Design and Validation of a Virtual Environment for Experimentation inside the Small Intestine

Liang Mi, Guanqun Bao and Kaveh Pahlavan
Center for Wireless Information Network Studies
Worcester Polytechnic Institute
Worcester, MA, 01609, USA
(Imi, gbao, Kaveh)@wpi.edu

ABSTRACT
Designing a precise and reliable localization system for wireless capsule endoscopy (WCE) has always been a challenging problem due to the complicated in-body environment and uncontrollable movement of body tissue. Knowing the motion information of the capsule would greatly enhance the localization accuracy. However, design and validate any motion tracking algorithm inside small intestine faces a lot of difficulties since any experimentation on the human being is extremely costly and restricted by law. Having a virtual environment that looks and functions exactly like small intestine would facilitate the process of verifying the performance of existing algorithms without going into the real human body. In this paper, we established a virtual testbed that emulates the contraction of intestinal lumen and the transition of endoscopic capsule. Under this emulation environment, a velocity estimation algorithm based on a feature detection algorithm (ASIFT) and a velocity estimation algorithm (MDR) was implemented and its performances were evaluated. Experimental results showed that our proposed emulation environment is able to provide reliable platform for motion detection validation.

Categories and Subject Descriptors
I.4 [Programming Language]: Matlab

General Terms
Design

Keywords
Wireless capsule endoscopy (WCE), endoscopic testbed, motion detection, velocity estimation

1. INTRODUCTION
Body Area Networks (BAN) is expected to be the next breakthrough for medical applications due to its great potential in minimizing implanted wireless communication devices [1-2]. Among those small devices, wireless endoscopic capsule (as shown in Figure 1) is one of the most innovative inventions since it can provide a noninvasive way to examine the human's digestive system [3]. However, this technology is unable to localize itself when an abnormality is found by the wireless camera. Having a precise and reliable localization system for the wireless capsule would greatly enhance the benefits of WCE by allowing doctors administrating immediate therapeutic operations [4].

During the past few years, many attempts have been made to developing accurate localization systems for the WCE. Kaveh Pahlavan and his colleagues used Time of Arrival (TOA) and Received Signal Strength (RSS) based techniques for RF localization inside human tissue [5]. Chao Hu and his research group developed a linear algorithm for tracing magnet position by using magnetic sensors [6]. However, to utilize image source to track the position of WCE is still understudied and has a great potential to enhance the accuracy of existing localization infrastructures by data fusion. In our previous work [7], we developed a novel image processing based motion tracking algorithm for the WCE and explore the potential of combining the movement of the wireless capsule with existing RF localization infrastructure to enhance the in-body localization accuracy [8]. Another research group [9] reported methods for fast interpretation, motion tracking and velocity estimation of the capsule based on endoscopic images. Many other methods using endoscopic images have been reported.

However, validation of these motion tracking algorithms are very challenging since any experimentation inside human being is extremely difficult due to the unpredictable movement of body tissue and individual difference. Once the wireless capsule is ingested by a patient and passes through the gastrointestinal (GI) tract, there is no mechanism to control the capsule’s speed or direction as it traverses the GI tract [10-11]. It is also particularly difficult to measure the capsule’s location or orientation during its traversal within the human body [12-13]. Plus carrying our experiments on real human beings is extremely costly and restricted by law [14-15]. Thus, an alternative way to do algorithm validation is to build up an emulation testbed.

Figure 1. Wireless endoscopic capsule. It is a pill-shaped capsule with built-in video camera, light-emitting diodes, video signal transmitter and battery [9]. There are also other sizes of capsule used in clinic [16].

This research was funded by the National Institute of Standards and Technology (NIST), USA, under contract FON 2009-NIST-ARRA-MSE-Research-01, entitled “RF Propagation Measurement and Modeling for Body Area Networking”.

Figure 2. The dimensions of the capsule endoscope. It is 11mm in diameter and 25mm in length.
To simulate the visual effect inside a human small intestine, most researchers assumed the small intestine was static and with no contraction. Hence an easy way to develop a testbed of small intestine was to create a virtual straight tube with a certain radius and features mapped inside the tube. Then a virtual camera was set to move along the tube with a certain movement speed and perhaps with a certain rotation velocity. But the reality is that the small intestine is bent with complicated shape. The capsule would experience many types of motion inside small intestine such as translation, tilt, and rotation in a three-dimensional space.

