

The Use of the Focus of Expansion for Automated Steering of Flexible Endoscopes

N. van der Stap, R. Reilink, *Student Member, IEEE*, S. Misra, *Member, IEEE*, I.A.M.J. Broeders and F. van der Heijden

Abstract— Colon cancer screening remains a time-consuming and expensive clinical process. Automating flexible endoscopy has the potential to increase screening efficiency. In this research the images captured by the camera at the endoscope tip are used to find the heading direction of the endoscope. Comparing the current heading direction to the desired target direction in a computer algorithm is expected to allow automated steering of the endoscope. Heading direction determination is achieved using an estimation of the focus of expansion (FOE) from the optical flow field. The resulting heading direction is compared to results obtained manually by several human observers. From our experiments it becomes clear that the FOE can be used as a reliable estimator for heading direction in human colonoscopy images. Additionally, the automated results have an intraclass correlation of 89% with the manual results, demonstrating that the algorithm works as expected. It is anticipated that the final steering algorithm can be used in a variety of motorized flexible endoscope applications.

I. INTRODUCTION

FLEXIBLE endoscopy can be performed to obtain a diagnosis and to perform small interventions in the human body without leaving scars. In screening for colorectal cancer [1], flexible endoscopy of the large bowel (colonoscopy) is performed to diagnose and, if necessary, remove lesions of the bowel wall. It is expected that the number of colonoscopies performed yearly in the Netherlands will increase by 60.000 after the start of the population screening program in 2013 [2]. This rise motivates hospitals to increase their colonoscopy capacity and efficiency drastically.

Currently, an endoscope driver (endoscopist) needs to perform a large number of procedures to reach the level of competence. Estimates of this number vary from 100 to 500 colonoscopic procedures over periods of 1 to 3 years

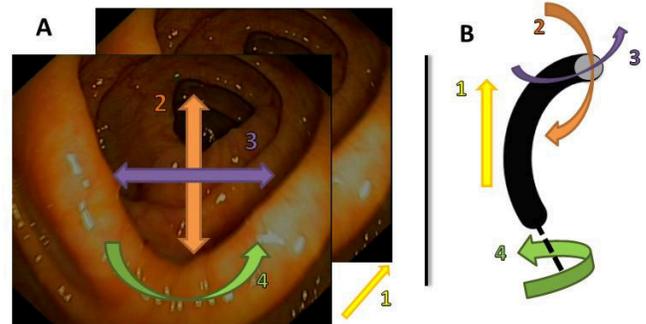


Fig. 1. A) consecutive images of the human bowel. B) schematic endoscope. Possible degrees of freedom are depicted and numbered in both. The endoscope is inserted in direction 1. It can be rotated along the shaft (direction 4), and the tip can bend up/down (direction 2) and left/right (direction 3).

[3]–[5]. Shortening the learning curve by facilitating the procedure would be one way to increase efficiency in colonoscopy departments. A second approach could be to have less-trained personnel perform the routine procedures, while leaving the more complicated ones to more highly trained gastroenterologists or surgeons. Finally, the introduction time of colonoscopes could be reduced to increase productivity. This research is aimed at a simplification of the steering mechanism of flexible endoscopes in order to bring about efficiency increases.

A. The Endoscope

A flexible endoscope consists of a hand grip attached to a flexible tube with a camera, an illuminator, and one or more working channels [6], [7]. The camera consists of a lens and a charged coupled device (CCD) or complementary metal oxide semiconductor (CMOS) chip. Manual control of flexible endoscopes is considered impractical and unintuitive [8]. The endoscopist folds one of his or her hands around the hand grip in order to steer the instrument. Two steering knobs on the hand piece can be turned to gradually bend the tip in the directions 2 and 3 (Fig. 1). The tip of the endoscope has four degrees of freedom: longitudinal translation and rotation (represented by arrows 1 and 4), and two combined rotations/translations (arrows 2 and 3). A predominantly bimanual steering method frequently results from the limiting design of the device, because most hands are too small for single-handed steering. However, bimanual steering does not allow simultaneous insertion or retraction of the endoscope by the same person.

