

NORMAL RECONSTRUCTION FROM SPECULARITY IN THE ENDOSCOPIC SETTING

Karim Makki, Adrien Bartoli

EnCoV-Institut Pascal, CNRS/Université Clermont-Auvergne, Clermont-Ferrand 63000, France

ABSTRACT

We show that for a plane imaged by an endoscope the specular isophotes are concentric circles on the scene plane, which appear as nested ellipses in the image. We show that these ellipses can be detected and used to estimate the plane’s normal direction, forming a normal reconstruction method, which we validate on simulated data. In practice, the anatomical surfaces visible in endoscopic images are locally planar. We use our method to show that the surface normal can thus be reconstructed for each of the numerous specularities typically visible on moist tissues. We show results on laparoscopic and colonoscopic images.

Index Terms— Specular, reconstruction, endoscopy.

1. INTRODUCTION

Specularities typically appear as moving white spots on the tissues in endoscopy; they are usually considered as nuisance [1]. In contrast, we propose to use them as 3D reconstruction cues arising from a single image. They may then be used as valuable assets to integrate in existing multi-image endoscopic 3D reconstruction systems [2, 3, 4]. Although shape-from-specularity has been researched in computer vision, our proposal of using the specular isophotes for 3D reconstruction is new. Existing work is generally based on tracking specularities or reflections over multiple images. The analysis in [5] shows that local surface convexity or concavity can be determined from the image motion of specularities with a known camera trajectory. The method in [6] reconstructs the shape mesostructure from a dense specularity field from hundreds of images. The method in [7] uses the specular flow for estimating scene structure from 2D image motion. The method in [8] uses a quadric surface model, an orthographic camera, and requires the object to rotate around the optical axis to reconstruct the shape of a mirror surface from sparse correspondences. The method in [9] reconstructs the pose of a moving plane from three images of its reflection on a specular surface. These methods rely on assumptions which cannot be fulfilled in interventional medical images.

In contrast, we provide theoretical and practical insights into the 3D reconstruction of the observed tissue’s normal

from an isolated specularity. We show that once a specularity is detected, the shape normal at the specularity’s Brightest Point (BP) can be estimated in closed-form. We derive these results under the so-called endoscopic setup, which is that the light is a point-light source collocated with the camera centre [10, 11]. Our main assumption is then that the observed surface is locally flat and that the reflectance is modelled by Phong’s specular model. We study the isophotes, which are the curves of equal incoming intensities on the camera’s retinal plane. We show that under the above assumptions the isophotes are concentric circles on the scene plane and ellipses in the image. From this key result, we use well-known 3D vision results on conics, of which the ellipse is a special case. Specifically, we show that using a single ellipse isophote detected in an image, the scene plane’s normal can be reconstructed up to a two-fold ambiguity. This leads to a reliable method of local normal reconstruction for mildly curved surface patches. We show results on synthetic and real images of the liver and the colon.

2. THEORY AND METHODS

2.1. General Specular Isocurve Model

A model of specular isophotes on planes was introduced [12] with Phong’s reflection [13], where it was shown that the specular isophotes are subset of quartic isocurves on the scene plane. We briefly recall and adapt the derivation of this model and specialise it to the endoscopic setup. The scene plane is chosen as the XY -plane with 2D points $p^\top = (x, y)$ corresponding to 3D points $P^\top = (x, y, 0)$. Defining $V, L \in \mathbb{R}^3$ as the viewpoint and light source position, the cosine formula [13] gives specular reflection as:

$$\hat{I}_s(p) \stackrel{\text{def}}{=} c \max(0, \cos(\beta(p)))^n, \quad (1)$$

where $n > 0$ characterises the surface roughness, $c > 0$ is an unknown constant factor modelling the surface albedo, light intensity and linear camera response coefficient, and $\beta(p)$ is the angle between the viewing direction $\mu(V - P)$ and the direction of perfect reflection $-\mu(R - P)$ such that $\mu(U) \stackrel{\text{def}}{=} \frac{U}{\|U\|}$ for $U \in \mathbb{R}^3$ and $R^\top = (L_X, L_Y, -L_Z)$. By substituting these components in equation (1), simplifying and introducing $\tau = \sqrt[n]{\frac{L}{c}}$, it was shown [12] that, on an infinite scene

