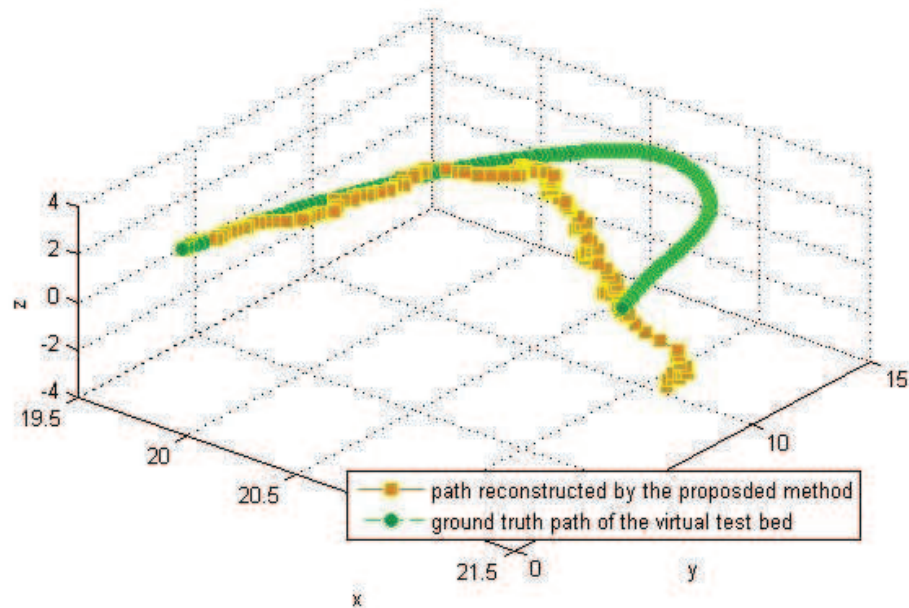
FIGURE 4.20: Result of the motion tracking compared with ground truth
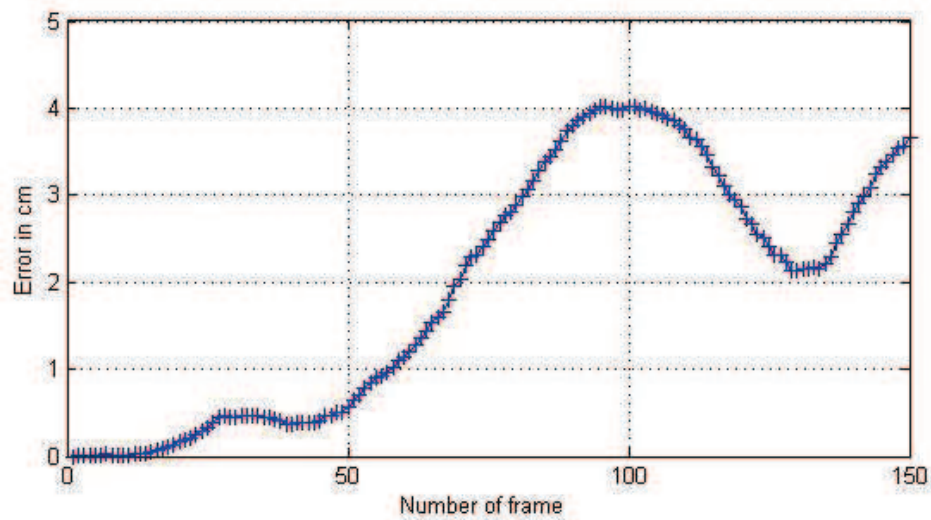
FIGURE 4.21: Mean square error (MSE) in the motion racking process

The results of motion tracking using images, RF localization and the proposed hybrid localization are given in Fig. 4.22. It can be seen from Fig. 4.22 (a) that the results of RF localization (represented in green triangles) are scattered all around the small intestine with relative large error (6.8 cm on average). This is because the RF channel suffers shadow fading and non-homogeneity of the body tissues. However, the good part of RF localization is its independent characteristics. Each measurement is an isolated procedure which cannot be affected by the previous measurements. Therefore, the localization error would not accumulate as the capsule moves along (shown in Fig. 4.22(b)).

The result of the camera motion tracking algorithm is shown in gray line in Fig. 4.22 (a). It shows that when using this algorithm alone, the estimated positions are continuous and the overall trend of the trajectory matches the ground truth path (shown in black line in Fig. 4.22 (a)) of the small intestine. However, as the capsule moves along, the localization error increases. It can be seen from Fig. 4.22 (b), after about 15 steps, the localization errors of motion tracking reaches almost the same level of RF localization and it keeps growing until explode (up to 70 cm). This is due to the accumulative characteristic of the motion tracking algorithm. In the motion tracking algorithm, the next motion state

is highly dependent on the current motion state plus the current transition information, which is estimated from the displacements of FPs between consecutive image frames. If an error happens during the estimation of this transition matrix, even with very little magnitude, the error would accumulate and the overall error would keep increasing. This is what we call "drifting effect" in Robotics. Thus, we cannot judge the performance of a motion tracking algorithm based on the overall accuracy. Instead, we should evaluate the performance of an algorithm by measuring the accuracy of estimation within each step. Statistics in Table I show that our proposed algorithm worked accurately in calculating the transition speed of the capsule and direction angles for each step. The average distance error was 0.04 cm which was way below the unit step size, and the average rotation error was 1.8^o, which was also very small compared with average rotation angle of 7.8^o. The estimation of tilt was not as accurate because the differences in motion vectors were not always obvious and the range of smaller motion vectors may cover up to more than 45 degree of x' axis.
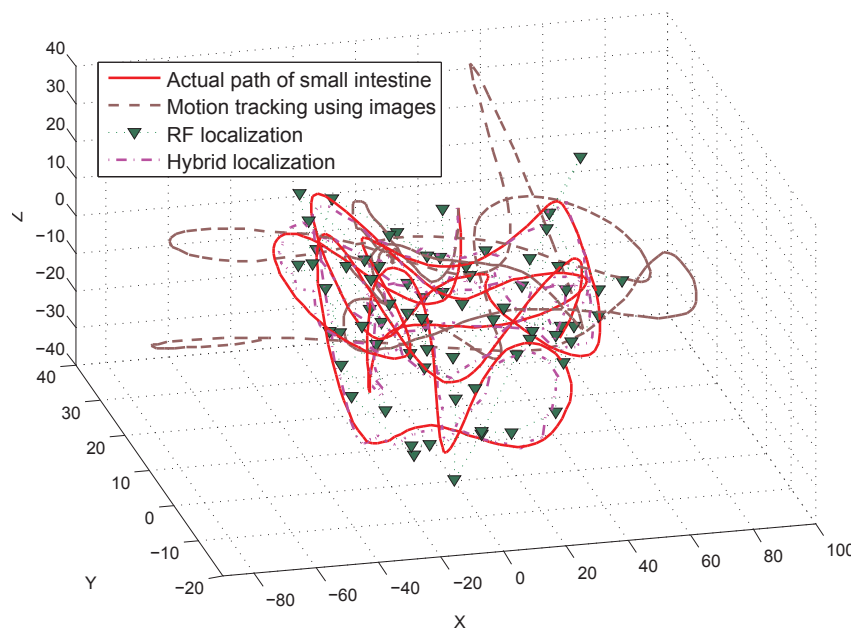
(a) Localization results of different algorithms
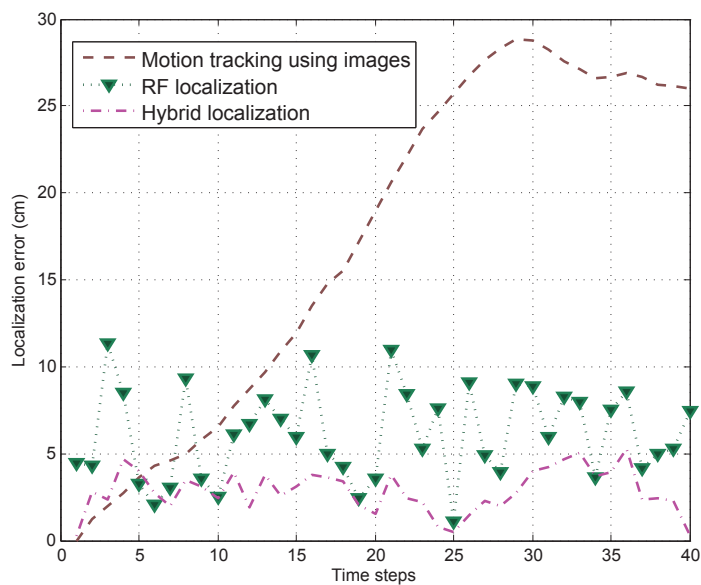
(b) Evolution of localizatoin error as the capsule moves

