

# 3D Surface Profilometry using Phase Shifting of De Bruijn Pattern

Matea Donlić, Tomislav Petković and Tomislav Pribanić  
University of Zagreb Faculty of Electrical Engineering and Computing  
Unska 3, HR-10000 Zagreb, Croatia  
{matea.donlic, tomislav.petkovic.jr, tomislav.pribanic}@fer.hr

## Abstract

*A novel structured light method for color 3D surface profilometry is proposed. The proposed method does not require color calibration of a camera-projector pair and may be used for reconstruction of both dynamic and static scenes. The method uses a structured light pattern that is a combination of a De Bruijn color sequence and of a sinusoidal fringe. For dynamic scenes a Hessian ridge detector and a Gaussian mixture model are combined to extract stripe centers and to identify color. Stripes are then uniquely identified using dynamic programming based on the Smith-Waterman algorithm and a De Bruijn window property. For static scenes phase-shifting and De Bruijn window property are combined to obtain a high accuracy reconstruction. We have tested the proposed method on multiple objects with challenging surfaces and different albedos that demonstrate usability and robustness of the method.*

## 1. Introduction

Structured light projection is a widely studied topic in the field of 3D imaging where reconstruction needs to be fast, noninvasive and inexpensive. Non-contact 3D surface profilometry based on structured light is the method of choice in reverse engineering, industrial quality control of product parts, object recognition, health-care applications and others. One of the main challenges is improving the accuracy of the reconstruction while projecting as few patterns as possible thus enabling reduction of acquisition time and reconstruction of moving objects.

Based on the scene type structured light methods may be classified into dynamic and static scene reconstruction methods. For dynamic scenes usually only one image is projected, therefore methods are also called one-shot methods. For static scenes multiple images are projected over time, therefore are called multiple-shot methods. Multiple-shot methods produce superior reconstruction than single-shot methods in terms of both resolution

and accuracy, but have a substantially longer acquisition times and are mostly unsuitable for dynamic scenes. There are approaches in which multiple images are used for dynamic scene reconstruction but expensive high-speed hardware is required [11, 26]. A more detailed classification of various reconstruction methods and structured light patterns is given in [18].

Several widely used one-shot methods [25, 29] are based on color stripe patterns that satisfy a window property. A finite sequence of length  $L$  satisfies a window property if every sub-sequence of length  $N$ , called a window, appears exactly once. The window property enables extraction of spatial position through identification of unique subsequences in a color sequence.

Usual problems that such approaches must solve are robust and precise detection of color stripe centers (or edges) and accurate color that does not depend on camera-projector color crosstalk, object albedo, and ambient lightning. Many have tried to solve these problems by using some form of color calibration to eliminate the undesired effects. Such approaches usually involve projection of multiple solid color patterns from which a crosstalk matrix, albedo coefficients, and ambient illumination are estimated [2, 25, 29]. Alternatively, a tedious photometric calibration of a camera-projector pair may be performed [9].

A large number of state-of-the art multiple-shot methods are based on a combination of (multiple) phase-shifting, Gray coding, and color coding [18]. In (multiple) phase-shifting a set of sinusoidal patterns is projected over time and pixel codification is based only on time-analysis. Phase estimation, a core of such methods, returns a phase estimate that is ambiguous by an unknown multiple of  $2\pi$  and is therefore called wrapped-phase. The main difference between various methods is how the problem of phase unwrapping is solved. This is usually achieved by projecting multiple spatial frequencies from which the required phase offset may be computed [16, 15]; or by using additional images that contain Gray-coded [19] or a color-coded pattern [6] that uniquely identifies a required multiple of  $2\pi$ . Some authors have proposed combining different patterns

in multiple channels of a color image [13, 24], effectively reducing number of required images and shortening acquisition time; such approaches give accurate unwrapped phase but cannot cope with discontinuities on objects with complicated surfaces [28].

In this paper we propose a novel one and multiple-shot approach for the 3D reconstruction of both dynamic and static scenes that is based on the proposed color structured light pattern combining a De Bruijn color sequence and a sinusoidal fringe. Compared to the current methods described in the literature the proposed method does not require precise color calibration of the camera-projector system as: (1) for one-shot reconstruction the colors are classified and recognized using Gaussian mixture models (GMM), and (2) for multiple-shot reconstruction acquired images are self-corrected by using a specific property of the proposed structured light pattern. There has been some previous work in various fields on the topic of image segmentation using GMM with images in HSI [7], HSV [8], RGB [23] color spaces. To the best of our knowledge GMM have not been applied to 3D surface profilometry. For one-shot reconstruction we also propose a novel approach to dynamic programming (DP) using bio-informatics inspired Smith-Waterman algorithm [21]. Dynamic programming is often used in stereo vision [20] and was applied to structured light to reconstruct correspondence between the projected pattern and the recorded image [12, 25, 29]. In both cases DP directly uses pixel intensities or some function of the intensities, however, we propose using detected color labels directly thus significantly simplifying the construction of the used cost function.

## 2. The Proposed Method

The proposed method uses a novel structured light pattern that enables both coarse one-shot reconstruction and fine multiple-shot reconstruction. In following sections we first present the proposed structured light pattern and describe how it is constructed. Then we give descriptions of the one-shot reconstruction for dynamic scenes and of the multiple-shot reconstruction for static scenes.

### 2.1. Structured Light Pattern

Proposed structured light pattern is comprised of multiple color images represented in the RGB color space that are produced by combining a De Bruijn color sequence and a sinusoidal fringe pattern.