plane, the specular isophote $\hat{I}_s(p) = t$, $0 < t \leq c$ is a subset of the quartic isocurve (that is, given by nullifying a bivariate polynomial of degree four) $Q(p) = 0$, called the *general specular isocurve*, with:

$$Q(p) = ((R-P)^\top(V-P))^2 - \tau^2 \|R-P\|^2 \|V-P\|^2. \quad (2)$$

We point out the distinction that exists between an *isophote* $\hat{I}_s(p) = t$, which is a curve on the scene plane of constant intensity as observed in the image, and its model as a quartic *isocurve* $Q(p) = 0$. The quartic isocurve is a pair of curves; parts of which form the isophote [12], in a non-trivial and non-systematic way. We show that simplifications exist in the endoscopic setup.

2.2. Endoscopic Specular Isophote Model

In the endoscopic setup, the light source and the camera are co-located and we thus set $L = V$. We substitute $\kappa = 1 - \tau^2$ and $R^\top = (V_X, V_Y, -V_Z)$ in equation (2), which after simplification yields the *endoscopic specular isocurve* $E(p) = 0$, with:

$$E(p) = \kappa((x - V_X)^2 + (y - V_Y)^2)^2 + 2V_Z^2(\kappa - 2)((x - V_X)^2 + (y - V_Y)^2) + \kappa V_Z^4. \quad (3)$$

This is still a quartic equation, albeit of a simpler form than the general quartic isocurve (2). Our main results are next given as two propositions, which characterise the geometric curve represented by this quartic.

Proposition 1. *The endoscopic specular isocurve $E(p) = 0$ represents a pair of concentric circles with centre $(V_X, V_Y)^\top$ and radii $r_\pm = |V_Z| \sqrt{\frac{2-\kappa}{\kappa} \pm \sqrt{(\frac{2-\kappa}{\kappa})^2 - 1}}$.*

Proof. The quartic (3) factorises into the product of two quadratics as:

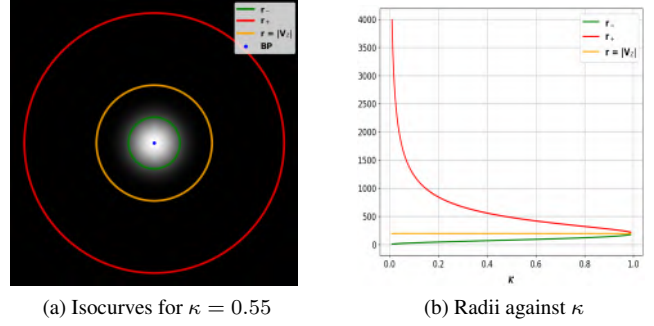
$$E(p) = q_+(p) q_-(p), \quad (4)$$

where $q_\pm(p) = (x - V_X)^2 + (y - V_Y)^2 - r_\pm^2$ and q_+ , q_- represent the sought circles. \square

Proposition 2. *The endoscopic specular isophote is the circle of smallest radius r_- from the endoscopic specular isocurve.*

Proof. For $0 < t < c$ we have $0 < \kappa < 1$ and $r_+ > |V_Z| > r_- > 0$. It was shown [12, lemma 27] that, in the general case, the isophote lies in a so-called ‘outer ring’, a circle of radius $|V_Z|$ and with centre $(V_X, V_Y)^\top$. Therefore, in the endoscopic case, the isophote must be the smallest of the two circles of the isocurve. \square

Figure 1(a) illustrates the isocurves obtained for simulated data, as described in section 3. We clearly see that the specularity is radial on the scene plane and that the isophote is given by the circle of radius r_- (in green), lying inside the outer ring (in orange), whilst the circle of radius r_+ lies outside it (in red). This is verified for all isophotes in figure 1(b).



(a) Isocurves for $\kappa = 0.55$

(b) Radii against κ

Fig. 1: Endoscopic specular isocurves on the scene plane.

