# Single-Shot Dense 3D Reconstruction using Self-Equalizing De Bruijn Sequence

T. Petković, *Member, IEEE,* T. Pribanić, *Member, IEEE,* and M. Đonlić, *Student Member, IEEE*

***Abstract*—Single-shot dense 3D reconstruction using colored structured light is a difficult problem due to the undesired effects of ambient lighting, object albedo, non-equal channel gains, and channel cross-talk. We propose a novel single-shot dense 3D reconstruction using colored structured light. Our method combines the self-equalizing De Bruijn sequence, scale-space analysis, and band-pass complex Hilbert filters to achieve insensitivity to ambient lighting, object albedo, and non-equal channel gains. The proposed method reconstructs about 85% of points compared to time-multiplexing structured light strategies and the decoding error in the recovered projector coordinate is less than one projector pixel for about 90% of reconstructed points.**

***Index Terms*—3D scanning, structured light, phase estimation, De Bruijn sequence, Hilbert transform filter**

## I. Introduction

THE three-dimensional surface profile of an object may be measured using various non-invasive optical techniques [1] including structured light. Structured light (SL) utilizes a projector to actively illuminate an object of interest using a designed pattern. The surface profile of an object is then recovered by examining the deformation of the projected pattern. There are two major classes of patterns used in SL: *single-shot* patterns that enable reconstruction from a single image and *multi-shot* patterns that require more than one image. Multi-shot patterns require the object to be stationary and are mostly used when an accurate high-resolution surface profile is needed, for example in industrial quality control. Single-shot patterns are of particular interest as they enable 3D reconstructions of moving objects making them applicable in various fields such as body structure analysis, biometrics, and human-computer interaction, to name but a few. Among single-shot patterns of particular interest are those using *color* as the majority of available projectors and cameras use color. Such a pattern would enable construction of simple 3D scanners comprised of ubiquitous low-cost components: a projector, a camera, and a personal computer.

The largest group of single-shot SL methods using color extends the well-known grayscale sinusoidal fringe patterns [2]. We divide these depending on the approach to color processing: (1) available color channels are analyzed in a device colorspace (e.g. RGB); and (2) available color channels are analyzed in a perceptual colorspace (e.g. HSV). In the first approach fringes having different phases and/or frequencies are embedded in each of device color channels: Wust and Capson [3] project same-frequency fringes shifted in phase and use classical three-step decoding; and Zhang et al. [4] project multi-frequency fringes and extract both 3D surface and rough texture as a mixture of object albedo and ambient lighting. In the second approach most often a sinusoidal fringe is placed in the value channel and fringes are colored to enable absolute pattern decoding: Chen et al. [5] color the fringes by changing color exactly at the maximum of each fringe thus making each fringe two-colored; Fernandez and Salvi [6] color the fringes by changing color exactly at the minima between fringes thus making each fringe single-colored; and Su [7] colors several adjacent stripes using the same color and changes their maximum intensity to enable absolute decoding. In addition, we mention grid patterns pioneered by Salvi et al. [8] and extended by Kawasaki et al. [9], which provide additional robustness compared to sinusoidal fringes as both projector coordinates are encoded. However, dense reconstruction using grid patterns requires interpolation as information about surface changes between projected grid lines is lost. For complete codification of single-shot SL patterns using color, including M-arry codes and frequency multiplexing, we refer the reader to a review by Salvi et al. [10].

Major problems when using color in SL are: ambient lighting, object albedo, non-equal color channel gains, and cross-talk between color channels. The color imaging model of Caspi et al. [11] is sufficient to capture and mitigate all of the aforementioned effects; however, for the model to be applicable all of its parameters must be known. This is difficult in practice as only some parameters are determined by the projector and camera, while the remaining crucial parameters depend on object albedo and ambient lighting. Most of the aforementioned works require a color calibration to get all parameters of the model. Specifically, to obtain the parameters describing object albedo and ambient lighting, partial color calibration must be performed prior to every imaging to achieve robustness; this fact severely limits the practical applicability of the mentioned SL methods. There are approaches that try to avoid this by inferring the missing model parameters from the data under the assumption of a constant object albedo [12], [13]. However, such an assumption is severely limiting in practice. Once the parameters of the color imaging model are known the color is most often decoded by analyzing the hue channel of HSV representation [5], [9], [12], [14]–[16]; a minority of works decodes colors in RGB

Manuscript received March 8, 2016; revised June 21, 2016; accepted July 31, 2016. This work was supported by the Croatian Science Foundation under Project IP-11-2013-3717. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jing-Ming Guo.

The authors are with the Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb 10000, Croatia (e-mail: tomislav.petkovic.jr@fer.hr).

Digital Object Identifier 10.1109/TIP.2016.2603231

colorspace [17].

In this paper we propose a novel method for *dense* 3D reconstruction from a single image using colored SL. The proposed method is an extension of our previous works [15], [16], where we described how to obtain dense multi-shot reconstruction of static objects and sparse single-shot reconstructions of moving objects. In [15] we have introduced the *self-equalizing* De Bruijn sequence and in [16] we have shown how to apply it to obtain sparse 3D reconstruction from video using a hybrid approach. However, the previous works are lacking in the sense that *dense* 3D reconstruction from a single image was not achieved, which is the main contribution of this paper. To summarize, besides using the *self-equalizing* De Bruijn sequence which enables removal of ambient lighting, cancellation of object albedo, and equalization of color channel gains, the extensions compared to [15], [16] are: (1) the method enables *dense* surface profile reconstruction from a single image; (2) the optimal parameters of the scale-space Hessian matrix analysis are automatically determined from the input data; (3) the improved spatial-equalization scheme allows better recovery of V-channel data; (4) the band-pass complex Hilbert filter allows accurate spatial phase estimation; and (5) the robust phase-unwrapping based on direct $L_1$ decoding of De Bruijn windows enables imaging of pure color objects.

Combining the scale-space Hessian matrix analysis and the band-pass (BP) complex Hilbert filter is a key improvement which enables *dense* wrapped phase estimation and *dense* reconstruction of a surface profile. Such an approach to phase estimation is superior to the existing discrete Fourier transform (DFT) based methods: phase estimation is localized, which provides insensitivity to discontinuities, and discretization effects such as spectral leakage are minimized. Also, the new optimal color decoding scheme using $L_1$ minimization combined with spatial-equalization allows robust color decoding, De Bruijn window identification, and phase unwrapping. Noting that the De Bruijn sequence of order $n$ is also that of all orders higher than $n$, we show how the proposed $L_1$ decoding may be applied to a single color channel thus enabling reconstruction of pure color objects.

This paper is structured as follows: Section II gives a brief overview of spatial phase estimation methods. Section III presents the proposed novel dense single-shot surface reconstruction method. Section IV describes the system calibration procedure. Experimental results and discussion are presented in Section V. We conclude in Section VI.

## II. A Brief Overview of Phase Estimation Methods

Some of the most common SL patterns are sinusoidal fringe patterns where the phase encodes the information required for 3D reconstruction. Such patterns differ mostly in the type of information encoded in phase. For example: the phase may be directly proportional to the object's height or depth [18]–[20]; the phase may be related to the object's height above some reference plane [21]–[23]; the phase may directly encode the vertical or horizontal projector planes [6]. Regardless of which information is encoded in phase, the essential step of 3D reconstruction is a phase estimation. The phase estimation in SL may be classified as either temporal or spatial; the former estimate the phase from three or more images while the latter estimate the phase from a single image. We focus on methods applicable to a single image which may be further classified as follows: (1) estimation using signal decomposition into in-phase and quadrature components [24]–[29], (2) least-squares estimation [20], [30], [31], and (3) transform domain methods using Fourier transform [21]–[23], [32], windowed Fourier transform [6], [33], [34], and wavelet transform [35]. There also exist other approaches, e.g. using a phase-locked loop [36].

The signal of interest for SL methods utilizing phase may be represented as

$$I(x,y) = A(x,y) + B(x,y)\cos\big(\phi(x,y)\big), \tag{1}$$

where $(x,y)$ are coordinates in the camera image, $I(x,y)$ is the intensity observed by the camera, $A(x,y)$ and $B(x,y)$ are jointly describing effects of projector and camera transfer functions, of object albedo, and of ambient lighting, and where the phase $\phi$ carries the information of interest. Therefore, the problem of spatial phase estimation in SL is the extraction of the phase $\phi(x,y)$ from the signal of Eq. (1).

Important concepts in phase estimation are the *analytic signal* and the *instantaneous phase* [37]. Let $I(x)$ be a real valued signal. Then $I_a(x) \stackrel{\text{DEF}}{=} I(x) + j\mathcal{H}\big[I(x)\big]$, where $\mathcal{H}[\cdot]$ is the Hilbert transform, is the analytic representation of a real valued signal $I(x)$, and $\phi_i(x) \stackrel{\text{DEF}}{=} \text{Arg}\big(I_a(x)\big)$ is the instantaneous phase of $I(x)$. If $I(x)$ is a signal of Eq. (1) with a fixed $y$ coordinate then the phase $\phi$ is equal to the instantaneous phase $\phi_i$ only if $A$ and $B$ are constants. However, in SL both $A$ and $B$ are spatially varying and this unwanted variation must be taken into consideration.

### A. Spatial Domain Methods

The early work describing spatial phase estimation method was by Toyooka and Iwasa [24], who consider the phase of the form $\phi = \omega_0 x_p + \varphi$, where $\omega_0$ is the carrier frequency that must be known, $x_p$ is the vertical projector coordinate (column), and $\varphi$ encodes the object height variation. They split the image into horizontal scanlines so every scanline is a 1D signal described by Eq. (1) with a fixed row $y$. The scanline is then decomposed into in-phase and quadrature components. In [24] the authors also show that frequency filtering sufficiently removes the influence of terms $A$ and $B$. The method essentially estimates the phase by computing the instantaneous phase of the BP filtered analytic signal of the observed real valued intensity signal for each scanline of the recorded image. This and similar methods [25]–[29] are collectively called spatial phase detection (SPD).

Another approach to phase estimation is the least squares phase estimation. The signal of Eq. (1) is spatially sampled using sufficient number of samples to utilize the least-squares approach. For most SL applications the phase is of the form $\phi = \omega_0 x_p + \varphi$. If the carrier frequency $\omega_0$ is known then linear least-squares are used; alternatively, if $\omega_0$ is unknown then non-linear least squares must be used. Least-squares

approaches to phase estimation are commonly used in digital phase-shift shadow Moiré topography [20], [31].

### B. Transform Domain Methods

Transform domain phase estimation methods process the signal in the transform domain using a Fourier or wavelet basis. In the transform domain there usually exists a clear separation between the signal of interest $\phi$ and the unwanted terms $A$ and $B$, a fact which simplifies phase estimation. Additionally, if the phase has the form $\phi = \omega_0 x_p + \varphi$ then the carrier frequency $\omega_0$ is easily detectable in the transform domain, a significant advantage compared to SPD.

