Proceedings of The 2018 IEEE International
Conference on Real-time Computing and Robotics
August 1-5, 2018, Kandima, Maldives

A Robust Feature Detection Method for an Infrared Single-Shot Structured Light System

Chu Shi, Jianyang Feng, Suming Tang, Zhan Song*, *Member, IEEE*

***Abstract*— This paper introduces a robust feature detection method based on an infrared structured light system (SLS). Subject to the absorption of infrared light by human face, the projected grid-pattern is difficult to be extracted directly. In the proposed image preprocessing stage, two nonlinear filters are firstly applied and their subtraction is calculated. Then, a local adaptive threshold filtering procedure is implemented to boost the contrast around grid-lines and to suppress the image background. Based on the enhanced structured light image, a cross filter is applied to extract candidate grid-points. Finally, a symmetry-based feature detector is introduced to localize the grid-points precisely. Real human face is used in the experiments. And results show that, most of the grid-points can be accurately detected and 3D model of human face can be successfully retrieved via single projection of an infrared grid-line structured light pattern.**

***Index Terms* - Structured light; grid-point detection; image processing; 3D reconstruction.**

I. INTRODUCTION

Structured light-based 3D sensing technology has been widely used in various applications such as robot navigation, 3D inspection and human-machine interaction etc. To realize 3D reconstruction of dynamic object and scene, spatially encoded structured light methods are usually adopted, which can perform the coding procedure inside single projection pattern [1-2]. For all the single-shot structured light systems, feature detection is always a critical step, which directly affects the final decoding procedure and 3D reconstruction precision. In real applications, to make the projection invasive to human eyes as well as to improve the system robustness to ambient lights, the infrared light source is usually adopted like the Kinect and Realsense products did [3]. Compared with the image quality under visible wavelength, image quality and contrast are usually lower in the infrared waveband. And thus makes the detection of projected feature more challenging in IR structured light images.

In this work, a single-projection structured light means was investigated for the 3D reconstruction of human face under IR light. The projected pattern is designed as grid-lines, and some geometrical shapes are imbedded into the grid cells as pattern elements. The coding strategy is designed based on pseudorandom array, and grid-point is defined as the pattern feature points. Subject to the low image contrast of infrared images, traditional feature detectors like LOG (Laplacian of Gaussian) [4], HARRIS [5], SIFT (Scale-Invariant Feature Transform) [6], and grid detector [7] failed to extract the grid-points successfully. To localize the grid-points precisely in the low contrast IR structured light images, a two-step grid-point detection algorithm is proposed in this work. The low quality IR image is firstly enhanced in the preprocessing stage based on a local adaptive filtering operation. Based on the enhanced structured light image, a symmetry-based feature detector is introduced for the precise localization of feature points. Experimental results with real human faces demonstrate that the proposed method is robust and efficient, and the grid-points with very low contrast can be successfully detected with high precision.

Rest of the paper is organized as follows. The proposed preprocess algorithm for IR structured light images is presented in Section II. The local symmetry-based grid-point detection method is introduced in Section III. Experimental results are provided in Section IV and the conclusion is offered in Section V.

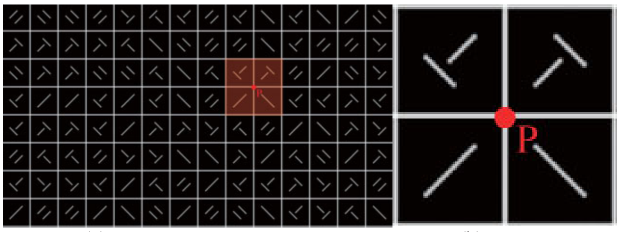
II. PREPROCESS OF IR STRUCTURED LIGHT IMAGE

The structured light pattern that proposed in our previous work [8] was used for 3D reconstruction, while the light source is changed to an IR one which works at 850 nm. The pattern is composed with some grid-lines and 8 different geometrical shapes. The patter size is 65×63 and the coding window size is 2×2. Grid-point is defined as the coding feature points as shown by Fig. 1. Each coding window is different from the others in the pattern, and that makes the feature point P is unique from all the other feature points. The pattern is projected on the target surface via the IR light source, and a camera with IR filter is used to capture the structured light image. By means of recognizing the grid-points and the pattern information in each coding window, the depth information of object can be calculated via system calibration parameters and triangulation approach [9-10].

