A Compact High-Definition Low-Cost Digital Stereoscopic Video Camera for Rapid Robotic Surgery Development

Jay CARLSON, Jędrzej KOWALCZUK, Eric PSOTA, and Lance C. PÉREZ
University of Nebraska-Lincoln, Department of Electrical Engineering

Abstract. Robotic surgical platforms require vision feedback systems, which often consist of low-resolution, expensive, single-imager analog cameras. These systems are retooled for 3D display by simply doubling the cameras and outboard control units. Here, a fully-integrated digital stereoscopic video camera employing high-definition sensors and a class-compliant USB video interface is presented. This system can be used with low-cost PC hardware and consumer-level 3D displays for tele-medical surgical applications including military medical support, disaster relief, and space exploration.

Keywords. Stereo imager, surgical robotics, minimally-invasive surgery

Introduction

A common problem surgeons face with in-vivo surgery, performed either traditionally or robotically, is the low-quality vision associated with existing imaging systems based on low-resolution analog imagers. In addition to being low-resolution, these systems typically use a single imager, and are incapable of providing the user with depth perception. Recent advances in image sensors have enabled the use of high-quality digital imaging systems for in-vivo surgical applications; meanwhile, trends in the miniaturization of sensors have created the opportunity for the development of simulated binocular vision which is capable of providing the surgeon with depth perception [10].

Stereoscopic video cameras aid medical professionals in many areas of treatment, diagnosis, pre-operative planning, surgery, and surgical training and teaching [15]. Although early tests showed mixed results [2], it is now well-established that surgeons – especially those inexperienced with performing laproscopic surgery – perform better with a stereoscopic camera system [6,1,11]. Stereoscopic vision enables better judgement of depth, as well as better perception of surface curvature and material properties [4]. In [12], stereoscopic vision has also been shown to improve the naturalness of images.

In this work, a stand-alone, compact stereoscopic digital camera is presented for in-vivo surgical robot systems. This device uses low-power high-definition (HD) CMOS sensors, and can be integrated with a variety of consumer-class 3D displays through a standard USB interface.

1. Background

Stereoscopic vision systems offer the surgeon realistic depth perception, and also lend themselves to a myriad of image processing applications [13,14]. It has been demonstrated that stereoscopic vision can be used to track robotic movements with significantly improved accuracy compared to traditional motor-based tracking [8]. Stereoscopic images paired with modern stereo matching techniques allow for full 3D reconstruction of the operating scene [9]. Emerging applications of stereoscopic reconstruction suggest that stereoscopic vision systems have the potential to assist and automate standard surgical procedures, thus reducing operating time and human error [3].

The adoption of a stereoscopic camera system necessitates further design considerations when compared to a monocular solution. Stereoscopic systems require a dual-imager device, a frame interleaver/interlacer, and an output display capable of targeting each eye separately. Existing imaging systems used in surgical robots have been based on off-the-shelf standard-definition analog CCD technology that has traditionally offered lower noise than the same generation of CMOS technology; however, CMOS has made great strides in quality over the last two decades [17], and because CMOS camera designs are much simpler, high-quality cameras are now available in small form factors.

In [13], Stoyanov *et al.* use a standard-definition analog camera for soft tissue 3D depth recovery. Hu and Miller developed an insertable stereoscopic imaging device [7] and built and tested their prototype [5] that used off-the-shelf CCD camera modules. Although their insertable system is small, it relies on expensive analog CCD imagers that offer poor resolution.

2. System Design

To improve upon existing analog camera systems used by many surgical robots, a digital low-cost CMOS-based USB high-definition stereoscopic camera is introduced. This system was designed to increase access to high-quality digital imaging while testing and developing software algorithms for stereoscopic reconstruction, tool tracking, and other surgical applications requiring 3D camera systems. The design of the camera was driven by a need[16] for researchers to have access to a high-quality, low-cost 3D video camera that could be deployed on multiple generations of robots, as well as for researchers working on 3D reconstruction and other image processing applications to have a comfortable development environment that can easily be replicated and used on desktop and laptop computers without bulky external camera control and video capture hardware found in traditional analog camera systems.

2.1. Design Overview

Figure 1 shows the proposed stereoscopic camera positioned next to a U.S. quarter. The camera is designed to be used with insertable robots for in-vivo surgery, and measures 2.4 centimeters in length and 2.6 centimeters in width, with a baseline separation between the image sensors of 1.5 centimeters. Given that the average distance between human eyes is 6.0 centimeters, the separation between sensors has the effect of providing $4\times$ depth magnification to the user when viewing the output of the video camera.

