Multi Model Tracking for Localization in Wireless Capsule Endoscope

Babak Moussakhani
Department of Electronics and Telecommunications
NTNU, N-7491
Trondheim, Norway
babak@iet.ntnu.no

Raul Chavez-Santiago
Interventional Center, Rikshospitalet University Hospital
N-0027
Oslo, Norway
Raul.Chavez-Santiago@rr-research.no

Ilango Balasingham
Interventional Center, Rikshospitalet University Hospital
N-0027
Oslo, Norway
ilangkob@medisin.uio.no

ABSTRACT
In this work, the tracking of a Wireless Capsule Endoscope (WCE) is considered. The special shape of the movement track in the small intestine has been taken advantage of to model it as a multi mode random process. Then two Multi-Model (MM) target tracking methods have been adapted to follow a source which is moving based on the proposed random process. Thus the goal is to study the adjustability of the MM methods to the intestinal track. Through simulation it is shown that such methods are able to reach far better accuracy than a simple triangulation which is currently used in the prevalent WCEs.

Keywords
Wireless Capsule Endoscope, Kalman Filtering, Localization, Multi-Model

1. INTRODUCTION

Until the year 2000, endoscopic and radiological investigations of the small intestine had a limited diagnostic yield. Therefore, in patients with suspected intestinal disease, the diagnosis was often delayed. That situation frequently would lead to the late discovery of diseases at a stage of worsened prognosis, especially in the case of a tumor suspicion. Wireless capsule endoscopy (WCE) [7], introduced by Given Imaging Ltd. is a procedure which has enabled for the first time a painless diagnosis inside the gastrointestinal (GI) tract, specifically in the small intestine. The ingestable capsule moves along the GI tract in the same manner as food, with the normal peristaltic movement of the gut, and transmits the images outside the body. However, the current WCE technique cannot provide an accurate location of the disease that is found in the capsule endoscopic video therefore making it difficult to apply further medical operation or other treatments such as biopsy or drug delivery. Thus localizing the WCE has been gaining attention in recent years. Meng et al. suggested using magnetic localization techniques for WCE [5], while the original localization methods for WCE were based on the Received Signal Strength (RSS) [4]. Recent research shows that higher localization accuracy is achievable in RSS-based methods using ultra-wideband techniques [1]. Regardless of the measurement type, the WCE localization problem could be considered an estimation problem with prior knowledge or in other words, a tracking problem. However none of the methods so far take advantage of the specific movement characteristics of the WCE. The small intestine of a human being is a 21-foot long coiled organ that connects the stomach to the large intestine. Being narrow and twisted, it restricts the movement of the WCE to a specific type. In this paper, we take advantage of the special characteristics in WCE movement to present it as a multi mode random process. Then the performance of two multi model tracking methods is evaluated based on the movement model.

The following section introduces the movement model. In section 3 after a short introduction on multi model tracking, two very simple multi model methods are presented. The detection performance and its effect on overall tracking is considered in section 4 and simulation results are presented in section 5. Finally the paper is concluded in section 6.

2. THE MOVEMENT MODEL

The capsule endoscopes which are currently in use in clinical practices depend on normal peristaltic movement of the gut for their movement. Therefore the movement behavior is totally governed by stress and strain cycle of the intestines which itself is controlled by the nerve system. Clinical observations show that the stress and strain cycle in the guts can be considered constant as far as the patient is under normal condition [10]. In addition to how often the intestine muscles are stimulated by the nerve system, the physical features of the endoscope is also a main factor in its speed. Previous investigations showed that the relative ratio of the diameter of the WCE and the inner diameter of the intestine is the most important physical feature in defining the speed [12]. The diameter of the WCE is fixed throughout the endoscopy however the inner diameter of the intestine varies for differ-
ent branches in the intestine [12, 10]. Moreover, the WCE is expected to linger in intestine bends for random amount of time based on the geometry of the bend and the capsule. There is also a very low possibility of stricture (1%) which generally happens in the bends [10]. Thus the movement of WCE is expected to have small variations in speed as long as WCE is moving inside a certain intestine branch with a certain inner diameter. So in the modeling procedure it is assumed that the movement in the small intestine consists of different tubes in which the WCE moves with almost constant speed while different tubes are expected to impose different speeds on WCE. Also at the end of each tube the WCE stops for a random amount of time. Assuming that there exist \( N \) speed modes in the whole track (which indicates existing of \( N \) inner diameter), the WCE speed at time instant \( i \) is approximated as a random variable with a Gaussian Mixture pmf, such that:

