Shape-Enhanced Surgical Visualizations and Medical Illustrations with Multi-Flash Imaging

Kar-Han Tan[1], James Kobler[2], Rogerio S. Feris[3], Paul Dietz[1], and Ramesh Raskar[1]

[1] Mitsubishi Electric Research Laboratories, Cambridge, MA, USA
[2] H. P. Mosher Laryngological Research Laboratory,
Massachusetts Eye and Ear Infirmary, Boston, MA, USA
[3] University of California at Santa Barbara, CA, USA

Abstract. We present a novel approach for enhancing images and video used in endoscopic surgery so that they are better able to convey shape. Our method is based on *multi-flash imaging*, in which multiple light sources are strategically positioned to cast shadows along depth discontinuities. We describe designs for achieving multi-flash imaging using multiple endoscopes as well as in single endoscopes. Multi-flash photography can also be used for creating medical illustrations. By highlighting the detected edges, suppressing unnecessary details, or combining features from multiple images, the resulting images convey more clearly the 3D structure of the subject. The method is easy to implement both in software and hardware, and can operate in realtime.

1 Introduction

In many medical applications like minimal-invasive surgery with endoscopes it is often difficult to capture images that convey the 3D shape of the organs and tissues being examined [1]. Perhaps for the same reason, medical textbooks and articles frequently resort to hand drawn illustrations when depicting organs and tissues. In this paper we propose the use of multi-flash imaging to address this problem.

A multi-flash imaging system captures additional shape information compared to traditional cameras and therefore has the potential to enhance visualization and documentation in surgery and pathology. The raw shadowed images can be processed to create finely detailed images that are comparable to medical illustrations or they can be used to enhance edge features for quantitative measurement. Alternatively, the shadowed images can be combined to generate shadowless images, which are often desirable for documentation of specimens in the field of pathology.

Surgery. Most endoscopic procedures are now performed with the surgeon observing monitor displays rather than the actual tissue. This affords the possibility of interposing image manipulation steps which, if they can run in close to

real time, can enhance the surgeon's perception. Depth perception is an obvious deficit when using monocular endoscopes. Three-dimensional imaging using stereoscopic methods has been explored with mixed results. A 1999 study found that stereoendoscopic viewing was actually more taxing on the surgeons than monocular viewing [2]. Structured lighting is also under investigation as a means for calibrating endoscopic images [3, 4], but this technique does not provide real-time enhancement of 3-D structures. Application of enhanced shadow information to augment surgical perception has not been exploited previously. Shadows normally provide clues about shape, but with the circumferential ("ringlight") illumination provided by most laparoscopes, this information is diminished. Similarly, the intense multisource lighting used for open procedures tends to reduce strong shadow effects. Loss of shadow information may make it difficult to appreciate the shapes and boundaries of structures and thus more difficult to estimate their extent and size. It may also make it more difficult to spot a small protrusion, such as an intestinal polyp, if there are no clear color differences to set it apart. The ability to enhance the borders of lesions so that they can be measured will become more useful as endoscopes begin to incorporate calibrated sizing features.

Pathology. Documentation of surgical and autopsy specimens is an important service provided by pathology departments. Systems for photography of such specimens involve special methods for eliminating unwanted shadows, usually by placing the specimens on glass plates suspended over black cavities. Using the multiflash system and processing to obtain the MAX composite image produces a view in which almost all shadows are eliminated.

Medical and Biological Illustration. Often it is desirable to generate black and white illustrations of medical and natural history specimens in which salient details are emphasized and unnecessary clutter is omitted [5]. The most important details to emphasize are those that convey the shape of the object. Typically shape is conveyed by emphasizing edges and using stippling for shading. This type of illustration is seen less frequently nowadays because of the expense involved in having artists create these graphics. With multi-flash imaging we found that images very similar to this kind of illustration can be generated quickly and easily. The medical or biological illustrator could use these images as the basis for their graphics.

2 Related Work

The creation of stylized renderings from images without 3D model reconstruction has recently received a great deal of attention. The majority of the available techniques for image stylization involve **processing a single image** as the input applying morphological operations, image segmentation, edge detection and color assignment. Some of them aim for stylized depiction [6] while others enhance legibility. Interactive techniques for stylized rendering such as rotoscoping have been used as well [7, 8].

