3D scanning by means of dual-projector structured light illumination

Daniel L. Lau[a] and Ying Yu[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

As one of the non-contact 3D shape measurement techniques, the structured light illumination (SLI) has been known for its high resolution and high speed.[1] Conventional SLI systems consist of one projector, one camera and one processing unit which is usually a computer. The projector presents patterns that are encoded with some information related to the pixel locations on the object to be measured. If the projection is on a flat surface, the patterns seen should be identical to its original design. However, with the presence of the non-planar object under the projection, what is actually seen from the camera is the distorted patterns on the surface of the object. By comparing and analyzing the distortion in the images taken by the camera, the 3D surface of the object can eventually be reconstructed in a processing unit.

Within the scope of this fundamental idea and system structure, a lot of research has been done over the past several decades.[2] Different practical implementations have been proposed, from pattern design to system calibration, then to the algorithms that are used to decode the captured patterns. One widely studied area is the one-shot SLI strategy in which only one static pattern is projected onto the object. It employs color pattern,[3] binary grid pattern,[4] gray-scale pattern[5] or even composite pattern.[6] Since there is only one image need to be processed for a scan, the 3D reconstruction can be completed fast. The one-shot strategy is ideal for high speed applications such as real time scanning. But in terms of accuracy, it is not so promising compared to some multi-shot SLI strategy in which the 3D reconstruction is derived by projecting a sequence of patterns and processing multiple images.[7]

According to the method of encoding the information containing the pixel location into the pattern, phase measuring profilometry (PMP)[8] is a subset of SLI techniques that acquires the depth value by translating the phase data from the projected patterns pixelwise. The PMP has some advantageous features including its insensitivity to ambient light and its high accuracy.[9,10] Higher frequency patterns have been introduced to PMP systems to reduce the effects of the noise and achieve higher accuracy.[11] Nonetheless, higher frequency patterns also produces the phase ambiguity which requires the system to execute some extra computation called phase unwrapping.[12] Furthermore, adding high frequency patterns increases the projection time as well as the number of images to be processed, which accordingly decreases the overall speed of the system. Liu *et al.* proposed an ingenious dual-frequency pattern strategy which combines a high-frequency pattern and a unit-frequency pattern into one composite pattern,[13] it improved the accuracy without increasing the scanning time.

In this paper, we propose a dual-projector, dual-frequency SLI scheme and its practical implementation. Provided with a wider projection angle by two projectors, our system not only takes advantage of the Liu *et al.*'s pattern design, but also overcomes the issues like muti-path[14] and occlusion[15] that single- projector systems suffer from. Additionally, if two projectors are positioned side by side and aimed at the same projection area, we can get

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

PIN	DATA
Data2+, Data2 Shield, Data2-	red pixel component, CTL2, CTL3 and auxiliary data
Data1+, Data1 Shield, Data1-	green pixel component, CTL0, CTL1 and auxiliary data
Data0+, Data0 Shield, Data0-	blue pixel component, HSYNC, VSYNC and auxiliary data
Clock+, Clock Shield, Clock-	pixel clock
SCL, SDA	DDC channel, the source reads the EDID from the sink
CEC	data or commands from remote control
Reserved/HEAC+	reserved for v1.3 and before, Ethernet and audio since v1.4
HOT PLUG DETECT/HEAC-	indicate the hot plug or paired with HEAC+
+5V, Ground	power from external or HDMI source, ground

Table 1: HDMI pinout

double the luminance. For instance, suppose the projectors we use are two identical Optoma ML750ST which has around 500 lumens each,[16] by using two projectors, we obtain a total 1000 lumens. Higher light intensity gives rise to the higher signal-to-noise ratio (SNR),[17] which makes the system less susceptible to noise, consequently the system becomes more robust and accurate.

2. HDMI OVERVIEW

HDMI is the abbreviation of High-Definition Multimedia Interface, it is one of the most popular display interfaces. The newest release, HDMI Version 2.1 supports up to $10K$ video at $120Hz$. A standard HDMI connector has 19 pins as listed in Table 1.

