Photometric stereo endoscopy

Vicente Parot
Daryl Lim
Germán González
Giovanni Traverso
Norman S. Nishioka
Benjamin J. Vakoc
Nicholas J. Durr

Photometric stereo endoscopy

Vicente Parot,[a,b] **Daryl Lim,**[a,b] **Germán González,**[a,b] **Giovanni Traverso,**[c,d] **Norman S. Nishioka,**[c] **Benjamin J. Vakoc,**[b] **and Nicholas J. Durr**[a,b]

[a]Madrid-MIT M+Visión Consortium, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
[b]Wellman Center for Photomedicine, Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114
[c]Harvard Medical School, Massachusetts General Hospital, Division of Gastroenterology, Boston, Massachusetts 02114
[d]Massachusetts Institute of Technology, Koch Institute for Integrative Cancer Research, Department of Chemical Engineering, Cambridge, Massachusetts 02139

Abstract. While color video endoscopy has enabled wide-field examination of the gastrointestinal tract, it often misses or incorrectly classifies lesions. Many of these missed lesions exhibit characteristic three-dimensional surface topographies. An endoscopic system that adds topographical measurements to conventional color imagery could therefore increase lesion detection and improve classification accuracy. We introduce photometric stereo endoscopy (PSE), a technique which allows high spatial frequency components of surface topography to be acquired simultaneously with conventional two-dimensional color imagery. We implement this technique in an endoscopic form factor and demonstrate that it can acquire the topography of small features with complex geometries and heterogeneous optical properties. PSE imaging of *ex vivo* human gastrointestinal tissue shows that surface topography measurements enable differentiation of abnormal shapes from surrounding normal tissue. Together, these results confirm that the topographical measurements can be obtained with relatively simple hardware in an endoscopic form factor, and suggest the potential of PSE to improve lesion detection and classification in gastrointestinal imaging. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.JBO.18.7.076017]

Keywords: three-dimensions; endoscopy; tissues; computer vision; stereoscopy.

Paper 130261R received Apr. 18, 2013; revised manuscript received Jun. 14, 2013; accepted for publication Jun. 21, 2013; published online Jul. 17, 2013.

1 Introduction

Many clinically important lesions of the gastrointestinal tract have unique surface topographies. For some lesions, this topographical contrast is more pronounced than their coloration. However, tissue coloration is the primary source of contrast in conventional endoscopic cameras. In fact, endoscope camera lighting strategies minimize rather than accentuate topographical contrast. For lesions with pronounced topographical features, it may be possible to improve detectability and classification accuracy by modifying the imaging hardware to enable acquisition of both conventional color-based images and surface topographical data.

Colon cancer screening by colonoscopy is an obvious candidate for topographical imaging. Precancerous lesions of the colon (polyps) can be difficult to detect using conventional color imagery. These lesions, however, can protrude (or be recessed) significantly relative to the surrounding mucosa.[1,2] Conventional colonoscopy relying primarily on coloration is estimated to miss one out of every four lesions.[3,4] A method to provide endoscopic measurement of surface topography paired with conventional color imagery could reduce this high missed lesion rate. Specific evidence for the value of topographical contrast in colon cancer screening comes from both computed tomography colonoscopy (CTC) and chromoendoscopy studies. In CTC, which relies entirely on topographical contrast, multiple studies have reported high lesion detection sensitivity.[5,6] In chromoendoscopy, surface dyes that highlight topography are used, and

lesion detection rates have been documented to increase.[7,8] Despite their recognized advantages, CTC and chromoendoscopy are infrequently used due to the increased cost associated with follow-up colonoscopy for CTC and the significant time required for chromoendoscopy.

Here, we introduce photometric stereo endoscopy (PSE), an optical approach to topographical imaging that can be implemented in a straightforward manner with conventional endoscopic cameras. PSE is based on photometric stereo (PS) concepts established in computer vision. In PS, surface normals of an imaged object are calculated by comparing a set of images acquired from the same camera but with differing illumination conditions.[9,10] If one assumes a Lambertian remission of the light, i.e., isotropic scattering, the surface normal of each pixel can be calculated by solving a system of linear equations relating the intensities of that pixel from each source. By integrating the associated gradients, the three-dimensional (3-D) topography of the field of view (FOV) can be reconstructed. Most endoscopes already include multiple light sources, but these light sources are used simultaneously to reduce shadowing and increase ambient luminosity and color contrast. In principle, with modifications only to light source cycling and camera frame synchronization, the traditional endoscope could be modified to acquire images with unique illuminations conditions, and PS algorithms could be applied to these images to extract surface topography.