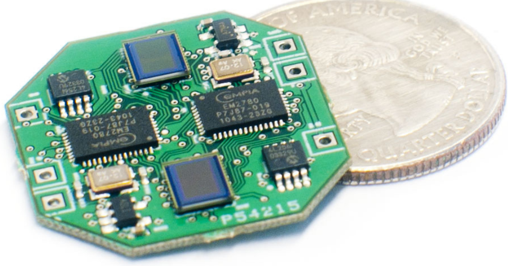
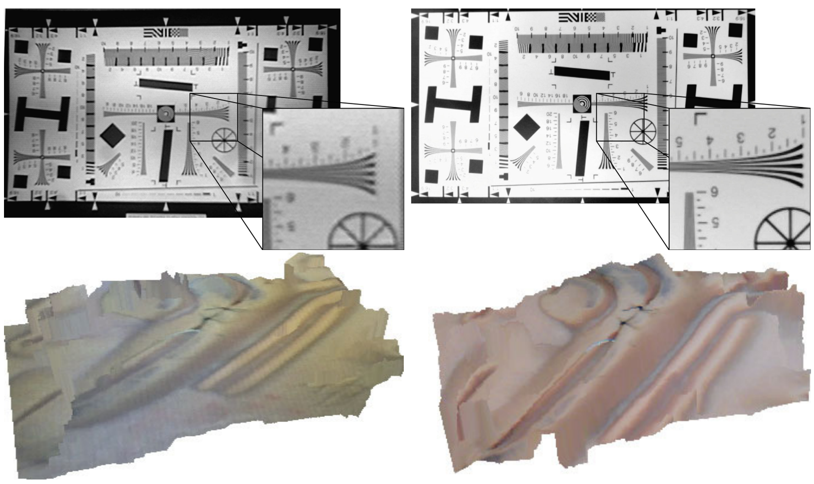


Figure 1. High definition 720p digital CMOS dual-imager circuit board, measuring 2.4 by 2.6 cm.

The camera system was designed around two consumer-class OmniVision OV9710 1/4" 1 megapixel CMOS digital imagers that output 720p HD video at 30 frames-per-second. Each camera is interfaced to a separate USB Video Class (UVC) compliant controller, the EETI EM2780, that provides control and bulk endpoints to the USB host. The EM2780 uses a 3.3V 0.18 micron process Intel 8051 core, resulting in low power consumption. The core is clocked at 60 MHz with a high-speed DMA buffer mated to the CPU and USB peripheral, and can stream HD video at 30 frames per second. Ancillary power management circuits are shared between the two imagers, which helps reduce the overall size of the imaging board to less than one square inch.

Because the proposed camera uses a standard USB interface, it can operated from any USB host. To view scenes in 3D, low-cost consumer-class polarization or LCD shutter monitors can easily be interfaced to the computer. Many high-end laptops feature integrated 3D-enabled displays that, when coupled with the proposed camera, enable high-quality 3D imaging in a lightweight, low-cost environment, thus providing an accessible platform for research and development of stereoscopic image processing algorithms.



(a) Analog NTSC video camera (b) Proposed 720p HD digital video camera

Figure 2. Resolution test using the ISO 12233 resolution chart for electronic still cameras (top), and three-dimensional reconstructions of a simulated surgical environment (bottom).

3. Results

The proposed digital stereoscopic video camera was holistically compared with a standard-definition stereoscopic analog camera. A consistent daylight-balanced florescent light source, measuring 2491 lux, was used to light the scene. Sunex DSL756 3.8mm f/2.8 quarter-inch format lenses, which provide a wide-angle field of view of 60°, were used for both the analog and the digital imagers.

The resolution of both the analog and digital cameras was tested using the ISO 12233 resolution chart. The results given in Figure 2 illustrate the improved sharpness and local contrast achieved using the proposed camera. The magnified sections of the test images show that the digital camera achieves approximately twice the resolving capabilities of the analog camera.

To test the capabilities of the analog and digital cameras when used for image processing, stereo images of a simulated surgical environment were captured using both systems. After capturing the stereo images, the 3D processing algorithm given in [9] was used to compute the textured mesh model of the simulated surgical environment. Figure 2 indicates the improved reconstruction capabilities of the digital camera when compared to the analog camera. There is a noticeable amount of discontinuity and a lack of detail in the model extracted from the analog stereo image pair, while the 3D reconstruction generated using images from the proposed digital camera accurately captures details in the surface of the scene. The resulting flatness of the reconstruction from the analog stereoscopic camera is likely due to a lack of local contrast, high noise levels, and the limited resolution of the image sensors.

4. Conclusion

A dual-imager camera was developed to address the need for a compact, low-cost, high definition video camera system, capable of providing depth perception in surgical situations. When compared to an analog camera, significant improvements in both local contrast and image resolving capabilities were demonstrated using the proposed camera. Additionally, the two systems were tested for their suitability to image processing applications. Using a stereo matching algorithm as a benchmark for comparison, the existing analog system was unable to match the depth and detail of the proposed digital system. These results indicate that the proposed system is capable of providing higher-quality video for robotic surgery applications – both in terms of conveying realistic depth perception, and allowing for the integration of advanced image processing.

