SLI in AR applications

Daniel L. Lau[a] and Ying Yu[a] and Matthew P. Ruffner[a]

[a]University of Kentucky, Address, Lexington, US;

ABSTRACT

This document shows the desired format and appearance of a manuscript prepared for the Proceedings of the SPIE. It contains general formatting instructions and hints about how to use LaTeX. The LaTeX source file that produced this document, `article.tex` (Version 3.3), provides a template, used in conjunction with `spie.cls` (Version 3.3).

Keywords: Manuscript format, template, SPIE Proceedings, LaTeX

1. INTRODUCTION

Structured light illumination is a 3D scanning technique based on projecting a series of striped line patterns and then reconstructing a target surface based on the warping of the stripes as measured by a digital camera.[1] Compared to alternative schemes like passive stereovision[2] or time-of-flight,[3] structured light is far superior in terms of accuracy,[4] but as a triangulation scheme, SLI cannot easily scan at long ranges.[5] Furthermore, as a scanning process, SLI is not typically associated with real-time applications with moving targets.[6]

Scanning time is a major issue in structured light research as numerous papers have been published that focus either on using high speed projector/camera pairs[7] or using unique pattern codecs[8–10] that minimize the number of projected patterns. In fact, single pattern SLI techniques do exist[11–13] that are quite popular in human computer interfacing applications[14] but which perform poorly in bright ambient light conditions.[15]

With regards to performing structured light scanning with commodity components, a major issue is the inability to properly synchronize a camera and projector such that one can associate a specific pattern to a specific captured image without including some type of key-frame scheme and analyzing the pixels of the captured image. As clearly documented by graphics card manufacturers, GPUs are not real-time systems, and there is no way for the CPU to know when a requested pattern shows up on the HDMI port.[16] And while it is possible to guarantee that no particular pattern is dropped, it is not possible to guarantee that a particular pattern is not repeated multiple times in a row.

As a means of addressing one's ability to properly synchronize the camera and projector, there are commercially available projectors that include electronics to embed patterns into the projector which include appropriate GPIO to interface with a camera's trigger;[17–19] however, these projectors are inordinately expensive on a per lumens basis. As of the time of this writing, the TI Lightcrafter 4500 retails for $1,250 with only 150 lumens brightness[20] while an LED-based Optoma ML750ST has approximately 500 lumens brightness[21] with the same spatial light modulator for $550.[22] Admittedly, the TI Lightcrafter 4500 can project patterns at up to 4,225 Hz, but the 120 Hz of 3D ready projectors is more than adequate for most machine vision cameras limited in frame rate by either the bus to the CPU (USB 3.0, for instance) or by the exposure time required by the sensor.

As a means of ensuring perfect projector/camera synchronization using commodity components, we propose a novel approach to machine vision camera architectures by incorporating HDMI pass-through where the camera reads the incoming HDMI video to extract hand-shaking messages embedded into the top-left pixel. Such handshaking could be used to then trigger the camera to capture the projected frames, ignoring duplicate or don't care frames. Without an incoming HDMI signal, the machine vision camera can alternately produce its own video signal, converting any HDMI projector into an instant structured light scanner which it instantly syncs.

Further author information: (Send correspondence to A.A.A.)
A.A.A.: E-mail: aaa@tbk2.edu, Telephone: 1 505 123 1234
B.B.A.: E-mail: bba@cmp.com, Telephone: +33 (0)1 98 76 54 32

Aside from simple synchronization to a projector, the incorporation of HDMI into a machine vision sensor produces many application possibilities. For instance, cameras could take advantage of EDID[23] to tell a host CPU what kind of camera it is and what its capabilities are including its resolution, color space, and frame rates. Multiple cameras can also by daisy chained together and instantly synchronized together using commodity cables purchased at any big box store with cable lengths of 15, 25, or even 50 feet[24] or longer with appropriate repeaters.[25] And now with the inclusion of ethernet over HDMI,[26] one could easily see GigE cameras converted to HDMI in the near future.

Now an application for our HDMI pass-through cameras that we are particularly interested is projector-based augmented reality such as the AR Sandbox[27] and the AR Pool Table.[28] In the case of the AR Sandbox, a digital projector paired with a PrimeSense RGB+D camera with the RGB+D camera used to monitor the shape of sand in a sandbox placed under the projector as depicted in Fig. **??**. The shape of the sand is then used to generate a psuedo-color relief map which is then projected back onto the sand in perfect register. The result system can then be used to visualize water flow through simulated rivers and streams.[29]

In the AR Pool Table, a digital projector is used to direct the player on where and how to strike the que ball in order to sink the other balls in the various pockets. A machine vision camera monitors the balls' positions on the table as well as the pose of the stick and project augmenting lines onto the table in order to show the player how well their stick alignment matches the alignment posed by the CPU. Working in reverse, the AR Pool Table is now being used to train the PC the play pool using a robotic stick.[30]

In either of the above or other projector-based AR systems,[31–33] a common trait of all systems is their inclusion of low-accuracy 3D vision with a commodity projector. Our proposed machine vision camera with HDMI pass through could be instantly inserted between the projector and computer to give that system the option of highly accurate (millimeter) 3D scanning when the RGB+D sensors give only centimeter accuracy. Such high resolution would not only enhance scene understanding in the computer (having highly accurate scans) but also enhance registration between the target surface with the projector as the camera can derive sub-pixel alignment with the projector.