The impractical way of steering has motivated scientists to consider automated flexible endoscope steering. Khan and

Manuscript received January 31, 2012.

N. van der Stap is with the Robotics and Minimally Invasive Surgery Group, MIRA - Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands (corresponding author: 0031-30-850-2571; e-mail: n.stap@utwente.nl).

R. Reilink and S. Misra are with the Control Engineering Group, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands.

I. A. M. J. Broeders is head of the Robotics and Minimally Invasive Surgery Group, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands.

F. van der Heijden is with the Signals and Systems Group of MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, The Netherlands.



Fig. 2. Examples of human colonoscopy images. The first two images consist of a lumen view of very different human colons. The third image (right) is an example of a blurred image, caused by irrigation to clean the lens.

Gillies [9] tried steering a flexible endoscope automatically using computer vision. Edge detection and the wall contour shape of the bowel were used to estimate depth and steer the endoscope. Although the steering itself was successful, the interpretation of images of the wall remained problematic. Also, many computer hardware adjustments were needed.

Reilink et al. [10] also investigated image-guided flexible endoscope steering. They used an ‘image intensity’-based method and an ‘optical flow’-based method to steer the endoscope tip towards the center of the lumen (the cavity enclosed by the walls of the bowel). Optical flow (OF) refers to the vectors that represent two-dimensional (2D) displacements of (groups of) pixels in two consecutive images in the image plane. These pixel displacements are caused by the three-dimensional (3D) motion of scene and camera. If the camera is shifted in the longitudinal direction of the tube, objects that are far away will show small or even zero displacements in the image plane. Objects in a more lateral position or nearby will have larger displacements. Theoretically, steering away from the larger displacements will lead the endoscope towards the most distant point in space. The ‘image intensity’ algorithm proposed by Reilink et al. [10] is based on different areas of brightness in the image. The colon lumen is assumed to correspond to a dark area in the image. Steering the endoscope by keeping this dark area in the center of the image will lead to a centered path through the colon. Reilink et al. [10] conducted experiments on both simulated images and images from the inside of an anatomical model. The conclusion was that automatically steering a flexible endoscope was feasible, but more work was needed to design a robust system.

Other research on automated flexible endoscope steering focused mainly on lumen centralization of the tip. In [11], [12] and [13], a variety of methods for lumen detection are proposed.

B. The Images

The images produced by a colonoscope camera viewing the inside of the human colon (real images) are more difficult to use for automated steering than the images used by Reilink et al. [10]. Real images often suffer from motion blur, out-of-focus artifacts and even less texture on wall surfaces than in the anatomical model. Furthermore, in many images the lumen is not located in the field of view (Fig. 2), so it is not enough to assume that centralization of the darkest region is enough to steer the colonoscope.

Although the objective of Mori et al. [14] slightly differed from ours, their approach was also to use optical flow. The detected features and displacement information were combined to reconstruct a 3D replica of the environment of the camera without the need for a lumen view. This led to

the conclusion that a combination of dark region detection (DRD) with OF techniques might provide the necessary information to steer the endoscope automatically. However, the assumption that large displacements indicate nearby objects (Reilink et al., red.) is based on the premise that the tip only moves in the longitudinal direction in a corridor-like environment. Yet in flexible endoscopy, the tip movement and the environment are more complex (Fig. 1). Optical flow displacements form a pattern of vectors that radially emerge from or point to a specific point, depending on the direction of movement. This point is called the focus of expansion (FOE) [15]. The distance of a displacement vector to this FOE also influences vector size, especially when sideward and forward motions are combined (Fig. 3). This fact is not discussed in the other work on flexible endoscope navigation.

C. Research and Paper Outline

The current research is designed around the same driving unit that Reilink et al. [10] describe. The aim of this research is to develop a suitable computer vision algorithm to apply to human colonoscopy images. This algorithm can then be used to automatically steer flexible endoscopes. Two research objectives are defined in this paper. Firstly, OF and the FOE estimation are implemented in a mathematical model. Secondly, we test whether calculation of the OF and a corresponding FOE in human colonoscopic images is feasible.

