A feasibility study of optical flow-based navigation during colonoscopy

N. van der Stap, R. Reilink, S. Misra, I.A.M.J. Broeders, F. van der Heijden

Purpose

Colonoscopy is performed in the screening for colorectal cancer to obtain a diagnosis and to perform small interventions in the large bowel. The manual control of flexible endoscopes is impractical and not very intuitive. The current research aims at computer assisted colonoscopic navigation in the human bowel. An interesting option then is to use the camera as a motion sensor. This allows using existing endoscopes without a need for additional sensors in the tip. A possibility for visual navigation is optical flow (OF), i.e. the apparent motion of objects in the image plane. OF is useful since it encodes the heading direction of the tip by means of the so-called focus of expansion (FOE). Additionally, the relative depth of possible fiducial points inside the bowel can be reconstructed using OF. As such, OF provides a clue for finding the target direction. However, calculation of the OF is not straightforward if the scene lacks texture. This paper presents a method to calculate the OF during colonoscopy. The feasibility of the method is demonstrated by estimating the FOE and comparing this with manually annotated FOEs.

Methods

Theory: At least three parameters are needed to steer an endoscope: the target direction (from the tip towards the lumen), the actual heading direction, and the direction towards which the tip is pointing (as defined by the optical axis OA). These are directions in the three-dimensional space. Provided that the camera has been calibrated, they can be represented as two-dimensional positions in the image plane. The FOE is found as the point from where all optical flow vectors emerge (Figure 1). Provided that rotational motions of the tip are not too severe, the FOE corresponds to the heading direction.

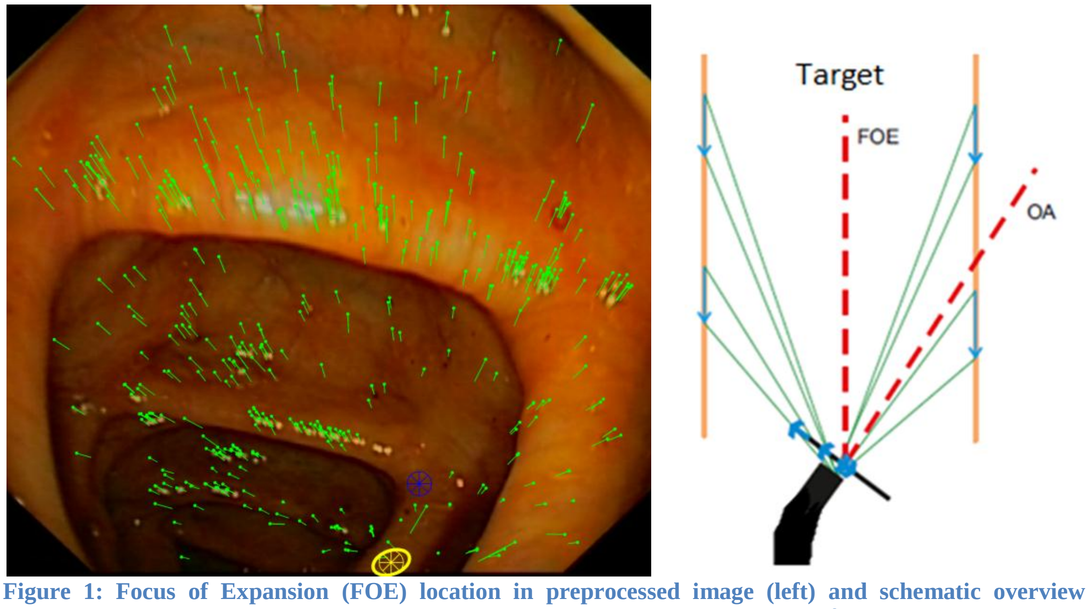

Figure 1: Focus of Expansion (FOE) location in preprocessed image (left) and schematic overview of heading direction, target direction (Target) and tip direction (Optical Axis, OA). *Left: the FOE (yellow: automatic; blue: manual) as the origin of the green OF vectors. Right: the bent colonoscope within the bowel tube, with the image plane before it. The two relevant directions are pictured in red. Note the blue optical flow arrows in the image plane and their origin on the bowel wall.*

OF estimation: The bowel lacks textured surfaces. Therefore, the usual method for OF calculation, e.g. Lucas-Kanade, is difficult to use. Instead, we use the SIFT point feature detecting and matching algorithm to find corresponding pixels in consecutive frames (Figure 2). OF vectors are found as the displacement between pairs of corresponding points. To maximize the number of detected points, preprocessing was applied. It consists of i) principal component feature extraction applied to the RGB color space, and ii) nonlinear image filtering made up by a linear high pass filter followed by a global histogram equalization. Such a filter approximates local histogram equalization, and enhances the textures of surfaces.