\[
v(i) \approx \sum_{n=-N}^{N} p(n)f(v; u_n, q_n),
\]

where \( f(v; u_n, q_n) \) is a Gaussian PDF with mean \( u_n \), covariance \( q_n \), and probability \( p(n) \). For simplicity we assume

\[q_n = q_1 \quad \forall q_n,
\]

where \( q_1 \) is considered to be very small compared to \( u_n \). For passive movements \( u_n \) is assumed to be less than \( 0.5 \text{mm/s} \).

A discrete review of the model introduced here reveals its similarities with the Random Way point model (RWP), which is a very prevalent movement model for Ad-Hoc networks [3]. Therefore, the results from our work also applies to RWP with small modifications.

Kalman Filtering is widely used for the tracking applications [8] and it is proven to be a very strong tool as long as the assumptions in the structure are met in the actual scenario. To adopt Kalman structure for any problem one must express it in two equations: first the plant equation and second the observation equation. Since it is assumed that the WCE has a constant speed with very minor deviations \( q_1 \) in each tube as in (1) the state vector at time instant \( i \) can be written as:

\[
X(i) = [x(i) \ y(i) \ z(i) \ \nu_x(i) \ \nu_y(i) \ \nu_z(i)]^T, \quad (3)
\]

where \((x, y, z)\) represents the three dimensional location vector and \((\nu_x, \nu_y, \nu_z)\) is the three dimensional speed vector. It is assumed that only distance measurements are available. Based on this assumption we are only able to observe a function of the location at each sensor in White Gaussian Noise (WGN). In other words the \( n \)th sensor at time instant \( i \) will observe:

\[
Z_n(i) = h_n(X(i)) + w_n(i) \quad (4)
\]

where

\[h_n(i) = \sqrt{(x_n - x(i))^2 + (y_n - y(i))^2 + (z_n - z(i))^2},
\]

In (4) \( w_n(i) \) is the observation noise at time \( i \) for the sensor \( n \) and follows a \( N(0, R_n(i)) \) distribution; \((x_n, y_n, z_n)\) represents the location of the sensor. An observation equation in the form of (4) is justified if the RSS measurements are used at the sensor nodes. The reason is that the radio signal inside the human body or any absorbing environment is proven to decay exponentially so the power in \( \text{dB} \) format is linearly dependent on the distance [11]. Equation (5) is also valid if the measurements are static magnetic field strengths as it is proven by [6]. In order to keep the math simple in the rest of the paper it is assumed that the tracking is done only in one dimension and using only one sensor which is located at \( x_n = 0 \). Such an assumption would reduce the observation equation in (4) to:

\[
Z(i) = HX(i) + w(i) \quad \text{where } H = [
\begin{bmatrix}
1 & 0
\end{bmatrix}
\]

On the other hand, the plant equation is not straightforward to write due to random jumps in speed. However, as long as the WCE is moving inside a specific tube with constant inner diameter, there will be very small variation in the speed. Thus, in the absence of maneuver the target is modeled as a constant velocity object in a plane with some process noise to model the slight changes in the velocity. The plant equation for this quiescent model, discretized over time intervals of length \( t_s \), is:

\[
\begin{bmatrix}
x(i+1) \\
\nu_x(i+1)
\end{bmatrix} = F
\begin{bmatrix}
x(i) \\
\nu_x(i)
\end{bmatrix} + G\nu(i+1) \quad (7)
\]

