RECONSTRUCTING THE NORMAL AND SHAPE AT SPECULARITIES IN ENDOSCOPY

Karim Makki and Adrien Bartoli

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

ABSTRACT

Specularities are numerous in endoscopic images. They occur as many white small elliptic spots, which are generally ruled out as nuisance in image analysis and computer vision methods. Instead, we propose to use specularities as cues for 3D perception. Specifically, we propose a new method to reconstruct, at each specularity, the observed tissue's normal direction (*i.e.*, its orientation) and shape (*i.e.*, its curvature) from a single image. We show results on simulated and real interventional images.

Index Terms— Specular, shape operator, endoscopy

1. INTRODUCTION

Specularities on smooth reflecting surfaces depend on the surface's local orientation and curvature, on the light source and on the camera position [1, Chapter 2]. In the general setting, all these contain far too many unknowns to allow one to exploit specularities effectively in 3D reconstruction. However, endoscopy is a particular case where the *light source and camera form a collocated rig*. Specularities are numerous in endoscopy, where the moist tissue causes specular reflection. However, they are generally considered as artifacts in endoscopic 3D reconstruction methods, both in classical methods such as Shape-from-Shading (SfS) [2] and learning-based methods such as direct depth estimation [3]. More recently, it has been understood that specularities could form useful cues for 3D reconstruction. The study [4] has shown that training a learning-based reconstruction model from synthetic data which can eventually deal with real data requires one to generate specularities. Concretely, we have recently shown that the surface normal could be reconstructed at each specularity [5, 6], with a two-fold ambiguity. The reconstruction method uses the *specular isophote*, which is the curve enclosing the specularity in the image. Based on the assumption that the surface underlying the specularity is flat, we have used a mathematical model to show that the isophote is elliptic. This approach has advantages: the method runs fast and can process high resolution images, in contrast to the learning-based approach which is currently limited to lower image resolution to keep runtime reasonable (the available datasets such as the EndoSLAM [7] and UCL [8] ones are

at lower resolution). However, it also has strong limitations. First, the reconstruction is up to a two-fold ambiguity: only one of the two directions found at each point is the correct normal. Second, the local planarity assumption is unreasonable, as specularities typically appear near highly curved surface patches. Is an ambiguous surface normal all one can reconstruct from a specularity?

We propose to move a step ahead by using a more advanced local surface model. Specifically, we leave the restrictive flat surface to adopt a curved surface model. We show that with this model, the specular isophote is still elliptic in the image, albeit that its eccentricity is not just due to surface orientation as in past work, but to both surface orientation and curvature. We show that one can reconstruct the surface normal from just the specularity's *Brightest Point* (BP) and the surface curvature from the elliptic isophote. Specifically, we show that the principal curvature directions (the directions of minimum and maximum curvatures on the tangent plane) and the curvature ratio can be reconstructed. Therefore, *the specular isophote contains much more information* than one could think from past work. We propose an automatic reconstruction method, which, from an endoscopic image and the camera's intrinsic parameters, detects the specularities and reconstructs the local surface orientation and curvature ratio at each specularity. This reconstruction can be used to implement downstream tasks such as computer-aided navigation, to bootstrap other reconstruction methods such as Shape-from-Shading (SfS) and to train depth estimation networks. We evaluate the method on realistic simulated colonoscopic data from Blender and surgical laparoscopy images.

2. THEORY AND METHODS

2.1. Mathematical Model

We consider a calibrated perspective camera model with intrinsic parameters in matrix $\mathbf{K} \in \mathbb{R}^{3 \times 3}$ and a point light source near the camera centre located at the origin. We consider a smooth specular surface and assume that the brightness reflected along the sightline direction V is a function of $V \cdot P$, where P is the *specular direction*, also known as the direction of a perfect specular reflection. This assumption is commonly used in reflection models such as Blinn-Phong [9] and the physics-based BSDF model popularised

by Disney [10]. In these models, each specularity has a BP, approximately located at its centre, where the reflected light energy is the highest. Let L be the lighting direction; it is known that, from the fact that the vectors L, V and P have unit length [1, 9, 10], the half-angle direction $H = (L+V)/2$ is collinear with the surface normal $\mathbf{N}$ at the specularity's BP. In endoscopy, as we have $L \approx V$, the surface normal is thus collinear with the sightline V.