In Section II of this paper, both relevant terms and the theory that we developed are discussed. In Section III we present an experimental set-up for evaluating FOE estimation in flexible endoscopy. The results in terms of the found accuracy of the FOE and its use in navigation will be given in Section IV. Section V discusses these results, while in Section VI the conclusions of this paper are presented.

II. ENDOSCOPE NAVIGATION

At least three parameters are needed to steer an endoscope (Fig. 3): the direction in which the endoscope is currently heading (heading direction), the direction of the target (target direction) and the direction in which the tip is pointing (tip direction or Optical Axis (OA) as shown in Fig. 3). These are all directions in 3D space, but they can conveniently be represented as 2D positions in the image plane.

It is known that the FOE (Fig. 3) corresponds to the heading direction for translational movements [15], provided that the effect of angular velocities of the tip is negligible. We need to assess whether the FOE holds as a measure for heading direction in our case. Section II.C discusses this aspect in more detail.

The target direction is derived from the darkest region in the image. This assumes no other dark region exists. We observed that the lumen is not visible in all images. However, for now we assume that the darkest region is the most distant point in the colon, which subsequently makes it the target. In the case of the disease diverticulosis, bulges exist in the bowel wall that can be mistaken for the colon lumen when observed from the inside. At the moment, our

algorithm has no way of coping with this situation, but a possible solution is proposed in Chettaoui et al. [11].

The tip direction corresponds to the optical axis and can be determined continuously by measuring steering knob position, as is performed by the driving unit. The control algorithm will be focused on keeping the FOE near the darkest region by correcting the tip direction when needed. This approach differs from other approaches because it prevents correctional steering in the situation where the endoscope is heading towards the lumen as a whole but the tip is pointing in another direction (Fig. 3, right).

The following three subsections accomplish the first research objective which was to explore the feasibility of using the FOE in a navigational model for flexible endoscopy. Firstly, Section II.A describes the dark region detection in more detail. Section II.B then describes the calculation of the OF field and finally, in Section II.C, FOE estimation is discussed.

A. Dark Region Detection

Finding the lumen in an image is performed using DRD. The dark region is found recursively by searching the minimum of the low-pass-filtered intensity (gray value) image in an area around the previously found location. Before the algorithm can start, the lumen is indicated manually in the first image.

B. Calculating Optical Flow

Calculation of OF is commonly based on a block motion model. The common assumptions are that a) the OF within a small block (window) in the image plane is constant; and that b) the pixels within corresponding blocks in consecutive images do not change. The latter is often referred to as the constant brightness assumption. These assumptions are used in, for instance, the Lucas-Kanade tracker [16], but do not hold well in human endoscopy images. The Horn-Schunck method is based on the assumption that the OF vectors only change smoothly along the image plane [17]. This assumption also does not hold in human colon images, especially not in the vicinity of the FOE. Therefore, a different method of OF calculation is needed. Another

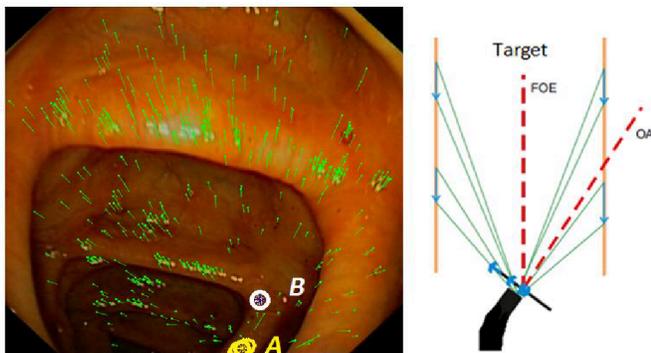


Fig. 3. Example of calculated FOE (A) and manual FOE (B) within OF field in real image (left). On the right, a schematic figure of OF arrow origin (arrows) and FOE direction vs. target and tip direction (OA) is depicted. Note that the target direction is not always along the optical axis, which makes steering superfluous in this example case.

problem comes in when trying to perform adequate feature detection in these images. Lack of texture and specular reflections cause problems detecting enough features with high reliability. As a result, the risk of mismatches is high.

