Fringe Projection Profilometry using Color

Tomislav Petković

University of Zagreb
Faculty of Electrical Engineering and Computing

IEEE Croatia Section
Signal Processing Society Chapter

Zagreb, December 15, 2016

Contents

Contents

Contents

Contents

3D Surface Scanning

3D surface scanning is a procedure which yields the position of object's surface (shape).

- 3D surface scanning
  - also known as 3D profilometry, range finding, depth sensing
  - techniques that acquire position of a surface
  - we measure coordinates (x, y, z) of a surface
  - sometimes surface reflectance or albedo is also measured
  - examples: fringe projection profilometry
- 3D imaging
  - techniques that acquire true 3D data
  - we measure some property p for each point (x, y, z) within a finite volume
  - examples: CT, MRI, 3DRA

3D Surface Scanning

3D surface scanning is a procedure which yields the position of object's surface (shape).

- 3D surface scanning
  - also known as 3D profilometry, range finding, depth sensing
  - techniques that acquire position of a surface
  - we measure coordinates (x, y, z) of a surface
  - sometimes surface reflectance or albedo is also measured
  - examples: fringe projection profilometry
- 3D imaging
  - techniques that acquire true 3D data
  - we measure some property p for each point (x, y, z) within a finite volume
  - examples: CT, MRI, 3DRA

3D Surface Scanning

3D surface scanning is a procedure which yields the position of object's surface (shape).

- 3D surface scanning
  - also known as 3D profilometry, range finding, depth sensing
  - techniques that acquire position of a surface
  - we measure coordinates (x, y, z) of a surface
  - sometimes surface reflectance or albedo is also measured
  - examples: fringe projection profilometry
- 3D imaging
  - techniques that acquire true 3D data
  - we measure some property p for each point (x, y, z) within a finite volume
  - examples: CT, MRI, 3DRA

3D Surface Scanning

3D surface scanning is a procedure which yields the position of object's surface (shape).

- 3D surface scanning
  - also known as 3D profilometry, range finding, depth sensing
  - techniques that acquire position of a surface
  - we measure coordinates (x, y, z) of a surface
  - sometimes surface reflectance or albedo is also measured
  - examples: fringe projection profilometry
- 3D imaging
  - techniques that acquire true 3D data
  - we measure some property p for each point (x, y, z) within a finite volume
  - examples: CT, MRI, 3DRA

3D Surface Scanning

3D surface scanning is a procedure which yields the position of object's surface (shape).

- 3D surface scanning
  - also known as 3D profilometry, range finding, depth sensing
  - techniques that acquire position of a surface
  - we measure coordinates (x, y, z) of a surface
  - sometimes surface reflectance or albedo is also measured
  - examples: fringe projection profilometry
- 3D imaging
  - techniques that acquire true 3D data
  - we measure some property p for each point (x, y, z) within a finite volume
  - examples: CT, MRI, 3DRA

A Key Concept: Triangulation

Triangulation means determining the location of a point by forming triangles to it from known points.

To reconstruct a point on the surface of an object we must observe the object from at least two different viewpoints.

Most often two cameras are used: *stereo vision*.

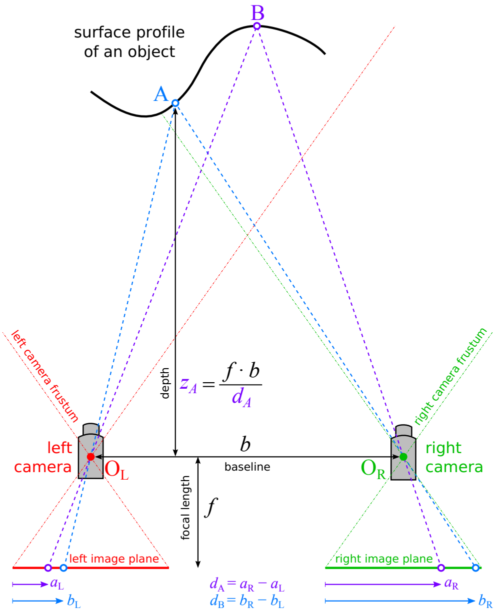
Stereo Vision

Uses two cameras to measure depth.

Processing steps:

1. undistort images;
2. stereo rectification;
3. find corresponding points;
4. triangulate.

Disadvantage: does not work for texture-less surfaces.



