Real-time 3D shape measurement using 3LCD projection and deep machine learning

5 authors, including:

Andrew-Hieu Nguyen
National Institute on Drug Abuse

33 PUBLICATIONS   **626** CITATIONS

Zhaoyang Wang
The Catholic University of America

89 PUBLICATIONS   **4,001** CITATIONS

Real-time 3D shape measurement using 3LCD projection and deep machine learning

HIEU NGUYEN,[1] NICOLE DUNNE,[1] HUI LI,[2] YUZENG WANG,[3] AND ZHAOYANG WANG[1],*

[1]Department of Mechanical Engineering, The Catholic University of America, Washington, DC 20064, USA
[2]RvBust Ltd., 16 Keyuan Rd., Shenzhen 518057, China
[3]School of Mechanical Engineering, Jinan University, Jinan 250022, China
*Corresponding author: wangz@cua.edu

For 3D imaging and shape measurement, simultaneously achieving real-time and high-accuracy performance remains a challenging task in practice. In this paper, a fringe-projection-based 3D imaging and shape measurement technique using a three-chip liquid-crystal-display (3LCD) projector and a deep machine learning scheme is presented. By encoding three phase-shifted fringe patterns into the red, green, and blue (RGB) channels of a color image and controlling the 3LCD projector to project the RGB channels individually, the technique can synchronize the projector and the camera to capture the required fringe images at a fast speed. In the meantime, the 3D imaging and shape measurement accuracy is dramatically improved by introducing a novel phase determination approach built on a fully connected deep neural network (DNN) learning model. The proposed system allows performing 3D imaging and shape measurement of multiple complex objects at a real-time speed of 25.6 fps with relative accuracy of 0.012%. Experiments have shown great promise for advancing scientific and engineering applications. © 2019 Optical Society of America

https://doi.org/10.1364/AO.58.007100

1. INTRODUCTION

As technologies evolve at an ever-increasing pace, noncontact 3D imaging and shape measurements have become a leading topic of research and development in a wide range of fields. Typical applications include machine vision, reverse engineering, quality assurance, 3D printing, and entertainment. At present, there are many techniques capable of performing 3D imaging and shape measurements, such as the time-of-flight method, photogrammetry method, laser scanning method, stereo vision method, moire method, interferometry method, and structured-light method [1–6]. In spite of the various advantages that these techniques have, simultaneously achieving real-time and high-accuracy performance remains challenging for all of them in practice.

So far, the structured-light techniques have led to the largest number of commercialization in the field of noncontact 3D imaging and shape measurement. The structured patterns employed by those techniques generally fall into two categories: (1) speckle or dot patterns and (2) fringe patterns. Representative products in the first category include Intel RealSense sensor [7] and Apple TrueDepth camera [8], and they have the ability of acquiring 3D images at a fast speed to deal with dynamic scenes. Their primary shortcoming is the low accuracy. On the contrary, the techniques in the second category have the capability of capturing accurate 3D images,

but the measurement speed is substantially limited when high accuracy is demanded. The ZEISS COMET scanner and GOM ATOS scanner [9,10] are two well-known products of this type. Considering the above facts, the research in this paper uses a structured-light technique with fringe patterns, named fringe-projection-based (FPB) technique to achieve a real-time and high-accuracy 3D imaging and shape measurement scheme.

The running time of the FPB technique is mainly comprised of the image acquisition time and the image processing time. In general, the image processing time for each 3D image reconstruction can be very short with several advanced approaches such as fast phase extraction algorithm, simple yet rigorous governing equation, and parallel computation. Consequently, the speed of the FPB 3D imaging and shape measurement system highly depends on the frame rates of the camera and the projector as well as their synchronization. In the past two decades, the speed of the FPB technique has been greatly improved, advancing from a few frames per second (fps) to a few thousand fps [11–14]. For instance, Rusinkiewicz et al. [15] introduced an approach of capturing four-frame binary-coded patterns at 60 fps to achieve a final 3D imaging speed of 15 fps, but its limitation was the low image resolution. Zhang et al. [16] modified a DLP (digital-light-processing) projector to obtain real-time 3D shape measurement at a speed of 30 fps, and the drawback was the relatively low accuracy as seen

from their experimental results. The method proposed by Liu et al. [17] used dual-frequency patterns and a look-up-table (LUT) scheme to acquire 3D shape reconstruction results at 228 fps with a variance of phase error at 0.031, and the approach reported in Ref. [18] proposed a method of using four designated patterns and a modified DLP projector to obtain 3D point clouds at 120 fps with an error of 0.527%.

The common low-accuracy issue of the fast speed FPB techniques generally originates from the notable errors in the phase distributions extracted from the captured images. For this reason, a number of approaches have been proposed to reduce the phase-detection errors [19–23]. The investigation in Ref. [24] compared and discussed different phase error compensation measures, including the LUT method, intensity correction method, and gamma- correction method. Yao et al. [25] believed that system vibration can also be a key factor of the phase errors. Recently, Lu et al. [26] presents a self-correction technique to suppress the effect of the projector nonlinearity by a filtering and curve-fitting technique. Most of these approaches, except the LUT method, often notably attenuate the measurement speed and processing speed of the FPB technique.

In the machine learning field, artificial deep neural networks (DNNs) have found numerous applications in object detection, image segmentation, image classification, scene understanding, medical image analysis, and natural language processing. Over the past year, the utilization of the DNN framework for fringe pattern analysis, such as pattern denoising and phase distributions determination, has drawn much attention. For instance, Feng et al. [27] used the DNN to perform phase demodulation from a single fringe pattern with enhanced accuracy. A number of investigations [28–31] have shown promising results on using the DNN models to improve the estimation and detection of phase distributions. Jeught [32] proposed a neural network with a large simulation dataset in the training process to acquire the height information of an object from a single-shot fringe pattern.