2. HARDWARE STRUCTURE

The entire system is based on a Zynq 7000 SoC. The Zynq 7000 consists of two parts, processing system (PS) and programmable logic (PL). PS is dual-core ARM Cortex-A9 processor, and PL is equivalent to a Xilinx 7 series FPGA. After powering on, an 80Mhz clock signal from the PL is not immediately fed to the LUPA300 until the PS boots up the operating system and configures the PL. However, the LUPA300 requires that the clock is supplied earlier than the power and the reset, so a delay logic is necessary to make sure that the power is supplied after the clock. Meanwhile, an EDID simulator is set to standby in the DDC channel of the HDMI once the PL is configured. Upon receiving an EDID request from the PC, it responses with a 128-byte or 256-byte EDID data package. Once this data is recognized by the PC, PC starts to streams video to the PL through the HDMI cable. In this case, the PL just passes whatever it receives to the projector. On the other hand, if the the PL doesn't detect a clock from the PC in 3 seconds, which indicates that there's no valid HDMI connection, a multiplexer will route the structured light patterns generated by the SLI pattern generator to the projector. The SLI pattern generator has its clock from an external oscillator (Si514). On the LUPA300 side, there is a control logic which includes a SPI controller and a synchronization module. The SPI controller configures the LUPA300 on startup by loading a bunch of internal registers with user defined values through a SPI bus. The synchronization module synchronizes the LUPA300 with a Vertical Sync signal either from the PC or the SLI pattern generator. To start a scan, the PC issues a command to the PS via Ethernet, the command as well as two parameters, exposure time and trigger delay are sent to the PL via Xillybus which is a dedicated IP core that provides some APIs for FPGA and Linux operating system to communicate. Then the PL begins to trigger the LUPA300 according to these parameters along with the synchronization module. After that the output data bus of the LUPA300 sequentially transmits pixel data to the FIFO, the Xillybus, the PS, and eventually the host PC. Fig. 1 illustrates the system diagram discussed above.

!t

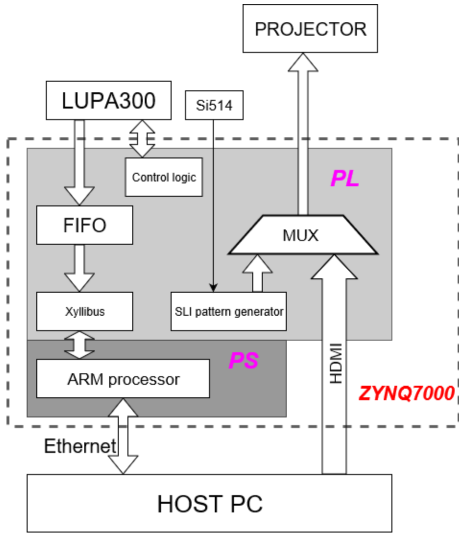

Figure 1. System diagram

ACKNOWLEDGMENTS

This unnumbered section is used to identify those who have aided the authors in understanding or accomplishing the work presented and to acknowledge sources of funding.