Stereo techniques including passive and active illumination are generally de-

signed to compute depth values or surface orientation rather than to detect depth edges. Depth discontinuities present difficulties for traditional stereo: it fails due to *half-occlusions*, which confuse the matching process [9]. Few techniques try to model the discontinuities and occlusions directly [10] [11] [12]. Active illumination methods have been proposed for depth extraction, shape from shading, shape-time stereo and photometric stereo but are unstable around depth discontinuities [13]. An interesting technique has been presented to perform logical operations on detected intensity edges, captured under widely varying illumination, to preserve shape boundaries [14] but it is limited to uniform albedo scenes. Using photometric stereo, it is possible to analyze the intensity statistics to detect high curvature regions at **occluding contours** or *folds* [15]. But the techniques assume that the surface is locally smooth which fails for a flat foreground object like a leaf or piece of paper, or view-independent edges such as corner of a cube. They detect regions near occluding contours but not the contours themselves.

Techniques for **shape from shadow** (or darkness) build a continuous representation (*shadowgram*) from a moving light source from which continuous depth estimates are possible [16, 17]. However, it involves a difficult problem of estimating continuous heights and requires accurate detection of start and end of shadows. Good reviews of shadow-based shape analysis methods are available in [18] [19].
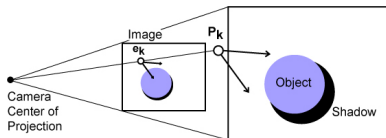

Fig. 1. Imaging geometry. Shadows of the gray object are created along the epipolar ray. We ensure that depth edges of all orientations create shadow in at least one image while the same shadowed points are lit in some other image.

3   Multi-Flash Imaging

This technique for detecting shape features in images was first described in [20, 21]. For completeness we review the basic idea here. The method is motivated by the observation that when a flashbulb illuminates a scene during image capture, thin slivers of cast shadow are created at depth discontinuities. Moreover, the position of the shadows is determined by the relative position of the camera and the flashbulb: when the flashbulb is on the right, the shadows are create on the left, and so on. Thus if we can shoot a sequence of images in which different light sources illuminate the subject from various positions, we can use the shadows in each image to assemble a depth edge map using the shadow images.

Imaging Geometry In order to capture the intuitive notion of how the position of the cast shadows are dependent on the relative position of the camera and light source, we examine the imaging geometry, illustrated in figure 1. Adopting a pinhole camera model, the projection of the point light source at P_k is at pixel e_k on the imaging sensor. We call this *image* of the light source the *light epipole*. The images of (the infinite set of) light rays originating at P_k are in turn called the *epipolar rays* originating at e_k. We also refer to the images of depth discontinuities as depth edges.

Removing and Detecting Shadows Our approach for reliably remove and detect shadows in the images is to strategically position lights so that every point in the scene that is shadowed in some image is also imaged without being shadowed in at least one other image. This can be achieved by placing lights strategically so that for every light there is another on the opposite side of the camera so that all depth edges are illuminated from two sides. Also, by placing the lights close to the camera, we minimize changes across images due to effects other than shadows. To detect shadows in each image, we first compute a **shadow-free image**, which can be approximated with the MAX composite image, which is an image assembled by choosing at each pixel the maximum intensity value from among the image set. The shadow-free image is then compared with the individual shadowed images. In particular, for each shadowed image we compute the *ratio image* by performing a pixel-wise division of the intensity by that of the MAX image. The ratio image is close to 1 at pixels that are not shadowed, and close to 0 at pixels that are shadowed. This serves to accentuate the shadows and also remove intensity transitions due to surface material changes.