This work was supported in part by the National Key R&D Program of China (No. 2017YFB1103602), Shenzhen Science Plan (KQJSCX20170731165108047, JCYJ20170413152535587), the National Natural Science Foundation of China (Nos. 61773363, 51705513, U1613213, U1713213,), and Shenzhen Engineering Laboratory for 3D Content Generating Technologies (No. [2017] 476).

C. Shi and F.J. Yang are with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China and Southwest Petroleum University, ChengDu, China ({chu.shi; jy.feng}@siat.ac.cn).
Suming Tang and Zhan Song are with CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. Shenzhen, China (tel: 86-755-86392100, fax: 86-755-86392299, e-mail: {sm.tang; zhan.song}@siat.ac.cn).

978-1-5386-6869-6/18/$31.00 © 2018 IEEE



(a) (b)

Fig. 1. (a) a part of the grid-line structured light pattern; (b) grid-point P is defind as the coding feature point.

Subject to the absorption of IR lights by the human skin, the captured grid-line pattern images are usually blurred with very low contrast. To detect the grid-points precisely, the IR structured light images must be preprocessed before the implementation of feature detection algorithm. Inspired by the method proposed in [11-12], a nonlinear filtering operator is introduced to boost the high frequency image component and compress the image low frequency. And thus, contrast between grid-lines and background in the IR image can be improved.

Suppose I_0 indicate the original structured light image, by the implementation of bilateral filter, we can get image I_1. The usage of bilateral filter is to remove image noise and preserve grid-line features. Based on the two images, we can define two nonlinear functions $G1(x)$ and $G2(x)$ as:

$$G1(x)=\frac{1}{2}\times(\sqrt{\frac{x}{1-x}}-\sqrt{\frac{1-x}{x}}), \tag{1}$$

$$\Delta=G1(x_1)-G1(x_2), \tag{2}$$

$$G2(\Delta)=\frac{1}{2}\times\left(\frac{\Delta}{\sqrt{1+(\Delta)^2}}+1\right), \tag{3}$$

where x is the normalized image intensity at image point (i, j), x_1 and x_2 refer to the image intensity values in I_0 and I_1 respectively. In the structured light images, the projected grid-lines usually have larger intensity compared with neighboring image areas. As illustrated in Fig. 2a, while the input image pixel value x>0.5, the image intensity can be stretched or increased especially for the image point with intensity range of [0.9, 1]. And for the image point with low intensity, their values can be further decreased by the function of $G1(x)$.

For the function $G2(\Delta)$, we used a constant a to replace $G1(x_2)$ in (2) as:

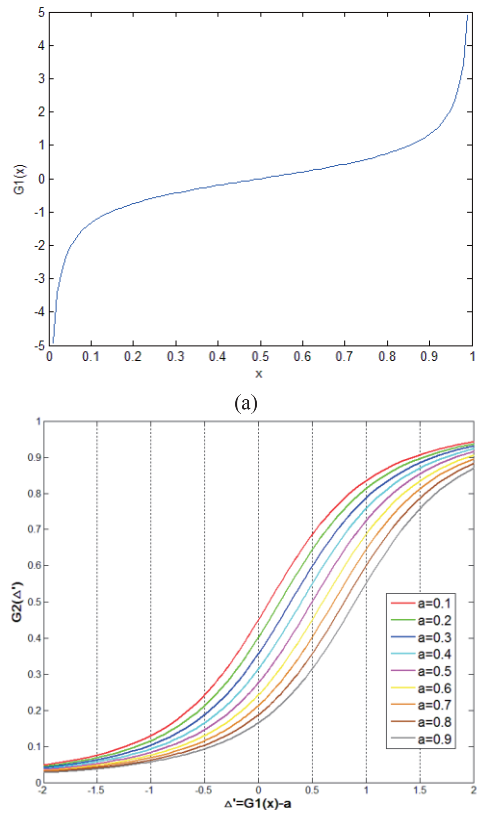
$$\Delta'=G1(x_1)-a. \tag{4}$$

It's hard to plot the real character curve of difference between two nonlinear functions, because the type of Δ function after subtraction between two nonlinear functions may be nonlinear or linear. An alternative means is to enumerate a series of a value. With different a value, the rough character curve can be e estimated approximately from (3). From Fig. 2b, we can see that, all curves show similar trend. Small value of Δ means small difference between two input images. Value of Δ around 0 shows steep slope on the curve, and this part will be stretched by $G2(\Delta)$. After nonlinear subtraction operation, contrast around grid-lines in the image can be improved.

