Proceeding of the 2015 IEEE
International Conference on Information and Automation
Lijiang, China, August 2015

3D Reconstruction of Specular Surface via a Novel Structured Light Approach

Haibo Lin and Zhan Song
Shenzhen Institutes of Advanced Technology
Chinese Academy of Sciences & The Chinese University of Hong Kong, Shenzhen, China
{hb.lin & zhan.song}@siat.ac.cn

***Abstract* – 3D measurement of specular surface is a difficult issue for optical-based structured light techniques. Due to strong reflectivity of target surface, the projected patterns are usually invisible to the camera. In this work, we present a novel structured light sensing approach by exploiting the high dynamic range imaging techniques. The target surface is scanned by the structured light device multiple times with various exposure levels. And then, images of the same pattern index are fused together to get a high dynamic range radiance map. By compressing the radiance map to the image space, a new set of images with more homogenous image intensity distribution and well exposure quality can be obtained. And finally, the 3D reconstruction is applied to these reconstructed images. A stainless steel sheet with very strong reflectance is experimented. And the results show that, the specular surface can be well reconstructed with satisfied accuracy.**

Index Terms – High Dynamic Range imaging, structured light sensing, 3D reconstruction, specular reflectance.

I. INTRODUCTION

Structured Light Sensing (SLS) is an important computer vision technology for 3D information acquisition [1]. The underlying principle of SLS is to project some features onto the target surface so as to solve the correspondence problem between the camera and projector. An ideal SLS system assumes that the light comes from the projector is directly projected to the target surface and the camera captures it without any loss or interference. However, the projecting and imaging procedures are usually affected by various factors in real applications. For instances, global illumination of the environment affects the patterns on the surface, and shiny surface may produce specular reflection and dark regions in the captured images. As a result, such difficult regions cannot be well reconstructed. Actually, 3D reconstruction of reflective or shiny surface is still an open issue for all optical-based 3D sensing technologies [2].

According to the classification in [3], current SLS coding schemes can be generally divided into temporally-based and spatially-based method. Spatial coding method only needs a single pattern to project, so it is simple and suitable for dynamic 3D applications. But the complexity in the decoding process reduces the robustness of spatial coding method. Due to one encoded pattern, the coding density is also limited. For temporal coding methods, a sequence of pattern images is usually demanded. Target surface is encoded by various illuminations along the time axis. The temporal SLS approaches can usually retrieve high density and accurate depth information [4]. For this reason, it is also widely adopted by off-the-shelf 3D scanners.

According to coding strategy, temporal coding method can be further categorized into binary code, *n*-ary code and hybrid code. For binary code, the patterns only have two illumination levels, which are coded as *0* and *1*. The codeword of each pixel is formed by a sequence of *0*s and *1*s, which is corresponding to every projected pattern. And the typical binary coding is *Gray code*. The advantage of Gray code is that consecutive codewords have a Hamming distance of one, and it is robust against noise. In [5], a temporally encoded structured light system based on Gray code is presented. Lens distortion is considered in the SLS model for both camera and projector, so as to obtain a balance in the accuracy of all components in the system. A sub-pixel feature detector was proposed to improve the 3D reconstruction accuracy. An *n*-ary code was proposed to reduce the number of patterns. Based on *n*-ary coding scheme, a color pattern was proposed in [6]. The mathematical basis for generating *n*-ary code was developed and the illumination model of SLS was analyzed. The *n*-ary code shares the same characteristic of Gray code, which has a Hamming distance of one between consecutive codewords. The hybrid code was proposed to improve the performance of 3D reconstruction, which is usually combined with the Gray code pattern and the phase shifting patterns [7]. However, the use of phase shifting method like sinusoidal patterns makes the whole coding scheme sensitive to the surface reflections. In our previous work, a pure binary coding scheme was investigated [8]. The SLS pattern was constructed by the combination of Gray code and binary strip shifting patterns. Edges of binary strips (alternate dark and bright strips) were detected by the proposed feature detectors with sub-pixel accuracy. We have shown that such a binary coding method can achieve much higher robustness to deal with reflective surfaces. However, if the target surface is very shiny and makes the projected strips broken, the method in [8] is also not applicable. How to directly scan the very shiny surface is still a challenging issue in SLS domain.

