

Received 29 July 2022, accepted 11 August 2022, date of publication 19 August 2022, date of current version 31 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3200388

RESEARCH ARTICLE

Multiprojector Multicamera Structured Light Surface Scanner

TOMISLAV PETKOVIĆ^{ID}, (Member, IEEE), AND TOMISLAV PRIBANIĆ^{ID}, (Senior Member, IEEE)

University of Zagreb, Faculty of Electrical Engineering and Computing, 10000 Zagreb, Croatia

Corresponding author: Tomislav Petković (tomislav.petkovic.jr@fer.hr)

This work was supported by the Croatian Science Foundation under Project IP-2019-04-9157 and Project IP-11-2013-3717.

ABSTRACT Structured light surface scanners are usually limited to a single projector and cannot provide a full surface profile in a single scan. Multiple projectors must be employed to extend the imaging area over the whole surface of the imaged object. Using multiple projectors is a challenging problem due to inter-projector interference which makes pattern decoding hard. We describe a practical approach to multi-projector structured light surface scanning based on sinusoidal fringe patterns and temporal phase shifting which allows simultaneous projector employment. In our approach a different set of temporal phase shifts is assigned to each projector to enable an efficient and robust separation of observed combined patterns into the contributions of each projector. The proposed approach does not place any restrictions on projector placement nor on the upper limit of the number of projectors used. We demonstrate the applicability of the proposed approach on a structured light surface scanner comprised of three projectors and six cameras.

INDEX TERMS Structured light, 3D imaging, multi-projector structured light, temporal phase shifting, frequency-division multiplexing.

I. INTRODUCTION

Structured light (SL) surface scanning [1], [2] is an important topic in computer vision and industrial metrology with different applications such as 3D sensing, object recognition, industrial inspection, reverse engineering, clothing design, to name a few. Most often a digital projector and a digital camera are used; together they form a calibrated projector-camera pair which is called a SL surface *scanner*. In SL surface scanning a pattern is projected onto an object of interest, is then imaged, and the deformation of the pattern due to object shape is observed. In such a setup the projector is an *active* device which emits modulated light, a SL pattern, onto the scanned object. The camera is a *passive* device which observes the light reflected from the object's surface. The image acquired by the camera contains the deformed SL pattern which carries all the information required to reconstruct the surface profile. How that information is extracted from the acquired image(s) depends on the pattern type. For example, in fringe projection profilometry (FPP) [3] pattern's phase is

the information carrier, while for binary coded patterns [4] information is encoded using pattern intensity.

Many practical applications of SL profilometry require a full surface profile of an object [5]. Obtaining the whole surface profile in a single scan using a SL surface scanner comprised of a single projector-camera pair is impossible due to a limited field-of-view (FOV) and due to inevitable occlusions and shadows. Therefore, multiple scans must be performed to obtain many partial surface profiles which must then be stitched together to construct the full surface of an object. Such multiple partial scans of a single object may be achieved in two ways: (1) by moving either the object or the scanner, and (2) by adding more projectors and cameras to the scanner itself. In the first approach partial scans are always obtained *sequentially* one after another with movement performed in-between the scans to change the FOV. Movement may be done by hand [6], however most often a rotating table [5], [7], [8], [9], [10] is used to move the object. Sometimes a robotic arm [11], [12] is used to move the scanner. The use of rotating tables or robotic arms simplifies the stitching of partial scans by providing an accurate information about the relative movement. If no such information is available then

The associate editor coordinating the review of this manuscript and approving it for publication was Olutayo O. Oyerinde^{ID}.

scan stitching becomes a tedious task, e.g. when movement is performed by hand. In the second approach all projectors and cameras may be used *simultaneously*. If all projectors and cameras are calibrated to the same reference coordinate system then such *simultaneous* employment eliminates the stitching step: the full surface profile is obtained immediately. Also, in contrast to the first approach, no movement whatsoever is required which may be a significant advantage in some applications.

What are the problems when adding more cameras and projectors to a SL surface scanner? Adding more cameras to the scanner is trivial as they are *passive* devices. Adding projectors is much more difficult as they are *active* devices which mutually interfere. A careful design of a SL pattern is necessary to overcome the undesired effects of inter-projector interference. Unfortunately, almost all SL patterns currently described in the literature [13], [14], [15], [16] are designed for a single projector systems. Using such SL patterns limits multi-projector scanners to the simplest acquisition strategy where projectors are used in-turns, a so-called *sequential* projector employment [17], [18]. The *sequential* projector employment is not desirable as it unnecessarily prolongs the scanning time and as it increases the amount of acquired images proportionally to the number of projectors used. A true multi-projector multi-camera scanner must employ projectors *simultaneously*, hence it requires a SL pattern which allows clear, efficient and robust separation of acquired images into the contributions of each projector.

Surface scanners using more than one projector or camera may be classified into three distinct groups: (a) single-projector multi-camera scanners; (b) multi-projector single-camera scanners; and (c) true multi-projector multi-camera scanners.

Single-projector multi-cameras scanners are frequently presented in the literature. The most common approach is active stereo [19], [20], [21], [22], [23] where a simple texture which improves the robustness of stereo-matching algorithms is projected. However, the projected texture is never decoded, hence all such methods are not true SL methods. Approaches that both project and decode a SL pattern are also plenty: a fringe SL pattern is used in [24], [25], [26], [27], and [28], Gray code in [29], and line patterns in [30] and [31]. More specifically: in [24], [25], and [26] a stereo-pair is used to directly unwrap a single wrapped phase; in [27], [28], and [31] a decoded SL pattern is used to drive the stereo-matching algorithm; in [29] the projected Gray code is used to directly establish the correspondences between the two cameras of a stereo-pair; and in [30] a stereo-pair is used to determine line order and drive SL decoding. Interestingly, all aforementioned scanners use only two cameras which are spatially arranged to form a stereo-pair with a maximum possible FOV overlap between cameras, i.e. additional cameras are not used to increase the size of the visible surface.

There are several multi-projector single-camera scanners described in the literature. The most common approach is assigning different spatial orientations to SL pattern of each

projector which enables pattern separation: Woolford and Burnett [32] describe a two-projectors one-camera system which uses one horizontal and one vertical sinusoidal fringe; Je *et al.* [33] describe a two-projectors two-cameras system which uses one horizontal and one vertical color stripe pattern; and Sagawa *et al.* [34] and Furukawa *et al.* [35], [36] describe a two- and three-projectors single-camera systems which use colored line pattern with different spatial line orientation for each projector. Another approach is by Servin *et al.* [37], [38] who describe a multi-projector single-camera system which relies on a very precise spatial placement of projectors, a so called co-phased profilometry, and therefore has limited general applicability. Similarly, Jia *et al.* [39] describe a four-projectors one-camera system specifically designed to scan indoor scenes; their design places the projectors in a circle facing outward so there exists no overlap between projectors' FOV.

Multi-projector multi-camera scanners are not often described in the literature. Two main reasons may be the difficulty of the multi-projector interference problem and the build complexity of such scanners. Yan *et al.* [40] provide a theoretical analysis of a multi-projector multi-camera system which combines De Bruijn color sequences and random dot patterns; unfortunately, their work provides simulated results only making practical applicability dubious. Garcia and Zakhor [17], [41] present a three-projector nine-camera system for human body scanning which uses hardware synchronization to implement *sequential* acquisition strategy which neatly sidesteps the projector interference problem at the cost of increased acquisition time. Petković *et al.* [42], [43] present a three-projector six-camera system where phase shifts are carefully assigned to each projector to enable efficient separation of projected patterns by using discrete Fourier transform.

Considering the state-of-the art a majority of multi-projector scanners described in the literature use some form of spatial pattern multiplexing where different spatial orientations are used to separate the patterns. However, enforcing different spatial orientations in a multi-projector systems places undesirable limits on projector placement. Furthermore, it is not easily extensible to an arbitrary number of projectors due to the fact that spatial orientation of a SL pattern changes from projection to recording depending on the object's surface normal and on the position of the camera. Using temporal pattern multiplexing [42] is a more attractive solution as it enables robust separation of projected patterns without placing constraints on projector placement.

In this work we describe a practical approach to multi-projector multi-camera SL surface scanners which is an extension of the method we have proposed in [42] and [43]. The proposed approach uses gray-level sinusoidal fringe patterns such that each projector has its own set of temporal phase shifts which together allow robust separation of projected patterns when projectors are employed *simultaneously*. The ability to employ projectors *simultaneously* is an advantage compared to the scanners which use *sequential*

employment [17], [18], [41]. Also, the proposed approach does not place any restrictions on projector placement nor on the upper limit of the number of projectors used; an arbitrary spatial arrangement is allowed. That property is a clear advantage compared to the scanners which use spatial orientation to separate patterns [32], [33], [34], [35], [36]. Compared to our previous work [42], [43]:

- We extend the method from fixed phase shifts mandated by the use of the discrete Fourier transform to arbitrary phase shifts.
- We give clear instructions and analytical expressions on how to select phase shifts to achieve an efficient computational separation.
- We discuss the conditions which phase shifts must satisfy for the separation to be robust to unwanted harmonic interference.
- We discuss in detail the photometric calibration of the scanning system.

This article is organized as follows: in Section II we first give a short review of single projector FPP using temporal phase shifting which is followed by the description of the proposed method. Experimental results and discussion are presented in Section III. We conclude in Section IV.

II. MULTI-PROJECTOR MULTI-CAMERA SCANNER

Before discussing the proposed method we first review a classical temporal phase shifting strategy which uses one projector only [44], and then we then extend this strategy to an arbitrary number of projectors.

A. CLASSICAL TEMPORAL PHASE SHIFTING

In classical temporal phase shifting one projector is used. The projector projects a set of N sinusoidal fringe patterns all having the same spatial frequency ω and a different temporal phase shift $\varphi[n]$. Each of N projected fringe patterns is imaged by the camera yielding a set of N frames.

The intensity of the n th sinusoidal fringe in the projector's coordinate system (x_{PRJ}, y_{PRJ}) is:

$$I_{PRJ}(x_{PRJ}, y_{PRJ}) = I_0(1 + \cos(\omega x_{PRJ} - \varphi[n]))/2, \quad (1)$$

where I_0 is projector's intensity and where $n = 1, \dots, N$ is a frame number (time-step). Note that Eq. (1) states that the projector coordinate x_{PRJ} is encoded in the phase using the spatial frequency ω . Also, instead of x_{PRJ} we may encode y_{PRJ} or a spatial coordinate along an arbitrary axis in the projector's image plane.

The intensity of the n th projected fringe as observed by the camera and expressed in camera's coordinate system (x_{CAM}, y_{CAM}) is:

$$I_{CAM}(x_{CAM}, y_{CAM}) = I_{AMB} + hI_{PRJ}, \quad (2)$$

where I_{AMB} is a constant ambient illumination and where $h, 0 \leq h < 1$, models a pixel-dependent channel loss. Combining Eqs. (1) and (2) yields:

$$I_{CAM}(x_{CAM}, y_{CAM}) = a + b \cos(\omega x_{PRJ} - \varphi[n]), \quad (3)$$

where

$$a = I_{AMB} + \frac{1}{2}hI_0 \quad (4)$$

is a time-invariant component comprised of both ambient and averaged projector illumination and where

$$b = \frac{1}{2}hI_0 \quad (5)$$

is the amplitude of the observed fringe or *contrast*. Eq. (3) defines a spatio-temporal signal where $\varphi[n]$ are temporal phase shifts and where $\omega x_{PRJ} = \Phi$ is a spatial phase. Note that in Eq. (3) the projector coordinate x_{PRJ} is encoded in the phase for each camera pixel (x_{CAM}, y_{CAM}) ; recovering the signal phase enables the extraction of x_{PRJ} and subsequent surface reconstruction via triangulation. Hence, all processing is temporal and is performed on a set of N frames for each camera pixel (x_{CAM}, y_{CAM}) independently. From here on spatial camera coordinates are omitted for clarity.

Expanding Eq. (3) for the n th frame yields

$$I_{CAM}[n] = a + b \cos(\omega x_{PRJ}) \cos(\varphi[n]) + b \sin(\omega x_{PRJ}) \sin(\varphi[n]). \quad (6)$$

We must recover the projector coordinate x_{PRJ} from the set of N temporal observations $I_{CAM}[n]$. That is usually done using the least-squares approach as explained in [44], [45], and [46]. We say the signal $I_{CAM}[n]$ is decomposed into parameters according to Eq. (6).