A De Bruijn sequence of order  $N$  over an alphabet of size  $K$  is a pseudo-random cyclic sequence of length  $K^N$  [10] in which every possible substring (window) of length  $N$  appears exactly once. The proposed pattern uses  $N = 3$  and  $K = 6$  colors. Used colors are red, yellow, green, cyan, blue, and magenta (Table 1). We also impose two additional constraints on the sequence: (1) any two adjacent colors in

Color	Letter	(b,g,r)	Hue
Red	R	(0,0,1)	0°
Green	G	(0,1,0)	120°
Blue	B	(1,0,0)	240°
Yellow	Y	(0,1,1)	60°
Cyan	C	(1,1,0)	180°
Magenta	M	(1,0,1)	300°

Table 1: An alphabet of six colors from which a De Bruijn sequence is constructed.

the sequence must be different, i.e. subsequences RR, GG, YY, BB, MM, and CC are prohibited; and (2) in every subsequence of length  $N$  both minimum and maximum in all three channels must be achieved at least once, e.g. substring BMC is not allowed as all values in blue channel are 1 (at least one 0 is required), similarly substring RGB is allowed as both 0 and 1 appear at least once in all channels. The first condition improves the robustness of color decoding for dynamic scenes. The second condition is required for static scenes as it enables pixel-based color equalization that effectively eliminates object albedo and enables temporal color decoding. Therefore, the proposed sequence satisfying the second condition is effectively color self-correcting. Consequently, no color calibration is needed.

Required sequence is generated as a Hamiltonian path in a De Bruijn graph of order  $N$  where all graph vertices that do not satisfy the aforementioned two conditions are removed. The maximal length of a sequence for  $N = 3$  and  $K = 6$  that satisfies both conditions is  $L = 90$ . An example sequence is:

$$S = \left\{ \begin{array}{l} \text{RYBRGCRGBRCRCYRCGRCBYRBYGBYCMR} \\ \text{GMRCMYGMYBYBGRBGYBCRBCYBMGRMGY} \\ \text{MGCMMGCMRCYMCMBYMBGMGBMYCBRYC} \end{array} \right\} \quad (1)$$

Intensity of a sinusoidal fringe pattern for one row of the image is defined as

$$I(x) = I_{\max} \left( \frac{1}{2} - \frac{1}{2} \cos \left( \frac{2\pi}{P} x - \varphi_i \right) \right), \quad (2)$$

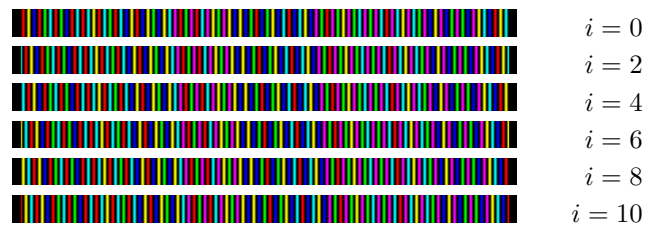


Figure 1: Six out of twelve images from the proposed structured light pattern for  $L = 90$  and  $N_p = 4$

where  $I_{\max}$  is the maximum intensity,  $x$  is image column index,  $P$  is the period in pixels, and  $\varphi_i$  is phase. Intensity (2) is then repeated for all rows.

To obtain one image of the proposed pattern we combine the color sequence (1) and the fringe pattern (2) in the HSV color space. The sequence (1) defines the hue: same hue is used for one period of the fringe pattern, therefore there are exactly  $L$  color fringes (periods) in the final image. The fringe pattern (2) defines the value. The saturation is always set to 1. A minimal pattern sequence is comprised of  $N \cdot N_p$  phase shifted images, where  $N$  is a De Bruijn window length and  $N_p$  is the number of phase shifts per one period of Eq. (2). For the  $i$ th image the colors in  $S$  are rotated by  $\text{mod}(i + N_p - 1, N_p)$  and  $I(x)$  is shifted by  $\varphi_i = (i \frac{2\pi}{N_p})$ , where  $i = 0, 1, \dots, N \cdot N_p - 1$ . An example of the proposed pattern sequence is shown in Fig. 1; note how the entire pattern shifts to the right and how used colors wrap around from the end to the beginning of the sequence.

The proposed pattern may be classified as a multi-slit pattern as the intensity modulation given by Eq. (2) has a side effect of introducing black slits between two color stripes. Every image from the pattern may be used for one-shot reconstruction of dynamic scenes, and whole  $N \cdot N_p$  images may be used for multiple-shot reconstruction of static scenes.

## 2.2. One-Shot Reconstruction

Proposed method for one-shot reconstruction is comprised of three steps: (a) detection of stripes and its centers in the acquired frame; (b) identification of stripe color; and (c) spatial decoding of the projected De Bruijn code.

### 2.2.1 Stripe Segmentation

A multi-scale vesselness measure based on an 2D ridge detector and Hessian matrix was used for detection of the stripe centers in the image. Detection is done in the following steps: (a) input image is converted from RGB to HSV color space and V channel is extracted, (b) the vesselness map is computed from V channel as described in [5], (c) stripe centers are extracted by finding the local maxima in the vesselness map, and (d) sub-pixel position of stripe centerline is obtained as described in [22]. Selected scales are such to cover the whole range of stripe widths. Parameters used for vesselness map computation are selected as described in [14].

### 2.2.2 Color Identification

For the color identification and stripe labeling we propose clustering in the RGB color space using only pixel data of found stripe centers. A similar approach based on clustering was used in [4], where pixel data was fitted to prototype color lines inside of the RGB cube.

Clustering approach mitigates a difficult color matching problem between the projector and camera that is usually solved by the color calibration procedure. Clustering will only be effective if colors can be separated into required number of distinct groups. A convenient fact that boosts the effectiveness of the clustering approach is *a priori* knowledge of the number of colors that are projected. This information about the number of colors enables the use of the Gaussian mixture model (GMM) and Expectation-maximization (EM) algorithm to separate all observed colors. For a detailed explanation of GMM and EM and their use for clustering see [1].