In this work, and High Dynamic Ranging (HDR)-based structured light method is investigated for the 3D

This work was supported in part by the National Natural Science Foundation of China (61375041), Shenzhen Science Plan (JCYJ20120903092425971, JCYJ20130402113127502, JCY20140509174140685), and the Introduced Innovative R&D Team of Guangdong Province-Robot and Intelligent Information Technology R&D Team (201001D0104648280).

reconstruction of very reflective surfaces. Gray code and striping shifting method is adopted as the coding strategy. Since the target surface is quite reflective, the captured images contain a large area of saturation and darkness. To recover 3D information in these areas, the surface is scanned multiple times with different camera exposing times. And then an HDR method is proposed for the fusion of structured light images. We use different exposed images to recover the response function of the camera. With the response function, multiple exposed images can be fused into a single, high dynamic range radiance map. The pixel values in the radiance map are proportional to the true radiance values of the scene. With the image fusion process, over-exposure and under-exposure regions can be adjusted to a reasonable image intensity range, which can provide useful image information for the decoding process. The experiments are conducted on a stainless steel sheet which has very strong reflective surface. And the results show that difficult regions can we well reconstructed by the proposed SLS method with satisfied accuracy.

II. BINARY CODING STRATEGY

Gray code pattern is widely used for its simplicity and robustness. If a binary Gray code pattern of codeword length n is to be used, an image sequence consisting of n+1 binary strip patterns need to be projected sequentially to encode the scene. With that, the scene image can be separated into 2^n distinguishable regions each with a unique different codeword, and the image positions are encoded with only pixel accuracy. Phase shifting and line shifting are two techniques often used with Gray code to achieve sub-pixel accuracy. In our system, strip shifting is used in place of the line shifting strategy, which is proposed in [8]. Edges of binary strips are easier to detect in image data than lines. With edges more precisely located, higher reconstruction accuracy and stronger robustness can be attained.

In our coding scheme, a series of n Gray code patterns are first projected in order to divide the target surface into 2^n sub-regions and each region has a unique n-bit long Gray codeword. Suppose that each of such sub-regions is m pixel width on the projector's pattern generation plane. Then a strip pattern in half of the finest strip width of the Gray code sequence is shifted one pixel m-1 times, each with a separate image capture by the camera, to encode each sub-region with additional bits. Such a periodic pattern by itself has periodic ambiguity of m projector pixels in distinguishing different points of the target surface. However, by combining the two codewords together, one from the Gray code projection and the other from the strip shifting pattern, the m sub-divisions can be introduced to each of the 2^n sub-regions to achieve finer 3D reconstruction with sub-pixel accuracy. The procedure can be expressed as:

$$G = 2^n,\ S = m,\ P = G + S \tag{1}$$

where G is the Gray codeword, S is the local codeword generated by the strip shifting pattern and P is the final unique codeword. Then, with the unique sub-pixel accuracy codeword, we can obtain accurate 3D information.

The pattern sequence is composed of two type patterns, as shown in Fig. 1. The top patterns are produced by Gray code and the bottom patterns are produced by strip shifting.

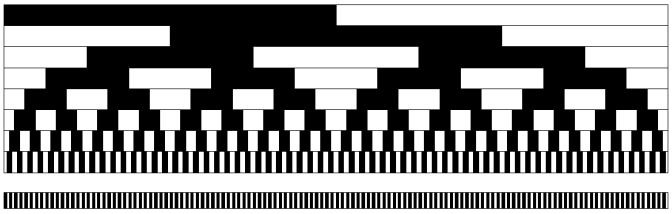

Fig. 1. The use of Gray code together with strip shifting for 3D reconstruction.