Here we give a brief overview of the signal decomposition problem in the context of Hilbert spaces. Recall that the dot product between two vectors $c[n]$ and $s[n]$, $n = 1, \dots, N$, in N -dimensional space over the real numbers is

$$\langle c[n], s[n] \rangle = \langle s[n], c[n] \rangle = \sum_{n=1}^N s[n]c[n]. \quad (7)$$

Let $V = \{1[n], c[n], s[n]\}$ be a set of three N -dimensional vectors which appear in Eq. (6) so

$$\begin{aligned} 1[n] &= 1 \\ c[n] &= \cos(\varphi[n]) \\ s[n] &= \sin(\varphi[n]) \end{aligned} \quad (8)$$

for $n = 1, \dots, N$. Then Eq. (6) may be interpreted to state we are observing a linear combination (a weighted sum) of three N -dimensional vectors in V with the weights

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_c \\ w_s \end{bmatrix} = \begin{bmatrix} a \\ b \cos(\omega x_{PRJ}) \\ b \sin(\omega x_{PRJ}) \end{bmatrix} \quad (9)$$

which we have to recover. Let \mathbf{G} be the 3×3 Gramian matrix of all dot-products in V ,

$$\mathbf{G} = \begin{bmatrix} \langle 1, 1 \rangle & \langle 1, c \rangle & \langle 1, s \rangle \\ \langle c, 1 \rangle & \langle c, c \rangle & \langle c, s \rangle \\ \langle s, 1 \rangle & \langle s, c \rangle & \langle s, s \rangle \end{bmatrix}, \quad (10)$$

where the time-step n was omitted for clarity. If the vectors in V are linearly independent then $\dim(\text{span}(V)) \geq 3$, the Gramian matrix \mathbf{G} has an inverse, and signal decomposition

of Eq. (6) may be solved for the unknown weight vector \mathbf{w} using the orthogonality principle as follows:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_c \\ w_s \end{bmatrix} = \mathbf{G}^{-1} \begin{bmatrix} \langle I_{CAM}, 1 \rangle \\ \langle I_{CAM}, c \rangle \\ \langle I_{CAM}, s \rangle \end{bmatrix}. \quad (11)$$

If all temporal phase shifts $\varphi[n]$ are different then it is sufficient to have $N \geq 3$ for the invertible matrix \mathbf{G} . Also, due to Hilbert's projection theorem the weights given by Eq. (11) are optimal in the least-square sense and are equivalent to solutions presented in [44], [45], and [46].

Once the weights \mathbf{w} are recovered using Eq. (11) a wrapped spatial phase $\phi \equiv \Phi \pmod{2\pi}$ may be recovered using

$$\phi = \text{atan2}(w_s, w_c), \quad (12)$$

and the contrast b may be recovered using

$$b = \sqrt{w_c^2 + w_s^2}. \quad (13)$$

To reconstruct the surface profile first the contrast b is used to determine which areas are illuminated by the projector by comparing it to some threshold T . Pixels for which

$$b > T \quad (14)$$

are considered illuminated. The threshold T is best set as a percentage of $\frac{1}{2}I_0$, e.g. setting T to 10% of $\frac{1}{2}I_0$ requires that channel loss h of Eq. (5) is no lower than 0.1. Next, the wrapped spatial phase ϕ is unwrapped for the illuminated pixels only using one of the well known phase unwrapping algorithms [47], [48], [49] to obtain Φ . The spatial projector coordinate x_{PRJ} is then recovered using

$$x_{PRJ} = \Phi/\omega. \quad (15)$$

Finally, the surface profile is reconstructed via triangulation [50].

In practice the phase shifts $\varphi[n]$ may be freely selected as long as the Gramian matrix is invertible. A selection of particular interest is one which diagonalizes the Gramian matrix \mathbf{G} of Eq. (10) and which makes Eq. (11) trivial to solve. One possible choice which is the standard in FPP [47] is

$$\varphi[n] = 2\pi(n-1)/N, \quad n = 1, \dots, N. \quad (16)$$

The additional advantage of phase shifts given by Eq. (16) is that they allow the user to design a customized phase extraction filter using the results of signal processing theory as was proposed by Servin *et al.* [51].

B. MULTI-PROJECTOR TEMPORAL PHASE SHIFTING

Let us now present the proposed solution to the problem of *simultaneous* employment of multiple projectors, which is an extension of our previous work [42].

Let P be the total number of projectors. When all projectors are employed *simultaneously* a camera observes an additive combination of all projected sinusoidal fringes for each of

N recorded frames. Adding the contributions of additional $P - 1$ projectors to Eq. (3) yields

$$I_{CAM} = a + \sum_{p=1}^P b_p \cos(\omega_p x_{PRJ,p} - \varphi_p[n]), \quad (17)$$

where b_p , ω_p , and $\varphi_p[n]$ are contrast, spatial fringe frequency, and temporal phase shift of p th projector. The term a of Eq. (3) which contains the ambient illumination and the average intensity of each projector is similarly transformed to

$$a = I_{AMB} + \sum_{p=1}^P \frac{1}{2} h_p I_p, \quad (18)$$

where I_p is intensity of p th projector and where $0 \leq h_p < 1$ models a pixel-dependent channel loss from p th projector to the camera. The contrast terms b_p of Eq. (17) are now different for each projector,

$$b_p = \frac{1}{2} h_p I_p. \quad (19)$$

Similarly to the decomposition of Eq. (6) the observed intensity given by Eq. (17) may be decomposed into a linear combination of sine and cosine functions:

$$I_{CAM}[n] = a + \sum_{p=1}^P b_p \cos(\omega_p x_{PRJ,p}) \cos(\varphi_p[n]) + \sum_{p=1}^P b_p \sin(\omega_p x_{PRJ,p}) \sin(\varphi_p[n]). \quad (20)$$

Let $V = \{1[n], c_p[n], s_p[n] : p = 1, \dots, P\}$ be a set of $2P + 1$ vectors of N elements each so

$$\begin{aligned} 1[n] &= 1 \\ c_p[n] &= \cos(\varphi_p[n]) \\ s_p[n] &= \sin(\varphi_p[n]). \end{aligned} \quad (21)$$

Then Eq. (20) may be interpreted as a linear combination of $2P + 1$ vectors with $2P + 1$ weights

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_{c,1} \\ w_{s,1} \\ \vdots \\ w_{c,P} \\ w_{s,P} \end{bmatrix} = \begin{bmatrix} a \\ b_1 \cos(\omega_1 x_{PRJ,1}) \\ b_1 \sin(\omega_1 x_{PRJ,1}) \\ \vdots \\ b_P \cos(\omega_P x_{PRJ,P}) \\ b_P \sin(\omega_P x_{PRJ,P}) \end{bmatrix}. \quad (22)$$

The weights \mathbf{w} must be recovered from the N -element observation vector $I_{CAM}[n]$.

Recovery may be performed in the same way as was done in Section II-A for the case of one projector. The Gramian matrix \mathbf{G} is a $2P + 1 \times 2P + 1$ matrix:

$$\begin{bmatrix} \langle 1, 1 \rangle & \langle 1, c_1 \rangle & \langle 1, s_1 \rangle & \dots & \langle 1, c_P \rangle & \langle 1, s_P \rangle \\ \langle c_1, 1 \rangle & \langle c_1, c_1 \rangle & \langle c_1, s_1 \rangle & \dots & \langle c_1, c_P \rangle & \langle c_1, s_P \rangle \\ \langle s_1, 1 \rangle & \langle s_1, c_1 \rangle & \langle s_1, s_1 \rangle & \dots & \langle s_1, c_P \rangle & \langle s_1, s_P \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \langle c_P, 1 \rangle & \langle c_P, c_1 \rangle & \langle c_P, s_1 \rangle & \dots & \langle c_1, c_P \rangle & \langle c_P, s_P \rangle \\ \langle s_P, 1 \rangle & \langle s_P, c_1 \rangle & \langle s_P, s_1 \rangle & \dots & \langle s_1, c_P \rangle & \langle s_P, s_P \rangle \end{bmatrix}. \quad (23)$$

If all vectors in V are linearly independent then $\dim(\text{span}(V)) \geq 2P + 1$ and the Grammian matrix \mathbf{G} has an inverse. The least-squares estimate of weights is then

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_{c,1} \\ w_{s,1} \\ \vdots \\ w_{c,P} \\ w_{s,P} \end{bmatrix} = \mathbf{G}^{-1} \begin{bmatrix} \langle I_{\text{CAM}}, 1 \rangle \\ \langle I_{\text{CAM}}, c_1 \rangle \\ \langle I_{\text{CAM}}, s_1 \rangle \\ \vdots \\ \langle I_{\text{CAM}}, c_P \rangle \\ \langle I_{\text{CAM}}, s_P \rangle \end{bmatrix}. \quad (24)$$

The weights obtained by Eq. (24) are clearly separated by projector index p , therefore the signal separation between projectors when they are employed *simultaneously* is achieved. Due to clear signal separation the unknown projector's coordinate $x_{\text{PRJ},p}$ of p th projector is easily recoverable using Eqs. (12) and (15), as explained in Section II-A.

C. SELECTING TEMPORAL PHASE SHIFTS

A key difference between one projector and multiple projectors is in the selection of temporal phase shifts $\varphi[n]$. For one projector the selection is a simple one as explained in Section II-A. For P projectors the phase shifts $\varphi_p[n]$ have to be selected to allow simple, efficient, and robust wrapped phase recovery.

A similar problem was studied in signal processing which gives one important result regarding wrapped phase selection [52]. Let the temporal phase shifts for the p th projector have the form

$$\varphi_p[n] = \Omega_p(n - 1) \quad (25)$$

where $0 < \Omega_p$ and $\Omega_{p_1} \neq \Omega_{p_2}$ for $p_1 \neq p_2$, i.e. temporal phase shifts $\varphi_p[n]$ are selected from an increasing linear sequences each having a different *temporal frequency* Ω_p . For such choice of temporal phase shifts if $N \geq 2P + 1$ then all $2P + 1$ vectors in V are linearly independent [52]. Then the Grammian matrix \mathbf{G} given by Eq. (23) is invertible, which is a sufficient condition to separate observed frames into the contributions of each projector.

This result means one may randomly assign different temporal frequencies Ω_p to each of p projectors to realize a multi-projector surface scanner which allows *sequential* projector employment, or even more generally phase shifts $\varphi_p[n]$ may be randomly assigned.¹ However, a random choice of temporal frequencies and/or phase shifts will in general yield a Grammian matrix \mathbf{G} which is not *diagonal*; having a *diagonal* Grammian matrix \mathbf{G} significantly reduces computational complexity and makes separation *efficient*. Also, in addition to efficiency we may also require some resistance to signal *interference* to achieve *robustness*.

In the remainder of this section we first discuss conditions required to make separation *efficient*, and then we discuss conditions to make separation *robust*.

¹For a completely random assignment one must always numerically verify that the Grammian is invertible.

1) LINEAR PHASE SHIFTS FOR EFFICIENT SEPARATION

The Grammian matrix \mathbf{G} given by Eq. (23) is comprised of several different dot products. Three of these must always evaluate to zero for the Grammian matrix to be diagonal. The dot products which must evaluate to zero are:

$$\begin{aligned} \langle 1, c_p \rangle &= \langle c_p, 1 \rangle = \sum_{n=1}^N c_p[n] = 0 \\ \langle 1, s_p \rangle &= \langle s_p, 1 \rangle = \sum_{n=1}^N s_p[n] = 0 \\ \langle c_{p_1}, s_{p_2} \rangle &= \langle s_{p_2}, c_{p_1} \rangle = \sum_{n=1}^N c_{p_1}[n]s_{p_2}[n] = 0 \end{aligned} \quad (26)$$

Eq. (26) must hold for all p, p_1 , and $p_2, p_1 \neq p_2$. Combining Eq. (25), which guarantees the invertibility of the Grammian matrix if all Ω_p are different, with Eq. (26), which guarantees the Grammian matrix is diagonal, gives the conditions on the set of phase shifts required to achieve an *efficient* separation of projectors' contributions.

Evaluating all three sums of Eq. (26) as a finite sums of geometric series of complex exponentials yields the following equations:

$$\begin{aligned} 0 &= \sin(\Omega_p N/2), \quad p = 1, \dots, P \\ 0 &= \sin((\Omega_{p_1} \pm \Omega_{p_2})N/2), \quad p_1 \neq p_2 \end{aligned} \quad (27)$$

Solving Eq. (27) together with conditions of [52] then gives the sufficient conditions for invertible and diagonal Grammian matrix:

$$\begin{aligned} N &\geq 2P + 1 \\ \Omega_p &= 2k_1\pi/N, \quad k_1 \in \mathbb{N} \setminus \{0\} \\ \Omega_{p_1} - \Omega_{p_2} &= 2k_2\pi/N, \quad p_1 \neq p_2 \wedge k_2 \in \mathbb{Z} \setminus \{0\} \end{aligned} \quad (28)$$

To summarize, the sufficient conditions to obtain the invertible and diagonal Grammian matrix which is required for *efficient* separation of projectors are: (1) the number of projected fringes/recorded frames N must be at least twice the number of projectors plus one, (2) phase shifts of p th projector must form a discrete linear sequence whose temporal frequency Ω_p is a multiple of $2\pi/N$, and (3) temporal frequencies of phase shift sequences must be different for each projector.