The essential of an endoscopy testbed is to provide realistic endoscopic images for algorithm validation. To this end, we designed a test system that looked extremely similar with human intestinal endoscopic environment and then we developed a velocity estimation algorithm to analyze the performance of our test system. First we built a physical model of a piece of small intestine with PVC tube. Then the physical model was measured and mapped into computer. Matlab computer vision toolbox was used to generate and visualize the virtual model. Second, we put a wired camera and a virtual camera inside the physical and virtual model respectively, and got the consecutive endoscopic images. After this we added more details on the virtual model such as intestinal contraction. Then we implemented a velocity estimation algorithm based on Affine Scale-Invariant Feature Transform (ASIFT) algorithm [17] and Model of Deformable Rings (MDR) algorithm [9] to analyze the performance of our test system. Experimental data was then achieved to assist our further research and other research on endoscopic image processing technique.

The major contribution of this paper is that we built a systematic intestinal endoscopy test environment for algorithm validation. We built both physical and virtual testbeds and used a velocity estimation algorithm to evaluate the performance of the testbeds. Another contribution is that we analyzed the effect ofintestinal contraction on a velocity estimation algorithm.

The rest of the paper is organized as follows: In section 2, we describe in detail the setup process of testbeds. In section 3, we introduce the methodology of developing a velocity estimation algorithm. In section 4, we give the results of analyzing the performance of testbeds using our proposed method. Finally in section 5, we talk about conclusion and future works.

2. ALTERNATIVES FOR EMULATING SMALL INTESTINE

It is very difficult to verify the performance of an endoscopic image processing algorithm, because doing experiments on human body is costly and restricted, and because of limitation of control of the capsule. Thus a testbed becomes practically useful for algorithm validation. In this section we talk about setup of intestinal environment in detail.

2.1 Physical Testbed Setup

One way to simulate the environment of a small intestine is to build a physical model. According to clinical pictures, we built a physical testbed whose appearance is shown in Figure 2 (a). It was created by bending and twisting a 1.5 meter long 3 centimeter diameter Polyvinyl Chloride (PVC) tube. The outer surface of the tube was painted 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 to prevent the light of camera from escaping outside.

Clinical data showed that the average length of a human small intestine was 7-9 meters long and the capsule stayed in the small intestine for about 3-4 hours. During the few hours when the capsule was in the small intestine, it took pictures at 2 frames / second. If we assume the capsule travels at a constant speed, then the average step distance between two consecutive frames could be roughly calculated as 0.03 cm. To simulate the transition of the endoscopic capsule, we inserted a wired endoscopy camera equipped with four LED lights 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 3 (a) indicates a test pictures take from inside the physical testbed. We can see that it is extremely similar with real pictures taken from small intestine.

One advantage of the physical testbed was that it is intuitional. It provided realistic endoscopic images for postprocessing. Another big advantage was that we can put the whole physical model into liquid to simulate the whole intestinal environment, not only inside, but also outside and around the body area. In this way, we could also add an antenna on the camera, measure the measure propagation characteristic of wireless BAN (WBAN) channel [18], and analyze its influence on endoscopic images.

However, there were still several drawbacks of a physical testbed:

- A big drawback was its restriction in camera control. After the camera was inserted into the tube, the only possible movement of the camera was along the tube with linear proceeding distance, rather than tilt and rotation.
- Besides, to emulate the complicated shape (especially the sharp turn) of the small intestine was very difficult because there was barely a kind of materials could be as smooth and soft as small intestine and had similar toughness at the same time. And the camera would get stuck at the corner.
- Plus adding features on the physical model to make it exactly the same as small intestine was also unpractical. Hence an alternative way to emulate the inside environment of the small intestine was to build a virtual testbed which will be introduced in the next subsection.

Figure 2 Emulation testbeds. (a) shows the inside and out appearance of physical testbed, as well as coordinates definition; (b) shows the corresponding virtual testbed.
In this 3D space, a virtual camera was placed inside the virtual model. The camera’s position was set at the axis of the cylinder and the camera’s target was set at the axis of the cylinder with a certain distance to the camera’s position, which was to emulate the movement of the camera inside the physical testbed as shown in Figure 3 (b). Also a Phong point light source was set at the camera’s position to simulate the illumination effect.