Unfortunately, conventional PS algorithms assume a narrow-angle FOV wherein the directional vectors from each object pixel to each light source are known (a vector field which changes with every movement of the object relative to the sources). In PS, this is achieved by placing the light sources far away from the sample so that the directional vectors are

Address all correspondence to: Nicholas J. Durr, Madrid-MIT M+Visión Consortium, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139. Tel: 617-324-4227; Fax: 617-643-9208; E-mail: ndurr@mit.edu

approximately constant, a solution that is incompatible with endoscopic imaging. Recent work toward 3-D monocular laparoscopy[11] has addressed this problem by performing an off-line calibration of the illumination system, but still requires a calibration target to be placed on the object at the time of imaging. Here, we demonstrate PSE, an endoscopic implementation of PS that employs additional algorithms to mitigate errors induced by using more compactly positioned light sources, and does not require a calibration target in the FOV. Specifically, in PSE, low spatial frequency components of measured gradients are rejected, and surface orientation maps are derived based on higher spatial frequency topographical measurements that are less sensitive to light placement.

In a series of experiments evaluating PSE, we demonstrate a qualitative agreement between the acquired high-frequency morphology and the shape of the imaged object. For these purposes, we have used a controlled benchtop prototype, with which we show how the morphology reconstruction performs with complex geometries present in an anatomical colon phantom. In another test with the same prototype, we assess the robustness of the technique when the object has heterogeneous optical properties, imaging a variety of *ex vivo* human gastrointestinal lesions. Finally, to confirm the scalability of PSE to an endoscopic size system, we demonstrate reconstructing the topography of a simple phantom using a modified commercial endoscope (with some implementation limitations) and compare it with the benchtop prototype imaging of the same phantom.

2 Methods

PSE calculates the surface normal for each pixel in an image from a set of images acquired with different lighting conditions. In PSE, the direction normal to the surface is represented by $\hat{n}$, the vector $\hat{s}_i$ represents the direction to light source i, and the image intensity is proportional to $\cos \theta = \hat{n} \cdot \hat{s}_i$ [Fig. 1(a)]. We implemented PSE in two systems: a benchtop prototype and a modified commercial endoscope. The benchtop prototype [Fig. 1(b)] allowed more flexible control over illumination and image capture conditions. In this system, we were able to access raw image data from our sensor, synchronize source

illumination with the frame rate, and introduce cross-polarizers to reduce specular reflections. However, the minimum source separation we could achieve with this system was 35 mm. By comparison, the distal tip of typical commercial colonoscopes ranges in diameter from 11 to 14 mm. To address this limitation, we implemented a second PSE setup using a commercial gastroscope. External light sources were attached around the scope housing [Fig. 1(c)]. In this system, we were able to reduce source separation to 12 mm. However, because the interface between the commercial camera sensor and the processing console was inaccessible, images were obtained after secondary processing by the console. In addition, the miniaturized size of the endoscope tip, combined with its wide-angle FOV and flat front design made it challenging to add polarizer layers. For this reason, we did not include cross-polarizers to reduce specular reflections. With the combination of these two setups, we were able to first analyze PSE deployments using optimized imaging hardware and transparent data processing but unrealistic size scales (benchtop prototype), and second confirm that these results are maintained qualitatively as the device scales to a more realistic size (modified endoscope).

2.1 Benchtop Prototype

Our benchtop PSE system was constructed with four light sources mounted around a camera with a fish-eye lens [Fig. 1(b)]. The size of the housing was 30×30 mm^2, and the four sources were oriented at equal angles about a circle with a 35 mm diameter. We used a dragonfly®2 remote head camera with a 1/3 in color, 12 bit, 1032×776 pixel CCD (Point Grey Research, Inc., Richmond, BC, Canada). The image was created with a 145 deg FOV board lens (PT-02120, M12 Lenses). The FOV of this system was partially occluded by the imaging head. We used white LEDs for illumination (Mightex FCS-0000-000), coupled to 1 mm diameter, 0.48 NA multimode fibers. Sources were synchronized to the camera frame rate of 15 Hz. A holographic light shaping diffuser was placed at the end of each source to efficiently spread illumination light (80 deg Circular Diffuser, Luminit). Linear
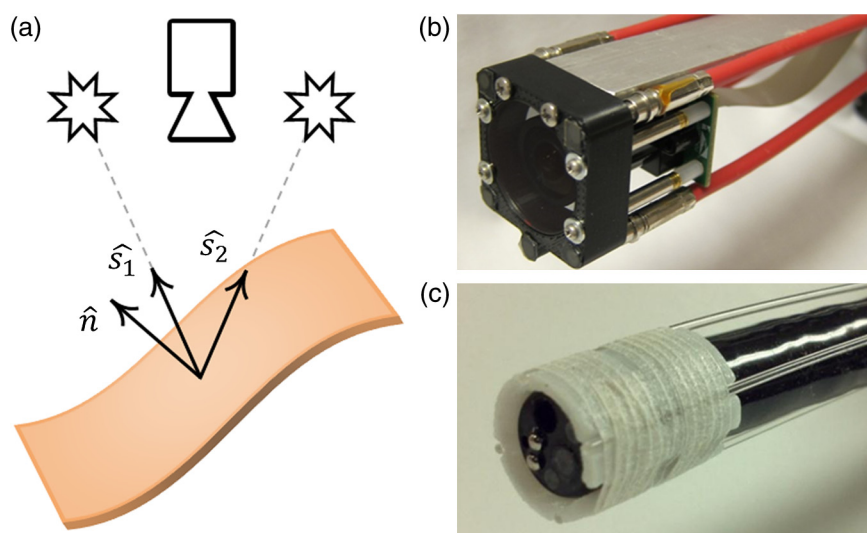

