34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Measuring the Size of Neoplasia in Colonoscopy using Depth-from-Defocus

François Chadebecq
Institut Pascal, ISIT
Clermont-Ferrand, France

Christophe Tilmant
Institut Pascal
Clermont-Ferrand, France

Adrien Bartoli
ISIT
Clermont-Ferrand, France

Abstract— **Colonoscopy is the reference medical examination for the diagnosis and treatment of neoplasia in gastroenterology. During the examination, the expert explores the colon cavity with a gastroscope in order to detect neoplasias - abnormal growths of tissue - and to diagnose which ones could be malignant. The *Paris classification of superficial neoplastic lesions* is the gold standard set of criteria for this type of diagnosis. One of the major criteria is the size. However, this is tremendously difficult to accurately estimate from images. This is because the absolute scale of the observed tissues is not directly conveyed in the 2D endoscopic image.**

We propose an image-based method to estimate the size of neoplasias. The core idea is to combine Depth-From-Focus (DFF) and Depth-From-Defocus (DFD). This allows us to recover the absolute scale by automatically detecting the blur/unblur breakpoint while the expert pulls the gastroscope away from a neoplasia. Our method is passive: it uses the image data only and thus does not require hardware modification of the gastroscope. We report promising experimental results on phantom and patient datasets.

I. Introduction

Colorectal cancer is the fourth cause of death by cancer according to the World Health Organization. Most colorectal cancer cases however are preventable. Massive screening campaigns facilitate preventive surveillance. The improvement of medical techniques facilitates early diagnosis and treatment of colorectal diseases. The diagnosis of neoplasia (an abnormal growth of tissue) relies on a gold standard defined in the *Paris classification of superficial neoplastic lesions* [12]. Several criteria such as texture and shape allow one to objectively evaluate a neoplasia's potential malignant decline. Amongst these criteria, a neoplasia's size (area and volume) is particularly important to the diagnosis. It directly influences the decision of resection.

Non-invasive diagnosis examinations can be classified into two categories: CT-colonography (also called 'virtual colonoscopy') and capsule endoscopy. CT-colonography is mainly used as a preliminary diagnosis examination. It is based on scanning. It allows one to browse a 3D reconstruction of the colon cavity and thus to measure the size of the neoplastic lesions. The capsule allows for the acquisition of a video sequence inside the colon cavity but is not controllable and requires time-consuming work of analysis in order to ensure that no lesion is overlooked.

Colonoscopy is a minimally invasive interventional technique. It is now the reference medical examination for diagnosis and treatment of colon diseases [6]. Its main drawback is the loss of the depth information preventing one from directly estimating the size of neoplasias. Most passive 3D reconstruction techniques suffer from a scale ambiguity [4]. Active techniques such as electromagnetic tracking or structured lighting resolve this ambiguity but require modifying the gastroscope's hardware.

We propose to estimate the size of neoplasias using Depth-From-Focus (DFF) and Depth-From-Defocus (DFD). These passive methods estimate the amount of blur in images to infer the scene depth. Assuming that a neoplasia is frontoparallel to the gastroscope's tip we can further infer its area. Our experimental results on a phantom model and on patient datasets show promising results. The order of accuracy matches the requirement of gastroenterological experts for the diagnosis of neoplastic lesions. Section II presents a state of the art on DFF and DFD. Section III gives the optical modeling of the colonoscope and defines a model for blur estimation. Section IV proposes a clinical protocol to apply our method. Finally, section V presents the results obtained with the new proposed method for estimating the size of neoplastic lesions.

II. State of the Art

DFF, also called 'software focus', consists in taking a sequence of images by controlling camera parameters such as the depth of field. The goal is to define the best focused points in each image [9]. The main drawback of DFF is the requirement of a large set of images and an accurate control of camera parameters. DFD techniques were introduced in [10] with the aim of estimating a reliable depth map. They are based on the estimation of the amount of defocus, generally between two images of the same scene taken at a distance d with different camera parameters. Both require an accurate modeling of the blur process.