For simplicity equation (7) is written only for one dimension but its extension to three dimensions is trivial. In (7) \( \nu(i+1) \) is the speed variation between time instants \( i+1 \) and \( i \), and it is modeled as a \( \mathcal{N}(0, q_1) \). The process equation assumed in (7) is the same as the model used in [2] to track a moving target with constant speed. It is important to note from (2) that \( q_1 \) is a very small number in comparison to \( u_n \) in (1). This kind of modeling enables the tracking system to follow small variations in the speed and track the location with small errors. However, whenever a maneuver happens and the speed jumps from one value to zero or the other way, the Kalman filter fails to track the drastic change. The reason is such sudden maneuvers are not considered in the Kalman filtering frame work. Due to the fact that the process noise variance \( q_1 \) is small, in response to the speed jump, the Kalman filter approaches the new speed very slowly which produces a very large error in the location estimation. To address the sudden changes in the WCE tracking problem we suggest using Multiple-Model (MM) tracking methods.

3. **MULTI MODEL TRACKING METHODS**

Multiple-model tracking methods have been generally considered for maneuvering target tracking under motion-mode uncertainty. The target motion-mode uncertainty exhibits itself in the situations where a target may undergo a known or unknown maneuver during an unknown time period. In a simple form for MM the target at time instant \( i \) is assumed to be in mode \( m^{(i)} \in M \), were \( M \) is the set of all possible modes.

\[M = \{m_1, m_2, \ldots, m_l\},
\]

where \( l \) is the number of all possible modes. Thus the estimation problem consists of two parts: estimating the current mode and the current state. Due to such a nature MM problems are also called **hybrid** estimation problem. If all possible modes of the system are known, a multiple model method can be designed for tracking with following plant equation:

\[
X(i+1) = F^{m_{(i)}}X(i) + G^{m_{(i)}}\nu(i+1). \quad (9)
\]
In (9) \( X \) is the state vector and \( \nu \) is the process noise. The superscript \( m \) in (9) indicates that the \( m \)th model is being used while the target is moving based on the \( m \)th mode.

In the WCE tracking problem, the movement is limited to two modes. The first one is the normal movement in which the WCE is moving or it is stalled at a bend in the intestine. While the second one is the maneuver mode in which the WCE suddenly stops moving or starts moving again. Thus the maneuvering mode is the transition between the moving and the stopping episodes. There are many ways to track a maneuvering target [2, 9] however in this paper we adopt Variable Noise Level and Variable State Dimension methods due to their simplicity. At each time instant both methods detect the mode first and then estimate the state. Therefore both methods deal with the hybrid estimation problem as a detection-estimation problem assuming that there is only one correct mode at time instant \( i \).

### 3.1 Variable Noise Level

For the tracking problem with a plant equation formulated in (7), using a very large process noise variance than \( q_1 \) would enable tracking in the presence of sudden changes in the speed. However, allowing a larger process noise would mean accepting a larger estimation error while the tracking is done under the normal circumstances. In addition, setting the process noise to \( q_1 \) would result in smaller error in normal mode and a very large error under the maneuvering mode. In order to balance such a trade-off, Variable Noise Level (VNL) method suggests using two noise variances. The first noise level, \( q_1 \) is used as long as the WCE speed is constant and variations in the speed are small. As soon as a maneuver is detected the tracking system sets the noise variance to a \( q_2 \), where \( q_2 > q_1 \).

The VNL algorithm simply starts a Kalman filter for normal mode, i.e. the process noise is set to a very small number. After detecting a maneuver at time instant \( i \), the method resets the Kalman filter by setting the \( \hat{X} = X^m \) and \( P(i) = P^m(i) \) and sets the process noise to \( q_2 \). Thus selecting the right metric for the maneuver detection plays a very important role in the method. One of the marvels of the Kalman filter is its ability to produces the estimation accuracy measure during the estimation process. therefore as the first candidate, using the estimation covariance matrix as a detection metric might be considered. However in the maneuvering target case, such a performance measure cannot be used to detect the maneuver since the error covariance matrix is calculated not using the actual observations but the statistics of the prior and the observations [8]. While a good measure to detect the maneuver must put into account both the actual measurements and the predicted state values. Such a property is only found in the innovation process in the Kalman filter.