REFERENCES

1. Li, E. B., Peng, X., J. Xi, J. F. C., Yao, J. Q., and Zhang, D. W., "Multi-frequency and multiple phase-shift sinusoidal fringe projection for 3d profilometry," *Opt. Express* **13**(5), 1561–1569 (2005).
2. Curless, B. and Levoy, M., "A volumetric method for building complex models from range images," *Proceedings on computer graphics and interactive techniques* , 303–312 (1996).
3. Gokturk, S. B., Yalcin, H., and Bamji, C., "A time-of-flight depth sensor-system description, issues and solutions," *Conference on Computer vision and pattern recognition workshop* , 35–35 (2004).
4. Li, J., Hassebrook, L. G., and Guan, C., "Optimized two-frequency phase measuring profilometry light sensor temporal noise sensitivity," *J. Opt. Soc. Am.* **20**, 106–115 (2003).
5. An, Y., Bell, T., Li, B., Xu, J., and Zhang, S., "Method for large-range structured light system calibration," *Appl. Opt.* **55**(33), 9563–9572 (2016).
6. Deetjen, M. E., Biewener, A. A., and Lentink, D., "High-speed surface reconstruction of a flying bird using structured light," *J. Expe. Biology* **220**, 1956–1961 (2017).
7. Gong, Y. and Zhang, S., "Ultrafast 3-d shape measurement with an off-the-shelf dlp projector," *Opt. Express* **18**(19), 19743–19754 (2010).
8. Zhang, S. and Huang, P., "High-resolution, real-time 3d shape acquisition," *Conference on Computer Vision and Pattern Recognition Workshop* **3**, 28–37 (2004).
9. Su, W. H. and Liu, H., "Calibration-based two-frequency projected fringe profilometry: a robust, accurate, and single-shot measurement for objects with large depth discontinuities," *Opt. Express* **14**(20), 9178–9187 (2006).
10. Liu, K., Wang, Y., Lau, D., Hao, Q., and Hassebrook, L. G., "Dual-frequency pattern scheme for high-speed 3-d shape measurement," *Opt. Express* (1 March 2010).
11. Takeda, M. and Mutoh, K., "Fourier transform profilometry for the automatic measurement of 3-d object shapes," *Appl. Opt.* **22**(24), 3977–3982 (1983).
12. Wust, C. and Capson, D. W., "Surface profile measurement using color fringe projection," *Machine Vision and Applications* **4**, 193–203 (1991).

13. Koninckx, T. P. and Gool, L. V., "Real-time range acquisition by adaptive structured light," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(3), 432–445 (2006).
14. Guan, C., Hassebrook, L. G., Lau, D. L., and Yalla, V. G., "Composite pattern structured light projection for human computer interaction in space," *Proc. SPIE* **5798** (2005).
15. Yuan, J., Du, S., Zhao, D., Yu, Y., and Zhou, Y., "Adaptive color calibration based one-shot structured light system," *Sensors* **12**(8), 10947–10963 (2012).
16. NVIDIA, *CUDA runtime API reference manual* (July 2018). vRelease Version.
17. "Greenlight dlp projection optics." `https://www.greenlightoptics.com/intro.aspx?id=1000`.
18. "Ajp-4500 structured light dmd projector." `https://ajile.ca/ajp-4500/`.
19. "Lightcrafter 4500." `http://www.ti.com/tool/DLPLCR4500EVM`.
20. Texas Instruments, *TI DLP LightCrafter 4500 Evaluation Module User's Guide* (July 2017).
21. "Road test: optoma ml750st
." `https://www.projectorcentral.com/optoma-ml750st-projector-review-road-test.htm`.
22. "Ml750st." `https://www.bhphotovideo.com/c/product/1213546-REG/optoma_technology_ml750st_700_lumen_wxga.html?ap=y&gclid=EAIaIQobChMIv7GHra--3gIVxoSzCh2oAwfyEAQYAyABEgKHsfD_BwE&smp=y`.
23. Video Electronics Standards Association (VESA), *E-EDID standard* (February 2000). Rev. 1.
24. "50 feet cable." `https://www.monoprice.com/product?p_id=2678&gclid=EAIaIQobChMI5tiHhbS-3gIVTb7ACh2v6QdtEAQYAiABEgLkLvD_BwE`.
25. "Hdmi repeater." `https://nerdtechy.com/best-hdmi-signal-amplifier-boosters`.
26. HDMI Licensing, LLC, *High-Definition Multimedia Interface Specification* (November 2009). Version 1.4.
27. Woods, T. L., Reed, S., Hsi, S., Woods, J. A., and Woods, M. R., "Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs," *Journal of Geoscience Enducation* **64**(3), 199–214 (2016).
28. Uchiyama, H. and Saito, H., "Ar supporting system for pool games using a camera-mounted handheld display," *Advances in Human-Computer Interaction* **2008** (2008).
29. Afrooz, A., Ballal, H., and Pettit, C., "Implementing augmented reality sandbox in geodesign, a future for geodesign," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **IV-4**, 5–12 (2018).
30. Greenspan, M., Lam, J., Godard, M., Zaidi, I., Jordan, S., Leckie, W., Anderson, K., and Dupuis, K., "Toward a competitive pool-playing robot," *Computer* **41**(1), 46–53 (2008).
31. Miyata, K., Shiroishi, R., and Inoue, Y., "Applying ar technology with a projector-camera system in a history museum," *Proc. SPIE* **7866**(1), D0–D6 (2011).
32. Kemmoku, Y. and Komuro, T., "Ar tabletop interface using a head-mounted projector," *IEEE International Symposium on Mixed and Augmented Reality* , 288–291 (2016).
33. Doshi, A., Smith, R., Thomas, B., and Bouras, C., "Use of projector based augmented reality to improve manual spot-welding procision and accuracy for automotive manufacturing," *The International Journal of Advanced Manufacturing Technology* **89**(5), 1279–1293 (2017).