2.3. Reconstruction Method

Proposition 2 shows that the specular isophotes are concentric circles on the scene plane. By property of perspective projection, these circles form ellipses on the image plane. We use this property to reconstruct the plane normal from the specularly observed in the image in closed-form. We denote as $C \in \mathbb{R}^{3 \times 3}$ the symmetric matrix representation of the isophote ellipse as a so-called point-conic in the image. Background on ellipses and conics is given in [14, Chapter 8]. We proceed with a given isovalue $0 < t < m$, where m is the maximum image intensity (typically $m = 255$ for 8-bit images), a given region of interest Ω containing the specularity and a given intrinsic camera parameter matrix $K \in \mathbb{R}^{3 \times 3}$ obtained by classical calibration [14, Chapter 6]. Our method follows four steps. First, we find the image pixels pertaining to the isophote in Ω using the Marching Squares algorithm, the 2D version of the popular Marching Cubes algorithm. Second, we fit the isophote ellipse using the closed-form solution [15]. This gives us matrix C . Third, we transfer the ellipse to the normalised camera plane as $C' = K^\top C K$. Fourth, we use the circle pose estimation method [16, Appendix A], or so-called backprojection problem for one conic correspondence, which provides us with two possible normals N_\pm for the scene plane. These two normals are due to a concave/convex surface ambiguity, under the assumption that the isophote lies on the tangent plane to the observed surface at the BP. This ambiguity cannot generally be resolved from a single viewpoint [5]. If the circle radius were known, the camera to plane distance could be reconstructed. However, as opposed to observing physical circles, the circle radius related to a specularity is never known, and the camera to plane distance is thus unrecoverable.

3. EXPERIMENTAL RESULTS

3.1. Synthetic Data

Figure 2 illustrates the three-stage evaluation pipeline with synthetic data. In stage 1, we simulate the imaging process with the following three steps. 1) We compute the scene

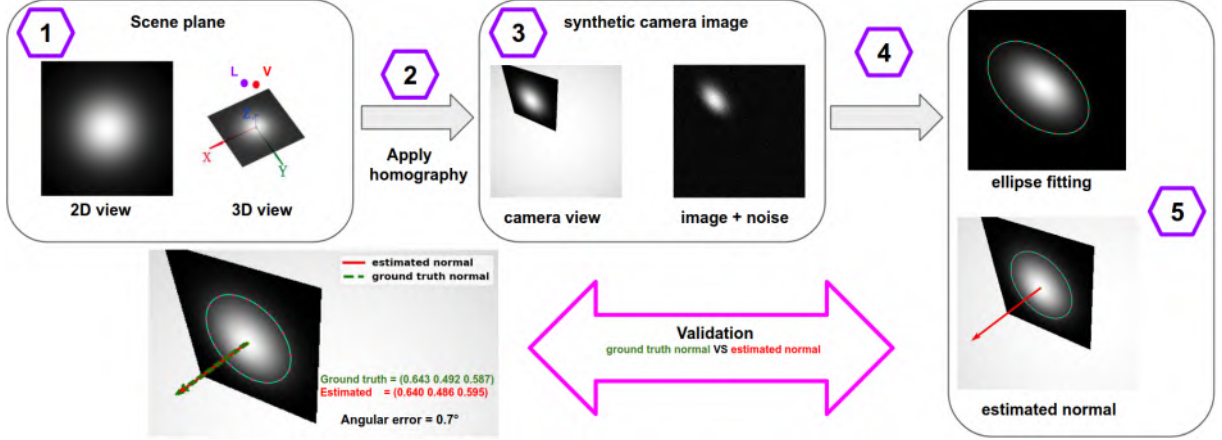


Fig. 2: Pipeline for the validation of normal reconstruction from specular isophotes on simulated data.