This paper presents an effective FPB technique to achieve real-time and high-accuracy 3D imaging and shape measurement of objects using a 3LCD projector. In the proposed system, the 3LCD projector is modified to display the red, green, and blue (RGB) channels consecutively. Accordingly, the technique uses a three-step phase shifting algorithm, and the three phase-shifted fringe patterns are individually put in the RGB channels to form a composite image to be fed into the 3LCD projector. For accurate 3D shape measurement, the three RGB encoded fringes must be captured at the specific time and in the correct order. Therefore, an external trigger circuit is designed to control the three LCD panels and synchronize them with a fast-speed camera. In addition, the 3D measurement accuracy is dramatically improved by using a flexible calibration method and an advanced phase determination approach based on a deep machine learning process. The proposed system allows carrying out 3D imaging and shape measurement with high accuracy at a real-time speed. An illustration of the proposed system is shown in Fig. 1.

The rest of the paper is organized as follows: Section 2 describes the fundamental concept of the involved FPB technique
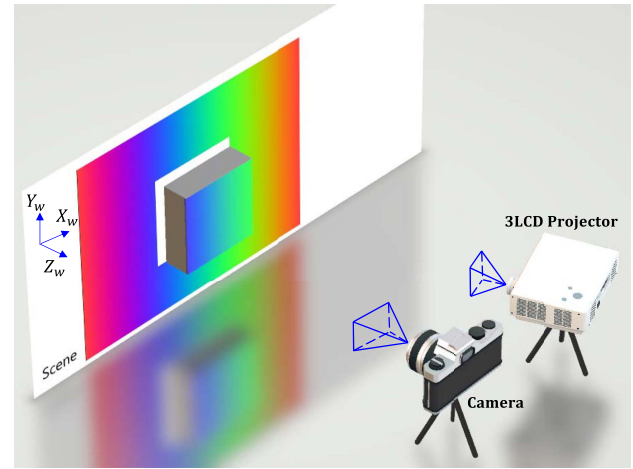
Fig. 1. Schematic of the proposed 3D imaging and shape measurement system.

and the governing equation for the out-of-plane coordinates determination; Section 3 presents the hardware implementation as well as the deep machine learning approach for phase extraction; Section 4 demonstrates a number of experiments to validate the proposed technique. A summary with a brief discussion is outlined in the last section.

2. PRINCIPLE OF THE FPB TECHNIQUE

The FPB 3D imaging and shape measurement system is typically comprised of a camera, a projector, and a computer. The technique involves projecting a set of phase-shifted sinusoidal fringe patterns from the projector onto the objects, where the surface depth or height information is naturally encoded into the camera-captured fringe patterns for subsequent determination of the 3D coordinates. Technically, the FPB technique is similar to the stereo-vision-based technique, which can retrieve 3D information based on geometric triangulation. The distinction between these two methods is that FPB technique uses a projector to replace one of the stereo cameras, and the fringe patterns help efficiently establish the correspondences between the captured image and the original reference image projected by the projector. In practice, the fringe patterns allow converting the calculation of geometric triangulation to be in a simplified format, which directly determines the height or depth map from the phase distributions of the captured fringe patterns. Here, the computation of the phase distributions relies on a phase-extraction process that normally uses phase-shifted fringe patterns.

In general, the original fringes are straight, evenly spaced, and oriented vertically (or horizontally in some cases). They are generated in a numerical way with the following function:

$$I_j(u, v) = I_0[1 + \cos(\phi(u, v) + \delta_j)]$$
$$= I_0\left[1 + \cos\left(2\pi f \frac{u}{w} + \delta_j\right)\right], \quad \textbf{(1)}$$

where I is the intensity value of the pattern at pixel coordinate (u, v); I_0 is a constant coefficient indicating the value of the intensity modulation; f is the number of fringes in the pattern;

w is the width of the digital image; ϕ is the fringe phase; $\delta_j = 2\pi\frac{(j-1)}{m}$ is the phase-shift amount; the subscript j denotes the jth phase-shifted image with $j = \{1, 2, \ldots, m\}$; and m is the number of the phase-shift steps. It is noteworthy that to achieve high accuracy for the 3D coordinate determination, f is usually assigned a large number, such as 100. Figure 2 shows a few representative fringe patterns.

The commercial projectors often apply a gamma modification value to the intensity output to enhance the visual effect for human eyes. However, the change in projected intensities introduces a nonlinear effect to the sinusoidal fringe intensities that affects the accuracy of the FPB measurement system. Although gamma correction can be applied at the camera side by carefully adjusting the exposure time and contrast of the images, such a manual correction affects the automatic nature of the FPB system. Currently, there are a number of existing methods available for compensating the gamma nonlinearity distortion [33–35], but they are computationally complex and often notably affect the image processing speed. Here, to cope with the nonlinear intensity-distortion issue, an easy yet very effective method is employed. It simply pre-encodes or modifies the initial fringe patterns described in Eq. (1) as

$$I_j(u, v) = 2I_0\left[\frac{1}{2} + \frac{1}{2}\cos\left(2\pi f\frac{u}{w} + \delta_j\right)\right]^{\frac{1}{\gamma}}, \quad (2)$$

where the value of γ is generally in a range 1.5–3.5, and it can be determined in advance using two or more sets of projected fringe patterns with a three-step or larger-step phase-shifting method [36]. Knowing such handling can solve the gamma problem at the source before the projection; no extra calculation is required for the FPB system. This scheme helps increase the quality of the fringe patterns while maintaining the fast processing speed.