Structured Light

Replace one camera with a projector: *structured light*.

Projector projects a specially crafted image which is called *structured light pattern*.

The principle of structured light is to reconstruct the surface profile from the distortion of the projected structured light pattern.

Compared to stereo vision in structured light 3D scanning the task of finding corresponding points in two images is replaced by the pattern decoding from one image.

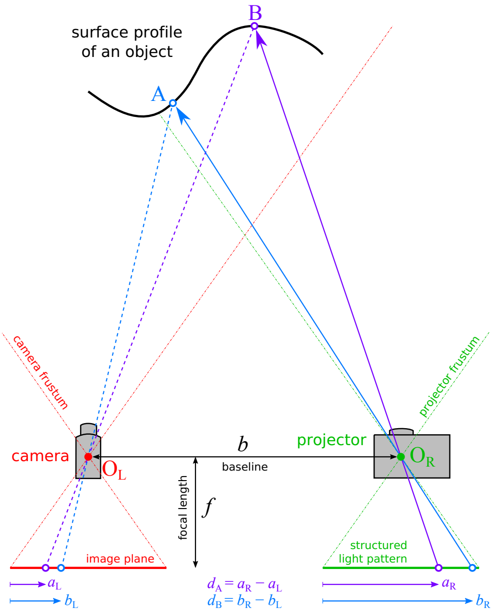
Structured Light

Observes deformation of projected pattern image.

Processing steps:

1. decode pattern (to get projector coordinates);
2. undistort camera coordinates;
3. triangulate.

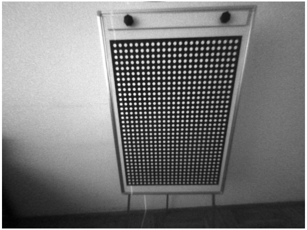
Issues: ambient light and object albedo.



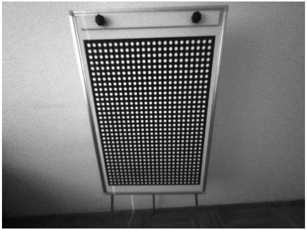
Structured Light vs. Stereo Vision

Stereo Vision

left camera

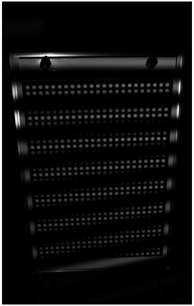


right camera

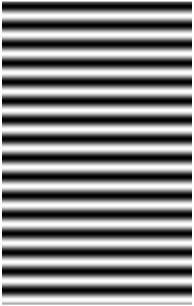


Structured Light

camera



pattern (projector)



Structured Light Patterns

In structured light surface scanning we may project one or more patterns:

1) one-shot patterns

- reconstruction from a single image
- object may move
- spatial pattern decoding
- reconstruction is sparse/low-resolution

2) multi-shot patterns

- multiple images are projected in time
- object must be stationary
- temporal pattern decoding
- reconstruction is dense/high-resolution

Structured Light Patterns

In structured light surface scanning we may project one or more patterns:

1) one-shot patterns

- reconstruction from a single image
- object may move
- spatial pattern decoding
- reconstruction is sparse/low-resolution

2) multi-shot patterns

- multiple images are projected in time
- object must be stationary
- temporal pattern decoding
- reconstruction is dense/high-resolution

Structured Light Patterns

In structured light surface scanning we may project one or more patterns:

1) one-shot patterns

- reconstruction from a single image
- object may move
- spatial pattern decoding
- reconstruction is sparse/low-resolution

2) multi-shot patterns

- multiple images are projected in time
- object must be stationary
- temporal pattern decoding
- reconstruction is dense/high-resolution

Structured Light Patterns

In structured light surface scanning we may project one or more patterns:

1) one-shot patterns

- reconstruction from a single image
- object may move
- spatial pattern decoding
- reconstruction is sparse/low-resolution

2) multi-shot patterns

- multiple images are projected in time
- object must be stationary
- temporal pattern decoding
- reconstruction is dense/high-resolution

Structured Light Patterns

In structured light surface scanning we may project one or more patterns:

1) one-shot patterns

- reconstruction from a single image
- object may move
- spatial pattern decoding
- reconstruction is sparse/low-resolution

