

Article

# 3D Reconstruction with Single-Shot Structured Light RGB Line Pattern

Yikang Li  and Zhenzhou Wang 

College of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255000, China; 14111103010@stumail.sdut.edu.cn

\* Correspondence: wangzz@sdut.edu.cn

**Abstract:** Single-shot 3D reconstruction technique is very important for measuring moving and deforming objects. After many decades of study, a great number of interesting single-shot techniques have been proposed, yet the problem remains open. In this paper, a new approach is proposed to reconstruct deforming and moving objects with the structured light RGB line pattern. The structured light RGB line pattern is coded using parallel red, green, and blue lines with equal intervals to facilitate line segmentation and line indexing. A slope difference distribution (SDD)-based image segmentation method is proposed to segment the lines robustly in the HSV color space. A method of exclusion is proposed to index the red lines, the green lines, and the blue lines respectively and robustly. The indexed lines in different colors are fused to obtain a phase map for 3D depth calculation. The quantitative accuracies of measuring a calibration grid and a ball achieved by the proposed approach are 0.46 and 0.24 mm, respectively, which are significantly lower than those achieved by the compared state-of-the-art single-shot techniques.

**Keywords:** 3D reconstruction; single shot; slope difference distribution; line indexing; structured light; phase map



**Citation:** Li, Y.; Wang, Z. 3D Reconstruction with Single-Shot Structured Light RGB Line Pattern. *Sensors* **2021**, *21*, 4819. <https://doi.org/10.3390/s21144819>

Academic Editors: Marco Leo and Marco Marcon

Received: 11 June 2021  
Accepted: 11 July 2021  
Published: 14 July 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Non-contact optical surface 3D imaging technology has made tremendous progress in the last decades. Traditional methods are challenged by more flexible alternatives [1]. Nevertheless, robust 3D reconstruction of moving and deforming objects remains challenging. To reconstruct a dynamic scene, the optimum choice is the single-shot method, and state-of-the-art single-shot methods can be divided into four categories. The first category is time of flight (TOF), the second category is passive stereo vision (PSV), the third category is active stereo vision (ASV), and the fourth category is structured light (SL). Each category has its representative product. For instance, Kinect V2 is the representative product of the first category. Despite its wide use and great popularity, its performance is not stable. For instance, it is reported that Kinect V2 is robust in reconstructing rigid scenes while it fails in providing a reasonable reconstruction result on non-rigidly deforming scenes [2]. The representative product of both PSV and ASV is Intel D435. The passive mode of Intel D435 performs poorly while measuring objects that lack textures [3]. The active mode of Intel D435 adopts an infrared radiation (IR) laser projector to generate the laser speckle patterns [4]. It performs better than other methods for face identification and face expression recognition as reported by Luca et al. [5]. The accuracies of ASV and SL are higher than those of TOF and PSV. In addition, the cost of SL is lower than that of ASV. Therefore, SL has been favored for reconstructing dynamic scenes for many decades, and a great number of single-shot SL methods have been proposed based on the different types of the coded patterns, as reviewed by Wang [1].

Fourier transform profilometry (FTP) [6] is one of the most popular single-shot structured light methods. It decodes a single high-frequency sinusoidal pattern in the frequency domain to calculate the phase map and the depth information. The carrier frequency needs

to be designed as high as possible to reduce the aliasing error. A band-pass filter is used to obtain the modulated phase in the frequency domain. The drawbacks of FTP include (1) sensitivity to noise, light interference, and the optical properties of the measured objects; (2) requirement of a non-trivial phase unwrapping process; (3) low accuracy in measuring complex surfaces. To overcome the drawbacks of FTP, the structured light line pattern profilometry (LPP) was proposed to reconstruct the 3D profile by single-shot in [7]. Different from FTP, LPP calculates the depth information solely in the spatial domain. LPP extracts the distorted line pattern with a series of image processing methods and then indexes the distorted lines uniquely. Because the extracted lines instead of the full-resolution plane are used, LPP is more resistant to noise and light interference. A typical example of the high resistance to noise and light interference can be found in [8] where the laser lines are extracted robustly under extremely strong arc light while measuring the geometry of the weld pool. Because of advances in the image segmentation method, the lines can be extracted robustly while measuring the objects with different optical properties, as shown by Wang et al. [8]. In addition, the phase map calculated by LPP is not wrapped and thus does not require the phase unwrapping process. To measure complex surfaces, more subtle line clustering methods are required by LPP. For instance, a specific line clustering method was proposed to reconstruct the deforming human face in [9], and the line clustering method is designed based on the characteristics of the human face. Consequently, it cannot be used for measuring the 3D shapes of other complex deforming objects, although it performed better than other methods [10–14] in 3D face reconstruction.

In this paper, we propose an exclusion-based line clustering method that can be used for measuring any type of deforming shape. To facilitate the clustering process, we designed the lines in red, green, and blue respectively instead of using a single white color [8,9,15]. The lines in different colors are separated from each other at equal intervals. The minimum interval between the lines in the same color is three times as large as the minimum interval between the adjacent lines in the coded pattern. Therefore, the lines in the same color have little intersection after they are distorted by depth because of their large intervals. After the projected RGB line pattern is captured by a color charge coupled device (CCD) camera, the lines in red, green, and blue are segmented in the hue saturation value (HSV) color space. The segmented red lines, green lines, and blue lines are clustered independently with the proposed line clustering method. To verify the effectiveness of the proposed approach, we conducted both quantitative and qualitative experiments that included reconstructing the static objects for 3D reconstruction accuracy evaluation and reconstructing the moving and deforming object for the line clustering (indexing) robustness evaluation.

In summary, the major contributions of this paper include:

- (1) A fully automatic color stripe segmentation method based on SDD is proposed to extract the RGB line pattern robustly. State-of-the-art methods rely on manually specified thresholds obtained by trial-and-error analysis or the classic classification methods to segment the color stripes. The trial-and-error analysis method is time-consuming and cannot guarantee that the optimal thresholds are always selected for each image. In addition, the demonstration in the paper shows that none of the compared classic classification methods can separate all the adjacent stripes from each other robustly. Due to the bottleneck problem of image segmentation, there are no 3D surface imaging products that are based on the segmented structured light patterns yet, even though the related theory and idea had been proposed for more than three decades. The proposed SDD segmentation method has the potential to solve this bottleneck problem for the single-shot dot-pattern- or line-pattern-based structured light methods.
- (2) An exclusion-based line clustering method is proposed to index the segmented stripes robustly. State-of-the-art methods have included different ways to index the segmented lines or dots uniquely. However, they may fail when complex and deforming objects are reconstructed. In [7], a line clustering method was proposed to index the lines from top to bottom, but it is only robust for simple and rigid objects. In [8,9],

the line-slope-based clustering method was proposed to index the lines distorted by complex and deforming objects. However, its accuracy is easily affected by occlusion and large discontinuity. The line clustering method proposed in this paper is more robust than the previous methods [7–9] in indexing the lines distorted by complex and deforming objects, which is validated by the 3D reconstruction results of a deforming face and a deforming hand in the attached videos.

- (3) An extensive comparison of the proposed method with state-of-the-art single-shot 3D surface imaging methods (products) was conducted by reconstructing the deforming face and the deforming hand continuously, which is a necessary condition to verify the effectiveness of a single-shot structured light method.

## 2. Related Work

The structured light 3D imaging technique reconstructs the sampled points on the surface of the measured object based on different working principles. Based on the sampling resolution, state-of-the-art structured light methods can be divided into three categories: (1) dot-pattern-based method [15–24]; (2) line (stripe)-pattern-based method [7–9,25–30], and (3) plane-pattern-based method [6,31,32]. The dot-pattern-based methods sample the surface of the object sparsely with the specifically coded dot patterns. Image segmentation [16–20] or feature detection [21–24] are used to extract the dot pattern from the acquired pattern image. Each extracted dot is then indexed uniquely for 3D depth calculation with different principles. The line-pattern-based methods sample the surface of the object significantly more densely than the dot-pattern-based methods. Similarly, image segmentation [26,29,30] or feature detection [25,27,28] can be used to extract the line pattern from the acquired pattern image. The plane-pattern-based methods sample the surface of the object with the resolution of the projector. The correspondence between the camera and the projector is pixel to pixel. To distinguish the pixels from each other, the pixels in the coded plane pattern vary according to the function of a sine wave. Unlike the dot pattern and the line pattern, the plane pattern is not identifiable spatially within a single image. As a result, image segmentation or feature detection cannot extract the phase map from a single image solely based on the spatial information. Currently, there are two popular ways for phase map calculation. One is FTP [6] and the other is phase-shifting profilometry (PSP) [30,31]. FTP calculates the phase map in the frequency domain by selecting the fundamental component of the Fourier transformed pattern image with a band-pass filter. To reduce the aliasing error, the carrier frequency of FTP is required to be as high as possible. In reality, alias is unavoidable for FTP to measure the non-planar objects, and consequently the phase map generated by FTP is usually not accurate enough for depth calculation. PSP calculates the phase map in the temporal domain by solving the united equation of the acquired pattern images with different phase-shifting values. In theory, PSP yields an exact analytical solution of the phase map. The major drawback of PSP is that it requires the object to be static when projecting multiple plane patterns onto it. Consequently, PSP is not suitable for measuring deforming and moving objects by nature. To reconstruct deforming and moving objects, the dot pattern or line pattern is usually preferred and used for the structured light method development.

One major challenge of dot-pattern- or line-pattern-based structured light methods is to extract the dots or the lines automatically and robustly. Since the color dots or the color lines are designed distinctively, threshold selection or color classification should be able to segment them robustly. Yet few works have described in detail how the dots or lines were segmented. In [17], the valleys and peaks of the histogram were analyzed and a threshold was selected manually based on the trial-and-error analysis. Obviously, the manually selected threshold cannot always guarantee that the dots or lines are segmented optimally, especially while reconstructing complex deforming objects. In [19], the authors demonstrated that the reconstruction was incomplete because of the challenge of color dot segmentation. In [26], the authors mentioned that the stripes could be segmented by color classification or edge detection, and they basically advocated color classification.

Unfortunately, they did not disclose what color classification method was used by them. In [27,28], the authors utilized the local maxima to segment the stripes. In reality, the captured images are usually affected by color crosstalk, specular reflection, and light interference. As a result, the local maxima-based method could not segment the stripes as accurately as the threshold selection methods or the classification methods. In [29], the authors specified the thresholds for the red, green, and blue stripes as 125, 90, and 95, respectively. Together with the comparison of the gray levels in different channels, the stripes could be segmented effectively. However, the segmentation accuracy was affected by the leakage in the color separation filters and poor color alignment. To our best knowledge, it is difficult to segment the color lines solely based on single-threshold selection or color classification. In this paper, we will address this challenge by proposing a robust RGB line pattern segmentation method based on SDD.

The second challenge of dot-pattern- or line-pattern-based structured light methods is to index the extracted dots or lines uniquely and robustly. Although the stripe indexing issue has been tackled for many decades, it has not yet been solved for the following two reasons: (1) there are no line (stripe) indexing-based single-shot 3D imaging products yet, (2) the proposed line (stripe) indexing methods have not yet been evaluated with deforming and moving objects extensively. Consequently, the robustness of the proposed methods cannot be verified. For instance, only one reconstruction result of a face was demonstrated in [27,28]. Eight reconstruction results of a deforming face were demonstrated in [26], and 18 reconstruction results of a deforming hand were demonstrated in [30]. There are even some single-shot structured light methods that have not been tested with deforming and moving objects at all. In this paper, we will validate the proposed method extensively with the deforming hand and the deforming face. In addition, we extensively compare the 3D reconstruction results of the moving face and hand by the proposed method with those reconstructed by state-of-the-art products in the attached videos.

Although the line (stripe) indexing problem remains unresolved, many interesting and promising line indexing methods have been proposed in the past decades. For instance, the flood fill algorithm was proposed by Robinson et al. to index the stripes in [27]. A maximum spanning tree-based algorithm was proposed by Brink et al. to index the stripes in [28]. The maximum spanning tree of a graph was used to define potential connectivity and adjacency in stripes. Despite a significant improvement in accuracy over existing methods, e.g., the flood fill-based method [27], the maximum spanning tree algorithm could not address the discontinuity in surface depth effectively. A survival of the fittest criterion was proposed by Boyer and Kak to index the stripes uniquely in [29]. The stripe indexing was treated as a labeling problem and formulated as an energy minimization problem by Koninckx and Van Gool in [30]. To solve the minimization problem, a weighted least squares solution was proposed to cast the problem into the search for a minimal cut on a graph. Though the theory sounds good, the reconstructed deforming hands had serious flaws, e.g., the black holes in the reconstructed palm. In this paper, we will address this challenge by proposing a robust line clustering method to index the stripes based on exclusion. With the assumption that the reconstructed surface is continuous and smooth, a line model is computed for each clustered line to make the reconstructed surface continuous and smooth. As a result, the occlusion problem, the discontinuity problem, and other similar problems, e.g., the black hole problem as shown in [30], can be solved effectively.