Our proposed structured light pattern uses six colors (Table 1), so GMM will use six mixtures that are labeled  $\{R, G, B, Y, C, M\}$ .

Used six colors are vertices of the RGB cube so we set the initial means of the components to the middle of color axes in the RGB cube. The covariance matrices of each component are chosen in the shape of an elongated ellipsoid that is oriented in the direction of the specific color axis. After applying the EM algorithm we assign each stripe center to a corresponding color cluster based on the resulting probability that the center belongs to the cluster, thus we effectively label all found stripes in the image.

### 2.2.3 Stripe Decoding

After color labels are assigned to the detected stripes they must be matched to the projected pattern. A common solution is comparison of each scanline (one row in the captured image) to the projected De Bruijn sequence of colors using a dynamic programming (DP) approach, so image is processed row-wise. We retain this approach, but instead of applying DP directly to pixel intensity values we propose to use color labels.

Since we have assigned a color label to each detected stripe center we need to find the optimal local alignment between two strings of color labels, one given by Eq. (1) that represents projected pattern and another that is extracted from the image for every scanline. The Smith-Waterman algorithm [21], a DP algorithm often used in bio-informatics for local alignment of two nucleotide string sequences, is ideal for this purpose. The Smith-Waterman algorithm optimizes a local alignment of two strings taking in account possible changes between two strings.

Let  $t$  be the projected color sequence and let  $s$  be a string of assigned color labels. When aligning strings  $s$  and  $t$  there four possible cases: (a) a character in  $s$  may match a character in  $t$  (action *match*); (b) a character in  $s$  may mismatch a character in  $t$  (action *mismatch*); (c) an empty character '-' must be inserted in  $t$  (action *delete*); and (d) an empty character '-' must be inserted in  $s$  (action *insert*). First step

of DP algorithm is creation of the scoring matrix  $M$ :

$$M(i, j) = \begin{cases} 0, & \text{if } i = 0 \text{ and } j = 0 \\ \max \begin{cases} M(i-1, j-1) + w(s_i, t_j) \\ M(i, j-1) + w_{\text{insert}} \\ M(i-1, j) + w_{\text{delete}} \end{cases}, & \text{else} \\ 0 \end{cases} \quad (3)$$

where  $0 \leq i \leq \text{length}(s)$ ,  $0 \leq j \leq \text{length}(t)$ ,  $s_i$  is  $i$ th character in  $s$ ,  $t_j$  is  $j$ th character in  $t$ , and  $w$  is a scoring function defined as

$$w(s_i, t_j) = \begin{cases} w_{\text{match}}, & s_i = t_j \\ w_{\text{mismatch}}, & \text{else} \end{cases}. \quad (4)$$

Values  $w_{\text{mismatch}}$ ,  $w_{\text{delete}}$ ,  $w_{\text{insert}}$ , and  $w_{\text{match}}$  are predefined scoring costs.

DP proceeds with initialization of the matrix  $M$  as follows: In the case of character alignment movement in the scoring matrix  $M$  is diagonal. If  $s_i$  and  $t_j$  are equal a reward  $w_{\text{match}}$  is added. If  $s_i$  and  $t_j$  are different the current score is penalized with  $w_{\text{mismatch}}$ . If a stripe has not been detected, e.g. due to occlusion or shadow, a string  $s$  will be missing a color label. In that case the movement in the scoring matrix  $M$  is horizontal and is penalized with  $w_{\text{insert}}$ . If an extra stripe is detected, e.g. due to an error in the stripe detection process, a string  $s$  will have an additional color label. In that case the movement in the scoring matrix  $M$  is vertical and is penalized with  $w_{\text{delete}}$ . During this initialization of matrix  $M$ , beside assigning the alignment score to each cell, the direction of the performed movement (diagonal, horizontal, vertical) must be recorded for each cell. Recorded directions are required for *backtracking* the alignment. *Backtracking* begins from the maximal value in the matrix  $M$  (the best alignment score), and by taking movement direction backward, the optimal alignment of two given strings is found.

For example, let the projected color sequence be  $t = \text{GMRYBYBGRBGYBCRBCYBMGRMGY}$  and let the detected color sequence be  $s = \text{YBGRRCRBCRBGMG}$ . The scoring values are set to  $w_{\text{mismatch}} = -3$ ,  $w_{\text{delete}} = -5$ ,  $w_{\text{insert}} = -2$ , and  $w_{\text{match}} = 3$ . Two strings will then be aligned as follows:

```
t  GMRCMYGMRYBYBGRBGYBCRBCYBMGRMGY
s  -----YBGR-----CRBCRBGMG-----
```

Notice the gap between two string groups in  $s$ . The gap is where stripes were not detected due to occlusion. Also, notice how the strings were correctly aligned even though one character in the string  $s$  was incorrectly labeled as R instead of Y. This example demonstrates the strength of DP in finding the optimal alignment between the projected pattern and the detected color sequence even when some labels in the sequence were incorrectly assigned.

## 2.3. Multiple-Shot Reconstruction

Multiple-shot reconstruction is based on the phase-shifting paradigm. Both phase and color subsequence that are required for the 3D reconstruction are derived using temporal analysis for each pixel independently, therefore there are no assumptions about spatial characteristics of the imaged object. Proposed reconstruction is done in three steps: (a) channel equalization; (b) estimation of wrapped phase; and (c) phase unwrapping based on temporal color decoding.