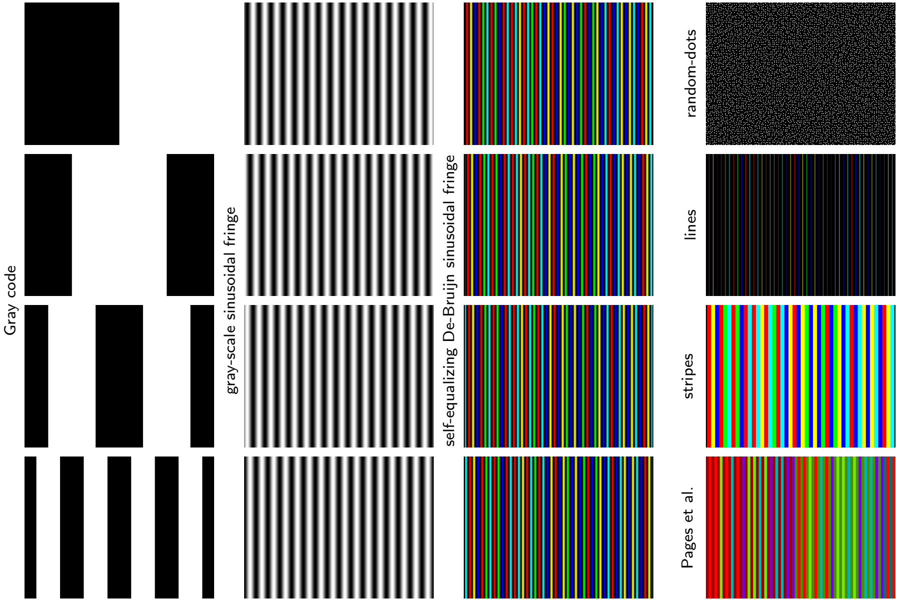
2) multi-shot patterns

- multiple images are projected in time
- object must be stationary
- temporal pattern decoding
- reconstruction is dense/high-resolution

Structured Light Patterns: Examples

Multi-Shot Patterns

One-Shot Patterns



Contents

Sinusoidal Fringe

The most commonly used gray-scale sinusoidal fringe pattern is

$$I_p(x, y) = I_0\big(1 - \cos(2\pi x/\lambda + \varphi_n)\big)/2, \qquad (1)$$

where (x, y) are projector pixel coordinates, I_0 is the maximum intensity, λ is fringe wavelength (in px), and φ_n is phase shift.

For one-shot patterns only one image is used ($\varphi_n = 0$).

For multi-shot patterns a set of N images is used ($0 \leq n < N$). Each image has different phase shift φ_n.

$n = 0,\ \varphi_n = 0$

$n = 1,\ \varphi_n = \frac{\pi}{2}$

$n = 2,\ \varphi_n = \pi$

$n = 3,\ \varphi_n = 3\pi/2$

The Importance of Phase

Multi-shot gray-scale sinusoidal fringe patterns are:

$$I_{p,\mathrm{col}}(x,y) = I_0\big(1-\cos(2\pi x/\lambda_{\mathrm{col}}+\varphi_n)\big)/2, \quad 0 \le n < N$$
$$I_{p,\mathrm{row}}(x,y) = I_0\big(1-\cos(2\pi y/\lambda_{\mathrm{row}}+\varphi_m)\big)/2, \quad 0 \le m < M$$

The projector column x and row y are encoded in phase.

Encoding the projector coordinates in phase provides several key advantages:

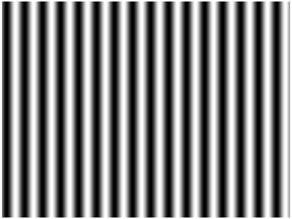
1. insensitivity to ambient light;
2. insensitivity to object albedo (color); and
3. insensitivity to blur/out-of-focus effects.

Phase Recovery

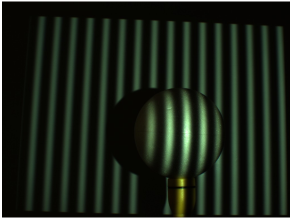
Phase is recovered:

1. for multi-shot patterns using temporal analysis:
   1.1 three-step algorithm (special case for $N = 3$);
   1.2 least-squares algorithm;
   1.3 Schwider-Hariharan algorithm (weighted LS);
2. for one-shot patterns using spatial analysis:
   2.1 spatial-domain methods;
   2.2 transform-domain methods.

