

Colored Structured Light Patterns with Binocular Cameras for Fast and Low-cost 3-D Reconstruction

Hailun Xia

1. Institute of Automation, Chinese Academy of Sciences, Beijing, China
2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China
xiahailun2017@ia.ac.cn

Jie Tan, Ke Wu

Institute of Automation,
Chinese Academy of Sciences
Beijing, China
tan.jie@163.com, wuke2016@ia.ac.cn

Abstract—Gray code is a commonly used coding strategy in structured light techniques. However, projecting the large number of patterns required by it slows down the reconstruction. In this paper, a colored Gray code was generated to reduce the number of projected patterns and cut down the projecting time, which appeared to be a major time-consuming part especially for low-cost projectors. Meanwhile, two cameras were used to improve the performance of color stripes matching. Therefore, the complex considerations about the spectral characteristics of devices and the calibration of projectors could be avoided, making the system easy to use. Inverse colored patterns were also projected to assist the decoding procedure. The effects of the combination of colored line-shifting methods and different threshold parameters were also studied. A 130mm*100mm statue was reconstructed into a dense point cloud with 214,4848 points using our approach. Evaluations showed that the system had an average deviation of 0.02mm, and its speed can be further elevated with GPUs.

Keywords—structured light; 3-D reconstruction; colored code; stereo cameras

I. INTRODUCTION

Structured light methods are widely used especially for reconstructing texture-less surfaces[1]. With cameras and projectors getting cheaper, structured light systems are expected to be an affordable solution for more 3-D reconstruction tasks[2]. However, low-cost devices have their own characteristics. The differences between professional projectors and low-cost ones can make it harder to reach certain requirements in terms of accuracy, resolution, speed, and ease of use.

As computing becomes faster with parallel processing and the use of GPUs, the projecting period could appear as the bottleneck of speed elevating. With low-cost projectors, the time needed to load pictures is even longer, making the picture gathering procedure unbearably long. Therefore, it would be beneficial to finish the reconstruction with less patterns.

Gray code is one of the most commonly used structured light approaches[1]. Considering how time-consuming it is, adding extra information may be a good thought. Color information is widely used in spatial coding methods[1], but not quite popular in temporal ones. Because it is easier to identify a color if there are sufficient different colors around

it to form a spatial codeword together. Meanwhile, the decoding algorithm can search for the best correspondences based on similarity, reducing the influences caused by color distortions. Yet in temporal coding strategies, every pixel is decoded without neighbouring information or similarity calculation. The results can be severely affected, because the color stripes projected by the projector and captured by the camera could be extremely different from the original ones[3].

However, with one more camera, the color stripes matching problem between the camera and the projector will be changed to the matching problem between two cameras, thus becoming significantly simpler[4]. In fact, the colored Gray code method with eight basic colors also gained better resolution, comparing with traditional Gray code in our experiments. Because although the thinnest black-and-white stripes of Gray code may get mixed and blurring, the thinnest color stripes still have a special color changing pattern and can form correspondences in binocular cameras.

In this paper, a structured light 3-D reconstruction method with binocular cameras and a low-cost projector was presented. Temporal coding method was used to get better resolution. Speed was gained by colored Gray code patterns, which reduced the number of patterns by three times thus significantly shortened the projecting procedure, while the decoding time did not increase much. Each of the column had a Hamming distance of 1 with its adjacent columns to suppress mismatching cost. Experiments combining colored Gray code with colored line-shifting method were done, attempting to gain better resolution. Comparisons between the presence and absence of inverse color patterns were made as well. Some of the experiment pictures are shown in Fig. 1. Finally, a system was built with satisfying resolution, better speed, cheap devices and simple procedures, friendly to beginners attempting to use it for various tasks.

II. RELATED WORK

Plentiful methods emerged for the purpose of 3-D reconstruction[5]. Among them, coded structured light methods have been well studied and used for being cheap and reliable for surfaces with less texture. Patterns with coding information are actively projected onto the surfaces to label each region, thus adding sufficient information for correspondence calculation. There are mainly two types of coding strategies, space coding and temporal coding.