Fig. 1 PSE is an endoscopic deployment of photometric stereo imaging. Using PSE, the object surface normal map can be reconstructed by comparing images of a sample taken under different illuminations. (a) In photometric stereo, the direction normal to the surface is represented by $\hat{n}$, vector $\hat{s}_i$ represents the direction to light source i, and the pixel intensity is proportional to $\cos \theta = \hat{n} \cdot \hat{s}_i$ (b) Photo of our 30×30 mm PSE prototype. (c) Photo of our 14 mm outer diameter PSE prototype.

polarizers were placed in front of the sources and objective lens in a cross-configuration to minimize specular reflection. Images in raw data format were processed with a de-mosaicing interpolation algorithm to provide full resolution RGB images from Bayer-patterned raw images. The pixel intensities were then estimated by a weighted average of the three color channels.

2.2 Modified Commercial Endoscope

To generate images with smaller light, camera, and object separations, we modified a commercial endoscope to use added light sources [Fig. 1(c)]. We used a Pentax EG-2990K gastroscope with a Pentax EPK-1000 video processor. For illumination, we used fibers with an integrated light diffuser (NA = 0.5, $D = 1$ mm, Doric Lenses Inc.), and no polarization filters. The four fibers were secured at equal angles in a 12-mm diameter circle around the endoscope tip, resulting in an external diameter of 14 mm. Components were held together with tape wrapping. We acquired uncompressed video in NTSC format at 8 bit, 720×486 pixel resolution, and 29.97 interlaced frames per second using a video capturing device (Blackmagic Intensity Shuttle). Light sources were alternated at 60 Hz, synchronized with the video signal to deinterlace a sequence of RGB frames captured with only one light source active at a time. Frames were then interpolated in every other horizontal line to obtain full resolution images. The image intensity was estimated as the weighted average of the three color channels.

2.3 PSE Algorithm

Our algorithm relies on the approximation that the light remitted from our sample surface follows the Lambertian reflectance model. For each pixel, the apparent brightness or luminous intensity (I) is determined by the surface irradiance (I_0), the cosine of the angle between the unit vector normal to the surface ($\hat{n}$) and the unit vector indicating the direction to the illumination source ($\hat{s}$), and a constant amplification factor (A) that includes the reflection coefficient or albedo:

$$I = A \cdot I_0 \cdot \hat{s} \cdot \hat{n}. \qquad (1)$$

Neglecting cast shadows and specular reflections, with the light source i coming from the direction $\hat{s}_i$, the measured intensity m_i at one pixel can then be represented as

$$m_i = \hat{s}_i \cdot \vec{n}, \qquad (2)$$

where $\vec{n} = [n_x, n_y, n_z]^T$ is a nonunitary vector with magnitude $A \cdot I_0$ and direction $\hat{n}$. An example sequence of images under different illumination is shown in Fig. 2(a)–2(d).

Under PS assumptions, a sequence of three measurements of the same sample can be written as

$$\begin{bmatrix} s_{x1} & s_{y1} & s_{z1} \\ s_{x2} & s_{y2} & s_{z2} \\ s_{x3} & s_{y3} & s_{z3} \end{bmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}. \qquad (3)$$

This is a linear system of equations that can be solved for $\vec{n}$ if the light sources matrix is nonsingular. This condition is equivalent to requiring that the three light sources and the sample do not lie in the same plane. If more than three measurements are acquired, the normal vector can be estimated by minimizing the residual error given the measurements and the source directions:

$$\arg\min_{n_x, n_y, n_z} \sqrt{\sum_i (\hat{s}_i \cdot \vec{n} - m_i)^2}. \qquad (4)$$

Our PSE approach uses the conventional framework of PS imaging as described above, but includes an additional filtering of lower spatial frequency components of the surface normal data. This is necessitated by close placement of the sources, camera, and imaging object. Specifically, the traditional PS assumption that $\hat{s}_i$ is constant for all pixels in the image becomes especially inaccurate when the FOV is large relative to the distance from the object to the light sources. We empirically found this effect to induce only a low frequency error on the reconstructed normals, on the order of 1 cycle/FOV. This result is expected, given that the source directions change slowly across the FOV. As the motivation for PSE is to obtain useful information about the lesions and texture present in an endoscopic setting, which are often high frequency topographies, we can preserve these features while filtering out the inaccurate information by applying a high-pass filter to our reconstructed normal map.