One possible choice which satisfies the conditions of Eq. (28) is selecting the temporal frequency of the phase shift sequence using the projector index p :

$$\Omega_p = 2p\pi/(2P + 1), \quad p = 1, \dots, P. \quad (29)$$

Then phase shifts for the p th projector are

$$\varphi_p[n] = \Omega_p(n - 1), \quad n = 1, \dots, 2P + 1, \quad (30)$$

and the two vectors in V corresponding to the p th projector are

$$\begin{aligned} c_p[n] &= \cos(2p\pi/(2P + 1)) \\ s_p[n] &= \sin(2p\pi/(2P + 1)). \end{aligned} \quad (31)$$

The choice of temporal phase shifts given by Eq. (29) significantly simplifies the phase recovery as Eqs. (12) and (13) transform into

$$\begin{aligned} \phi_p &= \text{atan2}(\langle I_{\text{CAM}}[n], s_p[n] \rangle, \langle I_{\text{CAM}}[n], c_p[n] \rangle) \\ b_p &= \sqrt{(\langle I_{\text{CAM}}[n], s_p[n] \rangle)^2 + (\langle I_{\text{CAM}}[n], c_p[n] \rangle)^2} \end{aligned} \quad (32)$$

Furthermore, note that due to the diagonal \mathbf{G} the right-hand side of Eq. (24) reduces to dot products of $I_{\text{CAM}}[n]$ and all vectors in V , i.e. we are representing $I_{\text{CAM}}[n]$ using vectors in V as a basis. The basis in V is comprised of one constant function which models the ambient illumination and of sine and cosine functions which model each projector. All elements of V may then be decomposed into sums of complex exponentials [53],

$$\begin{aligned} 1[n] &= W_{2P+1}^0 \\ c_p[n] &= \frac{1}{2}(W_{2P+1}^{-p(n-1)} + W_{2P+1}^{p(n-1)}) \\ s_p[n] &= \frac{1}{2j}(W_{2P+1}^{-p(n-1)} - W_{2P+1}^{p(n-1)}) \end{aligned} \quad (33)$$

where

$$W_A^B = \exp(-2\pi jA/B) \quad (34)$$

denotes a time-discrete complex exponential with the period B and the exponent A ; j is an imaginary unit. The decomposition of vectors in V into terms of complex exponentials allows the use of the fast Fourier transform (FFT) algorithm [54] for signal decomposition as $2P + 1$ complex exponentials $W_{2P+1}^{\pm(n-1)p}$, $p = 0, \dots, P$ and $n = 1, \dots, 2P + 1$ are the basis of the discrete Fourier transform in $2P + 1$ points [53]. Hence, both wrapped phase ϕ_p and contrast b_k may be computed using FFT. Let

$$\hat{I}_{\text{CAM}}[k] = \text{DFT}_{2P+1}[I_{\text{CAM}}[n]], \quad k = 0, \dots, 2P, \quad (35)$$

be the discrete Fourier transform of $I_{\text{CAM}}[n]$ in $2P + 1$ points computed using the FFT algorithm. Then

$$\phi_p = -\text{Arg}(\hat{I}_{\text{CAM}}[p]) = -\text{Arg}\left(\frac{2P+1}{2} b_k e^{-j\phi_k}\right) \quad (36)$$

and

$$b_p = \frac{2}{2P+1} |\hat{I}_{\text{CAM}}[p]|. \quad (37)$$

Therefore, selecting the linear phase shifts using Eqs. (29) and (30) allows one to use the FFT algorithm to efficiently separate the contributions of P projectors from $2P + 1$ observed frames.

2) TEMPORAL ALIASING AND ROBUSTNESS

The choice of linear phase shifts using Eqs. (29) and (30) may lead to temporal aliasing problems when projectors do not project a perfect sinusoid. This may happen due to many reasons; a common one in digital fringe projection profilometry is gamma transformation [55], [56], [57], [58].

Consider first a case of a single projector. Denote by $g(\cdot)$ the non-linear amplitude transfer function of the projector, i.e. for an input I the projector outputs $g(I)$. Eq. (2) then becomes

$$I_{\text{CAM}}[n] = I_{\text{AMB}} + hg(I_{\text{PRJ}}[n]). \quad (38)$$

The second term in Eq. (38) may be decomposed using the Fourier series expansion,

$$\begin{aligned} I_{\text{CAM}}[n] &= I_{\text{AMB}} + hg\left(\frac{1}{2}I_0\right) \\ &\quad + \sum_{k=0}^{+\infty} c_k \cos(\omega x_{\text{PRJ}} - (k+1)\varphi[n]), \end{aligned} \quad (39)$$

where c_k are the coefficients of Fourier series expansion, i.e. c_0 is the amplitude of the fundamental harmonic of the projected fringe and c_k , $k > 0$, are amplitudes of harmonic overtones. The value of coefficients c_k depends on $g(\cdot)$. In practice there usually exists some finite k_0 such that values of all c_k for $k > k_0$ are almost zero. Hence, the harmonics overtones for all $k > k_0$ have a negligible contribution and may be discarded from the sum.

Consider now a case of multiple projectors where phase shifts $\varphi_p[n]$ of p th projector are chosen using Eq. (29). Then Eq. (20) becomes

$$\begin{aligned} I_{\text{CAM}}[n] &= a + \sum_{p=1}^P c_{0,p} \cos(\omega_p x_{\text{PRJ},p}) \cos(\varphi_p[n]) \\ &\quad + \sum_{p=1}^P c_{0,p} \sin(\omega_p x_{\text{PRJ},p}) \sin(\varphi_p[n]) \\ &\quad + \sum_{k=1}^{k_0} \sum_{p=1}^P c_{k,p} \cos(\omega_p x_{\text{PRJ},p} - (k+1)\varphi_p[n]). \end{aligned} \quad (40)$$

In Eq. (40) coefficients $c_{0,p}$ correspond to contrasts b_p of Eq. (20). Coefficients $c_{k,p}$, $1 \leq k \leq k_0$, are unwanted harmonic overtones in the mixture which are normally called *harmonic distortion*. One additional issue of harmonic distortion in multi-projector systems which is not present in single projector systems is that harmonic distortion of a particular projector is not limited to itself, instead it affects all other projectors as well. For example, in two-projector system the choice of parameters given by Eq. (28) causes the first harmonic overtone of the first projector $c_{1,1}$ to affect the fundamental harmonic of the second projector $c_{0,2}$ due to $2\Omega_1 n = 2\frac{2\pi}{2P+1}n = \Omega_2 n$. Even worse, for a finite number of N frames all harmonic overtones whose temporal frequency Ω is larger than π will *alias* into the baseband making determination of affected fundamental harmonics a non-trivial task.

Consider a general case of an overlap between one fundamental harmonic and one harmonic overtone for N phase shifts (frames) and P projectors. Let Ω_{p_1} and Ω_{p_2} be temporal frequencies of phase shifts for two projectors p_1 and p_2 .² Then, the k th harmonic overtone of p_1 will affect the fundamental harmonic of p_2 if

$$(k+1)\Omega_{p_1} \equiv \pm\Omega_{p_2} \pmod{2\pi}, \quad 1 \leq k \leq k_0 \quad (41)$$

holds. Expanding this condition using Eq. (28) yields two Diophantine equations,

$$(k+1)p_1 - lN = p_2 \quad \text{and} \quad (k+1)(N-p_1) - lN = p_2. \quad (42)$$

²Note that there is no requirement that $p_1 \neq p_2$, i.e. a harmonic overtone may also affect its own fundamental harmonic.

If Eq. (42) has a solution (k, l) in whole numbers for $1 \leq k \leq k_0$ then k th harmonic overtone of projector p_1 affects the fundamental harmonic of projector p_2 .

By repeating Eq. (42) for all possible projector pairs we obtain a system of Diophantine equations [59] which, to achieve the insensitivity to the first k_0 harmonics, must have no solutions in whole numbers for $1 \leq k \leq k_0$. One way to achieve such insensitivity is to increase the number of projected fringes/recorded frames N from the minimal value of $2P + 1$ and to neglect the temporal frequency assignment given by Eq. (29). Then the temporal frequencies Ω_p of phase shifts may be selected to shift the unwanted harmonic overtones into the unused³ basis vectors of the discrete Fourier transform basis. Therefore, to achieve robustness to interference a separation between selected frequencies must be maintained.

Unfortunately, an analytical solution to Eq. (42) is not known to us (if it even exists), so we can only solve the frequency selection numerically. An example for three projectors, $P = 3$, and $k_0 = 1$: a particular N for which Eq. (42) has no solution for $1 < k \leq k_0$ in all projector-pairs is $N = 12$ using temporal frequencies $\Omega_1 = 1 \frac{2\pi}{N}$, $\Omega_2 = 3 \frac{2\pi}{N}$, and $\Omega_3 = 5 \frac{2\pi}{N}$.

III. RESULTS AND DISCUSSION

To demonstrate the applicability the proposed approach we have constructed a prototype scanner comprised of three projectors denoted $P1$ – $P3$ and of six cameras denoted $C1$ – $C6$. The prototype is shown in Fig. 1. The first projector $P1$ is Canon LV-WX310ST while the remaining two, $P2$ and $P3$, are Acer S1383WHne. All cameras are PointGrey's Grasshopper3 GS3-U3-23S6C-C; the first four cameras $C1$ – $C4$ are equipped with Fujinon HF12.5SA-1 lenses while the remaining two, $C5$ and $C6$, are equipped with Kowa LM8JCM lenses. All projectors and cameras are operated at their native resolutions: 1280×800 for projectors and 1920×1200 for cameras.

The architecture of the assembled prototype using three vertical poles on wheeled carts provides flexibility when testing various spatial arrangements. We have used two spatial arrangements. In the first arrangement all three carts were placed directly in front of an object to maximize the overlap between all projectors (Fig. 1). In the second arrangement each cart was placed on a vertex of (an equilateral) triangle with the object in the triangle's center. The first arrangement which maximizes inter-projector interference is used to test the proposed separation approach. The second arrangement which maximizes surface coverage is used to showcase how the (almost) complete surface may be acquired in a single scan.

³Using $2P+1$ projected frames with P projectors uses up all basis vectors, however increasing the number of projected frames introduces additional basis vectors which are not assigned to any projector and into which we may shift unwanted harmonics.

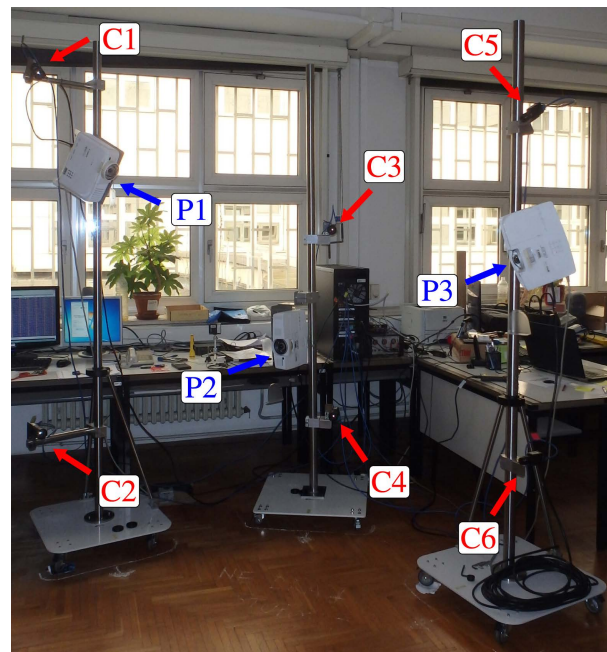


FIGURE 1. Prototype multi-projector multi-camera structured light 3D scanner comprised of three projectors and six cameras mounted on three vertical poles on wheeled carts. One projector and one camera-pair are always mounted on the same vertical pole as follows: $C1$ – $P1$ – $C2$ on the first pole, $C3$ – $P2$ – $C4$ on the second, and $C5$ – $P3$ – $C6$ on the third.

The remainder of this section is organized as follows: We first briefly explain how the scanner was calibrated, starting with geometric calibration and continuing with photometric calibration. Then, we present the results of spectrum estimation for both photometrically calibrated and uncalibrated projectors to demonstrate that the photometric calibration is necessary and cannot be skipped. Next, we present experiments to demonstrate the effectiveness of the proposed signal separation scheme and experiments to showcase a single scan surface reconstruction. Finally, we conclude this Section with a discussion.