Table (1) shows the steps in VNL algorithm for a system introduced with equation (6)and (7). The innovation process is the difference between the measured data and its predicted value. Since the prediction is done based on the chosen model, the amount of the innovation shows how close the model based prediction is to the actual measurement. If the model is selected correctly the innovation process should be a \( N(0, S(i)) \) process because both noises are considered to be Gaussian. However, after a sudden change in the speed, since the prediction is still done by the old model, the innovation would be larger than expected and as long as the maneuver is not detected the innovation increases in mean. In table (1) to reduce the effect of noise a moving average of \( \epsilon_p(i) \) is used as the test statistic for the maneuver detection and whenever \( \epsilon_p(i) > \tau_1 \) the maneuver is detected. After detecting the maneuver tracking continues with \( q_2 \) for \( r \) time steps and then switches back to the normal tracking model and monitors the \( \epsilon_p(i) \). The value \( r \) is chosen long enough so that the estimated state tracks the new speed, and \( q_2 \) is selected in proportion with \( r \) to enable a fast change in the speed. Setting \( q_2 \) to the variance of the GMM in (1) would be a good candidate.

### 3.2 Variable State Dimension

In the variable State Dimension (VSD) method the detection procedure is the same as VNL however the reaction to the maneuver is different. In VSD upon detecting the maneuver the state vector is changed to:

\[
X^m = \begin{bmatrix} x & v_x & a_x \end{bmatrix}, \tag{10}
\]

and the plant equation is changed to:

\[
X(i + 1) = F^m X(i) + G^m v^m(i), \tag{11}
\]

where

\[
F^m = \begin{bmatrix} 1 & ts & ts^2/2 \\ 0 & 1 & ts \\ 0 & 0 & 1 \end{bmatrix}, \quad G^m = \begin{bmatrix} ts^2/4 \\ ts^2 \\ 1 \end{bmatrix}, \tag{12}
\]

and the observation equation is the same as (6) with an additional column:

\[
H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}. \tag{13}
\]

Also after detecting the maneuver, the tracking system resets the Kalman filter by setting \( X(i) = X^m \) and \( P(i) = P^m(i) \). To choose \( X^m \) and \( P^m(i) \) in [9] it is assumed that the onset of the maneuver is \( L \) steps before its detection, so the filter goes back \( L \) steps and re-runs the estimation and the \( X^m \) and \( P^m(i) \) values are set accordingly. Using the method in [9] is helpful if the assumption on the maneuver onset time is correct. However, the method produces a very large error if the maneuver detection happens long after the onset of the maneuver.
1. Prediction (Model-Based):
   (for m=1,2)
   Predicted state: \( X^+(i) = F \hat{X}(i-1) \)
   Predicted covariance: \( P^+(i) = FP(i-1)F^T + BQ_1B^T \)
   Innovation process: \( \pi(i) = Z(i) - X^+(i) \)
   Innovation covariance: \( S(i) = HP(i)H^T + R(i) \)
   Filter gain: \( K(i) = P^+(i)H^T S(i)^{-1} \)

2. Mode update:
   If the system is not in maneuvering mode
   Normalized innovation: \( \epsilon_p(i) = \pi(i) S(i)^{-1} \pi(i) \)
   Test statistics: \( \epsilon_p(i) ^2 = \frac{1}{2} \sum_{l=0}^{\infty} \epsilon_p(i-l) \)
   Decide Mode(m=2): \( \epsilon_p(i) > \tau_1 \)

2. Correction:
   Overall estimate: \( \hat{X}(i) = X^+(i) + K(i) \pi(i) \)
   Overall covariance: \( P(i) = (I - K(i))H P^+(i) \)

\( F = F^m, G = G^m, H = H^m \)
\( X^m(i) = \begin{bmatrix} Z(i) \\ 0 \\ 0 \end{bmatrix}, \quad P^m(i) = \begin{bmatrix} R & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & a_0 \end{bmatrix} \)

If the system is in maneuvering mode
estimation significance: \( \delta_a(i) = a^T P^m(i) a \)
Test statistics: \( \delta_p^2(i) = \sum_{l=0}^{\infty} \delta_a(i-l) \)
Decide Mode(m=1): \( \delta_p^2(i) < \tau_2 \)

