Paper Review

Paper: J. Liu, K. Subramanian, T. Yoo, R.Van Uitert, "A Stable Optic-Flow Based Method for Tracking Colonoscopy Images," CVPRW, 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.1–8, June 2008.

Background and Relevance to Project:

This paper describes a new algorithm developed by Liu et al. for determining the position of "an endoscopic camera (during a colonoscopy procedure)" [1] using optic-flow and other optimization techniques. Using this technique, it is possible to integrate Virtual Colonoscopy (VC, where a 3D reconstruction of the colon is generated from CT scans) with Optical Colonoscopy (OC, where an endoscope is used to view the colon). This way, clinicians can navigate the colon to perform operations on regions of interest they have identified in the virtual model.

In our project, we aim to create a 3D reconstruction of the cochlear canal. Since we are using the EyeRobot to drive our endoscope, assuming our endoscopic probe is rigid, we already know the position and the orientation of our camera. However, we need to know the scale factor between the camera and world coordinates in order to be able to generate an accurate 3D model. The method described in this paper determines this scale factor and then calculates the position and the orientation of the endoscopic camera. We believe that we can use this algorithm to extract the necessary information for our calculations.

Algorithm Summary

The algorithm has three main steps for finding the camera position. First, an anisotropic Gaussian filter is applied to the image in order "to account for the different sampling rates across the spatial and temporal dimensions" [1]. Then, by applying the following Harris matrix (Equation 4),

$$\mu = \begin{bmatrix} L_x^2 & L_x L_y \\ L_x L_y & L_y^2 \end{bmatrix}$$

corners in the image are detected, to be used as features for tracking. The spatial and temporal scales are determined by minimizing a scale-space metric, Θ (Equation 6), where the numerator is a measure of how (dis)similar two corresponding points in two consecutive frames are, and the denominator is a measure of "how distinct the selected features are in their local neighborhood" [1]. Using Taylor series approximation, the numerator takes the form of the Lucas-Kenade algorithm (see Equation 7). 'Good' tracking features are chosen from the detected corners by applying a threshold to their 'distinctness'.

Next, using these scales, the optic-flow field of the image is computed using Horn's method, where the optical flow vector is calculated by minimizing the following energy function:

$$E = \iint \left[(I_x u + I_y v + I_t)^2 + \alpha^2 (|\nabla u|^2 + |\nabla v|^2) \right] \mathrm{d}x \mathrm{d}y$$ [2]

The authors claim that estimating the motion parameters for the camera using a linear system (Equation 10) is impractical since "there are errors in the optical flow estimation," and also that this system is computationally costly (a 6x6 linear system). To overcome this, the rotational parameters will be removed from the system, resulting in a 3x3 system that is less sensitive to errors. To be able to determine the rotational parameters, the focus of expansion (FOE) needs to be found. The focus of expansion is computed by making use of the theory that a line passing through two corresponding points in two subsequent frames will also pass through the focus of expansion.

The rotation and translation parameters for the camera motion are computed using the focus of expansion. In order to calculate the rotation parameters, the flow vectors for the 'good' tracking features found in the first step are "transformed into polar coordinates with the FOE at the origin" [1], eliminating the translation component. The translation parameters can then be computed by subtracting the rotation parameters from the optical flow equations. To reduce error, the median of multiple translation parameters is chosen, and outliers are removed via thresholding. A diagram of this algorithm can be seen in Figure 1.

Evaluation

The authors have tested their algorithm using both a colon phantom image sequence and a virtual colonoscopy image sequence. They compared the performance of their algorithm to that of Bruss and Horn's passive navigation method (considered to be one of the more superior methods), and found that the proposed method is more accurate and stable. This means that the method described in this paper is more suitable for use in our project.

Critique

The authors seem to have successfully improved upon the conventional optical flow methods to create a more robust algorithm for motion parameter estimation of an endoscopic camera. They have compared their method to a mainstream counterpart, and have demonstrated improvements. They were usually able to clearly explain how they used different theorems and equations, however there were also a few points which lacked details and are ambiguous.

The authors usually neglected to specify the thresholds used for eliminating outliers and for choosing good tracking features. These thresholds are usually determined empirically, however, they could have provided some guidelines for choosing these thresholds. A similar lack of providing values is seen when the authors talk about constants α and β in the scale-space metric Θ in Equations 6 and 7. We also see in these equations a window function G; the authors don't specify the type of this function either, but I assume that this is a Gaussian function since they talk about the Anisotropic Gaussian in Equation 3.

Another lacking point is seen in Figure 7. The authors have neglected to explain how they have registered the phantom images, CT scans, and the endoscopic camera, or how they generated these graphics.

Another crucial analysis that the authors have forgot to perform is how the computational speed of the algorithm compares to other methods'. In order for this algorithm to be feasible, it has to run in real-time.

Implications for Our Project

The goal of our project is to build a 3D reconstruction of the cochlear canal and generate a safe insertion path using a borescope. The method described in this paper can be reverse-engineered in order to get the scale factors, which can be then used to manipulate the image to form a 3D circular shape.

Our borescope currently has a non-rigid probe, and this poses a problem in determining the position of the tip of the probe. If our current plans for making the probe rigid do not work, this method can be used to determine the position of the tip.

References

[1] Jianfei Liu, Kalpathi Subramanian, Terry Yoo, Robert Van Uitert, "A Stable Optic-Flow Based Method for Tracking Colonoscopy Images," CVPRW, 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp.1–8, June 2008.

[2] Horn-Schunck Method. Wikipedia. <http://en.wikipedia.org/wiki/Horn%E2%80%93Schunck_method>. Accessed 19 May 2011.

[3] A. R. Bruss and B. K. P. Horn. Passive Navigation. Computer Vision, Graphics and Image Processing, 21:3–20, 1983.