Algorithm Codifying the ideas discussed above we arrive at the following algorithm [21]: Given n light sources positioned at $P_1, P_2...P_n$,
- Capture n pictures I_k, $k = 1..n$ with a light source at P_k
- For all pixels x, $I_{max}(\text{x}) = \max_k(I_k(x))$, $k = 1..n$
- For each image k,
  ▷ Create a ratio image, R_k, where $R_k(x) = I_k(x)/I_{max}(x)$
- For each image R_k
  ▷ Traverse each epipolar ray from epipole e_k
  ▷ Find pixels y with step edges with negative transition
  ▷ Mark the pixel y as a depth edge

4   Multi-Flash Imaging with Endoscopes

Unlike many traditional 3D shape recovery methods where the imaging apparatus need to be placed at large distances apart, in multi-flash imaging the light sources can be placed near to the camera. This allows compact designs that can be used in tightly constrained spaces.

Multiple-Instrument Designs The simplest way to implement multi-flash imaging is to use multiple instruments. One design is shown in figure 2(a), where instead of inserting one endoscope, three are inserted. The middle instrument
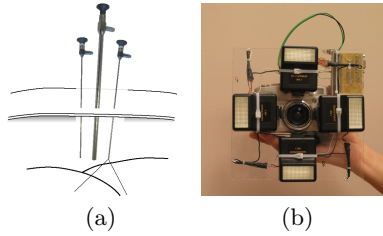
Fig. 2. Multi-flash imaging setup (a) for minimal-invasive operation (b) for medical illustrations.

acts as the camera while the two on the side act as light sources. By synchronizing the light sources with the image capture process for the middle endoscope, the entire setup would act as a multi-flash camera. While it may involve inserting more imaging instruments, it is a way to systematically illuminate the subject, and potentially could reduce the amount of adjustments required during an operation to produce images that convey the required 3D information necessary.

Single-Instrument Designs In many scenarios, it is more useful to have a single instrument capable of multi-flash imaging. For example in situations where flexible endoscopes are needed, it may be very difficult or impossible to insert and align multiple flexible light sources with the endoscope. Fortunately it is possible to implement multi-flash imaging on a single instrument. Figure 3 shows an R. Wolf *Lumina* laryngeal laparoscope endoscope modified to achieve multi-flash imaging. At the tip of the Lumina, as shown in figure 3(c), there is an imaging lens and two lights on opposite sides of the lens. These lights are made of numerous optic fibers that are bundled together at the base of the laparoscope. By illuminating this fiber bundle, the light is transmitted to the tip, serving as illumination near the imaging lens.

Clearly, if we can turn on each of the two lights independently, the Lumina would be capable of multi-flash imaging. Conventionally, both lights at the tip are turned on and off together, and standard equipment does not make any provisions for turning on only one light. However if one could identify which optic fibers lead to which light, then by only illuminating the fibers leading to one of the lights we can turn one light on without the other. If the fibers leading to the two lights are intertwined at the end of the fiber bundle, one would have to unbundle the fibers and sort them into two bundles manually. Fortunately, we found that the fibers from the two lights are grouped roughly into two parts at the fiber bundle. This is shown in figure 3(b). By illuminating the fibers at one of the lights and not the other, we can see that approximately one half of the fiber bundle is illuminated, with the other half corresponding to fibers leading to the other light. Thus by covering one or the other side of the fiber bundle, we can turn on and off the two lights individually.
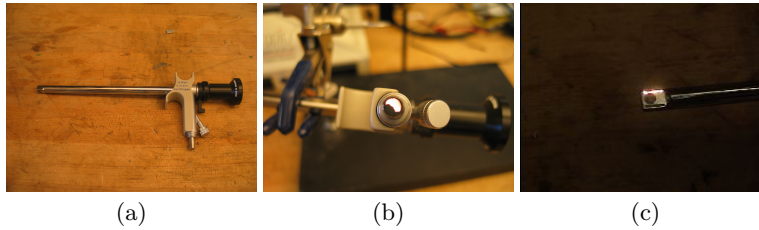
(a)   (b)   (c)

Fig. 3. Single-endoscope Multi-flash Setup. (a) R. Wolf Lumina laparoscope. (b) Fiber bundle with one end illuminated. The tip of the laparoscope shown (c) with left light turned on.

Fig. 4. Calf larynx displayed under (Left) Conventional imaging (Center) MAX composite (Right) enhanced with multi-flash imaging.