### 2.3.1 Temporal Channel Equalization

For every pixel and for every channel let  $I_{\min}$  and  $I_{\max}$  be the observed minimal and maximal values for the recorded  $N \cdot N_p$  frames. Then each of the RGB channels is separately equalized via affine mapping

$$I_{\text{eq}} = 255 \frac{I(i) - I_{\min}}{I_{\max} - I_{\min}}, \quad i = 0, \dots, i_{\max}, \quad (5)$$

where  $i_{\max} = N \cdot N_p - 1$ . This temporal channel equalization effectively removes albedo for static pixels and eliminates effects of unknown ambient light and camera channel gains. It is supported by the property we have imposed on the De Bruijn sequence: in every subsequence of length  $N$  both minimum and maximum in all three channels must be achieved at least once; observed values  $I_{\min}$  and  $I_{\max}$  are these extrema.

### 2.3.2 Estimation of Wrapped Phase

For every pixel the wrapped phase is estimated from the V channel of the HSV color space as

$$\psi_w = \text{atan2} \left( \sum_i -V_{\text{eq}}(i) \sin(\varphi_i), \sum_i V_{\text{eq}}(i) \cos(\varphi_i) \right), \quad (6)$$

where the index  $i$  goes over all  $N \cdot N_p$  recorded frames. Note that the V channel must be computed from the equalized channel data given by (5) as  $V_{\text{eq}}(i) = \max(B_{\text{eq}}(i), G_{\text{eq}}(i), R_{\text{eq}}(i))$ , otherwise the result will be affected by the object albedo.

### 2.3.3 Phase Unwrapping

To unwrap the wrapped phase  $\psi_w$  we must identify a De Bruijn subsequence of length  $N = 3$  for every pixel independently. Combining Eqs. (2) and (6) for any pixel yields a three parameter temporal intensity model of the V channel,

$$V(i) = V_0 + V_1 \cos(\psi_w + \varphi_i), \quad (7)$$

where  $V_0$ ,  $V_1$ , and  $\psi_w$  are model parameters and where  $i$ th recorded frame has phase  $\varphi_i = \frac{2\pi}{N_p} i$ . Temporal intensity

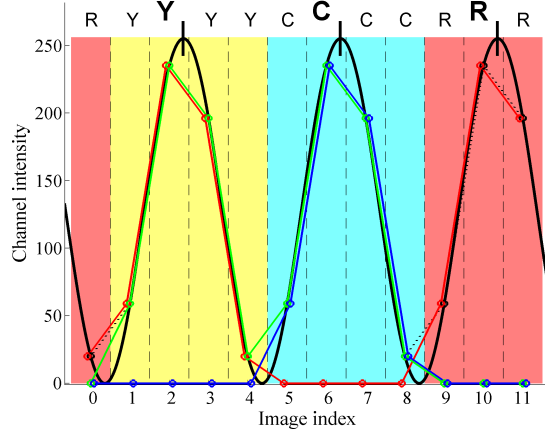


Figure 2: Example of temporal color decoding on simulated data. Phase offset is given by the position of RCY substring in the used De Bruijn sequence.

given by Eq. (7) achieves its maxima at  $\psi_w + \varphi_i = 2k\pi$ . To identify a length  $N$  subsequence we must use  $N$  samples at phases  $\varphi_i = -\psi_w + 2k\pi, k \in \mathbb{Z}$ , which correspond to real values  $i_R = -\frac{\psi_w}{2\pi}N_p + kN_p, k \in \mathbb{Z}$ , of index  $i$  that fall into  $[0, i_{\max}]$  interval. By taking two frames with indices  $i$  closest to the three chosen values of  $i_R$  and by linearly interpolating between them we obtain three RGB triplets that correspond to the three maxima in V channel. Color is then identified by computing the hue from RGB values and selecting the closest color. Time-reversed color sequence then identifies the phase offset needed for unwrapping, e.g. if a time-reversed triplet starts at position  $m$  in the sequence of Eq. (1), then the phase offset is  $2\pi(m + 1)$  and unwrapped phase is  $-\psi_w + 2\pi(m + 1)$ . An example of temporal color decoding for  $N_p = 4$  and  $P = 11$  is shown in Fig. 2. A function given by Eq. (7) fitted to the V channel is shown as a solid black line. Three sample points where colors are decoded are marked with vertical black lines. Decoded color substring is YCR, and its time-reversed value RCY is found at the position  $m = 12$  so phase offset is  $2\pi(12 + 1)$ .

### 3. Results and Discussion

All presented experiments were performed using Acer X1260 DLP projector with the resolution  $1024 \times 768$  that was paired with PointGrey DragonFly2 DR-HICOL camera with the resolution  $1024 \times 768$  that was fitted with Fujinon HF9-HA1B lens. The system operates at 30 frames per second. Camera-projector base distance was  $\sim 11$  cm and working distance was  $\sim 1$  m, giving a calibration volume of  $\sim 500 \times 500 \times 500$  mm. Geometric calibration of the camera (projector) was performed using the procedure described in [27, 16].

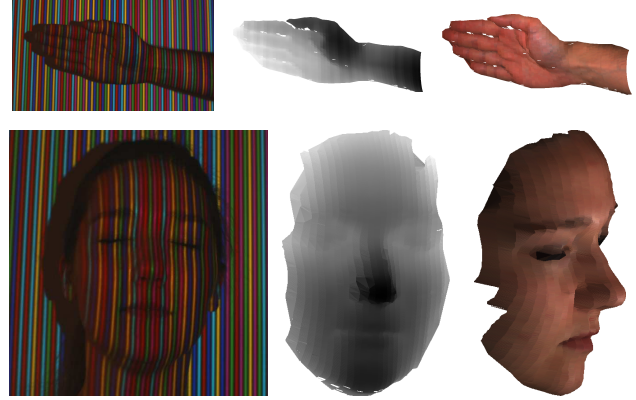


Figure 3: Examples of the 3D reconstruction for one-shot method for two different objects. From left to right: input image, depth, and textured reconstruction.