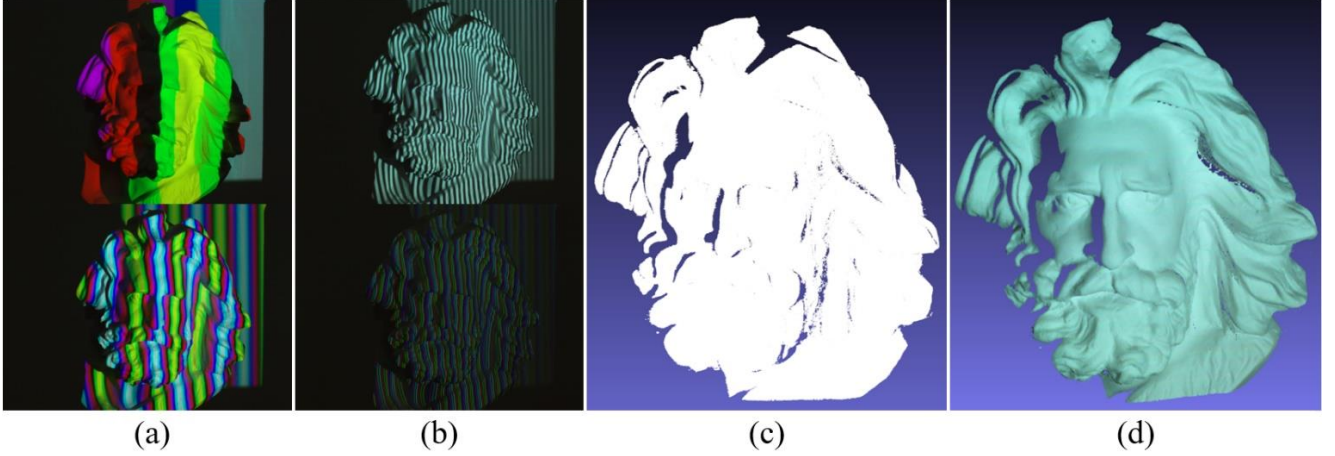


Figure 1. The results of the reconstruction. (a) The statue under colored Gray code patterns. (b) Top: Projecting traditional Gray code patterns as comparison. Bottom: Projecting colored line-shifting patterns to increase resolution. (c) The reconstructed point cloud using colored Gray code. (d) The reconstructed statue with added color values.

Space coding methods assign a codeword to a point according to its neighbouring points. These methods only require one shot of the scene to reconstruct, with color information widely used. Reconstructions can be quickly done with them, yet the point cloud is usually not very dense[6].

Temporal coding methods project a series of patterns onto the surfaces. The codes at each pattern are concatenated in a roll to generate a whole codeword. Binary code patterns were presented in 1981. A few years later, Gray code appeared in structured light coding method, and became popular[1]. With every codeword having a Hamming distance of 1 with its neighbours, Gray code is more robust to noise. Generally, temporal methods can achieve better resolution and accuracy compared with space coding methods. Even better results can be acquired together with other methods like phase-shifting[7] or line-shifting[8], for they can provide sub-pixel resolution. Yet temporal methods are suitable to run in real-time.

In 1998, Caspi et al. added information of color and intensity into Gray code with a thoroughly explored of the mathematical model, taking spectral and reflectance characteristics into account[3].