We have searched for an existing technique that can detect blob-like features and at the same time has some kind of match verification between those features. The scale invariant feature transform (SIFT) point feature detector [18] fulfills both requirements. A set of candidate 2D image points to track is selected in each of two consecutive images. Matching is based on feature vectors that are extracted from the neighborhood of the detected points. The features of the current frame are matched to those of the previous frame and a displacement vector between them can then be determined. These vectors can be regarded as alternative OF vectors, and together they form the OF field.

C. Estimating the Focus of Expansion

Estimation of the FOE, denoted by \mathbf{q} , is achieved by means of weighed Least Square Estimation (LSE). For that, we need a linear observation model of each pair of matched points:

$$\mathbf{z} = \mathbf{M}\mathbf{q} + \mathbf{n}$$

(1)

where \mathbf{z} is an observation, \mathbf{M} is the measurement matrix, and \mathbf{n} is the measurement noise. Such a linear model is based on the fact that the FOE is found on the line that connects the two matching points. This is true in the case of a translational motion and a static scene. Thus, if \mathbf{p}_1 is a point found in the first image, and \mathbf{p}_2 is the corresponding matching point in the second image, then \mathbf{q} , \mathbf{p}_1 and \mathbf{p}_2 are collinear. If $\mathbf{q} = [x_q \ y_q]^T$ then:

$$z = ax_q + by_q \quad \text{for} \quad a^2 + b^2 = 1.$$

(2)

where z , a , and b follow from \mathbf{p}_1 and \mathbf{p}_2 using their collinearity.

The definition of $\mathbf{M} = [a \ b]$ then results in eq. (1). The noise \mathbf{n} originates from the uncertainty of the point locations. The errors in the locations propagate through eq. (2) to \mathbf{n} . Eq. (1) provides us with the model for a single pair of matching points, but in reality there are N pairs of matching points leading to N equations resembling eq. (1). These equations can be stacked to yield:

$$\mathbf{z} = \mathbf{H}\mathbf{q} + \mathbf{n}$$

(3)

where \mathbf{H} is the N times stacked version of \mathbf{M} in eq (1). The noise vector \mathbf{n} is quantified by its covariance matrix \mathbf{C}_n . The matrix will be diagonal, assuming that errors in the point locations are independent. From the geometry of the problem (radial expansion originating in \mathbf{q}) we assume that the standard deviation of \mathbf{n} is inversely proportional to the distance d_n between \mathbf{p}_1 and \mathbf{p}_2 , and proportional to the

distance D_n from the key point to \mathbf{q} . The covariance matrix then is described by:

$$\mathbf{C}_n = \text{diag} \left(\sigma_n^2 \frac{D_n^2}{d_n^2} \right). \quad (4)$$

If we assume equal uncertainties for all measurements, and no prior knowledge on \mathbf{q} , the Least Square Estimator (LSE, [19]) $\hat{\mathbf{q}}_{LSE}$ is:

$$\hat{\mathbf{q}}_{LSE} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{z}. \quad (5)$$

If the uncertainties are unequal, according to eq (4), the maximum likelihood estimator (MLE) $\hat{\mathbf{q}}_{MLE}$ becomes:

$$\hat{\mathbf{q}}_{MLE} = (\mathbf{H}^T \mathbf{C}_n^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}_n^{-1} \mathbf{z}, \quad (6)$$

which can be regarded as an LSE in which each measurement is weighed in inverse proportion to its uncertainty. Note that the weighed LSE estimator cannot be used directly since knowledge of \mathbf{q} is required to calculate D_n . This leads to an iterative approach in which first an estimate is obtained using eq. (5), which then is iteratively used in eq (4) to obtain an estimate that fits the requirements using eq. (6). In practice, a single pass of eq. (5), followed by eqs. (4) and (6) suffices. In the next Section, the methods used for testing the FOE and OF field calculations are described.