With above process, local contrast around grid-lines can be improved and dynamic range of background can be compressed. In addition, the range of foreground and background also become similar. Then, an adaptive local threshold is used to separate foreground grid-lines from the background. For each image point, we calculate the average intensity of local image patch for example 10×10 pixel. And the average value is used as the local threshold. In this way, the pixel values will be compressed by attenuation coefficient β if they are lower than their local average value. Otherwise, this image point can be enhanced by coefficient γ, which can be set empirically.

$$I(i,j)=\begin{cases}\beta\times I(i,j) & I(i,j)<mean(i,j)\\ \gamma\times I(i,j) & I(i,j)>mean(i,j)\end{cases} \tag{5}$$

$$\beta=\sqrt{1-\frac{I(i,j)}{mean(i,j)}} \tag{6}$$



(a)

(b)

Fig. 2. (a) plot of function $G1(x)$; (b) plot of function $G2(\Delta')$.

Generally, the original image is firstly applied with a bilateral filter to remove noise and to preserve local grid-line edges. And then a nonlinear subtraction is used to get high frequency image areas. Finally, a local adaptive filter is used to improve the contrast around grid-lines furtherly. With above procedure, background in the image can be greatly compressed and the local contrast of grind-lines can be distinctly improved.

III. Grid-Point Detection from IR SLS Image

By the observation of grid pattern structure, we can find that, the local symmetry property preserved for grid-point in a local image area. Such property also can be preserved under perspective projection. Based on this observation, a local symmetry based grid-point feature detector is used for the grid-point localization.

Firstly, a cross-mask is first defined and used as a filter. With cross filtering, response values at all image points can be calculated, and a map named response image can be generated. With the cross filter, most non grid-points can be filtered out. And then, we used the K-means method to separate remaining candidate points. By the means of Non-maximum suppression, candidate feature points around the true grid-points can be extracted.

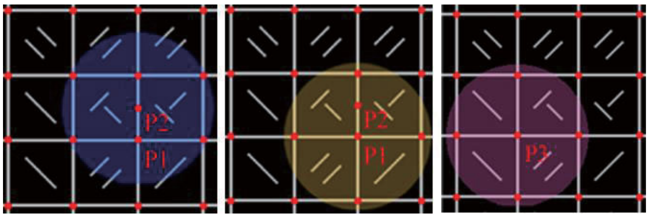


Fig. 3. Real grid-point like P1 and P3 have stronger local symmetry property than other candidate feautre points like P2.

According to the local symmetry property of grid-point as shown by Fig. 3, a circular mask is defined to remained feature points obtained after cross filtering. To measure the strength of the rotational symmetry, we used the coefficient of correlation between the circular window and the 90° rotation of it. We can write the correlation coefficient δ as:

$$\delta = \frac{\sum_m \sum_n \left(C_{mn} - C'_{mn}\right)^2}{\sum_m \sum_n \left(C_{mn} - \bar{C}\right)^2} \quad , \tag{7}$$

where m, n refer to individual element position of the circular mask centered by candidate point, C is the circular mask and C' is the mask rotated 90˚ around candidate point, $\bar{C}$ is the circular mask filled with mean value of the mask C. Eqn. (7) can express the difference between mask C and C' by calculating the variance of the corresponding position pixel value in each circular mask, and then normalize the results by calculating the variance of mask C itself. The value of δ represents the level of local symmetry on each candidate point. The smaller the value δ on each point, the better the symmetry. By setting a threshold of δ, e.g. δ=0.65, to find out the real grid-points. Based on the extracted feature points, decoding methods can be implemented for their 3D coordinate recovery [13].

IV. Experimental Results

The structured light pattern is proejcted by an DLP projector, which is equiped with an IR LED works at 850 nm. And the camera is also equiped with an IR filter. The image resolution is set to 2000×1500 pixels. Fig. 4a shows the projected grid-line patter on a real human face. We can see that, contrast of grid-lines w.r.t human face is relative low and most grid-points cannot be recognized even by human eyes. Without the implemention of preprocessing, the propsoed feature detection is applied directly on the orginal structured light image. The grid-point detection reuslt is as shown by Fig. 4b, we can see that, a lot of grid-points are missed.