III. FUSION OF MULTIPLE-EXPOSURE SLS IMAGES

A. Generation of Radiance Maps

High dynamic range radiance map is an important technology in computer vision and photography [9]. With the HDR technique, images with different exposing times can be fused into a single, high dynamic range radiance map. To get the radiance map, it assumes the scene is static and the lighting changes can be ignored. The irradiance value for each pixel is assumed constantly. So the imaging model can be written as:

$$Z_{ij} = f(E_i \Delta t_j) \tag{2}$$

where E_i denotes the irradiance of the ith pixel, Δt_j denotes the exposure time of the j index, and Z_{ij} denotes the gray level of the ith pixel at j index exposure. And f is a function of the system. The function f is assumed monotonic and invertible, and let function $g=\ln f^{-1}$, (2) can be rewritten as:

$$g(Z_{ij}) = \ln E_i + \ln \Delta t_j \tag{3}$$

In (3), the unknowns are the irradiance E_i and the function g. To recover the irradiances, we have to know the function g. On the other hand, we can see the function g has the finite number of values due to the finite number values of Z_{ij}. Then we can adopt a least-squares fitting algorithm with pixel values in different exposed photographs to solve the function g. After the response function of the camera system was gained, the irradiances can be easily recovered. Thus, the multiple images can be fused together into a single, high dynamic range radiance map. The pixel values in the radiance map are proportional to the true radiance values of the scene.

B. High Dynamic Range Compression

The range of radiance map is high and we need normal images to recover 3D information, so we have to translate the radiance map to an image. If we directly map the radiance map to the image linearly, a lot of details will be lost. A feasible method is to compress the gradients of the high dynamic range radiance map [10]. Main steps of the process include: 1) generate a gradient image from the radiance map,

2) compress the gradient image, and 3) produce a new image from the gradient image.

Firstly, we get a gradient image from the radiance map. For each pixel, we use the difference between adjacent pixels to denote the gradient as:

$$\nabla H\ (x,y) \approx (H(x+1,y)-H(x,y), H(x,y+1)-H(x,y)) \quad (4)$$

where $H(x, y)$ indicates the pixel value at image point (x, y).

Then, an attenuation function is used to compress the gradient image as expressed by (5). So that the gradients of larger magnitude are attenuated and gradients of small magnitudes can be slightly magnified.

$$G(x,y) = \nabla H(x,y)\Phi(x,y) \quad (5)$$

In (5), the attenuation function $\Phi(x,y)$ is computed from the Gaussian pyramid, and the gradient of each level k in Gaussian pyramid is denoted as $\nabla H_k(x,y)$. A scaling factor is used to compute $\Phi(x,y)$ as shown in (6). And the attenuation function of each level in the pyramid can be computed by multiplying the scale factor with the attenuation function at higher level.

$$\varphi_k(x,y) = \frac{\alpha}{\|\nabla H_k(x,y)\|}\left(\frac{\|\nabla H_k(x,y)\|}{\alpha}\right)^{\beta} \quad (6)$$

And then, the Poisson equation is used to produce a new image as:

$$\nabla^2 I = divG \quad (7)$$

To reduce the use of the computer memory when computing the new image, we use the Gauss-seidel method to solve the equations, where the radiance map is used as the initial value.

C. *Refine of the HDR Images*

After the image fusion, we can get a new set of structured light images. But it is difficult to acquire 3D information with it, since the generated HDR images do not comply with a unified surrounding condition after fusion. Therefore, it is necessary to refine the new image sequence.

We select images at an appropriate exposure time to adjust the new images. The formulation can be expressed as:

$$I(x,y) = I_0(x,y)\exp(I_1(x,y) - I_0(x,y)) \quad (8)$$

where x, y indicate the index of images, I_0 means the pixel value of the original image, I_1 means the pixel value of the new image, and I is the pixel value of the final image. And the range of the pixel value in (8) is set to 0~1. Equation (8) means that new images are used to adjust original images. If the pixel value in the new image is larger than the original one, the value of exponential function is larger than 1, so the value in the final image is larger than the original one.

We will prove the range of pixel values in final images is also 0 to 1 as follows. Firstly, we can see the smallest value of I is 0 when I_0 is 0, because the exponential value cannot be negative. Then, we can find the largest value of I. While I_1 remains unchanged, the first derivative of I versus I_0 is shown in (9). So, the first derivative is positive when the range of I_0 is 0 to 1. Therefore, pixel values in the final image range from 0 to 1.

$$I' = (1-I_0)e^{I_1-I_0} \quad (9)$$

After the final refined HDR images are obtained, we can use them to decode the 3D information of the shiny surface.

IV. Experimental Results

The experimental setup is configured with a DLP projector with resolution of 1024×768 pixels, and a camera with resolution of 2080×1552 pixels. The algorithm is implemented with the Matlab 2012a. The system is calibrated by the method proposed in [11]. And the working distance ranges from 800mm to 1000mm.