III. EXPERIMENTAL SET-UP

The second research objective revolves around the question: ‘‘Is calculation of the OF and a corresponding FOE in human colonoscopy images feasible?’’. We designed the experiment described below to find an answer.

A. Image Acquisition

The images used were acquired during human colonoscopy procedures. They were produced by a colonoscope camera (types CF Q 165 L and 160/180 AL of Olympus, Tokyo, Japan) and captured using an ADV55 video capture device (GrassValley, San Francisco, CA, USA). Storage of images was on a laptop. From the stored movie files, a clip was selected from which completely red frames (wall views) were absent. This series was 76 frames in length.

B. Processing

To obtain as much information from the images as possible, preprocessing was performed. The main goal here was to find as many 2D image points to track (key points) as possible. The large black area that surrounded the image of interest, was excluded from processing by using a similarly shaped mask. The subsequent three main steps that were taken are described in the following subsections. In the flow charts (Fig. 4., Fig. 5., and Fig. 6.), each preceding step is represented as a single block.

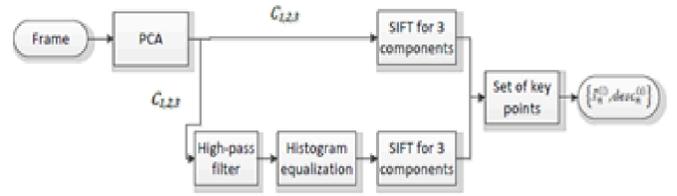


Fig. 4. Flow chart describing steps in SIFT feature extraction.

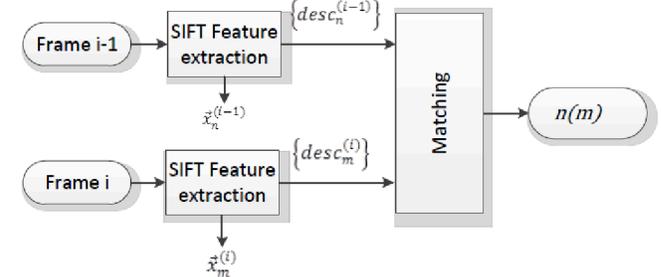


Fig. 5. The set of SIFT key points of two consecutive frames is matched.

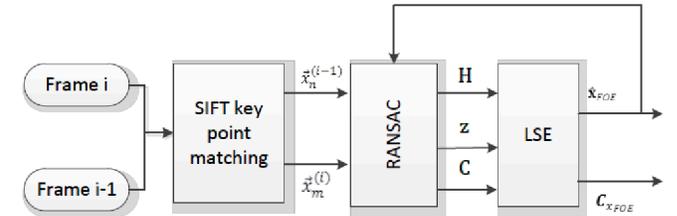


Fig. 6. Final estimation using RANSAC and iteration.

1) SIFT Key Point Extraction

Two sequential color images were read. On both images, a Principal Components Analysis (PCA) [20] was performed ($C_{1,2,3}$) on the three color channels and SIFT key points were obtained. This PCA was performed, because previous (unpublished) tests showed that more key points were found when information from the three channels was obtained separately. In addition, on $C_{1,2,3}$ feature enhancement in the form of high-pass filtering followed by histogram equalization was applied to improve the texture even further. From the resulting three preprocessed components, the SIFT key points were also obtained (Fig. 4). All the key points found were pooled per image and the resulting two sets were used to find matches.

2) SIFT Key Point Matching

The two sets of the sequential images frame i and frame $i-1$ were compared. A match was sought using the similarity metric of the SIFT algorithm [18] (Fig. 5). The features are named \vec{x}_{number}^{frame} and their properties are described by descriptor $\{desc_{number}^{frame}\}$. This descriptor is used in the similarity metric and thus to match the key points from two different frames.

3) Estimation

To estimate the FOE robustly, the Random Sample Consensus (RANSAC) [21] algorithm was used (Fig. 6). This iterative algorithm estimates the FOE and accepts the estimation if the requirements are met. Data obtained from RANSAC are the values for \mathbf{H} , \mathbf{z} , and the covariance matrix \mathbf{C} of \mathbf{H} . The covariance matrix was used as a measure for the accuracy of a and b and for calculation of the MLE.