The first transform domain method proposed by Takeda and Mutoh [21] is based on the Fourier transform (FT). The image is decomposed into horizontal scanlines on which a FT is applied. The phase is estimated by selecting a band of positive frequencies around the carrier followed by the inverse FT of the selected frequency band only. The inverse FT yields a complex signal whose phase is the desired phase estimate. In the framework of analytic signals such a procedure is equivalent to computing the instantaneous phase of a BP filtered analytic signal corresponding to the observed real valued signal for each scanline separately. The phase estimation step of [21] is thus in principle equivalent to SPD of [24] with an additional benefit of simpler carrier frequency detection. A review of Fourier transform profilometry (FTP) methods is given in [32].

Other transform domain methods use the same principle as FTP while trying to eliminate its shortcomings by extending scanline processing to full 2D and by replacing the infinite support Fourier basis with a finite support basis. The extension to 2D processing enables better suppression of noise and background. Using a finite support basis enables processing of images with discontinuities and images where the fringe field does not cover the entire field-of-view (FOV). Techniques of interest include the windowed Fourier transform (WFT) [33], [34] and the wavelet transform (WT) [38].

In our opinion a major issue with the current literature regarding transform domain phase estimation in SL is the absence of discussion regarding discretization, i.e. all formal derivations are almost always presented in the continuous domain leaving the details of actual discrete domain implementations open to interpretation. Although the effects of discretization are well known in the signal processing community [39], [40], the specifics that should be discussed in the context of SL are the implicit periodicity of the signal and the spectral leakage. We discuss these topics in Section V-C5.

## III. Single-Shot Dense 3D Reconstruction

We propose a novel single-shot dense surface reconstruction method using colored SL. The method utilizes the *self-equalizing* De Bruijn sequence which is combined with the sinusoidal fringe pattern so each fringe has one color only. The pattern's phase directly encodes either column or row projector coordinate.

The processing of the recorded image is as follows: (1) a scale-space ridge detection is performed; (2) based on the scale-space analysis, the input image is decomposed into optimally placed scanlines; (3) each scanline is spatially equalized to remove effects of ambient lightning, object albedo, and non-equal channel gains; (4) from the equalized data the wrapped phase is estimated using a BP complex Hilbert filter; (5) from the equalized data the De Bruijn window is decoded using $L_1$ optimization; and (6) the phase is unwrapped and the surface is reconstructed using triangulation. We first describe how to construct the pattern; the following subsections describe each of the processing steps.

### A. Self-Equalizing De Bruijn Sequence

We construct a *self-equalizing* De Bruijn sequence as described [15]. Briefly, a $k$-ary De Bruijn sequence of order $n$ is a cyclic sequence of length $L = k^n$ over an alphabet of $k$ symbols in which every subsequence of length $n$, called a window, appears exactly once in the cycle [41]. For RGB colorspace the symbols of the De Bruijn alphabet are all on/off channel combinations which may be represented as 3-digit binary numbers; these are black (000, K), red (100, R), green (010, G), blue (001, B), yellow (110, Y), magenta (101, M), cyan (011, C), and white (111, W). Pure black and white symbols are prohibited; therefore, the alphabet over which a De Bruijn sequence is constructed has $k = 6$ symbols. The *self-equalizing* constraint requires all channels to span the full available dynamic range in every De Bruijn window, or, more precisely [15]: (a) the $j$th binary digit should attain value 0 for at least one symbol in every De Bruijn window and (b) the $j$th binary digit should attain value 1 for at least one symbol in every De Bruijn window; where $j \in \{1, 2, 3\}$. The full De Bruijn sequence is constructed as an Eulerian cycle in the pruned De Bruijn graph where edges belonging to windows that do not satisfy the *self-equalizing* constraint are removed.

The pattern is constructed as follows: let $c_0 c_1 \cdots c_{L-1}$ be the *self-equalizing* De Bruijn sequence of length $L$ and let $c_l$ be a 3-digit binary number. The intensity of the red channel is

$$I_R(x_p, y_p) = I_{\mathrm{M}} \sum_{l=0}^{L-1} \mathbf{1}_R(c_l) \Big( \frac{1}{2} - \frac{1}{2} \cos\big(2\pi \frac{x_p - lP}{P}\big) \Big) \cdot \cdot \mathrm{rect}\Big( \frac{\tilde{x}_p}{P} - \frac{1}{2} \Big), \quad (2)$$

where $x_p$ is the projector column (projector width should be greater or equal to $L \cdot P$), $y_p$ is the projector row, $P$ is the spatial period of the fringe, $I_{\mathrm{M}}$ is the maximum intensity, and $l$ is the stripe index. The variable $\tilde{x}_p = x_p - lP \mod L \cdot P$ ensures that rect is periodically extended. The function $\mathbf{1}_R$ is the indicator function of the red channel, i.e. $\mathbf{1}_R(c_l) = 1$ if $c_l$ is 1 in the red channel and 0 otherwise. Intensities for the blue and green channels are obtained by replacing the indicator function $\mathbf{1}_R(\cdot)$ with $\mathbf{1}_G(\cdot)$ or $\mathbf{1}_B(\cdot)$.

Comparing Eq. (2) to Eq. (1) yields $\phi = 2\pi \frac{x_p - lP}{P}$, i.e. the phase directly encodes the column $x_p$ and has a carrier frequency $\frac{2\pi}{P}$. The V channel of Eq. (2), defined as $V = \max\{R, G, B\}$, is a continuous sinusoid whose phase encodes the projector column. All channels are coherent in phase; this property significantly aids in both spatial analysis and phase estimation.

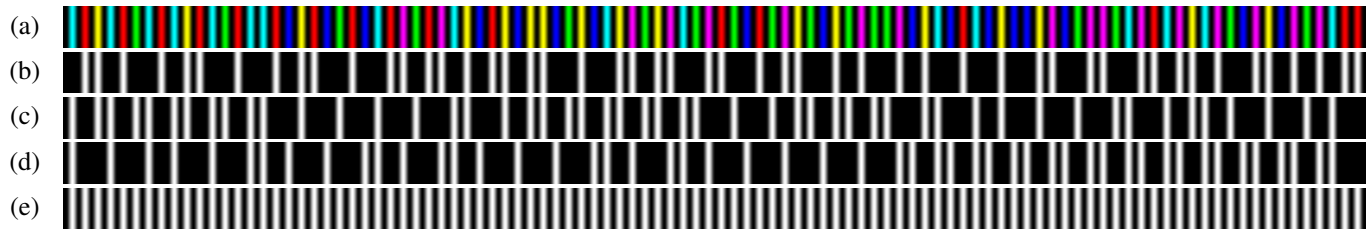


Fig. 1. Structured light pattern for $n = 3$ and $k = 6$ that contains $L = 102$ color stripes: (a) shows color pattern; (b) shows red channel; (c) shows green channel; (d) shows blue channel; and (e) shows V channel of HSV representation.

A construction example: for a De Bruijn sequence of order $n = 3$ the full digraph has exactly $k^n = 6^3 = 216$ edges; after pruning using the *self-equalizing* constraint exactly 102 edges remain. One possible sequence is CRYCRGCRCYRCGRCCRBYRBGRBCRMGRMCYBRYBYYBGY-BCYMGYMCGMRGBRGMYGBYGMGGMBYCBRCBYBBYMBGMM-GCMRCMYCMGBMYBMGMCRR for which the pattern is shown in Fig. 1. Note the sequence is of order $n = 3$ and that each channel also satisfies the unique window property with orders $n_R = 16$, $n_G = 27$, and $n_B = 35$; this enables stripe decoding even for pure-colored objects (see Section III-E).

### B. Scale-Space Ridge Detection

Scale-space ridge detection is a well established tool that is both robust and insensitive to noise. We combine the approaches of Steger [42] and Frangi et al. [43]. The ridge detection is performed on a single channel image which is obtained as a sum of all color channels. It is comprised of the following steps: (a) the scales of scale-space analysis are determined; (b) the Hessian matrix is computed at all selected scales; (c) the modified vesselness measure is calculated from the eigenvalues of the Hessian matrix; and (d) the ridges are extracted as the local extrema of the vesselness measure. The results of the ridge detection are then used to optimally decompose the image into scanlines.

To determine the scales at which the scale-space analysis is performed we sum all channels and divide the resulting image into rows. The power spectral density is estimated by computing the periodogram [44] of all rows. The first significant peak in the periodogram is a good estimate of the carrier frequency; we denote the angular frequency of the first peak by $\omega_o$. Similarly to [42], it can be shown the best scale for the detection of cosine shape ridges is $\sigma \approx 0.45w$, where $w = \frac{2\pi}{\omega_0}$ is the period of the cosine. To allow for pattern deformation and to improve robustness we select several additional scales around the optimal scale.

Once scales are selected the Hessian matrix and its eigenvalues $\lambda_{1,\sigma}$ and $\lambda_{2,\sigma}$, $|\lambda_{1,\sigma}| > |\lambda_{2,\sigma}|$, are computed at each scale $\sigma$. The modified vesselness, an extension of the vesselness defined by Frangi et al. [43], is computed as:

$$\mathcal{V} = \operatorname{sign}(-\lambda_{1,\sigma}) \max_{\sigma_{\min}<\sigma<\sigma_{\max}} \left\{ \exp\left(-\alpha \left|\frac{\lambda_{2,\sigma}}{\lambda_{1,\sigma}}\right|\right) \cdot \left(1 - \exp\left(-\beta(\lambda_{1,\sigma}^2 + \lambda_{2,\sigma}^2)\right)\right) \right\}, \quad (3)$$

where $\alpha, \beta \in \mathbb{R}^+$ are constants. The modified vesselness simultaneously extracts both bright color stripes for $\mathcal{V} > 0$ and dark black slits for $\mathcal{V} < 0$. The constants $\alpha$ and $\beta$ are automatically chosen to optimize the ridge detection with the threshold of $\pm 0.5$ as:

$$\alpha = \log(2) \quad \text{and} \quad \beta = 2\pi \log(2) \frac{\sigma^6}{h^2 w^2} \exp\left(\frac{w^2}{\sigma^2}\right), \quad (4)$$

where $h$ is the expected ridge height, $w$ is the expected ridge width, and $\sigma$ is the scale at which the vesselness map is computed. We set $h$ equal to the inter-quartile range of the input amplitudes and $w = \frac{2\pi}{\omega_0}$ is determined from the periodogram. An example input image and the corresponding vesselness are shown in Fig. 2 (a) and (b).

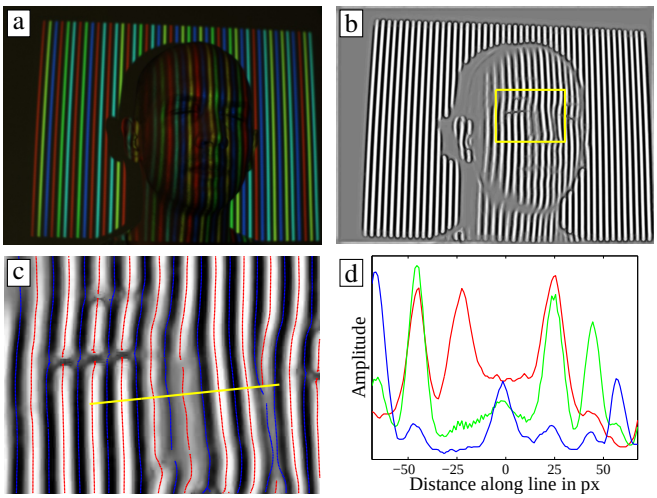


Fig. 2. Example of scale-space analysis and scanline placement. (a) shows input image; (b) shows vesselness of (a) with ROI marked; (c) shows enlarged ROI of (b) with marked subpixel ridge positions and one scanline; and (d) shows amplitudes of color channels along the scanline of (c).