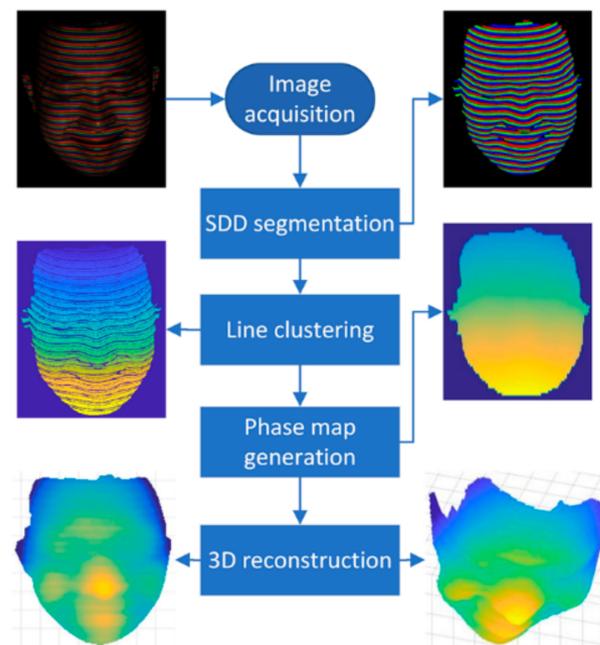
The major differences between the proposed method and the existing methods [26–30] are summarized in Table 1. Compared to the proposed method, the major weaknesses of the existing methods include: (1) the existing methods were not able to segment the lines (stripes) as robustly as the proposed method was able to; (2) the existing methods were not able to index the lines (stripes) as robustly as the proposed method was able to. Consequently, their results in reconstructing deforming and moving objects are inferior to those of the proposed method, which can be verified by comparing the reconstruction results of the deforming hand in this paper and those shown in [30].

**Table 1.** Comparison of the major differences between the proposed method and the existing methods [26–30].

Methods	Line Segmentation	Line Indexing
[26]	Color classification	Permutation
[27]	Local maxima	Flood fill
[28]	Local maxima	Spanning tree
[29]	Manual	Unique word
[30]	Color classification	Graph cut
Proposed	SDD thresholds	Exclusion

### 3. The Proposed Approach

The flowchart of the proposed approach is shown in Figure 1. Firstly, the acquired image is segmented in the HSV color space by the SDD image segmentation method [33,34]. The segmented lines in different colors are then indexed by the proposed exclusion-based line clustering methods. The indexed lines in different colors are fused together to generate the final indexing lines. The phase map is computed from the final indexed lines and the 3D shape is reconstructed with the phase map.



**Figure 1.** The flowchart of the proposed approach.

#### 3.1. The Proposed Pattern Extraction Method

The coded RGB pattern is illustrated in Figure 2. The red lines, green lines, and blue lines are in the same intervals, parallel and repeated in cycles to form the RGB pattern. Between the adjacent lines, there are dark parts to separate them from each other. The intensities in different channels of the RGB color space contain both the color information and the lighting information, which makes the color segmentation problematic. On the other hand, color and brightness are separated in different channels as hue and value, which makes the color segmentation straightforward. Therefore, the RGB color space is transformed into the HSV color space for line segmentation. The transformation from the RGB color space to the HSV color space is formulated as follows.

$$V(i, j) = \max(R(i, j), G(i, j), B(i, j)), \quad (1)$$

$$M(i, j) = \min(R(i, j), G(i, j), B(i, j)), \quad (2)$$

$$S(i, j) = \frac{V(i, j) - M(i, j)}{V(i, j)}, \quad (3)$$

$$H(i, j) = \begin{cases} 60 \times \frac{G(i, j) - B(i, j)}{V(i, j) - M(i, j)}, & R(i, j) = V(i, j) \\ 120 + \frac{B(i, j) - R(i, j)}{V(i, j) - M(i, j)}, & G(i, j) = V(i, j) \\ 240 + \frac{R(i, j) - G(i, j)}{V(i, j) - M(i, j)}, & B(i, j) = V(i, j) \end{cases} \quad (4)$$

where  $(i, j)$  denotes the index of the pixel in the acquired pattern image.  $H$  represents the  $H$  channel of the pattern image,  $S$  represents the  $S$  channel of the pattern image, and  $V$  represents the  $V$  channel of the pattern image.  $R$  represents the  $R$  channel of the pattern image,  $G$  represents the  $G$  channel of the pattern image, and  $B$  represents the  $B$  channel of the pattern image. From the above HSV color model, it is seen that the coded RGB pattern has the following characteristics: (1) the RGB color is distinguished in the  $H$  channel, (2) the difference between the RGB line part and the adjacent dark part is highlighted in the  $S$  channel, (3) the brightness of the RGB line part is highlighted in the  $V$  channel. The illustrated HSV channels in Figure 3 also correlate with these three characteristics highly. Figure 3a shows the RGB color image acquired by the CCD camera. Figure 3b–d shows the  $H$  channel, the  $S$  channel, and the  $V$  channel of the acquired image, respectively.

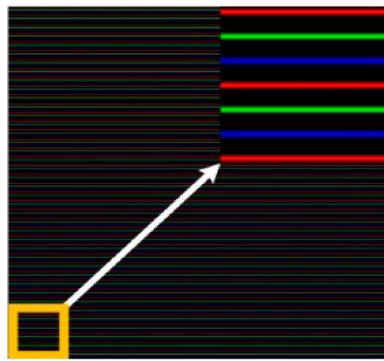


Figure 2. Illustration of the coded pattern.

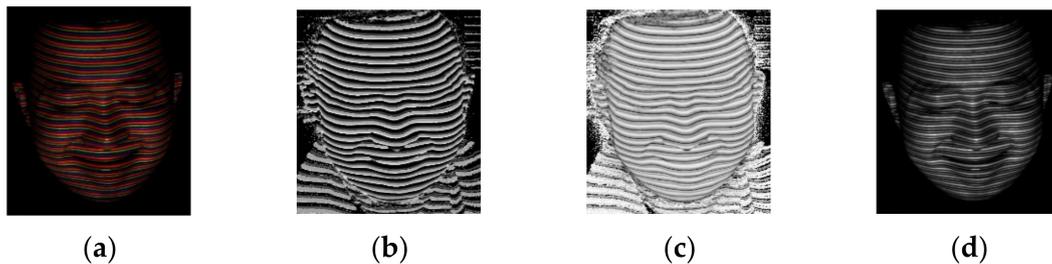
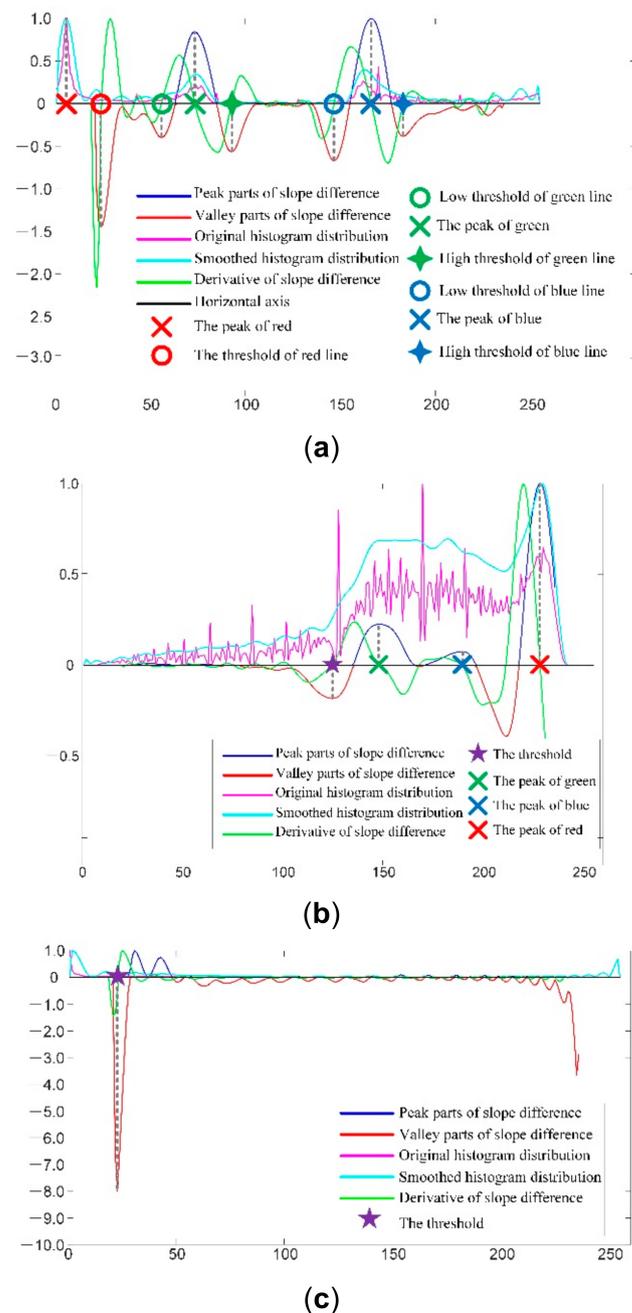


Figure 3. Illustration of the acquired pattern image in  $H$ ,  $S$ , and  $V$  channels. (a) The acquired RGB pattern image; (b) The  $H$  channel of the acquired pattern image; (c) The  $S$  channel of the acquired pattern image; (d) The  $V$  channel of the acquired pattern image.

The line segmentation method is proposed based on these three characteristics: (1) the lines in red, green, and blue are distinguished in the  $H$  channel. Thus, there should be three clustering centers in the histogram of the  $H$  channel that correspond to the red pixels, the green pixels, and the blue pixels, respectively. The clustering centers are calculated by SDD as described in [33,34]. For each clustering center, there are two corresponding SDD valleys. One is on its left and the other is on its right. They are selected as the thresholds to segment the lines in the corresponding color. The process of selecting the double thresholds for the red lines, blue lines, and green lines in the  $H$  channel by SDD is illustrated in Figure 4a; (2) the RGB line parts and the dark parts are distinguished in the  $S$  channel. Thus, there should be one global threshold that can distinguish the RGB line part and the dark part. The global

threshold is calculated with SDD as described in [33,34]. The process of the selecting the global threshold to separate the line part from the dark part in the S channel by SDD is illustrated in Figure 4b; (3) because the brightness of the RGB line part is highlighted in the V channel, the region of interest (ROI) of the RGB lines can be obtained from the V channel by a global threshold. The global threshold is calculated with SDD as described in [33,34]. The process of the selecting the global threshold to obtain the ROI of the RGB lines in the S channel with SDD is illustrated in Figure 4c.



**Figure 4.** Illustration of the SDD threshold selection process for the acquired pattern image in H, S, and V channels. (a) The SDD threshold selection process in the H channel; (b) The SDD threshold selection process in the S channel; (c) The SDD threshold selection process in the V channel.

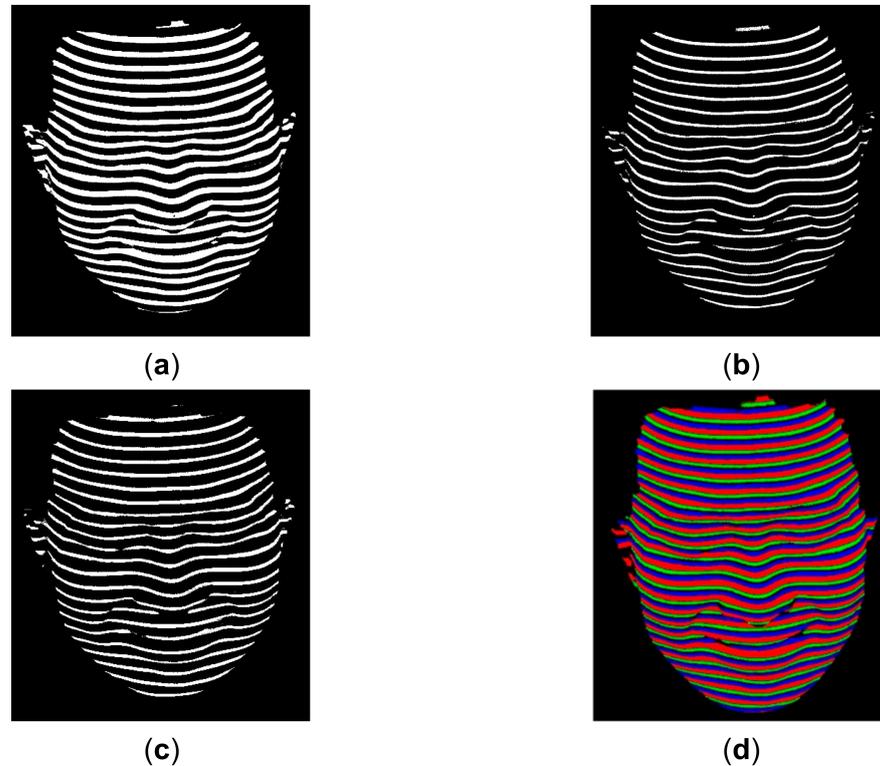
With all the thresholds computed in the HSV channels, the RGB lines are segmented with the following equations.

$$L_R(i, j) = \begin{cases} 1, T_{R1} < H(i, j) < T_{R2}, S(i, j) > T_S, V(i, j) > T_V \\ 0, else \end{cases}, \quad (5)$$

$$L_G(i, j) = \begin{cases} 1, T_{G1} < H(i, j) < T_{G2}, S(i, j) > T_S, V(i, j) > T_V \\ 0, else \end{cases}, \quad (6)$$

