Image-based Localization of the Active Wireless Capsule Endoscope inside the Stomach

M. Aghanouri, A. Ghaffari, and N. Dadashi

***Abstract*—Wireless capsule endoscopy (WCE) is a new technology which is introduced in recent years for examining the Gastrointestinal (GI) tract. Active WCE is a class of WCE in which the capsule's movement is controllable. Since there is a large open space in the stomach, steering the capsule in the proper distance from the walls of the stomach in order to achieve suitable images, is a challenge. Another challenge of using the active capsule in the stomach is its localization. The localization can be used as the feedback of the controller. In this paper for the first time for localizing the active capsule endoscope inside the stomach, a method is proposed in which, the rotation angle about the main axis of the capsule and the scales of the images are obtained by using Speeded-Up Robust Features (SURF), Fast Library for Approximate Nearest Neighbors (FLANN) and M-estimator SAmple and Consensus (MSAC). Also a new method is presented that modifies the distance between the capsule and the walls of the stomach by using scales of the images and fuzzy rules. The results of this work show a good estimation of the rotation angle of the capsule and scales of the images. The findings of this paper can be utilized for improving the control of the active capsule endoscope inside the stomach.**

I. INTRODUCTION

Gastrointestinal (GI) endoscopy is a minimally invasive procedure which examines GI tract. Conventionally this procedure is done by the flexible endoscope device which is not very comfortable for either the physician or the patient [1]. So a new method, wireless capsule endoscopy (WCE), is released in recent years. This technique is already used for small intestine [2], while its application for the upper GI specially stomach, because of existence of the large open space in the stomach, is a challenge. For examining different parts of the stomach, active WCE, in which the endoscopic capsule is steerable, is needed [1]. One of the most important parts of steering system is localization of the capsule which can be used for feedback of the steering system.

In the literature, different methods of localization for active and passive wireless endoscopic capsule are presented. Than et al. reviewed different types of localization method [3]. Among them for the active capsules, magnetic localization methods can derive the most number of DOFs of the capsule however they can't measure the rotation angle about the main axes of the capsule. Besides, these methods are very complicated and expensive, while image based methods are more accessible and less expensive. The works based on images [4-7], are developed for the capsule endoscope in the small intestine and never are developed for the stomach. These methods are generally based on image's local feature extraction and matching

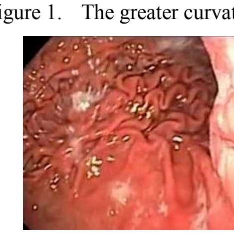
In this paper, an approach for localizing active wireless endoscopic capsule in the stomach, based on feature matching is proposed. Since this localization will be used as the feedback of the steering system (controller), features are matched between two consecutive frames. Two important parameters related to the position and orientation of the capsule will be derived. The first parameter is the rotation angle along the main axis of the capsule which cannot be calculated from any other types of localization methods. The second one is the information about the distance between the camera and surface of the walls of the stomach. This information is really important for steering system, because if the distance between the capsule and the surface is not proper, the images will not be useful due to the existence of open space in the stomach. For achieving these goals, four steps should be followed. In the first step, features are detected and described. Then feature matching is applied. In the next step outliers are eliminated and the relative rotation and scaling of the frames is estimated. Finally orientation information of the wireless capsule endoscope and image's scale factor are derived and evaluated and the method of adjusting the distance between the capsule and the walls of the stomach by defining the fuzzy rules is presented. The rest of this paper consists of 4 sections. Section II includes details about the proposed method. In section III, experimental results are used to evaluate the method, and finally in section IV, the conclusion is drawn.

II. THE METHOD FOR LOCALIZATION OF THE ACTIVE WIRELESS CAPSULE ENDOSCOPE INSIDE THE STOMACH

A. Feature Extraction

Since for achieving enough matching inliers, the number of detected interest points (features) should be large, a good feature detector should be selected according to nature of the walls which are covered by folds. In this paper the part of the walls which is called the greater curvature is considered and evaluated (Fig. 1).

Figure 1. The greater curvature



M. Aghanouri is with the Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran (corresponding author to provide phone: +98-913-2159054; e-mail: mehrnaz.aghanouri@gmail.com).

A. Ghaffari, is with the Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran (e-mail: ghaffari@kntu.ac.ir).

N. Dadashi is with the Medical Image and Signal Processing Research Center, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran (e-mail: nm.dadashi@yahoo.com).