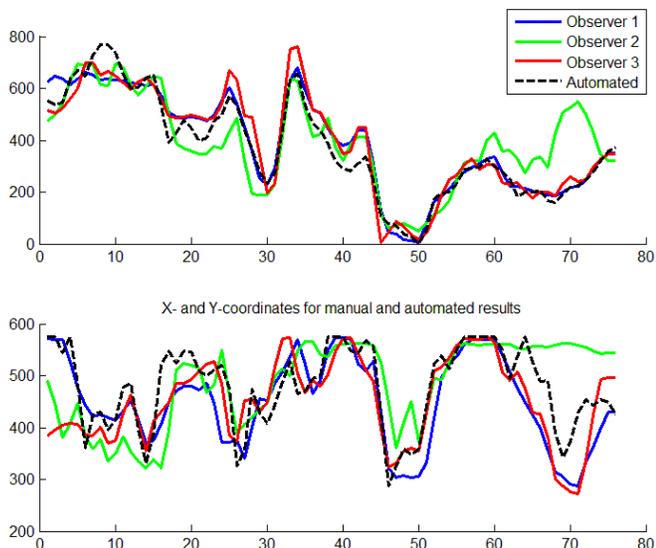


Fig. 7. Results of manually indicated FOE by the two observers of group A and automated FOE detection plotted in one figure. The y-axis represents the pixel value of the coordinate.

A. Determining Accuracy

The automatically determined FOE location in each frame was compared to manually determined FOE locations by calculating the intraclass correlation (ICC) between the observers [22]. Independently of each other volunteers indicated the point where they thought the FOE was by clicking it in a window. This window iteratively showed two alternating sequential frames until a location was chosen by the volunteer.

IV. EXPERIMENTAL RESULTS

A. Feature Detector Performance

The SIFT feature extraction algorithm was implemented according to [18]. Localization and matching of the key points was computationally intensive, and therefore analysis could not be done in real-time. Both the RANSAC and the estimation algorithm worked as expected. One of the parameters that could be set was the minimum number of inliers above which the estimation would be accepted. If RANSAC could not find a result within the acceptance threshold of this parameter, the set size was decreased. Again, this was done iteratively but only down to a minimum of 20% of the total set. If no result was found at this minimum threshold, then the frame was ignored. However, this situation did not occur during our experiments.

On average, 507 SIFT key points were matched per frame. The average number of inliers that was included by the RANSAC algorithm was 257 (50.7%). The frame with the least found key points only showed 49 displacement vectors, 25 of which were inliers. This frame appeared to possess a lot of motion blur. The computational time for one frame was approximately 60 seconds.

B. FOE Estimation Performance

The FOE was indicated manually by three observers.

When the manual information from the human observers is tested for consistency, the ICC is 88%. The ICC for the observers and the automated algorithm was 89% (Fig. 7).

V. DISCUSSION

SIFT key point detection is known to be a successful method of obtaining features from images that contain a lot of noise [18], and is better than, for instance, the Shi-Tomasi feature detector in finding blob-like features [23]. Human endoscopy images are not only inherently noisy, but also are lacking texture. Our preprocessing, consisting of PCA and a method resembling local histogram equalization, enhances image texture. This resulted in sufficient key points in each image with that had a high enough matching rate. However, SIFT key point detection is also known for its computational cost. Unfortunately, in this research this property of SIFT also surfaced. The algorithm is currently too computationally intensive. It was found that most of the computational time is spent in detecting and matching SIFT key points. For these operations, real-time implementations exist that use field programmable gate array (FPGA) or graphics processing unit (GPU) programming. In order to develop these, more detailed knowledge about the steering algorithm first needs to be acquired. Another, easier, approach could be to explore alternative feature detectors such as SURF [24] or FAST [25].