Assuming a continuous sample that can be described as $z = f(u, v)$ with z the distance from the objective to the sample and (u, v) the pixel coordinates, its directional gradients can be obtained by scaling the direction normal to the surface:
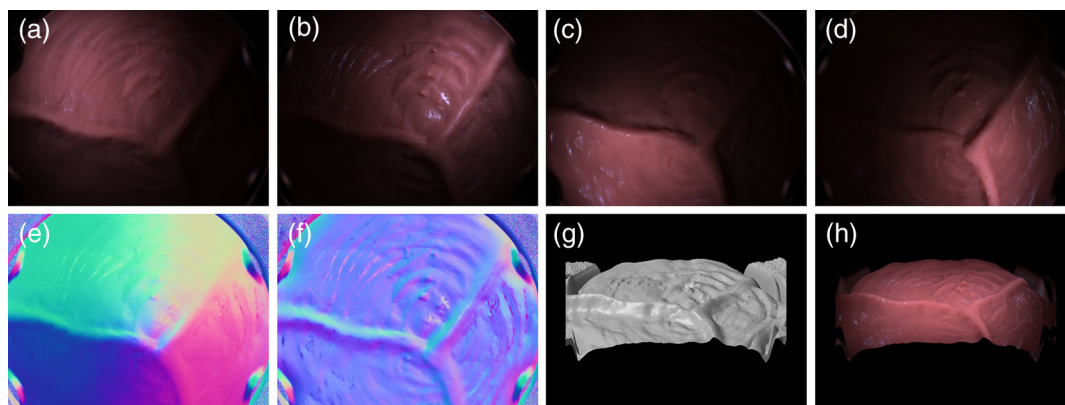

Fig. 2 Example of PSE workflow. The algorithm includes acquiring a series of images of the same FOV under different illumination (a–d), calculating the surface orientations using conventional photometric stereo, which results in severe topography artifacts (e), and performing additional spatial frequency filters (f). The resulting topography can be integrated and 3-D rendered without texture (g) and with textured color (h). Surface orientation maps in (e) and (f) show color coded RGB maps, where normalized x, y, and z components in the interval ($-1 \ldots 1$) have been mapped to R, G, and B values in the interval ($0 \ldots 1$).

$$\left[-\frac{\partial f}{\partial u} \quad -\frac{\partial f}{\partial v} \quad 1 \right]^T = \frac{\vec{n}}{n_z} = \left[\frac{n_x}{n_z} \quad \frac{n_y}{n_z} \quad 1 \right]^T. \quad (5)$$

Because both the spatial frequency filter and the differentiation are linear operations on (u, v), these operations are interchangeable, and the high-pass filter of $-\partial f / \partial u$ is equivalent to the gradient in direction u of the high-passed surface. Thus, by high-pass filtering each of the directional gradients, we can obtain the gradients of the high frequencies of the shape. For each directional gradient, we applied a high-pass filter by subtracting the low frequency component of the signal, which was calculated as a convolution of the original gradient with a Gaussian kernel with $\sigma = 40$ pixels in image space. This kernel's full width at half maximum value was approximately 8 cycles/FOV. Applying this filter to the normal map in Fig. 2(e) results in the normal map as shown in Fig. 2(f). In this work, we represent normal maps using a color coded RGB image, as is commonly used in computer vision,[12] where the surface normal vector is normalized and its x, y, and z components are mapped to values of red, green, and blue, respectively.

To calculate height maps, the filtered gradients were integrated using a multigrid solver[13] for the Poisson equation[14] that minimizes integration inconsistency errors [Fig. 2(g)]. To visualize both the color information and the acquired topography, we can overlay the color image on the calculated height map [Fig. 2(h)].

2.4 Tissue Models

2.4.1 Anatomical phantom

We tested the performance of PSE imaging in a tubular environment using a silicon colon phantom (Colonoscopy Trainer, The Chamberlain Group). This phantom has been used in a previous study for investigating lesion detection rates in colonoscopy.[15] The overall shape of the colon including curvature and haustra is represented in the phantom. Fabrication artifacts provide features comparable in size to subtle colon lesions. The material has a homogeneous color, and the surface is smooth and shiny. This model emulates the geometry of the colonoscopy environment to evaluate effects such as the tubular shape, wide FOV, cast shadows, varying working distance, and nonuniform illumination.

2.4.2 Ex vivo human tissue

Specimens from gastrointestinal tissue resections (for any indication) were identified in the Department of Pathology, Massachusetts General Hospital. Specimens with abnormalities were selected for imaging. All tissue samples were imaged within 24 h of resection, either fresh or after preservation in formalin for less than 24 h. The protocol was reviewed and approved by the Partners Institutional Review Board.

2.4.3 3-D printed phantom

PSE signal relies on intensity variation due to illumination from different light source positions. Intuitively, if the sources are moved closer together, there will be less intensity variation between images taken with different sources, and the signal-to-noise ratio (SNR) of the surface normal estimation will decrease. To evaluate the performance of PSE with a light source separation and working distance realistic for endoscopic purposes, we tested the ability of both the benchtop prototype and the modified endoscope to capture the morphology of a diminutive bump. We used a 3-D printed phantom with a 1-mm tall, 0.5-mm radius bump that was created using stereolithography (Quickparts.com). This phantom represents a challenging lesion topology—flat lesions, which exhibit a particularly subtle height change compared to the surrounding tissue, have typical elevations of 2.2 mm.[16] The sample was painted with pink tempera paint to reduce specular reflection. The working distance was set to 35 mm in the modified endoscope and adjusted to 87 mm in direct proportion to the light source separation for the benchtop prototype.