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = A \begin{bmatrix} k_R & 0 & 0 \\ 0 & k_G & 0 \\ 0 & 0 & k_B \end{bmatrix} P \begin{bmatrix} r \\ g \\ b \end{bmatrix} + \begin{bmatrix} R_0 \\ G_0 \\ B_0 \end{bmatrix} \quad (1)$$

$$A = \begin{bmatrix} a_{RR} & a_{RG} & a_{RB} \\ a_{GR} & a_{GG} & a_{GB} \\ a_{BR} & a_{BG} & a_{BB} \end{bmatrix} \quad (2)$$

In (1), $[R, G, B]^T$, $[R_0, G_0, B_0]^T$, and $[r, g, b]^T$ are the ambient camera readings, the camera readings, and the projection instructions. A represents the convolutional matrix of the spectral response of camera channels and the spectrum of projector channels as (2) shows. K and P represent surface reflectance matrix and the color calibration matrix of the projector. This work is meaningful to the colored projector-camera model, but the parameters in the model need some effort to obtain[9]. On the other hand, Chen et al. indicated that with two cameras, the complicated problem of lighting-

image correspondence can be substituted with an easier problem of image-image correspondence[10]. And with the artificial texture as assistance, the problem can be simplified even further[11].

With requirements of practical daily tasks, we have to choose methods to make a balance between resolution and speed, accuracy and simplicity, price and ease of use. We chose temporal coding methods to gain better resolution, and naturally the speed would not be as fast as one-shot approaches. Therefore, we added color information to accelerate, and the robustness would suffer to some extent. To deal with that and stay simple, another camera was added, and inverse colored patterns were also projected. Meanwhile, the colored codewords were generated with a Hamming distance of 1 to restrain noise.

III. PROPOSED METHOD

A. The Reconstruction System

The system consisted of a low-cost projector and two cameras, as previously explained, and the projector was placed in between as Fig.2 shows. The two cameras should be connected so that they can take shots at the same time.

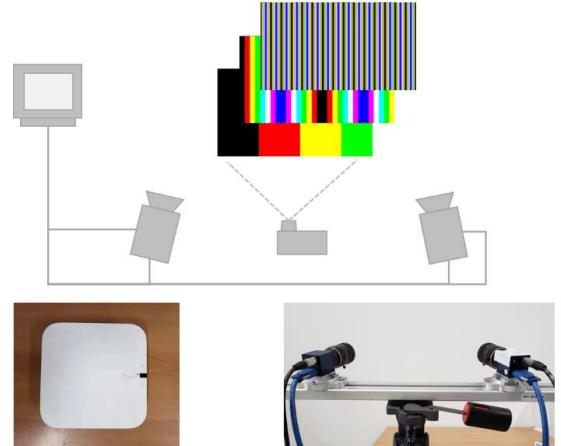


Figure 2. The Reconstruction System

With the target object placed at the right place, patterns were projected and captured simultaneously as fast as the system could. The key problem is how to use special code patterns to increase information into each shots, and to gain better speed and resolution.

B. Generation of the Colored Gray Code

Gray code is robust and simple as a reconstructing method. Yet with only two colors, it needs more projected patterns than any other methods. Gray-Level coding was used to reduce the patterns with M gray levels. Yet the bigger M is, the harder the identification becomes. It would be more robust to merely identify 0s and 1s. It is natural to consider about adding color information, for there would be three 0s and 1s in separated channels.

Considering there are three channels in color patterns, and each channel has two states, we used these eight basic colors: black (0,0,0), red (0,0,1), green (0,1,0), yellow (0,1,1), blue (1,0,0), magenta (1,0,1), cyan (1,1,0), and white (1,1,1). If there are N regions need to be distinguished, only $\log_8 N$ patterns are needed.

Yet it is not the best way to arrange colors in binary order. We should take noise into account and minimize the mismatching cost as far as we could. Therefore, colors in colored Gray code patterns are arranged as in Fig. 3. When we join the RGB values of the camera readings in the same order as the projected patterns, all codewords will be formed with a Hamming distance of 1 to their neighbours.