The fringe patterns captured by the camera are slightly different from the original patterns sent to the projector, and they can be expressed as

$$I_j(u, v) = A(u, v) + B(u, v)[1 + \cos(\phi(u, v) + \delta_j)], \quad (3)$$
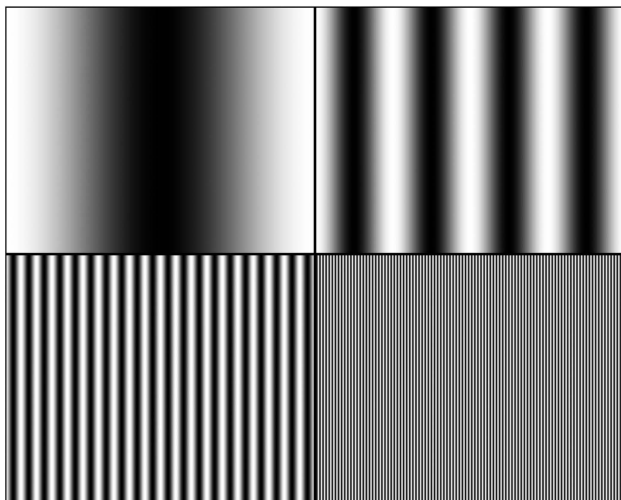
Fig. 2.   Example of sinusoidal fringe patterns with 1, 4, 20, and 100 fringes, respectively.

where A and B indicate the background intensity and fringe intensity modulation, respectively. After capturing the images, the next step is to obtain the phase distributions of the fringe patterns in the images. The fringe phase can be directly calculated from a standard phase-shifting algorithm, such as the three-step one:

$$\phi^w(u, v) = \arctan\frac{\sqrt{3}(I_3(u, v) - I_2(u, v))}{2I_1(u, v) - I_2(u, v) - I_3(u, v)}, \quad (4)$$

where the superscript w denotes the wrapped phase because the calculated phase value is wrapped in a range of $0 - 2\pi$. If $f = 1$, the wrapped phase is identical to the actual phase; however, the phase distributions in such a case normally contain large errors. On the other hand, if f is a large number, which is essential in practice for accurate measurements, the calculated phase must be unwrapped to obtain the true phase. Phase unwrapping is a challenging task in the 3D imaging and shape-measurement system, especially when it deals with objects with complex shapes and geometric discontinuities (e.g., multiple separated objects). Although a number of researchers [37–41] have addressed the phase unwrapping issue with a variety of solutions, they are not suitable for the fast-speed 3D imaging system. In this work, a deep machine-learning approach is proposed to determine the phase distributions with high accuracy from the case of $f = 1$, and the details will be described in the following section. Because the deep machine-learning model uses supervised datasets for training, validation, and test purposes, an accurate phase unwrapping algorithm with a large f number is still required by the deep learning process to create the labels of accurate phase distributions. After the learning is completed, phase unwrapping is no longer needed.

An efficient and robust scheme of using multifrequency fringe patterns is adopted to fulfill the phase unwrapping task. The scheme uses a set of fringe patterns with various frequencies, where the highest-frequency fringes provide the desired phase distributions for 3D coordinate determination, and the lower-frequency fringes provide the required integer fringe-order offsets for the higher-frequency fringes. The unwrapped phase distributions can be consecutively calculated from:

$$\phi_i(u, v) = \phi_i^w(u, v) + \text{INT}\left(\frac{\phi_{i-1}\frac{f_i}{f_{i-1}} - \phi_i^w}{2\pi}\right)2\pi. \quad (5)$$

In the equation, the subscript i indicates the ith fringe-frequency pattern with $i = \{2, 3, \ldots, n\}$, and n is the number of fringe frequencies; INT represents the function of rounding to the nearest integer; f is again the number of fringes in the projection pattern; and $f_n > f_{n-1} > \cdots > f_1 = 1$. A practical example is $n = 4$ with $f_4 = 100$, $f_3 = 20$, and $f_2 = 4$, as shown in Fig. 2.

The essential task of the FPB technique is to retrieve the out-of-plane depth or height map from the afore-calculated phase distributions of the highest frequency fringes with the highest possible accuracy. The governing equation for a generalized setup where the system components can be arbitrarily positioned [42] is:

$$z = \frac{\mathbf{CP}}{\mathbf{DP}}$$

$$\mathbf{C} = \{\, 1 \quad c_1 \quad c_2 \quad c_3 \quad \cdots \quad c_9 \quad c_{10} \quad c_{11} \,\}$$

$$\mathbf{D} = \{\, d_0 \quad d_1 \quad d_2 \quad d_3 \quad \ldots \quad c_9 \quad d_{10} \quad d_{11} \,\}$$

$$\mathbf{P} = \{\, 1 \quad \phi \quad u \quad u\phi \quad v \quad v\phi \quad u^2 \quad u^2\phi \quad v^2 \quad v^2\phi \quad uv \quad uv\phi \,\}^\mathsf{T},$$

$$\text{(6)}$$

where z is the out-of-reference-plane height or depth at the point corresponding to the pixel (u, v) in the captured images; ϕ is the unwrapped phase of the highest-frequency fringe pattern at the same pixel location; and $c_1 - c_{11}$ and $d_0 - d_{11}$ are coefficients that must be determined by a calibration process.

The calibration of the system, i.e., the determination of $c_1 - c_{11}$ and $d_0 - d_{11}$, can be accomplished by using a number of gage blocks with known and different heights. However, such a calibration method is not convenient in practice due to the cost and difficulties of manufacturing precise 3D calibration targets. In this paper, a flexible calibration technique that uses a planar board as a gage object is employed [43]. The board has checker or concentric-circle patterns on its surface, as demonstrated in Fig. 3.