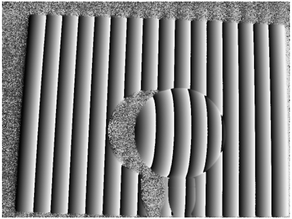
Phase Recovery: Example



I_p for $n = 0$



observed I_p for $n = 0$



recovered phase

Phase Unwrapping

The phase is measurable modulo-2π only.

Formally, let Φ denote the true phase value and let ϕ denote the phase measured modulo-2π; then

$$\Phi = \phi + 2\pi k, \tag{2}$$

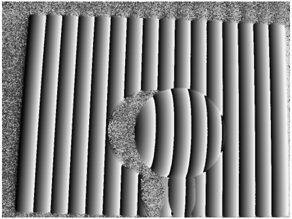
where Φ is the true or absolute phase, ϕ is the wrapped or principal phase, and $k \in \mathbb{Z}$ is an unknown integer which models the phase ambiguity and is sometimes called period-order or fringe-order number.

The task of phase unwrapping is to unwrap the wrapped phase ϕ and obtain the true phase Φ.

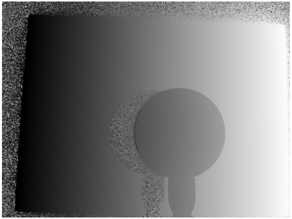
Phase Unwrapping

Approaches to phase unwrapping may be classified into two distinct groups:

1. spatial phase unwrapping (for one-shot patterns),
2. temporal phase unwrapping (for multi-shot patterns).

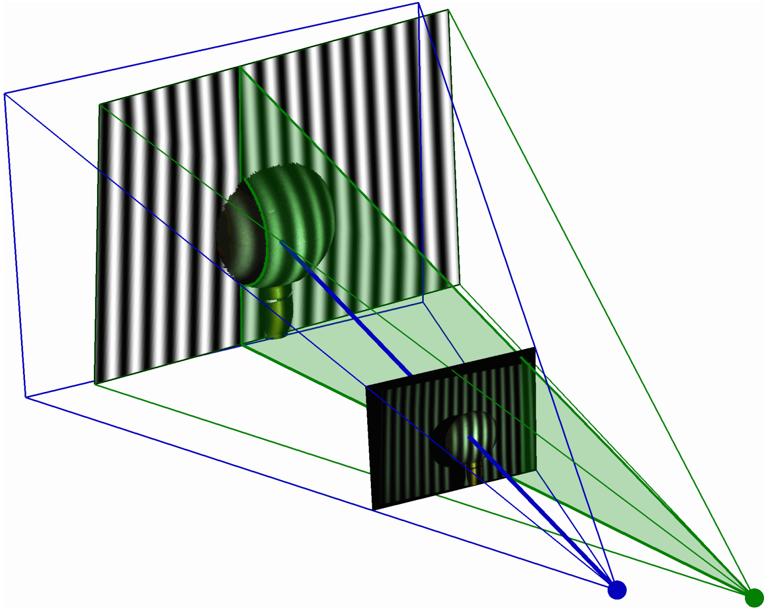


wrapped phase ϕ



unwrapped phase Φ

Triangulation



Contents

The remainder of this presentation is describes our work in IEEE Transactions on Image Processing, Vol. 25, No. 11, November 2016, doi:10.1109/TIP.2016.2603231:

IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 25, NO. 11, NOVEMBER 2016 5131

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 bandpass 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 Sagawa 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

Manuscript received March 8, 2016; revised June 21, 2016; accepted July 31, 2016. Date of publication August 26, 2016; date of current version September 13, 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).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2016.2603231

T. Petković, T. Pribanić, M. Đonlić

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

Motivation

- Structured light is a robust approach to 3D profilometry
- We want to enable robust 3D scanning using low cost consumer grade electronics
  - projector and camera embedded in a mobile device
  - home DLP projector and web camera
- Almost all such devices use color

The Problem of Color

Using colored structured light is difficult:

1. the projector must accurately render colors
2. the camera must accurately capture colors
3. the object must not shift colors
4. the system must be insensitive to ambient lighting

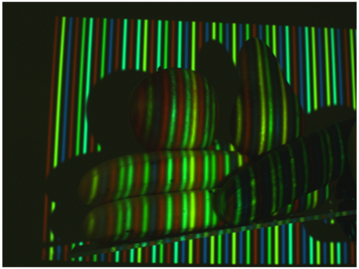
too much green

auto white balance (AWB)