The advantages of using a virtual testbed are as follows:

- It has better camera control, more realistic interior texture, and quicker processing time, compared with a physical testbed.
- It can be distorted into any kinds of shape like a real small intestine.

Anatomy results show that a human small intestine is compressed in the lower abdomen. The intestinal tissue allows itself to bend into different kinds of shapes and at the same time to maintain its toughness and elasticity. Moreover, intestinal motility is also an important factor that should be considered when simulating the virtual environment of small intestine. Clinical data shows that the average diameter of small intestine is 2-3 cm. When it is contracted, the diameter can be as small as 7 mm which is smaller than the diameter of the wireless endoscopic capsule [19]. The average frequency of intestinal contraction is 9-10 min⁻¹ consistent with localization, which means during several hour transition of a wireless capsule, it may come across a few dozen times of contraction. Thus it is necessary to study the possible influence of intestinal contraction on the endoscopic images captured by a wireless capsule.

The field of view of a wireless capsule can be as large as 156°, so objects around the capsule, not only those in front of the capsule, take up a large space in the endoscopic images. Because of limited light strength, there would form a black hole approximately in the center of the images. Objects which are relatively far from the capsule have little influence on the images, including a possible contraction. Therefore, to simplify the problem, we assume the walls of small intestine are stick to the front cover of the capsule, which is shown in Figure 3 (b). Since the front cover of the wireless endoscopic capsule is semispherical, we used the shape of the cover to generate a same virtual model to represent a piece small intestine with contraction.

### 3. IMPLEMENTATION OF MOTION DETECTION ALGORITHM ON VIRTUAL ENVIRONMENT

Velocity estimation based on endoscopic images is a hot research topic especially since the early 21st century, when wireless endoscopic capsules were invented and practically used in clinic. During the past few years, there have been some attempts to estimate the velocity of the endoscopic capsule. Usually velocity estimation based on endoscopic images consists of two steps. First some feature points are detected in two consecutive images. Then the displacement of the feature points (which is relatively the displacement of the capsule) is calculated based on geometry analysis. The displacement will then be divided by the elapsed time for deducing the velocity. We talk about these two steps one by one.

#### 3.1 Feature Point Detection

The purpose of feature point detection is to track the transformations such as translation, rotation, and scaling between images, which reflects the motion of the camera. It is important that the feature points extracted from the reference image can be accurately detected in the second image. According to the literature, in the WCE application, more feature points can be detected by the affine SIFT algorithm then other algorithms [18]. Therefore we choose affine SIFT to do feature point detection.

The distance between a feature point and the capsule is very important for velocity estimation. When taking account into intestinal contraction, it is more important for researchers to understand how the contraction influences the movement of feature points in an endoscopic image. Since the capsule is surrounded by the walls of small intestine, these walls are projected to the 2-D image plane as a bunch of circular rings (Figure 4). Under this image acquisition system, the points that are closer to the capsule lie on circles with a larger radius compared to the points that are farther down the intestine. Therefore to better study the influence of intestinal contraction, we use basic idea of MDR algorithm to generate a bunch of rings with a certain amount of nodes in the endoscopic images, as shown in Figure 5 (a).

![Figure 5](image)

**Figure 5.** Feature point projection. L refers to the lens of the capsule. V indicates the closest view that can be captured by the camera, forming the angle of view α. P1 and P2 refers to two feature points, forming two angular depth θ1 and θ2 respectively. r1 is the radius of the tube, rL and r2 are the distances between P1 and P2 to the center of the image.
Every node is assigned a pair of index to, \( p = 1, 2, 3, ..., P \) and \( q = 1, 2, 3, ..., Q \), where \( p \) indicates the ring number and \( q \) indicates the amount of nodes in each ring. We use the following equation to transform the indexes into Cartesian coordinates:

\[
[x_{p,q}, y_{p,q}] = rw^{p-1}\left[\cos\left(\frac{2\pi q}{Q}\right), \sin\left(\frac{2\pi q}{Q}\right)\right]
\]

where \( x_{p,q}, y_{p,q} \) are a pair of Cartesian coordinates of a node, \( r \) is the radius of the smallest ring, \( r \) is the ratio of the radius of two adjacent rings.