plane texture map as an $M \times M$ image using equation (1) with material roughness n . Without loss of generality, we set $V_X = V_Y = 0$ so that the BP lies at the image center. We thus have that $V_Z > 0$ is the camera to scene plane distance. We simulate a non-perfect endoscopic setup by choosing $\epsilon > 0$ and setting $L = V + \epsilon(\cos(\alpha), \sin(\alpha), \beta)^T$, where α and β are uniformly distributed in $[0, 2\pi[$ and $[-0.5, 0.5]$, respectively. 2) We apply a projective transformation to the scene plane by means of a homography H , to form its sought perspective image I_s . We use a slanted view, as typical of colonoscopy, with a plane normal to camera axis angle $\theta = 58^\circ$. We set the camera intrinsics as $f = M$ for the focal length and $c_x = c_y = M/2$ for the principal point. 3) We add white Gaussian noise to the simulated image with variance σ^2 . In stage 2, we proceed with the reconstruction process with the following two steps. 4) We use Gaussian smoothing to reduce the effect of noise. We then normalise the image intensities so that the intensity at the BP is 1. 5) Given the isovalue $0 < t < 1$, we run the proposed reconstruction method. In stage 3, we compute the angular error on each of the estimated normals. The default parameters are $\epsilon = 0$, $M = 406$, $V_Z = 1000$, $n = 50$, $\sigma = 5\%$, and $t = 0.1$; they lead to the illustration from figure 2.

The error plots for 1000 independent realisations are shown in figure 3. We vary five important parameters within their range of interest while freezing the others to their default values. First, the noise σ is varied from 0% to 10% of the intensity range. We observe that the normal error increases with noise, in average and in standard deviation. It is on average kept below 1.25° , including the case of very noisy images. Second, the plane to camera angle θ is varied from 0 (fronto-parallel view) towards $\frac{\pi}{2}$ (extreme grazing view). We observe a strong growth of the error, with nonetheless a maximum value kept under 7° . Third, the material roughness n is varied from 30 to 120. The higher this parameter, the steeper the intensity profile of the specularity, making isophote detection more sensitive to noise. We indeed observe that the

normal error increases linearly with roughness, with a very small slope of approximately 10^{-2} , without exceeding 1.75° . Fourth, the isovalue t is varied from 0.02 and 0.8. Recall that for a low value the isophote is wide and close to the outer ring whilst for a high value it shrinks closely to the BP. Therefore the stable configurations should be between these two extremes. We observe that this indeed happens, with a globally U-shaped normal error curve, reaching a minimum at about 0.5 to 0.6. Fifth, the imperfect co-location of the light source and camera centre ϵ is varied between 0 and 800, which, as for V_Z is in number of pixels. Specifically, we first centre the scene plane's window at the BP and choose $n = 100$ to explore larger space of model imperfections while ensuring that the model errors are not mixed with errors related to partial view. Recall that for a large value of ϵ the isophote is no longer a circle in the scene plane. We observe that the normal error remains stable for a light source and camera located inside a cylinder of radius 200 mm and height 200 mm (*i.e.*, for $\epsilon \leq 200$). This is much larger than the distance which may exist in the design of most endoscopes.

3.2. Real Data

We used our method in two concrete cases, laparoscopic liver resection and colonoscopy. In both cases, the practical problem is to obtain an interventional 3D model from images. This is difficult because the observed structures are very deformable, precluding the use of many vision techniques such as Structure-from-Motion. It is thus tremendously important to be able to exploit single-image 3D reconstruction, for which our method is well-adapted. First, the camera can be easily calibrated when the procedure starts. Second, because the observed structures are moist, the images contain many specularities. Third, the specularities are typically small in size, and thus supported by a locally planar surface region.

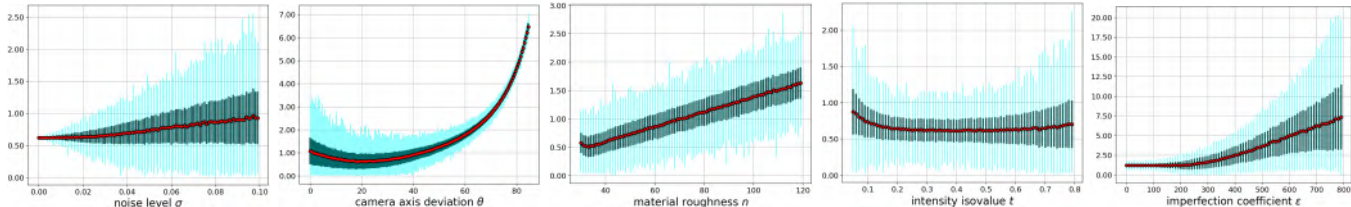


Fig. 3: Evaluation results on synthetic data, giving the average normal error \bar{E} in degrees. The black bars represent the error standard deviation and the cyan bars represent the interval $[\bar{E} - \min(E), \max(E) - \bar{E}]$.