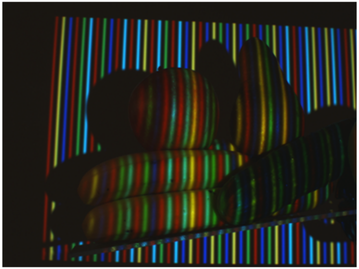
too much blue

calibrated using SpyderCUBE

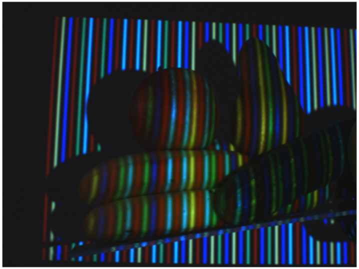
too much green



auto white balance (AWB)



too much blue



calibrated using SpyderCUBE

Color Model

Model for RGB color-space imaging in structured light:

$$\underbrace{\begin{bmatrix} R \\ G \\ B \end{bmatrix}}_{I_c} = \underbrace{\begin{bmatrix} a_{RR} & a_{RG} & a_{RB} \\ a_{GR} & a_{GG} & a_{GB} \\ a_{BR} & a_{BG} & a_{BB} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} k_R & 0 & 0 \\ 0 & k_G & 0 \\ 0 & 0 & k_B \end{bmatrix}}_{K} f\Bigg(\underbrace{\begin{bmatrix} r \\ g \\ b \end{bmatrix}}_{I_p} \Bigg) + \underbrace{\begin{bmatrix} R_0 \\ G_0 \\ B_0 \end{bmatrix}}_{I_0}, \quad (3)$$

where I_c is the color recorded by the camera, I_p is the color instruction to the projector, A is the channel transfer matrix, K is the albedo matrix, f is a monotonic function modeling projector's non-linearity, and I_0 is the ambient lighting.

Using color for fringe projection profilometry is hard as every object and scene has a different and generally unknown albedo K.

Solution: Construct a structured light pattern I_p in such way that all relevant parameters of the model (3) may be estimated from the recorded image.

Self-Equalizing De Bruijn Sequence

De Bruijn Sequence

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.

Self-equalizing constraint

The *self-equalizing* constraint requires all color channels to span the full available dynamic range in every De Bruijn window.

Example sequence of length 42:

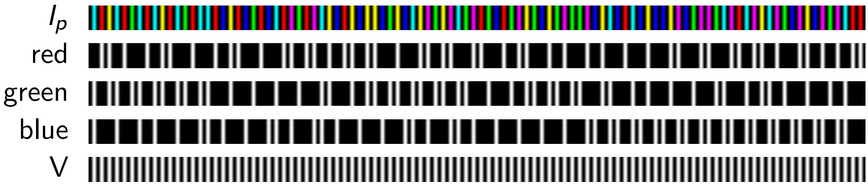
`CRMCYRCYBCYMGYMCGMYCBYMBYCMRCMGCMYBMYGMCRY`

Structured Light Pattern

A full-length De Bruijn sequence of order $n = 3$ has exactly $k^n = 6^3 = 216$ elements.

The *self-equalizing* sequence we have proposed removes problematic windows for which color model cannot be solved leaving a sequence of length 102:

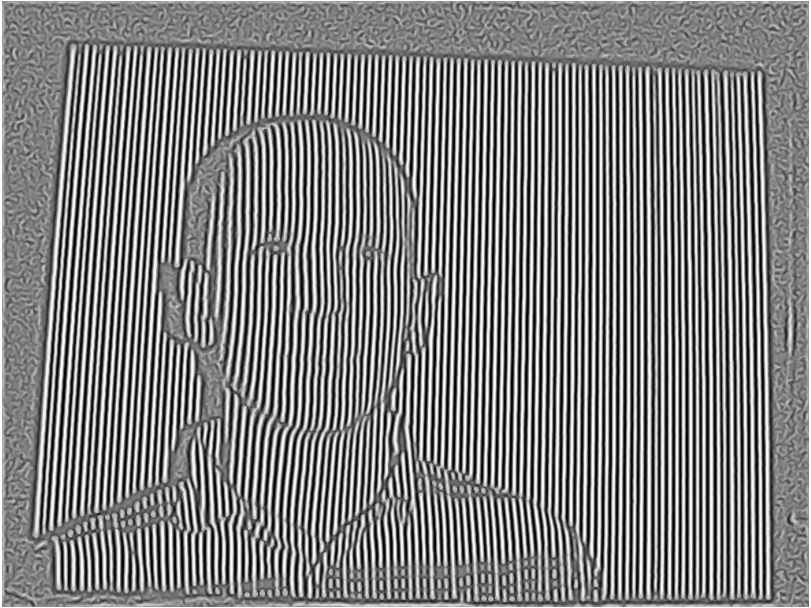
CRYCRGCRCYRCGRCCRBYRBGRBCRMGRMCYBRYBYYBGYBCYMGYMCG-
MRGBRGMYGBYGMGGMBYCBRCBYBBYMBGMMGCMRCMYCMGBMYBMGMC-
RR