Now we regard these nodes as feature points and use affine SIFT algorithm to detect the positions of the feature points in the next image. Nodes are connected as the same topology as before, as shown in Figure 5 (b). As seen in the picture, as the virtual camera move forward along the tube, all the nodes moved outward resulting in a larger radius of each ring.

### 3.2 Displacement Estimation

The velocity of the wireless capsule can be estimated by measuring the displacement of the feature points. As shown in Figure 4 and Figure 5, feature points (nodes) that are in the same ring have the similar displacements outward if the capsule is moving straight forward. Thus we measure the displacements of the feature points in different groups each of which consists of feature points in one ring. After averaged, each ring gives an estimate showing the reflecting the displacement of the capsule. We use Figure 6 to illustrate the procedure of estimating displacement of the capsule by measuring the estimate of each ring.

In Figure 6 (a), \( L \) refers to the initial position of the capsule and \( L' \) refers to the position after the capsule moves forward for a distance of \( D \). During this process, the projection of a feature point \( P \) moves from \( P_1 \) to \( P_2 \), forming two angles \( \theta_1 \) and \( \theta_2 \). According to our previous analysis, the displacement \( D \) can be calculated by the equation (2).

\[
D = \frac{r_s}{\tan \theta_2} \left( 1 - \frac{\tan \theta_1}{\tan \theta_2} \right)
\]

Figure 6 (b) reflects the situation where the intestine is contracted and stick to the front cover of the capsule. In this scenario, as the capsule moves, the distance between the feature point and the lens of the capsule stays the same, yet still forming two angles \( \theta_1 \) and \( \theta_2 \). Plug in the expressions of \( \theta_1 \) and \( \theta_2 \), the displacement \( D \) can be calculated by equation (3).

\[
D = r_s (\theta_2 - \theta_1)
\]

\( \theta_1 \) and \( \theta_2 \) can be calculated by equation (4), according to Figure6.

\[
\theta = \tan^{-1}\left(\frac{r_s}{r} \tan \alpha\right)
\]

Finally the velocity of the capsule can be deduced as follows:

\[
v = \frac{D}{dt}
\]

One of the biggest influences of intestinal contraction on velocity estimation algorithm is that objects in different regions of endoscopic images moves in different modes, compared to that without contraction. In next section, we give the test estimation results of both situations.

### 4. USING VIRTUAL ENVIRONMENT FOR PERFORMANCE EVALUATION

Two groups of experiments were respectively conducted in two kinds of situations. In the first situation, the virtual tube was straight with a constant radius; in the second situation, the virtual tube is straight but the radius was variable.

#### 4.1 Velocity Estimation in a Straight Tube without Contraction

To verify the performance of our proposed method, we created a straight virtual tube and set a virtual camera to move along the axis of it, using the same method introduced in section 2. Velocity estimation algorithm was implemented with Matlab and used under the testbed. The length of the tube was 100 cm, with a constant radius of 2 cm. The far end of the tube was covered with a black circle. The velocity of the camera’s displacement was set to be 0.7 mm/s. During the transition of the camera, it took pictures at resolution of 512 x 512 pixels. Field of view was 62°, which was the same as that of the wired camera we used in the physical testbed. Thus \( \alpha = 31° \). To generate the nodes, we also set \( P = 7 \), \( Q = 128 \), \( r = 130 \), \( w = 1.1 \).

Figure 7 shows the results of feature point detection. As expected, all the nodes moved outward resulting in a larger radius of each ring. Also the structure of the rings was slightly distorted because of some error in detection.

According to our previous analysis, when the tube is straight, the estimated displacement of the camera calculated based on different rings, should be the same. And the estimation result verified this as shown in Figure 9 (a) (red line). The result was normalized and average displacement of 7 rings was 1.02. Average error is 2%.
We set a constant $P = 7$ and changed the amount of nodes in each ring, $Q$, and obtained the results shown in Figure 9 (b) (red line). It reflects that when $Q$ is bigger than 4, the estimated result doesn’t change a lot as $Q$ increases.