3.2.1. Laparoscopic Liver Resection

We ran our method on five images with a total of 26 specularities extracted from five procedures collected in our hospital under ethical approval IRB00008526-2019-CE58 issued by CPP Sud-Est VI in Clermont-Ferrand, France. A representative example is shown in figure 4. We disambiguated the normal-pairs manually and performed a comparison with [17], a method which registers a preoperative 3D model to the image, obtaining an angular difference of $7.11^\circ \pm 4.86^\circ$. This shows a very good quantitative agreement between our method and [17].

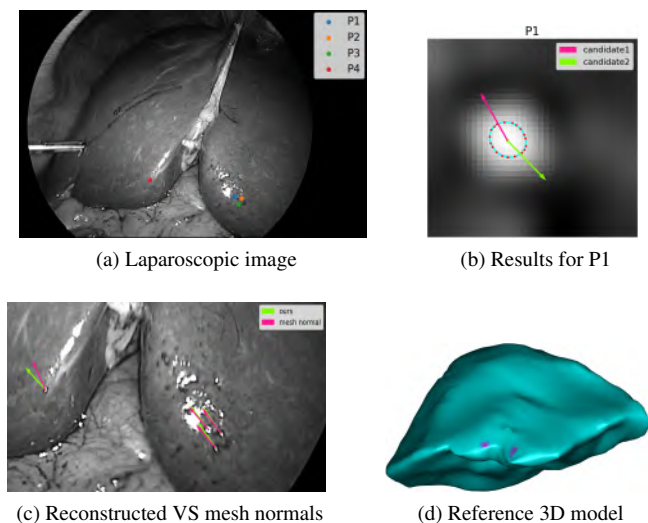


Fig. 4: Example of normal reconstruction for a 2D laparoscopic image. (a) The four specularities used. (b) Close-up and reconstructed normals for P1. (c) Reconstructed normals for our method compared to preoperative model registration [17]. (d) Location of the four specularities on the registered preoperative model.

3.2.2. Colonoscopy

We ran our method on 1248×1080 images extracted from colonoscopy procedures within the IRB approved protocol of the Endomapper project. We show representative examples in figure 5 with 11 reconstructed normals. We did not find a method which would provide a satisfying 3D reconstruction for comparison. Nonetheless, the reconstructed normals we

obtained are qualitatively sound, showing the potential applicability of our method to colonoscopy. We considered using the simulated data from EndoSLAM [3], for which however the surface is made of fronto-parallel planes whose normal is thus always $(0, 0, 1)^\top$. In contrast, [18] provides smoother depth maps but the surface triangulation is not dense enough to create realistic local specularities, as shown in figure 6.

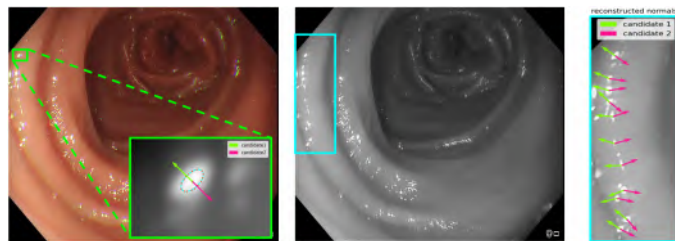


Fig. 5: Sparse normal reconstruction for the colon surface. We show the two possible orientations of each local tangent plane containing an isolated specular highlight.

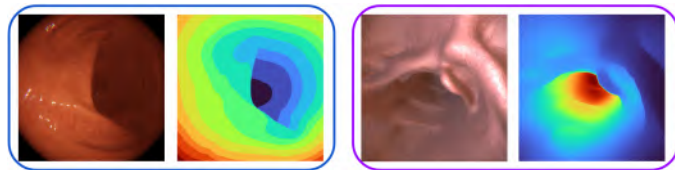


Fig. 6: Excerpts of publicly available datasets, left: EndoSLAM [3] and right: [18].