A. CALIBRATION

The 3D surface scanner must be calibrated prior to use. The calibration is usually separated into *geometric* and *photometric* calibrations which may be performed independently. The purpose of *geometric* calibration is to retrieve the intrinsic and extrinsic parameters of all projectors and cameras. The purpose of *photometric* calibration is to retrieve the intensity transfer functions for all projectors and cameras. For our prototype (Fig. 1) the *photometric* calibration is performed once while *geometric* calibration must be performed each time three carts carrying projectors and cameras are rearranged in space.

1) GEOMETRIC CALIBRATION

We have performed geometric calibration exactly as explained in our previous work [60].

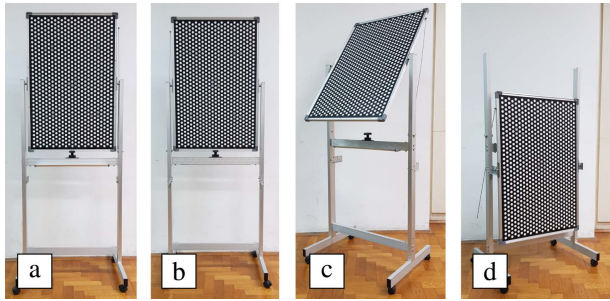


FIGURE 2. Double-sided planar calibration board: (a) front side; (b) back side; (c) tilted; and (d) lowered. Note that calibration board contains markings to automatically determine which side is observed.

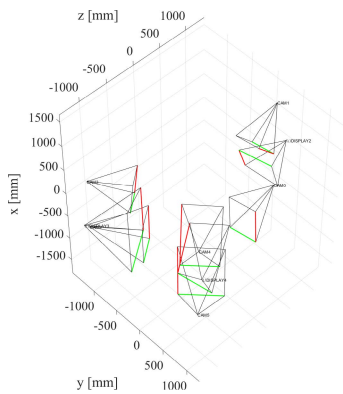


FIGURE 3. Spatial arrangement of cameras and projectors in a common coordinate system.

Briefly, projectors are modeled as inverse cameras and the approach of Zhang and Huang [61] is used to enable each projector to “capture” the images of a calibration board like a camera, thus reducing the calibration problem to the well known case of calibrating a multiple-camera system.

The calibration object is a double-sided calibration board which is shown in Fig. 2 and which was assembled by gluing two large vinyl stickers containing the calibration pattern onto both sides of a whiteboard mounted on wheels. A rigid transformation (rotation and translation) between the front and the back calibration pattern was then determined using MicroScribe G2LX digitizing arm. To calibrate the system in practice the double-sided calibration board is recorded in many positions. The coordinates of white circles are then automatically extracted the whole system is calibrated as a multi-camera system using bundle adjustment [50], [62]. The final result of geometric calibration are intrinsic and extrinsic parameters of each camera and projector, where extrinsic parameters refer to a common coordinate system. A calibration result for the second spatial arrangement where projectors and cameras surround an area of interest is shown in Fig. 3.

2) PHOTOMETRIC CALIBRATION

We have performed the photometric calibration of the system to correct the non-linear intensity transfer function of the

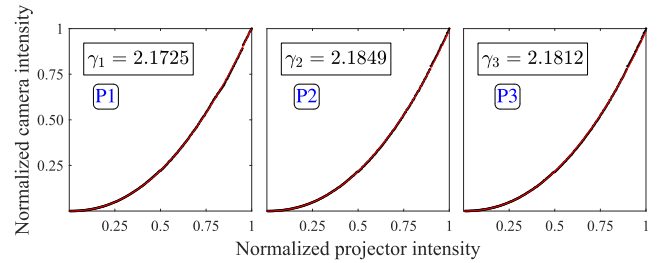


FIGURE 4. Intensity transfer functions for each of three projectors using least-squares power-law fitting. Measured samples are shown as black dots and a fitted power-law model is shown as a thin red line which overlap almost perfectly. The value of factor γ for each projector is indicated in the upper left corner.

projectors, specifically to compensate for projector’s unknown gamma factors. For each projector we must measure its gamma factor γ . Once γ is known a pre-corrected image

$$I_{PRJ}(x_{PRJ}, y_{PRJ}) = I_0 \left(\frac{1}{2} + \frac{1}{2} \cos(\omega x_{PRJ} - \varphi[n]) \right)^{1/\gamma} \quad (43)$$

is sent to the projector to compensate for projector’s gamma transformation [55].

The factor γ of Eq. (43) may be determined from the intensity transfer function using the least-squares power-law fitting [63] to determine the parameters a and γ of the model $I_{CAM} = aI_{PRJ}^\gamma$. The gray-level intensity transfer function of each projector was sampled in 255 points by recording a planar white board in a dark room. To this purpose we have used PointGrey’s Dragonfly2 DR2-HICOL camera which was operated in 12 bit acquisition mode. This camera was intentionally used for the intensity calibration of the projectors as it has better intensity resolution than Grasshopper3 GS3-U3-23S6C-C. After the least-squares power-law fitting (see Fig. 4) the estimated values of γ factors for three used projectors are: $\gamma_1 = 2.1725$ for **P1**, $\gamma_2 = 2.1849$ for **P2**, and $\gamma_3 = 2.1812$ for **P3**.

B. PHOTOMETRIC CALIBRATION AND HARMONIC DISTORTION

To evaluate the effectiveness of the photometric calibration regarding harmonic distortion we have estimated temporal power spectral densities (PSD) using Bartlett’s method [64]. Temporal PSD were estimated for each projector separately using a white board and PointGrey’s Dragonfly2 DR2-HICOL camera which was operated in 12bit acquisition mode.

To estimate the temporal PSD of a sinusoidal fringe pattern comprised of N phase shifts using some temporal frequency Ω we have to record a total of $N \cdot M$, $M \gg N$, frames using temporal phase shifts $\varphi[n] = \Omega(n - 1)$, $n = 1, \dots, N \cdot M$. This produces an ensemble of temporal signals, one temporal signal for each camera pixel, all having exactly $N \cdot M$ samples. Applying the Bartlett’s method of periodogram averaging [64] to such ensemble produces the desired temporal PSD estimate. Note that the temporal PSD estimate obtained in this way is invariant with regard to the spatial fringe frequency ω .

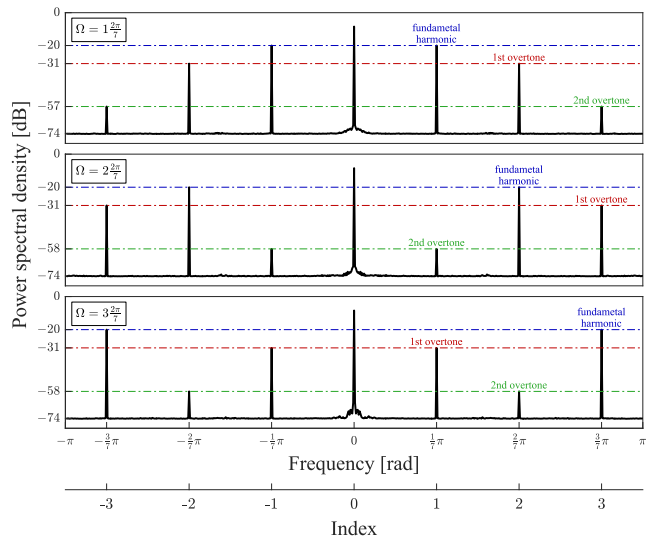


FIGURE 5. Temporal PSD for Acer S1383WHne (P3) without compensating for projector's gamma. Note the 1st harmonic overtone is quite prominent at about -11 dB compared to the fundamental harmonic. Also note the aliasing of harmonic overtones.

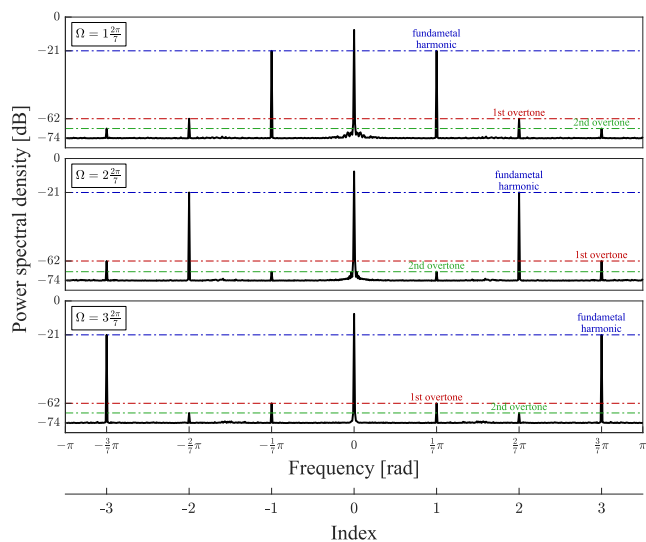


FIGURE 6. Temporal PSD for Acer S1383WHne (P3) with applied compensation for projector's gamma. Note the significant reduction in the power of harmonic overtones compared to Fig. 5.

We have performed two different experiments to evaluate the effect of harmonic distortion caused by projector's gamma: (1) estimation of temporal PSD for the case of $N = 7$ phase shifts using $M = 90$; and (2) estimation of temporal PSD for the case of $N = 12$ phase shifts using $M = 70$. In both experiments temporal PSDs were estimated twice, first without compensating for projector's gamma and then with compensation given by Eq. (43). Also note that values $M = 90$ and $M = 70$ are only used to obtain reliable PSD estimates; such a huge number of images is not required for normal 3D scanning.

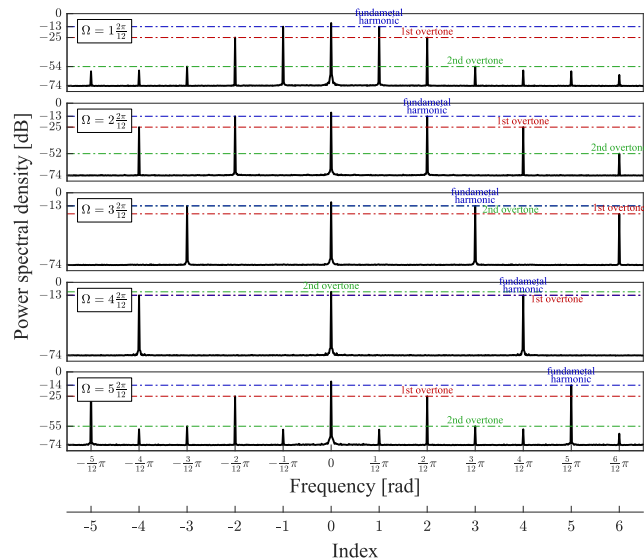


FIGURE 7. Temporal PSD for Cannon LV-WX310ST (P1) without compensating for projector's gamma. Note the prominence of 1st harmonic overtone compared to the fundamental harmonic. Also note the aliasing of harmonic overtones.

TABLE 1. Attenuation of 1st and 2nd harmonic overtones compared to the fundamental harmonic without compensation for projector's gamma.

Projector	Temporal frequency	Attenuation [dB]	
		1st overtone	2nd overtone
(P1)	$\Omega = 1 \frac{2\pi}{7}$	-11.18	-39.92
	$\Omega = 2 \frac{2\pi}{7}$	-11.18	-39.88
	$\Omega = 3 \frac{2\pi}{7}$	-11.18	-40.03
(P2)	$\Omega = 1 \frac{2\pi}{7}$	-11.01	-37.21
	$\Omega = 2 \frac{2\pi}{7}$	-11.03	-37.30
	$\Omega = 3 \frac{2\pi}{7}$	-11.04	-37.47
(P3)	$\Omega = 1 \frac{2\pi}{7}$	-11.07	-37.34
	$\Omega = 2 \frac{2\pi}{7}$	-11.08	-37.46
	$\Omega = 3 \frac{2\pi}{7}$	-11.09	-37.52

1) PSD FOR $N = 7$ PHASE SHIFTS

Three temporal PSDs for $N = 7$ without compensating for projector's gamma are shown in Fig. 5; all are for (P3) using three available choices of temporal frequency Ω . Note that 1st harmonic overtone is always attenuated about -11 dB compared to the fundamental harmonic. In terms of amplitudes used in Eqs. (39) and (40) such attenuation gives $c_1/c_0 \approx 0.28$, i.e. 1st harmonic overtone is significant as its amplitude is 28% of the amplitude of the fundamental harmonic. The 2nd harmonic overtone is about -37 dB attenuated ($c_2/c_0 \approx 0.014$) and may be neglected. A complete list of attenuations for 1st and 2nd harmonic overtones compared to the fundamental harmonic for all three projectors is listed in Table 1.