The flexible calibration of the FPB technique include two steps: (1) camera calibration and (2) FPB system calibration. The camera calibration process is to determine the intrinsic and extrinsic parameters of the camera as well as the lens distortion coefficients that must be considered in practice. The FPB system calibration is to obtain the coefficients $c_1 - c_{11}$ and $d_0 - d_{11}$ for the governing equation, Eq. (6). The calibration process requires capturing a series of fringe patterns projected on the planar calibration board with different poses at various positions (e.g., 15–20 positions). The calibration image at each position can be obtained by averaging the captured fringe images to remove the fringes and reduce the noise. The 3D world coordinates of the control points (such as the corner points on chess board or the circle centers on the concentric-circle board) can be determine using the bundle-adjustment-based camera calibration process. These control points can then serve as gage points to calibrate the FPB 3D imaging system using a nonlinear least-squares approach as follows:

$$S = \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\frac{\mathbf{CP}}{\mathbf{DP}} - z_{kl} \right)^2, \qquad \text{(7)}$$

where z_{kl} is the height or depth of lth control point on the calibration board obtained at the kth calibration position; and K and L are the total numbers of the board positions
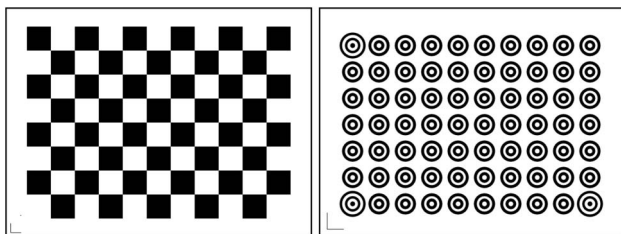

Fig. 3. Example of two commonly used planar calibration targets.

and the control points, respectively. The constant coefficients $c_1 - c_{11}$ in $\mathbf{C}$ and $d_0 - d_{11}$ in $\mathbf{D}$ can be determined by using the well-known Levenberg–Marquardt algorithm. The initial guess of the coefficients is obtained by minimizing a linear least-squares cost function of $S' = \sum_{k=1}^{K} \sum_{l=1}^{L} (\mathbf{CP} - \mathbf{DP}z_{kl})^2$.

3. ADVANCED METHODOLOGY

A. Hardware Design and Enhancement

DLP, LCD, and liquid crystal on silicon (LCoS) are three popular technologies used in designing digital projectors [44,45], and they each have their own advantages and disadvantages depending on the applications. At present, DLP- and LCD-based projectors hold the majority of the market share. A DLP projector typically uses three individual light sources or a color wheel to control the projection of the RGB channels of every color image in sequence, and it relies on human color perception to mix and interpret the correct color for each pixel. In contrast, an LCD projector normally uses three LCD panels to handle the RGB channels of every color image, respectively, and then mix them together simultaneously right before the projection. For this reason, an LCD projector is usually called a 3LCD projector, and it is adopted in this work because 3LCD projectors dominate the market in recent years. To take advantage of the projection mechanism of the 3LCD projector, the proposed system uses three phase-shifted fringe patterns to determine the full-field phase distributions.

The optical parts inside the 3LCD projector include the light source, dichroic mirrors, mirrors, RGB LCD panels, prism, and lens. The light source generates uniform light that passes through the dichroic mirrors to become three different RGB light channels, and the regular mirrors lead the light channels to the LCD panels. Each of the three LCD panels then processes and modulates the light according to the incoming video signal. After that, the RGB modulated lights from three LCD panels are recombined at the prism to generate the color images.

The working principle of the 3LCD projector can considerably facilitate the proposed approach. In general applications, the 3LCD panels are always turned on to project the color images. In the proposed approach, on the other hand, a trigger circuit is designed to turn on and off the LCD panels so that each of the RGB lights can be controlled to go through the prism and lens individually. In this way, the three phase-shifted fringe patterns can be respectively coded into the three RGB channels of a color image. When the static image is projected by the 3LCD projector, it can be controlled to display the three fringe patterns sequentially for the camera to capture them one by one. This scheme avoids continuously replacing the projection image sent to the projector, so it allows the system to perform faster and more reliably. It is noted that three individual images instead of a composite image are captured; the reason is that separating a captured composite image into three desired images is very difficult in practice. Figure 4 illustrates the schematic of the 3LCD projection mechanism and the three RGB fringe patterns.

Testing shows that each of the three fringe patterns requires 10.5 milliseconds (ms) to be stable on the scene for the camera to capture after the trigger is turned on. To capture each of the RGB images, a fast-speed complementary-metal-oxide-semiconductor
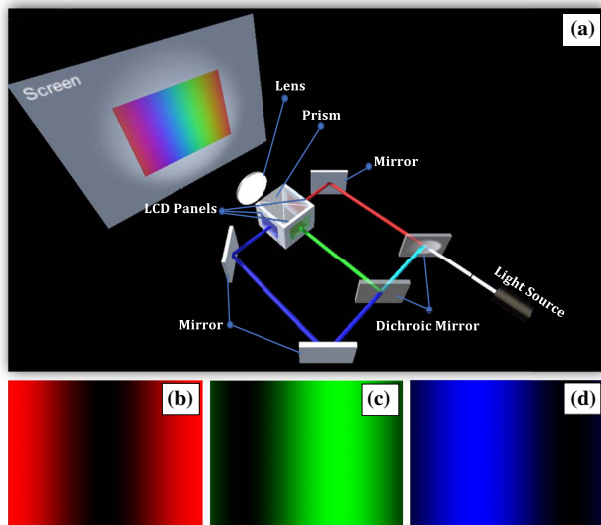
Fig. 4. (a) Illustration of the proposed 3LCD projection mechanism (see Visualization 1); (b)–(d) the three consecutive images projected on the scene.
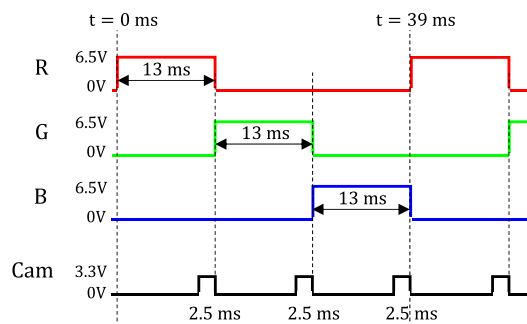
Fig. 5. Synchronization timeline of the projecting and capturing processes.