FIGURE 4.22: Localization results of different algorithms and performance evaluation

Finally we evaluated the performance of the proposed Body-SLAM localization algorithm. The results are shown in purple line in Fig. 4.22 (a). It shows after the combination of motion tracking and RF signals, the hybrid localization is able to achieve more continuous position estimation of the capsule and the reconstructed path that the capsule has traveled matches the ground truth path of the small intestine very well. From Fig. 4.22 (b) we can see that, compared with the existing RSS based localization system, the localization error of hybrid localization stays stable at a very low level (2.3 cm on average) and the error would not increase as the capsule moves along. The error distribution and CDF plot of the above three algorithms are given in Fig. 4.23 and Fig. 4.24, respectively. From both statistical plots we can see that the localization accuracy of the proposed hybrid localization is much better than the traditional RSS based RF localization. Since the diameter of small intestine is approximately 2.5-3 cm, the localization accuracy that the proposed hybrid localization provides meets the requirement of WCE application.
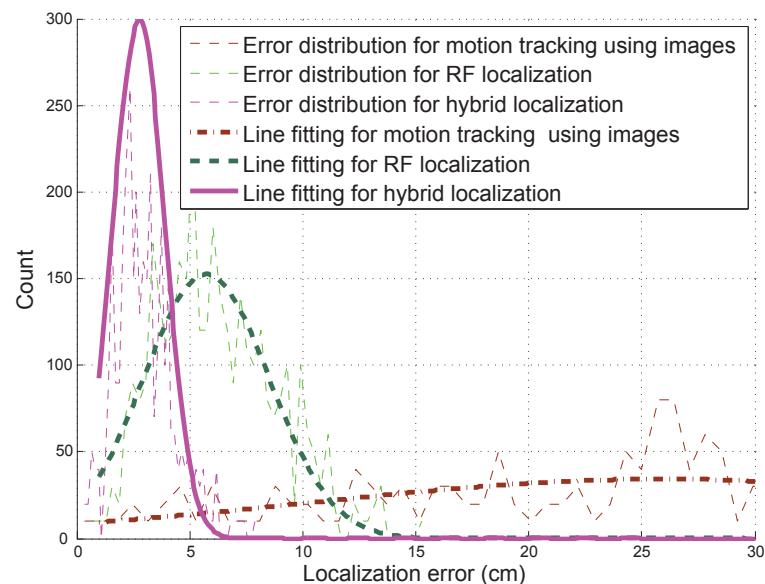
FIGURE 4.23: Error distributions of different algorithms
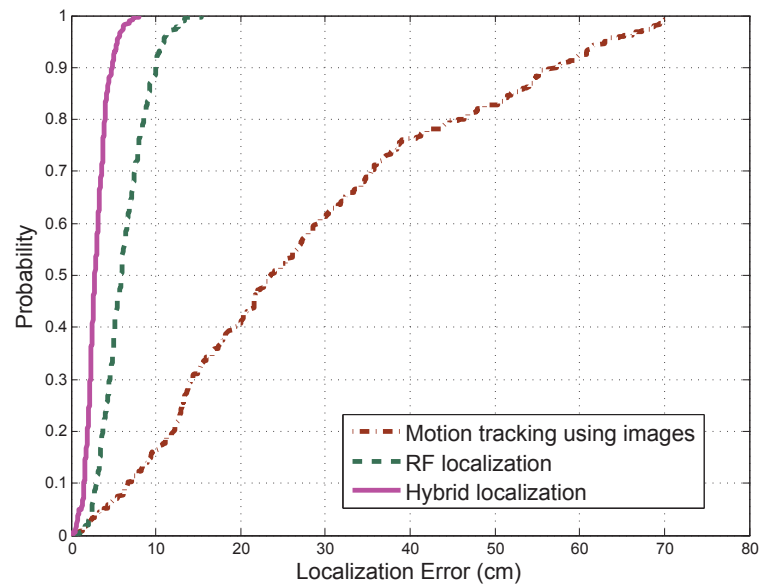
FIGURE 4.24: Performance evaluation by CDF plot of different algorithms

Chapter 5

Conclusion and Future Direction

This chapter provides an overall conclusion of the dissertation and points out the possible future directions of our research.

5.1　Conclusion

In this dissertation, we presented a Body-SLAM technique that integrates motion information extracted from the video source and RF signal emitted by the capsule to enhance to the localization accuracy of the WCE and meanwhile reconstruct the path the capsule has traveled. The major contribution of this work is that we demonstrated the potential of using endoscopic images to aid the RF localization and possibility to map inside of human body. The proposed motion tracking technique is purely based on the image sequence that captured by the video camera which is already equipped on the capsule, thus, no extra components such as Inertial measurement units (IMUs) or magnetic coils are needed. Another contribution is that we provided an uniform performance evaluation platform to test the performance of localization algorithms inside

human body. Both empirical and experimental results have shown that by combining the motion information with RF measurements, the proposed Body-SLAM algorithm is able to provide accurate, smooth and continuous localization and mapping results that meet the requirement of WCE application.

5.2 Future Direction

In the future, we will focus on refining this Body-SLAM algorithm in the following ways:

- Enhance the performance of the Body-SLAM algorithm using more complicated Extended Kalman Filter (EKF).

- Verify the performance of the Body-SLAM algorithm using more clinical data, including experiments on animal and human subjects.

- Refine the virtual testbed by adding intestinal motilities.

A proposal based on the material of this dissertation has been filed to the NSF to pursuit funding to continue this research.

Appendix A

Appendix: Full Publication List

A.1   Related to the Thesis

1. K. Pahlavan, G. Bao, Y. Ye, S. Makarov, U. Khan ... K. Sayrafian, "Rf localization for wireless video capsule endoscopy". International Journal of Wireless Information Networks, Vol.19 (4), pp.326-340, 2012.

2. G. Bao, Y. Ye, U. Khan, X. Zheng and K. Pahlavan, "Modeling of the Movement of the Endoscopy Capsule inside GI Tract based on the Captured Endoscopic Images", The 2012 International Conference on Modeling, Simulation and Visualization Methods (MSV), Las Vegas, USA, July, 2012.

3. G. Bao and K. Phalavan, "Motion Estimation of the Endoscopy Capsule using Region-based Kernel SVM Classifier", 2013 IEEE International Conference on Electro/Information Technology (EIT), Rapid City, SD, May 9-11, 2013.

4. G. Bao, L. Mi and K. Phalavan, "Emulation on Motion Tracking of Endoscopic Capsule inside Small Intestine", 2013 World Congress in Computer Science,Computer Engineering, and Applied Computing (WORLDCOMP'13), Las Vegas, USA, 2013.