Three temporal PSDs for $N = 7$ after compensating for projector's gamma are shown in Fig. 6 for projector (P3) only while the list of attenuations for 1st and 2nd harmonic

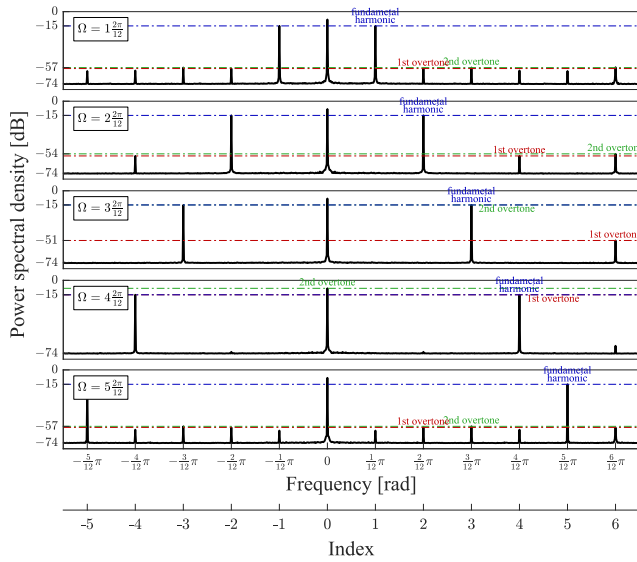


FIGURE 8. Temporal PSD for Canon LV-WX310ST (P1) with applied compensation for projector's gamma. Note the significant reduction in the power of harmonic overtones compared to Fig. 7.

TABLE 2. Attenuation of 1st and 2nd harmonic overtones compared to the fundamental harmonic with applied compensation for projector's gamma.

Projector	Temporal frequency	Attenuation [dB]	
		1st overtone	2nd overtone
P1	$\Omega = 1 \frac{2\pi}{7}$	-42.07	-41.46
	$\Omega = 2 \frac{2\pi}{7}$	-41.82	-41.62
	$\Omega = 3 \frac{2\pi}{7}$	-41.65	-41.40
P2	$\Omega = 1 \frac{2\pi}{7}$	-42.91	-48.75
	$\Omega = 2 \frac{2\pi}{7}$	-42.61	-49.02
	$\Omega = 3 \frac{2\pi}{7}$	-42.86	-48.53
P3	$\Omega = 1 \frac{2\pi}{7}$	-41.55	-47.55
	$\Omega = 2 \frac{2\pi}{7}$	-41.62	-48.12
	$\Omega = 3 \frac{2\pi}{7}$	-41.58	-47.45

overtones compared to the fundamental harmonic for all three projectors is given in Table 2. Compensating for projector's gamma adds on average -30 dB of attenuation to the first harmonic overtone, i.e. after compensating for projector's gamma we have $c_1/c_0 < 9\%$ which is a significant improvement. Note that this result is consistent across all three projectors.

2) PSD FOR $N = 12$ PHASE SHIFTS

Five temporal PSDs for $N = 12$ are shown in Fig. 7 without compensating for projector's gamma and in Fig. 8 with compensation for projector's gamma; all PSDs are for P1 using five available choices of temporal frequency Ω . Similarly to the first experiment note the significant reduction in the power of harmonic overtones after the compensation for projector's gamma given by Eq. (43) is applied.

Of particular interest is the topmost PSD in Fig. 7 which gives attenuations of for first four harmonic overtones: -11.16 dB, -41.01 dB, -44.98 dB, and -45.82 dB. The most significant is the first harmonic overtone while other three are quite similar. This means we may in practice choose $k_0 = 1$ to cut the Fourier series expansion of Eq. (40) (Section II-C2). Also note that the choice of temporal frequency $\Omega = 4 \frac{2\pi}{12}$ for $N = 12$ is particularly bad as then the first harmonic overtone overlaps its own fundamental harmonic due to aliasing.

The case of $N = 12$ phase shifts is interesting as observed PSDs confirm that for the usual gamma factor a reliable separation can be achieved by shifting the dominant first harmonic into unused basis vectors as explained in the last paragraph of Section II-C2, which achieves no overlap between all possible fundamental harmonic and 1st overtone pairs.

C. SEPARATION OF PROJECTED PATTERNS

To showcase and test the proposed separation of projected patterns we have performed several experiments using the first arrangement of projectors which maximizes the unwanted effects of inter-projector interference. The subjects which were scanned are: a human male, a calibration board, and a mannequin.

To perform these experiments we have to choose a spatial frequency and select temporal phase shifts for each projector. We have selected three spatial frequencies, $\omega_A = 2\pi \frac{21}{W}$, $\omega_B = 2\pi \frac{31}{W}$, and $\omega_C = 2\pi \frac{15}{W}$, where $W = 1280$ px is projector width, with $N = 7$. To each of these three spatial frequencies we have assigned its own temporal phase shift sequence: $\varphi_A[n] = 2 \frac{2\pi}{7} n$, $\varphi_B[n] = 3 \frac{2\pi}{7} n$, and $\varphi_C[n] = 1 \frac{2\pi}{7} n$. These spatial frequencies and phase shifts are then assigned to projectors in a round-robin fashion: first P1 projects A, P2 projects B, and P3 projects C; then P1 projects C, P2 projects A, and P3 projects B; and finally P1 projects B, P2 projects C, and P3 projects A. This means each camera acquires a set of $7 \cdot 3 = 21$ frames. Such choice allows us to use multiple-phase shifting (MPS) unwrapping strategy [48] to recover projector coordinate from three wrapped phases at different spatial frequencies.

1) SEPARATION EXAMPLE

A scan of a human male is used as an example of the separation process. Each of three projectors ($P = 3$) projects seven patterns ($2P + 1 = 7$) which are in turn recorded by six cameras, so each camera concurrently [65] acquires exactly seven images. Separation of recorded images is performed as explained in Section II-B for each camera separately, and the separation result for one camera is shown in Fig. 9. Note that the obtained decomposition yields seven images. The first one is a constant component which is the combined contribution of ambient illumination and of averaged illumination of each of three used projectors. The remaining six images contain amplitudes and wrapped phases for each of three projectors. Note that in Fig. 9 the intensity scaling for components a ,

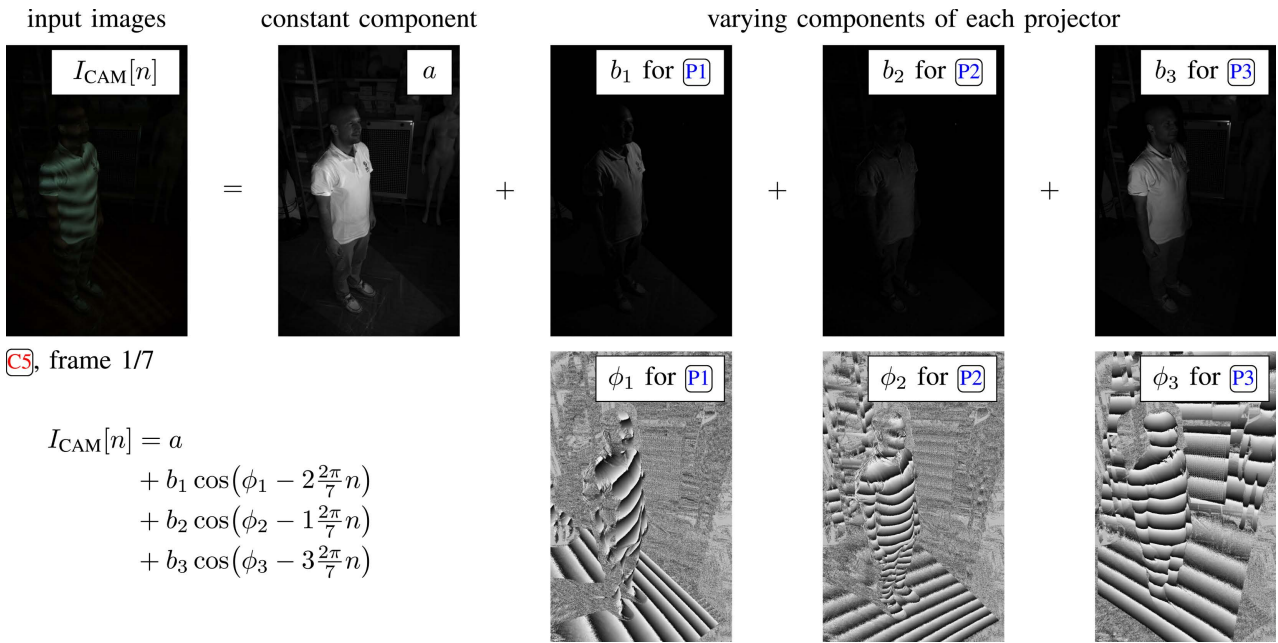


FIGURE 9. Separation of projector contributions for **CS**. Seven input frames are decomposed using the proposed method to obtain contributions of each individual projector. Note that ϕ_1 has spatial frequency $\omega_A = 2\pi\frac{21}{W}$, ϕ_2 has $\omega_B = 2\pi\frac{31}{W}$, and ϕ_3 has $\omega_C = 2\pi\frac{15}{W}$.

21 input frames in three groups of 7 frames are first decomposed into nine wrapped phases from which x_{PRJ} is computed.

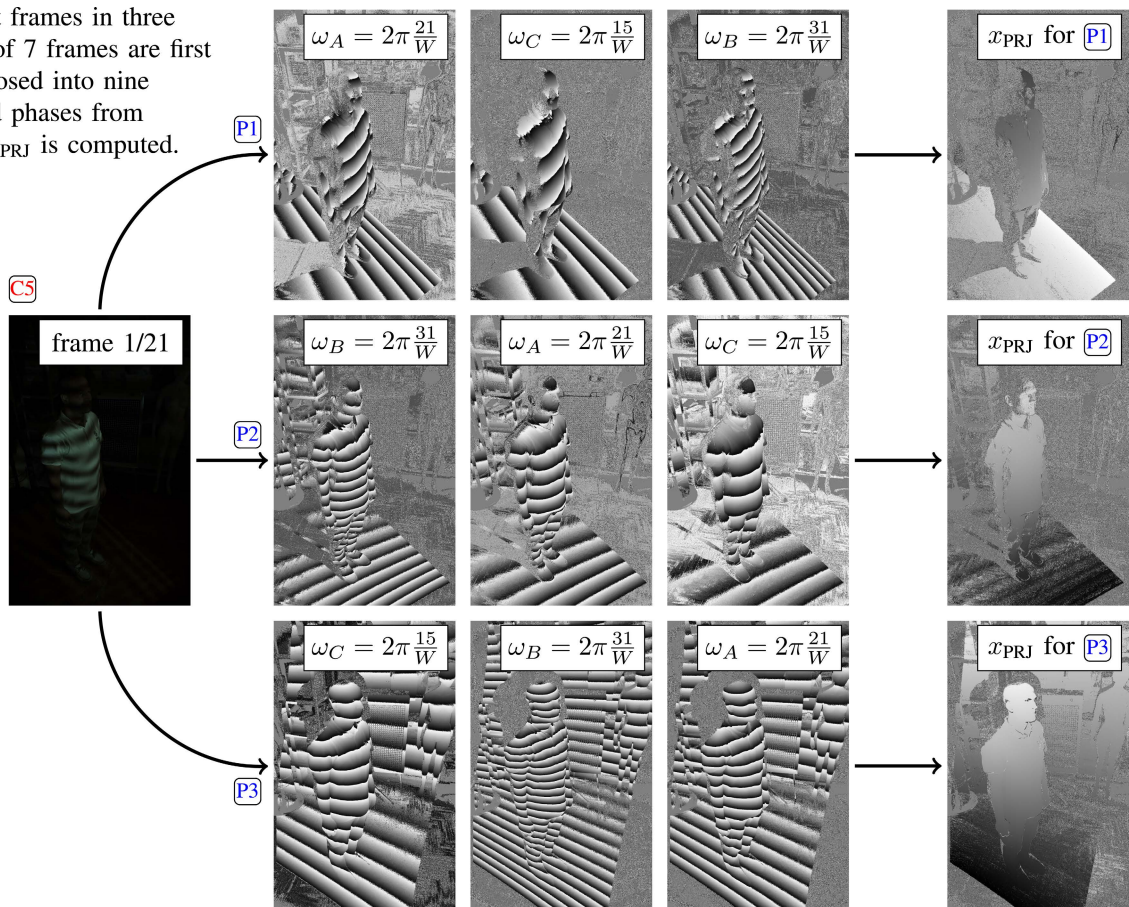


FIGURE 10. Recovered wrapped phases and projector coordinates for **CS**. Note different shadows in the projector coordinate x_{PRJ} for three projectors.