(CMOS) camera with a short exposure time of 2.5 ms is employed in this work. The camera is controlled by using a low-voltage transistor-transistor-logic (LVTTL) trigger input and strobe output module, which transmits the signal from a microcontroller. The purposed of using the microcontroller is to synchronize the projecting and capturing with a designated timeline pattern, as shown in Fig. 5. It is noted that in our work, the voltages of the triggering signals required by the camera and the projector are 3.3 and 6.5 V, respectively.

### B. Machine Learning Approach for Phase Determination

As described previously, a notable issue with the phase-shifting approach is the phase error induced by the nonlinearity of the projected light and the noise. A practically excellent solution to tackle this problem is using the multifrequency fringe patterns; however, the drawback is that both the capturing and the processing times will be substantially increased. As an example, Fig. 6 shows the phase distributions along an arbitrary line in a representative experiment using singe-frequency singe-fringe patterns and the desired phase distributions obtained by using

multifrequency fringe patterns. It is evident from the figure that phase distributions detected from using single-frequency single-fringe patterns must be corrected to get the desired phase distributions. The goal is to make the plot shown in Fig. 6(d) a straight line of equality.

A recent approach presented in Ref. [46] solves the problem by capturing two sets of fringe patterns and utilizing a LUT to compensate for phase errors. In this work, a machine learning approach using an artificial DNN is proposed to determine the phase distributions in real-time 3D imaging and shape measurement; whereas the method of using multifrequency fringe patterns is adopted for the system calibration and the generation of training, validation, and test datasets.

DNN-based machine learning has drawn considerable attentions in recent years because of its unprecedented outstanding behaviors in numerous applications. The basic principle of the proposed DNN phase determination approach is similar to the idea of using accurate unwrapped phase extracted from multifrequency fringe patterns to correct the phase distributions determined from single-frequency single-fringe patterns. One primary difference is that the DNN approach is an end-to-end process because it directly yields the desired phase from the captured three images. Another key difference is that instead of using a polynomial or a more complex function, a fully connected network is adopted and trained to accomplish the phase determination task.

Figure 7 illustrates the proposed DNN framework for phase determination. The DNN consists of eight network layers, including one input layer, six hidden layers, and one output layer. The hidden layers are arranged in a chain structure, and each layer is a mapping function of the layer that preceded it. In the network, the six hidden layers include 32, 64, 128, 64, 32, and 16 nodes, respectively. The network has a total of 21,441 parameters, including both weights and biases. The chain architecture has been found in this work to be an excellent network for the phase determination task via experimentation guided by monitoring the validation set error.
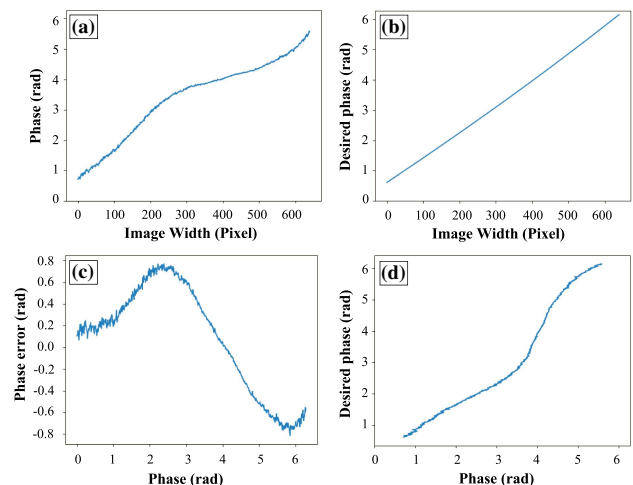
Fig. 6. Example of phase distributions along a line in a representative experiment. (a) Phase detected from single-frequency single-fringe patterns; (b) desired phase extracted using multiple-frequency patterns; (c) phase differences of (a) and (b); and (d) the phase plot of (b) versus (a).
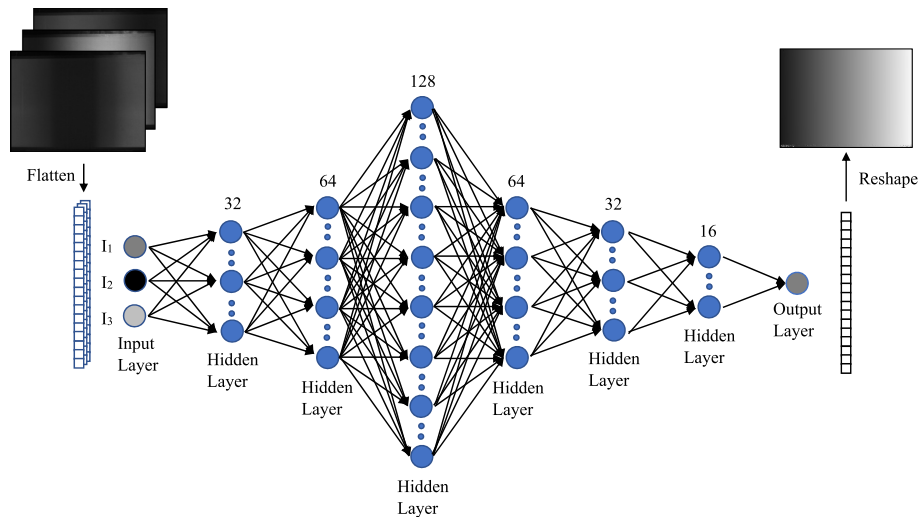
Fig. 7. Schematic of the artificial deep neural network for phase determination using three phase-shifted singe-fringe patterns.