#### 3.1. One-Shot Reconstruction

Examples of the one-shot reconstruction method are shown in Fig. 3, a female and a hand.

We have proposed to use clustering in the RGB color space for the color identification. Choice of the RGB color space was motivated by the fact that the hue component of the HSV color space cannot be used for robust color identification alone due to significant color changes caused by object's albedo. This effect is demonstrated in Fig. 4 for a human face. Note the significant hue color shifts and unclear color boundaries in the hue histogram, especially for the red and magenta colors that have merged together. Also note that color shifts are visible in the RGB scattergram, but a clear separation between six colors is preserved.

Initial positioning of the color clusters may affect the result of GMM clustering, e.g. if the number of data points of one color is significantly different than other colors, it may be necessary to adjust initial parameters. This problem may be mitigated by ensuring proper sequence properties: colors must be equally spatially distributed in any area of the image. Additionally, initial centers may be adjusted by using the equalization parameters  $I_{\min}$  and  $I_{\max}$  of the multiple-shot method as they provide sufficient information about color shifting and are available for every static pixel.

Most of the state-of-the-art approaches [3, 20, 25, 29] that use DP for the De Bruijn pattern decoding directly input pixel intensities or some function of intensities, e.g. gradient, into DP. This usually means a few to several parameters that control DP must be chosen, and those are almost certainly image and scene dependent as pixel intensities are used directly. For example, the work by Zhang et al. [25] proposes using a combined gradient of RGB channels to detect edges, and then extracts pixel gradients at found edges. This triplet is input into DP for which a cost function with several parameters must be defined, making DP step com-



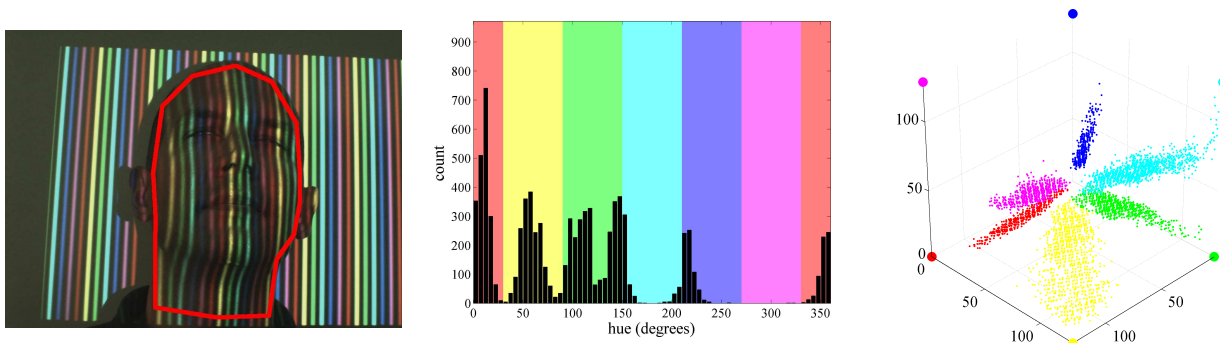


Figure 4: Left: Manually segmented human face showing the ROI where stripe centers were extracted. Middle: A histogram of hue channel for selected ROI. Right: A scattergram of RGB data for selected ROI.

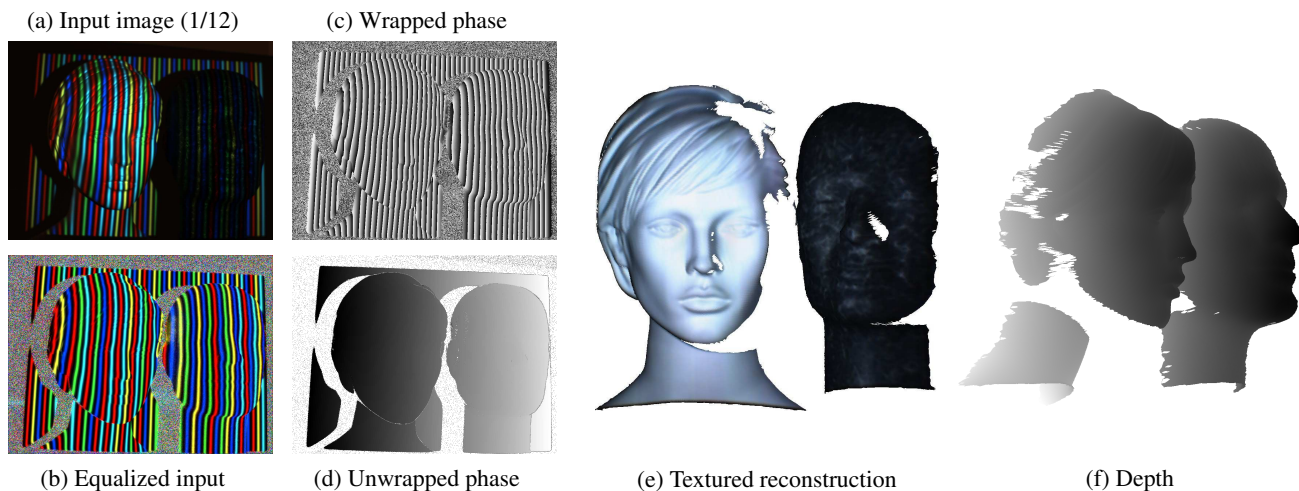


Figure 5: Two mannequins, a very bright female and a very dark male. Note a major difference in albedo.

plex and difficult to adjust. Our approach uses color labels directly and thus simplifies DP as: (a) parameters required for DP are no longer affected by pixel intensity values and may be chosen independently of the image; and (b) computational time required for DP is shorter as the cost is defined over color labels. We have implemented the edge-based method described in [25] and our approach is about three times faster when computing DP scoring matrix. Besides, choosing the optimal parameters for the edge detection step and for DP is not straightforward, thus influencing end results (details omitted due to the shortage of space).