5   Multi-Flash Imaging for Medical Illustrations

In many traditional medical and biology textbooks and articles, pen-and-ink drawings of organs and tissues are often used in the place of photographs because they are easier to comprehend. However the use of these stylized illustrations is increasingly scarce, primarily because they are more expensive than photographs, and take more time to create. In our experiments with multi-flash imaging, we have observed that the depth edge confidence maps frequently resemble hand drawn sketches. At the same time, since they are created from photographs, they retaining a high degree of realism without overwhelming amounts of detail. We felt that multi-flash photography may make it faster, easier, and cheaper for artists and researchers to create medical illustrations.

To verify this, we built a multi-flash camera based on a 4 megapixel Canon Powershot G2, as shown in figure 2(b). The 4 flashes mounted around the camera are triggered sequentially by a controller board based on a Microchip PIC16F876 which synchronizes them to the image capture process by sensing the flash signal from the camera hot shoe. Some results generated with this camera are shown in figure 5. The same figure also shows the typical results produced using a conventional intensity edge detector. We showed the results to medical professionals at the Massachusettes General Hospital, and received very positive feedback on the usefulness of these images. In addition, they also found the MAX composite im-

age to be most useful, as it is difficult to take shadow-free images with ordinary cameras. Anatomists and medical illustrators for whom we have demonstrated the system are most interested in the 'depth edge confidence' image, which can be easily converted into a detailed black and white drawing.
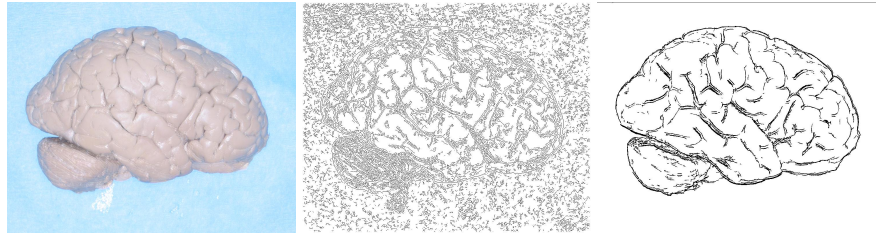

Fig. 5. Creating medical illustrations by multi-flash imaging. (Left) MAX Composite image. (Center) Canny Edges. (Right) Multi-flash Edges.

6 Conclusion

We have presented a promising approach to enhancing medical images and video that convey a heightened sense of 3D shape. Using cameras and laparoscopes with multiple sources of illumination, we described methods for computing a MAX composite image that is largely free of shadows and a depth edge confidence map that localize the 3D shape boundaries and provide a qualitative description of the local relative depth. Our contribution to the science of medical imaging are as follows:

▷ A design for implementing multi-flash imaging using multiple endoscopes.
▷ A second design for the same purpose, using a single modified endoscope.
▷ Experimental verification of the two designs.
▷ Experimental verification of the usefulness of both the MAX composite image and the depth edge confidence map for creating medical and biological illustrations and photography.

We believe that we have only scratched the surface of the potential enhancements in medical imaging that are enabled by the use of the multi-flash technique. By using light sources and imaging sensors placed in different spatiotemporal configurations as well as wavelengths, the technique can be adapted to augment many existing methods of medical imaging. We also hope that the technique can serve as a useful tool for medical illustrators and researchers. Additional results can be seen in the "Additional Material" supplement. and at `http://www.merl.com/people/raskar/MICCAI04`.