In the chain-based architecture, each layer is given by [47]

$$\mathbf{h}^{(i)} = g^{(i)}(\mathbf{W}^{(i)\mathsf{T}}\mathbf{x} + \mathbf{b}^{(i)}), \tag{8}$$

where $\mathbf{h}$ is the output vector; i is the layer index which spans from 1–6; function g is called an activation function; $\mathbf{x}$ is a vector of input; the parameters $\mathbf{W}$ in a matrix form and $\mathbf{b}$ in a vector form are optimized by the learning process with a mean squared error (MSE) cost function.

The input of the training or validation dataset is a two-dimensional array of size $s \times c$, where s is the number of the training or validation samples, and c is the channel dimension with $c = 3$. Therefore, the intensity values at the valid pixels in the three phase-shifted images are flattened to become the input array. Similarly, the desired phase values corresponding to the intensity inputs are flattened to be a label array of size $s \times 1$. The layers in the network are fully connected, and each hidden layer learns the input features from the previous layer and produces the output features for the next layer where the depth axis changes following the number of nodes. A nonlinear activation function named rectified linear unit (ReLU), expressed as $\max(0, x)$, is adopted in each hidden layer to accelerate the learning convergence. The output layer with a single node is attached to the last hidden layer to transform the features to the desired phase value.

The optimization adopts a total of 100 epochs with a mini-batch size of 500 samples. The learning rate is reduced by 10 times whenever the validation loss does not improve within 10 epochs. In each epoch, there is no need to shuffle the training data because the mapping function is not complicated and does not have a notable over-fitting problem. The learning process uses the Adam optimizer as well as the MSE as the loss function. Furthermore, the input values are normalized to a range between 0 and 1 to improve the performance of the DNN model.

The overview procedure of the deep machine learning approach is summarized as follows:

1. Capture the first set of images that involve three-step phase-shifted patterns with a single fringe (i.e., the fringe frequency $f = 1$). The three grayscale intensities at the same pixel in the three captured images are the input of the DNN model.

2. Capture the second set of images that include multifrequency three-step phase-shifted patterns with a few different frequencies starting from 1 and ending at a relatively high frequency (e.g., a set of four: 1, 4, 20, and 100). The phase calculated from Eq. (5) is the output label for the DNN learning process.

3. A few additional sets of images can be acquired by repeating the previous two steps.

4. The datasets are then used to train, validate, and test the DNN model.

5. Use the trained DNN model to generate a LUT. It is noteworthy that a LUT is adopted because the maximum number of possible different input cases is limited. For instance, the theoretical upper limit is $2^8 \times 2^8 \times 2^8 = 16{,}777{,}216$ cases for 8-bit grayscale images. The actual number is much smaller since the three images are phase shifted images.

6. The LUT is employed to determine the phase values in subsequent 3D imaging and shape measurement applications, which uses three phase-shifted single-fringe patterns.

Unlike the commonly used polynomial regression algorithm for phase-error compensation which usually involves a small number of coefficients, the DNN phase-determination model uses ground-truth datasets to train a much larger number of coefficients in the fully connected networks. For comparison purposes, many tests have been conducted in this work to find the best polynomial regression models for the task of phase-error compensation. These regression models include linear regression, Tikhonov regularization (Ridge), least absolute shrinkage and selection operator (Lasso), and Elastic Net [48–51]. The comparison reveals that the performance of the DNN model is superior to those of the polynomial regression models. Some results will be presented in the following section.

4. EXPERIMENTS AND RESULTS

A number of experiments have been carried out to demonstrate the validity and applicability of the proposed technique.
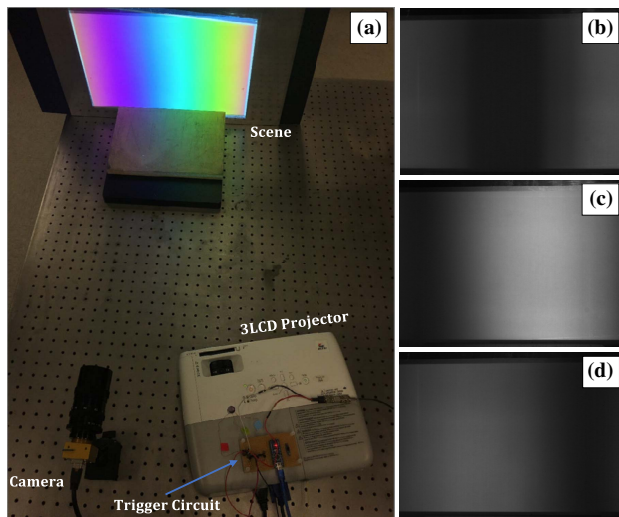
Fig. 8. (a) Experimental setup (see Visualization 2); (b)–(d) three grayscale fringe images captured by the camera.

The experimental system consists of a CMOS camera with a resolution of 800×600 pixels, a 3LCD projector, a trigger circuit, and a computer with an Intel Core i7-860 2.8 GHz processor and 16 GB RAM. In the experiments, the camera and projector are arbitrarily positioned to form a generalized setup. The distances among the camera, projector, and the reference plane as well as the field of view are adjusted according to the dimensions of the objects of interest in each experiment. The image capturing and processing program is written in C++ language, and the machine learning part is written in Python language. Figure 8 demonstrates the experimental set up with an example of captured three phase-shifted fringe images.

As described previously, it takes 13 ms to perform the task of projecting and capturing a fringe-pattern image. Because the processing speed is ultrafast and the task is conducted in parallel with the projecting and capturing process, the processing does not affect the overall running time, and each measurement cycle requires 39 ms to complete the reconstruction of a 3D image. Consequently, the proposed real-time FPB 3D imaging and shape measurement has a speed of 25.6 fps.