Another advantage over referenced De Bruijn based one-shot methods is the use of a multi-scale 2D ridge detector: it outputs subpixel precise ridge locations and as its parameters are automatically adjusted. A single-shot De Bruijn pattern with color ridges is advantageous to a pattern with color edges as the ridge location is not affected by projector defocus and is not shifted due to blooming effect. An edge between two colors will always shift toward the color with the darker intensity while the ridge widens but remains

centered, so the blooming effect widens the ridges, but does not displace them. A disadvantage of the used ridge detector is increased computational time, however, computational complexity per pixel is fixed.

The advantage of computing stripe ridges is also recognized and demonstrated in [12]. However, a method of [12] requires an explicit color calibration before each scanning which is then tied to the particular subject color (albedo). We note that during the course of this work the authors of [12] were generous to share their code with us. While testing their method we have noticed more than six different parameters that need to be adjusted for stripe detection only (e.g. stripe width, filter order, thresholds for 2nd derivative and saturation intervals etc.), which is not an easy task to do. Moreover, when the lightning conditions and/or subject change it is necessary to readjust at least several parameters. Perhaps most importantly, the method [12] and similar are aimed at dynamic object reconstruction without the possibility to adapt in the case of static scenarios to provide dense reconstruction.

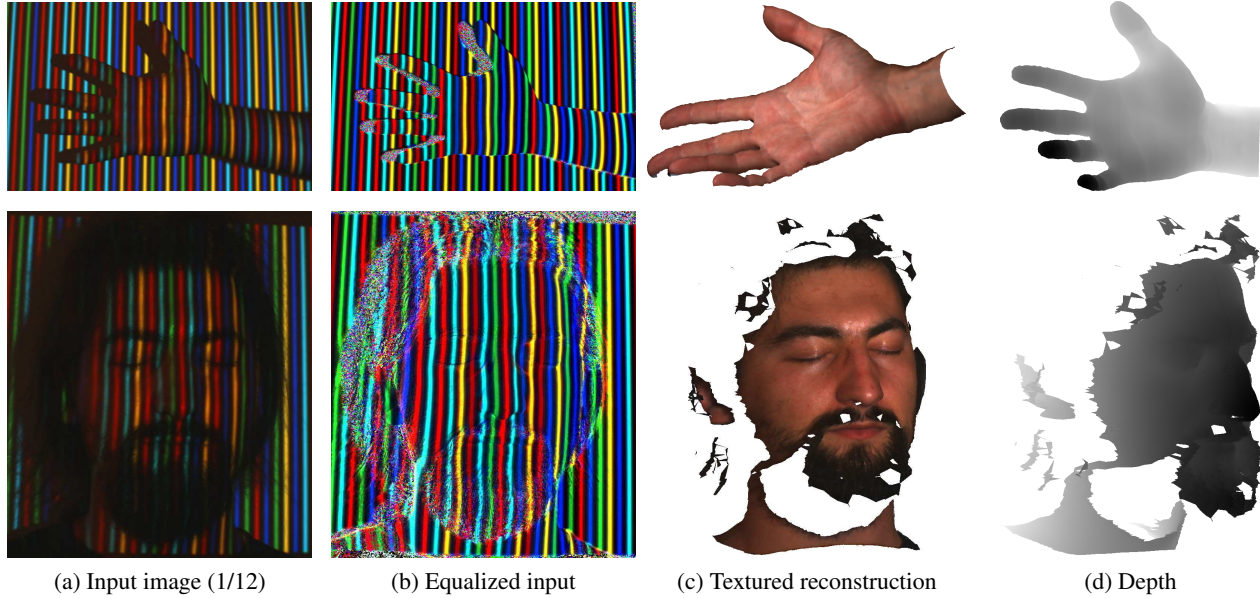


Figure 6: Reconstruction using multiple-shot method of two different objects, a hand and a face. Note that the beard was partially reconstructed.

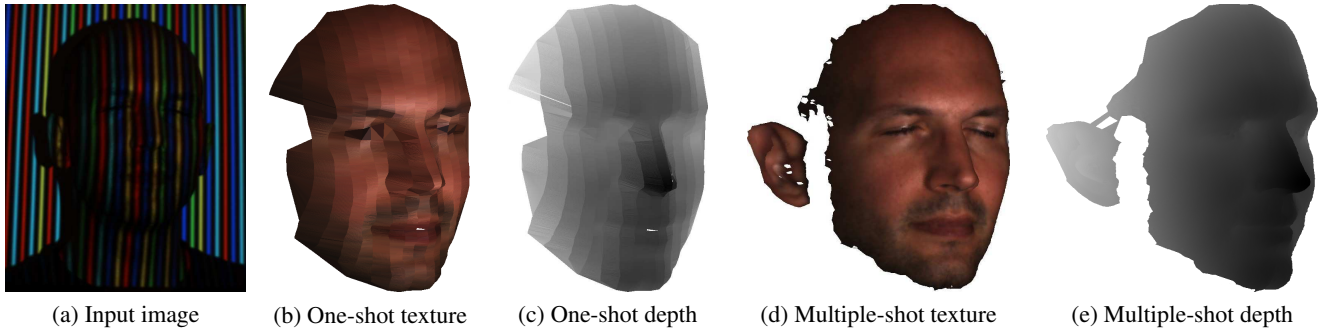


Figure 7: Comparison of one- and multiple-shot reconstruction of a human face. Note the lower spatial resolution in the horizontal direction for the one-shot method.