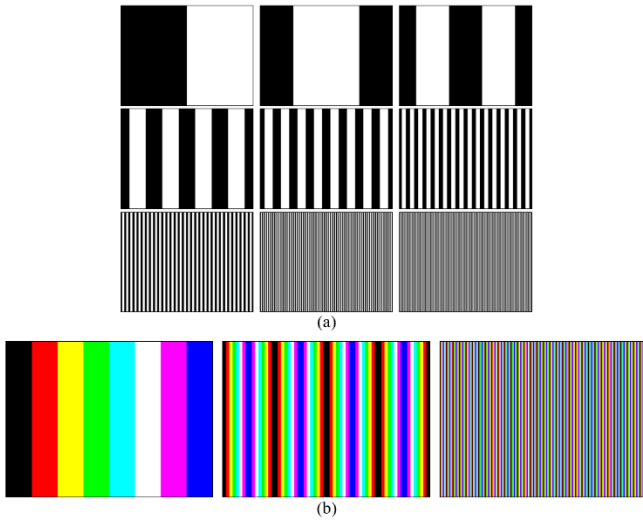


Figure 3. The coded structured light patterns for projectors with 512×384 resolution. (a) Traditional Gray code. The number of the patterns is 9. (b) Colored Gray code. Only 3 patterns are needed. Every column has a Hamming distance of 1 with its neighbours.

We also generated inverse color patterns to test their influences on color judgement. As long as the measured surface is not a monochromatic, flat plane, surface points would have different camera readings even if the illumination is uniform. If we project both the original color patterns and their inverse, where every channel of every point takes the opposite value, we could dynamically judge every point for every projection by comparing the two values, instead of using

a fixed preset threshold. Inverse patterns are widely used[4] for their effects. And we did experiments to test whether it is worthy to double the projecting time for them.

Colored line-shifting patterns and their inverse were also generated, attempting to improve the resolution. Black was chosen to be the background color. Lines appeared once every six pixels, just like the original work[8]. However, lines were generated with the rest seven colors, not just white. Hence every 42 pixels are unambiguous instead of 6.

IV. EXPERIMENTS

Our experiment system was composed of two projectors and two cameras. The projecting time was tested with the professional projector (Texas Instruments DLP LightCrafter 4500, 912×1140 resolution, \$1299), and the low-cost projector (XGIMI Z6, 1920×1080 resolution, \$400). TI 4500 can switch patterns as fast as the user desires. Images can be changed and captured every 0.35s. XGIMI Z6, on the other hand, needs 1.26s to load a new picture. And without any software to control it, we had to switch pictures manually with a remote control. The images were captured every 1.30s. (We did not connect it as a second screen to the computer. Because although it could project images in real-time in this way, its resolution suffered severely. Instead, we loaded patterns by plugging a USB flash disk into it.)

Our cameras were of the same type, in order to have the same spectral responses (DAHENG MER-U3C, 3840×2748 resolution, \$476). Calibrations were done in advance. A $130\text{mm} \times 100\text{mm}$ statue and a $139.88\text{mm} \times 139.88\text{mm}$ cube were used in our experiments.

Three sets of patterns were projected for each object: (a) traditional Gray code, (b) colored Gray code, and (c) colored Gray code with colored line-shifting patterns. The inverse patterns were also projected. Moiré effect could be observed when the thinnest stripes were projected by XGIMI Z6, but not TI 4500, and it was more obvious in Gray code.

A. Projecting Time

With traditional Gray code and the TI 4500 projector, 10 column patterns, 11 row patterns and their inverse were generated. An all-black and an all-white pattern were also projected to calculate the shadowed regions, which could be removed from the reconstruction. With XGIMI Z6, 22 patterns were projected for columns and for rows, and 2 black-and-white patterns were also added.

As we said before, it took 0.35s for TI 4500 and 1.30s for XGIMI Z6 to switch a pattern. The time needed to collect those 44-46 images was approximately 15.40s for TI 4500 and 59.80s for XGIMI Z6, the latter was unbearably long.

Yet with colored Gray code, only 10 patterns were required (18 if the inverse ones were projected). And the projecting time were just 13.0s for XGIMI Z6 (and 10.4 more seconds for inverse ones). That was much more efficient. And if line-shifting methods were used, the time needed would be 44.2s for XGIMI Z6 with 34 patterns.