We model the observed tissue as a collection of 3D surface patches. We model each patch by three components, as shown in figure 1 (left): 1) its *normal direction* $\mathbf{N}$, 2) its *tangent plane*, with basis vectors $\mathbf{u}_1$, $\mathbf{u}_2$ chosen along the patch's principal directions, and 3) its *curvature ratio*, which is the ratio of the minimal to the maximal curvatures. We instantiate one patch per specularity. We assume the specularity to conform to the elliptic model, which is widely applicable in endoscopic images, where they occur as small isolated white blobs [1, 5].
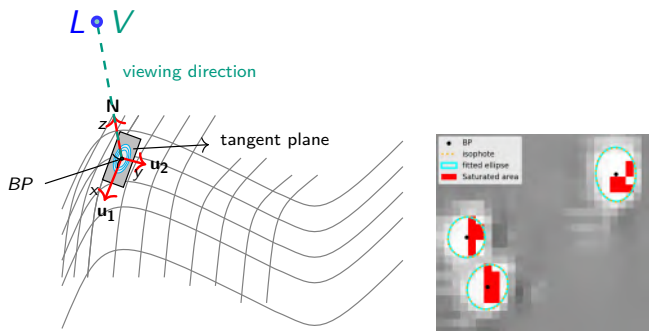
Fig. 1: **Left:** local coordinate system at the BP oriented using the normal and principal directions. The light and camera are collocated. Elliptic specular isophotes are shown in cyan. **Right:** example of three neighbouring surface patches from a real endoscopic image, showing the difficulty to determine the BP directly from image intensities.

2.2. Reconstruction Method

Our reconstruction method has three main steps given as Algorithm 1. The first step, which is as in [6], uses a learning-based specularity detector and ellipse fitting to retrieve an isophote for each specularity. We perform a simple ellipticity test based on the fitting residual. The second and third steps are geometry-based reconstruction given directly below.

2.2.1. Sightline-based Normal Reconstruction

We exploit the property that, at the BP, the surface normal is collinear with the sightline. We approximate the BP by the elliptic isophote's centre (x_0, y_0) in image pixel coordinates, as shown in figure 1 (right). The sightline is trivially found using the camera model, giving the surface normal direction $\mathbf{N} \in S^2$ as $\mathbf{N} \propto \mathbf{K}^{-1}(x_0, y_0, 1)^\top$, where S^2 is the

set of direction vectors in 3-space. Obviously, the reconstruction accuracy directly depends on the accuracy of the BP estimate. However, determining the exact location of the BP is challenging because of camera saturation at specularities. As seen in the right side of figure 1, all the red pixels have the maximal intensity for the selected surface patch. In contrast, it is much easier to detect an isophote whose centre and the unknown BP are nearly coincident.