We now review blur estimation which is a keypoint of DFD and DFF methods. Blur estimation approaches can be classified in three categories: frequency domain, spatial domain and statistical approaches. All of these require a calibration stage in order to define an absolute depth map. Frequency domain approaches use the Fourier transform

This work has been supported by the French National Research Agency (ANR) through the TecSan program. The authors thank Yansys Medical®, as well as Professor Thierry Ponchon and Doctor Vincent Lepilliez from the gastroenterology department of Edouard Herriot Hospital.

and require at least two images [10]. These methods are global and do not allow for the computation of an accurate depth map. Various space-frequency approaches have been proposed to overcome this problem [8]. However, these techniques trade-off between frequency and spatial accuracy. Early spatial domain approaches rely on the study of sharp edges to compute local degrees of defocusing [10]. A spatial convolution and deconvolution transform (STM) [13] can be used for autofocusing. More general DFD by diffusion methods based on the well-known heat diffusion equations also exists [3]. This approach takes into account the spatially varying nature of blur. A global DFD technique based on these approaches, for a camera whose internal camera parameters are fixed, exists [14]. Statistical approaches use Markov Random Fields for modeling blur in images [11]. Both of these blur estimation approaches rely on strong assumptions (presence of edge, Lambertian reflectance of surfaces) and are sensitive to changing illumination conditions.

Colonoscopy images present specific difficulties due to the nature of the explored environment (such as the complex specular reflectance of biological tissues [5]) and the characteristics of the mono-focal optical system. In our work, we consider a video in order to robustly estimate the depth of a Region of Interest (ROI) by combining DFF and DFD. Unlike DFD, our approach does not requires one to actively control the parameters of the camera displacement [15]. We propose instead to estimate blur at each frame of the input video. In our current implementation, the ROI correspondence is manually established between the frames. We validated our approach experimentally and obtained promising results.

III. Single Depth Estimation

A. Geometrical optics modeling (GOM)

Consider a point p at a distance d from a thin lens with focal length f that models the optical system of the camera. The in-focus image point must be at a distance s such that:

$$\frac{1}{d} + \frac{1}{s} = \frac{1}{f} \tag{1}$$

If the image sensor is not at the same location as the in-focus plane but at a distance e of the lens, the image of a point becomes a spot, called circle of confusion (figure 1). The radius R of this spot is given by the law of similar triangles (2) and depends on the radius r of the lens:

$$R(d) = r\left|\frac{e}{f} - \frac{e}{d} - 1\right| \tag{2}$$

B. Physical optics modeling (POM)

Considering the depth of field of a colonoscope and the diameter of the colon cavity, we must use a physical modeling to explain the effect of diffraction. In other words, the impulsional response of the optical system called 'Point Spread Function' (PSF) has to be carefully taken into account. However, physically modeling a camera's optical system is a hard task because of the various components which are required to acquire an image. Diffraction effects due to the aperture's shape, the lenses as well as sensor integration and sampling require an accurate calibration process. Advanced PSF modeling showed that the generalized Gaussian is a suitable model [2]. It encompasses the classical pillbox and 2D Gaussian models. The generalized Gaussian is particularly efficient at modeling the Edge Spread Function (ESF; extention of the PSF to edges). However, a fine regularized numerical differentiation is required. It has been experimentally observed [10] that, considering a polychromatic light, the PSF could be approximated by a 2D isotropic Gaussian. Our approach relies on this standard assumption generally accepted in DFD and DFF. An observed ROI I_d is expressed as the convolution of the corresponding focused ROI I_{if} with the 2D isotropic Gaussian PSF g_{σ_s} with variance σ_s^2 as:

$$I_d(x,y) = I_{if} * g_{\sigma_s}(x,y) \tag{3}$$

where σ_s is proportional to the distance from the camera to the observed ROI [10]. In order to combine both geometrical and physical modeling of the PSF, a simple model based only on the thin lens equation is considered and equation (2) can be rewritten in the spatial domain as:

$$\sigma_s(d) \propto R(d) + \mu d \tag{4}$$

Parameter μd represents the width of the PSF profil.

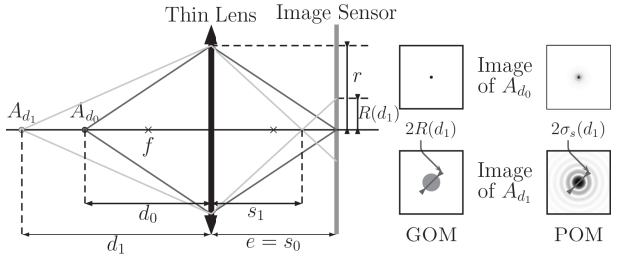

Fig. 1. Image formation by a thin lens: with GOM the image of a point (A_{d_0}) is a point if A_{d_0} lies on the in-focus plane, but is otherwise a spot. With POM these patterns are convolved by the PSF of the optical system.

