Elliptical Specularity Detection in Endoscopy with Application to Normal Reconstruction

Karim Makki[1*], Kilian Chandelon[1] and Adrien Bartoli[1]

[1]EnCoV, IP, UMR6602 CNRS/UCA, Clermont-Ferrand, France.

*Corresponding author(s). E-mail(s): karim.makki@uca.fr;

Abstract

Purpose. To detect specularities as ellipses in endoscopic images to facilitate normal reconstruction. **Methods.** A pipeline combining deep learning with handcrafted steps. **Results.** Convincing results in detection and reconstruction on synthetic and real images. **Conclusion.** First fully automatic method to exploit specularities in endoscopic 3D reconstruction. Promising results towards future integration with learning-based depth inference and SfM methods.

Keywords: specularity, ellipse, detection, endoscopy, 3D reconstruction.

1 Purpose

Specularities are generally considered as nuisance, even in conditions where they are numerous. However, they have recently been shown to bring useful information in endoscopy, with as main assumptions that the light source is to some extent an isotropic point collocated with the camera centre, leading to the following results [1]: *in the endoscopic setting, 1) specularities are near-elliptical blobs and 2) the surface normal can be reconstructed from the ellipse coefficients.* The ability to reconstruct surface normals from a single image is a potential boost for learning-based depth inference, which suffers high uncertainty in specularities, and multi-image 3D reconstruction methods such as Structure-from-Motion. The state of the art lacks an effective method to fully exploit the information contained in specularities. The main reason is the lack of a reliable method to detect elliptical specularities. The problem involves four main challenges: C1) the local intensity strongly varies; C2) the overlap

of neighbouring blobs; C3) the small size of many blobs, which make ellipse fitting unstable; C4) the unwanted blobs off the tissues, for instance on the interventional instruments. We propose a method which addresses these challenges by leveraging deep learning to reach a global understanding of the image and ellipse fitting to achieve high accuracy. Experimental results show that our method produces stable normal estimates in a fully automatic manner from a single endoscopic image.

2 Methods

2.1 Pipeline

Our method follows the pipeline of figure 1, which combines a neural network and handcrafted steps, taking the best of both. The neural network finds relevant blobs but has limited precision whilst the handcrafted steps, exploiting the elliptical prior, bring high precision.
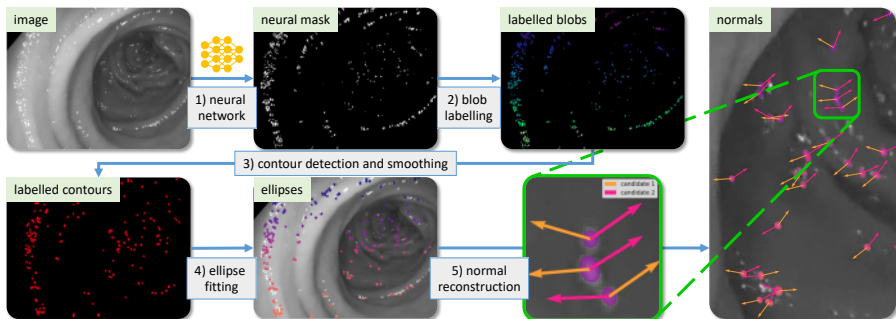

Fig. 1: Proposed pipeline for specularity detection and normal reconstruction, combining deep learning with handcrafted steps exploiting the elliptical prior.

2.2 Neural Network

We achieve *Step 1)* with a fully convolutional network to produce an initial *neural mask M* which specifically finds specular pixels on the colon tissue, dealing with challenges C1) and C4). Our network has 18 layers with ReLU activation. We created a dataset of 2240 images extracted from 13 colonoscopy procedures from the Endomapper dataset, which we manually annotated using adaptive thresholding. We split the dataset in 2000 training, 200 validation and 40 test images. We used data augmentation using standard image transformations, namely horizontal and vertical flipping and realistic brightness adjustment within range $[-0.3, 0.3]$. We trained with Adam for 20 epochs with a batch size of 32, a learning rate of 0.001 and a binary cross-entropy loss function. We obtain a Dice score of 80.05% with standard deviation 23.06% on the test set.