5. G. Bao, L. Mi and K. Phalavan, "A Video Aided RF Localization Technique for the Wireless Capsule Endoscope (WCE) inside Small Intestine", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

6. L. Mi, G. Bao and K. Phalavan, "Design and Validation of a Virtual Environment for Experimentation inside the Small Intestine", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

7. R. Fu, G. Bao and K. Pahlavan, "Activity Classification with Empirical RF Propagation Modeling", 8th International Conference on Body Area Networks, Boston, Massachusetts, United States, September 30 - October 2, 2013.

8. L. Mi, G. Bao and K. Pahlavan, "Geometric Estimation of Intestinal Contraction for Motion Tracking of Video Capsule Endoscope", SPIE Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling, San Diego, California, February 15-20, 2014.

9. G. Bao, L. Mi, Y. Geng and K. Pahlavan, "A Computer Vision based Speed Estimation Technique for Localizing the Wireless Capsule Endoscope inside Small Intestine," submitted to Signal Processing Letters, IEEE, April, 2014.

10. G. Bao, L. Mi, Y. Geng, M. Zhou and K. Pahlavan, "A Video-based Speed Estimation Technique for Localizing the Wireless Capsule Endoscope inside Gastrointestinal Tract, " submitted to IEEE Engineering in Medicine and Biology Society (EMBC 14), March, 2014.

11. M. Zhou, G. Bao and K. Pahlavan, "Mutual Information based Motion Tracking Technique for the WCE inside Large Intestine", submitted to IEEE Engineering in Medicine and Biology Society (EMBC 14), March, 2014.

12. G. Bao, L. Mi and K. Pahlavan, "Hybrid Localization of Micro-robotic Endoscopic Capsule inside Small Intestine by Data Fusion of Vision and RF Sensors", submitted to Sensor Journal, IEEE, March, 2014.

13. K. Pahlavan, G. Bao and L. Mi. "Body-slam: Simultaneous localization and mapping inside the human body". Keynote speech, 8th International Conference on Body Area Networks (BodyNets), Boston, MA, September 30-October-2, 2013.

A.2   Not Related to the Thesis

1. Fu R, Bao G, Pahlavan K. "Activity classification with empirical RF propagation modeling in body area networks," Proceedings of the 8th International Conference on Body Area Networks. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013: 296-301.

2. Fu R, Bao G, Ye Y, et al. "Heterogeneous Cooperative Localization for Social Networks," International Journal of Wireless Information Networks, 2013, 20(4): 294-305.

3. Zheng X, Bao G, Fu R, et al. "The Performance of Simulated Annealing Algorithms for Wi-Fi Localization using Google Indoor Map", Vehicular Technology Conference (VTC Fall), 2012 IEEE. IEEE, 2012: 1-5.

4. Salari E, Bao G. "Super-resolution using an Enhanced Papoulis-Gerchberg Algorithm," IET image processing, 2012, 6(7): 959-965.

5. Salari E, Bao G. "Pavement Distress Detection and Classification using Feature Mapping," Electro/Information Technology (EIT), 2010 IEEE International Conference on. IEEE, 2010: 1-5.

6. Salari E, Bao G. "Pavement Distress Detection and Severity Analysis," SPIE Electronic Imaging, International Society for Optics and Photonics, 2011: 78770C-78770C-10.

7. Salari E, Bao G. "Automated Pavement Distress Inspection based on 2D and 3D Information," Electro/Information Technology (EIT), 2011 IEEE International Conference on. IEEE, 2011: 1-4.

Appendix B

Appendix Tutorial

```
% To download the path , please go to http://www.cwins.wpi.edu/personnel/guanqun.html


clc;

close all;

clear all;

step =3;

Ps=load('path_small_intestine.txt'); %load the path for small intestine

Si=load('small_intestine.txt'); %load the model for small intestine


C=load('ex.txt');

li_x=C(1:step:end);

li_y=C(2:step:end);

li_z=C(3:step:end);

%h1=plot3(li_x,li_y,li_z,'r.');

p1=[li_x' li_y' li_z'];


figure;

hold on;


[t1]=MyCrustOpen(p1);

surf3=trisurf(t1,p1(:,1),p1(:,2),p1(:,3),'facecolor','b','edgecolor','b')%plot della superficie

alpha(surf3,0.01)
```

```matlab
% rotation and translation matrix

Rz=[0 -1 0; 1 -1 0;0 0 1];

Ry=[0.86 0 -0.5; 0 1 0; 0.5 0 0.86];

Rx=[1 0 0; 0 0.94 0.34; 0 -0.34 0.94];

T=[369 -400 -478];


Ps=load('path_small_intestine.txt'); %load the path for small intestine

Si=load('small_intestine.txt'); %load the model for small intestine


% Ps(:,1)=Ps(:,2);

Temp=ones(length(Ps),3);

Ps=Ps*Rz*Ry*Rx*1.6+Temp(:,3)*T;


% Si(:,1)=Si(:,2)*0.7;

Temp=ones(length(Si),3);

Si=Si*Rz*Ry*Rx*1.6+Temp(:,3)*T;


[t1]=MyCrustOpen(Si);

temp=rgb('darkgreen');

surf3=trisurf(t1,Si(:,1),Si(:,2),Si(:,3),'facecolor','white','edgecolor',[temp]); % plot della s

alpha(surf3,0.01);


Pl=load('path_large_intestine.txt');%load the path for large intestine

Li=load('large_intestine.txt');%load the model for large intestine


% Pl(:,1)=Pl(:,2)*0.7;

Temp=ones(length(Pl),3);

Pl=Pl*Rz*Ry*Rx*1.6+Temp(:,3)*T;


% Li(:,1)=Li(:,2)*0.7;

Temp=ones(length(Li),3);

Li=Li*Rz*Ry*Rx*1.6+Temp(:,3)*T;
```

```
[t2]=MyCrustOpen(Li);

temp=rgb('darkgreen');

surf3=trisurf(t2,Li(:,1),Li(:,2),Li(:,3),'facecolor','white','edgecolor',[temp]); %plot della su

alpha(surf3,0.01);


plot3(Ps(:,1),Ps(:,2),Ps(:,3),'r .')

plot3(Pl(:,1),Pl(:,2),Pl(:,3),'r .')


view(-60,60)

view(40,35)
```
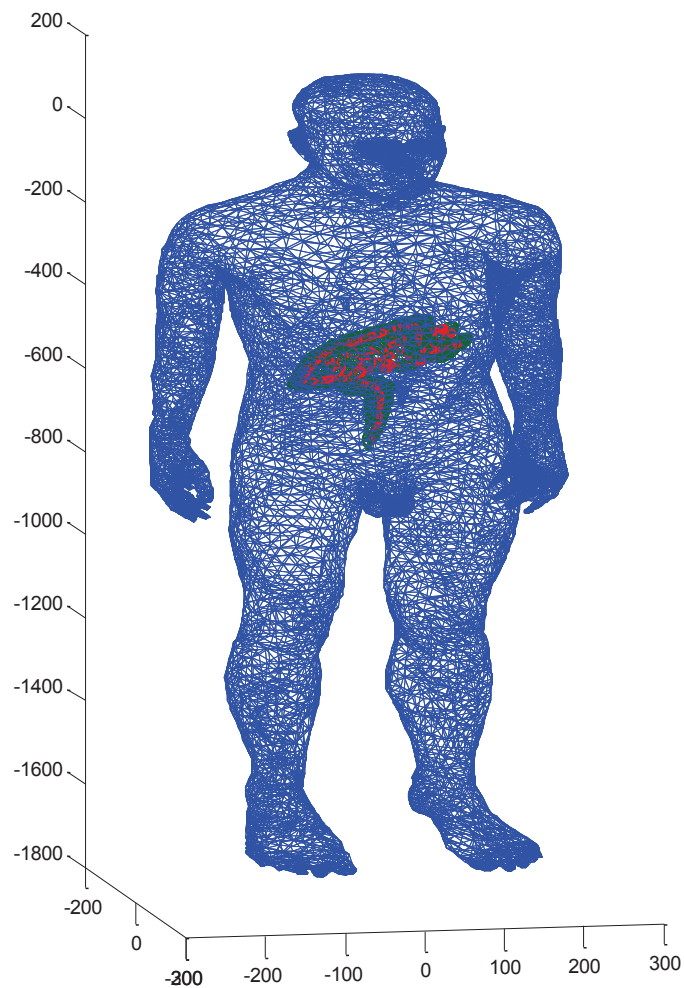