C. Model fitting on optical blur

We propose to use a video to robustly estimate the depth of a ROI and later estimate its size assuming that this ROI is planar and frontoparallel to the distal end of the gastroscope. This strong assumption allows us to easily derive a measurement thanks to a preoperative calibration stage (Section IV) but could lead to the underestimation of tumor size (SectionV). As the optical system of most colonoscopes is fixed, the closest in-focus frame gives the depth. Our approach relies on a coherent video sequence corresponding to a push or pull back of the colonoscope's tip towards the ROI. The PSF measurement used for this study relies on [16] (Section IV.A). The PSF measurement is done in the frequential domain by studying the effect of reblurring salient features. The inverse of the generic model (4) can thus be fitted to the PSF measures, by solving the following optimisation problem:

$$\min_{\mathbf{p}\in\mathbb{R}^6} \sum_{d_i=1}^{n} \left(\eta(d_i;\mathbf{p}) - \sigma_f(d_i)\right)^2 \tag{5}$$

where d_i is the frame number of the video sequence linked to the distance by the coherent motion of the gastroscope. The parametric generic model (with $\mathbf{p} = \{p_j\}_{1 \leq j \leq 6}$) derives from (4) and is defined by:

$$\eta(d; \mathbf{p}) = \frac{p_1}{|p_2 - \frac{p_3}{d - p_6}| - p_4(d - p_6)} + p_5 \qquad (6)$$

The maximum of the fitted model corresponds to the minimum σ_s of the PSF (the in-focus plane). The initialization of model parameters, which is critical, is given by the following relationships (assuming a pull-back of the colonoscope):

(i) p_5 and p_6 relate respectively to the value at the origin and the amplitude of the data to fit;
(ii) $\max(\sigma_f(d_i))$ is at a singular point as: $p_2 = \frac{p_3}{d - p_6}$;
(iii) $\lim_{d \to +\infty} \sigma_f(d_i) - p_5 = \frac{p_1}{|p_2| + p_4(p_5 - p_6)}$;
(iv) equation (6) can be rearranged and differentiated leading to:
$\lim_{d \to +\infty} \frac{dU}{dD} = p_4$ Where $U = \frac{\sigma_f - p_6}{p_1}$ and $D = d - p_6$.

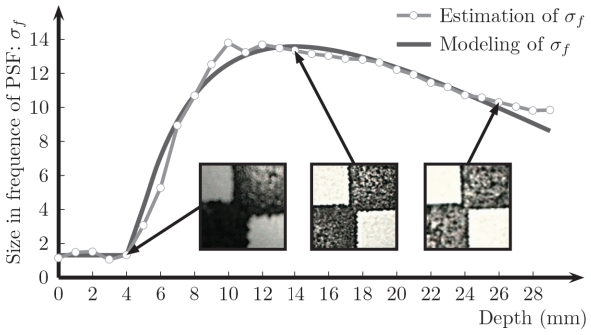

Fig. 2. Validation of the PSF model: the light gray curve represents the estimate of the size in frequencie of the PSF σ_f of varying controlled depths. The dark gray curve is the model.

D. Experimental validation of the PSF model

A first set of experiments was done in a laboratory with an Olympus® CV-160 colonoscope. The distal end of the colonoscope was fixed to the support of a motorized linear stage and an object was placed in contact with it. The motorized stage allowed us to control the pull back of the colonoscope from the distance $d = 0$ to $d = 100$ mm which is the suitable range of distances in colonoscopy. Experiments showed that the proposed model suitably fits the PSF estimation. Different textured objects (chessboard pattern, red pepper ...) were used in order to test the modeling. Experiments have lead to similar results with a single optimum depth. Experiments (section V) were also carried out in a preoperative context with two differents Olympus® CV-180 colonoscopes. They have been realized by acquiring images of a chessboard pattern at predefined distance (with a step of 5 mm). The results validate the proposed modeling (figure 3) despite a rough calibration.

IV. Clinical Protocol and Implementation Details

A. PSF estimation method and calibration

The estimation of the standard deviation (STD) of the PSF was indirectly obtained thanks to the method proposed in [16]. This method relies on the Gaussian reblurring of edges and relies on the fact that a blurred edge is inherently less sensitive to low-pass filtering than a sharp edge. The gradient ratio between the input and reblurred images allows us to evaluate the STD of the PSF.