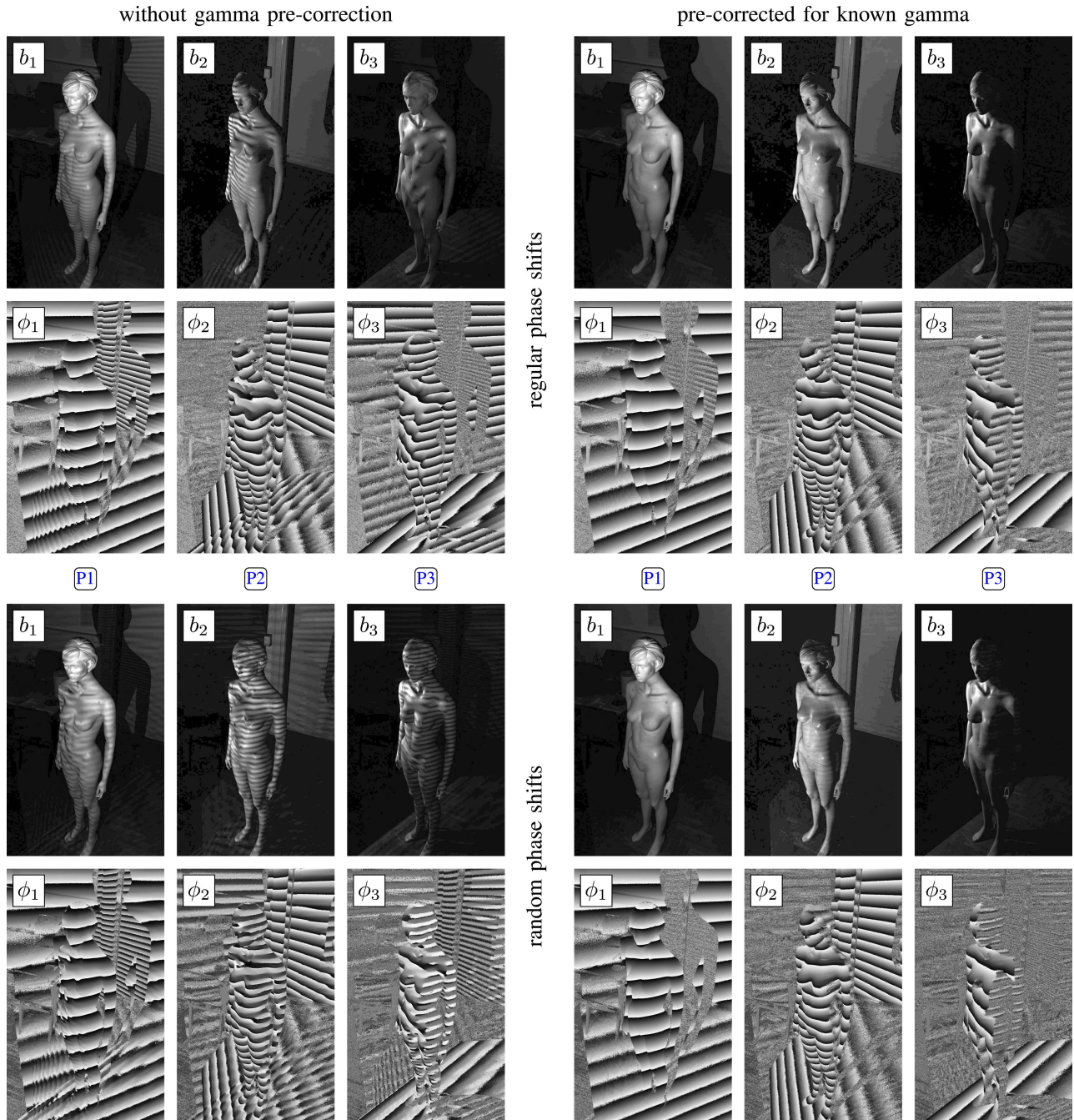


FIGURE 11. Separation of projector contributions for $C1$ into amplitudes and wrapped phases with and without gamma pre-correction when using seven SL patterns. All amplitude images $b_{1,2,3}$ are linearly contrast-stretched to make the interference more noticeable. Three leftmost columns show separation without gamma pre-correction: amplitudes $b_{1,2,3}$ are all affected by other components due to interference and wrapped phases $\phi_{1,2,3}$ exhibit unwanted artifacts. Three rightmost columns show separation with applied gamma pre-correction: amplitude components are free of interference and wrapped phases are correct.

b_1 , b_2 and b_3 and the input image is the same, and due to this a appears much brighter than individual components b_1 , b_2 and b_3 (as it is their sum). Phases ϕ_1 , ϕ_2 and ϕ_3 are wrapped phases and have to be unwrapped to obtain projector coordinates, which is done by recording another set of images using different frequencies in a round robing fashion as explained earlier and as illustrated in Fig. 10. The phases are unwrapped using the method of [48].

2) REGULAR AND RANDOM PHASE SHIFTS

To showcase how separation performs for regular and for random phase shifts we use a mannequin which is stationary during multiple acquisitions thus facilitating an easier qualitative comparison.

The mannequin was recorded four times in the same position using the minimal number of SL patterns: (i) using regular phase shifts without gamma pre-correction; (ii) using

regular phase shifts with gamma pre-correction; (iii) using random phase shifts without gamma pre-correction; and (iv) using random phase shifts with gamma pre-correction. Separation results are shown in Fig. (11). Note that the amplitude components b_1 , b_2 and b_3 should in the ideal case depend only on the output illumination of the projector and on the albedo of the object; they should not contain visible traces of the fringe patterns projected by other projectors. However, if projector's non-linear amplitude transfer function is not accounted for using the gamma pre-correction then this causes interference which affect both amplitudes and wrapped phases. The unwanted interference in amplitude images manifests as a fringe pattern having a different spatial frequency, and in wrapped phase images as an unwanted ripples.

Also note that the robustness of separation scheme for random phase shifts may depend on the selected phase values, i.e. phase shift values which are not evenly spread over the whole 2π interval of wrapped phases may lead to numerical instabilities in Eq. (24), and which is a drawback of a random choice. However, one advantage of randomly selected phase shifts is that they may be independently selected while retaining a high probability that separation is possible, which may be of use when e.g. a projector is added to an existing closed system.

Considering results shown in Fig. (11) and taking into considerations the observed temporal PSDs to reduce the effects of unwanted harmonic distortion all projectors in a multi-projector SL scanner must be photometrically calibrated to project a clear sinusoid with harmonic overtones maximally suppressed. If such photometric calibration is for any reason not possible then surface scanning may be performed using $N > 2P + 1$ and $k_0 \geq 1$, i.e. unwanted harmonic components may be shifted into spectral bands which are unused as explained in Section II-C2. An example is shown in Fig. 12 which demonstrates that unwanted artifacts can be significantly suppressed by frequency and phase shift selection which places dominant overtones into unused basis vectors. We also note that such an approach will not work with randomly selected phase shifts as the random selection causes spectral leakage, i.e. if random shifts or random frequencies are used then all projectors must be photometrically calibrated.

3) QUANTITATIVE RESULTS

A flat board (shown in Fig. 13) and a mannequin are used for quantitative comparison between sequential and simultaneous acquisitions due to the fact that they are stationary and are easily recorded multiple times, allowing for a straightforward comparison between different SL patterns. As the most important output of the proposed method are wrapped phases we compare them using the absolute circular distance $d_{\text{circular}}(\cdot, \cdot)$ as the error metric:

$$d_{\text{circular}}(\phi, \phi_{\text{MPS}}) = \left| \text{Arg}(e^{i\phi} / e^{i\phi_{\text{MPS}}}) \right|, \quad (44)$$

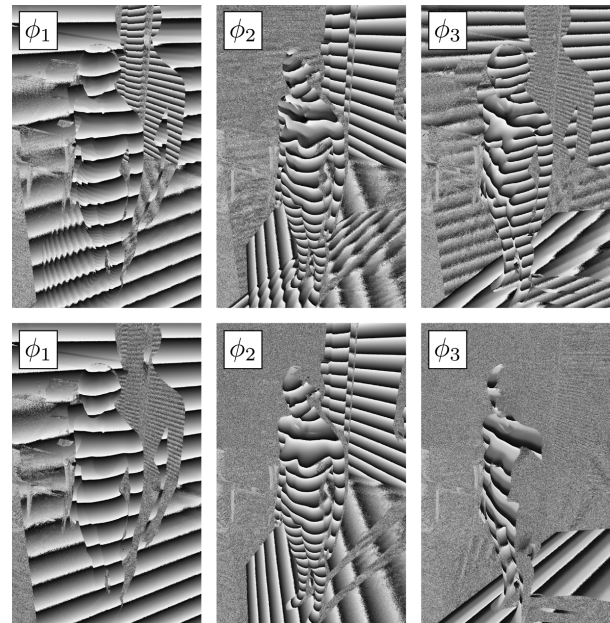


FIGURE 12. Wrapped phases using 12 regular shifts and no gamma pre-correction. Top row shows wrapped phases using standard frequency and phase shift assignment of Eqs. (29) and (30). Bottom row shows wrapped phases using proposed assignment of Section II-C2. Note the missing artifacts for the bottom row.

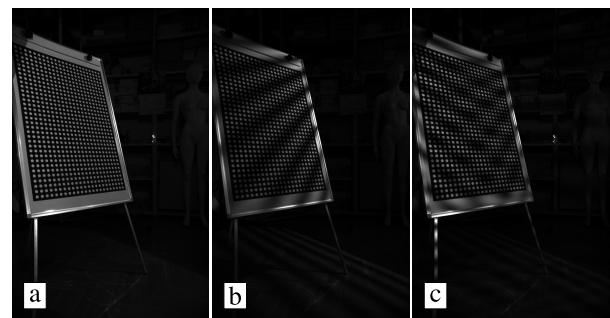


FIGURE 13. A flat board as seen by CS: (a) board under white illumination; (b) board under illumination by P1 only using the spatial frequency $\omega_A = 2\pi \frac{21}{W}$; and (c) board under illumination by all three projectors simultaneously. Note the difference between the pure fringe in (b) and the multi-projector interference pattern in (c).

where ϕ is the wrapped phase obtained using Eq. (36) and ϕ_{MPS} is the wrapped phase obtained using a classical sequential approach which is taken to be the ground truth. The error metric was computed for the illuminated pixels only using the threshold T of 5%, see Eq. (14).

Average errors in the form of mean and deviation are listed in Table 3 for the calibration board and in Table 4 for the mannequin. Listed errors are averages over all cameras for all projectors; the results are identical for individual camera and projector pairs so we omit those. Sequences denoted with “regular” use phase shifts listed at the start of Section III-C, while “random” sequences use random phase shifts selected from the uniform distribution over a 2π interval. Note that errors in the wrapped phase are between 1° and 2° and more

TABLE 3. Errors for calibration board in degrees and in pixels for four different sequences with gamma pre-correction.

Sequence	Wrapped phase error [°]	Coordinate error [px]
$N = 7$, regular	1.52 ± 1.45	0.17 ± 1.35
$N = 7$, random	2.07 ± 2.04	0.38 ± 5.92
$N = 12$, regular	1.13 ± 0.96	0.13 ± 1.10
$N = 12$, random	1.36 ± 1.26	0.20 ± 2.92

TABLE 4. Errors for mannequin in degrees and in pixels for four different sequences with gamma pre-correction.

Sequence	Wrapped phase error [°]	Coordinate error [px]
$N = 7$, regular	1.56 ± 1.30	0.16 ± 0.19
$N = 7$, random	2.00 ± 1.70	0.21 ± 0.48
$N = 12$, regular	1.17 ± 0.97	0.12 ± 0.22
$N = 12$, random	1.32 ± 1.10	0.14 ± 0.26

TABLE 5. Errors for mannequin in degrees and in pixels for two different sequences without gamma pre-correction.

Sequence	Wrapped phase error [°]	Coordinate error [px]
$N = 12$, regular	1.10 ± 0.92	0.13 ± 1.58
$N = 12$, random	3.94 ± 4.43	3.60 ± 29.53

importantly, when transformed to errors in decoded projector coordinate in pixels, all listed mean errors are in the subpixel range. Note that the errors for the mannequin have a somewhat lower spread than the errors for the calibration board, which is probably due to different albedo: the mannequin is uniformly white while the calibration board is not.

We have also performed the experiment with and without gamma pre-correction for which errors are listed in Table 5. Note that errors for the $N = 12$ using strategy of Section II-C2 are acceptable for “regular” phase shifts.

The last experiment is the least-squares fitting of a plane to the reconstructed point cloud of the calibration board. The median absolute distance of the reconstructed 3D points from the fitted plane is 0.97 mm, the average is 1.45 mm with the deviation 1.61 mm. Considering the scanner’s working space which spans several meters the error in the millimeter range is more than acceptable.