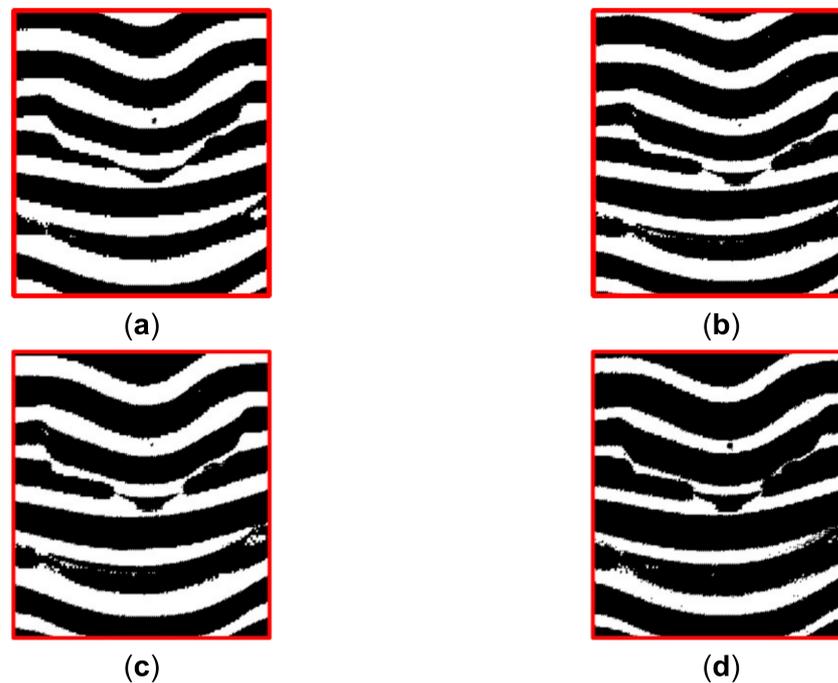
$$L_B(i, j) = \begin{cases} 1, T_{B1} < H(i, j) < T_{B2}, S(i, j) > T_S, V(i, j) > T_V \\ 0, else \end{cases} \quad (7)$$

where  $L_R$  represents the segmentation result of the red lines,  $L_G$  represents the segmentation result of the green lines, and  $L_B$  represents the segmentation result of the blue lines.  $T_{R1}$  denotes the low threshold and  $T_{R2}$  denotes the high threshold for red line segmentation.  $T_{R1} = 0$  and  $T_{R2}$  is denoted as the red circle in Figure 4a.  $T_{G1}$  denotes the low threshold and  $T_{G2}$  denotes the high threshold for green line segmentation.  $T_{G1}$  is denoted as the green circle and  $T_{G2}$  is denoted as the green tetragonal star in Figure 4a.  $T_{B1}$  denotes the low threshold and  $T_{B2}$  denotes the high threshold for blue line segmentation.  $T_{B1}$  is denoted as the blue circle and  $T_{B2}$  is denoted as the blue tetragonal star in Figure 4a.  $T_S$  denotes the global threshold to separate the RGB line part from the dark part in the S channel, and it is denoted as the purple pentagram in Figure 4b.  $T_V$  denotes the global threshold to separate the RGB line ROI from the dark background in the V channel, and it is denoted as the purple pentagram in Figure 4c. The segmented RGB lines with Equations (5)–(7) are illustrated in Figure 5. By comparing the segmented lines in Figure 5 with the acquired lines in Figure 3, it is seen that the SDD segmentation method is robust.



**Figure 5.** Illustration of the segmented lines in different colors. (a) The segmented lines in red; (b) The segmented lines in green; (c) The segmented lines in blue; (d) The combination of the RGB lines.

As reviewed in the second section, two existing methods [26,30] use the classification methods to segment the lines. To demonstrate the superiority of the proposed line segmentation method, the classic classification methods were tested to segment the H channel of the same illustrated image, and the line classification results were refined by the SDD segmentation results of S and V channels in the same way. The compared classic classification methods included K-means [35], Otsu's method [36], and expectation maximization (EM) [37]. Zoomed-in parts of the classification results around the mouth and nose by different classic methods are shown in Figure 6. As can be seen, only the segmented adjacent lines by the SDD method are separate from each other. There are adjacent lines adhesive to each other in the results of the compared classic classification methods, which will make the subsequent line clustering fail.



**Figure 6.** Qualitative comparison of the line segmentation results by state-of-the-art classification methods and the result by the proposed SDD method. (a) Result of the SDD method; (b) Result of K-means; (c) Result of Otsu; (d) Result of EM.

### 3.2. The Proposed line Clustering Method

The proposed line clustering method contains the following steps.

**Step 1:** The starting line is selected based on its length and centroid in the top to bottom direction. Then, it is fitted as:

$$y = a_1x^7 + a_2x^6 + a_3x^5 + \dots + a_7x + a_8 \quad (8)$$

where  $x$  denotes the coordinate of the pixel on the starting line in the horizontal direction and  $y$  denotes the coordinate of the pixel on the starting line in the vertical direction. The coefficient of the fitted line is computed by the least squares error method.

$$[a_1, a_2, \dots, a_7, a_8]^T = (\psi^T \psi)^{-1} \psi^T Y, \quad (9)$$

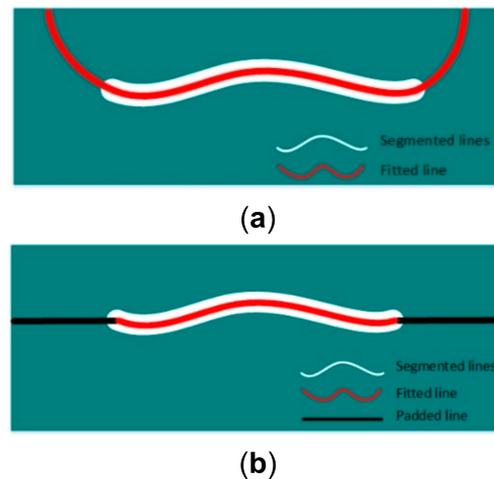
$$\psi = \begin{bmatrix} x_1^7 & x_1^6 & \dots & x_1 & 1 \\ x_2^7 & x_2^6 & \dots & x_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^7 & x_m^6 & \dots & x_m & 1 \end{bmatrix}, \quad (10)$$

$$Y = [y_1, y_2, \dots, y_m]^T \quad (11)$$

where  $m$  denotes the total number of pixels on the starting line. The leftmost point of the starting line and the rightmost point of the starting line are computed as  $(x_l, y_l)$  and  $(x_r, y_r)$ , respectively. The padded line is computed as follows.

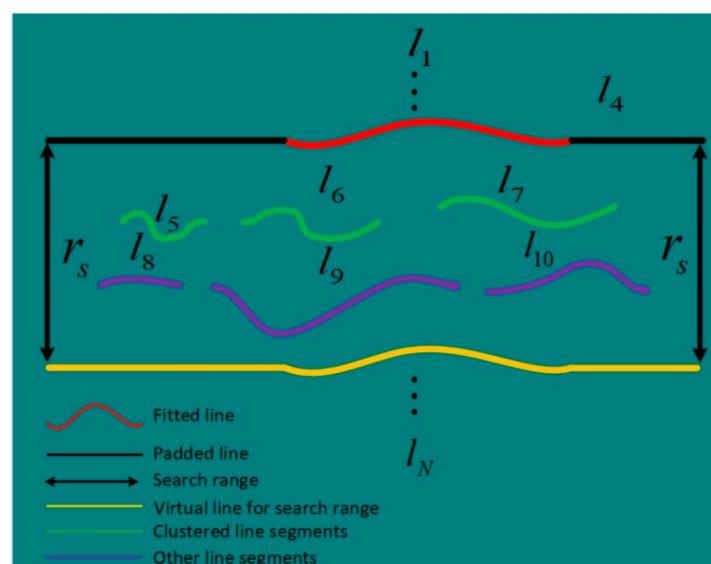
$$y = \begin{cases} y_l, & x < x_l \\ y_r, & x > x_r \end{cases} \quad (12)$$

The fitted line and the padded line are illustrated in Figure 7a,b, respectively.



**Figure 7.** Illustration of the fitted line and the padded line. (a) The fitted line in red superimposed on the segmented line; (b) The non-overlapping part of the fitted line is cut off and replaced with the padded line.

**Step 2:** A search range  $r_s$  is defined based on the fitted line and the padded line.  $r_s$  is computed as twice the average distance between adjacent lines. As illustrated in Figure 8, the search range is denoted by the virtual yellow line. All the line segments within the search range will be selected for clustering the next line. In this illustration, the starting line is determined as the fourth segmented line. Consequently, the clustering process will include downward clustering and upward clustering.



**Figure 8.** Illustration of the search range.

**Step 3:** The line segments are clustered in the search range. Firstly, all the line segments in the search range are denoted as 1, i.e., all of them belong to the next clustered line by default. The exclusion method works as follows.

The centroid, the coordinate of the leftmost point, and the coordinate of the rightmost point are computed as  $(C_x^i, C_y^i)$ ,  $(x_l^i, y_l^i)$  and  $(x_r^i, y_r^i)$ , respectively, for the  $i$ -th line segment  $l_i$  in the search range. The  $i$ -th line segment  $l_i$  is combined with any one of the remaining line segments,  $l_j$  and  $j \neq i$ . If one of the following conditions is met, the  $i$ -th line segment  $l_i$  will be determined as not belonging to the next clustered line.

- (1) If the leftmost point of the  $i$ -th line segment  $l_i$  falls between the length range  $[X_l^j, X_r^j]$  of the  $j$ -th line segment  $l_j$ , i.e.,  $X_l^i > X_l^j$  and  $X_l^i < X_r^j$ , the  $j$ -th line segment  $l_j$  is fitted with the line model by Equation (8) as  $f_j$ . If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is below the position of the fitted line model  $f_j$  for the downward clustering, i.e.,  $Y_l^i - f_j(X_l^i) > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is above the position of the fitted line model  $f_j$  for the upward clustering, i.e.,  $Y_l^i - f_j(X_l^i) < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.
- (2) If the rightmost point of the  $i$ -th line segment  $l_i$  falls between the length range  $[X_l^j, X_r^j]$  of the  $j$ -th line segment  $l_j$ , i.e.,  $X_r^i > X_l^j$  and  $X_r^i < X_r^j$ , the  $j$ -th line segment  $l_j$  is fitted with the line model by Equation (8) as  $f_j$ . If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is below the position of the fitted line  $f_j$  for the downward clustering, i.e.,  $Y_r^i - f_j(X_r^i) > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is above the position of the fitted line  $f_j$  for the upward clustering, i.e.,  $Y_r^i - f_j(X_r^i) < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.
- (3) If the centroid of the  $i$ -th line segment  $l_i$  falls between the length range  $[X_l^j, X_r^j]$  of the  $j$ -th line segment  $l_j$ , i.e.,  $C_x^i > X_l^j$  and  $C_x^i < X_r^j$ , the  $j$ -th line segment  $l_j$  is fitted with the line model by Equation (8) as  $f_j$ . If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is below the position of the fitted line  $f_j$  for the downward clustering, i.e.,  $C_y^i - f_j(C_x^i) > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the  $i$ -th line segment  $l_i$  is above the position of the fitted line  $f_j$  for upward clustering, i.e.,  $C_y^i - f_j(C_x^i) < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.
- (4) If the leftmost point of the  $j$ -th line segment  $l_j$  falls between the length range  $[X_l^i, X_r^i]$  of the  $i$ -th line segment  $l_i$ , i.e.,  $X_l^j > X_l^i$  and  $X_l^j < X_r^i$ , the  $i$ -th line segment  $l_i$  is fitted with the line model by Equation (8) as  $f_i$ . If the position of the leftmost point of the  $j$ -th line segment  $l_j$  is below the position of the fitted line  $f_i$  for the downward clustering, i.e.,  $f_i(X_l^j) - Y_l^j > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the  $j$ -th line segment  $l_j$  is above the position of the fitted line  $f_i$  for the upward clustering, i.e.,  $f_i(X_l^j) - Y_l^j < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.
- (5) If the rightmost point of the  $j$ -th line segment  $l_j$  falls between the length range  $[X_l^i, X_r^i]$  of the  $i$ -th line segment  $l_i$ , i.e.,  $X_r^j > X_l^i$  and  $X_r^j < X_r^i$ , the  $i$ -th line segment  $l_i$  is fitted with the line model by Equation (8) as  $f_i$ . If the position of the leftmost point of the  $j$ -th line segment  $l_j$  is below the position of the fitted line  $f_i$  for the downward clustering, i.e.,  $f_i(X_r^j) - Y_r^j > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the

$j$ -th line segment  $l_j$  is above the position of the fitted line  $f_i$  for the upward clustering, i.e.,  $f_i(X_r^j) - Y_r^j < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.

- (6) If the centroid of the  $j$ -th line segment  $l_j$  falls between the length range  $[X_l^j, X_r^j]$  of the  $i$ -th line segment  $l_i$ , i.e.,  $C_x^j > X_l^i$  and  $C_x^j < X_r^i$ , the  $i$ -th line segment  $l_i$  is fitted with the line model by Equation (8) as  $f_i$ . If the position of the leftmost point of the  $j$ -th line segment  $l_j$  is below the position of the fitted line  $f_i$  for the downward clustering, i.e.,  $f_i(C_x^j) - C_y^j > 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line. If the position of the leftmost point of the  $j$ -th line segment  $l_j$  is above the position of the fitted line  $f_i$  for the upward clustering, i.e.,  $f_i(C_x^j) - C_y^j < 0$ , the  $i$ -th line segment  $l_i$  is determined as not belonging to the next clustered line.