A. Calibration and Phase Determination

After the hardware setup is complete, the system is first calibrated to determine the camera intrinsic and extrinsic parameters as well as the coefficients of the governing equation. This calibration process involves capturing 15–20 sets of multifrequency phase-shifted fringe images of the calibration board at different positions, where four frequencies with $f = 1, 4, 20, 100$ and three-step phase shifting scheme are employed. After the calibration, additional groups of fringe images are captured to carry out the learning task of the DNN model. One group is similar to the aforementioned sets of phase-shifted multifrequency fringe images for labeling the datasets, and another group is similar to the one in 3D imaging and shape measurement applications, which involves three phase-shifted single-fringe images.

In the training process, 80%, 10%, and 10% of the valid points in the datasets serve as the training, validation, and test

Table 1.   Detection Errors of the Training and Test Data Using Different Models^a

| Model	Parameter	Training		Test	
		MAE	RMSE	MAE	RMSE
Linear	17	0.0258	0.0333	0.0595	0.0720
Ridge	20	0.0257	0.0332	0.0586	0.0687
Lasso	20	0.0895	0.1072	0.0930	0.1135
Elastic Net	20	0.0895	0.1071	0.0929	0.1135
DNN	21,441	0.0221	0.0280	0.0242	0.0309

^{a}Unit: radian.

data, respectively. In this work, there are a total of 1,550,000 samples in the datasets. If required, additional datasets can always be obtained with ease by capturing extra groups of fringe patterns, so the data augmentation process commonly adopted in deep machine learning to artificially expand the dataset size is unnecessary here. The data samples employed in the DNN model and polynomial regression analysis can be obtained by capturing the 3D shape of a plain board without specific requirements on the surface profile and color. In addition to the proposed fully connected DNN model, a few other commonly used and previously mentioned regression models are adopted for comparison purposes. The results are listed in Table 1, where both mean-absolute error (MAE) and root-mean-squared error (RMSE) are utilized to evaluate the performance of the models. The results show that each of the linear, Ridge, Lasso, and Elastic Net polynomial regression models performs the best with 17–20 parameters, whereas the proposed DNN model with 21,441 parameters is superior to all the regression models for both the training and test datasets. The number of parameters used in the polynomial regression models has been determined using the early stopping rule, where the training process is terminated when the validation error cannot be further reduced over a prespecified threshold. The learning process of the DNN model is quite fast and takes less than 2 min to complete.

Figure 9 shows four selected plots of phase distributions for visual comparison. The results again reveal that the DNN model performs the best, which is expected because other models use a global polynomial function for the phase error compensation.
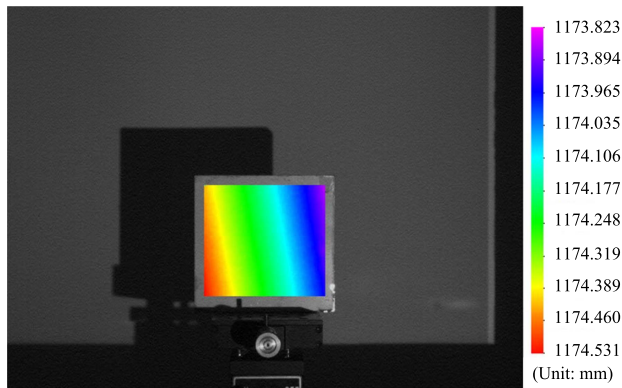
B. Accuracy Test

The accuracy test aims to check the measurement accuracy of the proposed 3D imaging and shape measurement system. Because the key dimension of the 3D shape measurement is the depth or height information, a distance measurement is conducted in this test. In the experiment, the distance from the camera to the virtual reference plane is determined as 1200.33 mm from the camera calibration, and the field of view is 381.0 mm at the reference plane. A white flat-gage block of 101.6 mm × 101.6 mm × 25.4 mm is placed on an optical translation stage driven by a differential adjuster with micrometer resolution. From its initial position, the block is moved towards the camera for 1 mm (which is the working range of the translation stage at high resolution) with an increment of 0.1 mm in each of the 10 movement steps. Figure 10 shows one of the measurement results, where the color map represents
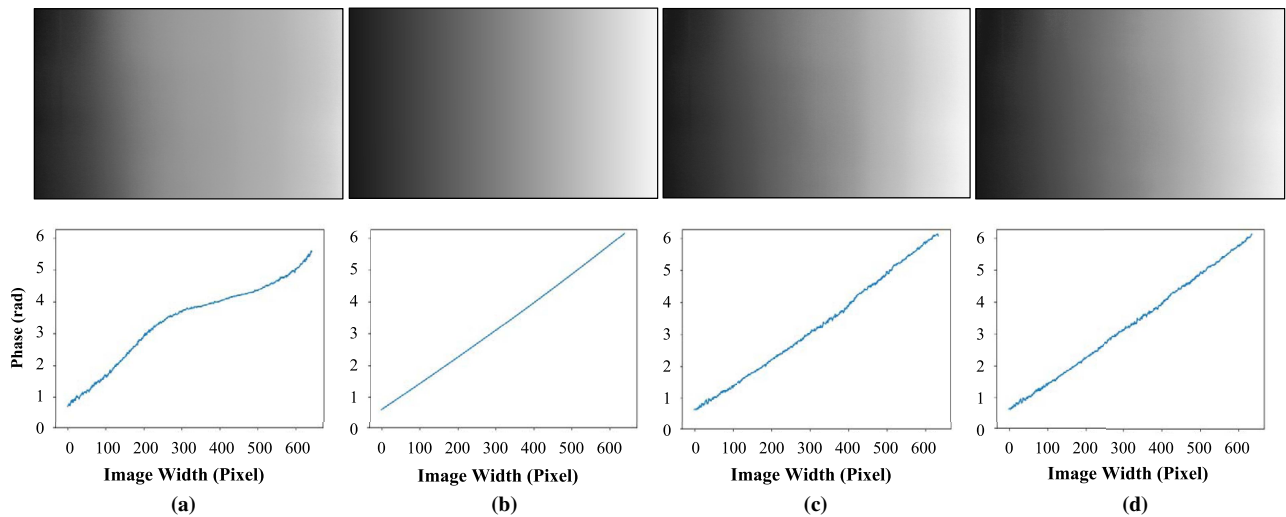
Fig. 9. Phase determination results. (a) Original phase extracted from three single-fringe images; (b) desired phase extracted using multiple-frequency fringe images; (c) phase in (a) corrected using a degree-16 linear polynomial regression model; and (d) phase determined from the same images used by (a) with the proposed DNN model.