B. Reconstruction Results Comparison between Gray Code and Colored Gray Code

The time of decoding and correspondence calculating on a CPU (Intel Core i7-8700, 3.20GHz) was recorded. Calculations can be accelerated if a GPU or multiple CPUs are available, because the algorithm is highly supportive to parallel computing. It would be beneficial to make some budget for computing power especially when the camera resolution is really high (10 million resolution in our case).

Inverse patterns were used for both Gray code and colored Gray code methods. A threshold was set in advance to identify whether a point was noisy. For example, if the difference between the values in original patterns and inverse patterns was less than 7, the point would be removed from the reconstruction. With the same threshold, the time and reconstructed points number for each method were recorded in Table 1.

TABLE I. THE PERFORMANCES OF TRADITIONAL GRAY CODE AND COLORED GRAY CODE

| Methods | Number of Patterns | Projecting Time | Reconstructed Points | Decoding Time |
|-----------------------|--------------------|-----------------|----------------------|---------------|
| Traditional Gray Code | 46 | 59.8s | 156,2973 | 181.48s |
| Colored Gray Code | 18 | 23.4s | 214,4848 | 229.41s |

Although the decoding time seems so long that it could wipe our effort to reduce the projecting time, it could be cut down significantly if regular-resolution-cameras or GPUs were used. The average decoding time was about one-fifth of that when the program ran on a GPU for 10 times (NVIDIA GeForce GTX 1060 6GB). Note that the time would be longer at first for the initialization but much less after that, so the average performances were much better.

In fact, camera resolution and computing devices does not matter, because we have to pay some attention to the “Reconstructed Points” column. The decoding time for colored Gray code was longer, just because the number of reconstructed points was bigger. And this is unexpectedly good news, indicating that colored Gray code not only have less projecting time and similar decoding time for each point, but also better resolution in comparison. The point clouds can be seen in Fig. 6. The one generated by colored Gray code appeared less noisy than the others.

We expected to lose some resolution, because traditional Gray code should be more robust with only two colors. Yet the result showed other possibilities. There were two possible reasons. First, we did observe that the projected white stripes separated into three colors when they got really thin. In that case, it would be noisy to identify white from black, yet more robust to identify the three basic colors. Second, if we take a good look at the thinnest stripes of monochrome patterns and color patterns as Fig. 4 shows, we could notice that the monochrome stripes got blurring and gray after getting mixed with each other. If the projector cannot focus precisely, or the surface cannot stay in the focusing range, the thinnest stripes would hardly be recognizable. On the contrary, the color ones still had a special color changing pattern, even if the color stripes were mixed. And the points in each camera could still find their correspondences, resulting in 37% denser point cloud.

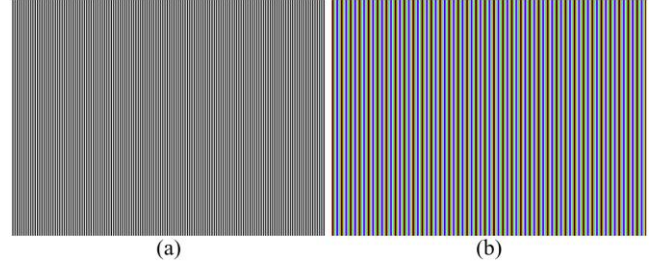


Figure 4. The thinnest stripes. (a) Traditional monochrome Gray code. The thinnest stripes were mixed and hard to recognize. (b) Colored Gray code. The mixed thinnest color stripes could form new colors instead of blurring, so the color matching procedure would not be affected as much as in traditional Gray code.

C. The Effect of Inverse Patterns

The inverse patterns are shown in Fig. 5. When we removed the inverse patterns, and set the judging threshold based on the average channel values of the all-black and all-white images, the reconstructed point cloud was much noisier (Fig. 6 (c)). So it is probably a bad idea to save time by deleting the inverse patterns.