3 Results and Discussion

3.1 PSE Imaging of Anatomic Shapes: Phantom Imaged by Benchtop Prototype

We performed PSE imaging on several regions of the silicon anatomical phantom with the benchtop prototype. The expected surface normals were recovered across the FOV, as shown in the frontal view of the cecal wall presented in Fig. 3(a). The flat regions of the cecum generated regions with normal components pointing primarily in the z-direction, and bumps and ridges yielded normals such that they were correctly reconstructed after integration. We emphasize that the topographical data presented in the surface normal map and the 3-D rendering are complementary to the color information in the conventional image—this topography cannot be reconstructed from the conventional image alone. Three diminutive bumps that range in height from 0.5 to 1 mm were correctly identified as elevations in our reconstruction [white arrows, Fig. 3(a)], though it is difficult to appreciate based on the conventional color image alone.

The illumination intensity that reached the sample from the light sources was strongly affected by the working distance, which varied significantly within the FOV. For example, when imaged down a tubular shape, the pixels in the center of the image received much less light than those at the periphery. However, accurate normal reconstruction in PSE relies on intensity differences for each pixel in a series of images, and lighting changes that are consistent across the PSE image series should primarily affect the SNR of the measurements rather than introduce topographical bias. This was demonstrated in a PSE image of the transverse colon presented in Fig. 3(b). Though the light intensity reaching the surface down the tube was much lower than that illuminating the adjacent wall, the high frequency details in surface orientations of the object were still acquired.

There are several sources of topographical error in PSE that become amplified when imaging in a tubular geometry. The assumption that the source vectors are constant across the FOV becomes worse as the distance between each point in the object and the light source changes. Furthermore, any portion of the object that casts a shadow creates a nonlinear artifact; the region that is cast in shadow is reconstructed to have a surface normal that points more perpendicularly than it should to the direction of the light source that shadows the region. This artifact exaggerates slopes facing away from the lights. Qualitatively, this effect emphasizes ridges and sharp features, which may actually be helpful for the purpose of increasing lesion contrast. This effect was observed in the shadows cast by the muscular features and haustra of the colon phantom, shown in Fig. 3(b).
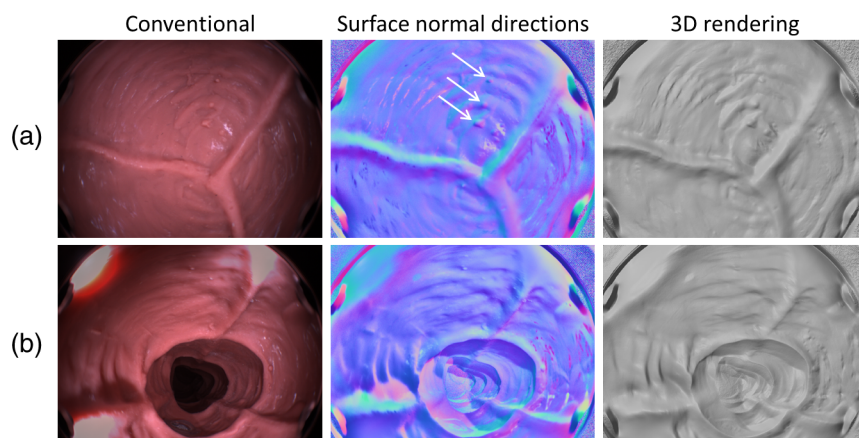
Fig. 3 (a) Conventional view, PSE normal map, and topography estimation in a silicon colon phantom. The surface normal directions and 3-D rendering of a cecum view captures the orientation of slopes and curvature, which are not contained in a conventional color image. Three diminutive bumps that are 0.5 to 1.0 mm in height are correctly identified as elevations on the normal map (white arrows) (Video 1, QuickTime, 6.3 MB) [URL: http://dx.doi.org/10.1117/1.JBO.18.7.076017.1]. (b) In a tubular region of the transverse colon, the high-frequency morphology shows the details of features at different working distances contained in the FOV. Cast shadow artifacts consistently exaggerate slopes from the feature generating the shadow.

3.2 PSE Imaging of Tissue Surfaces: Ex Vivo Human Tissue Imaged by Benchtop Prototype

We performed PSE on *ex vivo* human gastrointestinal tissue in order to evaluate performance on samples with heterogeneous optical properties, reflective surfaces, and clinically relevant lesions. Figure 4(a) presents the topography obtained from a right colectomy with a tattoo applied next to a benign ulcer that resulted from a polypectomy. Here, our normal map correlated to the visible folds in the conventional image. The ulcer, identified by a gastroenterologist at the time of the imaging, was reconstructed as a prominent indentation in the tissue, indicated with a dashed arrow in Fig. 4(a). However, the tattoo, which left

a concentrated point of indigo color at the site of the injection, did not register as a topographical change (solid arrow). This demonstrated that the PSE was able to separate a pixel's surface normal vector from its albedo.