Data channel 2, 1, 0 are mainly used to transfer red, green and blue components of the video respectively. The HDMI does not only transfer video data, but also some auxiliary data, for example audio data, packet header. The auxiliary data, video data as well as some control signals are encoded in data channel 2, 1, 0 and then digitally transmitted in serial. In between any two adjacent video periods, one or more data island period and control period are inserted. There are six important control signals, HSYNC indicates the beginning and end of a row in a frame of the video, VSYNC indicates the beginning and end of a frame, CTL0 CTL3 indicate the data type of the following data period. The three data channels are transmitted through a differential signaling technology called Transition-Minimized Differential Signaling (TMDS) to reduce the impact of electromagnetic interference and enable high clock skew tolerance. Another common application of TMDS is in the Digital Visual Interface (DVI).

The 5 volts power signal is provided by the HDMI source or an external source. After the HDMI sink receives an 5 volts signal at the corresponding pin, it immediately asserts the pin Hot Plug Detect. Once the HDMI source detects the presence of a sink by the assertion of the pin Hot Plug Detect, it sends an I^2C-based command of a read request to the sink. The pins SCL and SDA compose the display data channel (DDC) via which the Extended Display Identification Data (EDID) is read by the HDMI source from the sink as the response to the read request. The EDID is usually 128 or 256 bytes long, it contains various information related to the features of display system, including but not limited to, manufacturer ID, serial number, week and year of manufacture, screen size, supported timing, etc.

The pin CEC is used to add some advanced functionalities for the HDMI systems. Usually it is a remote control that issues different high-level commands to the devices connected by HDMI cables. CEC stands for Consumer Electronics Control, it is also a one-wire bus protocol, the implementation of CEC is optional, because not all the HDMI devices support this feature. Since HDMI 1.4, the previously reserved pin has become the HDMI Ethernet and Audio Returen Channel (HEAC). While it is in audio return channel mode, only the HEAC+ line is used to transmit audio data; in HDMI Ethernet channel mode, the HEAC+ line pairs up with the HEAC- line as a differential signal to establish a high speed Ethernet communication.

3. SYSTEM IMPLEMENTATION BASED ON FPGA

Our FPGA-based dual-projector Structured Light Illumination (SLI) system generates two synchronized SLI patterns which are then fed to two projectors via HDMI. Meanwhile, the projectors and the camera need to be

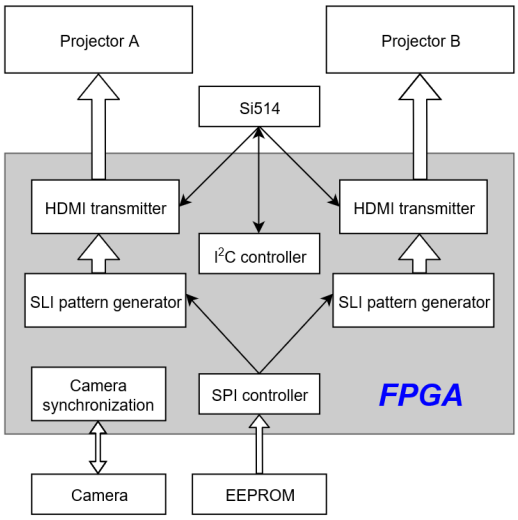

Figure 1: System diagram

synchronized to ensure that the camera images are taken at the right timing. The system diagram is shown in Figure 1.

The two SLI pattern generators output two synchronized phase-shifting fringe patterns which are later encoded into TMDS data streams by the HDMI transmitters and eventually move to projectors. The camera synchronization module controls the timing of the camera trigger signal, it detects the end of the camera's exposure time and makes sure that during the whole camera's exposure time, there is no different pattern projected and every unique phase-shifting fringe pattern is pictured sequentially by the camera. The projector in our SLI system operates at the resolution of $800x600$ and the refresh rate of $120Hz$. According to the document from VESA,[18] the HDMI timing should be set as Table 2 lists.