Positions of color stripes and dark slits are extracted as the local extrema of Eq. (3). The scale-space analysis also provides orientation of all ridges: the eigenvector associated with the eigenvalue $\lambda_1$ is orthogonal to the local ridge, and may be used to obtain precise subpixel positions of ridges [42].

Local orientation and precise positioning of the ridges are essential to the fringe pattern analysis: the input image is decomposed into scanlines that are orthogonal to the local ridge direction and that are spanning as many stripes as is the order of the De Bruijn sequence. Let $(x_0, y_0)$ be the position of a ridge center and let $\vec{v} = \hat{\imath}v_x + \hat{\jmath}v_y$ be a unit vector orthogonal to the ridge. Decomposing the input into scanlines means replacing $(x, y)$ of Eq. (1) with $x = x_0 + v_x t$ and $y = y_0 + v_y t$, where $t \in \mathbb{Z}$ is the sample index along the

scanline. The V channel of Eq. (1) is then transformed into equidistantly sampled scanlines:

$$I[t] = A[t] + B[t]\cos\big(\phi[t]\big). \tag{5}$$

The equidistant sampling is the necessary condition for the phase estimation. It allows the use of a single BP complex Hilbert filter for phase extraction. The input is in color so the sampling is performed for each color channel separately. Also, the input exists on a discrete spatial grid so values at sample points must be interpolated; we use bilinear interpolation. An example of scanline placement is shown in Fig. 2 (c) for the pattern which has $w \approx 24\,\text{px}$ and for a De Bruijn sequence of order $n = 3$; therefore to contain a full De Bruijn window each scanline must have about $n \cdot w \approx 3 \cdot 24 = 72$ samples.

Once all scanlines are extracted they are further analyzed and only valid scanlines are retained. In [15] we have introduced *spatial invariance conditions* which must be fulfilled for the De Bruijn window to be decodeable. In brief, the conditions are: (a) the distance $d$ between adjacent color stripes must be about $\sigma/0.45$, where $\sigma$ is the optimal detection scale selected by Eq. (3); (b) adjacent stripes must be parallel; and (c) exactly one black slit exists between every two adjacent color stripes. The conditions of [15] limit the spatial configuration only and omit limiting the color channel amplitudes. We extend these conditions to include the signal amplitudes. Let $d_R$ be the dynamic range of amplitudes in the red channel, $d_G$ in the green, and $d_B$ in the blue. Then, in addition to conditions (a-c), for each valid De Bruijn window we require: (d) the maximum of $d_R$, $d_G$, and $d_B$ must be larger than some threshold $T_d$; (e) the minimum of $d_R$, $d_G$, and $d_B$ must be larger than some threshold $T_e$; and (f) the maximum absolute difference of amplitudes at dark slits must be below some threshold $T_f$. The conditions (d,e) enable unbiased phase estimation and color decoding. The threshold $T_d$ ensures at least one channel has sufficient dynamic range for phase estimation; $T_d$ should be set sufficiently large, e.g. around half of the camera dynamic range. The threshold $T_e$ requires that all channels contain sufficient information for color decoding; $T_e$ should be set quite small, e.g. around the expected value of channel cross-talk. The condition (f) ensures the object albedo is constant within one De Bruijn window under the assumption of constant ambient light (at black slits only the ambient lighting is reflected). The threshold $T_f$ should be small.

### C. Spatial Equalization

Once the scanlines are extracted they must be equalized using the *self-equalization* property. The aim is to remove the influence of terms $A[t]$ and $B[t]$ of Eq. (5). We perform two separate equalizations: the first prepares the data for the phase estimation and the second prepares the data for the color decoding.

*1) Equalization for Phase Estimation:* The analysis of Section III-B yields precise positions of both dark slits and color stripes. Correlating the positions with the minima and maxima of the pattern given by Eq. (2) reveals that the projected intensity $(0, 0, 0)$ corresponds to the dark slits and

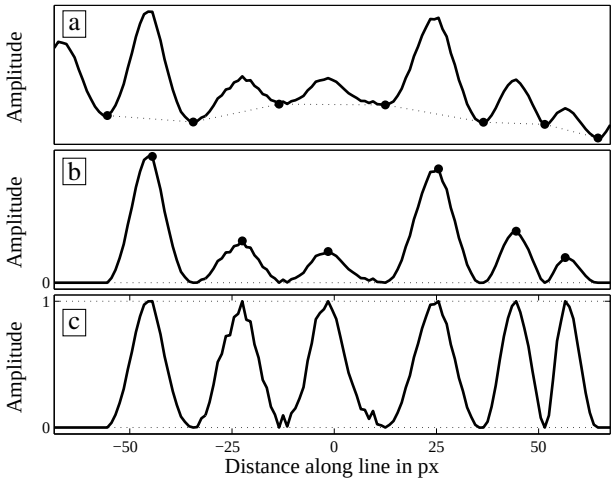


Fig. 3. Equalization for phase estimation for the scanline of Fig. 2: (a) shows the scanline data with minimum values marked; (b) shows the scanline with maximum values marked after trends are removed; and (c) shows equalized scanline after linear scaling of every fringe to $[0, 1]$ interval.

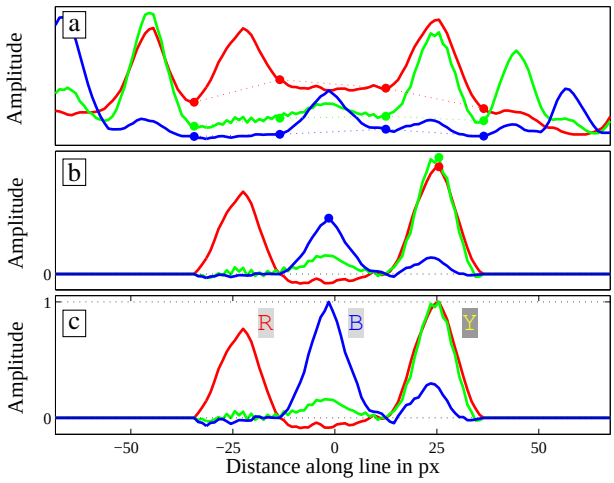


Fig. 4. Equalization for color decoding for the scanline of Fig. 2 for one De Bruijn window of length $n = 3$: (a) shows the scanline RGB data with marked positions of dark slits; (b) shows the scanline RGB data with one maximum value for each channel marked after trends are removed; and (c) shows equalized RGB scanline after linear scaling of maximum values to 1 for each channel.

the projected intensity $I_{\text{M}}$ corresponds to the color stripes maxima. All color channels are spatially coherent in phase: the phase of the V channel and the phase of the sum of all channels are equal. Hence, we equalize the sum of all channels and not the V channel due to improved signal-to-noise ratio. The spatial coherence of the color channels prevents the cross-talk from affecting the phase estimation.

Equalization is done on each scanline by segments. A segment is a part of the scanline between two adjacent minima which has exactly one maximum between the two minima. Equalization transforms amplitudes of both minima to 0 and of one maximum to 1 and it preserves the structure of amplitudes in-between. Therefore, it effectively removes ambient lighting and equalizes the channel gains and object albedo. This is achieved by subtracting a line interpolated between two adjacent amplitude minima, which removes the

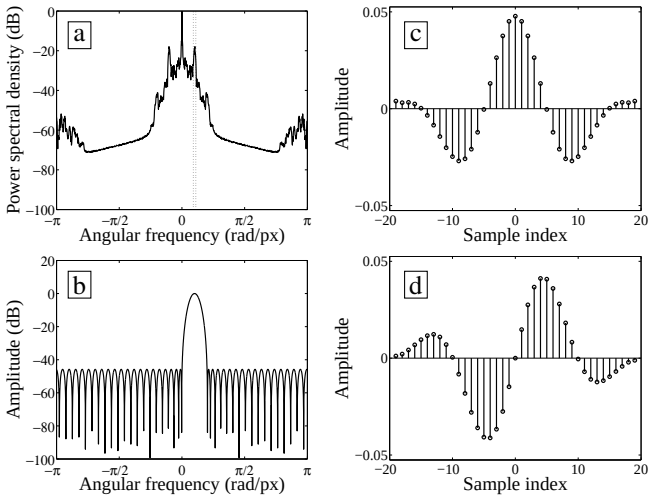


Fig. 5. Design of BP complex Hilbert filter of order 38 for central frequency $\omega_0 = 0.316$ rad/px and half-bandwidth $B = 0.03068$ rad/px: (a) power spectral density showing the band of interest; (b) amplitude-frequency characteristic of the designed filter; (c) real part of filter impulse response; and (d) imaginary part of filter impulse response.

trend, followed by linearly scaling the maximum values to 1, which preserves the structure of the signal. Note the line segments are overlapping only at minima (dark slits) which are always transformed to 0. The procedure therefore does not affect the continuity of the signal. Equalization steps are shown in Fig. 3. Overall, the equalization for phase estimation ensures the signal amplitudes are limited to $[0, 1]$ interval with exactly one maximum equal to 1 per fringe.

*2) Equalization for Color Decoding:* Equalization for color decoding must be different than the equalization for phase estimation as color information must be retained. Overall, the procedure is similar but is performed for each channel over a group of segments instead of on each segment separately.

A segment is a part of the scanline between two adjacent minima that has exactly one maximum between the two minima, where minima and maxima are taken from the sum of all channels. A group of segments is comprised of $n$ adjacent segments that form one De Bruijn window. Equalization for color decoding first removes ambient lighting by removing the trend for each color channel separately; after trend removal the values at the positions of the minima of sum of all channels are 0. Then, to equalize for channel gains and object albedo, the maximum of each color channel on the group of $n$ segments is linearly scaled to 1. The scaling on the group of $n$ segments is the key difference that ensures color information is retained. Equalization steps are shown in Fig. 4.

### D. Phase Estimation

After equalization the amplitudes of each scanline may be described by Eq. (5) with the variation in $A$ and $B$ removed. The phase may be estimated by computing the analytic signal directly by using the discrete Hilbert transform [45]. However, to improve the robustness to equalization errors we limit the signal to the frequency band of interest. Selecting the frequency band of interest and estimating the corresponding analytic signal may be achieved in one step by using the BP complex Hilbert filter.