Next, we imaged a benign precancerous sessile serrated polyp that was identified after a right colectomy [Fig. 4(b)]. In this experiment, the light source in the bottom right of the FOV did not diffuse as well as the other three sources. As a result, the image with this light source saturated the bottom right of the FOV and the topographies were poorly reconstructed in that region. Nonetheless, the sessile lesion encircled with a dashed oval in Fig. 4(b) clearly influenced the normal map. In the
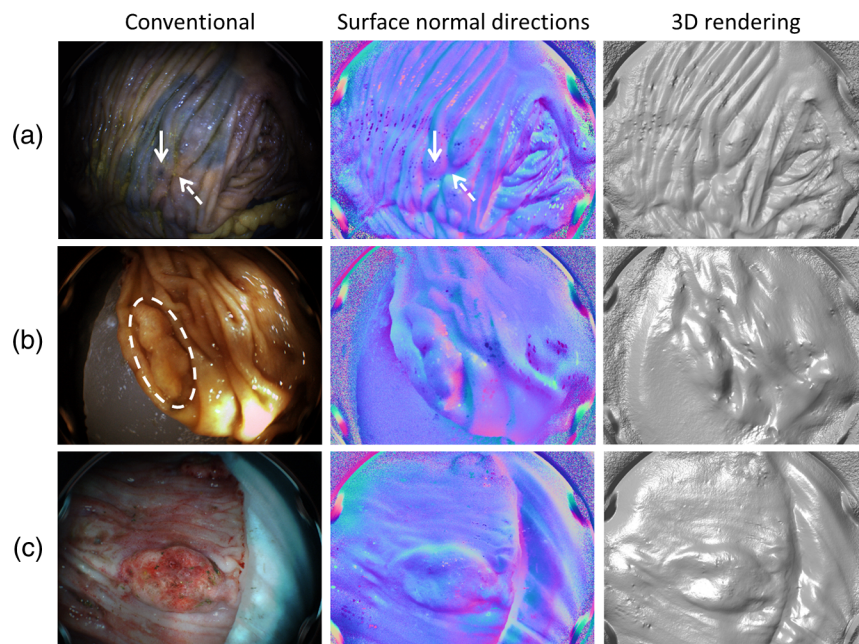

Fig. 4 PSE morphology estimation of *ex vivo* human tissue with heterogeneous optical properties. (a) The reconstructed morphology of a post-polypectomy benign ulcer and surrounding tissue folds in formalin fixed colon tissue correlates with the folds that are visible in the conventional image (Video 2, QuickTime, 8.1 MB) [URL: http://dx.doi.org/10.1117/1.JBO.18.7.076017.2]. As expected, an injection site with remnant ink (solid arrow) produces color contrast but not morphologic contrast, while the ulcer (dashed arrow) produces both color and morphologic contrast. The plateau shape of a benign precancerous sessile serrated polyp in the fixed *ex vivo* right colon tissue (b), and a metastatic melanoma lesion in fresh *ex vivo* small bowel tissue (c) are both prominent in the estimated morphology.
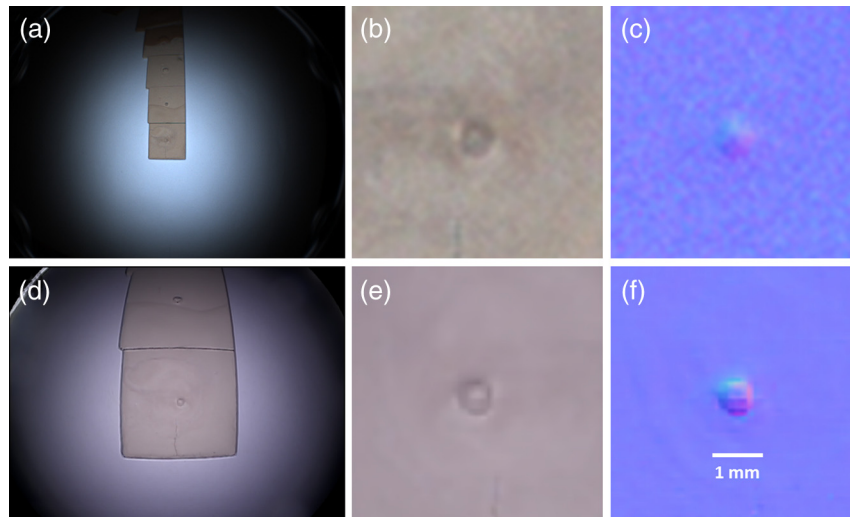
Fig. 5 PSE enabled topographical identification of a diminutive bump with the benchtop prototype at an 87 mm working distance [panels (a) to (c)] and the modified endoscope at a 35 mm working distance [panels (d) to (f)]. (a) and (d) show the full FOV of each system. The conventional images in (b) and (e) were insufficient to discriminate the feature as an elevation or a depression, while their morphology was revealed in the surface orientations (c) and (f). Insets (b), (c), (e), and (f) cover a 5×5 mm^2 region of interest.

surface rendering that was generated from the normal map, the lesion had a plateau-like topography that is characteristic of a sessile lesion, and that was observed during this experiment.