Future applications for the system include installation and testing on the robotic platform described in [16] in a real surgical environment. Additional components to be integrated with the camera include an LED light source, a voice-coil-powered autofocus lens assembly, and a pan-tilt mechanism. Future upgrades to the camera might include further miniaturization, increasing image resolution and frame rate, as well as integrating a high-speed uncompressed digital video interface.

Acknowledgements

This work was funded in part by TATRC2 grant W81SWH-09-2-0185.

References

[1] John C. Byrn, Stefanie Schluender, Celia M. Divino, John Conrad, Brooke Gurland, Edward Shlasko, and Amir Szold. Three-dimensional imaging improves surgical performance for both novice and experienced operators using the da vinci robot system. *The American Journal of Surgery*, 193(4):519–522, 2007.

[2] A. C. W. Chan, S. C. S. Chung, A. P. C. Yim, J. Y. W. Lau, E. K. W. Ng, and A. K. C. Li. Comparison of two-dimensional vs three-dimensional camera systems in laparoscopic surgery. *Surgical Endoscopy*, 11(5):438–440, 1997.

[3] J. Dumpert, A.C. Lehman, N.A. Wood, D. Oleynikov, and S.M. Farritor. Semi-autonomous surgical tasks using a miniature in vivo surgical robot. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 266 –269, sept. 2009.

[4] N. S. Holliman. *3D display systems*. IOP Press, 2006.

[5] Tie Hu, P.K. Allen, T. Nadkarni, N.J. Hogle, and D.L. Fowler. Insertable stereoscopic 3d surgical imaging device with pan and tilt. In *Biomedical Robotics and Biomechatronics, 2008. BioRob 2008. 2nd IEEE RAS EMBS International Conference on*, pages 311 –316, oct. 2008.

[6] G. Hubens, H. Coveliers, L. Balliu, M. Ruppert, and W. Vaneerdeweg. A performance study comparing manual and robotically assisted laparoscopic surgery using the da vinci system. *Surgical Endoscopy*, 17:1595–1599, 2003. 10.1007/s00464-002-9248-1.

[7] A. Miller, P. Allen, and D. Fowler. In-vivo stereoscopic imaging system with 5 degrees-of-freedom for minimal access surgery. *Studies in Health Technology and Informatics*, 98:234–240, 2004.

[8] E. T. Psota, K. Strabala, J. Dumpert, L. C. Pérez, S. Farritor, and D. Oleynikov. Stereo image-based arm tracking for in vivo surgical robotics. In *Studies in Technology and Informatics*, volume 163, pages 454–460, 2011.

[9] Eric T. Psota, Jedrzej Kowalczuk, Jay Carlson, and Lance C. Pérez. A local iterative refinement method for adaptive support-weight stereo matching. In *International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, July 2011.

[10] Brian J. Rogers. *Seeing in Depth: Depth perception*, volume 2. University of Toronto Press, 2002.

[11] Philip Servos, Melvyn A. Goodale, and Lorna S. Jakobson. The role of binocular vision in prehension: a kinematic analysis. *Vision Research*, 32(8):1513 – 1521, 1992.

[12] Pieter J. H. Seunti, Ingrid E. J. Heynderickx, Wijnand A. Ijsselsteijn, Paul M. J. Van Den Avoort, Jelle Berentsen, Iwan J. Dalm, Marc T. M. Lambooij, and Willem Oosting. Viewing experience and naturalness of 3d images. In *Three-Dimensional TV, Video, and Display IV, Proceedings of SPIE*, volume 6016, Bellingham, WA, ETATS-UNIS, 2005. Society of Photo-Optical Instrumentation Engineers.

[13] Danail Stoyanov, Ara Darzi, and Guang Yang. Dense 3d depth recovery for soft tissue deformation during robotically assisted laparoscopic surgery. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*, volume 3217 of *Lecture Notes in Computer Science*, pages 41–48. Springer Berlin / Heidelberg, 2004.

[14] Danail Stoyanov, Marco Scarzanella, Philip Pratt, and Guang-Zhong Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In Tianzi Jiang, Nassir Navab, Josien Pluim, and Max Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, volume 6361 of *Lecture Notes in Computer Science*, pages 275–282. Springer Berlin / Heidelberg, 2010.

[15] Maurice H. P. H. van Beurden, Gert van Hoey, Haralambos Hatzakis, and Wijnand A. Ijsselsteijn. Stereoscopic displays in medical domains: a review of perception and performance effects. In *Human Vision and Electronic Imaging XIV, Proceedings of the SPIE*, volume 7240, 2009.

[16] Tyler D. Wortman, Kyle W. Strabala, Amy C. Lehman, Shane M. Farritor, and Dmitry Oleynikov. Laparoendoscopic single-site surgery using a multi-functional miniature in vivo robot. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 7(1):17–21, 2011.