2.2.2. Ellipse-based Shape Reconstruction

We represent a specular isophote ellipse by its standard 3×3 real symmetric matrix form C. This satisfies $(x, y, 1)C(x, y, 1)^\top = 0$ for the image pixel coordinates (x, y) of an isophote point. Next, we map the conic C to the normalised image plane[1] as $C' = \mathbf{K}^\top C \mathbf{K}$, which is also a 3×3 real symmetric matrix. It thus has the following properties: 1) it is always diagonalisable, 2) it has real eigenvalues, and 3) its eigenvectors are orthogonal. We diagonalise C' as $C' = \mathbf{R}^\top \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)\mathbf{R}$, where $\{\lambda_1, \lambda_2, \lambda_3\}$ are the eigenvalues sorted in ascending order. The corresponding eigenvector set $\{\vec{v_1}, \vec{v_2}, \vec{v_3}\}$ forms an orthonormal basis of $\mathbb{R}^3$. We have that $\mathbf{R} = (\vec{v_1}, \vec{v_2}, \vec{v_3})$ is a 3D rotation of the standard camera coordinate system to the so-called *eigenvector-frame*. It aligns the principal directions on the tangent plane at the BP with the ellipse's major and minor axes, with the normal direction $\mathbf{N}$ being aligned with the cross product $\vec{v_3} = \vec{v_1} \times \vec{v_2}$. The principal directions on the tangent plane are therefore given by $\mathbf{u}_1, \mathbf{u}_2 = \vec{v_1}, \vec{v_2}$. The ratio of principal curvatures $|\lambda_1/\lambda_2|$ can be derived from the ellipse's eccentricity, defined as $\epsilon = \sqrt{1 - (\lambda_1/\lambda_2)^2}$.

3. EXPERIMENTAL RESULTS

3.1. Synthetic Colonoscopic Data

3.1.1. Model

We use the modelling and rendering software suite Blender to synthesise images using a wide-angle perspective camera of 10 mm focal length. We use a spotlight located at the camera centre with a radius of 0.1 mm and projected to a 120° cone. We generate a colon-like deforming surface with free-form deformation [11]. The generated surface is represented by a triangular mesh with 502,502 vertices and 1,002,000 faces. We deform it over 250 frames. The surface is imparted with a strong 'clearcoat' to boost specular reflections. We use the anisotropic principled BSDF model [10], mimicking real-world shading and specularities.

---

[1]The origin is at the principal point and the distances are measured in units of focal length.

Algorithm 1 Single-image specular normals and shapes reconstruction

Input: Image I, camera intrinsics $\mathbf{K}$
Output: Sets of Brightest Points (BP), normals $\mathbf{N}$, principal directions $\mathbf{u}_1, \mathbf{u}_2$, curvature ratios ϵ
    - Use the neural network to segment the specular mask M from I ▷ [6]
    for each connected component $c \in M$ **do**
        - find the specular isophote as the outer boundary of c
        - smooth the isophote using a cubic B-spline
        - fit an ellipse to the isophote, with matrix representation C and fitting residual r
        if $r \leq t$ **then**   ▷ ellipticity test with threshold t
          - find the BP (x_0, y_0) as the ellipse centre
          - compute normal direction $\mathbf{N} = \mathbf{K}^{-1}(x_0, y_0, 1)^\top$
          - normalise it as $\mathbf{N} = \mathbf{N}/\|\mathbf{N}\|$
          - transfer the ellipse to the normalised image plane as $C' \leftarrow \mathbf{K}^\top C \mathbf{K}$
          - retrieve the ellipse axes by eigendecomposition as $C' = \mathbf{R}^\top \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)\mathbf{R}$, with eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and eigenvectors $\mathbf{R} = (\vec{v_1}, \vec{v_2}, \vec{v_3})$
          - determine the principal directions on the tangent plane as $\mathbf{u}_1, \mathbf{u}_2 = \vec{v_1}, \vec{v_2}$
          - compute the curvature ratio from the ellipse eccentricity ϵ
        end if
    end for
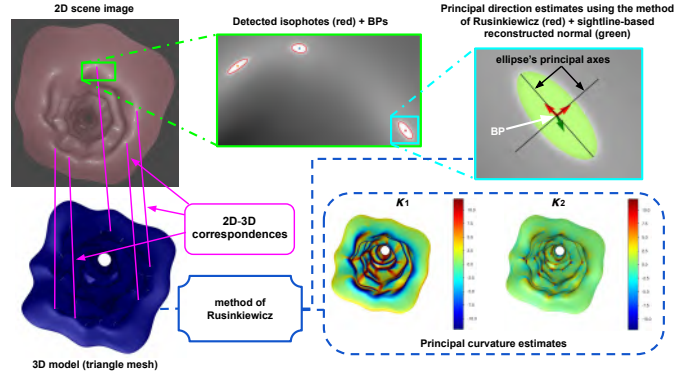
3.1.2. Normal Reconstruction

Our method found 1,500 elliptic specularities. The reconstructed normals using the proposed sightline-based method are compared with the true mesh normals, at mesh vertices corresponding to the set of detected BPs. The histogram of errors in normal estimates is shown in figure 3.