### 4.2 Velocity Estimation in a Straight Tube with Contraction

In this subsection, we verify of performance of our proposed method in a contracted virtual intestine. We built a contracted tube based on the method in section 2, as shown in Figure 3 (b). From the capsule to the far end, the radius of the tube varied from 6 mm (which is approximately the radius of the capsule) to 2 mm. We set $r = 50$, $w = 1.26$. Other parameters were the same as those in the previous subsection.

Feature point detection results are shown in Figure 8. As seen in the images, because of contraction, the walls of tube are closer to the lens of camera, resulting in a smaller black hole in the center, compared to the tube without contraction. Nodes of the rings moved a little longer if we compare Figure 8 (b) with 9 (b), which was corresponding to our previous analysis.

Use our proposed method for contraction model to estimate the displacement of the virtual camera. The results are as shown in Figure 9 (blue line). In Figure 9 (a), because the walls that were near to the center of endoscopic images were not stick to the front cover of the capsule, the calculated estimates relatively differed a lot from the expected result. Ring no. 5-7, which are corresponding to the feature points that far away from the center of the images had the relatively accurate estimates. The field of view of a real wireless endoscopic capsule can be as large as 156°, so our method can still provide accurate estimates of displacement of the capsule.

Furthermore, in most existing velocity estimation algorithm or motion detection algorithm the small intestine are all assumed to have a constant diameter. Hence we did another group of experiments to study that if the small intestine is contracted but we don’t know about it, how the contraction would influence the estimation or detection results. To this end we used the previous model in this subsection and the method in subsection 4.1.

Parameters were set the same as before. Figure 13 shows the results. In this scenario, set the camera is the origin, the included angle of Ring No.7 (which is the outer ring) and the axis is about 30° as shown in Figure 27. At this point, according to geometry, the actual displacement $D$ is approximately a quarter of estimated displacement $D'$, which is extremely similar with our result. More experiments showed that if the wireless endoscopic capsule has a field of view of 62°, the error that is about 4-5 times of the estimates should be considered. If the field of view is much larger than 62°, the error could be less than 2 times of the estimates.

Because of the big drawbacks of the endoscopic image processing technique, there have been a few attempts to combine the motion detection results from endoscopic images with Radio Frequency (RF) localization infrastructure [6] to enhance the estimated results.

![Figure 7](image7.png)

**Figure 7.** Feature point detection with a constant diametrical tube. Each node in the rings was regarded as a feature point, some of which didn’t have enough “feature” information for detection though. For each of these nodes, we set a threshold. If the detection result was far away from the original position, a real feature point near the node would substitute it for feature point detection.

![Figure 8](image8.png)

**Figure 8.** Feature point detection in the scenario with contraction. The detection strategy was the same as discussed in subsection 4.1, shown in Figure 7.

![Figure 9](image9.png)

**Figure 9.** Displacement estimates in a constant diametrical tube (red line) and a variable diametrical tube (blue line). In (a), the estimates in uncontracted situation was very accurate; the estimates in contracted situation was far from the expected value when $p$ was small, because in the central region of the endoscopic image the walls of intestine was not stick to the cover and the nodes moved slower than outer ones. (b) indicates that the estimates became accurate when the amount of nodes in each ring increased and after it reached 8, the estimates were already very accurate and didn’t change much. (c) As the ring number increased, the estimate became closer to the expected one. Results were normalized. The expected value of velocity were 1.
5. CONCLUSION AND FUTURE WORKS

In this paper, we presented a systematic method to design and validate the virtual environment for experimentation inside small intestine. Using this environment, we analyzed the influence of small intestinal contraction on motion detection algorithms. Experimental results precisely coincided with expectation.

In the future, we will focus on refining our test system according to the clinical data, and explore its potential application on other research directions such as localization inside human body. Also we will study the possibility to create a more realistic and reliable physical testbed and explore the potential to directly use the data from the physical testbed for velocity estimation and other applications.

6. ACKNOWLEDGMENT

The authors would like to thank Dr. David Cave at UMass Memorial Medical Center for his precious suggestions and the colleagues at the CWINS laboratory for their directly or indirectly help in preparation of the results presented in this paper.

7. REFERENCES