D. 3D RECONSTRUCTION

Once projector coordinates are known for each camera a 3D point cloud is obtained via triangulation [50] and no additional registration [66] is required.

Two examples of obtained point clouds are shown in Fig. 14 where points corresponding to different cameras are rendered using different randomly assigned colors. The top

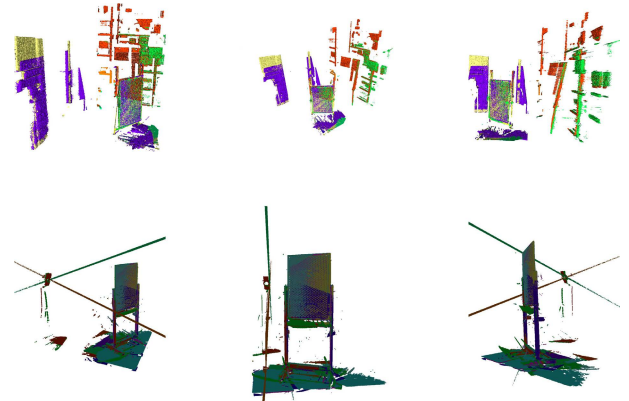


FIGURE 14. Colored point clouds showing contribution of each camera in different hue. Top row: three views of a calibration board when all projectors and cameras face the same way. Bottom row: three views of a calibration board with projectors and cameras surrounding the board.

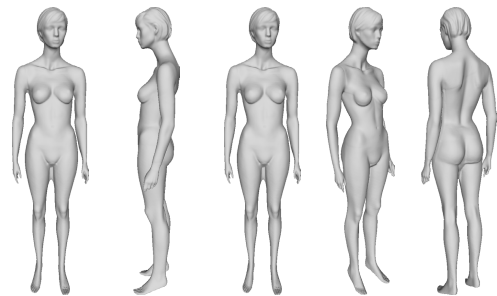


FIGURE 15. Several views of a reconstructed 3D mesh of a mannequin.

row in Fig. 14 shows reconstructed point clouds for the first spatial arrangement of wheeled carts where all projectors and cameras are facing in the same direction. The bottom row in Fig. 14 shows reconstructed point clouds for the second arrangement where wheeled carts surround the calibration board. Note that regardless of the spatial arrangement of wheeled carts contributions of each projector and camera pair are obtained in the same coordinate system which makes registration unnecessary. Also note two artifacts for the second spatial arrangements where two cameras $C1$ and $C2$ directly observe the lens of the opposing projector $P3$. A direct observation of projector’s lens produces an oversaturated patch in the image. For such an oversaturated patch camera coordinates are well defined but decoded projector coordinates are essentially randomly distributed and cannot be filtered out via thresholding using Eq. (14). As a consequence reconstructed points corresponding to such patches are constrained to a narrow set of camera rays. To avoid such unwanted artifacts position and/or orientation of projectors should be adjusted so no camera directly observes the lens of an opposing projector, or such patches should be identified during calibration and masked out.

Once point clouds are obtained the unwanted parts of the scan may be cut out and a surface may be produced using a standard meshing software such as MeshLab [67], [68]. One reconstruction example for a mannequin is shown in Fig. 15.

Regarding the scanning time our prototype scanner operates at 20 FPS [43], [65]. Therefore, if using ray-plane triangulation then three frequencies with seven shifts per frequency are required giving a total of 21 image to be projected and acquired, so the acquisition time is 1.05 seconds. The processing time for pattern decoding and triangulation is about 1 second per camera-projector pair, so for six camera three projector system the total image processing time is a bit less than half a minute. Note that reported times are for C++ implementation of the image acquisition, while image processing and 3D reconstruction is implemented in Matlab.

E. DISCUSSION

The proposed method uses sinusoidal fringes with a relatively simple and straightforward assignment of spatial frequencies and temporal phase shifts which offers many advantages as discussed before. Additionally, we note that sinusoidal fringes are naturally robust to projector and camera blur as the phase is usually unaffected by blurring due to the spatial symmetry of the blurring optical transfer function [69]. This is a highly desirable property and is a clear advantage of the proposed multi-projector FPP approach over non-FPP multi-projector approaches such as [34], [40], [70].

A possible drawback of the proposed method is the limited dynamic range of cameras, especially if all projectors are positioned to have overlapping FOVs. Camera pixels which observe an overlap of projector's FOVs receive more light, and if there are many projectors then overlap may require a reduction in the light sensitivity of a camera. On the other hand, reducing the light sensitivity of a camera to accommodate for increased illumination affects the signal-to-noise (SNR) ratio for pixels which are illuminated by a single projector only, and for pixels which observe parts of the object which are dark and do not reflect light well. For the proposed approach camera's available dynamic range must be shared between all projector's which particular camera observes, and this in turn defines an upper limit to the number of projectors which may simultaneously illuminate any part of the scene. Note that in practice an overlap of more than three projectors is not expected as we naturally want such a spatial arrangement which maximizes the coverage of the working space which is in turn achieved when FOV overlap is minimized. The available dynamic range can be estimated using estimated PSDs. Observe in Figs. 6, 5, 7 and 8 the existence of the noise floor which is at about -65 dB compared to the DC component. Also observe that after compensating projector's gamma almost all harmonic overtones are at about -42 dB compared to the DC component. Taking into consideration a standard formula for signal-to-noise ratio [71], $SNR = 6.02B + 1.76$ dB, where B is the number of bits, we obtain $B = 10.5$ bits for the camera and $B = 6.68$ for the projector, which is consistent with the equipment used as Dragonfly2 camera is declared to use 12 bits, and projectors to use 8 bits (note that we can never utilize the whole theoretical dynamic range). From our experience 4 to 5 bits

of contrast are required per projector which makes a standard 8 bit camera capable of handling three projectors easily.

Recovered 3D surface points may be textured using the value of a as the gray-level intensity. Furthermore, if cameras are recording in color then the DC spectral coefficient a may be computed for each color channel separately yielding a RGB texture with the sinusoidal fringe removed. To get an even better rendering of the final mesh note that photometric stereo [72] may be used to recover surface normals. Consider three images showing the recovered amplitudes b_k in Fig. 9 where it is clearly observable that the direction of incident illumination depends on the position of a particular projector. If a sufficient number of viewing directions are known then surface normals may be recovered from the coefficients b_k using photometric stereo with an additional advantage that view directions of camera and of projector are known in advance due to geometric calibration. Therefore, by introducing additional data processing the proposed method should be able to acquire both texture and normal information for each reconstructed 3D point with no need to project additional frames such as e.g. an all-white pattern or to acquire images with projectors turned off, and this is a future work we will pursue

IV. CONCLUSION

We have proposed a method for efficient SL scanning which is applicable to multi-projector multi-camera SL scanners. The proposed method is based on a computational separation between projected sinusoidal fringe patterns which is enabled by specific choices of temporal phase shifts: a different set of temporal phase shifts is assigned to each projector in a way which enables efficient and/or robust separation of the observed combined patterns into the contributions of each projector.

We have shown that for a system having P projectors the minimal number of phase shifts required for separation is $2P + 1$. Considering specific strategies to select the required phase shifts we have discussed two options, regular phase shifts and random phase shifts. The choice of regular phase shifts is such that phase shifts define basis vectors of the discrete Fourier transform in $2P + 1$ (or more) points which in turn enables an efficient separation. The choice of random phase shifts is one where phase shifts are chosen from the uniform distribution over a 2π interval.

The proposed method enables construction of large FPP systems comprised of many projectors and cameras which can obtain the whole surface profile of an object in a single scan.

REFERENCES

- [1] J. Geng, "Structured-light 3D surface imaging: A tutorial," *Adv. Opt. Photon.*, vol. 3, no. 3, pp. 128–160, Mar. 2011. [Online]. Available: <http://aop.osa.org/abstract.cfm?URI=aop-3-2-128>
- [2] G. Sansoni, M. Trebeschi, and F. Docchio, "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation," *Sensors*, vol. 9, no. 1, pp. 568–601, 2009. [Online]. Available: <http://www.mdpi.com/1424-8220/9/1/568>
- [3] S. S. Gorthi and P. Rastogi, "Fringe projection techniques: Whither we are?" *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 133–140, 2010.