3.1.3. Shape Reconstruction

We estimate the true principal curvatures κ_1 and κ_2 as the eigenvalues of the second fundamental form that we compute numerically [12], from which we obtain the true curvature ratio. We also compute the true principal directions, as the eigenvectors of the second fundamental form. A representative example is shown in figure 2 (top right), showing that the specular ellipse's axes match the principal directions. We perform a sanity check by verifying that the reconstructed normal is orthogonal to the reconstructed principal directions on the tangent plane, confirming that they together form an orthonormal local coordinate system. The histogram of errors in principal direction estimates is shown in figure 4.

3.2. Real Data

The applicability of our reconstruction approach to real data is evaluated in 2D laparoscopic views of the liver, where a

Fig. 2: Colon-like synthetic sequence. The available 2D-3D correspondences (in magenta) are used to compare between our reconstruction of the normal, the principal curvature ratio and the principal directions for the surface at the specularity's brightest point and the same quantities estimated from the 3D model using the method of Rusinkiewicz [12].
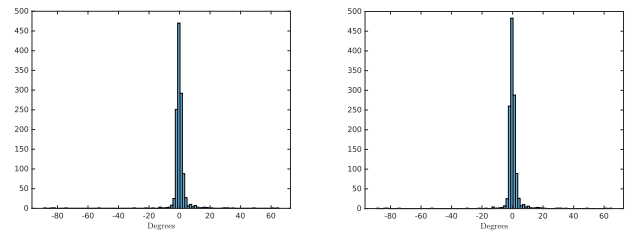
Fig. 3: Error histogram: angle between reconstructed and true normals (in degrees).

Fig. 4: angle between reconstructed normals and $\vec{v_1} \times \vec{v_2}$ (in degrees).

sparse set of specular highlights is detected and for which we have 3D *reference models*. These references models are obtained by 3D-to-2D registration of a virtual patient model segmented from a preoperative CT scan by experts. The registration is solved by the method [13] to a satisfying accuracy. We use three images extracted from three 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 5.

3.2.1. Normal Reconstruction

Contrary to [5] in which a small set of isolated specularities is manually selected to then detect isophotes using the marching squares, we automatically detect numerous specular isophotes from our neural network [6]. An example showing elliptic isophotes is shown in figure 5. A total of 193 small elliptic specularities are kept from the three selected images according to the ellipticity criterion and used to compare the reconstructed normals with two other estimates: 1) the normals produced by the pose from circle method [5], and 2) the normals directly determined from the registered virtual 3D
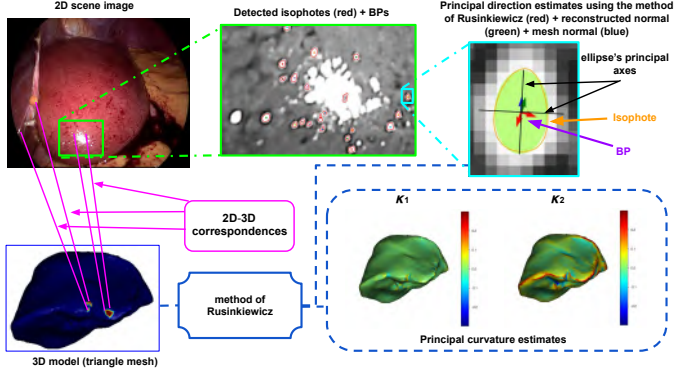
Fig. 5: Qualitative results for the liver data.