### 3.2. Multiple-Shot Reconstruction

Multiple-shot reconstruction may be used for all pixels that are static. For those pixels the proposed method uses the self-correcting property of the proposed De Bruijn sequence that makes color calibration and albedo estimation unnecessary. This is demonstrated in Fig. 5: note how the equalized image has significantly clearer colors than the raw image; dynamic range a dark male mannequin was about 40/255 and the colors were still successfully extracted. Examples of multiple-shot reconstruction are shown in Figs. 5 and 6. To assess the accuracy of the proposed multiple-shot reconstruction we have compared for several recorded objects wrapped phases difference between ours and values obtained using a traditional phase shifting, considered in this context as the ground truth. Phase differences converted

to projector pixels are  $-0.08 \pm 0.20$  px for two mannequins in Fig. 5 (absolutely static objects),  $-0.25 \pm 0.59$  px for the face and  $-0.15 \pm 0.30$  px for the hand in Fig. 6 (quasi static objects). An example histogram of phase errors is shown in Fig. 9.

A side-by-side comparison of one- and multiple-shot reconstruction is shown in Fig. 7.

### 3.3. Combined Static and Dynamic Scene

The crucial step for combining one and multiple-shot approach within a same scene is pixel classification either as static or dynamic. Classical frame-based motion detection methods that operate on pixel values directly may not be used due to the structured light pattern that constantly changes. Therefore, we propose to detect motion by tracking the unwrapped phase of each pixel. If the unwrapped

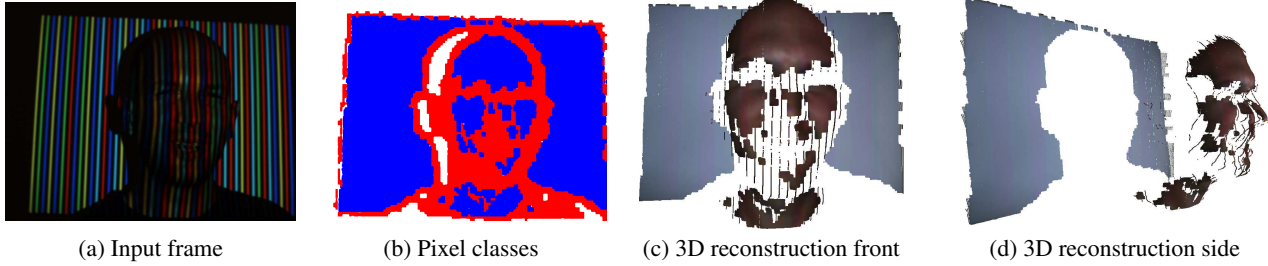


Figure 8: Reconstruction combining both methods. In (b) static pixels are blue, moving pixels are red, and background pixels are white. For this experiment we deliberately used a coarser stripe pattern to clearly show the difference in spatial density.

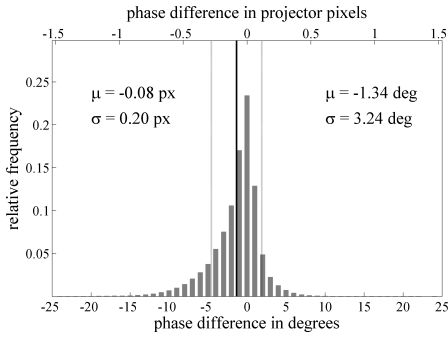


Figure 9: Histogram of phase errors for two mannequins from Fig. 5.

Object	A [mm]	B [mm]	C [mm]	D [mm]
Face 1	0.506	0.616	0.387	0.394
Face 2	0.508	0.650	0.386	0.407
Face 3	0.496	0.628	0.370	0.366
Hand	0.468	0.653	0.328	0.327
Mannequins	0.572	0.702	0.522	0.519

Table 2: Mean distances between point clouds after ICP registration, in mm. A is one-shot registered to PS+GC, B is PS+GC to one-shot, C is multiple-shot to PS+GC, and D is PS+GC to multiple-shot.

phase of a pixel does not change more than some predetermined threshold in a specified time interval the pixel is static; otherwise the pixel is dynamic. The combined 3D reconstruction for a dynamic scene of a person talking in front of a white wall is shown in Fig. 8. Note the dense reconstruction for static parts of the scene: the white wall and person’s forehead and cheeks.

### 3.4. ICP Registration

To qualitatively evaluate our method we have compared it against structured light scanning using a well established phase shifting and Gray code (PS+GC). We note that PS+GC is generally regarded as the method of choice for static object scenarios, due to its robustness, high accuracy,

and resolution of 3D reconstruction [18]. We have performed 3D registration of point clouds obtained by PS+GC and by our method using ICP [17]. The results are shown in Table 2. As can be seen the differences are negligible for the most purposes.

## 4. Conclusion

The main challenge in colored structured light is the correct matching of colors perceived by the camera to the colors projected by the projector, a task that is made even more difficult when ambient illumination and object albedo are introduced. We have proposed two techniques to solve this problem: using clustering techniques in one-shot method and using the proposed color self-correcting De Bruijn sequence in multiple-shot method.

We have demonstrated the proposed method usability on various objects with challenging surfaces and with different albedos. Our method successfully identifies colors without the need for color calibration of the camera-projector pair. One-shot reconstruction gives a lower spatial resolution than multiple-shot as expected, but only in horizontal direction. This shortcoming is mitigated by a multiple-shot method that enables dense 3D surface reconstruction for static objects. The proposed method successfully applies within the same scene (camera frame) one- and multiple-shot approach to dynamic and static objects respectively.