Let the $h[t] : \mathbb{Z} \to \mathbb{C}$ be the impulse response of the BP complex Hilbert filter. Then the BP analytic signal $I_a[t] : \mathbb{Z} \to \mathbb{C}$ of the equalized signal $I[t] : \mathbb{Z} \to \mathbb{R}$ is the convolution

$$I_a[t] = I[t] * h[t] = \sum\nolimits_i I[i]h[t-i]. \tag{6}$$

The wrapped instantaneous phase of the complex signal $I_a[t]$ is

$$\phi_i[t] = \operatorname{atan2}\big(\operatorname{Im}(I_a[t]), \operatorname{Re}(I_a[t])\big). \tag{7}$$

As both $h[t]$ and $I[t]$ have a finite number of samples only a certain central part of the wrapped phase estimate given by Eq. (7) is free of border effects. This effectively places an upper limit on the order of the filter $N$.

The desired BP complex Hilbert filter has the central angular frequency $\omega_0$ and bandwidth $2B$, which are extracted from the power spectral density as described in Section III-B. The filter is constructed using the approach of Reilly et al. [39]: (1) a real FIR filter with the pass-band of $[\omega_0 - B - \omega_S/4, \omega_0 + B - \omega_S/4]$ is designed, where $\omega_S$ is the sampling frequency; (2) the impulse response of the FIR filter is then modulated with a complex exponential of frequency $\omega_S/4$ recovering the frequency shift of $\omega_S/4$. The FIR filter of step (1) is designed using the Parks-McClellan algorithm [40]. The desired BP complex Hilbert filter is therefore an equiripple filter that has the best-possible suppression in all frequencies outside of the pass-band for the chosen filter order $N$.

The filter order $N$ must be even; this ensures the filter group delay is zero and eliminates spatial shifts in the estimated phase. The maximum viable filter order $N$ that eliminates undesired border effects may be computed by counting the number of samples between maxima in all De Bruijn windows for all scanlines. Let the minimum distance between two fringe maxima for all scanlines be $M$. Then the filter order $N$ is the first even number larger than or equal to $M \cdot 2\lfloor \frac{n-1}{2} \rfloor$, where $n \geq 3$ is the order of the De Bruijn sequence. The boundary effects are eliminated as for every scanline the wrapped phase estimate given by Eq. (7) is needed only for the segment demarcated by the two minima around the central maximum. As there are exactly $n$ maxima in one window and as each scanline has at least $M \cdot n$ samples then the largest filter order for odd $n$ is $M \cdot (n-1)$ and for even $n$ is $M \cdot (n-2)$.

In Fig. 5 we show a typical power spectral density together with amplitude-frequency characteristic of the designed filter. The real part of the impulse response is essentially the impulse response of a BP filter and the imaginary part resembles a convolution of the impulse response of a BP filter and the impulse response of the Hilbert transformer.

### E. De Bruijn Window Decoding

We propose to decode color in the device colorspace using an $L_1$ metric to select the best possible matching De Bruijn window. Let $w_{n,k} = c_{0,k}\, c_{1,k} \cdots c_{n-1,k}$, $0 \leq k < L$, be the string of colors for one of $L$ De Bruijn windows in the SL pattern. Then $c_{i,k} = (r_{i,k}, g_{i,k}, b_{i,k})$, $r_{i,k}, g_{i,k}, b_{i,k} \in \{0, 1\}$, is the color at position $i$, $0 \leq i < n$, for the De Bruijn

window at position $k$. Let $W_n = C_0\,C_1\,\cdots\,C_{n-1}$ be the string of observed RGB triplets after color equalization at $n$ adjacent bright ridges (color stripes). Then $C_i = (R_i, G_i, B_i)$, $R_i, G_i, B_i \leq 1$, is a triplet of equalized channel intensities at position $i$, $0 \leq i < n$, in the observed De Bruijn window. All colors of $W_n$ are then simultaneously decoded by selecting the window $w_{n,k}$ which is an $L_1$ minimizer of

$$\arg\min_k \frac{1}{n}\left|\left|W_n - w_{n,k}\right|\right|_1 = \arg\min_k \frac{1}{n}\sum_{i=0}^{n-1} |R_i - r_{i,k}| + |G_i - g_{i,k}| + |B_i - b_{i,k}|. \quad (8)$$

The proposed decoding scheme is resistant to a certain amount of noise or to color mismatches caused by cross-talk as all colors in every window are decoded simultaneously.

Every De Bruijn sequence of order $n$ is also a De Bruijn sequence of orders $k$, $n \leq k < L$, where $L$ is length of the sequence. Therefore, the decoding of Eq. (8) should not use the minimum order $n$; it must consider all orders between the minimum order $n$ and some maximum possible order $n_{\max}$, which is specific for each observed window. Spatial extension of Eq. (8) up to some maximum order $n_{\max}$ may be implemented by constructing connections between adjacent stripes. Recall the result of Section III-B are ridge centers and their normals. By following the normal direction to adjacent color stripes on both sides and extending this to include all stripes that satisfy the *spatial invariance conditions* of Section III-B, the maximum window extension is achieved and $n_{\max}$ is determined. Then, Eq. (8) becomes

$$\arg\min_{k,m} \frac{1}{m}\left|\left|W_m - w_{m,k}\right|\right|_1, \quad (9)$$

where $0 \leq k < L$ and $n \leq m \leq n_{\max}$. Eq. (9) is further extended to include vertical connectivity of color stripes by following directions orthogonal to ridge normal. Therefore, the final color decoding scheme uses robust $L_1$ minimization to decode a bundle of horizontally extended and vertically connected scanlines.

### F. Phase Unwrapping

Phase unwrapping requires combining the wrapped phase estimate $\phi$ of Eq. (7) with the position $k$ of the De Bruijn window of Eq. (9). Ideally, as every fringe covers one wrapped phase interval of $[-\pi, \pi]$, the unwrapped phase along one scanline is

$$\Phi[t] = \phi[t] + 2k\pi \quad (10)$$

for indices $t$ selecting the samples of the $k$th fringe (the samples between two dark slits around the $k$th color stripe). The naive unwrapping of Eq. (10) does not take into consideration the discontinuity of the $\mathrm{atan2}$ function. Ideally, the estimate of Eq. (7) should be exactly 0 at the positions of color stripes and should be $\pm\pi$ at the positions of dark slits. However, the wrapped phase around dark slits is not well-defined and becomes unstable under minimal noise. Therefore, applying Eq. (10) in practice always produces $2\pi$ unwrapping errors.

To mitigate the unwrapping errors we note the discontinuity of $\mathrm{atan2}$ may be shifted as needed, e.g. if the wrapped phase $\phi$ has discontinuities at the unwrapped phase $\Phi = (2k+1)\pi$ then the shifted phase $\phi' = (\phi + x + \pi \mod 2\pi) - \pi$ has discontinuities shifted by $x$ to $\Phi = (2k+1)\pi + x$. We rewrite Eq. (10) to use positions of both the central color stripe and of first adjacent left and right dark slits:

$$\Phi[t] = \begin{cases} (\phi[t] \mod 2\pi) + (2k-2)\pi, & \text{left slit} \\ \phi[t] + 2k\pi, & \text{central stripe} \\ (\phi[t] \mod 2\pi) + 2k\pi, & \text{right slit} \end{cases}, \quad (11)$$

where the left slit means the sample index $t$ is closest to the position of the left dark slit etc. Unwrapping of Eq. (11) effectively shifts the discontinuity of the $\mathrm{atan2}$ outside of the spatial interval which is being unwrapped, thus avoiding $2\pi$ unwrapping errors.

### G. 3D Reconstruction

The unwrapped phase of Eq. (11) directly encodes the projector's column $x_p$ (vertical plane). It is retrieved using $x_p = P\Phi/(2\pi)$, where $P$ is the spatial period of the sinusoidal fringe in projector pixels. The retrieved projector column $x_p$ and the coordinates $(x, y)$ in the camera image enable recovery of the 3D position which is the intersection between the camera ray and the projector plane.

## IV. Calibration

A SL imaging system is comprised of a camera and a projector in a fixed imaging geometry for which calibration must be performed. When using color the calibration is often split into geometric and colorimetric calibration. The goal of geometric calibration is to retrieve the intrinsic and extrinsic parameters for the camera and the projector. We have used a modified method of Moreno and Taubin [46] where the Gray code SL pattern was replaced by the combination of Gray code and phase-shifted fringe.

The goal of colorimetric calibration is to retrieve the color transfer functions for both camera and projector together with linear channel mixing coefficients using the color imaging model introduced by Caspi et al. [11]. A full colorimetric calibration may be performed using the procedure described by Juang and Majumder [47] which enables recovery of the spatially dependent channel transfer functions. For SL applications the spatial dependency is always dropped and the channel transfer functions are assumed to be the same for all camera/projector pixels; this is acceptable if the spatial variation is sufficiently small. In most of the previous works colorimetric calibration is used to compensate for channel cross-talk [6], [11], [12], [17], [48], [49]; only a few works extend the calibration to linearize channel transfer functions [3], [48].

Ideally, for the proposed SL pattern both projector and camera should have pure linear channel transfer functions with no cross-talk between channels. More relaxed and realistic requirements are non-decreasing channel transfer functions and minimal cross-talk. If these are fulfilled then, due to the properties of the *self-equalizing* De Bruijn sequence [15], unwanted effects of ambient lighting, object albedo, and unequal

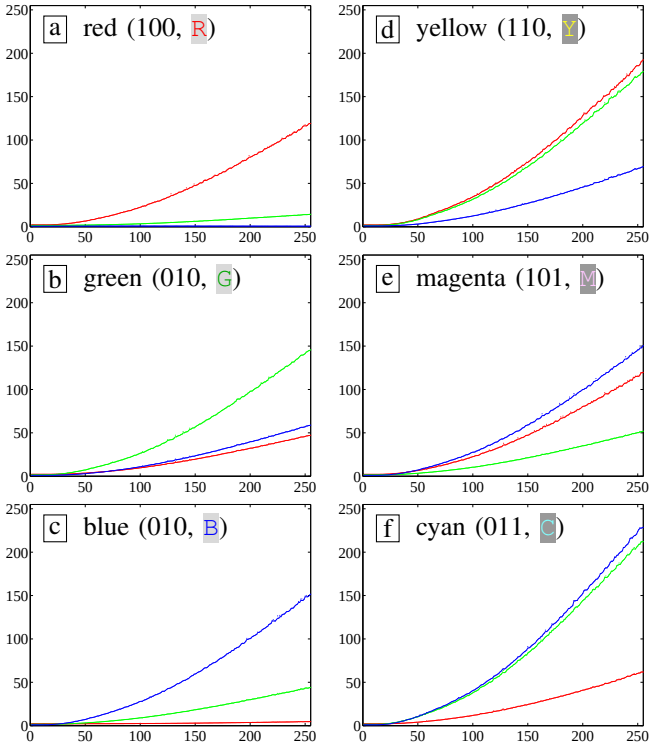


Fig. 6. Color channel transfer functions for Acer X1260 DLP projector for primary (a-c) and secondary (d-f) colors: (a) red transfer function; (b) green transfer function; (c) blue transfer function; (d) yellow transfer function; (e) magenta transfer function; and (f) cyan transfer function.

color channel gains are effectively removed making colorimetric calibration unnecessary. However, while testing several consumer-grade projectors we noticed that one-time projector intensity calibration is required to correct for inconsistent color intensities.