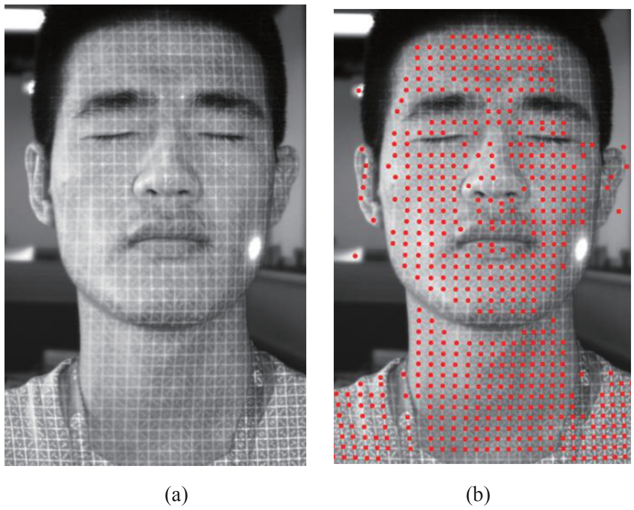


Fig. 4. (a) human face with grid-pattern illumiation; (b) feature points detection result without preprocessing.

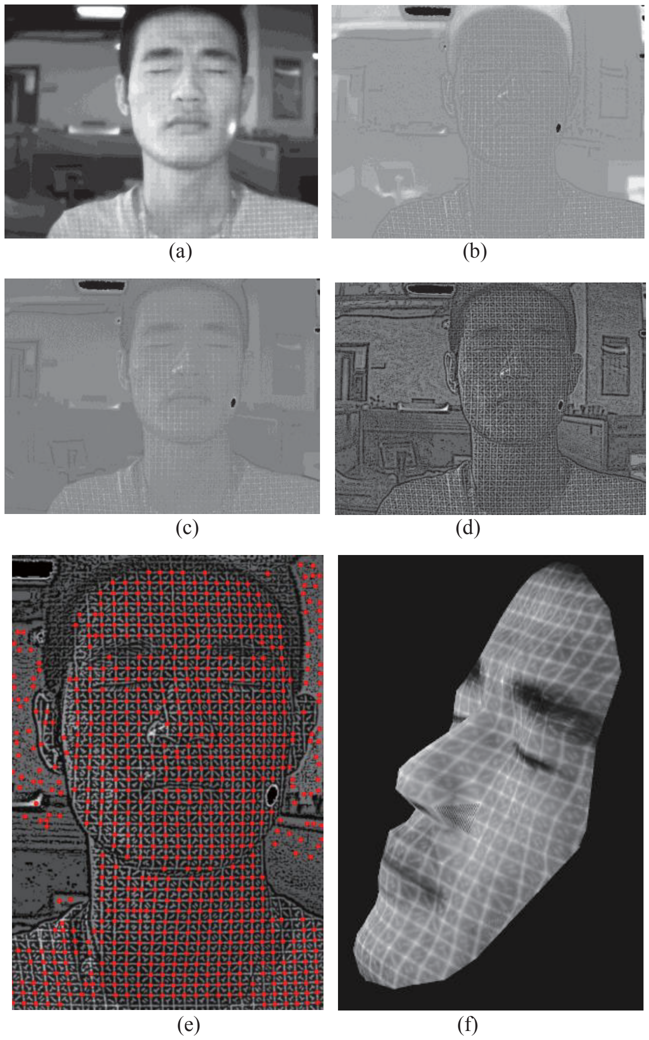


Fig. 5. Images obatined at each stage in the propsoed preprocessing algorithm.

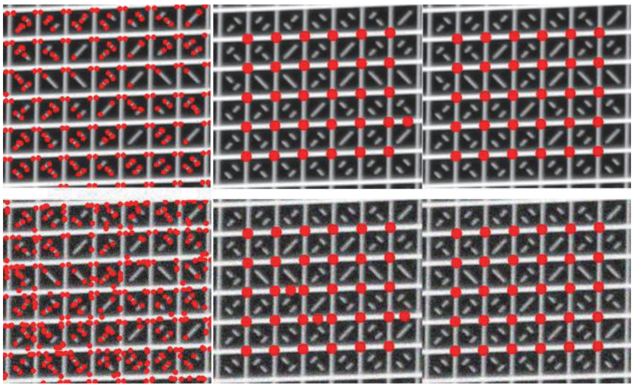
Fig. 5 shows the working flow and output images in the preprocessing algorithm at each processing stage. Fig. 5a show the image after bilteral filtering, the overall image looks blurred, but local constrat arround grid-lines can be preserved. Fig. 5b shows the image after substraction, where background image points can be greatly compressed. Fig. 5c shows the image after linear transformation. Fig. 5d shows that resutls after local adaptive threshold filerting. Based on the image as shown by Fig. 5d, the proposed feature detector was applied for grid-point detection as shown by Fig. 5e. Compared with the result shown by Fig. 4b, most of the grid-points can be successfully detected with precise localization. And the final 3D reconstructed model with texure is as shown by Fig. 5f. To verify robustness of the feature detector to color texture, a curvered color paper was used as shown by Fig. 6a. The detected grid-points and reconstructed 3D curve surface were shown by Fig. 6b and Fig. 6c. The results showed that, most of the grid-points can be detected correctly by the proposed feature detector.