4. CONCLUSION

We have introduced a geometric model of the specular isophotes in endoscopy, from which we have derived a normal reconstruction method. This is based on the general specular model [12], which we have specialised to endoscopy and extended to 3D reconstruction. We have obtained convincing results in two clinical applications. Our future work will focus on two main points. First, we will design automatic specularity detection by means of combining deep learning and the elliptic shape prior we have derived. This will increase the applicability of our method and cope with the mixed specularities of non-elliptical shape owing the inter-reflections. Second, we will incorporate our method into the tubular-NRSfM method applied to the colon [4].

Declarations Informed consent was obtained from all individual participants included in the study. This article does not contain any studies with animals performed by any of the authors.

5. REFERENCES

- [1] R. Daher, F. Vasconcelos, and D. Stoyanov, “A temporal learning approach to inpainting endoscopic specularities and its effect on image correspondence,” *arXiv preprint arXiv:2203.17013*, 2022.
- [2] CS. Bell, KL. Obstein, and P. Valdastrì, “Image partitioning and illumination in image-based pose detection for teleoperated flexible endoscopes,” *Artificial intelligence in medicine*, vol. 59, no. 3, pp. 185–196, 2013.
- [3] KB Ozyoruk et al., “EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos,” *Medical image analysis*, vol. 71, pp. 102058, 2021.
- [4] A. Sengupta and A. Bartoli, “Colonoscopic 3D reconstruction by tubular non-rigid structure-from-motion,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 16, no. 7, pp. 1237–1241, 2021.
- [5] A. Zisserman, P. Giblin, and A. Blake, “The information available to a moving observer from specularities,” *Image and vision computing*, vol. 7, no. 1, pp. 38–42, 1989.
- [6] T. Chen, M. Goesele, and H-P. Seidel, “Mesostructure from specularity,” in *CVPR’06*. IEEE, 2006, vol. 2, pp. 1825–1832.
- [7] S. Roth and M-J. Black, “Specular flow and the recovery of surface structure,” in *CVPR’06*. IEEE, 2006, vol. 2, pp. 1869–1876.
- [8] A-C. Sankaranarayanan, A. Veeraraghavan, O. Tuzel, and A. Agrawal, “Specular surface reconstruction from sparse reflection correspondences,” in *CVPR*. IEEE, 2010, pp. 1245–1252.
- [9] M. Liu, Kwan-Yee K. Wong, Z. Dai, and Z. Chen, “Pose estimation from reflections for specular surface recovery,” in *CVPR*. IEEE, 2011, pp. 579–586.
- [10] T Okatani and K Deguchi, “Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center,” *Computer vision and image understanding*, vol. 66, no. 2, pp. 119–131, 1997.
- [11] SY Yeung, Hung-Tat Tsui, and A Yim, “Global shape from shading for an endoscope image,” in *MICCAI*. Springer, 1999, pp. 318–327.
- [12] A. Bartoli, “The highlight ovals,” *Journal of Mathematical Imaging and Vision*, vol. 61, no. 7, pp. 919–943, 2019.
- [13] Bui Tuong Phong, “Illumination for computer generated pictures,” *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, 1975.
- [14] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge university press, 2003.
- [15] A. Fitzgibbon, M. Pilu, and R-B. Fisher, “Direct least square fitting of ellipses,” *IEEE PAMI*, vol. 21, no. 5, pp. 476–480, 1999.
- [16] D. Forsyth, J-L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, “Invariant descriptors for 3D object recognition and pose,” *IEEE PAMI*, vol. 13, no. 10, pp. 971–991, 1991.
- [17] B. Koo, E. Özgür, B. Le Roy, E. Buc, and A. Bartoli, “Deformable registration of a preoperative 3D liver volume to a laparoscopy image using contour and shading cues,” in *MICCAI*. Springer, 2017, pp. 326–334.
- [18] A Rau et al., “Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy,” *International journal of computer assisted radiology and surgery*, vol. 14, no. 7, pp. 1167–1176, 2019.