Pixel Clock	73.250MHz
Hor. Front Porch	48 pixels
Hor. Sync Time	32 pixels
Hor. Back Porch	80 pixels

Ver. Front Porch	3 lines
Ver. Sync Time	4 lines
Ver. Back Porch	29 lines

Table 2: HDMI timing of $800x600@120Hz$

To obtain the uncommon 73.25MHz pixel clock, we utilize a programmable oscillator Si514. By configuring the internal registers through I^2C bus, Si514 can generate any frequency from 100kHz to 250MHz with a tuning resolution of 0.026 ppb. Therefore, an I^2C master controller was incorporated into the system. One last module in the system is a lookup table (LUT) that is used to linearize the output of the projector. Ideally, the input digital values of the projector is proportional to the light intensity at the output side of the projector. The light intensity is measured by reading the pixel value of the photo taken by the camera. However, in practice, they are non-linear for most of the times due to some intrinsic characteristics of the projector, e. g. gamma distortion.[19] Applying a LUT to compensate the non-linearity is an effective way to address this problem, as shown in the Figure 2, the x-axis represents the 8-bit input digital value of the projector at a certain point of the scene, the y-axis represents the light intensity of the same point that is measured from camera's perspective. This LUT can be hard-coded into the configuration file of the FPGA, but the drawback is that once the lookup table changes, the FPGA configuration file has to be changed. It is quite inconvenient especially when the system needs to be often applied to a different projector, because generating a new FPGA configuration file requires special software tool and takes more time. We devise a method that the user stores the LUT in an separate EEPROM chip which can be erased and written by any computer via some general serial or USB tools, and the SPI master module in the FPGA reads the EEPROM every time the system is powered on. With this approach, users can load new

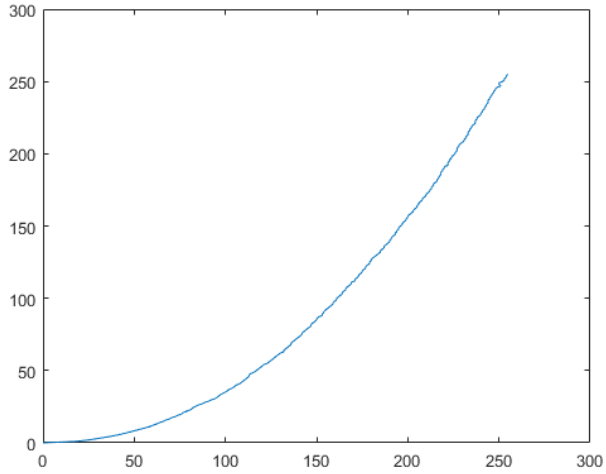

Figure 2: Non-linearity of the projector Optoma ML750ST

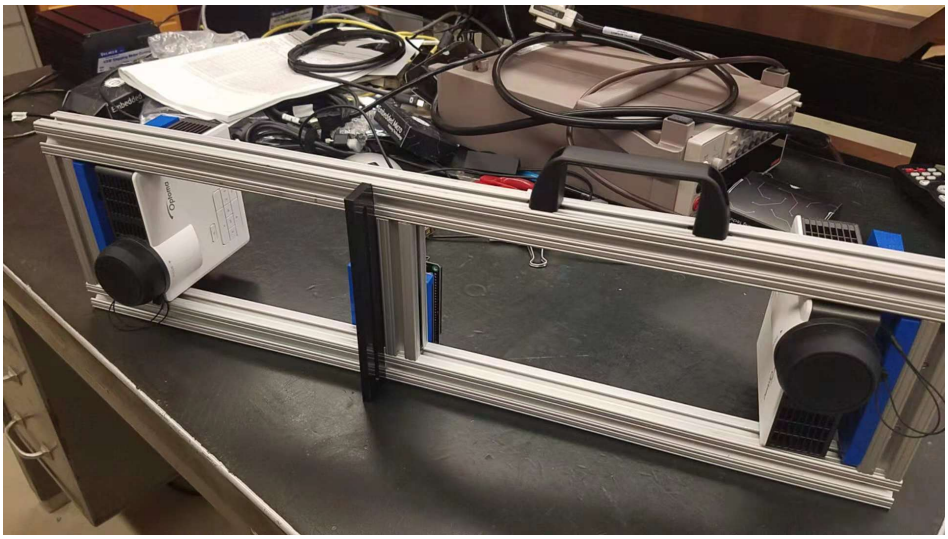

Figure 3: the dual-projector SLI system

LUTs much faster and easier.

Figure 3 shows a picture of our tested system.

4. DUAL-FREQUENCY PATTERN SCHEME

The dual-frequency pattern scheme was initially proposed by Liu *et al.*,[13] it is essentially an improved derivation of the phase measuring profilometry.[10] In the the traditional PMP a series of phase-shifting sinusoidal fringe patterns are projected onto the object to be measured, the fringe patterns can be either horizontal or vertical, the presence of the object distorts the patterns, then by analyzing the phase values in the deformed pattern images, the 3D depth values can be obtained.