Finally, we imaged a metastatic melanoma that was present in fresh *ex vivo* human small bowel tissue [Fig. 4(c)]. This feature was also identifiable in the normal map and reconstructed height profile. Note again that here PSE was able to distinguish between color changes of the tissue and actual folds that were present in the sample.

Because the *ex vivo* human tissue was wet, specular reflection was more prominent than what was observed in the silicon phantom. This led to artifacts in our surface normal reconstructions. Specifically, pixels that had specular reflections were reconstructed to have surface normals that pointed more toward the source that generated the specular reflection than they actually should.

3.3 Translation of PSE Imaging to Endoscopic Size: 3-D Printed Phantom Imaged by Modified Endoscope

These previous results show the ability of PSE to obtain high frequency topography in geometrical settings where traditional PS assumptions do not hold, and demonstrate that PSE topography records changes in surface orientation of tissues with heterogeneous optical properties. However, these results have been obtained using a light source separation distance larger than what fits in an endoscope.

To qualitatively confirm the scalability of PSE to endoscopic size, we compared results of imaging a simple object with the benchtop prototype and the modified endoscope. Because our modified endoscope setup was sensitive to specular reflections and had an inaccessible video digitizing system, we used a non-specular 3-D printed phantom with a single bump and compared with PSE reconstructions from our two prototypes. PSE consistently estimated the morphology of bumps and depressions with both systems. In all estimations, the surface normal directions correctly showed the elevation or depression as a region in which border surfaces were oriented outwards for elevations and inwards for depressions.

When using a small source separation of 12 mm and a large working distance of 35 mm with the modified endoscope, PSE was still able to recover surface orientations of a small bump. Figure 5 shows results comparing the performance of PSE in the benchtop prototype at an 87 mm working distance [Fig. 5(a)–5(c)] and the modified endoscope at a 35 mm working distance [Fig. 5(d)–5(f)]. Figure 5(a) and 5(d) display the full FOV of each system, while the remaining insets show a 5×5 mm^2 region of interest. The conventional images [Fig. 5(b) and 5(e)] were insufficient to discriminate the feature as an elevation or a depression, while their morphology was revealed in the surface orientations [Fig. 5(c) and 5(f)]. As noticeable in Fig. 5(c), the topography imaged at 87 mm distance resulted in an increased level of noise, possibly due to the low illumination intensity at that distance and a lower number of pixels covered by the 5×5 mm^2 area.

Noticeable artifacts present in these estimations included measurement noise, gradients amplitude scaling, discretization of the shape, and albedo variations. The albedo nonuniformities were caused by an uneven layer of paint and differences in surface orientation. The amplitude of the gradients decreased with the working distance. Discretization of object was noticeable and is attributable to the small portion of the FOV that the sample covered. For example, a 1-mm wide feature imaged at a 35 mm working distance covered approximately 9 pixels across the images acquired with the modified endoscope.

4 Conclusions

Here, we introduced PSE as a new endoscopic technique to simultaneously acquire both a color image and the associated high frequency topography of a FOV. We demonstrated the utility of PSE in samples with complex geometries, including tubular environments. PSE is also able to reconstruct normal maps that are correlated to color images in *ex vivo* human tissues with heterogeneous optical properties. This demonstrates the power of the technique to separate a pixel's surface normal vector from its albedo. We observed that very fine folds [such as those

present in Fig. 4(c)] are sometimes missed during reconstruction. These artifacts can be caused by deep, sharp folds in the tissue, where shadows are generated from multiple light sources, or by a poor signal to noise in the normal reconstruction. In both cases, a random error is introduced and the resulting reconstructed normal and detailed topography can be lost. Our technique also suffered from artifacts from specular reflection because our model assumes purely Lambertian remittance. More sophisticated reconstruction algorithms can actually use this specular information for more accurate normal map reconstructions.[17] Furthermore, implementing the technique with a higher resolution sensor, such as an HD endoscope, would significantly increase the ability of PSE to capture fine topographical detail.

We demonstrate that PSE can reconstruct normal maps of diminutive structures in geometries relevant for clinical endoscopy. The ability of PSE to reconstruct these normal maps is related to the difference in intensity that is registered for each pixel as it is illuminated from different light sources. Thus, if the light sources are moved closer together, the illumination of each pixel becomes more similar, and the normal reconstruction decreases in signal to noise. This is precisely what happens as the working distance is increased. We found that even with a low resolution image sensor, the signal to noise in the normal reconstruction was still sufficient to register topographical changes from a 1-mm tall bump at working distances of up to 40 mm. At this distance, the power from our light sources limits the ability to image. In practice, most lesions are identified during the withdrawal of the endoscope, where the new field appears at the periphery of the image.[18] Thus, in clinical situations, the endoscopist is typically examining regions that are significantly closer than 40 mm from the endoscope tip.