The key requirement for the *self-equalizing* property to be effective is that channel intensities sent to the projector are not modified by the projector, e.g. the intensity output for the red channel at the maximum intensity must be the same for all symbols R, Y, or M. This simple requirement is not fulfilled by many consumer-grade projectors, especially if they feature color enhancement technology. most DLP projectors using Texas Instruments' BrilliantColor™ [50] consistently output more light[1] in mixed (Y, M, C) than in pure colors (R, G, B), a fact which may be attributed to the color wheel of a DLP projector having additional white, cyan, and yellow segments. For SL applications such unwanted intensity variations must be compensated for. Hence, the proposed method requires one-time colorimetric intensity calibration of the projector only. It may be performed using any camera. The goal of projector intensity calibration is to make the maximum channel intensities equal regardless of which color is projected. To do this six color transfer functions for R, G, B, Y, M, and C are recorded using a white projection board and a color camera. Using the recorded curves one then finds projector intensities that minimize the maximum absolute intensity differences at stripe maxima, thus calibrating the projector.

[1]Increased light output produces a brighter image which is more appealing to potential buyers.

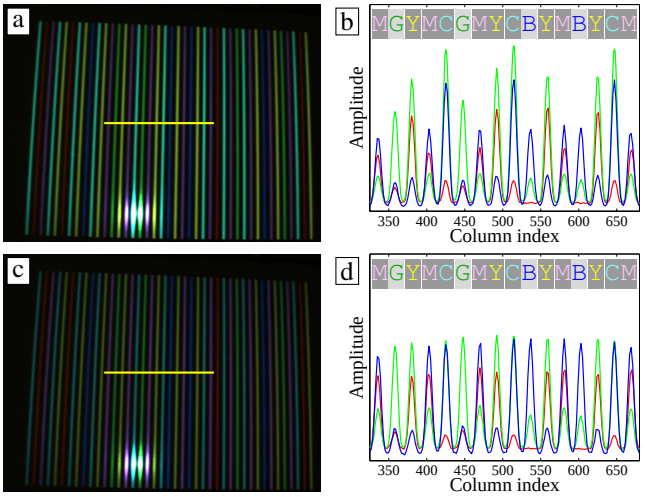


Fig. 7. Uncalibrated and calibrated SL pattern for Acer X1260 projected on a whiteboard and recorded using PointGrey DragonFly2 DR-HICOL camera: (a) uncalibrated SL pattern with a horizontal scanline marked in yellow; (b) scanline of (a); (c) calibrated SL pattern with a horizontal scanline marked in yellow; and (d) scanline of (c).

Examples of color transfer curves for an Acer X1260 DLP projector are shown in Fig. 6: note 62% difference in maximum intensities of the red channel comparing Y to R, 45% in the green channel comparing C to G, and 51% in the blue channel comparing C to B. Uncalibrated and calibrated SL patterns for this Acer X1260 are shown in Fig. 7; note the severe variation in maximum channel intensities in the uncalibrated pattern is almost non-existent in the calibrated pattern.

Note that we do not calibrate for cross-talk. After the intensity calibration the cross-talk from blue to green channel is largest at about 30%, and from green to both red and blue is about 23%. This does not present a problem for the proposed method. The proposed phase estimation is completely insensitive to cross-talk due to spatial coherence. The proposed $L_1$ color decoding is minimally affected by the present amount of cross-talk due to the high breakdown point of the $L_1$ estimators.

## V. Results and Discussion

All presented experiments were performed using an Acer X1260 DLP projector, a PointGrey DragonFly2 DR-HICOL camera, and a Fujinon HF9-HA1B lens. The resolution of both projector and camera is $1024 \times 768$. The camera-projector base distance is $\sim 11$ cm. The calibration volume is $\sim 0.5 \times 0.5 \times 0.5$ m. The system is able to reconstruct the depth in a range from $0.5$ m up to $5$ m in the intersection of the camera and projector view frustums. Besides the color sequence of length 102 described in Section III-A we also used a shorter sequence of length 42: CRMCYRCYBCYMGYMCGMYCBYMBYCMRCMGC-MYBMYGMCRY.

### A. Planar Surfaces

We have reconstructed static planar surfaces, a whiteboard and three colored checkerboards, in various positions within

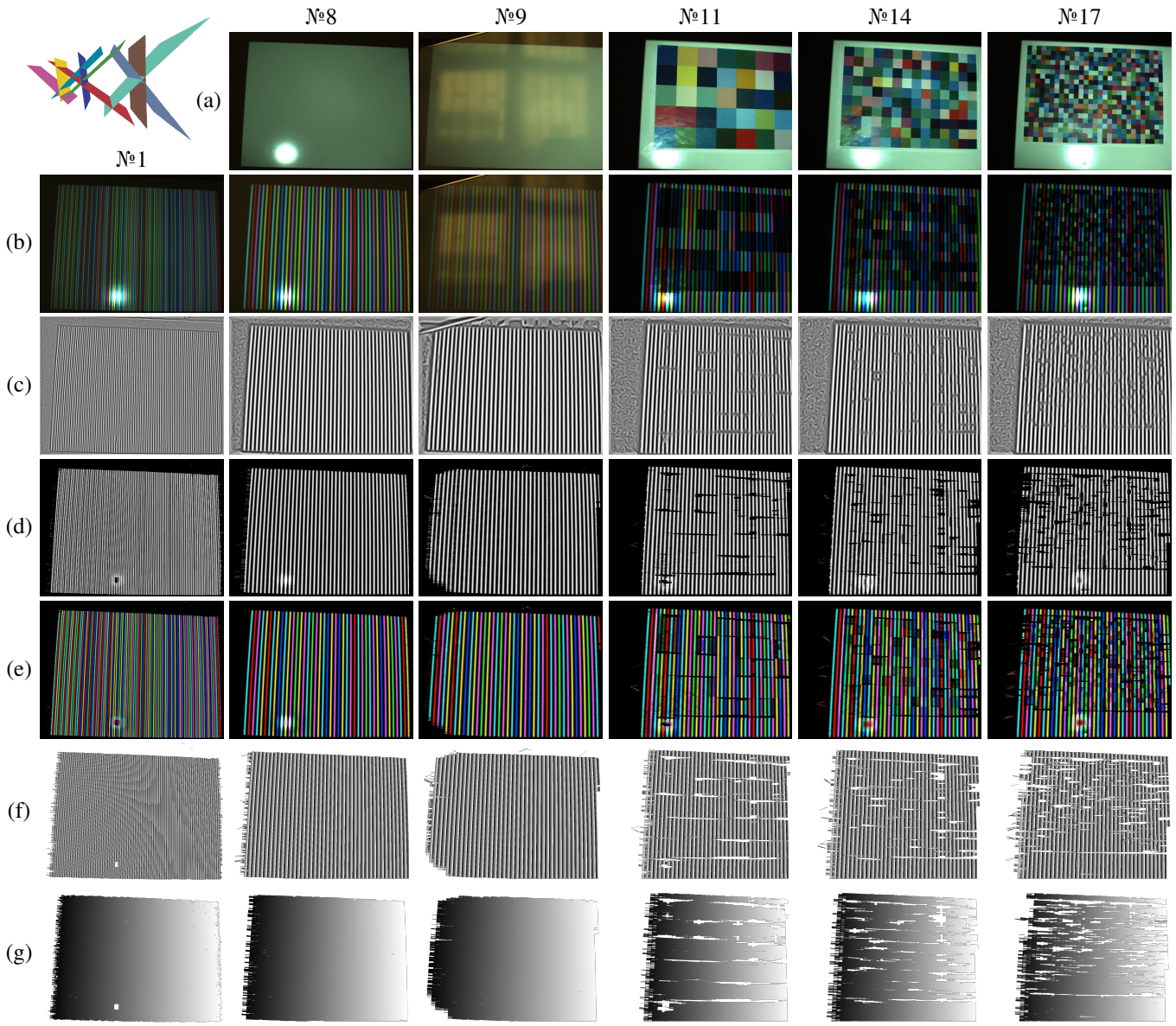


Fig. 8. Reconstruction of planar surfaces. In rows: (a) planar surface under ambient and white illumination; (b) planar surface under SL illumination; (c) vesselness; (d) equalization for phase estimation; (e) equalization for color decoding; (f) wrapped phase; (g) unwrapped phase/projector column. Top-left shows placements of white surfaces in calibration volume for measurements №2-10. For quantitative results see Table I.

the calibration volume for a total of 19 measurements: measurements №1-10 were of a whiteboard; measurements №11-13 were of a $5 \times 7$ checkerboard; measurements №14-16 were of a $10 \times 14$ checkerboard; and measurements №17-19 were of a $20 \times 29$ checkerboard. Regarding the distance from the camera-projector pair: №2-4 were placed in front of the calibration volume; №1,5-7,11-19 were placed in the middle of the calibration volume; and №8-10 were placed in the back of the calibration volume. Regarding orientation of the planar surfaces: №1,2,5,11,14,17 were oriented en-face; №3,6,9,12,15,18 were tilted left; and №4,7,10,13,16,19 were tilted right. For №1 we used the sequence of length 102; all other measurements used the sequence of length 42. For №1-10 ceiling lights were on and sunlight was coming through the window on the left thus affecting planar surfaces tilted left, e.g. Fig. 8 for №9; for №11-19 the window blinds were closed.

All planar surfaces were reconstructed using the proposed method and using a standard continuous single phase-shifting method (SPS) [10] at 28 shifts for precise phase estimation. SPS used the same fringe frequency to enable wrapped phase comparison. Quantitative results taking SPS as the ground truth are listed in Table I: columns $\mu_\phi$ and $\kappa_\phi$ (location and concentration of von Mises distribution) describe the wrapped phase error $\phi - \phi_{\text{SPS}}$; columns $\mu_{x_p}$ and $\sigma_{x_p}$ describe the projector coordinate error $x_p - x_{p,\text{SPS}}$; column RMSE is the root mean square error of a model plane fitted to the 3D data; and column % decoded is the percentage of pixels where SL illumination was successfully decoded compared to SPS. Note that $\sigma_{x_p}$ is high for some positions due to a small percentage of outliers; most were caused by directly reflected projector spotlight. Fig. 8 shows intermediate results for six selected