All the line segments that meet one of the above six conditions will be determined as not belonging to the next clustered line. For succinct expression and better understanding, the six conditions are formulated by the following equation.

$$l_i = \begin{cases} 0 & (X_l^j > X_l^i \ \& \ X_l^j < X_r^i \ \& \ Y_l^j - f_j(X_l^j) > 0) \\ 0 & (X_r^j > X_l^i \ \& \ X_r^j < X_r^i \ \& \ Y_r^j - f_j(X_r^j) > 0) \\ 0 & (C_x^j > X_l^i \ \& \ C_x^j < X_r^i \ \& \ C_y^j - f_j(C_x^j) > 0) \\ 0 & (X_l^j > X_l^i \ \& \ X_l^j < X_r^i \ \& \ f_i(X_l^j) - Y_l^j > 0) \\ 0 & (X_r^j > X_l^i \ \& \ X_r^j < X_r^i \ \& \ f_i(X_r^j) - Y_r^j > 0) \\ 0 & (C_x^j > X_l^i \ \& \ C_x^j < X_r^i \ \& \ f_i(C_x^j) - C_y^j > 0) \end{cases} \quad (13)$$

As illustrated in Figure 8, the combination of the line segments  $l_5$ ,  $l_6$  and  $l_7$  is clustered as the next line. Their combination is fitted by the line model (Equation (8)) and padded by Equation (12). Then, the newly fitted and padded line is used for the subsequent line clustering. As can be seen, the working principle of the proposed exclusion method is based on the selection of a large search range that includes line segments belonging to more than one line. If the search range is not large enough, it is easy to miss some line segments as illustrated in Figure 9.

**Step 4:** Repeat Steps 2–3 until all the lines are clustered. After all the lines are clustered, the fitted lines by Equation (8) are used to generate the phase map with the method proposed in [7].

### 3.3. Three-Dimensional Reconstruction Based on the Phase Map

Although the method of generating the phase map by a single line pattern was proposed in [7], its 3D reconstruction principle has not yet been explained systematically. This missing work is included in this section. As illustrated in Figure 10, the structured light 3D imaging system is composed of a projector and a camera. The camera coordinate and the world coordinate are related mathematically by the rule of the pinhole camera model [38] as follows.

$$Z_c \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = M_{int} M_{ext} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (14)$$

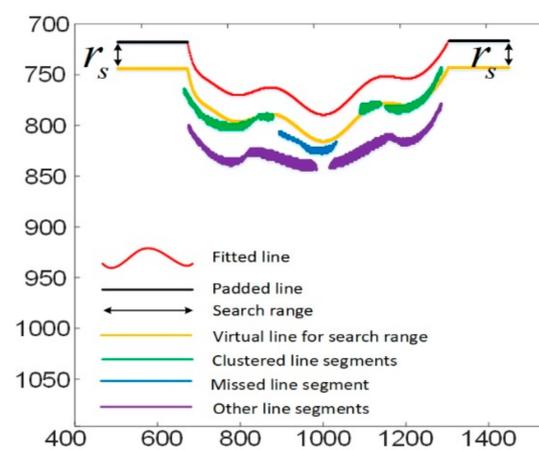
where  $(x_c, y_c)$  is the camera coordinate and  $(X_w, Y_w, Z_w)$  is the world coordinate.  $M_{int}$  is the intrinsic matrix and  $M_{ext}$  is the extrinsic matrix. They are formulated as [38]:

$$M_{int} = \begin{bmatrix} f_x & 0 & u_c \\ 0 & f_y & v_c \\ 0 & 0 & 1 \end{bmatrix}, \quad (15)$$

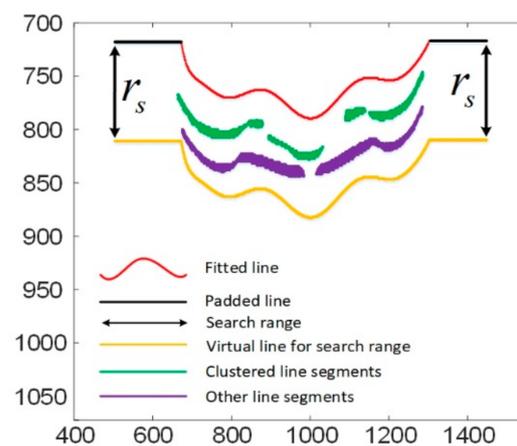
$$M_{ext} = \begin{bmatrix} r_{11}, & r_{12}, & r_{13}, & T_1 \\ r_{21}, & r_{22}, & r_{23}, & T_2 \\ r_{31}, & r_{32}, & r_{33}, & T_3 \end{bmatrix} \quad (16)$$

where  $(f_x, f_y)$  is the camera's focal length and  $(u_c, v_c)$  is the camera's principle point.  $[r_{11}, r_{12}, r_{13}; r_{21}, r_{22}, r_{23}; r_{31}, r_{32}, r_{33}]$  is the rotation matrix and  $[T_1, T_2, T_3]^T$  is the translation vector. For simplicity, Equation (14) can be rewritten as the following format [38].

$$Z_c \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^{wc}, & m_{12}^{wc}, & m_{13}^{wc}, & m_{14}^{wc} \\ m_{21}^{wc}, & m_{22}^{wc}, & m_{23}^{wc}, & m_{24}^{wc} \\ m_{31}^{wc}, & m_{32}^{wc}, & m_{33}^{wc}, & m_{34}^{wc} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (17)$$



(a)



(b)

**Figure 9.** Comparison of the clustering result with different search ranges. (a) The clustering result with a small search range; (b) The clustering result with a proper search range.

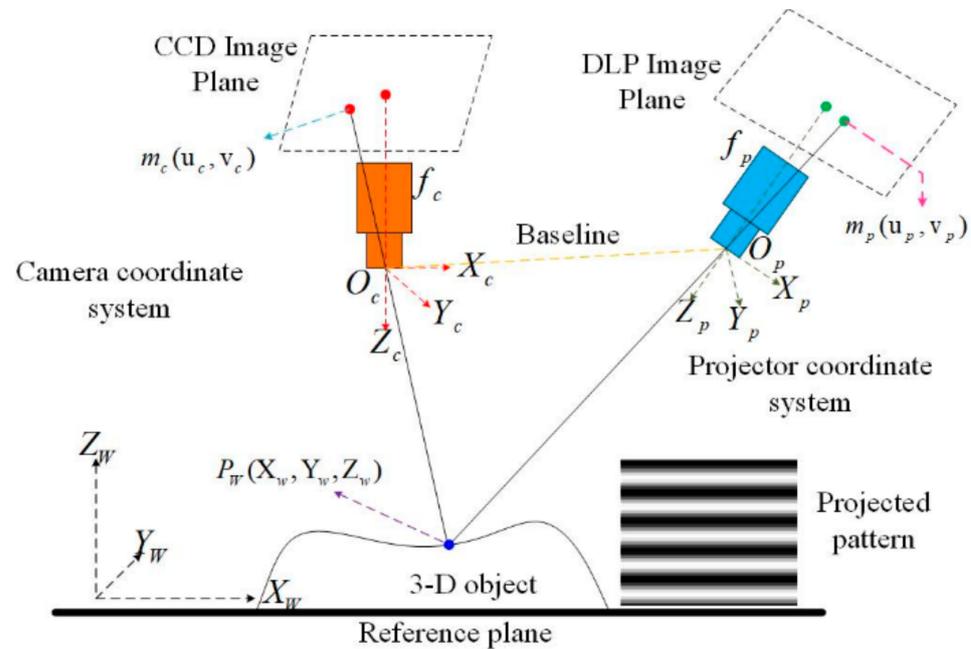


Figure 10. The illustration of the SL 3D imaging system.

To decrease the number of unknown parameters, we let  $m_{34}^{wc} = 1$ . Equation (17) is then expanded into the following three equations.

$$Z_c x_c = m_{11}^{wc} X_w + m_{12}^{wc} Y_w + m_{13}^{wc} Z_w + m_{14}^{wc}, \quad (18)$$

$$Z_c y_c = m_{21}^{wc} X_w + m_{22}^{wc} Y_w + m_{23}^{wc} Z_w + m_{24}^{wc}, \quad (19)$$

$$Z_c = m_{31}^{wc} X_w + m_{32}^{wc} Y_w + m_{33}^{wc} Z_w + 1. \quad (20)$$

Substituting Equation (20) into Equation (18), the following equation is obtained.

$$m_{11}^{wc} X_w + m_{12}^{wc} Y_w + m_{13}^{wc} Z_w + m_{14}^{wc} = x_c m_{31}^{wc} X_w + x_c m_{32}^{wc} Y_w + x_c m_{33}^{wc} Z_w + x_c. \quad (21)$$

Substituting Equation (20) into Equation (19), the following equation is obtained.

$$m_{21}^{wc} X_w + m_{22}^{wc} Y_w + m_{23}^{wc} Z_w + m_{24}^{wc} = y_c m_{31}^{wc} X_w + y_c m_{32}^{wc} Y_w + y_c m_{33}^{wc} Z_w + y_c. \quad (22)$$

Equation (21) can be rewritten into the following format [31]:

$$[X_w, Y_w, Z_w, 1, 0, 0, 0, 0, -x_c X_w, -x_c Y_w, -x_c Z_w] \theta_c = x_c \quad (23)$$

where

$$\theta_c = [m_{11}^{wc}, m_{12}^{wc}, m_{13}^{wc}, m_{14}^{wc}, m_{21}^{wc}, m_{22}^{wc}, m_{23}^{wc}, m_{24}^{wc}, m_{31}^{wc}, m_{32}^{wc}, m_{33}^{wc}]^T.$$

In the same way, Equation (22) can be rewritten into the following format [31]:

$$[0, 0, 0, 0, X_w, Y_w, Z_w, 1, -y_c X_w, -y_c Y_w, -y_c Z_w] \theta_c = y_c. \quad (24)$$

Equations (23) and (24) are combined into the following equation.

$$X_c \theta_c = Y_c \quad (25)$$

where  $X_c = \begin{bmatrix} X_w, Y_w, Z_w, 1, 0, 0, 0, 0, -x_c X_w, -x_c Y_w, -x_c Z_w \\ 0, 0, 0, 0, X_w, Y_w, Z_w, 1, -y_c X_w, -y_c Y_w, -y_c Z_w \end{bmatrix}$  and  $Y_c = [x_c, y_c]^T$ .

During SL 3D system calibration,  $X_c$  and  $Y_c$  are known, the unknown  $\theta_c$  can be solved by least squares method [31].

$$\theta_c = (X_c^T X_c)^{-1} X_c^T Y_c. \quad (26)$$

The mathematical relationship between the projector coordinate system and the world coordinate system also follows the rule of the pinhole camera model that is formulated as follows.

$$Z_p \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^{wp}, & m_{12}^{wp}, & m_{13}^{wp}, & m_{14}^{wp} \\ m_{21}^{wp}, & m_{22}^{wp}, & m_{23}^{wp}, & m_{24}^{wp} \\ m_{31}^{wp}, & m_{32}^{wp}, & m_{33}^{wp}, & m_{34}^{wp} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (27)$$

where  $(x_p, y_p)$  is the camera coordinate. To decrease the number of unknown parameters, we also let  $m_{34}^{wp} = 1$ . Equation (27) is then expanded in the same way as Equation (17) into three equations. Since the projected pattern only changes in the  $y$  direction, only  $y_p$  will be used for 3D measurement. Thus, only the following equation is used.

$$\begin{aligned} & m_{21}^{wp} X_w + m_{22}^{wp} Y_w + m_{23}^{wp} Z_w + m_{24}^{wp} \\ & = y_p m_{31}^{wp} X_w + y_p m_{32}^{wp} Y_w + y_p m_{33}^{wp} Z_w + y_p \end{aligned} \quad (28)$$

Equation (28) can be rewritten into the following format:

$$X_p \theta_p = Y_p \quad (29)$$

where  $X_p = [X_w, Y_w, Z_w, 1, -y_p X_w, -y_p Y_w, -y_p Z_w]$ ,  $\theta_p = [m_{21}^{wp}, m_{22}^{wp}, m_{23}^{wp}, m_{24}^{wp}, m_{31}^{wp}, m_{32}^{wp}, m_{33}^{wp}]^T$  and  $Y_p = [y_p]$ .

During calibration,  $X_p$  and  $Y_p$  are known, the unknown  $\theta_p$  can be solved by the least squares method [31].

$$\theta_p = (X_p^T X_p)^{-1} X_p^T Y_p. \quad (30)$$

After  $\theta_c$  and  $\theta_p$  are obtained, they are used for 3D measurement by the following equation.

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = H^{-1} \begin{bmatrix} x_c - m_{14}^{wc} \\ y_c - m_{24}^{wc} \\ y_p - m_{14}^{wp} \end{bmatrix} \quad (31)$$

$$\text{where } H = \begin{bmatrix} m_{11}^{wc} - x_c m_{31}^{wc}, & m_{12}^{wc} - x_c m_{32}^{wc}, & m_{13}^{wc} - x_c m_{33}^{wc} \\ m_{21}^{wc} - y_c m_{31}^{wc}, & m_{22}^{wc} - y_c m_{32}^{wc}, & m_{23}^{wc} - y_c m_{33}^{wc} \\ m_{11}^{wp} - y_p m_{21}^{wp}, & m_{12}^{wp} - y_p m_{22}^{wp}, & m_{13}^{wp} - y_p m_{23}^{wp} \end{bmatrix}.$$

The coded line pattern contains  $L$  indexed parallel lines. When the coded line pattern is projected onto the object, the parallel lines will be distorted by the depth of the object. The camera acquires the distorted pattern image with resolution  $M \times N$ . The distorted lines in the acquired pattern image are clustered according to their indexes in the coded line pattern.  $y_p$  is computed as the  $y$  coordinate of the pixel on the distorted line. The computed  $y_p$  is stacked into an  $N \times 1$  vector for each clustered line. All the vectors are then stacked into an  $L \times N$  matrix. To make  $y_p$  correspond to the  $(x_c, y_c)$  one by one, the  $L \times N$  matrix is transformed into an  $M \times N$  matrix by spline interpolation. As a result, the object can be reconstructed with the same resolution of the camera. Both the reconstruction resolution and the reconstruction accuracy can be increased when the number of indexed lines in the coded pattern is increased. On the other hand, it will be more challenging to segment and cluster dense lines.