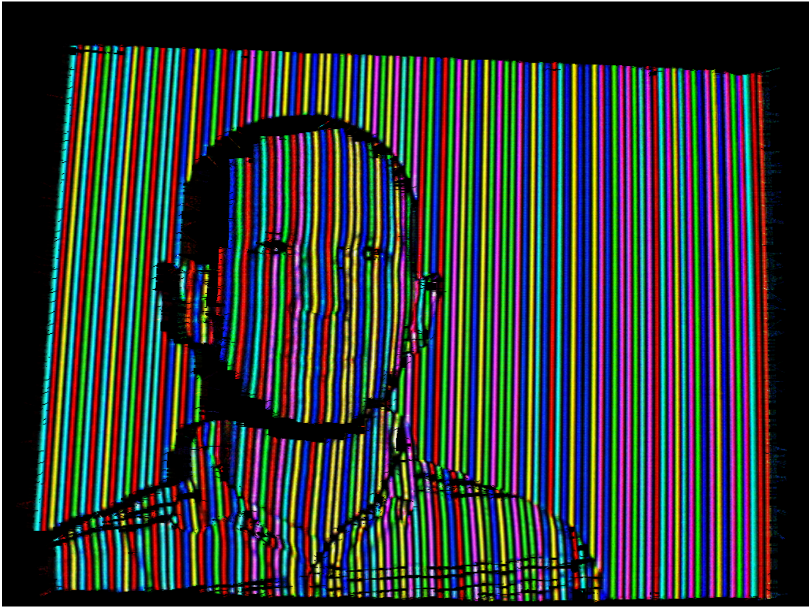
Input image

Sum of all channels

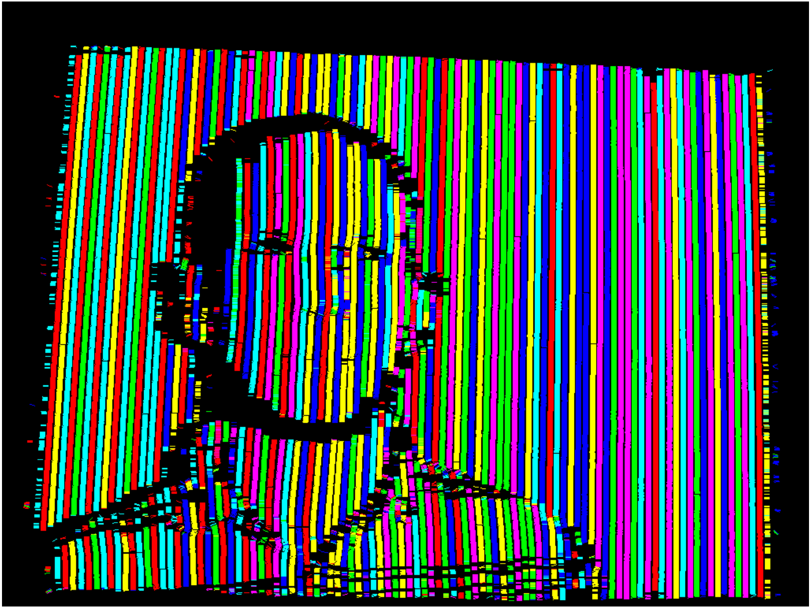
Vesselness map (fringe detection)