FIGURE B.1: Generating the path inside human body

```
% unrolling the image

% [xp, yp] are coordiniates of feature points after projection
```

```
% [x, y] are the coordinates of feature points in the original image

% [x0, y0] is the center of the original image

% B is the heigth of the unrolled image

% L is the length of the unrooled image


function [xp,yp] = inverse_proj(x,y,x0,y0,B,L)

        r = sqrt((x-x0)^2+(y-y0)^2);

        if(y>=y0&&x==x0)

            theta = 0;

        elseif(y>y0&&x<x0)

            theta = atan(abs((x-x0)/(y-y0)));

        elseif(y==y0&&x<=x0)

            theta = pi/2;

        elseif(y<y0&&x<x0)

            theta = atan(abs((y-y0)/(x-x0))) + pi/2;

        elseif(y<=y0&&x==x0)

            theta = pi;

        elseif(y<y0&&x>x0)

            theta = atan(abs((x-x0)/(y-y0))) + pi;

        elseif(y>=y0&&x==x0)

            theta = 3/2*pi;

        elseif(y>y0&&x>x0)

            theta = atan(abs((y-y0)/(x-x0))) + 3/2*pi;

        else

            theta=2*pi;

        end

        xp = round(L*theta/(2*pi))+1;

        yp = round(B-r)+1;

end
```

```
% Motion estimation for WCE


DistRatio= 1.25; % Increase value to decrease the number of matches

NumOrientBins= 16;


N_frame=6000;
```

Cylindrical image   "Unrolled" image
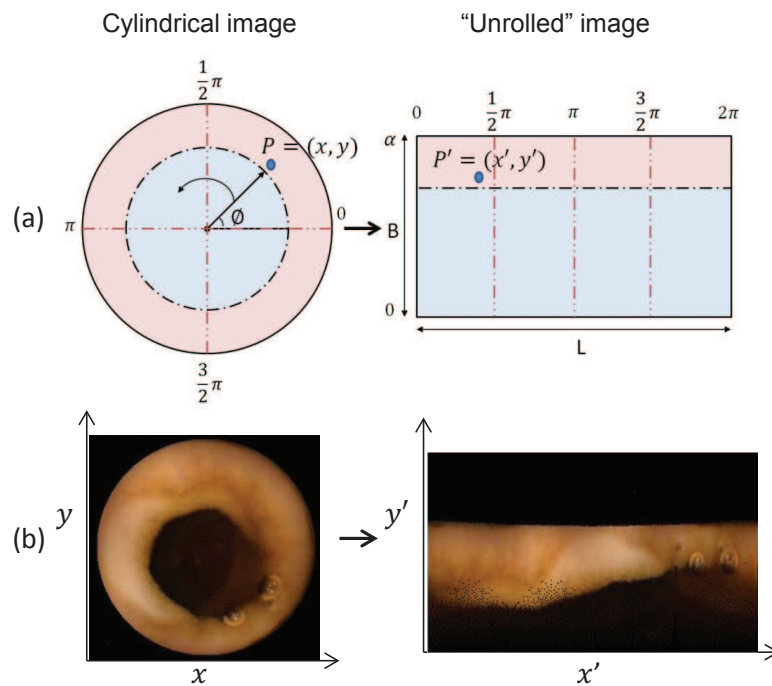
FIGURE B.2: Illustration of unrolling image

```
M_Vector = zeros(1, N_frame);

figure(1)


for n=1:N_frame

    % Load images

    rootname = 'motion/1 (';        % Root filename       % Root filename

    extension = '.jpg';

    filename1 = [rootname, num2str(n),')', extension];

    filename2 = [rootname, num2str(n+1),')', extension];


    A = imreadbw(filename1);

    B = imreadbw(filename2);


    J1 = imresize(A, 0.7);

    J2 = imresize(B, 0.7);


    I1 = J1(20:152,25:157); % get rid of the black part

    I2 = J2(20:152,25:157); % get rid of the black part
```

```
    I1=I1-min(I1(:));

    I1=I1/max(I1(:));

    I2=I2-min(I2(:));

    I2=I2/max(I2(:));


    fprintf('Computing frames and descriptors (~0.5 minutes).\n');

    [frames1,descr1]= sift(I1, 'NumOrientBins', NumOrientBins);

    [frames2,descr2]= sift(I2, 'NumOrientBins', NumOrientBins);


    [inds ratios] = SiftRatioMatch(sqrt(descr1), frames1, sqrt(descr2), frames2, DistRatio);


    matches= zeros(2,sum(inds~=0));

    i= 1;

    for r=1:size(descr1,2)

        if (inds(r)~=0)

            matches(1,i)= r;

            matches(2,i)= inds(r);

            i=i+1;

        end

    end


    M_Vector(n) = plotMotionVectors(I1, I2, frames1(1:2,:), frames2(1:2,:), matches, 'Stacking',

    drawnow;

    clf;

end
```

```
function [x,y,p] = sample_lds3d(F, H, Q, R, init_state, T, models, G, u)

% sample_lds3d model the movement of the WCE in a linear dynamical system.

% [x,y] = switching_lds_draw(F, H, Q, R, init_state, models, G, u)

%

%   x(t+1) = F*x(t) + G*u(t) + w(t),  w ~ N(0, Q),  x(0) = init_state

%   y(t) =   H*x(t) + v(t),  v ~ N(0, R)

%

% Input:

% F(:,:,i) - the transition matrix for the i'th model

% H(:,:,i) - the observation matrix for the i'th model
```

```
% Q(:,:,i) - the transition covariance for the i'th model

% R(:,:,i) - the observation covariance for the i'th model

% init_state(:,i) - the initial mean for the i'th model

% T - the num. time steps to run for

%

% Optional inputs:

% models(t) - which model to use at time t. Default = ones(1,T)

% G(:,:,i) - the input matrix for the i'th model. Default = 0.

% u(:,t)   - the input vector at time t. Default = zeros(1,T)

%

% Output:

% x(:,t)    - the hidden state vector at time t.

% y(:,t)    - the observation vector at time t.



if ~iscell(F)

  F = num2cell(F, [1 2]);

  H = num2cell(H, [1 2]);

  Q = num2cell(Q, [1 2]);

  R = num2cell(R, [1 2]);

end



M = length(F);

%T = length(models);



if nargin < 7,

  models = ones(1,T);

end

if nargin < 8,

  G = num2cell(repmat(0, [1 1 M]));

  u = zeros(1,T);

end



[os ss] = size(H{1});

state_noise_samples = cell(1,M);

obs_noise_samples = cell(1,M);
```

```
for i=1:M

  state_noise_samples{i} = sample_gaussian(zeros(length(Q{i}),1), Q{i}, T)';

  obs_noise_samples{i} = sample_gaussian(zeros(length(R{i}),1), R{i}, T)';

end


x = zeros(ss, T);

y = zeros(os, T);


m = models(1);

x(:,1) = init_state(:,m);

p(:,1) = init_state(:,m);

y(:,1) = H{m}*x(:,1) + obs_noise_samples{m}(:,1);


for t=2:T
%   m = models(t);

  p(:,t) = F{t}*p(:,t-1);

  x(:,t) = F{t}*x(:,t-1) + G{t}*u(:,t-1) + state_noise_samples{t}(:,t);

  y(:,t) = H{t}*p(:,t)  + obs_noise_samples{t}(:,t);

end
```

```
function [x, V, VV, loglik] = kalman_filter(y, A, C, Q, R, init_x, init_V, varargin)
% Kalman filter.
% [x, V, VV, loglik] = kalman_filter(y, A, C, Q, R, init_x, init_V, ...)
%
% INPUTS:
% y(:,t)   - the observation at time t
% A - the system matrix
% C - the observation matrix
% Q - the system covariance
% R - the observation covariance
% init_x - the initial state (column) vector
% init_V - the initial state covariance
%


[os T] = size(y);

ss = size(A,1); % size of state space
```

```
% set default params
% model = ones(1,T);
model = 1:T;
u = [];
B = [];
ndx = [];


args = varargin;
nargs = length(args);
for i=1:2:nargs
  switch args{i}
    case 'model', model = args{i+1};
    case 'u', u = args{i+1};
    case 'B', B = args{i+1};
    case 'ndx', ndx = args{i+1};
    otherwise, error(['unrecognized argument ' args{i}])
  end
end


x = zeros(ss, T);
V = zeros(ss, ss, T);
VV = zeros(ss, ss, T);


loglik = 0;
for t=1:T
  m = model(t);
  if t==1
    %prevx = init_x(:,m);
    %prevV = init_V(:,:,m);
    prevx = init_x;
    prevV = init_V;
    initial = 1;
  else
    prevx = x(:,t-1);
    prevV = V(:,:,t-1);
```

```matlab
      initial = 0;
  end

  if isempty(u)
    [x(:,t), V(:,:,t), LL, VV(:,:,t)] = ...
        kalman_update(A(:,:,m), C(:,:,m), Q(:,:,m), R(:,:,m), y(:,t), prevx, prevV, 'initial', i
  else

    if isempty(ndx)
      [x(:,t), V(:,:,t), LL, VV(:,:,t)] = ...
          kalman_update(A(:,:,m), C(:,:,m), Q(:,:,m), R(:,:,m), y(:,t), prevx, prevV, ...
                        'initial', initial, 'u', u(:,t), 'B', B(:,:,m));
    else

      i = ndx{t};
      % copy over all elements; only some will get updated
      x(:,t) = prevx;
      prevP = inv(prevV);
      prevPsmall = prevP(i,i);
      prevVsmall = inv(prevPsmall);
      [x(i,t), smallV, LL, VV(i,i,t)] = ...
          kalman_update(A(i,i,m), C(:,i,m), Q(i,i,m), R(:,:,m), y(:,t), prevx(i), prevVsmall, ..
                        'initial', initial, 'u', u(:,t), 'B', B(i,:,m));
      smallP = inv(smallV);
      prevP(i,i) = smallP;
      V(:,:,t) = inv(prevP);
    end

  end

  loglik = loglik + LL;

end
```