Acknowledgement

References

1. Vogt, F., Kruger, S., Niemann, H., Schick, C.: A system for real-time endoscopic image enhancement. In: MICCAI. (2003)
2. Mueller, M., Camartin, C., Dreher, E., Hanggi, W.: Three-dimensional laparoscopy. gadget or progress? a randomized trial on the efficacy of three-dimensional laparoscopy. Surg Endosc. **13** (1999)
3. Schade, G., Hess, M., Rassow, B.: Mglichkeit endolaryngealer morphometrischer messungen mit einem neuen laserlichtverfahren. HNO **50** (2002) 753–755
4. Rosen, D., Minhaj, A., Hinds, M., Kobler, J., Hillman, R.: Calibrated sizing system for flexible laryngeal endoscopy. In Schade, G., Mller, F., Wittenbeg, T., Hess, M., eds.: 6th International Workshop: Advances in Quantitative Laryngology (University of Hamburg) Advances in Quantitative Laryngology , Voice and Speech Research (Proceedings), Verlag (2003)
5. Hodges, E.R.S.: The Guild Handbook of Natural History Illustration. John Wiley and Sons (2003)
6. DeCarlo, D., Santella, A.: Stylization and Abstraction of Photographs. In: Proc. Siggraph 02, ACM Press. (2002)
7. Waking Life: Waking Life, the movie (2001)
8. Avenue Amy: Curious Pictures (2002)
9. Geiger, D., Ladendorf, B., Yuille, A.L.: Occlusions and binocular stereo. In: European Conference on Computer Vision. (1992) 425–433
10. Intille, S.S., Bobick, A.F.: Disparity-space images and large occlusion stereo. In: ECCV (2). (1994) 179–186
11. Birchfield, S., Tomasi, C.: Depth discontinuities by pixel-to-pixel stereo. International Journal of Computer Vision **35** (1999) 269–293
12. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In: International Journal of Computer Vision. Volume 47(1). (2002) 7–42
13. Sato, I., Sato, Y., Ikeuchi, K.: Stability issues in recovering illumination distribution from brightness in shadows. IEEE Conf. on CVPR (2001) 400–407
14. Shirai, Y., Tsuji, S.: Extraction of the line drawing of 3-dimensional objects by sequential illumination from several directions. Pattern Recognition **4** (1972) 345–351
15. Huggins, P., Chen, H., Belhumeur, P., Zucker, S.: Finding Folds: On the Appearance and Identification of Occlusion . In: IEEE Conf. on Computer Vision and Pattern Recognition. Volume 2., IEEE Computer Society (2001) 718–725
16. Raviv, D., Pao, Y., Loparo, K.A.: Reconstruction of three-dimensional surfaces from two-dimensional binary images. In: IEEE Transactions on Robotics and Automation. Volume 5(5). (1989) 701–710
17. Daum, M., Dudek, G.: On 3-D surface reconstruction using shape from shadows. In: CVPR. (1998) 461–468
18. Yang, D.K.M.: Shape from Darkness Under Error. PhD thesis, Columbia University (1996)

19. Kriegman, D., Belhumeur, P.: What shadows reveal about object structure. Journal of the Optical Society of America (2001) 1804–1813
20. Raskar, R., Yu, J., Ilie, A.: A non-photorealistic camera: Detecting silhouettes with multi-flash. In: SIGGRAPH Technical Sketch. (2003)
21. Raskar, R., Tan, K.H., Feris, R., Yu, J., Turk, M.: Stylized renderings using a multi-flash camera. (2004) Submitted for review.
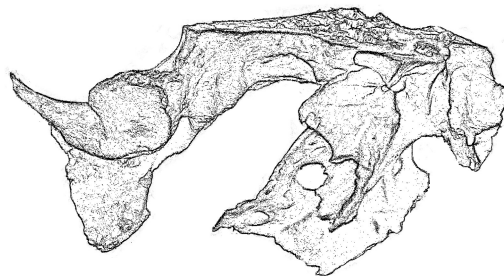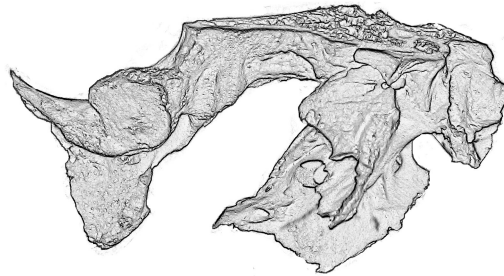
Additional Material

Fig. 6. Depicting a fragment of a human sphenoid bone. (Top) MAX Composite. (Middle) Depth Edge Confidence image. (Bottom) Applying a Photoshop pen-and-ink filter to the depth edge confidence image produces an image that resembles a traditional hand drawn illustration.