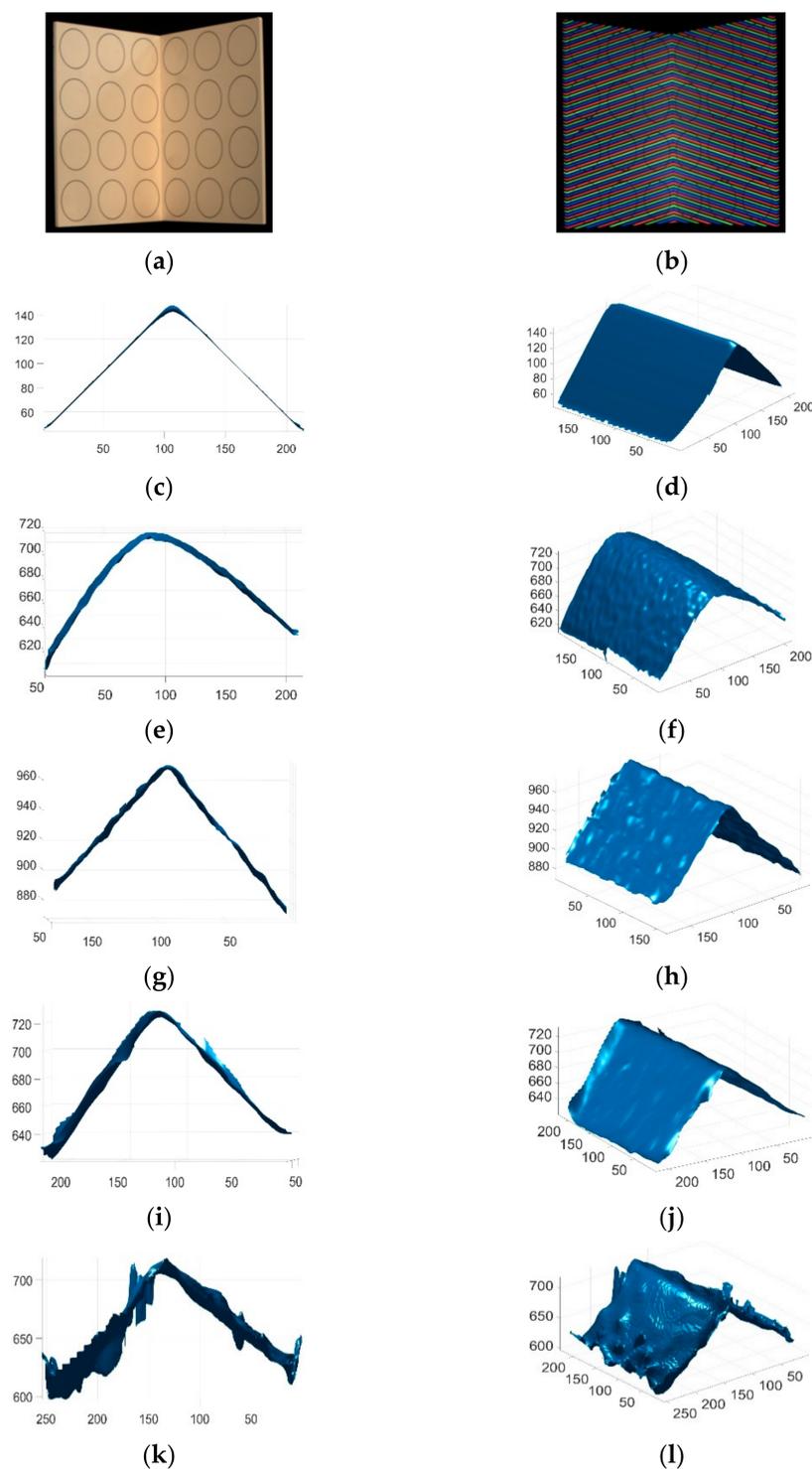
#### 4. Results and Discussion

In our established structured light 3D imaging system, a BENQ MU706 projector with the resolution  $1920 \times 1200$  was used to project the RGB line pattern. The projector was calibrated manually as follows: (1) the projection angle of the projector with reference to the object was selected as  $90^\circ$ ; (2) the distance between the projector and the object was adjusted to make sure that the projected pattern could cover the object completely. In the conducted experiments, the distance was chosen as 70 cm; (3) the focus of the projector was adjusted manually to make sure that the RGB lines were clearly shown on the surface of the object; (4) the contrast and brightness of the projector were adjusted manually to make sure that the RGB lines acquired by the camera were clear enough for line segmentation. A Blackfly BFLY-PGE-50A2C-CS camera with the resolution  $2592 \times 1944$  was used to acquire the pattern images. The camera was calibrated with the MATLAB camera calibrator toolbox to calculate the distortion parameters  $\{k_1, k_2, p_1, p_2\}$  and they were calculated as  $k_1 = -0.34995$ ,  $k_2 = 0.15753$ ,  $p_1 = -0.0003$ , and  $p_2 = 0.000377$ , respectively. The structured light 3D imaging system was calibrated by a calibration grid with Equation (26) and Equation (30) to compute the parameters  $\theta_c$  and  $\theta_p$ , respectively. The  $\theta_c$  computed by the calibration grid with 24 known marker points was  $[3.7747, 0.0174, -1.0601, 974.6918, -0.0827, 3.8175, -0.8671, 383.6297, -0.0000, -0.0000, -0.0010]$  and the  $\theta_p$  computed by the calibration grid with 24 known marker points was  $[-0.1332, 3.5434, 0.4618, 335.1717, -0.0000, -0.0002, -0.0009]$ . After 3D calibration, the calibration grid was used to compute the reconstruction accuracy of the structured light 3D imaging system and the computed mean root squared error (MRSE) was 0.4581 mm.

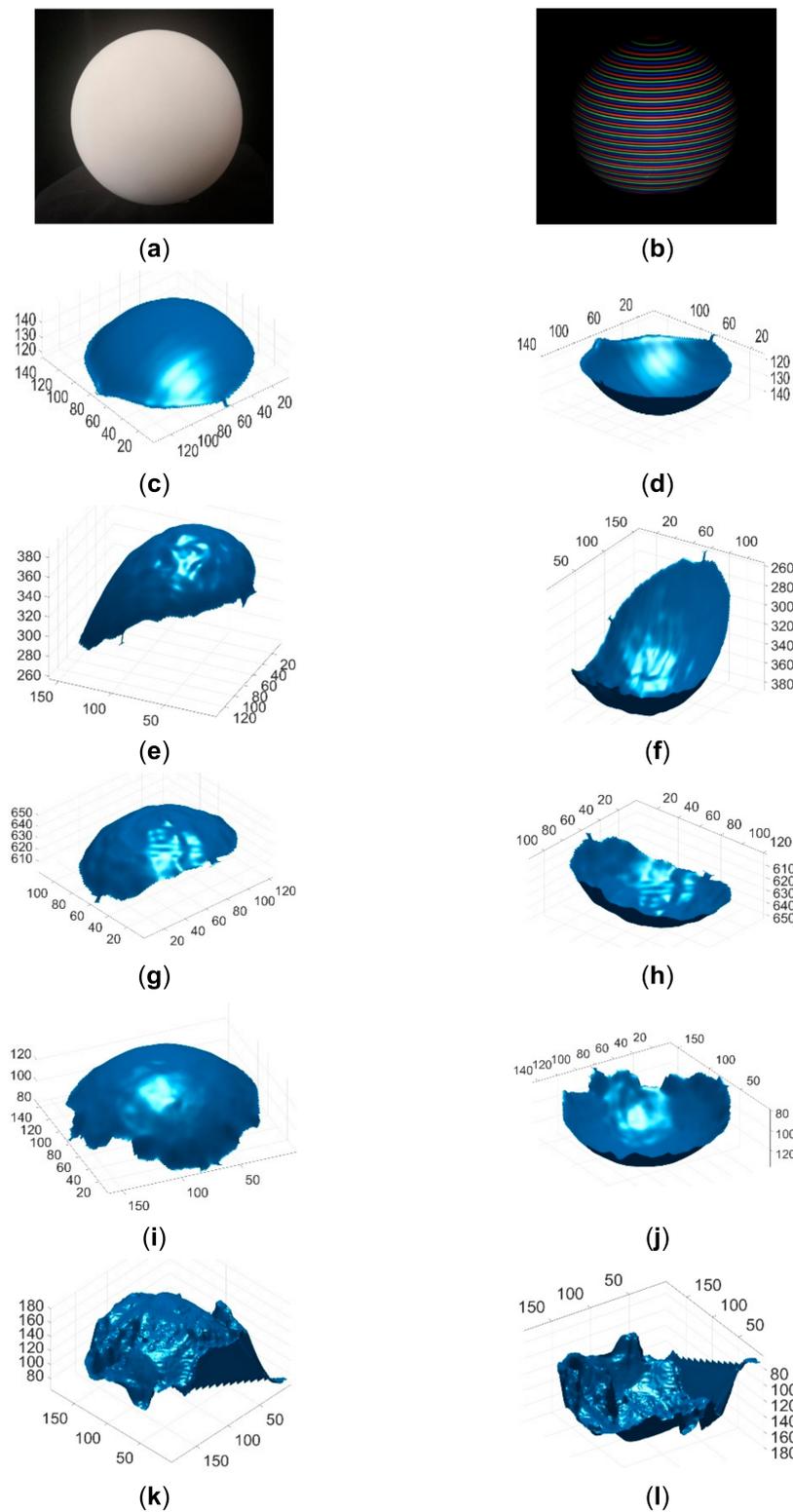
The proposed method was compared with state-of-the-art single-shot 3D surface imaging methods (products) in measuring both static and moving objects. The compared methods included: (1) passive stereo vision (Intel realsense D435 SV); (2) active stereo vision (Intel RealSense D435 ASV); (3) light coding (Kinect V1); and (4) time of flight (Kinect V2). To compare the reconstruction accuracy quantitatively, these single-shot imaging systems were used to reconstruct the calibration grid and a ball. The reconstruction error was evaluated with MRSE, and the reconstruction errors of the different methods are compared in Table 2. As can be seen, the accuracy of the proposed method was significantly better than those of existing methods. The reconstructed calibration grids and the reconstructed balls of different methods are shown in Figures 11 and 12, respectively. As can be seen, the reconstruction results of the existing methods contained significantly more noise. In addition, there were obvious distortions in the reconstruction results of existing methods. The angle between the two planes of the calibration grid was  $90^\circ$ , and the proposed method restored the angle much better than the other methods.

**Table 2.** Quantitative accuracy comparison of the proposed method with existing single-shot methods (products) (The bold values denote the best results).

Methods\Objects	Grid	Ball
Intel D435 SV	5.83 mm	5.3 mm
Intel D435 ASV	0.79 mm	0.86 mm
Kinect V1	14.36 mm	1.18 mm
Kinect V2	3.05 mm	0.36 mm
<b>Proposed</b>	<b>0.46 mm</b>	<b>0.24 mm</b>



**Figure 11.** Reconstructed calibration grids by different methods. (a) The image of the calibration grid; (b) The acquired pattern image of the calibration grid; (c) Reconstructed calibration grid by the proposed method in view 1; (d) Reconstructed calibration grid by the proposed method in view 2; (e) Reconstructed calibration grid by Kinect V2 in view 1; (f) Reconstructed calibration grid by Kinect V2 in view 2; (g) Reconstructed calibration grid by Kinect V1 in view 1; (h) Reconstructed calibration grid by Kinect V1 in view 2; (i) Reconstructed calibration grid by Intel D435 ASV in view 1; (j) Reconstructed calibration grid by Intel D435 ASV in view 2; (k) Reconstructed calibration grid by Intel D435 SV in view 1; (l) Reconstructed calibration grid by Intel D435 SV in view 2.



**Figure 12.** Reconstructed balls by different methods. (a) The image of the ball; (b) The acquired pattern image of the ball; (c) Reconstructed ball by the proposed method in view 1; (d) Reconstructed ball by the proposed method in view 2; (e) Reconstructed ball by Kinect V2 in view 1; (f) Reconstructed ball by Kinect V2 in view 2; (g) Reconstructed ball by Kinect V1 in view 1; (h) Reconstructed ball by Kinect V1 in view 2; (i) Reconstructed ball by Intel D435 ASV in view 1; (j) Reconstructed ball by Intel D435 ASV in view 2; (k) Reconstructed ball by Intel D435 SV in view 1; (l) Reconstructed ball by Intel D435 SV in view 2.

These single-shot imaging systems were used to reconstruct two colorless and static objects for qualitative comparisons. One colorless object was a simple cone with cylinder, and its reconstruction results by different methods are shown in Figure 13. The other colorless object was a complex face statue, and its reconstruction results by different methods are shown in Figure 14. As can be seen, the reconstruction results of the proposed method looked better than those of existing methods. In addition, these imaging systems were used to reconstruct three static objects with colors. The reconstruction results are shown in Figures 15 and 16, respectively. As can be seen, the reconstruction results of the proposed method also looked better than those of the existing methods. Finally, these single-shot imaging systems were used to reconstruct deforming and moving objects. The typical results of reconstructing a deforming face by different methods are shown in Figure 17, and more reconstruction results of deforming objects by different methods are available in the attached videos (Deforming face and Deforming hand). From all the reconstruction results by different methods, it is seen that the proposed method is significantly more accurate than state-of-the-art single-shot 3D imaging methods.