- [4] M. Mimou, T. Kanade, and T. Sakai, "A method of time-coded parallel planes of light for depth measurement," *IEICE Trans.*, vol. 64, no. 8, pp. 521–528, Aug. 1981.
- [5] X. Liu, X. Peng, H. Chen, D. He, and B. Z. Gao, "Strategy for automatic and complete three-dimensional optical digitization," *Opt. Lett.*, vol. 37, no. 15, pp. 3126–3128, Aug. 2012. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-37-15-3126>
- [6] T. Pribanić, T. Petković, M. Đonlić, V. Angladon, and S. Gasparini, *3D Structured Light Scanner on the Smartphone*. Cham, Switzerland: Springer, 2016, pp. 443–450.
- [7] F. Sadlo, T. Weyrich, R. Peikert, and M. Gross, "A practical structured light acquisition system for point-based geometry and texture," in *Proc. Eurographics/IEEE VGTC Symp. Point-Based Graph.*, Jun. 2005, pp. 89–145.
- [8] H. He, M. Chen, H. Guo, and Y. Yu, "Novel multiview connection method based on virtual cylinder for 3-D surface measurement," *Proc. SPIE*, vol. 44, no. 8, 2005, Art. no. 083605, doi: [10.1117/1.2011067](https://doi.org/10.1117/1.2011067).
- [9] H. Guo and M. Chen, "Multiview connection technique for 360-deg three-dimensional measurement," *Proc. SPIE*, vol. 42, no. 4, pp. 900–905, 2003, doi: [10.1117/1.1555056](https://doi.org/10.1117/1.1555056).
- [10] P. Zheng, H. Guo, Y. Yu, and M. Chen, "Three-dimensional profile measurement using a flexible new multiview connection method," *Proc. SPIE*, vol. 7155, Sep. 2008, Art. no. 715539, doi: [10.1117/12.814619](https://doi.org/10.1117/12.814619).
- [11] C. Chen and A. Kak, "Modeling and calibration of a structured light scanner for 3-D robot vision," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 4, Mar. 1987, pp. 807–815.
- [12] GOM. *ATOS ScanBox—Optical 3D Coordinate Measuring Machine*. Accessed: Jul. 2022. [Online]. Available: <https://www.gom.com/en/products/3d-measuring-machines>
- [13] J. Salvi, S. Fernandez, T. Pribanić, and X. Lladó, "A state of the art in structured light patterns for surface profilometry," *Pattern Recognit.*, vol. 43, no. 8, pp. 2666–2680, Aug. 2010.
- [14] J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognit.*, vol. 37, no. 4, pp. 827–849, Apr. 2004.
- [15] J. Batlle, E. Mouaddib, and J. Salvi, "Recent progress in coded structured light as a technique to solve the correspondence problem: A survey," *Pattern Recognit.*, vol. 31, no. 7, pp. 963–982, 1998.
- [16] Z. H. Zhang, "Review of single-shot 3D shape measurement by phase calculation-based fringe projection techniques," *Opt. Lasers Eng.*, vol. 50, no. 8, pp. 1097–1106, Aug. 2012.
- [17] R. R. Garcia and A. Zakhor, "Markerless motion capture with multi-view structured light," *Electron. Imag.*, vol. 28, no. 21, pp. 3DIPM-050.1–3DIPM-050.7, Feb. 2016. [Online]. Available: <http://www.ingentaconnect.com/content/ist/ei/2016/00002016/00000021/art00018>
- [18] W.-H. Su, C.-Y. Kuo, C.-C. Wang, and C.-F. Tu, "Projected fringe profilometry with multiple measurements to form an entire shape," *Opt. Exp.*, vol. 16, no. 6, pp. 4069–4077, Mar. 2008. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-16-6-4069>
- [19] C.-S. Chen, Y.-P. Hung, C.-C. Chiang, and J.-L. Wu, "Range data acquisition using color structured lighting and stereo vision," *Image Vis. Comput.*, vol. 15, no. 6, pp. 445–456, Jun. 1997.
- [20] S. Fernandez, J. Forest, and J. Salvi, "Active stereo-matching for one-shot dense reconstruction," in *Proc. 1st Int. Conf. Pattern Recognit. Appl. Methods (ICPRAM)*, vol. 2, 2012, pp. 541–545.
- [21] S. Cristina, K. P. Camilleri, and T. Galea, "Multi-view 3D data acquisition using a single uncoded light pattern," in *Proc. 8th Int. Conf. Inform. Control, Autom. Robot. (ICINCO)*, vol. 2, 2011, pp. 317–320.
- [22] S. Heist, P. Lutzke, I. Schmidt, P. Dietrich, P. Kühmstedt, and G. Notni, "GOBO projection-based high-speed three-dimensional shape measurement," *Proc. SPIE*, vol. 9868, May 2016, Art. no. 98680F, doi: [10.1117/12.2227091](https://doi.org/10.1117/12.2227091).
- [23] S. Heist, P. Lutzke, I. Schmidt, P. Dietrich, P. Kühmstedt, A. Tünnermann, and G. Notni, "High-speed three-dimensional shape measurement using GOBO projection," *Opt. Lasers Eng.*, vol. 87, pp. 90–96, Dec. 2016.
- [24] R. R. Garcia and A. Zakhor, "Consistent stereo-assisted absolute phase unwrapping methods for structured light systems," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 5, pp. 411–424, Sep. 2012.
- [25] P. Cong, Z. Xiong, Y. Zhang, S. Zhao, and F. Wu, "Dense single-shot 3D scanning via stereoscopic fringe analysis," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2013, pp. 3627–3631.
- [26] T. Tao, Q. Chen, J. Da, S. Feng, Y. Hu, and C. Zuo, "Real-time 3-D shape measurement with composite phase-shifting fringes and multi-view system," *Opt. Exp.*, vol. 24, no. 18, pp. 20253–20269, 2016. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-24-18-20253>
- [27] T. Pribanić, N. Obradović, and J. Salvi, "Stereo computation combining structured light and passive stereo matching," *Opt. Commun.*, vol. 285, no. 6, pp. 1017–1022, Mar. 2012.
- [28] K. Song, S. Hu, X. Wen, and Y. Yan, "Fast 3D shape measurement using Fourier transform profilometry without phase unwrapping," *Opt. Lasers Eng.*, vol. 84, pp. 74–81, Sep. 2016.
- [29] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Rationale, "Experimentation of structured light and stereo vision for underwater 3D reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 4, pp. 508–518, Jul. 2011.
- [30] F. Willomitzer, S. Ettl, C. Faber, and G. Häusler, "Single-shot three-dimensional sensing with improved data density," *Appl. Opt.*, vol. 54, no. 3, pp. 408–417, Jan. 2015. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-54-3-408>
- [31] W. Jang, C. Je, Y. Seo, and S. W. Lee, "Structured-light stereo: Comparative analysis and integration of structured-light and active stereo for measuring dynamic shape," *Opt. Lasers Eng.*, vol. 51, no. 11, pp. 1255–1264, Nov. 2013.
- [32] S. Woolford and I. S. Burnett, "Toward a one shot multi-projector profilometry system for full field of view object measurement," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 569–573.
- [33] C. Je, K. H. Lee, and S. W. Lee, "Multi-projector color structured-light vision," *Signal Process., Image Commun.*, vol. 28, no. 9, pp. 1046–1058, Oct. 2013.
- [34] R. Sagawa, R. Furukawa, and H. Kawasaki, "Dense 3D reconstruction from high frame-rate video using a static grid pattern," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 9, pp. 1733–1747, Sep. 2014.
- [35] R. Furukawa, R. Sagawa, A. Delaunoy, and H. Kawasaki, "Multiview projectors/cameras system for 3D reconstruction of dynamic scenes," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCV Workshops)*, Nov. 2011, pp. 1602–1609.
- [36] R. Furukawa, R. Sagawa, H. Kawasaki, K. Sakashita, Y. Yagi, and N. Asada, "One-shot entire shape acquisition method using multiple projectors and cameras," in *Proc. 4th Pacific-Rim Symp. Image Video Technol.*, Nov. 2010, pp. 107–114.
- [37] M. Servin, G. Garnica, J. C. Estrada, and A. Quiroga, "Coherent digital demodulation of single-camera N -projections for 3D-object shape measurement: Co-phased profilometry," *Opt. Exp.*, vol. 21, no. 21, pp. 24873–24878, Oct. 2013. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-21-21-24873>
- [38] M. Servin, M. Padilla, G. Garnica, and A. Gonzalez, "Profilometry of three-dimensional discontinuous solids by combining two-steps temporal phase unwrapping, co-phased profilometry and phase-shifting interferometry," *Opt. Lasers Eng.*, vol. 87, pp. 75–82, Dec. 2016.
- [39] T. Jia, B. Wang, Z. Zhou, and H. Meng, "Scene depth perception based on omnidirectional structured light," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4369–4378, Sep. 2016.
- [40] Z. Yan, L. Yu, Y. Yang, and Q. Liu, "Beyond the interference problem: Hierarchical patterns for multiple-projector structured light system," *Appl. Opt.*, vol. 53, no. 17, pp. 3621–3632, 2014. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-53-17-3621>
- [41] R. R. Garcia and A. Zakhor, "Geometric calibration for a multi-camera-projector system," in *Proc. IEEE Workshop Appl. Comput. Vis. (WACV)*, Jan. 2013, pp. 467–474.
- [42] T. Petković, T. Pribanić, M. Đonlić, and P. Sturm, "Efficient separation between projected patterns for multiple projector 3D people scanning," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2017, pp. 815–823.
- [43] T. Petković, T. Pribanić, M. Đonlić, and N. D'Apuzzo, "Multi-projector multi-camera structured light 3D body scanner," in *Proc. 8th Int. Conf. Exhib. 3D Body Scanning Process. Technol.*, Oct. 2017, pp. 319–326, doi: [10.15221/17.319](https://doi.org/10.15221/17.319).
- [44] K. Creath, "V phase-measurement interferometry techniques," *Prog. Opt.*, vol. 26, Jan. 1988, pp. 349–393.
- [45] J. E. Greivenkamp, "Generalized data reduction for heterodyne interferometry," *Proc. SPIE*, vol. 23, no. 4, Aug. 1984, Art. no. 234350, doi: [10.1117/12.7973298](https://doi.org/10.1117/12.7973298).

- [46] C. J. Morgan, "Least-squares estimation in phase-measurement interferometry," *Opt. Lett.*, vol. 7, no. 8, pp. 368–370, Aug. 1982. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-7-8-368>
- [47] C. Zuo, L. Huang, M. Zhang, Q. Chen, and A. Asundi, "Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review," *Opt. Lasers Eng.*, vol. 85, pp. 84–103, Oct. 2016.
- [48] T. Petković, T. Pribanić, and M. Đonlić, "Temporal phase unwrapping using orthographic projection," *Opt. Lasers Eng.*, vol. 90, pp. 34–47, Mar. 2017.
- [49] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. Hoboken, NJ, USA: Wiley, May 1998.
- [50] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [51] M. Servin, J. C. Estrada, and J. A. Quiroga, "The general theory of phase shifting algorithms," *Opt. Exp.*, vol. 17, no. 24, pp. 21867–21881, Nov. 2009. [Online]. Available: <http://www.opticsexpress.org/abstract.cfm?URI=oe-17-24-21867>
- [52] I. Dologlou, S. Bakamidis, and G. Carayannis, "Signal decomposition in terms of non-orthogonal sinusoidal bases," *Signal Process.*, vol. 51, no. 2, pp. 79–91, Jun. 1996.
- [53] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 3rd ed. New York, NY, USA: McGraw-Hill, Jan. 2005.
- [54] M. Frigo and S. G. Johnson, "FFTW: An adaptive software architecture for the FFT," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, vol. 3, May 1998, pp. 1381–1384.
- [55] Z. Wang, D. A. Nguyen, and J. C. Barnes, "Some practical considerations in fringe projection profilometry," *Opt. Lasers Eng.*, vol. 48, no. 2, pp. 218–225, Feb. 2010.
- [56] K. Liu, Y. Wang, D. L. Lau, Q. Hao, and L. G. Hassebrook, "Gamma model and its analysis for phase measuring profilometry," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 27, no. 3, pp. 553–562, 2010. [Online]. Available: <http://josaa.osa.org/abstract.cfm?URI=josaa-27-3-553>
- [57] T. Hoang, B. Pan, D. Nguyen, and Z. Wang, "Generic gamma correction for accuracy enhancement in fringe-projection profilometry," *Opt. Lett.*, vol. 35, no. 12, pp. 1992–1994, 2010. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-35-12-1992>
- [58] B. Pan, Q. Kemao, L. Huang, and A. Asundi, "Phase error analysis and compensation for nonsinusoidal waveforms in phase-shifting digital fringe projection profilometry," *Opt. Lett.*, vol. 34, no. 4, pp. 416–418, 2009. [Online]. Available: <http://ol.osa.org/abstract.cfm?URI=ol-34-4-416>
- [59] R. N. Greenwell and S. Kertzer, "Solving linear diophantine matrix equations using the Smith normal form (more or less)," *Int. J. Pure Appl. Math.*, vol. 55, no. 1, pp. 49–60, 2009.
- [60] T. Petković, S. Gasparini, and T. Pribanić, "A note on geometric calibration of multiple cameras and projectors," in *Proc. 43rd Int. Conv. Inf., Commun. Electron. Technol. (MIPRO)*, Sep. 2020, pp. 1157–1162.
- [61] S. Zhang and P. S. Huang, "Novel method for structured light system calibration," *Proc. SPIE*, vol. 45, no. 8, 2006, Art. no. 083601, doi: [10.1117/1.2336196](https://doi.org/10.1117/1.2336196).
- [62] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, *Bundle Adjustment—A Modern Synthesis*. Berlin, Germany: Springer, 2000, pp. 298–372.
- [63] E. W. Weisstein. *Least Squares Fitting—Power Law (Mathworld—A Wolfram Web Resource)*. Accessed: Apr. 2017. [Online]. Available: <http://mathworld.wolfram.com/LeastSquaresFittingPowerLaw.html>
- [64] M. H. Hayes, *Statistical Digital Signal Processing and Modelling*. Hoboken, NJ, USA: Wiley, 1996.
- [65] T. Petković, T. Pribanić, M. Đonlić, and N. D'Apuzzo, "Software synchronization of projector and camera for structured light 3D body scanning," in *Proc. 7th Int. Conf. 3D Body Scanning Technol.*, Nov. 2016, pp. 286–295, doi: [10.15221/16.286](https://doi.org/10.15221/16.286).
- [66] Y. Díez, F. Roure, X. Lladó, and J. Salvi, "A qualitative review on 3D coarse registration methods," *ACM Comput. Surv.*, vol. 47, no. 3, pp. 1–36, Apr. 2015, doi: [10.1145/2692160](https://doi.org/10.1145/2692160).
- [67] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "MeshLab: An open-source mesh processing tool," in *Proc. Eurograph. Italian Chapter Conf.*, Salerno, Italy, V. Scarano, R. D. Chiara, and U. Erra, Eds. Geneva, Switzerland: The Eurographics Association, 2008, pp. 129–136.
- [68] *MeshLab*. Accessed: Jul. 2022. [Online]. Available: <https://www.meshlab.net/>
- [69] B. Li, N. Karpinsky, and S. Zhang, "Novel calibration method for structured-light system with an out-of-focus projector," *Appl. Opt.*, vol. 53, no. 16, pp. 3415–3426, Jun. 2014. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-53-16-3415>
- [70] R. Sagawa, Y. Ota, Y. Yagi, R. Furukawa, N. Asada, and H. Kawasaki, "Dense 3D reconstruction method using a single pattern for fast moving object," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 1779–1786.
- [71] W. Kester, "Taking the mystery out of the infamous formula, 'SNR=6.02N+1.76 dB,' and why you should care," Analog Devices, Wilmington, MA, USA, Tech. Rep. MT-001 TUTORIAL, 2009.
- [72] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, vol. 19, no. 1, 1980, Art. no. 191139, doi: [10.1117/12.7972479](https://doi.org/10.1117/12.7972479).



TOMISLAV PETKOVIĆ (Member, IEEE) received the engineering, magister, and Ph.D. degrees in electrical engineering from the University of Zagreb, in 2002, 2006, and 2010, respectively. He is currently an Assistant Professor with the Faculty of Electrical Engineering and Computing, University of Zagreb. He teaches several graduate courses in the field of digital signal and image processing. His research interests include digital image processing and analysis, depth sensing, and 3D imaging. He is a member of ACM.



TOMISLAV PRIBANIĆ (Senior Member, IEEE) is currently a Professor with the Faculty of Electrical Engineering and Computing, University of Zagreb. He teaches several undergraduate and graduate courses in the field of algorithms and data structures, image processing, sensors, and human motion analysis. His research interests include computer vision and biomedical signal measurement and analysis. He has led a number of domestic and international scientific projects, collaborating with researchers from within and outside the EU. He was a Visiting Researcher at INRIA Rhone-Alpes, Grenoble, France, Fraunhofer IGD, Darmstadt, Germany, and a Fulbright Visiting Scholar at the University of Wisconsin–Madison, Madison, USA. The results of his research have been implemented in technological projects and he has also received recognition for innovations. He is a Senior Member of IFMBE and a Collaborating Member of the Croatian Academy of Engineering.

• • •