In our experiments, although the surfaces had diffused reflection properties, the camera readings of a point would still be highly affected by the illumination in the neighbouring areas. By providing dynamic judging criteria at every point and every projection, instead of using the same threshold under different projecting patterns, inverse patterns could distinctly improve the reconstruction accuracy.

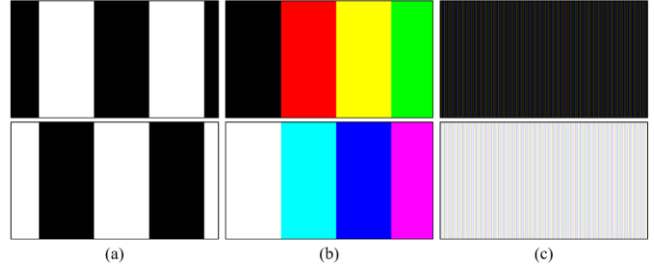


Figure 5. Original patterns (top) and their inverse (bottom). (a) Traditional Gray code patterns. (b) Colored Gray code patterns. (c) Colored Line-shifting patterns. The original patterns consist of black background and lines of 7 colors.

D. The Combination of Colored Line-Shifting Patterns

We attempted to use colored line-shifting patterns to increase resolution. Two colored stripes patterns and six colored line patterns were projected, along with their inverse patterns.

It appeared to be a huge mistake to use white-background inverse patterns (Fig. 5 (c) Bottom). Like we said before, the readings of every point in inverse patterns became much higher than the original ones because of the white stripes, making it impossible to get the correct color.

Therefore, we deleted the inverse patterns, and set thresholds for color judgement. The result was the worst one as Fig. 6 (d) shows. It appeared to be more robust to use continuous patterns as in colored Gray code.

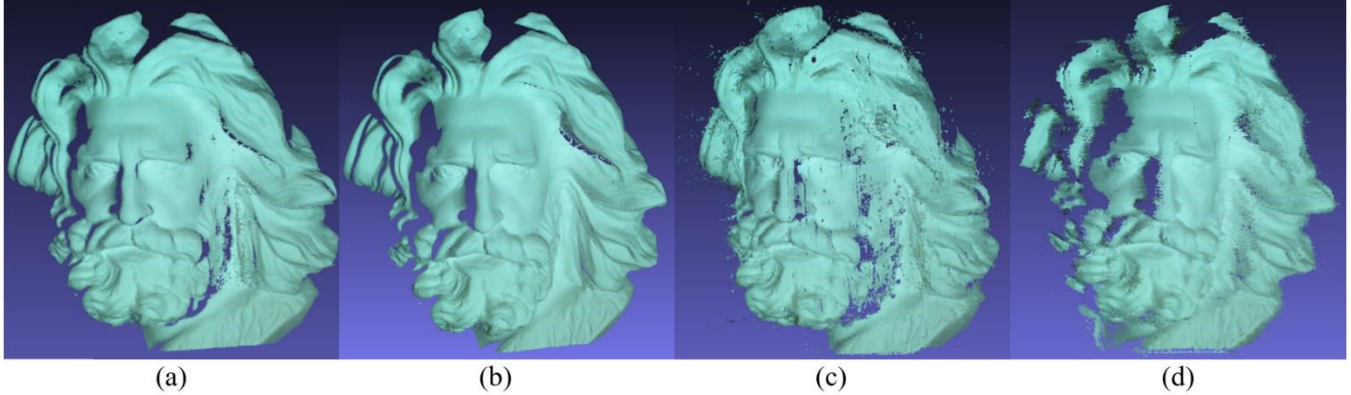


Figure 6. Results from different methods. (a) Traditional Gray code. (b) Colored Gray code with inverse patterns. (c) Colored Gray code without inverse patterns. (d) Colored Gray code with colored line-shifting patterns.