978-1-5090-4179-4/17/$31.00 ©2017 IEEE

Different well known methods, Speeded-Up Robust Features (SURF) [8], Features from Accelerated Segment Test (FAST) [9] and Harris Corner [10] detectors were applied on the images to find the local interest points. Among them SURF detector satisfies the large number of features condition, so it is selected as the suitable method. In the Surf algorithm, a fast approximation of the Hessian matrix that exploits integral images is adopted and the local maxima of Hessian matrix determinate are determined and specified as the interest points. In order to describe interest points, also SURF descriptor is used. This descriptor again exploits integral images for speed and describes a distribution of Haar-wavelet responses around the detected interest point. In this paper, in order to improve the accuracy, the length of descriptor is considered to be 128. It is important to say that SURF detector and descriptor are invariant to rotation, contrast and scale changes, and thus they are very suitable for endoscopy images. Also SURF algorithm is very fast and robust in comparison to the other similar methods. An example of extracting features by SURF method is shown in Fig. 2. Since the number of detected features is large, in Fig. 2, just the 50 strongest features are demonstrated.

B. Feature Matching between two consecutive frames

In this step, the corresponding features in two frames should be matched. The speed of WCE should be low to make possible accurate investigation in the stomach. So assuming similar features in two frames is not inconsequent. The features are matched which the distance between their describing vectors is less than a threshold. Because the number of features is large, for fast searching, Fast Library for Approximate Nearest Neighbors (FLANN) [11] is used. In this method the best algorithm and optimum parameters are chosen automatically based on dataset. Fig.3 shows the matched features between two frames.

C. Geometric Transformation Calculation and Outliers Elimination

The geometric transformation between two images should be calculated by use of corresponding features which are determined in previous step. Actually a transformation model should be fit on matched features. A well-known method, RANdom SAmple Consensus (RANSAC) [12], can be used to do this task. RANSAC estimates the parameters of a model by starting from a small set of data which may contain lots of outliers and then enlarging this set with feasible data which are compatible with the estimated model. In this paper M-estimator SAmple and Consensus (MSAC) [13] which is a variant of RANSAC is used. This algorithm yields more benefit to RANSAC without any additional computational burden [14].

In MSAC algorithm, a re-descending M-estimator is used in order to score inliers due to their fitness to the model. In this algorithm the outliers are weighted by a constant number. In this work, due to the requisites, the similarity transformation is calculated which needs at least 2 points to initialize the model.

Figure 2. The 50 strongest extracted features by SURF method.

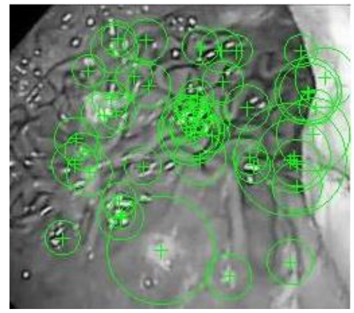
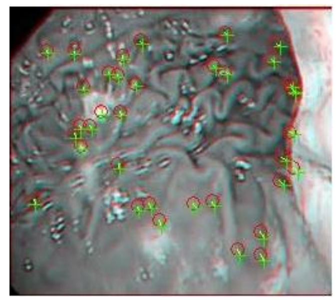


Figure 3. The matched features between two cnesequtive frames by using FLANN.



We can divide the algorithm of calculation of geometry transformation into 4 steps:

1- Two points are selected randomly and by using them, similarity transformation is estimated.
2- Inliers are scored according to their fitness to the model (if their distance to the estimated model is less than a predefined value) and outliers are specified.
3- Step 1 and 2 are iterated until the number of the iterations satisfies the predefined iteration value.
4- Finally the similarity transformation is selected which its number of inliers satisfies a predefined value.

The transformation matrix is determined as:

$$T = \begin{bmatrix} sc & -ss & 0 \\ ss & sc & 0 \\ t_x & t_y & 1 \end{bmatrix} \quad (1)$$

in which t_x and t_y are translations in x and y direction respectively and s is the scale factor and we have:

$$sc = s\cos(\theta), ss = s\sin(\theta) \quad (2)$$

Equations (3) and (4) show how we can find the values of rotation angle and scale factor from the transformation matrix.

$$s = \sqrt{(sc)^2 + (ss)^2} \quad (3)$$

$$\theta = \tan^{-1}(\frac{ss}{sc}) \quad (4)$$

D. Active wireless capsule endoscope Localization

The rotation angle about the main axis of the capsule is derived directly from (4). As it was mentioned in section I, the distance between the camera and the walls of the stomach is very important. To achieve a proper distance, the following concept is considered:

The camera is kept in a specified depth and two consecutive frames are considered (z-axis is considered along the optical axis of the camera). The scale factor of these frames is obtained. As an example, a scale factor greater than one means that the distance between the camera and the scene become lesser in the second frame. Since the camera is kept in a constant depth, it can be concluded that the scene is become closer to the camera in the second frame. Thus in this case to achieve a proper distance, the camera (capsule) should be moved in the opposite direction of the optical axis ($-\Delta z$).