References

[1] Kaveh Pahlavan. Wireless communications for office information networks. *Communications Magazine, IEEE*, 23(6):19–27, 1985.

[2] Min Chen, Sergio Gonzalez, Athanasios Vasilakos, Huasong Cao, and Victor C Leung. Body area networks: A survey. *Mobile Networks and Applications*, 16(2): 171–193, 2011.

[3] Arthur W Astrin, LI Huan-Bang, and Ryuji Kohno. Standardization for body area networks. *IEICE transactions on communications*, 92(2):366–372, 2009.

[4] K Pahlavan, G Bao, Y Ye, S Makarov, U Khan, P Swar, D Cave, A Karellas, P Krishnamurthy, and K Sayrafian. Rf localization for wireless video capsule endoscopy. *International Journal of Wireless Information Networks*, 19(4):326–340, 2012.

[5] Ruijun Fu, Guanqun Bao, and Kaveh Pahlavan. Activity classification with empirical rf propagation modeling in body area networks. In *Proceedings of the 8th International Conference on Body Area Networks*, pages 296–301. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.

[6] N Vakil. Measurement of lesions by endoscopy: an overview. *Endoscopy*, 27(09): 694–697, 1995.

[7] Guido Costamagna, Saumil K Shah, Maria Elena Riccioni, Francesca Foschia, Massimiliano Mutignani, Vincenzo Perri, Amorino Vecchioli, Maria Gabriella Brizi, Aurelio Picciocchi, and Pasquale Marano. A prospective trial comparing small bowel radiographs and video capsule endoscopy for suspected small bowel disease. *Gastroenterology*, 123(4):999–1005, 2002.

[8] Emanuele Rondonotti, Juan Manuel Herrerias, Marco Pennazio, Angel Caunedo, Miguel Mascarenhas-Saraiva, and Roberto de Franchis. Complications, limitations, and failures of capsule endoscopy: a review of 733 cases. *Gastrointestinal endoscopy*, 62(5):712–716, 2005.

[9] Daniel S Mishkin, Ram Chuttani, Joseph Croffie, James DiSario, Julia Liu, Raj Shah, Lehel Somogyi, William Tierney, Louis M Wong Kee Song, and Bret T Petersen. Asge technology status evaluation report: wireless capsule endoscopy. *Gastrointestinal endoscopy*, 63(4):539–545, 2006.

[10] Wael El-Matary. Wireless capsule endoscopy: indications, limitations, and future challenges. *Journal of pediatric gastroenterology and nutrition*, 46(1):4–12, 2008.

[11] Gavriel Iddan, Gavriel Meron, Arkady Glukhovsky, and Paul Swain. Wireless capsule endoscopy. *Nature*, 405:417, 2000.

[12] Baopu Li and Max Q-H Meng. Wireless capsule endoscopy images enhancement via adaptive contrast diffusion. *Journal of Visual Communication and Image Representation*, 23(1):222–228, 2012.

[13] Guobing Pan and Litong Wang. Swallowable wireless capsule endoscopy: Progress and technical challenges. *Gastroenterology research and practice*, 2012, 2011.

[14] Baopu Li and Max Q-H Meng. Texture analysis for ulcer detection in capsule endoscopy images. *Image and Vision computing*, 27(9):1336–1342, 2009.

[15] Douglas G Adler and Christopher J Gostout. Wireless capsule endoscopy. *Hospital Physician*, 39(5):14–22, 2003.

[16] Zhi-Zheng Ge, Yun-Biao Hu, Shu-Dong Xiao, et al. Capsule endoscopy in diagnosis of small bowel crohn's disease. *World Journal of Gastroenterology*, 10(9):1349–1352, 2004.

[17] DW Lee, AO Poon, and AC Chan. Diagnosis of small bowel radiation enteritis by capsule endoscopy. *Hong Kong medical journal= Xianggang yi xue za zhi/Hong Kong Academy of Medicine*, 10(6):419–421, 2004.

[18] Ruijun Fu, Guanqun Bao, Yunxing Ye, and Kaveh Pahlavan. Heterogeneous cooperative localization for social networks. *International Journal of Wireless Information Networks*, 20(4):294–305, 2013.

[19] Xin Zheng, Guanqun Bao, Ruijun Fu, and Kaveh Pahlavan. The performance of simulated annealing algorithms for wi-fi localization using google indoor map. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5. IEEE, 2012.

[20] Kaveh Pahlavan. *Principles of wireless networks: A unified approach.* John Wiley & Sons, Inc., 2011.

[21] Akram Alomainy and Yang Hao. Modeling and characterization of biotelemetric radio channel from ingested implants considering organ contents. *Antennas and Propagation, IEEE Transactions on*, 57(4):999–1005, 2009.