E. Evaluation

A cube was measured by a vernier caliper (with a precision of 0.02mm) and our colored Gray code reconstruction program as evaluation. The width of the cube is 139.88mm on average after 15 measurements at five positions. 260,7882 points were reconstructed in 293.28s. Without any filtering or outlier detections, the width was calculated as the distance of pairs of points who were on the edge at the same height. The measured average width was 139.90mm, with a maximum deviation of 0.032mm because of some noisy points at the edge of the cube.

Generally, the system using our method had satisfying accuracy, yet the resulting point cloud did need some filtering method to suppress outliers.

V. CONCLUSION

Efforts were made to present a 3-D reconstructing method, so that it could have high resolution, satisfying accuracy, better speed and ease of use. The idea of Gray code was appreciated for its simplicity and robustness. Color information was added. Basic colors increased from 2 to 8 to accelerate the projecting process and elevate the density of the point cloud. Another camera along with inverse patterns were used and proved to be necessary for color decoding. Colored line-shifting methods were tested to acquire better resolution, yet it did not help because colored lines cannot be identified robustly when they were surrounded by black stripes.

This work can provide help for people who would like to build an inexpensive and simple 3-D reconstruction system for their daily tasks. Because the characteristics of low-cost projectors were thoroughly considered and dealt with. The system can also handle a sequence of reconstruction tasks, for the projecting time have been cut down. And the decoding time can be shortened as well, when a GPU is used and the decoding time for each surface is apportioned.

Future work can be done to analyze the performances on surfaces with extreme reflectance and texture characteristics. Adjustments can be made to deal with different surface properties. Traditional line-shifting can be combined for better resolution. And we would like to make other attempts to

reduce the patterns. In conclusion, we want the system to be faster and cheaper, functioning at more scenarios.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grants U1701262 and U1801263.

REFERENCES

- [1] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. A state of the art in structured light patterns for surface profilometry. *Pattern recognition*, 43(8): 2666–2680, 2010.
- [2] Chunyu Li, Akihiko Torii, and Masatoshi Okutomi. Robust, precise, and calibration-free shape acquisition with an off-the-shelf camera and projector. In *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–6, 2018.
- [3] Dalit Caspi, Nahum Kiryati, and Joseph Shamir. Range imaging with adaptive color structured light. *IEEE Transactions on Pattern analysis and machine intelligence*, 20(5):470–480, 1998.
- [4] Scharstein D, Szeliski R. High-accuracy stereo depth maps using structured light[C]. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*. 2003.
- [5] Sam Van der Jeught and Joris JJ Dirckx. Real-time structured light profilometry: a review. *Optics and Lasers in Engineering*, 87:18–31, 2016.
- [6] Porras-Aguilar R , Falaggis K , Ramos-Garcia R . Error correcting coding-theory for structured light illumination systems[J]. *Optics and Lasers in Engineering*, 2017, 93:146-155.
- [7] Kayaba H , Kokumai Y . Non-contact Full Field Vibration Measurement Based on Phase-Shifting[C]. *IEEE Conference on Computer Vision & Pattern Recognition*. IEEE Computer Society, 2017.
- [8] Jens Gühring. Dense 3d surface acquisition by structured light using off-the-shelf components. In *Videometrics and Optical Methods for 3D Shape Measurement*, volume 4309, pages 220–232. International Society for Optics and Photonics, 2000.
- [9] Lin H , Song Z . 3D reconstruction of specular surface via a novel structured light approach[C]. *2015 IEEE International Conference on Information and Automation (ICIA)*. IEEE, 2015.
- [10] Chu-Song Chen, Yi-Ping Hung, Chiann-Chu Chiang, and Ja-Ling Wu. Range data acquisition using color structured lighting and stereo vision. *Image and Vision Computing*, 15(6):445–456, 1997.
- [11] Shuang Y , Baoyuan C , Lei Z , et al. Encoded light image active feature matching approach in binocular stereo vision[C]. *2016 11th International Forum on Strategic Technology (IFOST)*. IEEE, 2016.