Five steps should be done for modifying the distance:

1- The relation between changing the position of the capsule along the optical axis (Δz) and the scale factor, also the acceptable and proper range for the scale factor should be obtained by a kind of calibration. Due to the open space in the stomach and possible distance between its walls and the capsule, physicians believe that the proper distance between the camera and the surface of the stomach is about 4 cm (z_0=4 cm). As a result, Δz can be defined as constant numbers -3, -2, -1, 0, 1, 2 and 3. In the proposed calibration, first, the camera (capsule) is placed 4 cm far away from the walls (here the greater curvature) of the phantom model of the stomach and image 1 is acquired. This image has a suitable scale and quality due to the proper distance between the camera and the wall. After that, the capsule is shifted 1 cm away along the optical axis (Δz= -1), the new picture is acquired and the scale factor of this image and image 1 is obtained (s_{-1}). This procedure is repeated for Δz=-2, -3, 1, 2 and 3 related to the first position (Fig. 4), and five scale factors which are named s_{-2}, s_{-3}, s_1, s_2 and s_3 respectively, are determined. Now the ranges of the scale factors are defined by the triangular membership functions (Fig. 5). The calibration step should be done before starting the stomach examination. It is enough to do this step just one time for each type of the active capsule endoscope.

2- In this step, examining of the stomach is started. The capsule should be parked so that the proper image with acceptable scale is achieved. Since this image is obtained when the capsule is in a suitable position and proper distance from the surface of the stomach (for example the greater curvature), it can be concluded that the distance between the camera and the wall is about 4 cm.

3- The capsule is moved along the wall as Δx cm and 2 consecutive frames which are acquired in positions R_1 and R_2 (Fig. 6) are considered. Δx should be enough small so that there is enough corresponding feature in two acquired images.

4- The scale factor between two frames (s) is calculated.

5- The distance between the camera and the walls is modified according to the following fuzzy rules and by using zero-order Sugeno fuzzy model.

If s is Too Large then Δz is -3.
If s is Very Large then Δz is -2.
If s is Large then Δz is -1.
If s is Proper then Δz is 0.

Figure 4. The different position of the capsule related to the greater curvature in the proposed calibration.

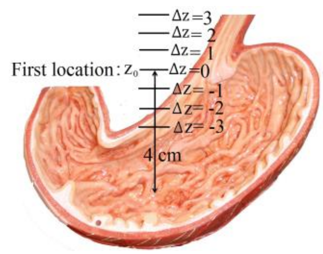
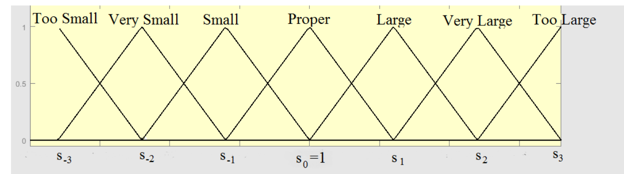


Figure 5. The triangular membership functions of the scale factors.



If s is Small then Δz is 1.
If s is Very Small then Δz is 2.
If s is Too Small then Δz is 3.
If s is NA then Δz is 3.

5- Steps 3, 4 and 5 are repeated until the examination is finished.

It should be noticed that the value of proper distance (z_0) and the values of Δz can be changed by the user.

III. EXPERIMENTAL RESULTS

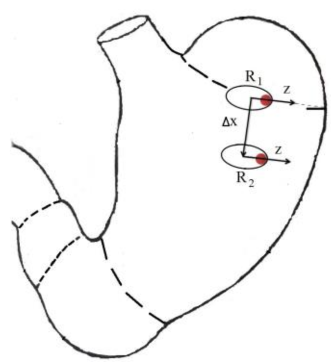
The method is applied on the frames of the video of upper GI endoscopy and the frames related to the greater curvature are considered. Because of the absence of the suitable WCE video from the stomach, and also since a proper wireless capsule endoscope maneuver should be similar to the flexible endoscope device, the endoscopic video of upper gastrointestinal which is recorded by the flexible endoscope [15] is used for evaluating the method.

In order to verify the method, the simulated rotation and scale transformation of the frames is used which also is adopted by [4].