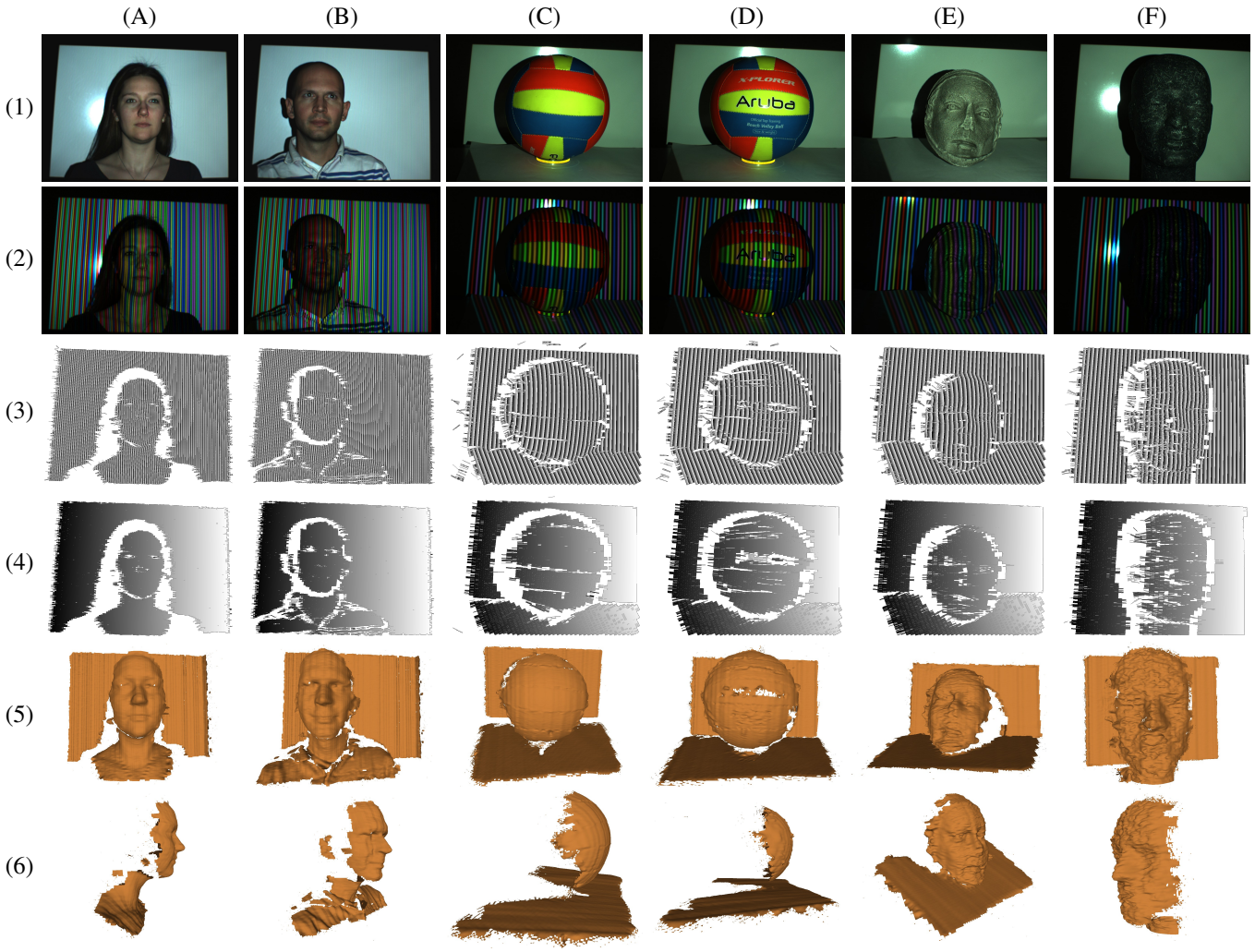


Fig. 9. Examples of 3D reconstruction for human faces and real-world objects. In rows: (1) face or object under white illumination; (2) face or object under SL illumination; (3) wrapped phase; (4) unwrapped phase/projector column; (5) front-view of reconstructed 3D surface; and (6) rotated/side view of reconstructed 3D surface. For quantitative results see Table II.

positions. The direct illumination from the window does not affect the reconstruction for №9. The checkerboards of $10 \times 14$ and $20 \times 29$ squares do not fulfill the *spatial invariance conditions* but are successfully reconstructed if the threshold $T_f$ is relaxed. The strong edges between the colored squares negatively affect ridge extraction.

### B. Human Faces and Real-World Objects

We have recorded human faces and real-world objects to demonstrate the versatility of the proposed method: (A) a female face; (B) a male face; (C) a colorful beach volley-ball; (D) a colorful beach volley-ball with text; (E) a plaster reproduction of female head; and (F) a textured male mannequin head. Human faces were recorded using the sequence of length 102; real-world objects were recorded using the sequence of length 42. Reconstruction results are shown in Fig. 9. Note that all reconstruction are dense. This property is more clearly demonstrated in Fig. 10 which shows a depth-colored point cloud of (A). most of the classical SL approaches using De Bruijn sequences produce sparse reconstructions which contain only the stripe maxima.

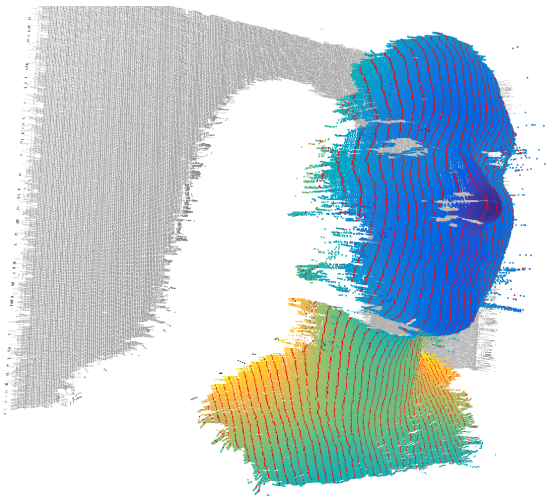

Fig. 10. Unfiltered point cloud of female face (A). Facial points are colored according to the distance from the camera; background points are gray. Centers of color stripes are marked red. The centers correspond to the standard De Bruijn SL coding; note the significantly higher number of points obtained using the proposed single-shot method.

TABLE I
QUANTITATIVE RESULTS FOR PLANAR SURFACE RECONSTRUCTION.

| № | $\mu_\phi$ [rad] | $\kappa_\phi$ | $\mu_{x_p}$ [px] | $\sigma_{x_p}$ [px] | RMSE [mm] | % decoded |
|---|---|---|---|---|---|---|
| 1 | 0.011 | 21.98 | 0.142 | 10.91 | 1.398 | 96.99 |
| 2 | 0.013 | 50.25 | −0.451 | 18.72 | 0.408 | 89.62 |
| 3 | 0.003 | 22.94 | −1.368 | 37.61 | 0.545 | 77.20 |
| 4 | 0.019 | 63.31 | 0.070 | 0.505 | 0.388 | 87.38 |
| 5 | 0.009 | 55.48 | −0.666 | 23.97 | 0.560 | 92.57 |
| 6 | 0.003 | 42.15 | 0.016 | 3.915 | 0.749 | 92.69 |
| 7 | 0.008 | 57.77 | 0.038 | 0.555 | 0.470 | 89.74 |
| 8 | 0.001 | 66.37 | −1.418 | 28.15 | 1.019 | 94.64 |
| 9 | −0.006 | 64.08 | −0.341 | 22.10 | 1.277 | 80.67 |
| 10 | 0.005 | 37.82 | 0.034 | 1.163 | 0.744 | 95.76 |
| 11 | 0.018 | 29.18 | −0.427 | 17.08 | 0.978 | 81.05 |
| 12 | 0.002 | 26.45 | 0.004 | 2.861 | 0.992 | 83.18 |
| 13 | 0.008 | 19.40 | 0.039 | 1.228 | 0.891 | 83.82 |
| 14 | 0.026 | 16.23 | −2.723 | 38.30 | 1.105 | 80.49 |
| 15 | 0.009 | 17.11 | −0.001 | 3.674 | 2.351 | 79.54 |
| 16 | 0.022 | 14.54 | 0.553 | 5.875 | 7.247 | 81.51 |
| 17 | 0.021 | 10.33 | −1.111 | 23.12 | 2.332 | 74.95 |
| 18 | 0.014 | 13.19 | −0.118 | 5.623 | 2.866 | 75.91 |
| 19 | 0.008 | 11.78 | 0.043 | 5.312 | 1.612 | 79.69 |

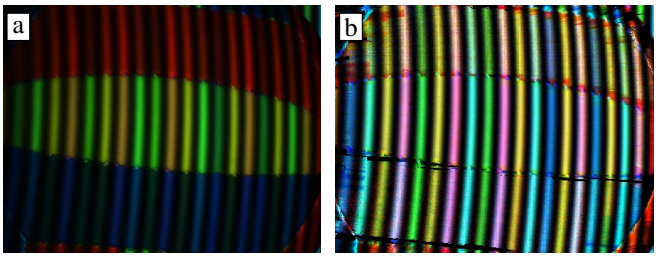


Fig. 11. Pure color surfaces: (a) a detail of the input image for the volley-ball (C); (b) equalization for color decoding. Note the colors are unrecognizable in (a) and are recognizable in (b) after equalization.

The human faces (A) and (B) reflect light in all channels and are less difficult to reconstruct. The beach volley-ball (C,D) is significantly harder to reconstruct as it contains pure colors and even text for (D): the reconstruction is possible only due to the proposed color decoding of Eq. (9) which maximally extends each De Bruijn window. Another problem for pure colored surfaces is cross-talk which is not compensated for in the current method. The effect of cross-talk is visible in Fig. 11: note in (b) how equalized colors are shifted toward white. The fine surface texture of the plaster female head (E) introduces significant noise into the estimated phase which manifests itself as a rippling of the reconstructed 3D surface. This problem is even more pronounced for the last object, a mannequin head (F), where the quite dark surface and fine white-streaked texture affect both the wrapped phase estimation and the color decoding: the wrapped phase estimate is even more noisy than (E). Nevertheless, the proposed method provides quite satisfactory results. We note that the mannequin head (F) is darker than any natural human skin color and is therefore a good indicator about results for dark skin colors. Quantitative results are shown in Table II. In addition to

TABLE II
QUANTITATIVE RESULTS FOR FACES AND REAL-WORLD OBJECTS.

| Id | $\mu_\phi$ [rad] | $\kappa_\phi$ | $\mu_{x_p}$ [px] | $\sigma_{x_p}$ [px] | % $\Delta_{x_p} < 1$ | % decoded |
|---|---|---|---|---|---|---|
| (A) | 0.057 | 7.989 | −0.168 | 17.11 | 96.15 | 90.11 |
| (B) | −0.050 | 8.139 | 0.455 | 28.62 | 95.30 | 87.25 |
| (C) | 0.008 | 12.01 | 0.100 | 28.42 | 96.45 | 76.06 |
| (D) | 0.008 | 10.01 | 1.540 | 29.82 | 93.25 | 78.48 |
| (E) | −0.010 | 14.77 | 0.388 | 10.25 | 93.46 | 82.07 |
| (F) | −0.005 | 7.837 | −0.280 | 36.72 | 87.74 | 77.84 |

parameters described for Table I we also list the percentage of points whose projector column error $\Delta_{x_p} = |x_p - x_{p,\text{SPS}}|$ is less than 1 px; note this value is about 90 %, therefore the majority of points are inliers and $\sigma_{x_p}$ is only large due to the small number of non-filtered outliers.

### C. Discussion

The proposed method has several strong advantages compared to other approaches to single-shot SL reconstruction: (1) the imaging geometry is not constrained; (2) the *self-equalizing* De Bruijn sequence removes most problems caused by using color; (3) the proposed color decoding scheme enables reconstruction of pure-colored objects; (4) the pattern is not sensitive to projector and camera blur; and (5) the proposed complex BP Hilbert filter is a better phase estimator than DFT based methods. The following sections discuss these aspects of our new work in more detail. We also discuss differences compared to a random-dots pattern.