[22] Neil Marya, Andrew Karellas, Anne Foley, Abhijit Roychowdhury, and David Cave. Computerized 3-dimensional localization of a video capsule in the abdominal cavity: validation by digital radiography. *Gastrointestinal endoscopy*, 2014.

[23] Douglas O Faigel and David R Cave. *Capsule endoscopy*. Saunders Elsevier Amsterdam, 2008.

[24] Yoshihiro Kusuda. A further step beyond wireless capsule endoscopy. *Sensor Review*, 25(4):259–260, 2005.

[25] Gastone Ciuti, Arianna Menciassi, and Paolo Dario. Capsule endoscopy: from current achievements to open challenges. *Biomedical Engineering, IEEE Reviews in*, 4:59–72, 2011.

[26] Rohit Chandra, Anders J Johansson, and Fredrik Tufvesson. Localization of an rf source inside the human body for wireless capsule endoscopy. In *Proceedings of the 8th International Conference on Body Area Networks*, pages 48–54. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.

[27] Trung Duc Than, Gürsel Alici, Hao Zhou, and Weihua Li. A review of localization systems for robotic endoscopic capsules. *Biomedical Engineering, IEEE Transactions on*, 59(9):2387–2399, 2012.

[28] Marco Salerno, Gastone Ciuti, Gioia Lucarini, Rocco Rizzo, Pietro Valdastri, Arianna Menciassi, Alberto Landi, and Paolo Dario. A discrete-time localization method for capsule endoscopy based on on-board magnetic sensing. *Measurement Science and Technology*, 23(1):015701, 2012.

[29] Chao Hu, Max Q-H Meng, and Mrinal Mandal. Efficient magnetic localization and orientation technique for capsule endoscopy. *International Journal of Information Acquisition*, 2(01):23–36, 2005.

[30] Federico Carpi, Stefano Galbiati, and Angelo Carpi. Controlled navigation of endoscopic capsules: concept and preliminary experimental investigations. *Biomedical Engineering, IEEE Transactions on*, 54(11):2028–2036, 2007.

[31] Federico Carpi, Nathan Kastelein, Michael Talcott, and Carlo Pappone. Magnetically controllable gastrointestinal steering of video capsules. *Biomedical Engineering, IEEE Transactions on*, 58(2):231–234, 2011.

[32] Xiaona Wang, MQ-H Meng, and Chao Hu. A localization method using 3-axis magnetoresistive sensors for tracking of capsule endoscope. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 2522–2525. IEEE, 2006.

[33] Gastone Ciuti, Pietro Valdastri, Arianna Menciassi, and Paolo Dario. Robotic magnetic steering and locomotion of capsule endoscope for diagnostic and surgical endoluminal procedures. *Robotica*, 28(2):199, 2010.

[34] Li Liu, Chao Hu, Wentao Cai, and MQ-H Meng. Capsule endoscope localization based on computer vision technique. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 3711–3714. IEEE, 2009.

[35] Yunxing Ye, Pranay Swar, Kaveh Pahlavan, and Kaveh Ghaboosi. Accuracy of rss-based rf localization in multi-capsule endoscopy. *International Journal of Wireless Information Networks*, 19(3):229–238, 2012.

[36] Mohammad Pourhomayoun, Zhanpeng Jin, and M Fowler. Accurate localization of in-body medical implants based on spatial sparsity. 2013.

[37] Perzila Arab, Michael Heimlich, and Eryk Dutkiewicz. Investigation of radar localization system accuracy for human gastro intestine (gi) tract. In *Medical Information and Communication Technology (ISMICT), 2013 7th International Symposium on*, pages 144–148. IEEE, 2013.

[38] NC Atuegwu and RL Galloway. Volumetric characterization of the aurora magnetic tracker system for image-guided transorbital endoscopic procedures. *Physics in medicine and biology*, 53(16):4355, 2008.

[39] Mingyuan Gao, Chengzhi Hu, Zhenzhi Chen, Honghai Zhang, and Sheng Liu. Design and fabrication of a magnetic propulsion system for self-propelled capsule endoscope. *Biomedical Engineering, IEEE Transactions on*, 57(12):2891–2902, 2010.

[40] W Weitschies, J Wedemeyer, R Stehr, and L Trahms. Magnetic markers as a noninvasive tool to monitor gastrointestinal transit. *IEEE transactions on biomedical engineering*, 41(2):192–195, 1994.

[41] Guanqun Bao, Liang Mi, and Kaveh Pahlavan. A video aided rf localization technique for the wireless capsule endoscope (wce) inside small intestine. In *Proceedings of the 8th International Conference on Body Area Networks*, pages 55–61. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.

[42] Dimitris K Iakovidis, Evaggelos Spyrou, Dimitris Diamantis, and Ilias Tsiompanidis. Capsule endoscope localization based on visual features. In *Bioinformatics*

and Bioengineering (BIBE), 2013 IEEE 13th International Conference on*, pages 1–4. IEEE, 2013.

[43] K Duda, T Zielinski, M Duplaga, M Grega, and M Leszczuk. Vq classification based on mpeg-7 visual descriptors for video endoscopic capsule localization in the gastrointestinal tract. *changes*, 1:2, 2007.

[44] Jaroslaw Bulat, Krzvsztof Duda, Mariusz Duplaga, Rafal Fraczek, Andrzej Skalski, Miroslaw Socha, Pawel Turcza, and TP Zielinski. Data processing tasks in wireless gi endoscopy: image-based capsule localization & navigation and video compression. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 2815–2818. IEEE, 2007.

[45] Doron Fischer, Reuvert Shreiber, G Meron, M Frisch, H Jacob, A Glukhovsky, and A Engel. Localization of the wireless capsule endoscope in its passage through the gi tract. In *Gastrointestinal Endoscopy*, volume 53, pages AB126–AB126. MOSBY, INC 11830 WESTLINE INDUSTRIAL DR, ST LOUIS, MO 63146-3318 USA, 2001.

[46] Doron Fischer, Reuven Schreiber, Daphna Levi, and Rami Eliakim. Capsule endoscopy: the localization system. *Gastrointestinal endoscopy clinics of north america*, 14(1):25–31, 2004.

[47] Taimoor Shah, Syed Mahfuzul Aziz, and Tharshan Vaithianathan. Development of a tracking algorithm for an in-vivo rf capsule prototype. In *Electrical and Computer Engineering, 2006. ICECE'06. International Conference on*, pages 173–176. IEEE, 2006.

[48] Lujia Wang, Chao Hu, Longqiang Tian, Mao Li, and MQ-H Meng. A novel radio propagation radiation model for location of the capsule in gi tract. In *Robotics and*

Biomimetics (ROBIO), 2009 IEEE International Conference on, pages 2332–2337. IEEE, 2009.

[49] Le Zhang, Yongxin Zhu, Tingting Mo, Jinlong Hou, and Guoguang Rong. Design and implementation of 3d positioning algorithms based on rf signal radiation patterns for in vivo micro-robot. In *Body Sensor Networks (BSN), 2010 International Conference on*, pages 255–260. IEEE, 2010.

[50] Kegen Yu, Ian Sharp, and Y Jay Guo. *Ground-based wireless positioning*, volume 5. John Wiley & Sons, 2009.

[51] Kaveh Pahlavan, Xinrong Li, and J-P Makela. Indoor geolocation science and technology. *Communications Magazine, IEEE*, 40(2):112–118, 2002.

[52] Alan Bensky. *Wireless positioning technologies and applications.* Artech House, 2007.

[53] K Pahlavan, F Akgul, Y Ye, T Morgan, F Alizadeh-Shabdiz, M Heidari, and C Steger. Taking positioning indoors wi-fi localization and gnss. *Inside GNSS*, 5 (3):40–47, 2010.