We conduct the experiments with a piece of stainless steel sheet as shown in Fig. 3. Surface of the stainless steel is quite reflective. While the patterns are projected on it, a large area of over-exposure and darkness can be captured by the camera. For these area, there are no information can be extracted from the images as well as lack of 3D information in the reconstructed 3D models.

Before the HDR image generation, we need to compute the high dynamic range radiance maps with the images at different exposures. The g curve in our system is as shown in Fig. 2. In general, the g curve is smooth as we expected. With the radiance maps, a new homogeneous and high-contrast image sequence can be generated.

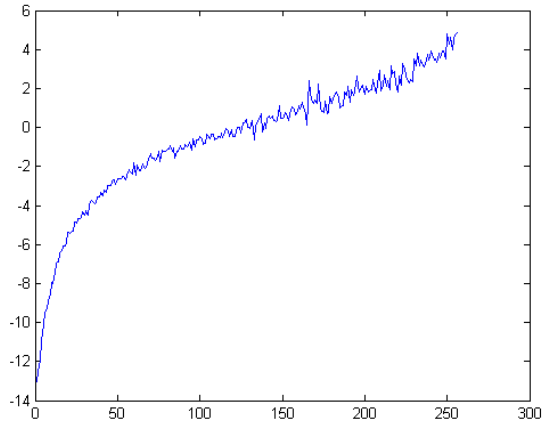

Fig. 2 The g curve is derived from digital values at one pixel for 6 different known exposures.

The SLS device scans the target with different camera exposing times, e.g. 2ms, 3ms, 4ms, 5ms, 6ms, 7ms. From the images we can see that, with the increase of exposure time, the over-exposure regions are expanded and the under-exposure regions are shrunk. For each exposure level, we reconstruct the steel sheet surface with the original image sequence for comparisons. The 3D reconstruction results are as shown in Fig. 3(a-f). From the results we can see that, with a single

exposure time, we cannot get a complete 3D reconstruction surface. When the exposure is low, the shiny area can be reconstructed, but the dark area is missing. When the exposure is high, the shiny area is saturated, and the 3D information of the shiny area is lost.

By implementing the proposed HDR method, a final image with homogenous intensity distribution and well exposure quality can be obtained as shown by Fig. 3g. Based upon these HDR images, all the surface parts with under-exposure and over-exposure in the original images can be well reconstructed.