The automated method of FOE determination has a good ICC compared to the manual FOE determination (89%). This result indicates the high reliability of our automated FOE detection algorithm. However, we observed large differences between human observers who have a theoretical background in this area and those who do not. These differences make the ground truth less reliable, but it currently remains the best option there is.

Furthermore, results are given for a carefully selected image sequence. In human endoscopy, the tip touches the bowel wall often, which did not happen once in our sequence. When there is wall contact, the produced images appear red and blurred, and therefore key points will be hard to find. We need to either filter those images out in the final, robust, steering algorithm, or prevent contacts from occurring.

Finally, the FOE is known to be suitable for determining heading direction in the translational direction [26]. During the experiments we also found a rotational component present in the camera movements of an endoscope. Nonetheless, rotational information can theoretically also be obtained from the images, as in the example described in [26]. ‘De-rotating’ the image before FOE determination then becomes feasible and this in turn will make the final steering algorithm more accurate.

Several products that could facilitate the colonoscopy procedure already exist. Most of them rely on some kind of inflatable balloon propulsion mechanism. These include the ColonoSight (Stryker Endoscopy, San Jose, CA, USA), its successor the ProtectiScope, the Aeroscope (GI-view, Ramat

Gan, Israel) and the Endo-Ease system (Spirus Medical Inc., Stoughton, MA, USA). Caterpillar-like propulsion is applied in the Endotics system (Era Endoscopy S.r.l., Peccioli, Pi, Italy), and inverted sleeve propulsion can be found in the Invendoscope (Invendo Medical, Kissing, Germany). Although these propulsion techniques are innovative, the steering mechanisms are still mostly based on steering knobs on a hand grip that are used to control the tip. Our steering algorithm could probably benefit these products. Additionally, the algorithm is generic and could be applied for endoscopy of other organs where navigation remains challenging. These include but are not limited to lungs, gall ducts and salivary glands.

VI. CONCLUSIONS AND FUTURE WORK

This study has shown that the use of the FOE for flexible endoscope navigation is feasible. The FOE can be used as a way of determining the heading direction of the endoscope mathematically, which can then be compared to a desired target direction in a steering algorithm. If these directions match, no steering movement is initiated, which keeps the number of movements to a minimum. This is the most important advantage of our approach: no unnecessary steering corrections.

Currently, it remains unknown what the influence is of the rotational component in the camera movements of the endoscope on our method. This will therefore be the subject of future work over the short term. Also, we will look into real-time improvement. Additionally, complete colonoscopy image sequences will be tested. Such sequences will consist of more images with little texture and more motion blur artifacts.

Future work will also need to determine feasibility in clinical practice. Image sequences of multiple patients will be evaluated in order to further improve the method. Over a longer term, motion estimation and the FOE will be included in a steering algorithm.