Suppose the phase-shifting patterns are vertical, which indicates that all the pixels in a same row have the identical intensity, and from top to bottom the value of intensities in each row form a sinusoidal function.

Therefore, this kind of PMP patterns can be generalized as:

$$I_n^p(x^p, y^p) = A^p + B^p \cos(2\pi f y^p - \frac{2\pi n}{N}), \tag{1}$$

where I_n^p is the intensity of the pixel at the coordinate (x^p, y^p) from the projector's point of view; A^p is background intensity which is considered as a constant; B^p is another constant which represents the fringe contrast compared to the background; f is the frequency of the fringe pattern set which is equal to the number of sinusoidal periods from top to the bottom of the projected patterns; N is the total number of the phase-shifting patterns of the same frequency in a set, n is the current number of the pattern within the range of $[0, N-1]$.

On the camera's side, the image of the object under the projection of each unique PMP pattern is captured as part of the data to reconstruct the 3D profile of the object. So the projected scene of the object can be described as eq. (2) from the camera's point of view,

$$I_n^c(x^c, y^c) = A^c(x^c, y^c) + B^c(x^c, y^c) \cos(\phi(x^c, y^c) - \frac{2\pi n}{N}), \tag{2}$$

where I_n^c is the intensity of the given pixel (x^c, y^c) in the camera's coordinates system; $A^c(x^c, y^c)$ is the average intensity of the given pixel over the N patterns; B^c can be considered as the average intensity of the PMP pattern seen by the camera, so for any given pixel (x^c, y^c) in a pattern set $[0, N-1]$, A^c and B^c are both constant, furthermore, $B^c(x^c, y^c)$ can be derived from eq. (2) as:

$$B^c(x^c, y^c) = \frac{2}{N} \sqrt{[\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})]^2 + [\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})]^2}. \tag{3}$$

In Equation 2, there is another important constant ϕ for any given pixel (x^c, y^c) in a given pattern set $[0, N-1]$, it is the phase value $\phi(x^c, y^c)$ of the distorted fringe patterns that is eventually used to calculate the corresponding depth value of the object. The expression of $\phi(x^c, y^c)$ can be inferred from eq. (3) as:

$$\phi(x^c, y^c) = \arctan \frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}. \tag{4}$$

The dual-frequency pattern scheme adds a second sinusoidal component to the phase-shifting patterns, as a result, the total number of patterns is still N, but each pattern has two sinusoidal components, one is of unit frequency f_u, the other is of a high frequency f_h that is used to reduce the impact of noises in the system. If we extend the single-frequency PMP pattern above to the dual-frequency, the intensities of the row pixels from top to bottom would be like an amplitude modulation in which the f_u is frequency of the modulating signal and f_h is the frequency of the carrier wave. According to Liu *et al.*, from the projector's point of view, the new dual-frequency pattern is expressed as:

$$I_n^p(x^p, y^p) = A^p + B_1^p \cos(2\pi f_h y^p - \frac{2\pi n}{N}) + B_2^p \cos(2\pi f_u y^p - \frac{4\pi n}{N}), \tag{5}$$

where the I_n^p is the intensity of the pixel (x^p, y^p) in the n^th pattern. Similarly the intensity equation of the camera coordinates used for reconstruction becomes:

$$I_n^c(x^c, y^c) = A^c(x^c, y^c) + B_1^c(x^c, y^c) \cos(\phi_h(x^c, y^c) - \frac{2\pi n}{N}) + B_2^c(x^c, y^c) \cos(\phi_u(x^c, y^c) - \frac{4\pi n}{N}). \tag{6}$$

Likewise the B_1^c, B_2^c and ϕ_h, ϕ_u can be derived from eq. (3) and eq. (4) respectively,

$$B_m^c(x^c, y^c) = \frac{2}{N} \sqrt{[\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(m\frac{2\pi n}{N})]^2 + [\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(m\frac{2\pi n}{N})]^2}, \tag{7}$$

where $m = 1$ *or* 2 in this case,

$$\phi_h(x^c, y^c) = \arctan \frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{2\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{2\pi n}{N})}, \tag{8}$$

$$\phi_u(x^c, y^c) = \arctan \frac{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \sin(\frac{4\pi n}{N})}{\sum_{n=0}^{N-1} I_n^c(x^c, y^c) \cos(\frac{4\pi n}{N})}. \tag{9}$$