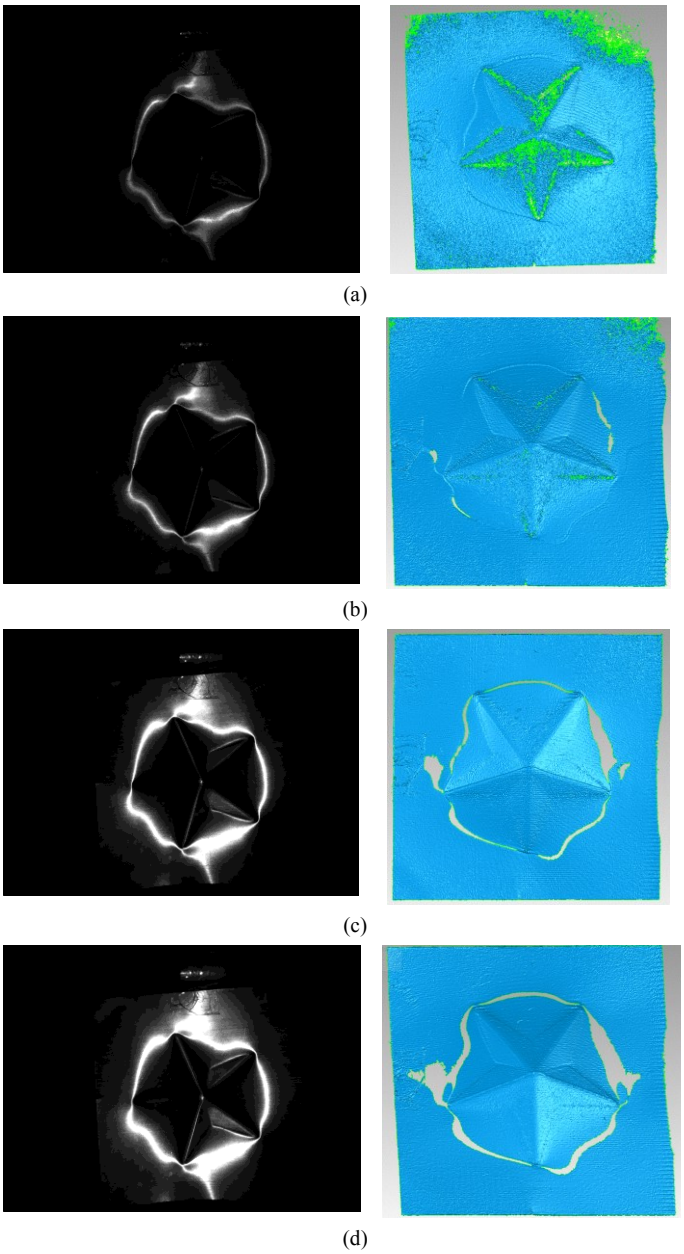

Fig. 3. 3D reconstruction results with different exposure levels. (a) the exposure time is 2 ms. (b) the exposure time is 3 ms. (c) the exposure time is 4 ms. (d) the exposure time is 5 ms. (e) the exposure time is 6 ms. (f) the exposure time is 7 ms. (g) the generated HDR image and its 3D reconstruction result.

V. Conclusion

This paper describes a novel structured light method for the 3D scanning of very reflective surfaces. The coding scheme of the structured light is based on gray code with strip shifting pattern. Binary strip edges are encoded for 3D reconstruction. While the patterns are projected on the surface with very specular reflection, the strips are usually broken and make the decoding unsuccessfully. To solve this problem, an HDR method is investigated. The surface is firstly scanned by the SLS with different exposure levels. And then, the captured images of the same pattern are fused together to create a new radiance map. After compressing the range of radiance map, a new image can be produced. Then, a proper original image sequence is selected and the fused image sequence is used to adjust it to get the refined HDR structured light images. Finally, the fused HDR image sequence is decoded to reconstruct the reflective surface. The experiment is conducted with a specular stainless steel sheet. And the results show that, the specular surface can be well reconstructed with satisfied accuracy. Future work can address how to reduce the image

numbers of various exposure levels, so as to boost the 3D scanning efficiency.

REFERENCES

[1] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, "A state of the art in structured light patterns for surface profilometry", *Pattern Recognition,* no. 43, vol. 8, pp. 2666-2680, 2010.

[2] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Structured light 3D scanning in the presence of global illumination," *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE,* pp. 713-720, 2011.

[3] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition*, vol. 37, no. 4, pp. 827-849, 2004.

[4] T. Pribanić, S. Mrvoš, and J. Salvi. "Efficient multiple phase shift patterns for dense 3D acquisition in structured light scanning," *Image and Vision Computing,* vol. 28, no. 8, pp. 1255-1266, 2010.

[5] R. J. Valkenburg, A. M. McIvor, "Accurate 3D measurement using a structured light system," *Image and Vision Computing,* vol. 16, no. 2, pp. 99-110, 1998.

[6] D. Caspi, N. Kiryati, and J. Shamir, "Range imaging with adaptive color structured light," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* vol. 20, no. 5, pp. 470-480, 1998.

[7] F. Sadlo, T. Weyrich, R. Peikert, and et al. "A practical structured light acquisition system for point-based geometry and texture," *Point-Based Graphics, 2005. Eurographics/IEEE VGTC Symposium Proceedings. IEEE,* pp. 89-145, 2005.

[8] Z. Song and R. Chung, "Off-the-shelf structured light-based system for accurate 3D reconstruction," *HKIE Transactions,* vol. 15, no. 4, pp. 44-51, 2008.

[9] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," *ACM SIGGRAPH 2008 classes. ACM,* pp. 31, 2008.

[10] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Transactions on Graphics (TOG). ACM,* vol. 21, no. 3, pp. 249-256, 2002.

[11] Z. Song and R. Chung, "Use of LCD panel for calibrating structured light-based range sensing system," *IEEE Trans. Instrumentation and Measurement,* vol. 57, no. 11, pp. 2623-2630, Nov. 2008.