model. The histograms of figure 6 show a good agreement between the three methods. The green histogram shows an angular error centred at 2.5°, which confirms the efficiency of both the proposed and the registration methods [14]. The improvement is partly due to the ellipticity, which is not present in [6]. Nevertheless, the magenta histogram shows an angular error centred at 3°, which shows a fair agreement between the proposed and the previous methods.
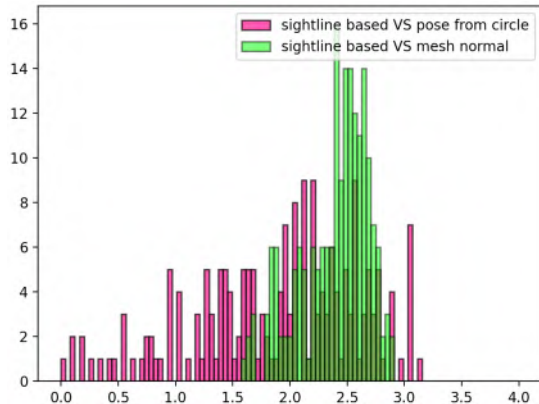

Fig. 6: Evaluation of normal reconstruction for the laparoscopic liver images. The histograms show the absolute angular difference (in degrees).

3.2.2. Principal Directions and Principal Curvature Ratio

First, in order to evaluate the quality of our reconstructed principal directions $\{\vec{v_1}, \vec{v_2}\}$, we compared them to the principal directions $\{\mathbf{u}_1, \mathbf{u}_2\}$ estimated using [12] from the triangle mesh. An easy way to do that is to quantify the angle between their cross products since both of them are supposed to be aligned with the normal direction. Let $\theta_e = \angle(\mathbf{u}_1 \times \mathbf{u}_2, \vec{v_1} \times \vec{v_2})$ be this angle expressed in degrees. We then consider its absolute value $|\theta_e|$ to quantify the degree of agreement between the two methods, leading to the histogram error illustrated in figure 7(a). This histogram shows a small

error which is always lower than 3° for the same specularities as in section 3.2.1 for normal reconstruction evaluation.

Second, we evaluated the differences between the two estimated curvature ratios (proportional to the eccentricity) as $d = ||\lambda_1/\lambda_2| - |\kappa_1/\kappa_2||$ where κ_1 and κ_2 are the minimum and maximum curvatures estimated using [12]. The histogram of figure 7(b) shows that the error was almost comprised between 0 and 0.2 (and mostly close to zero) which is still very acceptable since the eccentricity of an ellipse always lies between 0 (a perfect circle) and 1 (a straight line). The example presented in figure 5 (top right) show that the reconstructed normals and principal directions are reliable.
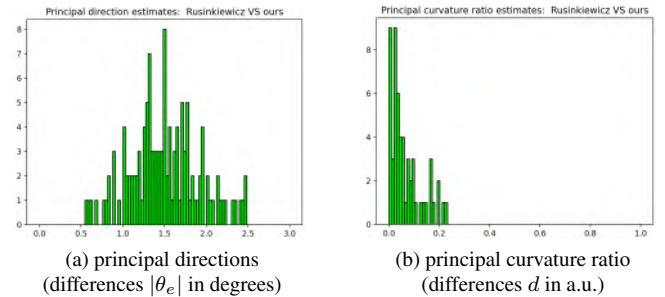

(a) principal directions (differences $|\theta_e|$ in degrees) (b) principal curvature ratio (differences d in a.u.)

Fig. 7: Evaluation of shape reconstruction: direct comparison between our method and [12].

Figure 8 shows qualitative results for a colonoscopic image extracted from the EndoMapper database [15] and for which high quality 3D models are not available.
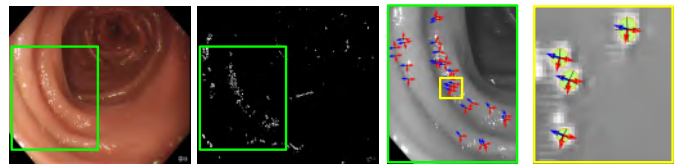

Fig. 8: Colonoscopy. From left to right: image I, neural mask M, the third and fourth images show some reconstructed principal directions (red) and normals (blue).