*1) Imaging Geometry:* In the proposed method the phase directly encodes the projector row (or column) thus reducing the 3D position recovery to a simple triangulation between projector planes and camera rays. In FTP methods [21], [32] the object's depth is recovered using a phase-to-depth conversion formula which requires at least one reference plane to be recorded during the system calibration. Furthermore, the phase-to-depth conversion formula is valid only for crossed-optical-axes or parallel-optical-axes imaging geometries. Such restrictive imaging geometries may be difficult to achieve if standard consumer-grade equipment is used. In contrast, the proposed method only requires a common FOV between projector and camera.

*2) Equalization:* The purpose of equalization is to remove the influence of terms $A$ and $B$ in Eqs. (1) and (5). Previous works proposed alternative data transformations with the same purpose. For example, in [12] the V channel is modified to $\max(R,G,B) - \min(R,G,B)$ partially suppressing ambient light for better color detection and in [51] a quadrature transform is used to remove amplitude variation of gray data for better phase estimation. Such approaches are quite general and do not impose a specific constraint on the SL pattern itself which would make the task simple. Our previous work [15] imposed the *self-equalizing* constraint which enables removal of spatial variations in $A$ and $B$. We extended [15] by noting that a significant difference exists between equalization for phase estimation and for color decoding: for phase estimation we equalize each fringe separately, while for color decoding

we must equalize a group of adjacent fringes. This difference makes the *spatial invariance conditions* of [15] different for the two tasks: the color detection requires larger spatially invariant patches having the same albedo compared to the phase estimation. We stress the fact that the channel cross-talk does not affect phase estimation at all due to spatial phase coherence of the proposed pattern. The channel cross-talk only affects color decoding, but undesired effects are mitigated by the proposed robust $L_1$ color decoding.

*3) Color Decoding:* The reconstruction of pure color objects is a huge advantage compared to previous works. Equalization and the proposed color decoding scheme enable reconstruction of pure-color objects if the De Bruijn windows are extensible to some minimum size required for unique decoding. As noted in Section III-A, each color channel of a De Bruijn sequence of order $n$ is also a De Bruijn sequence of an order higher than $n$. Therefore, pure-color objects are reconstructible if the decoding window is extensible to the order of a particular color channel. Fortunately, common real-world objects seldom have pure red, green or blue patches and may be recovered even for smaller window sizes as there is always some information present in other channels. This property is clearly demonstrated in Fig. 9 for objects (C,D).

The proposed color decoding may seem similar to the dynamic programming (DP) approaches of [12], [17], [49]. Most DP approaces require an assumption of depth-ordering, or monotonicity: the color code cannot repeat iteself more than once in each scanline. The proposed decoding has no assumption about non-repetition of the color code along each scanline and will therefore succeed even in situations where DP fails.

*4) Projector and Camera Blur:* A colored sinusoidal fringe is robust to projector and camera blur: the phase and color at fringe maxima are only affected under severe blurring. This property enables the use of consumer-grade projectors, including short-throw projectors, which have quite narrow depth-of-field. For example, the SL pattern was blurred for planar surfaces in the front (Pos. 2-4) and in the back (Pos. 8-10) of the calibration volume without affecting the 3D reconstruction. Other methods of encoding fringe order such as that of Zhang et al. [52], who combine a sinusoidal fringe with a speckle pattern, and that of Budianto and Lun [53], who embed texture, may be significantly affected if the pattern is blurred.

*5) BP Hilbert Filter vs. DFT:* In the continuous domain the phase of Eq. (1) is estimated for each row separately. In FTP

$$\phi(t) = \mathrm{Arg}\big(\mathrm{ICTFT}\big[\mathrm{CTFT}[I(t)]H(\Omega)\big]\big), \qquad (12)$$

is used and in SPD

$$\phi(t) = \mathrm{Arg}\big(I(t) * h(t)\big) = \mathrm{atan2}(I(t) * h_Q(t), I(t) * h_I(t)), \qquad (13)$$

is used, where $t$ is a continuous coordinate describing the camera column, $H(\Omega)$ is an ideal BP filter in positive frequencies only, and $h(t) = h_I(t) + jh_Q(t)$ is a complex filter which decomposes $I(t)$ into band-limited in-phase and quadrature components. Both methods are conceptually the same when analyzed in the continuous domain: if $h(t) \bigcirc\!\!-\!\!\bullet\, H(\Omega)$ is a Fourier transform pair then SPD and FTP produce the same results. However, differences appear when actual implementations are compared in the discrete domain.

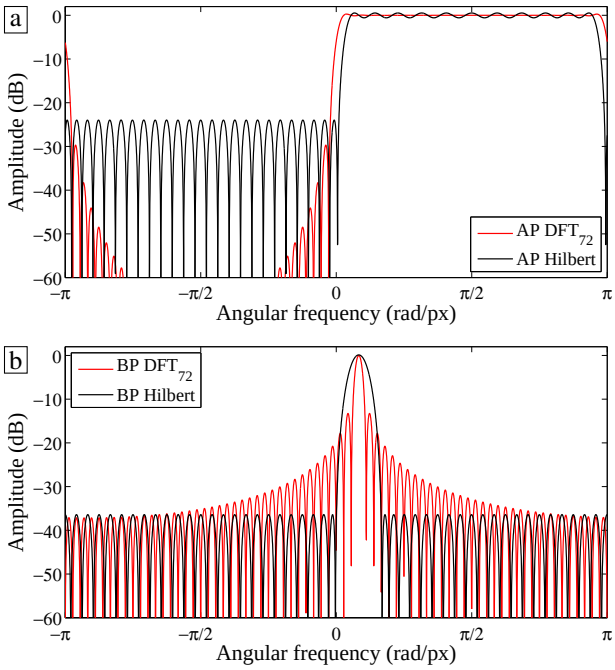


Fig. 12. Comparison of band-pass complex Hilbert filter of order 48 to $\mathrm{DFT}_{72}$: (a) all-pass AF characteristics; (b) band-pass AF characteristics for $\omega_0 = 0.2618$ rad/px and $B = 0.03$ rad/px.

In FTP Eq. (12) is usually discretized by replacing CTFT with $\mathrm{DFT}_N$, where $N$ is the total number of samples in one row. Such replacement introduces errors in phase estimation due to the circularity of $\mathrm{DFT}_N$ [40] and due to the spectral leakage in the negative frequencies [39]. Effects of circularity will be negligible if $I(t)$ has minimal discontinuity between outermost left and right samples. This may be achieved by using a dark background which produces wide dark borders in the image, or by using full-field imaging. Effects of spectral leakage in negative frequencies are significant for small orders $N$ of $\mathrm{DFT}_N$ and decrease as $N$ increases. In practice they are non-existent in full-field FTP where $N$ is large, but may be significant if the signal has discontinuities (due to shadows) or if WFT analysis on short windows is used.

Discretization of Eq. (13) in SPD is usually done by sampling $h(t)$. Such discretization has fewer side effects than replacement of CTFT with $\mathrm{DFT}_N$. however, sampling $h(t)$ directly is not the best approximation of Eq. (13), a fact long known in the signal processing community [40].

The approach we propose, using a BP complex Hilbert filter, approximates Eq. (13) in the discrete domain in the minimax sense. It is an optimal approximation and therefore produces superior phase estimation results, especially under windowed analysis with a low number of samples per scanline, as is the case for the proposed method. This is easily demonstrated by comparing the ideal amplitude-frequency (AF) characteristics, which is 1 in the band of interest and zero everywhere else, to the proposed BP complex Hilbert filter and to the classical approximation of Eq. (12) using $\mathrm{DFT}_N$. Fig. 12 shows AF characteristics of the proposed BP Hilbert filter

and of using $\mathrm{DFT}_N$ for phase estimation for a pattern with parameters $n = 3$, $w = 24\,\mathrm{px}$, $\omega_0 = 0.2618\,\mathrm{rad/px}$ and $B = 0.03\,\mathrm{rad/px}$, which allows for filter order 48 for the proposed method and $N = 72$ for $\mathrm{DFT}_{72}$. Note the chosen $\omega_0$ exactly matches one of $\mathrm{DFT}_{72}$ frequencies, something that would be difficult to achieve in practice. In Fig. 12 (a) all-pass (AP) AF characteristics are shown: compare the uneven suppression of AP $\mathrm{DFT}_{72}$ to the equiripple suppression of AP Hilbert filter in negative frequencies. In Fig. 12 (b) BP AF characteristics are shown: note the significant spectral leakage of BP $\mathrm{DFT}_{72}$ compared to BP Hilbert filter.

*6) Comparison with Random-Dots Pattern:* The random-dots pattern is a popular choice for single-shot 3D reconstruction and is used in today's consumer depth cameras such as Microsoft Kinect, ASUS Xtion, and Intel RealSense. Such devices operate using the *projected-texture* principle [54]: a projector illuminates the scene with a random-dots pattern; the acquired image is matched against the reference image(s) taken in advance using stereo-matching. Most often an infrared illumination is used to make the projected pattern invisible to humans. The main advantage of a random-dots pattern is computational simplicity of depth retrieval.

The random dots pattern is not smooth in intensity and is negatively affected by blurring. To mitigate against such blurring in practice a projector whose depth-of-field is larger than the depth range of the device must be used, e.g. a laser projector. Consumer-grade low-cost projectors which use visible light have quite narrow depth-of-field and are therefore a poor match for random-dots pattern. Our proposed pattern does not suffer from blurring effects as discussed in Section V-C4.

## VI. Conclusion

We have proposed a novel single-shot SL method combining a *self-equalizing* De Bruijn sequence, scale-space Hessian matrix analysis, and BP complex Hilbert filter. Although it is well known that De Bruijn patterns provide relatively *low density* of 3D reconstruction whose resolution is limited by the number of color stripes, our method provides a *very dense* 3D reconstruction. Compared to the upper bound of reconstructing every camera pixel, which may be achieved by time-multiplexing SL strategies such as SPS, the proposed method performs exceptionally well reconstructing between 80% and 90% of points. Equally important, about 90% of the reconstructed points have error in the recovered projector coordinate of less than 1 px compared to time-multiplexing SPS. The proposed method is extremely robust to object color. It is able to reconstruct even pure colored surfaces regardless of the general expectation that color patterns are usable only for relatively pale surfaces.

In our previous work [16] we have shown how to obtain dense reconstruction of static objects and sparse reconstruction of dynamic objects, including how to decompose a video frame into static and moving objects and how to apply the corresponding 3D reconstruction methods. Since this work proposes a method to obtain a dense reconstruction of dynamic objects as well, altogether it leads to a new framework for dense 3D reconstruction of video sequences.

## Acknowledgment

The authors are grateful to Dr. Fiona Berryman from the Royal Orthopaedic Hospital in Birmingham, United Kingdom, to Professor Joaquim Salvi from the University of Girona, Spain, and to Directeur de Recherche Peter Sturm from the Inria Grenoble Rhône-Alpes, France, whose suggestions improved the content and the presentation of this paper.

## References

[1] F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," *Optical Engineering*, vol. 39, no. 1, pp. 10–22, 2000.