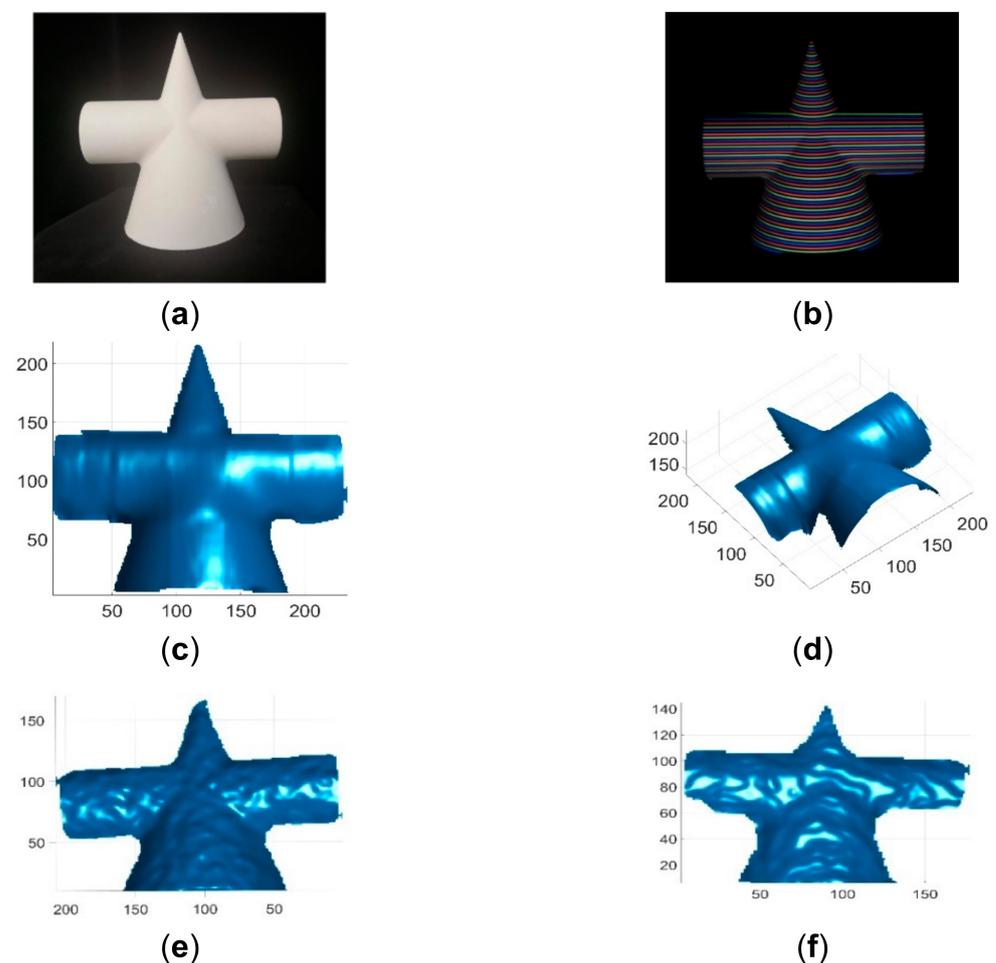
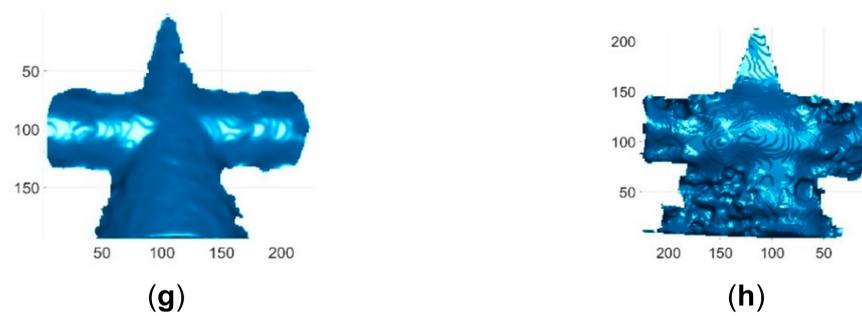
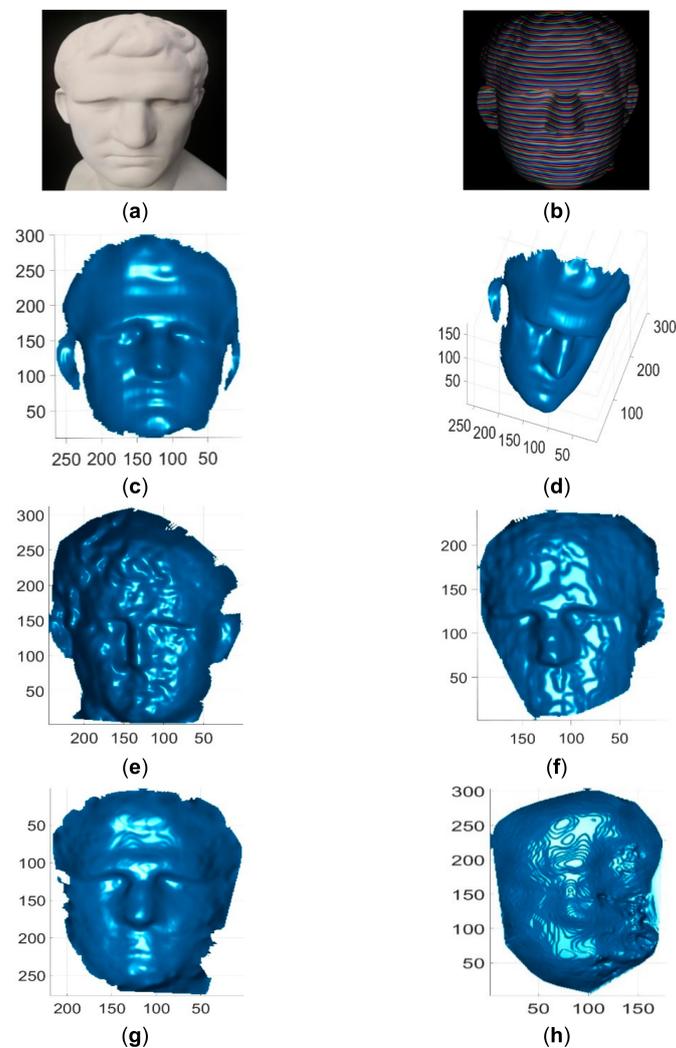


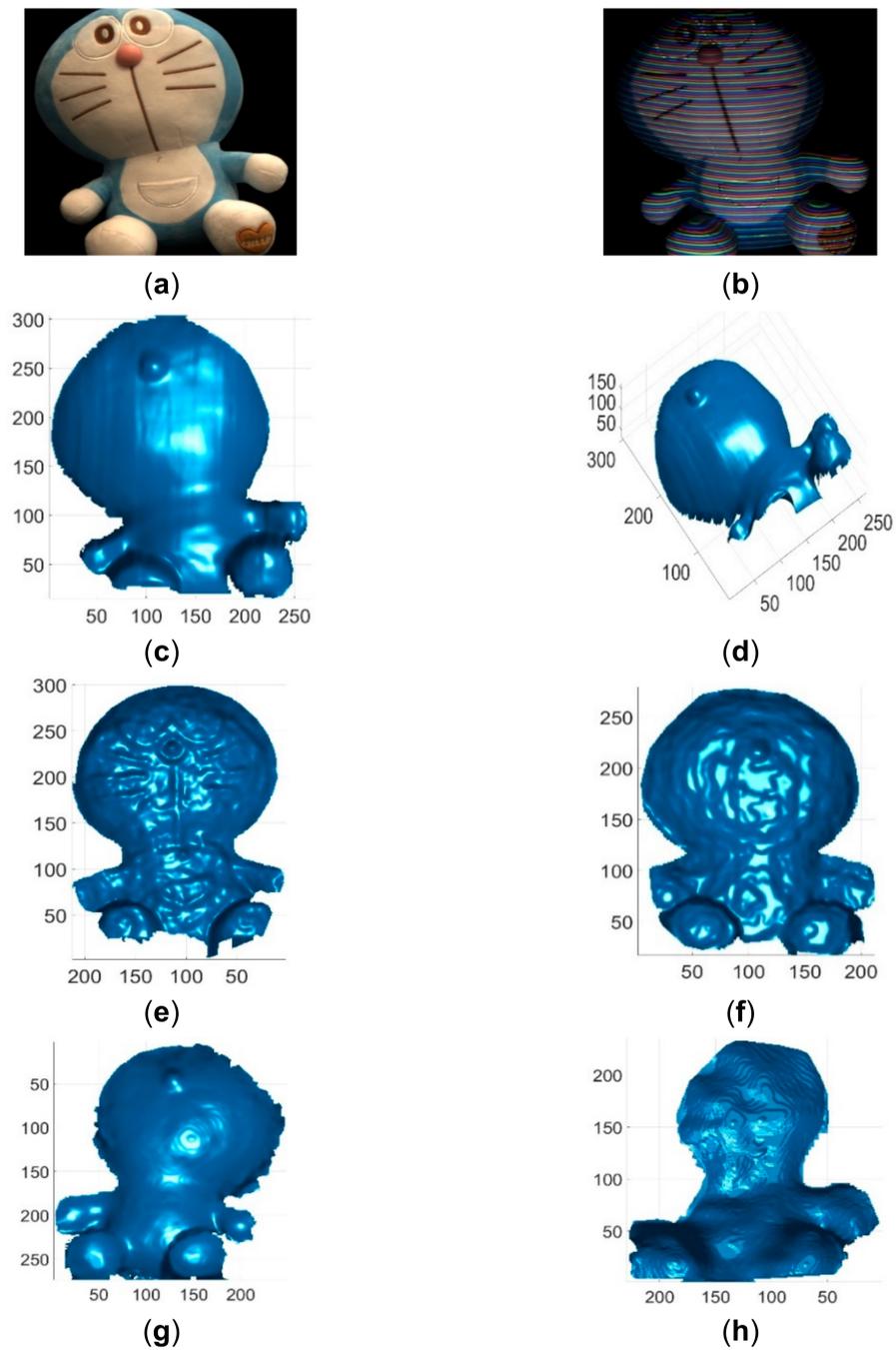
Figure 13. Cont.



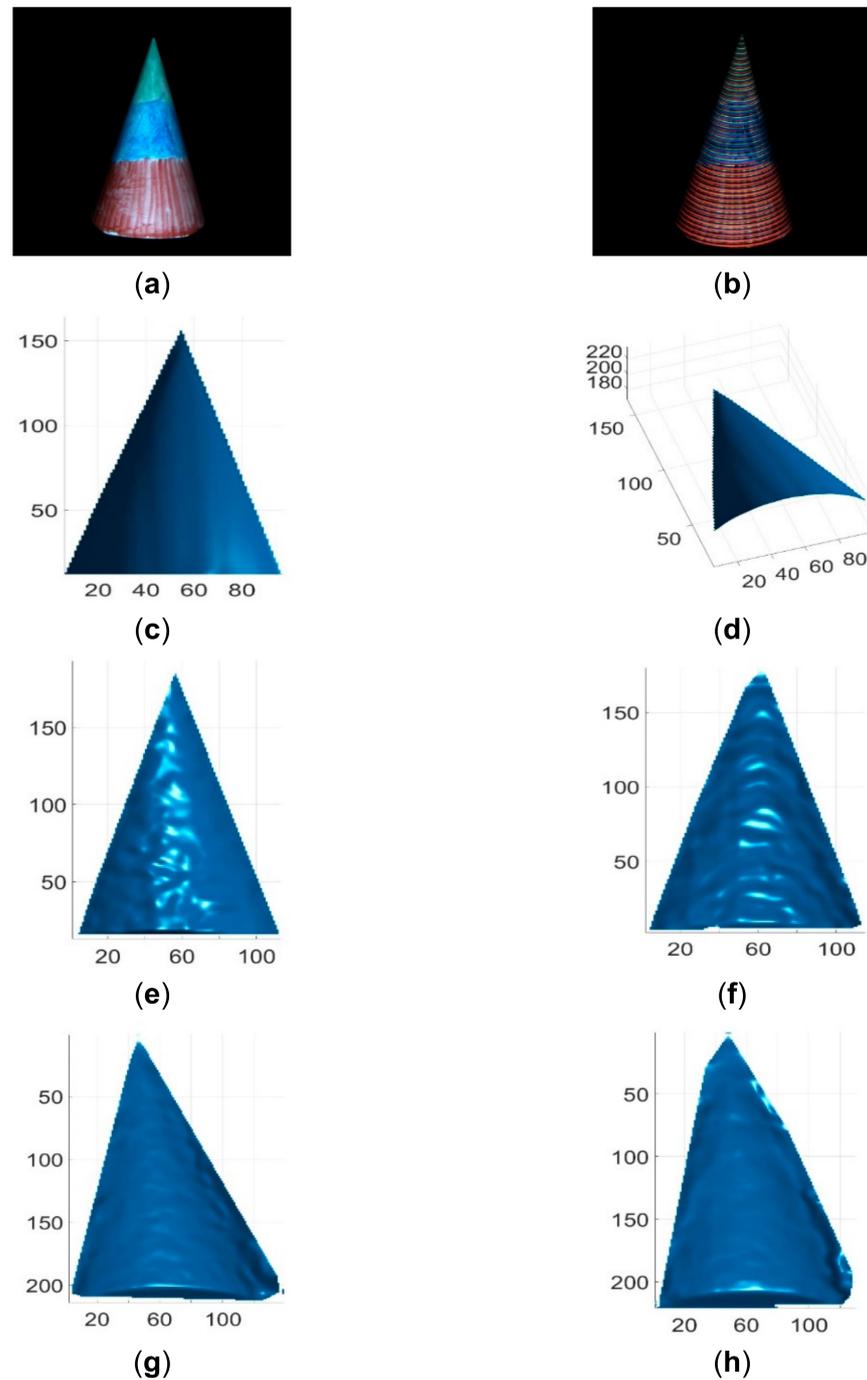
**Figure 13.** Reconstruction results of a cone with cylinder. (a) The cone with cylinder image; (b) The acquired pattern image of the cone with cylinder; (c) Reconstructed cone with cylinder by the proposed method in view 1; (d) Reconstructed cone with cylinder by the proposed method in view 2; (e) Reconstructed cone with cylinder by Kinect V2; (f) Reconstructed cone with cylinder by Kinect V1; (g) Reconstructed cone with cylinder by Intel D435 ASV; (h) Reconstructed cone with cylinder by Intel D435 SV.