4. CONCLUSION

We have presented a mathematical model based on a second-order model of the tissue surface at specularities. This model leads to a reconstruction method which can infer the local surface normal, principal directions and curvature ratio. The method is simple and fast. It produces normal estimates with accuracy on par with previous methods, solves the local two-fold ambiguity and estimates an additional level of local shape information. As it works from a single image, we plan to use it to boost the NRSfM method in the endoscopic case.

5. REFERENCES

[1] David A Forsyth and Jean Ponce, *Computer vision: a modern approach*, prentice hall professional technical reference, 2002.

[2] Gastone Ciuti, Marco Visentini-Scarzanella, Alessio Dore, Arianna Menciassi, Paolo Dario, and Guang-Zhong Yang, "Intra-operative monocular 3d reconstruction for image-guided navigation in active locomotion capsule endoscopy," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2012, pp. 768–774.

[3] Yuying Liu and Siyang Zuo, "Self-supervised monocular depth estimation for gastrointestinal endoscopy," *Computer Methods and Programs in Biomedicine*, p. 107619, 2023.

[4] Rema Daher, O León Barbed, Ana C Murillo, Francisco Vasconcelos, and Danail Stoyanov, "Cyclesttn: A learning-based temporal model for specular augmentation in endoscopy," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2023, pp. 570–580.

[5] Karim Makki and Adrien Bartoli, "Normal reconstruction from specularity in the endoscopic setting," in *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2023, pp. 1–5.

[6] Karim Makki, Kilian Chandelon, and Adrien Bartoli, "Elliptical specularity detection in endoscopy with application to normal reconstruction," *International Journal of Computer Assisted Radiology and Surgery*, pp. 1–6, 2023.

[7] Kutsev Bengisu Ozyoruk, Guliz Irem Gokceler, Taylor L Bobrow, Gulfize Coskun, Kagan Incetan, Yasin Almalioglu, Faisal Mahmood, Eva Curto, Luis Perdigoto, Marina Oliveira, 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.

[8] Anita Rau, PJ Eddie Edwards, Omer F Ahmad, Paul Riordan, Mirek Janatka, Laurence B Lovat, and Danail Stoyanov, "Implicit domain adaptation with conditional generative adversarial networks for depth prediction in endoscopy," *International journal of computer assisted radiology and surgery*, vol. 14, pp. 1167–1176, 2019.

[9] James F Blinn, "Models of light reflection for computer synthesized pictures," in *Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, 1977, pp. 192–198.

[10] Brent Burley and Walt Disney Animation Studios, "Physically-based shading at disney," in *Acm Siggraph*. vol. 2012, 2012, vol. 2012, pp. 1–7.

[11] Pushkar Joshi, Mark Meyer, Tony DeRose, Brian Green, and Tom Sanocki, "Harmonic coordinates for character articulation," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 71–es, 2007.

[12] Szymon Rusinkiewicz, "Estimating curvatures and their derivatives on triangle meshes," in *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004*. IEEE, 2004, pp. 486–493.

[13] Yamid Espinel, Lilian Calvet, Karim Botros, Emmanuel Buc, Christophe Tilmant, and Adrien Bartoli, "Using multiple images and contours for deformable 3d–2d registration of a preoperative ct in laparoscopic liver surgery," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 12, pp. 2211–2219, 2022.

[14] Bongjin Koo, Erol Özgür, Bertrand Le Roy, Emmanuel Buc, and Adrien Bartoli, "Deformable registration of a preoperative 3d liver volume to a laparoscopy image using contour and shading cues," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2017, pp. 326–334.

[15] Pablo Azagra, Carlos Sostres, Ángel Ferrández, Luis Riazuelo, Clara Tomasini, O León Barbed, Javier Morlana, David Recasens, Víctor M Batlle, Juan J Gómez-Rodríguez, et al., "Endomapper dataset of complete calibrated endoscopy procedures," *Scientific Data*, vol. 10, no. 1, pp. 671, 2023.