(a) (b)

(c)

Fig. 6. Experiment with a curvered color paper surface. (a) original image; (b) grid-point detection result; (c) reconstructed 3D model with texture.

Moreover, to test the robustness of the proposed algorithm, Gaussian noise with σ=0.03 is added to the image as shown in Fig. 7 in second row. From the results we can see that, even without extra noise, the Harris detector cannot extract the grid-points correctly as shown by Fig. 7a. Fig. 7b shows the feature detection results only using cross filter. We can see that, affected by image noise, some false detections aroused. Fig. 7c shows the feature detection by the proposed operator, it can precisely detect the grid-points under image noise.



(a) (b) (c)

Fig. 7. Grid-point detection results of three operators. (a) result by the Harris corner detector; (b) result using the cross filter; (c) result by the propsoed feature detector.

V. Conclusion and Future Work

This paper presents a novel feature detection method for grid-line structured light pattern. Subject to the low image quality obtained under infrared pattern illumination. An image preprocessing algorithm is introduced to improve the image quality. By the introduced nonlinear filters and their subtraction operation, local edges around grid-lines can be preserved while the background image can be compressed. A local adaptive filter is used to improve the contrast around grid-lines furtherly. Based on the improved IR structured light image, a cross template is applied to extract potential grid-points. And then a local symmetry-based feature detector is introduced to localize the grid-points precisely. Then experiments are conducted on real human face based on an IR structured light system. And the results show that, most of the grid-points can be successfully extracted, and 3D model of human face can be reconstructed via single IR pattern projection. Future works can address how to improve the pattern coding density so as to improve the reconstructed point cloud density.

References

[1] J. Salvi, J. Batlle, and E. M. Mouaddib, “A robust-coded pattern projection for dynamic 3D scene measurement,” *Pattern Recognition Letters*, vol.19, no.11, pp.1055-1065, 1998.

[2] A. O. Ulusoy, F. Calakli, and G. Taubin, “One-shot scanning using De Bruijn spaced grids,” *IEEE International Conference on Computer Vision Workshops*, pp.1786-1792, 2009.

[3] R. A. Newcombe, S. Izadi, O. Hillige, et al. “KinectFusion: Real-time dense surface mapping and tracking,” *IEEE International Symposium on Mixed and Augmented Reality*, pp.127-136, 2011.

[4] G. David and D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol.60, no.2, pp.91-110, 2004.

[5] C. Harris, “A combined corner and edge detector,” *Proc. of Alvey Vision Conference*, vol.3, pp.147-151, 1988.

[6] D. Lowe and G. David, “Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image,” *International Journal of Computer Vision*, vol. 64, no. 2, pp.91-110, 2004.

[7] Z. Song and R. Chung, “Grid point extraction exploiting point symmetry in a pseudo-random color pattern,” *IEEE International Conference on Image Processing*, pp. 1956-1959, 2008.

[8] H.L. Jiang and Z. Song, “A robust feature detection algorithm for the binary encoded single-shot structured light system,” *IEEE International Conference on Information and Automation*, 2017.

[9] Y. Lei, K. Bengtson, L. Li, and J. Allebach, “Design and decoding of an m-array pattern for low cost structured light 3D reconstruction

systems," *IEEE International Conference on Image Processing*, pp. 2168-2172, 2013.

[10] J. Xu, "Real-time 3D shape measurement system based on single structure light pattern," *IEEE International Conference on Robotics and Automation*, pp. 121-126, 2010.

[11] G. Deng, "A generalized unsharp masking algorithm," *IEEE Transactions on Image Processing*, vol. 20, no. 5, pp.1249-1261, 2011.

[12] H. Shvaytser and S. Peleg, "Inversion of picture operators," *Pattern Recognition Letters*, vol.5, no.1, pp.49-61, 1987.

[13] S. M. Tang, X. Zhang and Z. Song, "3D surface reconstruction via a robust binary shape-coded structured light method," *Optical Engineering*, vol. 56, no. 1, p. 014102, 2017.