[54] Yiu-Tong Chan, Wing-Yue Tsui, Hing-Cheung So, and Pak-chung Ching. Time-of-arrival based localization under nlos conditions. *Vehicular Technology, IEEE Transactions on*, 55(1):17–24, 2006.

[55] Nirupama Bulusu, John Heidemann, and Deborah Estrin. Gps-less low-cost outdoor localization for very small devices. *Personal Communications, IEEE*, 7(5): 28–34, 2000.

[56] Zhe Xiang, Song Song, Jin Chen, Hao Wang, Jian Huang, and Xingxin Gao. A wireless lan-based indoor positioning technology. *IBM Journal of research and development*, 48(5.6):617–626, 2004.

[57] Hui Liu, Houshang Darabi, Pat Banerjee, and Jing Liu. Survey of wireless indoor positioning techniques and systems. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 37(6):1067–1080, 2007.

[58] Jie He, Yishuang Geng, Yadong Wan, Shen Li, and Kaveh Pahlavan. A cyber physical test-bed for virtualization of rf access environment for body sensor network. 2013.

[59] Jie He, Kaveh Pahlavan, Shen Li, and Qin Wang. A testbed for evaluation of the effects of multipath on performance of toa-based indoor geolocation. 2013.

[60] Yishuang Geng, Jie He, and Kaveh Pahlavan. Modeling the effect of human body on toa based indoor human tracking. *International Journal of Wireless Information Networks*, 20(4):306–317, 2013.

[61] Kaveh Pahlavan, Yunxing Ye, Ruijun Fu, and Umair Khan. Challenges in channel measurement and modeling for rf localization inside the human body. *International Journal of Embedded and Real-Time Communication Systems (IJERTCS)*, 3(3): 18–37, 2012.

[62] Chao Hu, MQ-H Meng, and Mrinal Mandal. Efficient linear algorithm for magnetic localization and orientation in capsule endoscopy. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 7143–7146. IEEE, 2006.

[63] Guanqun Bao, Liang Mi, and Kaveh Pahlavan. Emulation on motion tracking of endoscopic capsule inside small intestine. In *14th International Conference on Bioinformatics and Computational Biology, Las Vegas*, 2013.

[64] Laurel R Fisher and William L Hasler. New vision in video capsule endoscopy: current status and future directions. *Nature Reviews Gastroenterology and Hepatology*, 9(7):392–405, 2012.

[65] Khalil Arshak and Francis Adepoju. Adaptive linearized methods for tracking a moving telemetry capsule. In *Industrial Electronics, 2007. ISIE 2007. IEEE International Symposium on*, pages 2703–2708. IEEE, 2007.

[66] Kamran Sayrafian-Pour, Wen-Bin Yang, John Hagedorn, Judith Terrill, and Kamya Yekeh Yazdandoost. A statistical path loss model for medical implant communication channels. In *Personal, Indoor and Mobile Radio Communications, 2009 IEEE 20th International Symposium on*, pages 2995–2999. IEEE, 2009.

[67] Jie He, Yishuang Geng, and Kaveh Pahlavan. Modeling indoor toa ranging error for body mounted sensors. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, pages 682–686. IEEE, 2012.

[68] Yishuang Geng, Yadong Wan, Jie He, and Kaveh Pahlavan. An empirical channel model for the effect of human body on ray tracing. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pages 47–52. IEEE, 2013.

[69] Yishuang Geng. *Modeling of Time-of-arrival for CM4 Body Area Networks Channel*. PhD thesis, Worcester Polytechnic Institute, 2013.

[70] Richard Fuller. Tutorial on location determination by rf means. In *Mobile Entity Localization and Tracking in GPS-less Environnments*, pages 213–234. Springer, 2009.

[71] Ahmet Çamlıca, Barış Fidan, and Mustafa Yavuz. Implant localization in human body using adaptive least-squares based algorithm. In *ASME 2013 International Mechanical Engineering Congress and Exposition*, pages V03BT03A015–V03BT03A015. American Society of Mechanical Engineers, 2013.

[72] Guowei Shen, Rudolf Zetik, and Reiner S Thoma. Performance comparison of toa and tdoa based location estimation algorithms in los environment. In *Positioning, Navigation and Communication, 2008. WPNC 2008. 5th Workshop on*, pages 71–78. IEEE, 2008.

[73] KW Cheung, Hing-Cheung So, W-K Ma, and Yiu-Tong Chan. Least squares algorithms for time-of-arrival-based mobile location. *Signal Processing, IEEE Transactions on*, 52(4):1121–1130, 2004.

[74] Ilan Ziskind and Mati Wax. Maximum likelihood localization of multiple sources by alternating projection. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(10):1553–1560, 1988.

[75] Joe C Chen, Ralph E Hudson, and Kung Yao. Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. *Signal Processing, IEEE Transactions on*, 50(8):1843–1854, 2002.

[76] Xinghong Kuang and Huihe Shao. Maximum likelihood localization algorithm using wireless sensor networks. In *Innovative Computing, Information and Control, 2006. ICICIC'06. First International Conference on*, volume 3, pages 263–266. IEEE, 2006.

[77] Harald Cramér. *Mathematical methods of statistics*, volume 9. Princeton university press, 1999.

[78] Calyampudi Radhakrishna Rao and Helge Toutenburg. *Linear models*. Springer, 1995.

[79] Neal Patwari, Alfred O Hero, Matt Perkins, Neiyer S Correal, and Robert J O'dea. Relative location estimation in wireless sensor networks. *Signal Processing, IEEE Transactions on*, 51(8):2137–2148, 2003.

[80] Neal Patwari, Joshua N Ash, Spyros Kyperountas, Alfred O Hero, Randolph L Moses, and Neiyer S Correal. Locating the nodes: cooperative localization in wireless sensor networks. *Signal Processing Magazine, IEEE*, 22(4):54–69, 2005.

[81] Evaggelos Spyrou and Dimitris K Iakovidis. Video-based measurements for wireless capsule endoscope tracking. *Measurement Science and Technology*, 25(1):015002, 2014.

[82] David R Cave, David E Fleischer, Christopher J Gostout, Douglas O Faigel, Jonathan A Leighton, Russell I Heigh, Virender K Sharma, Klaus Mergener, Knishka Bhattacharya, Elizabeth Rajan, et al. A multi-center randomized comparison of the endocapsule: Olympus inc and the pillcam sb: Given imaging in patients with obscure gi bleeding. *Gastrointestinal Endoscopy*, 65(5):AB125, 2007.

[83] Andrew Fort, Claude Desset, Philippe De Doncker, Piet Wambacq, and Leo Van Biesen. An ultra-wideband body area propagation channel model-from statistics to implementation. *Microwave Theory and Techniques, IEEE Transactions on*, 54(4):1820–1826, 2006.

[84] Julien Ryckaert, Philippe De Doncker, René Meys, Arnaud de Le Hoye, and Sophie Donnay. Channel model for wireless communication around human body. *Electronics letters*, 40(9):543–544, 2004.

[85] Guanqun Bao and Kaveh Pahlavan. Motion estimation of the endoscopy capsule using region-based kernel svm classifier. *accepted by 2013 IEEE International conference on Electro Information Technology*, 2013.

[86] Guanqun Bao, Yunxing Ye, Umair Khan, Xin Zheng, and Kaveh Pahlavan. Modeling of the movement of the endoscopy capsule inside gi tract based on the captured endoscopic images. In *Modeling, Simulation and Visualization Methods, The 2012 International Conference on*, 2012.

[87] Santosh Pandey and Prathima Agrawal. A survey on localization techniques for wireless networks. *Journal of the Chinese Institute of Engineers*, 29(7):1125–1148, 2006.

[88] Yunxing Ye, Umair Khan, Nayef Alsindi, Ruijun Fu, and Kaveh Pahlavan. On the accuracy of rf positioning in multi-capsule endoscopy. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2011 IEEE 22nd International Symposium on*, pages 2173–2177. IEEE, 2011.