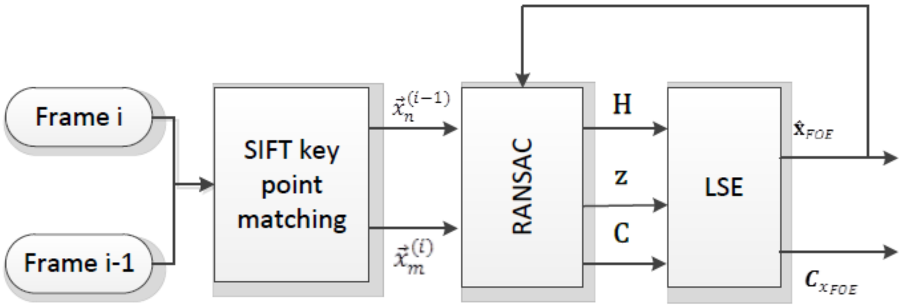

Figure 2: Flowchart describing algorithm steps. *Two consecutive frames are loaded, preprocessed and SIFT point features are extracted. Point features are matched and the FOE is estimated using the geometrical model, the LSE estimation and the RANSAC algorithm.*

FOE estimation: The FOE, $\mathbf{x}_{FOE} = \begin{bmatrix} x_{FOE} & y_{FOE} \end{bmatrix}^T$, is located at the intersection of N straight lines which are created by connecting the corresponding points. The estimation of the FOE is done by means of weighed Least Square Estimation (LSE). A linear observation model of each pair of corresponding points is:

$$\mathbf{z} = \mathbf{H}\mathbf{x}_{FOE} + \mathbf{n} \tag{1}$$

The vector $\mathbf{z}$ and the matrix $\mathbf{H}$ follow from the N line equations $h_{n,1}x_{FOE} + h_{n,2}y_{FOE} = z_n$ with $h_{n,1}^2 + h_{n,2}^2 = 1$. The vector $\mathbf{n}$ represents the uncertainty. From the geometry of the problem we assume that the uncertainty in each equation is inversely proportional to the distance d_n between corresponding points, and proportional to the distance D_n from the FOE to the nearest point feature. The covariance matrix of the noise is:

$$\mathbf{C_n} = diag\left(\sigma^2 \frac{D_n^2}{d_n^2}\right) \quad \text{with } diag(a_n) \text{ a diagonal matrix with elements } a_n \tag{2}$$

LSE estimation performs feebly in the presence of outliers. Since the SIFT algorithm frequently produces false matches, we apply outlier suppression using the RANSAC algorithm.

Evaluation: Movies that are used in this research are produced by an Olympus colonoscope camera (types CF Q 165 L and 160/180 AL) during human colonoscopy. From these movies, a clip of 75 frames was selected showing the insertion process of the tip. The optical flow is determined between each two frames. From that, the FOE is estimated. The results are compared to FOE results of the sequence obtained by manual annotation. The agreement of the FOE sequences is expressed in the intraclass correlation coefficient.

Results

The two manually annotated FOE's have an intraclass correlation of 97%, whereas the three sequences (the automatically estimated one included) have an intraclass correlation of 96% (Figure 3). The average number of found OF vectors per frame is 507. The average number of inliers, as determined by RANSAC, is 257. The most critical frame in the sequence, having the minimum number of 49 OF vectors, is heavily deteriorated by motion blur. The corresponding minimal number of inliers is 25; still enough to estimate the FOE with a reasonable accuracy. In the current implementation, the time for processing one frame is around 60 s.

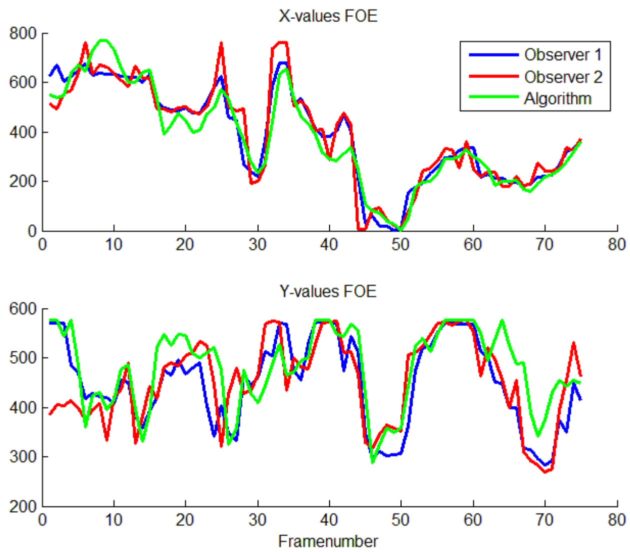

Figure 3: The X- and Y-values of the three observers.

Conclusion

This study has shown that the statistical agreement between the results of human observers is similar tothe agreement between observers when automated calculation of the FOE is included. This agreement is excellent. It demonstrates that OF and FOE calculation for flexible endoscope navigation is feasible. We think that the FOE is an important clue for computer assisted steering as it is strongly connected to the heading direction. In future work, the neutralization of the influence of rotations of the camera will be researched. Future work also includes determining the feasibility in clinical practice. Anatomical models can be used to validate the steering method, followed by animal and human subject testing. The current implementation is too slow for real-time performance. However, most of the computational time is spent in detecting the SIFT point features and the matching. For both algorithms real-time implementations exist using FPGA or GPU architectures.