REFERENCES

- [1] National Cancer Institute, (2011, Nov 30) "Colorectal cancer," [Online]. Available: <http://www.nlm.nih.gov/medlineplus/colorectalcancer.html>.
- [2] H. van Veldhuizen-Eshuis, M.E.M. Carpay, J.A. van Delden, L. Grievink, B. Hoebee, A.J.J. Lock, and R. Reij, "Uitvoeringstoets bevolkingsonderzoek naar darmkanker," Min VWS, Den Haag, 2011.
- [3] S.-H. Lee et al., "An adequate level of training for technical competence in screening and diagnostic colonoscopy: a prospective multicenter evaluation of the learning curve," *Gastrointestinal Endoscopy*, vol. 67, no. 4, pp. 683-689, 2008.
- [4] B. J. Spier, M. Benson, P. R. Pfau, G. Nelligan, M. R. Lucey, and E. a Gaumnitz, "Colonoscopy training in gastroenterology fellowships: determining competence.," *Gastrointestinal endoscopy*, vol. 71, no. 2, pp. 319-24, Mar. 2010.
- [5] P. S. Tassios, S. D. Ladas, I. Grammenos, K. Demertzis, and S. A. Raptis, "Acquisition of Competence in Colonoscopy: The Learning Curve of Trainees," *Endoscopy*, New York, vol. 31, no. 9, pp. 702-706, 1999.
- [6] J. D. Wayne, D. K. Rex, and C. B. Williams, *Colonoscopy: Principles and Practice*, 2nd ed. Chichester: Blackwell Publishing Ltd, 2009, pp. 267-345.
- [7] M. Liedlgruber and A. Uhl, "Endoscopic Image Processing - An Overview," in *Proc. 6th symp. on Image and Signal Processing and Analysis*, 2009, pp. 707-712.
- [8] N. Kuperij, R. Reilink, M. P. Schwartz, S. Stramigioli, S. Misra, and I. A. M. J. Broeders, "Design of a User Interface for Intuitive Colonoscopy Control," in *IEEE/RSJ Int'l Conf on Intelligent Robots and Systems(IROS)*, San Francisco, USA, 2011, pp. 937-942.
- [9] D. Gillies and G. Khan, "Vision based navigation system for an endoscope," *Image and Vision Computing*, vol. 14, pp. 763-772, 1996.
- [10] R. Reilink, S. Stramigioli, and S. Misra, "Image-Based Flexible Endoscope Steering," in *IEEE/RSJ Int'l Conference on Intelligent Robots and Systems(IROS)*, Taipei, Taiwan, 2010, pp. 2076-2082.
- [11] H. Chettaoui, G. Thomann, C. B. Amar, and T. Redarce, "Extracting and tracking Colon 's ' Pattern ' from Colonoscopic Images," in *IEEE Canadian Conference on Computer and Robot Vision*, 2006, pp. 65-71.
- [12] N. Masson, F. Nageotte, P. Zanne, and M. D. Mathelin, "In vivo comparison of real-time tracking algorithms for interventional flexible endoscopy," in *Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Boston, MA, USA, pp. 1350-1353.
- [13] C. S. Tan, "Computerized detection of lumen boundary in robotic colonoscopy for medical diagnosis," Ph.D. Dissertation, Nanyang Technological University, Singapore, 1995.
- [14] K. Mori, D. Deguchi, J. Ichi Hasegawa, and Y. Suenaga, "A Method for Tracking the Camera Motion of Real Endoscope by Epipolar Geometry Analysis and Virtual Endoscopy System," in *2001 International Conference on Medical Image Computing and Computer-Assisted Intervention*, Utrecht, The Netherlands, pp. 1-8.
- [15] W. Burger and B. Bhanu, "On Computing a 'Fuzzy' Focus of Expansion for Autonomous Navigation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 1989, no. 3, pp. 563-568.
- [16] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 1981, pp. 674-679.
- [17] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow," *Artificial Intelligence*, vol. 17, pp. 185-203, 1981.
- [18] D. G. Lowe, "Object Recognition from Local Scale-Invariant Features," in *International Conference on Computer Vision*, 1999, pp. 1-8.
- [19] F. van der Heijden, R. P. W. Duin, D. D. Ridder, and D. M. J. Tax, *Classification, Parameter Estimation and State Estimation*. Chichester: John Wiley & Sons Ltd, 2004, pp. 45-79.
- [20] I. T. Jolliffe, *Principal Component Analysis, Second Edition*, 2nd ed. New York: Springer-Verlag New York, Inc., 2002, pp. 1-61.
- [21] M. Fischler and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with applications to image analysis and automated cartography," *Comm. of the ACM*, Vol. 24, pp. 381-395, 1980.
- [22] E. M. Jellinek, "On the use of the intra-class correlation coefficient in the testing of the difference of certain variance ratios," *Journal of Educational Psychology*, vol. 31, no. 1, pp. 60-63, 1940.
- [23] J. Shi and C. Tomasi, "Good features to track," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 1994, Seattle, WA, USA, pp. 593-600.
- [24] H. Bay, a Ess, T. Tuytelaars, and L. Vangool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346-359, Jun. 2008.
- [25] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," *10th IEEE International Conference on Computer Vision*, 2005, Beijing, China, Vol. 2., pp. 1508-1515.
- [26] H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," in *Proceedings of the Royal Society London*, 1980, pp. 385-397.