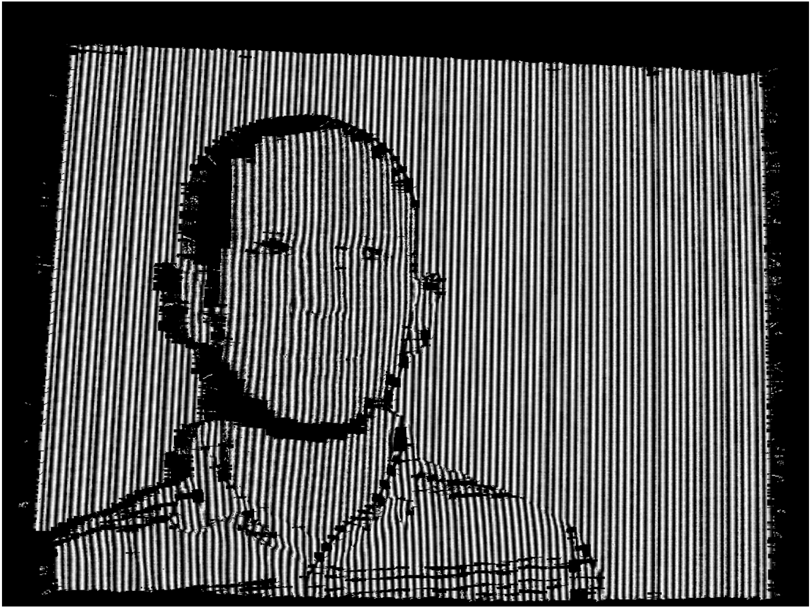
Color equalization (remove external factors)