## Acknowledgment

This work has been supported in parts by the Croatian Science Foundation’s funding of the project IP-11-2013-3717 and by the Croatian Ministry of Science, Education and Sport research “Sensor technologies for the ambient assisted living”. We also acknowledge the part of support and the inspiration gained through the bilateral Croatian-Chinese project “Single shoot structured light 3D reconstruction”. Finally, we are grateful to Joaquim Salvi, Professor of Computer Vision at University of Girona, Spain, who suggested improvements to the content and presentation of this paper.



## References

- [1] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [2] D. Caspi, D. Caspi, N. Kiryati, N. Kiryati, J. Shamir, and J. Shamir. Range Imaging with Adaptive Color Structured Light. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(5):470–480, 1998.
- [3] C.-S. Chen, Y.-P. Hung, C.-C. Chiang, and J.-L. Wu. Range data acquisition using color structured lighting and stereo vision. *Image and Vision Computing*, 15(6):445–456, 1997.
- [4] P. Fechteler and P. Eisert. Adaptive Color Classification for Structured Light Systems. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 1–7, 2008.
- [5] A. Frangi, W. Niessen, K. Vincken, and M. Viergever. Multi-scale Vessel Enhancement Filtering. In W. Wells, A. Colchester, and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention*, volume 1496 of *Lecture Notes in Computer Science*, pages 130–137. Springer Berlin Heidelberg, 1998.
- [6] S. S. Gorthi and K. R. Lolla. A New Approach for Simple and Rapid Shape Measurement of Objects with Surface Discontinuities. In *Optical Metrology*, pages 184–194. International Society for Optics and Photonics, Proceedings of SPIE, 2005.
- [7] N. Henderson, R. King, and R. H. Middleton. An Application of Gaussian Mixtures: Colour Segmenting for the Four Legged League Using HSI Colour Space. In U. Visser, F. Ribeiro, T. Ohashi, and F. Dellaert, editors, *RoboCup 2007: Robot Soccer World Cup XI*, volume 5001 of *Lecture Notes in Computer Science*, pages 254–261. Springer Berlin Heidelberg, 2008.
- [8] Z.-K. Huang and D.-H. Liu. Segmentation of color image using em algorithm in hsv color space. In *International Conference on Information Acquisition*, pages 316–319, 2007.
- [9] R. Juang and A. Majumder. Photometric self-calibration of a projector-camera system. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- [10] F. MacWilliams and N. Sloane. Pseudo-random sequences and arrays. *Proceedings of the IEEE*, 64(12):1715–1729, Dec 1976.
- [11] S. Narasimhan, S. Koppal, and S. Yamazaki. Temporal dithering of illumination for fast active vision. In D. Forsyth, P. Torr, and A. Zisserman, editors, *Computer Vision ECCV 2008*, volume 5305 of *Lecture Notes in Computer Science*, pages 830–844. Springer Berlin Heidelberg, 2008.
- [12] J. Pagès, J. Salvi, C. Collewet, and J. Forest. Optimised de bruijn patterns for one-shot shape acquisition. *Image and Vision Computing*, 23(8):707–720, 2005.
- [13] J. Pan, P. S. Huang, and F.-P. Chiang. Color phase-shifting technique for three-dimensional shape measurement. *Optical Engineering*, 45(1):013602–013602–9, 2006.
- [14] T. Petković and S. Lončarić. Using x-ray imaging model to improve guidewire detection. In *IEEE 10th International Conference on Signal Processing*, pages 805–808, Oct 2010.
- [15] T. Pribanić, H. Džapo, and J. Salvi. Efficient and low-cost 3D structured light system based on a modified number-theoretic approach. *EURASIP Journal on Advances in Signal Processing*, 2010(1):474389, 2010.
- [16] T. Pribanić, S. Mrvoš, and J. Salvi. Efficient multiple phase shift patterns for dense 3d acquisition in structured light scanning. *Image and Vision Computing*, 28(8):1255–1266, 2010.
- [17] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on*, pages 145–152, 2001.
- [18] J. Salvi, S. Fernandez, T. Pribanić, and X. Lladó. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.
- [19] G. Sansoni, A. Patrioli, and F. Docchio. OPL-3D: A novel, portable optical digitizer for fast acquisition of free-form surfaces. *Review of Scientific Instruments*, 74(8):2593–2603, 2003.
- [20] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, 2002.
- [21] T. Smith and M. Waterman. Identification of common molecular subsequences. *Journal of Molecular Biology*, 147(1):195–197, 1981.
- [22] C. Steger. An unbiased detector of curvilinear structures. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 20(2):113–125, 1998.
- [23] Y. Wu, X. Yang, and K. L. Chan. Unsupervised color image segmentation based on Gaussian mixture model. In *Proceedings of the Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing and Fourth Pacific Rim Conference on Multimedia*, volume 1, pages 541–544, 2003.
- [24] C. Wust and D. Capson. Surface profile measurement using color fringe projection. *Machine Vision and Applications*, 4(3):193–203, 1991.
- [25] L. Zhang, B. Curless, and S. M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *3D Data Processing Visualization and Transmission*, pages 24–37, 2002.
- [26] S. Zhang and P. S. Huang. High-resolution, real-time three-dimensional shape measurement. *Optical Engineering*, 45(12):123601–1–123601–8, 2006.
- [27] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.
- [28] Z. Zhang. Review of single-shot 3d shape measurement by phase calculation-based fringe projection techniques. *Optics and Lasers in Engineering*, 50(8):1097 – 1106, 2012.
- [29] Y. Zhou, D. Zhao, Y. Yu, J. Yuan, and S. Du. Adaptive color calibration based one-shot structured light system. *Sensors*, 12(8):10947–10963, 2012.