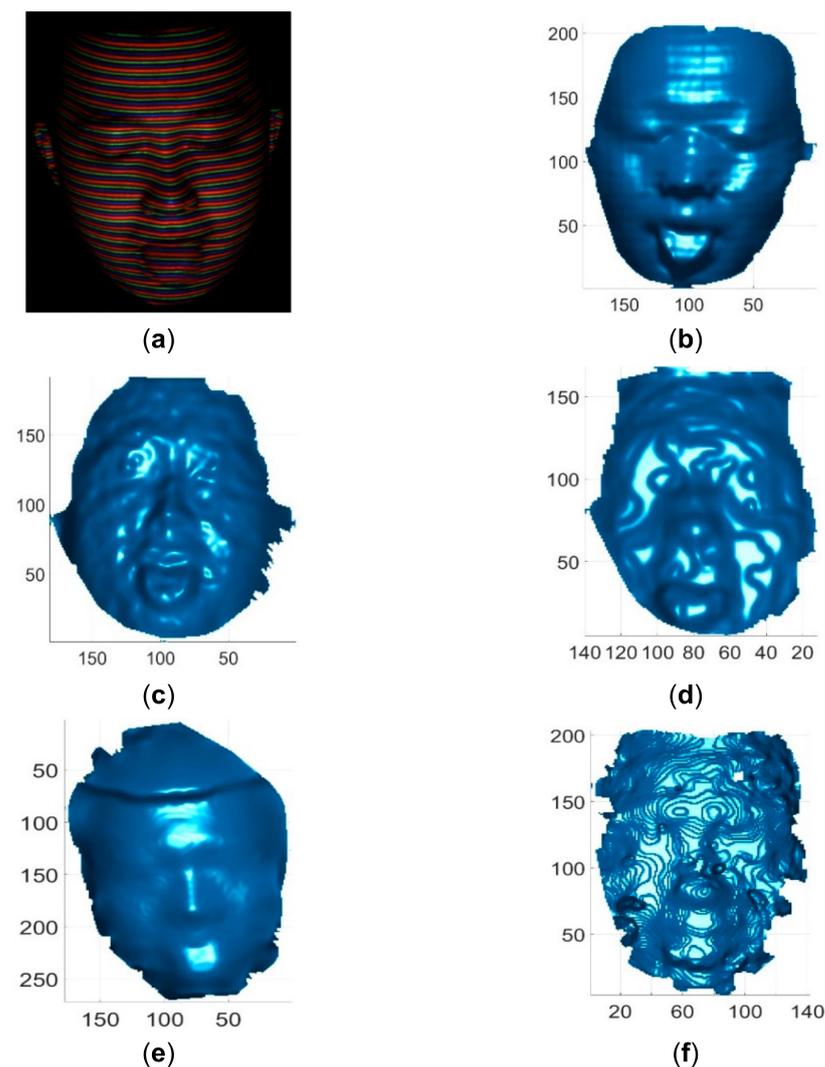
**Figure 14.** Reconstruction results of a face statue. (a) The face statue image; (b) The acquired pattern image of the face statue; (c) Reconstructed face statue by the proposed method in view 1; (d) Reconstructed face statue by the proposed method in view 2; (e) Reconstructed face statue by Kinect V2; (f) Reconstructed face statue by Kinect V1; (g) Reconstructed face statue by Intel D435 ASV; (h) Reconstructed face statue by Intel D435 SV.



**Figure 15.** Reconstruction results of a colorful toy. (a) The toy image; (b) The acquired pattern image; (c) Reconstructed colorful toy by the proposed method in view 1; (d) Reconstructed colorful toy by the proposed method in view 2; (e) Reconstructed colorful toy by Kinect V2; (f) Reconstructed colorful toy by Kinect V1; (g) Reconstructed colorful toy by Intel D435 ASV; (h) Reconstructed colorful toy by Intel D435 SV.

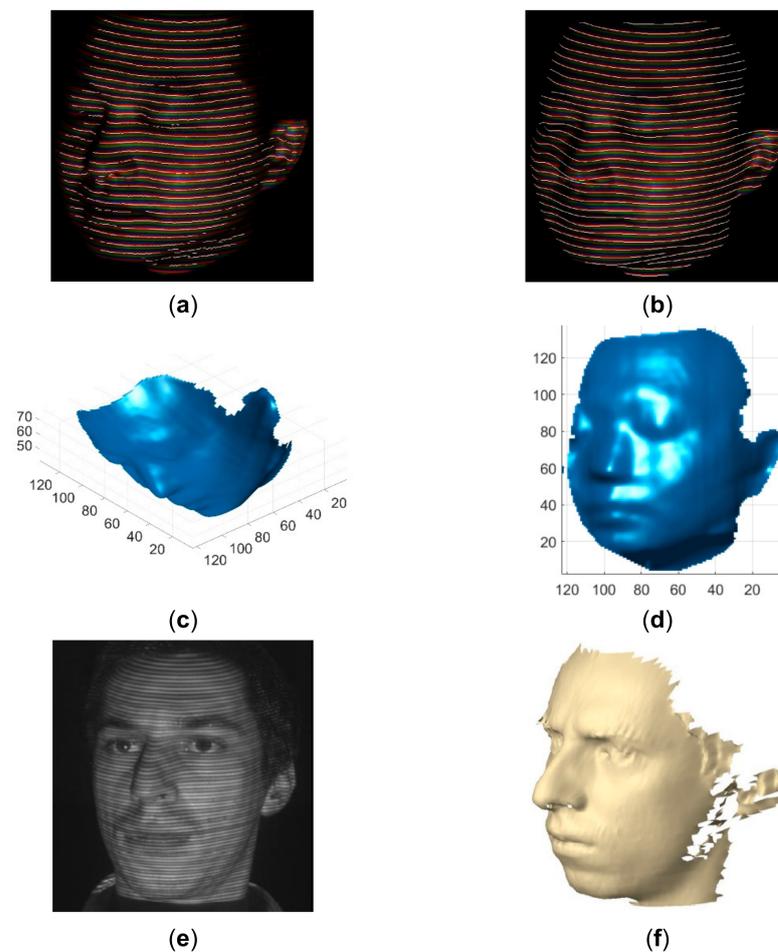


**Figure 16.** Reconstruction results of a cone with RGB color. (a) The RGB cone image; (b) The acquired pattern image of the cone; (c) Reconstructed cone by the proposed method in view 1; (d) Reconstructed cone by the proposed method in view 2; (e) Reconstructed cone by Kinect V2; (f) Reconstructed cone by Kinect V1; (g) Reconstructed cone by Intel D435 ASV; (h) Reconstructed cone by Intel D435 SV.

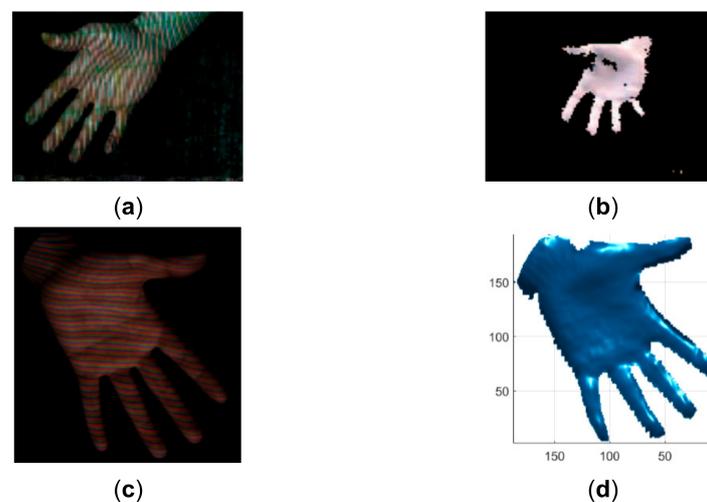


**Figure 17.** Reconstruction results of a deforming human face. (a) The acquired pattern image of the human face; (b) Reconstructed human face by the proposed method; (c) Reconstructed human face by Kinect V2; (d) Reconstructed human face by Kinect V1; (e) Reconstructed human face by Intel D435 ASV; (f) Reconstructed human face by Intel D435 SV.

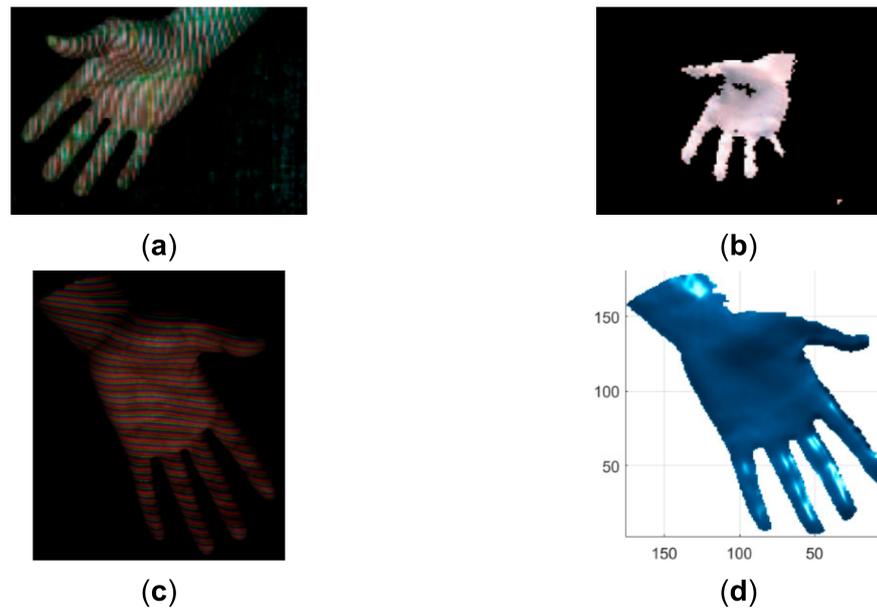
To verify the advantage of the proposed method over existing line indexing methods [28,30], the results of reconstructing similar objects are compared in Figures 18–23. The red lines detected by local maxima [28] are shown in Figure 18a and the red lines extracted by the proposed method are shown in Figure 18b. It is seen that the lines extracted by the proposed method are more accurate and more complete. In addition, the lines detected by local maxima could not be clustered robustly by the method proposed in this paper and the methods proposed previously in [7–9]. By comparing the reconstructed face by the proposed method in Figure 18c,d and the reconstructed face from [28] in Figure 18f, it is seen that the ear part reconstructed by the proposed method is significantly better. Similar to [30], a moving and deforming hand was reconstructed by the proposed method and the reconstructed hands were compared with those with similar gestures from [30] in Figures 19–23. As can be seen, the reconstruction results of the proposed method are more accurate and more complete than the reconstruction results from [30].



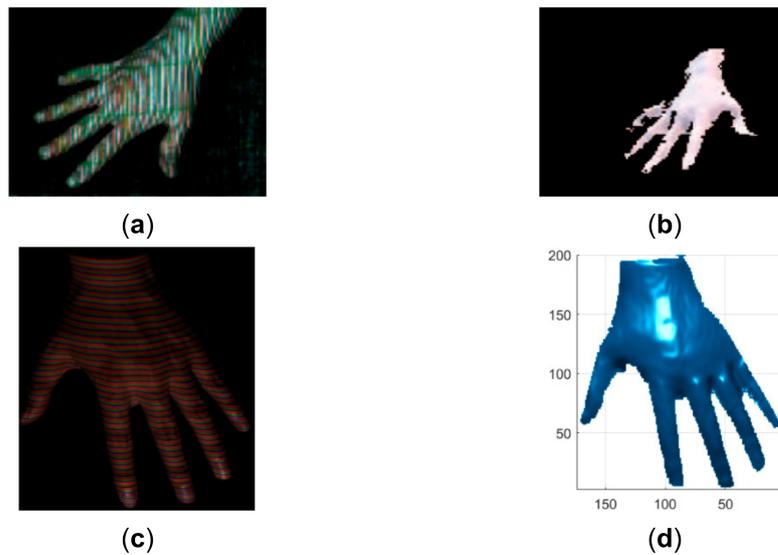
**Figure 18.** Qualitative comparison of the proposed method with the method proposed in [28]. (a) The detected red lines by local maxima superimposed on the acquired pattern image; (b) The extracted red lines by the proposed method superimposed on the acquired pattern image; (c) The reconstructed human face by the proposed method in view 1; (d) The reconstructed human face by the proposed method in view 2; (e) The acquired pattern image in [28]; (f) The reconstructed human face by the method in [28].