\( \frac{\text{after maneuver onset. Such a scenario is expected to happen whenever the Signal to Noise Ratio (SNR) is low. To address this problem in this paper we assign:}}{X^m(i) = \begin{bmatrix} Z(i) \\ 0 \\ 0 \end{bmatrix}, \quad P^m(i) = \begin{bmatrix} R & 0 & 0 \\ 0 & q_2 & 0 \\ 0 & 0 & a_0 \end{bmatrix}} \)

\( a_0 = q_2/(t_s)^2 \)
Changing the state vector as in (10)-(12) means that the system is assumed to have acceleration while it is in maneuvering mode. Thus as long as the estimated acceleration is statistically significant, the system is considered to be in the maneuvering mode. The test statistics for the significance of the estimated acceleration is:
\( \delta_a(i) = a^T P^m(i) a \)
\( \delta_p^2(i) = \sum_{l=0}^{\infty} \delta_a(i-l) \)
where \( a = \hat{x} \) is the estimated acceleration and \( P^m(i) \) is the respective error covariance. To reduce the effect of noise it is common to use a moving average of the \( \delta_a(i) \) as the test statistic. Table(2) shows the summary of the VSD method.

4. DETECTION PROBLEM FORMULATION

Regardless of which method being used, the maneuver detection criterion is the same. As mentioned in section (3) the innovation process the parameter is used to detect the maneuver and when the system model is correct, the innovation process \( \pi(i) \) is expected to have following characteristics:
\( \pi(i) \sim \mathcal{N}(0, S(i)), \quad S(i) = HP^+(i)H^T + R. \)

However the test statistics used for maneuver detection in both methods was \( \epsilon_p^L(i) \), which consists of adding \( L \) squared zero mean normalized Gaussian random variables. So the test statistics under the null hypothesis has the following distribution:
\( L \epsilon_p^L(i) \sim \chi_L^2. \)

Based on (17) the probability of False Alarm is
\( \Pr(\epsilon_p^L(i) > \tau_1|H_0), \)
where \( H_0 \) represents the null hypothesis. The threshold in (18) is decided based on the desired false alarm.

The statistical model in equation (16) is valid before the maneuver. Assuming the maneuver onset is \( i-1 \), the innovation process at \( i \) is written as:
\( \pi(i) = x(i-1) - \hat{x}(i-1) + t_s(v(i-1) - \hat{v}(i-1) + b(i)) + \epsilon(i), \)
\( \epsilon(i) \sim N(0, \lambda S_i) \),
\( \lambda = \frac{\sum_{i=0}^{L-1} b(i)}{S(i-1)} \).

The following step after the maneuver, only one of the Gaussian terms in \( \epsilon_p^L(i) \) has a non-zero mean, but after \( L \) observations all terms in \( \epsilon_p^L(i) \) are non-zero mean Gaussians. Also as long as the maneuver has not been detected, the mean of the innovation process will grow since the location prediction is still done based on the previous speed. Therefore the non-centrality parameter in (20) grows as well. Remembering the fact that the detection threshold is a function of false alarm and it is a fixed number, the probability of detection is expected to increase in time. Figure (1) shows the test statistics under three SNRs. The maneuver onset is \( i = 300 \) and after it, the test statistics starts to increase almost linearly for all cases however for larger SNRs the increase is sharper. So as the test statistics increases the detection probability for the same false alarm also increases.

On the other hand a late maneuver detection results in a very high. Another interesting behavior in figure(1) is that after some time the innovation starts to decline which is due to the fact that the Kalman filter is still able to track the change very slowly even when the used model is wrong.

5. SIMULATION RESULTS
Applying multi-model Kalman filtering to Wireless Capsule Endoscopy tracking problem is studied in this paper. The movement of the WCE is modeled as a multi mode random process based on the nature of the track. Two methods from maneuvering target tracking are adapted for WCE localization and through simulations it is shown that the localization performance would increases substantially if the prior knowledge was used in the localization. It is presented that the multi mode tracking problem is a combined detection-estimation problem and it is shown how selecting parameters on the detection side influences the overall estimation performance.

7. REFERENCES