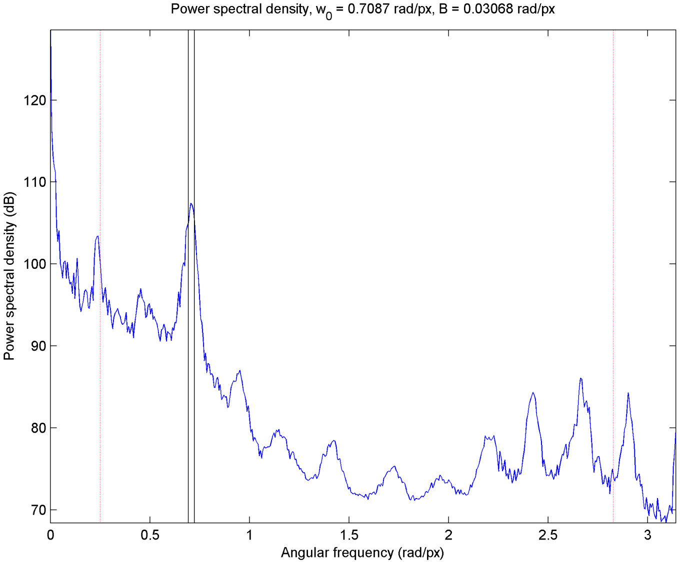
Recognized projected color



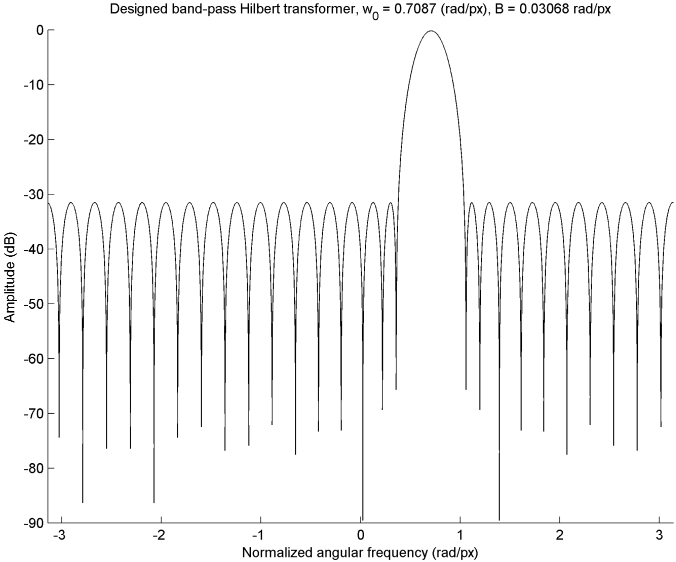
Extracted and equalized V channel



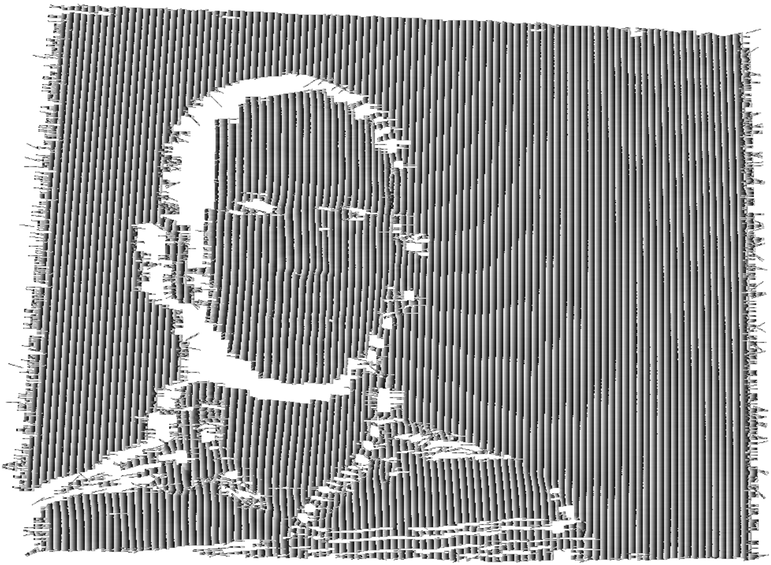
Designing band-pass Hilbert transformer (step 1)



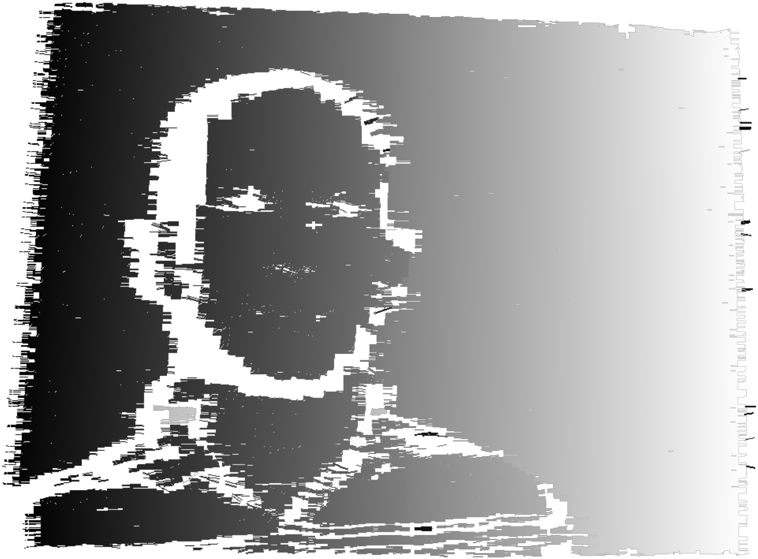
Designing band-pass Hilbert transformer (step 2)



Estimated wrapped phase

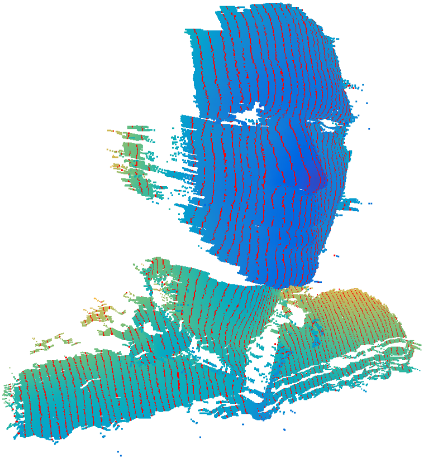


Unwraped phase (via De Bruijn property)



Final 3D reconstruction (surface)

Dense vs. Sparse 3D Reconstruction



Conclusion

Proposed *self-equalizing* De Bruijn sequence allows:

1. the removal of ambient lighting,
2. the removal of object albedo, and
3. the equalization of channel gains.

Proposed clever composition of standard signal processing elements (scale-space Hessian matrix analysis and BP complex Hilbert filtering) enables a *very dense* 3D reconstruction which yields between 80% and 90% of points compared to time-multiplexing approaches while retaining excellent precision as about 90% of the reconstructed points have error in the recovered projector coordinate of less than 1 px.