The practical implementation of 3-D imaging in an endoscopy system is an active field of research. Previous approaches include dual lens stereo vision, time-of-flight imaging, and structured illumination, among other techniques. Detailed reviews of these methods can be found in Refs. 19 and 20. Our technique has advantages over many of these alternatives in that it is simple to implement in an existing clinical endoscope. However, because PSE utilizes a high pass spatial frequency filter, our technique has the disadvantage of being qualitative in height reconstruction. PSE can be made more quantitative by combining it with an alternative approach such as structured illumination, or by using a calibration target in the field.[11] As the topography representation from PSE is limited to high frequencies, other 3-D imaging techniques may be more appropriate for applications that seek quantitative measurement of shape, intraoperative navigation, or volume reconstruction. The attractive value of PSE is its potential to increase detectability of clinically relevant lesions by capturing high-frequency morphology over a wide FOV in real time, while minimizing changes to existing endoscope technology.

With few changes to a commercial endoscope, PSE can be implemented to provide new information that is currently unavailable in conventional endoscopy. We have shown that this topography can be visualized as normal maps or renderings, but there are other possible use models for this data. The additional information provided in the normal maps may lead to better computer-aided detection algorithms for automatic lesion finding. PSE may also be useful for improved mapping of large regions of a sample (mosaicing), and generating novel morphology-based image enhancements (e.g., virtual chromoendoscopy).

References

1. D. A. Lieberman, "Screening for colorectal cancer," *N. Engl. J. Med.* **361**(12), 1179–1187 (2009).
2. R. M. Soetikno et al., "Prevalence of nonpolypoid (flat and depressed) colorectal neoplasms in asymptomatic and symptomatic adults," *JAMA* **299**(9), 1027–1035 (2008).
3. D. K. Rex, "Maximizing detection of adenomas and cancers during colonoscopy," *Am. J. Gastroenterol.* **101**(12), 2866–2877 (2006).
4. D. Rex et al., "Colonoscopic miss rates of adenomas determined by back-to-back colonoscopies," *Gastroenterology* **112**(1), 24–28 (1997).
5. R. E. van Gelder et al., "Computed tomographic colonography compared with colonoscopy in patients at increased risk for colorectal cancer," *Gastroenterology* **127**(1), 41–48 (2004).
6. H. Yoshida and J. Nappi, "Three-dimensional computer-aided diagnosis scheme for detection of colonic polyps," *IEEE Trans. Med. Imag.* **20**(12), 1261–1274 (2001).
7. J. C. Brooker et al., "Total colonic dye-spray increases the detection of diminutive adenomas during routine colonoscopy: a randomized controlled trial," *Gastrointest. Endosc.* **56**(3), 333–338 (2002).
8. D. P. Hurlstone et al., "Detecting diminutive colorectal lesions at colonoscopy: a randomised controlled trial of pan-colonic versus targeted chromoscopy," *Gut* **53**(3), 376–380 (2004).
9. R. J. Woodham, "Photometric stereo: a reflectance map technique for determining surface orientation from image intensity," *Proc. SPIE* **0155**, 136–143 (1979).
10. V. Argyriou and M. Petrou, "Photometric stereo: an overview," in *Advances in Imaging and Electron Physics*, P. W. Hawkes, Ed., Vol. 156, pp. 1–54, Elsevier, New York (2009).
11. T. Collins and A. Bartoli, "3D reconstruction in laparoscopy with close-range photometric stereo," in *Medical Image Computing, and Computer-Assisted Intervention (MICCAI 2012)*, A. Nicholas et al., Eds., pp. 634–642, Springer Berlin Heidelberg, Germany (2012).
12. T.-P. Wu and C.-K. Tang, "Photometric stereo via expectation maximization," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(3), 546–560 (2010).
13. G. Farneback et al., "Efficient computation of the inverse gradient on irregular domains," in *IEEE 11th Int. Conf. Computer Vision (ICCV 2007)*, pp. 1–8 (2007).
14. T. Simchony, R. Chellappa, and M. Shao, "Direct analytical methods for solving Poisson equations in computer vision problems," *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(5), 435–446 (1990).
15. G. Triadafilopoulos et al., "A novel retrograde-viewing auxiliary imaging device (third eye retroscope) improves the detection of simulated polyps in anatomic models of the colon," *Gastrointest. Endosc.* **65**(1), 139–144 (2007).
16. P. J. Pickhardt, D. H. Kim, and J. B. Robbins, "Flat (nonpolypoid) colorectal lesions identified at CT colonography in a U.S. screening population," *Acad. Radiol.* **17**(6), 784–790 (2010).
17. S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(10), 1239–1252 (2003).
18. J. D. Waye, D. K. Rex, and C. B. Williams, Eds., *Colonoscopy: Principles and Practice*, 1st ed., Wiley-Blackwell, London, United Kingdom (2003).
19. D. J. Mirota, M. Ishii, and G. D. Hager, "Vision-based navigation in image-guided interventions," *Annu. Rev. Biomed. Eng.* **13**(1), 297–319 (2011).
20. P. Mountney, D. Stoyanov, and G.-Z. Yang, "Three-dimensional tissue deformation recovery and tracking," *IEEE Signal Process. Mag.* **27**(4), 14–24 (2010).