**Figure 19.** Qualitative comparison of the proposed method with the method proposed in [30]. (a) The acquired pattern image in [30]; (b) The reconstructed hand in [30]; (c) The acquired pattern image in this study; (d) The reconstructed hand by the proposed method.



**Figure 20.** Qualitative comparison of the proposed method with the method proposed in [30]. (a) The acquired pattern image in [30]; (b) The reconstructed hand in [30]; (c) The acquired pattern image in this study; (d) The reconstructed hand by the proposed method.



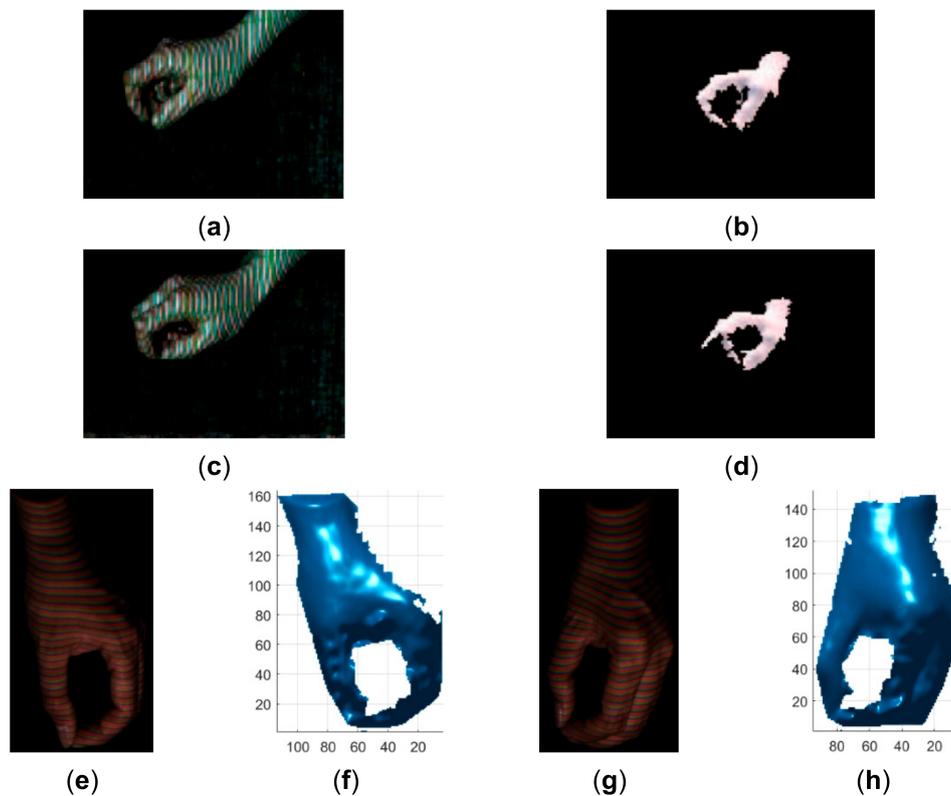
**Figure 21.** Qualitative comparison of the proposed method with the method proposed in [30]. (a) The acquired pattern image in [30]; (b) The reconstructed hand in [30]; (c) The acquired pattern image in this study; (d) The reconstructed hand by the proposed method.



**Figure 22.** Cont.



**Figure 22.** Qualitative comparison of the proposed method with the method proposed in [30]. (a) The acquired pattern image in [30]; (b) The reconstructed hand in [30]; (c) The acquired pattern image in this study; (d) The reconstructed hand by the proposed method.



**Figure 23.** Qualitative comparison of the proposed method with the method proposed in [30]. (a) The acquired pattern image in [30]; (b) The reconstructed hand in [30]; (c) The acquired pattern image in [30]; (d) The reconstructed hand in [30]; (e) The acquired pattern image in this study; (f) The reconstructed hand by the proposed method; (g) The acquired pattern image in this study; (h) The reconstructed hand by the proposed method.

As verified by the experimental results, the usability of the proposed method includes profile measurement, 3D motion reconstruction and other related or similar applications. To use the proposed method correctly and to reduce the measurement uncertainties, users must know the limitations of the proposed method as follows.

- (1) The measurement distance between the object and the structured light 3D imaging system is constrained in a relatively small range to make sure that both the camera and the projector are in focus, which is a common limitation of structured light technology.
- (2) The proposed method is not suitable for applications that require the reconstruction results to be full of accurate 3D details, such as defect detection.

## 5. Conclusions

In conclusion, a RGB line-pattern-based single-shot structured light method is proposed in this paper. The proposed method addresses the challenges of line segmentation and line indexing more effectively and systematically than existing methods [26–30]. The RGB lines are segmented by SDD multiple-threshold selection in the HSV color space, which is significantly more accurate than classic classification methods. The guaranteed line segmentation accuracy makes the subsequent line indexing feasible. In this paper, a more robust line clustering method is proposed to index the segmented lines by an exclusion scheme, which is more effective than existing methods in dealing with the occlusions and large discontinuities. The proposed single-shot structured light method was verified by extensive experimental results of reconstructing a deforming face and a deforming hand. In addition, the accuracy of the proposed approach in measuring a calibration grid was 0.46 mm, which is significantly better than the second-best accuracy, 0.79 mm achieved by Intel D435 ASV. The accuracy of the proposed approach in measuring a ball was 0.24 mm, which is significantly better than the second-best accuracy, 0.36 mm achieved by Kinect V2.

**Author Contributions:** Conceptualization, Z.W.; methodology, Y.L. and Z.W.; software, Y.L.; validation, Z.W.; formal analysis, Z.W.; investigation, Z.W.; resource, Z.W.; data curation, Y.L.; writing—original draft, Y.L.; writing—review & editing, Z.W.; visualization, Y.L.; supervision, Z.W.; project administration, Z.W.; funding acquisition, Z.W. Both authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Natural Science Foundation of Shandong Province, grant number ZR2020MF018. The APC was funded by Shandong University of Technology.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This research work is funded by Natural Science Foundation of Shandong Province with the grant no. ZR2020MF018.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wang, Z.Z. Review of real-time three-dimensional shape measurement techniques. *Measurement* **2020**, *156*, 107624. [[CrossRef](#)]
2. Al Ismaeil, K.; Aouada, D.; Solignac, T.; Mirbach, B.; Ottersten, B. Real-Time Enhancement of Dynamic Depth Videos with Non-Rigid Deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 2045–2059. [[CrossRef](#)] [[PubMed](#)]
3. Sabater, N.; Almansa, A.; Morel, J.-M. Meaningful Matches in Stereovision. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *34*, 930–942. [[CrossRef](#)] [[PubMed](#)]
4. Jepsen, G. *Projectors for Intel® RealSense™ Depth Cameras D4xx*. Intel Support; Intel Corporation: Santa Clara, CA, USA, 2018.
5. Ulrich, L.; Vezzetti, E.; Moos, S.; Marcolin, F. Analysis of RGB-D camera technologies for supporting different facial usage scenarios. *Multimed. Tools Appl.* **2020**, *79*, 29375–29398. [[CrossRef](#)]
6. Takeda, M.; Motoh, K. Fourier transform profilometry for the automatic measurement of 3-D object shapes. *Appl. Opt.* **1983**, *22*, 3977–3982. [[CrossRef](#)]
7. Wang, Z.; Yang, Y. Single-shot three-dimensional reconstruction based on structured light line pattern. *Opt. Lasers Eng.* **2018**, *106*, 10–16. [[CrossRef](#)]
8. Wang, Z.Z. Unsupervised Recognition and Characterization of the Reflected Laser Lines for Robotic Gas Metal Arc Welding. *IEEE Trans. Ind. Inform.* **2017**, *13*, 1866–1876. [[CrossRef](#)]
9. Wang, Z.Z. Robust three-dimensional face reconstruction by one-shot structured light line pattern. *Opt. Lasers Eng.* **2020**, *124*, 105798. [[CrossRef](#)]
10. Beumier, C.; Acheroy, M. Automatic Face Authentication from 3D surface. *Br. Mach. Vis. Conf.* **1998**, *45*, 45. [[CrossRef](#)]
11. Kemelmacher-Shlizerman, I.; Basri, R. 3D Face Reconstruction from a Single Image Using a Single Reference Face Shape. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 394–405. [[CrossRef](#)]
12. Garrido, P.; Valgaerts, L.; Wu, C.; Theobalt, C. Reconstructing detailed dynamic face geometry from monocular video. *ACM Trans. Graph.* **2013**, *32*, 1.
13. Moss, J.; Linney, A.; Grindrod, S.; Mosse, C. A laser scanning system for the measurement of facial surface morphology. *Opt. Lasers Eng.* **1989**, *10*, 179–190. [[CrossRef](#)]

14. You, Y.; Shen, Y.; Zhang, G.; Xing, X. Real-Time and High-Resolution 3D Face Measurement via a Smart Active Optical Sensor. *Sensors* **2017**, *17*, 734. [[CrossRef](#)] [[PubMed](#)]
15. Vilchez-Rojas, H.L.; Rayas, J.A.; Martínez-García, A. Use of white light profiles for the contouring of objects. *Opt. Lasers Eng.* **2020**, *134*, 106295. [[CrossRef](#)]
16. Chen, S.Y.; Li, Y.F.; Zhang, J. Vision Processing for Realtime 3-D Data Acquisition Based on Coded Structured Light. *IEEE Trans. Image Process.* **2008**, *17*, 167–176. [[CrossRef](#)]
17. Payeur, P.; Desjardins, D. Structured Light Stereoscopic Imaging with Dynamic Pseudo-random Patterns. In *Image Analysis and Recognition, ICIAR 2009; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2009; Volume 5627*, pp. 687–696.
18. Griffin, P.M.; Narasimhan, L.S.; Yee, S.R. Generation of uniquely encoded light patterns for range data acquisition. *Pattern Recognit.* **1992**, *25*, 609–616. [[CrossRef](#)]
19. Morano, R.; Ozturk, C.; Conn, R.; Dubin, S.; Zietz, S.; Nissano, J. Structured light using pseudorandom codes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 322–327. [[CrossRef](#)]
20. Ito, M.; Ishii, A. A three-level checkerboard pattern (TCP) projection method for curved surface measurement. *Pattern Recognit.* **1995**, *28*, 27–40. [[CrossRef](#)]
21. Vuylsteke, P.; Oosterlinck, A. Range image acquisition with a single binary-encoded light pattern. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 148–164. [[CrossRef](#)]
22. Wang, J.; Xiong, Z.; Wang, Z.; Zhang, Y.; Wu, F. FPGA Design and Implementation of Kinect-Like Depth Sensing. *IEEE Trans. Circuits Syst. Video Technol.* **2015**, *26*, 1175–1186. [[CrossRef](#)]
23. Khoshelham, K.; Elberink, S.O. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors* **2012**, *12*, 1437–1454. [[CrossRef](#)]
24. Zhang, Y.; Xiong, Z.; Yang, Z.; Wu, F. Real-Time Scalable Depth Sensing with Hybrid Structured Light Illumination. *IEEE Trans. Image Process.* **2013**, *23*, 97–109. [[CrossRef](#)]
25. Li, W.; Li, H.; Zhang, H. Light plane calibration and accuracy analysis for multi-line structured light vision measurement system. *Optik* **2020**, *207*, 163882. [[CrossRef](#)]
26. Je, C.; Lee, S.W.; Park, R.-H. Colour-stripe permutation pattern for rapid structured-light range imaging. *Opt. Commun.* **2012**, *285*, 2320–2331. [[CrossRef](#)]
27. Robinson, A.; Alboul, L.; Rodrigues, M. Methods for indexing stripes in uncoded structured light scanning systems. *J. WSCG* **2004**, *12*, 371–378.
28. Brink, W.; Robinson, A.; Rodrigues, M.A. Indexing Uncoded Stripe Patterns in Structured Light Systems by Maximum Spanning Trees. *BMVC* **2008**, *2018*, 1–10.
29. Boyer, K.L.; Kak, A.C. Color-Encoded Structured Light for Rapid Active Ranging. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *9*, 14–28. [[CrossRef](#)]
30. Koninckx, T.; Van Gool, L. Real-time range acquisition by adaptive structured light. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 432–445. [[CrossRef](#)]
31. Yalla, V.G.; Hassebrook, L.G. Very high resolution 3D surface scanning using multi-frequency phase measuring profilometry. *Def. Secur.* **2005**, *5798*, 44–53. [[CrossRef](#)]
32. Wang, Z. Robust measurement of the diffuse surface by phase shift profilometry. *J. Opt.* **2014**, *16*, 105407. [[CrossRef](#)]
33. Wang, Z. Robust segmentation of the colour image by fusing the SDD clustering results from different colour spaces. *IET Image Process.* **2020**, *14*, 3273–3281. [[CrossRef](#)]
34. Wang, Z. Automatic Localization and Segmentation of the Ventricles in Magnetic Resonance Images. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 621–631. [[CrossRef](#)]
35. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. [[CrossRef](#)]
36. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [[CrossRef](#)]
37. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.* **1977**, *39*, 1–38.
38. Hartley, R.; Zisserman, A.; Faugeras, O. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004. [[CrossRef](#)]