A preoperative calibration stage is required in order to obtain the absolute depth of a ROI. This calibration aims at obtaining a curve relating the depth of an object and the estimation of the STD of the PSF (figure 2). The maximum of this curve corresponds to the distance of the object plane in focus. As part of this study, an ideal calibration process was done for laboratory experiments using a motorized linear stage. However, in a medical context, such an accurate calibration is not suitable and a rough calibration was done in preoperative conditions by acquiring a set of images at different depths.

B. In-vivo protocol

Similar to the experiments presented in the previous sections, a peroperative protocol was defined. The distal end of the gastroscope is firstly far away from a neoplasia and progressively approched to touch it. In order to avoid issues due to distortion [1], the ROI was imaged in the center of the frames. The gastroscope has been moved slowly in order to avoid movement blur which could lead to bias in results. Several studies address the problem of motion blur detection [7] and should be considered to improve the robustness of the proposed approach. This study assumes no motion blur in order to focus on the accuracy of the depth estimation. Moreover the video treatment was realized of line.

V. Experimental Results

In order to validate the previously exposed method, two datasets were used. Laboratory experiments were carried out with an Olympus® CV-160 colonoscope on a phantom model and real colonoscopy sequences were also evaluated with two Olympus® CV-180 colonoscopes.

A. Results on ground truth dataset

The phantom model is composed of a tube within which a colon pig has been inserted and flattened against the cavity wall. Pig colons are frequently used by experts for training purposes due to their similarity to human colon. Two marbles of diameter 5.75 mm (blue marble) and 15.54 mm (transparent marble) were alternatively placed inside the colon cavity. A motorized linear stage was used in order to control the pull-back of the distal end of the colonoscope. Figure 3 shows the result for these two cases. The blurring model proposed fits the measurements of the PSF STD. Moreover, a singular maximum corresponding to the best in-focus frame of the video sequence allows us to compute

the absolute size of the marble. The sizes are underestimated as expected due to our initial assumptions.

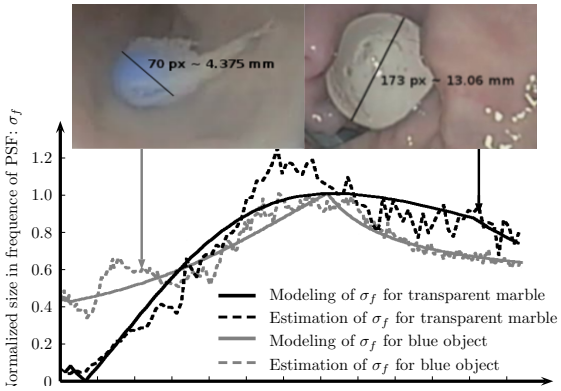

Fig. 3. Measurement of object size with controlled depths. The black curves are the estimate and modeling of the size in frequence of the PSF with a transparent marble object. The gray curves are the estimate and modeling with a blue object. The maximums of the two modelings are on the same point, because the same colonoscope was used.

B. Results on real datasets

Three real colonoscopy sequences have been acquired by an expert. In order to estimate the robustness of the proposed method, colonoscopy forceps was placed in contact with the observed polyp. This technique is frequently used by experts for size estimation of lesions. Tumor sizes are underestimated due to the initial assumptions but the order of accuracy matches the needs of experts. The model, which fits the measurement of the PSF STD, allows us to prevent from possible outliers due to changing illumination condition. The robustness of measurements could be improved thanks to an advanced calibration step. A fine tracking method of the tumor inside the video sequence as well as a local estimation of the PSF STD have also to be considered.

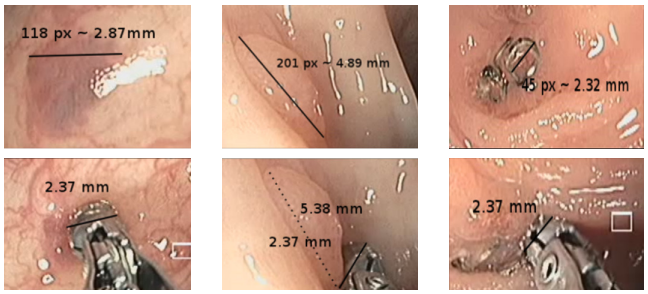

Fig. 4. Measurement of polyp size in *in vivo* sequences. The first row represents the images of 3 real datasetsets in the plane of focus. The second row shows the images which permit to compute the real size.