[2] Z. Zhang, "Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques," *Optics and Lasers in Engineering*, vol. 50, no. 8, pp. 1097–1106, 2012.

[3] C. Wust and D. Capson, "Surface profile measurement using color fringe projection," *Machine Vision and Applications*, vol. 4, no. 3, pp. 193–203, 1991.

[4] Z. Zhang, D. P. Towers, and C. E. Towers, "Snapshot color fringe projection for absolute three-dimensional metrology of video sequences," *Appl. Opt.*, vol. 49, no. 31, pp. 5947–5953, Nov 2010.

[5] H. J. Chen, J. Zhang, D. J. Lv, and J. Fang, "3-D shape measurement by composite pattern projection and hybrid processing," *Opt. Express*, vol. 15, no. 19, pp. 12 318–12 330, Sep 2007.

[6] S. Fernandez and J. Salvi, "One-shot absolute pattern for dense reconstruction using De Bruijn coding and windowed Fourier transform," *Optics Communications*, vol. 291, pp. 70–78, 2013.

[7] W.-H. Su, "Projected fringe profilometry using the area-encoded algorithm for spatially isolated and dynamic objects," *Opt. Express*, vol. 16, no. 4, pp. 2590–2596, Feb 2008.

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

[9] R. Sagawa, R. Furukawa, and H. Kawasaki, "Dense 3D reconstruction from high frame-rate video using a static grid pattern," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36, no. 9, pp. 1733–1747, Sept 2014.

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

[11] 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*, vol. 20, no. 5, pp. 470–480, 1998.

[12] Y. Zhou, D. Zhao, Y. Yu, J. Yuan, and S. Du, "Adaptive color calibration based one-shot structured light system," *Sensors*, vol. 12, no. 8, pp. 10 947–10 963, 2012.

[13] P. Fechteler and P. Eisert, "Adaptive color classification for structured light systems," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2008, pp. 1–7.

[14] J. Pagès, J. Salvi, and J. Forest, "A new optimised De Bruijn coding strategy for structured light patterns," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4, Aug 2004, pp. 284–287.

[15] T. Petković, T. Pribanić, and M. Đonlić, "The self-equalizing De Bruijn sequence for 3D profilometry," in *Proceedings of the British Machine Vision Conference (BMVC)*, X. Xie, M. W. Jones, and G. K. L. Tam, Eds. BMVA Press, September 2015, pp. 155.1–155.11.

[16] M. Đonlić, T. Petković, and T. Pribanić, "3D surface profilometry using phase shifting of De Bruijn pattern," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[17] 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*, 2002, pp. 24–37.

[18] X. Peng, S. M. Zhu, and M. M. Su, C. J. and; Tseng, "Model-based digital moiré topography," *Optik*, vol. 110, no. 4, pp. 184–190, 1999.

[19] C. F. Lo, X. Peng, and L. Cai, "Surface normal guided method for two-dimensional phase unwrapping," *Optik - International Journal for Light and Electron Optics*, vol. 113, no. 9, pp. 439–447, 2002.

[20] J. Degrieck, W. V. Paepegem, and P. Boone, "Application of digital phase-shift shadow Moiré to micro deformation measurements of curved surfaces," *Optics and Lasers in Engineering*, vol. 36, no. 1, pp. 29–40, 2001.

[21] M. Takeda and K. Mutoh, "Fourier transform profilometry for the automatic measurement of 3-D object shapes," *Appl. Opt.*, vol. 22, no. 24, pp. 3977–3982, Dec 1983.
[22] J. Li, X. Su, and L. Guo, "Improved Fourier transform profilometry for the automatic measurement of three-dimensional object shapes," *Optical Engineering*, vol. 29, no. 12, pp. 1439–1444, Dec. 1990.
[23] J.-F. Lin and X. Su, "Two-dimensional Fourier transform profilometry for the automatic measurement of three-dimensional object shapes," *Optical Engineering*, vol. 34, no. 11, pp. 3927–3302, Nov. 1995.
[24] S. Toyooka and Y. Iwaasa, "Automatic profilometry of 3-D diffuse objects by spatial phase detection," *Appl. Opt.*, vol. 25, no. 10, pp. 1630–1633, May 1986.
[25] K. Freischlad and C. L. Koliopoulos, "Fourier description of digital phase-measuring interferometry," *J. Opt. Soc. Am. A*, vol. 7, no. 4, pp. 542–551, Apr 1990.
[26] S. Tang and Y. Y. Hung, "Fast profilometer for the automatic measurement of 3-D object shapes," *Appl. Opt.*, vol. 29, no. 20, pp. 3012–3018, Jul 1990.
[27] J. Villa, M. Servin, and L. Castillo, "Profilometry for the measurement of 3-D object shapes based on regularized filters," *Optics Communications*, vol. 161, no. 1-3, pp. 13–18, 1999.
[28] C. A. Hobson, J. T. Atkinson, and F. Lilley, "The application of digital filtering to phase recovery when surface contouring using fringe projection techniques," *Optics and Lasers in Engineering*, vol. 27, no. 4, pp. 355–368, 1997.
[29] L. Chi-Fang and L. Chih-Yang, "A new approach to high precision 3-D measuring system," *Image and Vision Computing*, vol. 17, no. 11, pp. 805–814, 1999.
[30] H. Schreiber and J. H. Bruning, *Optical Shop Testing*, 3rd ed., ser. Willey Series in Pure and Applied Optics. Wiley-Interscience, July 2007, ch. Phase Shifting Interferometry, pp. 547–655.
[31] K. Creath, J. Schmit, and J. C. Wyant, *Optical Shop Testing*, 3rd ed., ser. Willey Series in Pure and Applied Optics. Wiley-Interscience, July 2007, ch. Optical Metrology of Diffuse Surfaces, pp. 756–807.
[32] X. Su and W. Chen, "Fourier transform profilometry: a review," *Optics and Lasers in Engineering*, vol. 35, no. 5, pp. 263–284, 2001.
[33] Q. Kemao, "Two-dimensional windowed Fourier transform for fringe pattern analysis: Principles, applications and implementations," *Optics and Lasers in Engineering*, vol. 45, no. 2, pp. 304–317, 2007.
[34] S. Fernandez, M. A. Gdeisat, J. Salvi, and D. Burton, "Automatic window size selection in windowed Fourier transform for 3D reconstruction using adapted mother wavelets," *Optics Communications*, vol. 284, no. 12, pp. 2797 – 2807, 2011.
[35] M. A. Gdeisat, A. Abid, D. R. Burton, M. J. Lalor, F. Lilley, C. Moore, and M. Qudeisat, "Spatial and temporal carrier fringe pattern demodulation using the one-dimensional continuous wavelet transform: Recent progress, challenges, and suggested developments," *Optics and Lasers in Engineering*, vol. 47, no. 12, pp. 1348–1361, 2009.
[36] R. Rodríguez-Vera and M. Servín, "Phase locked loop profilometry," *Optics & Laser Technology*, vol. 26, no. 6, pp. 393–398, 1994.
[37] S. Haykin, *Communication Systems*, 4th ed. Wiley, Apr. 2000.
[38] M. A. Gdeisat, D. R. Burton, and M. J. Lalor, "Spatial carrier fringe pattern demodulation by use of a two-dimensional continuous wavelet transform," *Appl. Opt.*, vol. 45, no. 34, pp. 8722–8732, Dec 2006.
[39] A. Reilly, G. Frazer, and B. Boashash, "Analytic signal generation-tips and traps," *Signal Processing, IEEE Transactions on*, vol. 42, no. 11, pp. 3241–3245, Nov 1994.
[40] S. K. Mitra, *Digital Signal Processing*, 3rd ed. McGraw-Hill, Jan. 2005.
[41] F. Ruskey, *Combinatorial generation*, preliminary working draft ed. University of Victoria, Victoria, 2003.
[42] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 113–125, 1998.
[43] A. Frangi, W. Niessen, K. Vincken, and M. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Interventation*, ser. Lecture Notes in Computer Science, W. Wells, A. Colchester, and S. Delp, Eds. Springer Berlin Heidelberg, 1998, vol. 1496, pp. 130–137.
[44] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, 1st ed. Wiley, Apr. 1996.
[45] J. Marple, S.L., "Computing the discrete-time "analytic" signal via FFT," *Signal Processing, IEEE Transactions on*, vol. 47, no. 9, pp. 2600–2603, Sep 1999.
[46] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization Transmission*, Oct 2012, pp. 464–471.
[47] R. Juang and A. Majumder, "Photometric self-calibration of a projector-camera system," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.
[48] L. Kinell, "Multichannel method for absolute shape measurement using projected fringes," *Optics and Lasers in Engineering*, vol. 41, no. 1, pp. 57–71, 2004.
[49] J. Pagès, J. Salvi, C. Collewet, and J. Forest, "Optimised De Bruijn patterns for one-shot shape acquisition," *Image and Vision Computing*, vol. 23, no. 8, pp. 707–720, 2005.
[50] D. C. Hutchison, "Introducing brilliantcolor(tm) technology," Texas Instruments, Tech. Rep., 2005.
[51] J. A. Quiroga, J. A. Gómez-Pedrero, and Á. García-Botella, "Algorithm for fringe pattern normalization," *Optics Communications*, vol. 197, no. 1-3, pp. 43–51, 2001.
[52] Y. Zhang, Z. Xiong, Z. Yang, and F. Wu, "Real-time scalable depth sensing with hybrid structured light illumination," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 97–109, 2014.
[53] Budianto and D. P. K. Lun, "Robust fringe projection profilometry via sparse representation," *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1726–1739, April 2016.
[54] K. Konolige, "Projected texture stereo," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, May 2010, pp. 148–155.

**Tomislav Petković** (S'03–M'11) received the engineer's degree, magister degree and Ph.D. degree in electrical engineering all from the University of Zagreb, in 2002, 2006 and 2010 respectively. He is currently a post-doctoral researcher at the University of Zagreb. His main fields of interest are digital image processing and analysis, depth sensing and 3D imaging.

**Tomislav Pribanić** (M'04) is an Associate Professor at the University of Zagreb Faculty of Electrical Engineering and Computing. He teaches several undergraduate and graduate courses in the field of algorithms and data structures, image processing, sensors and human motion analysis. His main research interests include computer vision and biomedical signal measurement and analysis. He has led a number of scientific domestic and international projects, collaborating with the researchers from EU and outside EU. He was a visiting researcher at INRIA Rhone-Alpes, Grenoble, France and Fraunhofer IGD, Darmstadt, Germany. Results of his research have been implemented in technological projects and he has received recognitions for innovations as well. He is a member of IEEE, IFMBE and a collaborating member of the Croatian Academy of Engineering.

**Matea Đonlić** (S'15) received her B.Sc. degree in Computing (2012) and M.Sc. degree in Information and communication technology (magna cum laude, 2014) at the University of Zagreb, Croatia. She is currently pursuing her Ph.D. in Computing at the University of Zagreb Faculty of Electrical Engineering and Computing, Croatia. Her current research interests in the computer vision focus on the shape analysis and the 3D reconstruction. She is a Student Member of IEEE.