2.3 Handcrafted Steps

The three handcrafted Steps 2) to 5) start from the neural mask M and reconstruct the normals. *Step 2) blob labelling.* The blobs are isolated, dealing with challenge C2), by giving a unique id to the connected components in M, with background pixels labelled as 0. We only keep those components whose area lies between 7×10^{-4} and $3 \times 10^{-3}\%$ of the image area. For our 1248×1080 images, the range is $[10, 40]$ px^2. This was found empirically to discard overly small blobs (unstable ellipses) and large blobs (breaking the local planarity assumption or consisting of overlapping specularities). *Step 3) contour detection and smoothing.* The isophote contours are extracted and smoothed. We use marching squares to detect the zero level set for the labelled blobs, leading to a set of closed contours. We fit a smoothing cubic B-spline, which reduces the effect of noise, and resample the countours with 1000 points. *Step 4) ellipse fitting.* We use [2], dealing with challenge C3). *Step 5) normal reconstruction.* We use [3], which takes the ellipse coefficients as input.

3 Experimental Results

Simulated data. We modified the synthetic image renderer of [1] to use a sphere instead of a plane, hence creating curvature. We use the same parameters but increase the roughness parameter to $n = 120$ for a better trade-off between specularity size and surface curvature. The average angular error on the estimated normals over 300 trials is shown in figure 2. First, we observe that the normal error remains in a reasonable range lower than 14 degrees. Second, we observe that the larger the sphere radius R, the smaller the normal error. This was expected as a larger radius flattens the surface, hence conforms it to the planar hypothesis made in normal reconstruction. Note that we took care of keeping the distance from the camera to the brightest point of the specularity unchanged. Third, we observe that the large the camera tilt θ with respect to the sphere's normal at the brightest point of the specularity, the larger the error. This was expected because a larger tilt creates more perspective, carrying the effects of curvature more strongly.
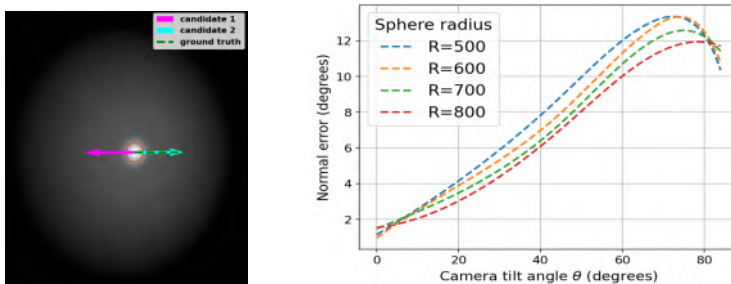

Fig. 2: Synthetic data. (left) Example image. (right) Normal errors.

Real data, segmentation evaluation. We selected 5 representative colonoscopic images from the Endomapper dataset, with various viewing conditions and the presence of interventional tools. We segmented the images with ITK-SNAP, creating 1252 specular blobs, which we ensured are isolated and in the range of size leading to stable normal reconstruction. We used this carefully labelled dataset to evaluate performance and perform an ablation study. First, we ran our pipeline without the neural network, replaced intensity thresholding, obtaining a mean Dice of 43%. Second, we ran our pipeline without the handcrafted steps, directly using the neural mask as final result, obtaining a mean Dice of 66%. Finally, we ran our full pipeline, obtaining a mean Dice of 84%. This shows that 1) both the learning-based and the handcrafted steps contribute and 2) the complete pipeline segmentation is complete and accurate.

Real data, reconstructed normals comparison. We compared the normals reconstructed by our method to the external depth reconstruction method [4]. This allowed us to perform a test on real colonoscopic images in challenging conditions, where ground truth is not available. Concretely, we used the 7 images and depthmaps showcased in [4], from which we extracted normal maps. Our method reconstructed 67 normals and we obtained an average angular discrepancy of $12.11°$ with standard deviation $9.86°$. This discrepancy indicates a satisfying agreement between the two methods. The low image resolution of 256×256 required by [4] explains the lower number of normals detected by our method and possibly the angular discrepancy.

4 Conclusion

We have proposed a method that combines learning-based and handcrafted steps to achieve elliptical specularity detection, from which robust and accurate normal reconstruction follows. We plan to combine our method with Structure-from-Motion and run advanced tests in colonoscopic navigation.

References

[1] Makki, K., Bartoli, A.: Normal reconstruction from specularity in the endoscopic setting. arXiv preprint arXiv:2211.05642 (2022)

[2] Halır, R., Flusser, J.: Numerically stable direct least squares fitting of ellipses. In: Proc. 6th Int. Conf in Central Europe on Computer Graphics and Visualization. WSCG, vol. 98, pp. 125–132 (1998). Citeseer

[3] Forsyth, D., et al.: Invariant descriptors for 3D object recognition and pose. IEEE PAMI **13**(10), 971–991 (1991)

[4] Rau, A., et al.: Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy. IJCARS **14**(7), 1167–1176 (2019)