C. Discussion

The order of accuracy matches the requirement of gastroenterological experts for the diagnosis of neoplastic lesions. Apart from the initial assumptions (plane tumor frontoparallel to the camera), the results obtained with our method are also underestimated because of the significant barrel distortion effect of colonoscopic optical system. Removing such a distortion requires a more accurate calibration process. The changing illumination conditions, and more particularly the saturation due to the non Lambertian reflectance of biological tissues, have an important influence on the blur estimation. The use of a coherent colonoscopic sequence which could be fitted to the blur model allows us, nevertheless, to estimation the size of neoplasias.

VI. Conclusions and Future Work

We have proposed a method to estimate the size of neoplasias *in-vivo*. Our method combines Depth-from-Focus and Depth-from-Defocus. It detects the closest in-focus plane and infers its distance from the colonoscope's tip using a pre-calibration step. The size of a neoplasia is then computed under the assumption that it is frontoparallel to the colonoscope. Our experimental results demonstrate the feasibility of this approach. Quantitative measurements compared to groundtruth for both lab and *in-vivo* cases reveal a slight systematic underestimation of size. Our future work will address local blur estimation so as to raise the frontoparallel assumption.

References

[1] Bräuer-Burchardt, C., Heinze, M., Munkelt, C., Kühmstedt, P., Notni, G.: Distance Dependent Lens Distortion Variation in 3D Measuring Systems Using Fringe Projection. *British Machine Vision Conference*, 327–336 (2006).

[2] Claxton, C.D., Staunton, R.C.: Measurement of the point-spread function of a noisy imaging system. *J. Opt. Soc. Am. A*, 25, 159–170 (2008).

[3] Favaro, P., Soatto, S., Burger, M., Osher, S.: Shape from Defocus via Diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 518–531 (2008).

[4] Hartley, R. I., Zisserman, A.: Multiple View Geometry in Computer Vision – 2nd Edition, *Cambridge University Press*, (2004).

[5] Kaufman, A., Wang, J.: Mathematics and Visualization, Visualization in Medicine and Life Sciences. *Springer*, (2008).

[6] Kim, D.H, Pickhardt, P.J., Taylor, A.J., Leung, W.K., Winter, T.C., Hinshaw, J.L., Gopal, D.V., Reichelderfer, M.H., Richard, H., Pfau, P.R.: CT Colonography versus Colonoscopy for the Detection of Advanced Neoplasia. *The New England Journal of Medicine*, 357, 1403–1412 (2007).

[7] Liu, R., Zhaorong, J.: Image partial blur detection and classification. *IEEE Conference on Computer Vision and Pattern Recognition*, 1–8 (2008).

[8] Minhas, R., Adeel Mohammed, A., Jonathan Wu, Q. M.: Shape from focus using fast discrete curvelet transform. *Pattern Recognition Letters*, 44, 839–853 (2011).

[9] Nayar, S.K., Nakagawa, Y.: Shape from Focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16, 824–831 (1994).

[10] Pentland, A.P.: A new sense for depth of field. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9, 523–531, (1987).

[11] Rajagopalan, A.N., Chaudhuri, S.: Optimal Recovery of Depth from Defocused Images Using an MRF Model. *International Conference on Computer Vision*, 1047–1052 (1998).

[12] Classification, E: Update on the Paris classification of superficial neoplastic lesions in the digestive tract. *Journal of Endoscopy*, 37, 570–578 (2005).

[13] Subbarao, M., Surya, G.: Depth from Defocus: A Spatial Domain Approach. *International Journal of Computer Vision*, 13, 271–294 (1994).

[14] Wei, Y., Dong, Z., Wu, C.: Global depth from defocus with fixed camera parameters. *Mechatronics and Automation, ICMA*, 24, 1887–1892 (2009).

[15] Xiong, Y., Shafer, S.A.: Depth from focusing and defocusing. *Proceedings of IEEE Computer Vision and Pattern Recognition*, 68–73 (1993).

[16] Zhuo, S., Sim, T.: On the Recovery of Depth from a Single Defocused Image. *Procedings of the 13th International Conference on Computer Analysis of Images and Patterns*, (2009).