In eq. (8) and (9), ϕ_h and ϕ_u are confined within $[-\pi, \pi]$, however, the actual range of ϕ_h is $[0, 2f_h\pi]$, the ϕ_h is called the wrapped phase. Unwrapping the ϕ_h to $\tilde{\phi_h} \in [0, 2f_h\pi]$ can lead to a more accurate conversion from phase to depth. In order to obtain $\tilde{phi_h}$, the value of ϕ_u is used as the following equation shows,

$$\tilde{\phi_h} = \phi_h + 2\pi \lfloor \frac{\phi_u}{2\pi/f_h} \rfloor, \tag{10}$$

where the $\lfloor \rfloor$ is the symbol of floor function which outputs the greatest integer less than or equal to the value enclosed.

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. Chen, F., Brown, G. M., and Song, M., "Overview of three-dimensional shape measurement using optical methods," *Opt. Eng.* **39**, 10–22 (2000).
2. Geng, J., "Structured-light 3d surface imaging: a tutorial," *IEEE Intelligent Transportation System Society* (31 March 2011).
3. Wust, C. and Capson, D. W., "Surface profile measurement using color fringe projection," *Machine Vision and Applications* **4**, 193–203 (1991).
4. Grin, P. M., Narasimhan, L. S., and Yee, S. R., "Generation of uniquely encoded light patterns for range data acquisition," *Pattern Recog.* **25**(6), 609–616 (1992).
5. Durdle, N. G., Thayyoor, J., and Raso, V. J., "An improved structured light technique for surface reconstruction of the human trunk,," *IEEE Canadian Conference of Electrical and Computer Engineering* **2**, 874–877 (1998).
6. Guan, C., Hassebrook, L. G., Lau, D. L., Yalla, V., and Casey, C., "Improved composite-pattern structured-light profilometry by means of postprocessing," *Opt. Eng.* **47**(9), 0972031–09720311 (2008).
7. Blais, F., "A review of 20 years of range sensor developement," *Proceedings of SPIE-IS&T Electronic Imaging* **5013**, 62–76 (2003).
8. Srinivasan, H.C., Liu, H. C., and Halioua, M., "Automated phase measuring profilometry: a phase mapping approachl," *Appl. Opt.* **24**, 185–188 (1985).
9. C., G., Hassebrook, G., L., and Lau, D. L., "Composite structured light pattern for three-dimensional video," *Opt. Express* **11**(5) (10 March 2003).
10. Halioua, M. and Liu, H., "Optical three-dimensional sensing by phase measuring profilometry," *Opt. Lasers Eng.* **11**, 185–215 (1989).
11. 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).
12. Song, J., Ho, Y. S., Lau, D. L., and Liu, K., "Universal phase unwarpping for phase measuring profilometry using geometry analysis," *Proc. SPIE* **10546**, 0B0–0B8 (2018).
13. 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).

14. O'Toole, M., Mather, J., and Kutulakos, K. N., "3d shape and indirect appearance by structured light transport," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(7), 1298–1312 (2016).
15. Lin, J., Jiang, K., and Chang, M., "A novel solution for camera occlusion in stereo vision technique," *Adv. Mech. Eng.* **2013**, 1–8 (2013).
16. "Road test: optoma ml750st." `https://www.projectorcentral.com/optoma-ml750st-projector-review-road-test.htm`.
17. Wang, Y., Liu, K., Lau, D. L., Hao, Q., and Hassebrook, L. G., "Maximum snr pattern strategy for phase shifting methods in structured light illumination," *J. Opt. Soc. Am.* **27**(9), 1962–1971 (2010).
18. Video Electronics Standards Association, *VESA and industry standards and guidelines for computer display monitor timing* (May 2007). Version 1.0.
19. Liu, K., Y.Wang, Lau, D. L., Hao, Q., and Hassebrook, L. G., "Gamma model and its analysis for phase measuring profilometry," *J. Opt. Soc. Am.* **27**(3), 553–562 (2010).