A. Orientation

For evaluating the method about the rotation, 10 rotation angles from 5° to 50° with the step of 5°, are tested. Table I, shows the rotation angles calculated by the presented method and the error between them and actual angles.

Figure 6. The position of the capsule and its moving along the wall of the stomach.



It is clear from table I that the maximum error is at 45° and is not really significant. It can be concluded that the proposed method is suitable for computing the rotation angle about the main axis of the capsule inside the stomach.

B. Scale factor

Different scale factors, which are shown in table II, are tested and the results obtained by using the presented method are compared with the simulated scales. According to table II, the results from scale 0.4 to 5 are satisfying while for the scales smaller than 0.4 enough corresponding points are not founded and the error increases a lot for scales greater than 6.

IV. CONCLUSION

In this paper, for the first time, an image-based localization method is presented for localizing the active wireless capsule endoscope inside the stomach. The rotation angle of the capsule and the scale factor are derived by using the SURF algorithm, FLANN and MSAC. Also a novel method based on fuzzy logic is suggested for modifying the distance between the capsule and the walls of the stomach. The results show that the presented method can successfully estimate the rotation angle. Since the scale factor greater than 6 will not happen during examination due to the stomach size, the proposed method for modifying the distance can be used successfully.

TABLE I. ROTATION ANGLES

Actual rotation angles (degree)	*Estimated rotation angles (degree)*	*Absolute Error (degree)*
5	4.9948	0.0052
10	10.0400	0.0400
15	14.9140	0.0860
20	20.0905	0.0905
25	24.8595	0.1405
30	29.8505	0.1495
35	34.7980	0.2020
40	39.9993	0.0007
45	44.7188	0.2812
50	50.0518	0.0518

TABLE II. SCALE FACTOS

Actual scale factors	*Estimated scale factors*	*Absolute Error*
0.2	N/A	-
0.4	0.4046	0.0046
0.6	0.6011	0.0011
0.8	0.7996	0.0004
1.2	1.2004	0.0004
1.4	1.4004	0.0004
1.6	1.6016	0.0016
1.8	1.8006	0.0006
2	1.9993	0.0007
2.5	2.4999	0.0001
3	2.9976	0.0024
4	4.0006	0.0006
5	4.9846	0.0154
6	0.8396	5.1604

REFERENCES

[1] H. Keller *et al*., "Method for navigation and control of a magnetically guided capsule endoscope in the human stomach," *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, Rome, 2012, pp. 859-865.

[2] Z. Liao, R. Gao, C. Xu, and Z. Li, “Indications and detection, completion, and retention rates of small-bowel capsule endoscopy: a systematic review,” Gastrointestinal Endoscopy, vol. 71, no. 2, pp. 280–286, 2010.

[3] T. D. Than, G. Alici, H. Zhou and W. Li, "A Review of Localization Systems for Robotic Endoscopic Capsules," in *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2387-2399, Sept. 2012.

[4] L. Liu, C. Hu, W. Cai, and M.-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. IEEE, 2009, pp. 3711–3714.

[5] D. K. Iakovidis, E. Spyrou, D. Diamantis, and I. Tsiompanidis, “Capsule endoscope localization based on visual features,” in Bioinformatics and Bioengineering (BIBE), 2013 IEEE 13th International Conference on. IEEE, 2013, pp. 1–4.

[6] E. Spyrou and D. K. Iakovidis, "Homography-based orientation estimation for capsule endoscope tracking," 2012 IEEE International Conference on Imaging Systems and Techniques Proceedings, Manchester, 2012, pp. 101-105.

[7] E. Spyrou and D. K. Iakovidis, “Video-based measurements for wireless capsule endoscope tracking,” *Measurement Science and Technology*, vol. 25, no. 1, pp. 015002, 2014.

[8] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool. Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, 2008.

[9] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, Beijing, 2005, pp. 1508-1515 Vol. 2.

[10] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," Proceedings of the 4th Alvey Vision Conference, 1988, pp. 147-151.

[11] M. Muja and D. G. Lowe, "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration." International Conference on Computer Vision Theory and Applications.VISAPP, 2009.

[12] M.A. Fischler and R.C. Bolles, “Random Sample Consensus – A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,” Communications of the ACM, 24(6):381-395, 1981.

[13] P.H.S. Torr and A. Zisserman, “MLESAC: A new robust estimator with application to estimating image geometry”, *Journal of Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138-156, 2000.

[14] M. Zuliani, "RANSAC for Dummies." *Vision Research Lab, University of California, Santa Barbara,* 2009.

[15] www.murrasaca.com