[89] Leopoldo Jetto, Sauro Longhi, and Giuseppe Venturini. Development and experimental validation of an adaptive extended kalman filter for the localization of mobile robots. *Robotics and Automation, IEEE Transactions on*, 15(2):219–229, 1999.

[90] Andrew Howard. Multi-robot simultaneous localization and mapping using particle filters. *The International Journal of Robotics Research*, 25(12):1243–1256, 2006.

[91] ITO Takahiro, Daisuke Anzai, and WANG Jianqing. Performance evaluation on rssi-based wireless capsule endoscope location tracking with particle filter. *IEICE Transactions on Communications*, 97(3):579–586, 2014.

[92] Geraldo Silveira, Ezio Malis, and Patrick Rives. An efficient direct approach to visual slam. *Robotics, IEEE Transactions on*, 24(5):969–979, 2008.

[93] José A Castellanos, José Neira, and Juan D Tardós. Multisensor fusion for simultaneous localization and map building. *Robotics and Automation, IEEE Transactions on*, 17(6):908–914, 2001.

[94] Andrew J Davison and David W Murray. Simultaneous localization and map-building using active vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):865–880, 2002.

[95] Kaveh Pahlavan, Guanqun Bao, and Mi Liang. Body-slam: Simultaneous localization and mapping inside the human body. *Keynote speech, 8th International Conference on Body Area Networks (BodyNets), Boston, MA, September 30-October 2*, 2013.

[96] Chee Khun Poh, That Mon Htwe, Liyuan Li, Weijia Shen, Jiang Liu, Joo Hwee Lim, Kap Luk Chan, and Ping Chun Tan. Multi-level local feature classification for bleeding detection in wireless capsule endoscopy images. In *Cybernetics and Intelligent Systems (CIS), 2010 IEEE Conference on*, pages 76–81. IEEE, 2010.

[97] Richard Nock and Frank Nielsen. Statistical region merging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(11):1452–1458, 2004.

[98] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.

[99] Andrea Vedaldi. An open implementation of the sift detector and descriptor. *UCLA CSD*, 2007.

[100] Jean-Michel Morel and Guoshen Yu. Asift: A new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2(2):438–469, 2009.

[101] Guoshen Yu and J-M Morel. A fully affine invariant image comparison method. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 1597–1600. IEEE, 2009.

[102] Tammam Tillo, EngGee Lim, Zhao Wang, Jianwei Hang, and Rongqiang Qian. Inverse projection of the wireless capsule endoscopy images. In *2010 International Conference on Biomedical Engineering and Computer Science (ICBECS)*. pp.1-4, 2010.

[103] Suchitra Sathyanarayana, Srikanthan Thambipillai, and Chris T Clarke. Real time tracking of camera motion through cylindrical passages. In *IEEE 15th International Conference on Digital Signal Processing*. pp.455-458, 2007.

[104] F De Iorio, C Malagelada, F Azpiroz, M Maluenda, C Violanti, L Igual, J Vitrià, and J-r Malagelada. Intestinal motor activity, endoluminal motion and transit. *Neurogastroenterology & Motility*, 21(12):1264–e119, 2009.

[105] Nathalie Schoofs, Jacques Devière, and André Van Gossum. Pillcam colon capsule endoscopy compared with colonoscopy for colorectal tumor diagnosis: a prospective pilot study. *Endoscopy*, 38(10):971–977, 2006.

[106] Cristiano Spada, Maria Elena Riccioni, Cesare Hassan, Lucio Petruzziello, Paola Cesaro, and Guido Costamagna. Pillcam colon capsule endoscopy: a prospective, randomized trial comparing two regimens of preparation. *Journal of clinical gastroenterology*, 45(2):119–124, 2011.

[107] Chen-yi Li, Bing-ling Zhang, Chun-xiao Chen, and You-ming Li. Omom capsule endoscopy in diagnosis of small bowel disease. *Journal of Zhejiang University SCIENCE B*, 9(11):857–862, 2008.

[108] Laure France, Julien Lenoir, Alexis Angelidis, Philippe Meseure, M-P Cani, François Faure, and Christophe Chaillou. A layered model of a virtual human intestine for surgery simulation. *Medical image analysis*, 9(2):123–132, 2005.

[109] Liang Mi, Guanqun Bao, and Kaveh Pahlavan. Design and validation of a virtual environment for experimentation inside the small intestine. In *Proceedings of the 8th International Conference on Body Area Networks*, pages 35–40. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013.

[110] Liang Mi, Guanqun Bao, and Kaveh Pahlavan. Geometric estimation of intestinal contraction for motion tracking of video capsule endoscope.

[111] Kálmán Palágyi and Attila Kuba. A 3d 6-subiteration thinning algorithm for extracting medial lines. *Pattern Recognition Letters*, 19(7):613–627, 1998.

[112] Sharmishtaa Seshamani, William Lau, and Gregory Hager. Real-time endoscopic mosaicking. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006*, pages 355–363. Springer, 2006.

[113] Piotr M Szczypiński, Ram D Sriram, Parupudi VJ Sriram, D Nageshwar Reddy, et al. A model of deformable rings for interpretation of wireless capsule endoscopic videos. *Medical Image Analysis*, 13(2):312, 2009.

[114] Zhuoran Liu, Jin Chen, Umair Khan, Bader Alkandari, and Kaveh Pahlavan. Wideband characterization of rf propagation for toa localization of wireless video capsule endoscope inside small intestine. In *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pages 326–331. IEEE, 2013.

[115] Amara Okoli, Nischala Ammannagari, Mohammed Mazumder, and Kiran Nakkala. When the dissolvable does not dissolve: An agile patency capsule mystery. *The American journal of gastroenterology*, 109(4):605–607, 2014.

[116] Pranay Swar, Kaveh Pahlavan, and Umair Khan. Accuracy of localization system inside human body using a fast fdtd simulation technique. In *Medical Information and Communication Technology (ISMICT), 2012 6th International Symposium on*, pages 1–6. IEEE, 2012.

[117] Jin Chen, Yunxing Ye, and Kaveh Pahlavan. Comparison of uwb and nb rf ranging measurements in homogenous tissue for ban applications. In *Wireless Telecommunications Symposium (WTS), 2013*, pages 1–5. IEEE, 2013.

[118] Meng Zhang, Eng Gee Lim, Zhao Wang, Tammam Tillo, Ka Lok Man, and Jing Chen Wang. Rf characteristics of wireless capsule endoscopy in human body. In *Grid and Pervasive Computing*, pages 700–706. Springer, 2013.

[119] X SEMCAD. Reference manual for the semcad simulation platform for electromagnetic compatibility, antenna design and dosimetry. *SPEAG-Schmid & Partner Engineering AG, http://www. semcad. com*, 2006.

[120] Dennis M Sullivan. *Electromagnetic simulation using the FDTD method.* John Wiley & Sons, 2013.

[121] Kane S Yee, Jei Shuan Chen, and Albert H Chang. Conformal finite-different time-domain (fdtd) with overlapping grids. *Antennas and Propagation, IEEE Transactions on*, 40(9):1068–1075, 1992.

[122] Kenichi Takizawa, Takahiro Aoyagi, J-i Takada, Norihiko Katayama, Kamya Yekeh, KY Yazdandoost, and T Kobayashi. Channel models for wireless body area networks. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 1549–1552. IEEE, 2008.

[123] Eng Gee Lim, Zhao Wang, Jin Hui Chen, Tammam Tillo, and Ka Lok Man. Investigation of em wave propagation of the wireless capsule in human body. In *East-West Design & Test Symposium, 2013*, pages 1–4. IEEE, 2013.