Fig. 10. Example of distance measurement in accuracy test.

Table 2. Distance Measurement of a Block[a]

Position	Nominal distance	Measured distance	Error
0	1173.877	1173.877	0
1	1173.977	1173.985	0.008
2	1174.077	1174.091	0.014
3	1174.177	1174.160	−0.017
4	1174.277	1174.305	0.028
5	1174.377	1174.421	0.044
6	1174.477	1174.502	0.025
7	1174.577	1174.542	−0.035
8	1174.677	1174.723	0.046
9	1174.777	1174.749	−0.028
10	1174.877	1174.910	0.033

[a]Unit: mm.

the distance from the block surface to the camera. It is noted that because the block has a very small incline angle with the virtual reference plane, the distances in the region of interest have a 0.754 mm variation.

The average distance from each block position to the camera is calculated for assessment, and the result is summarized in Table 2. This table documents that the largest error of measurement is 0.046 mm at the eighth movement position. There are many possible factors of errors including system error of the technique, translation stage error, noise, and vibration. Nevertheless, the small measurement error helps confirm the validity of the proposed 3D imaging and shape measurement approach. Particularly, the accuracy test substantiates that the relative accuracy, defined as the ratio of the measurement error in depth to the field width of the view, is 0.046 mm/381.0 mm ≈ 0.012%.

C. 3D Shape Measurement of Multiple Objects
A number of additional experiments have been carried out to measure the 3D surface profiles of multiple objects to demonstrate the capability of handling geometric discontinuities. Figure 11 shows one of the experimental results, where the two subfigures are the image of the objects of interest and the corresponding rendered plot of the 3D surface profiles, respectively. The experiment involves not only multiple separated objects but also objects with bright and dark surfaces, and this validates the effectiveness of the proposed technique for practical applications.

D. 360° 3D Shape Measurement
Furthermore, the proposed FPB 3D imaging and shape measurement technique has been employed to reconstruct the 360° shapes of a few objects with different surface roughness. In the experiments, the FPB system is positioned to measure the 3D shape of each object from a fixed view, and the measurements are then repeated by rotating the object to cover the entire surfaces. Typically, eight 3D surface shapes are acquired and stitched together to form a complete 360° 3D image. Figure 12 demonstrates the experimental results of two samples, where the first image in each row is the photo of the object, and the other images are selected views of the 3D image
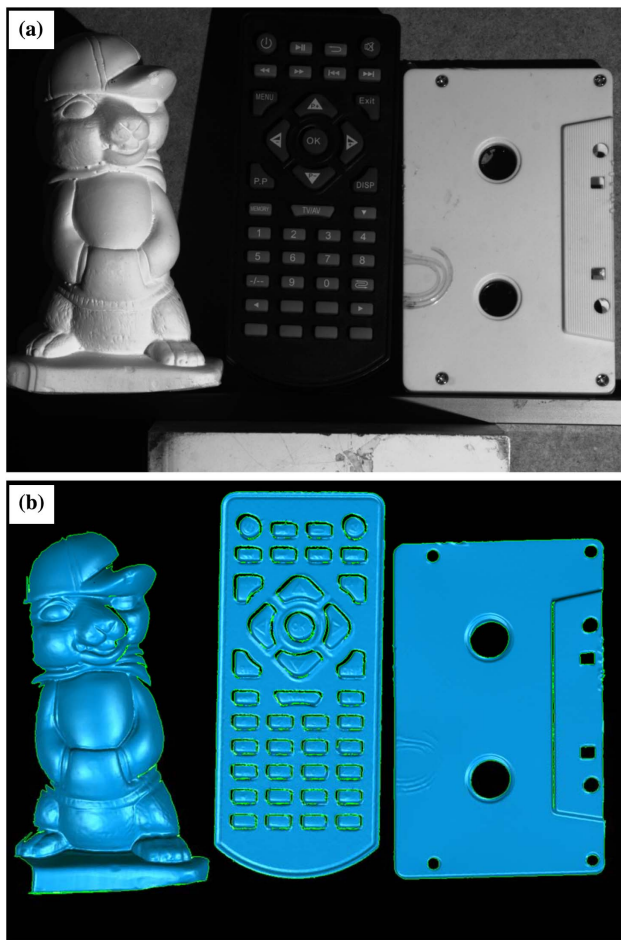
Fig. 11. 3D shape measurement of (a) multiple objects of interest and (b) the corresponding rendered plot of the 3D surface profiles.

reconstructed by the FPB system. Because the errors of the 3D coordinates in each view will accumulate and be magnified when stitching multiple views into a full 3D image, the errors must be very small for a successful 360° 3D reconstruction.

Accordingly, this experiment is an alternative and effective way to verify the accuracy and practicality of the proposed approach.

5. CONCLUSION

This paper presents a real-time, high-accuracy 3D imaging and shape measurement system based on the fringe projection technique. Through using an external trigger circuit to control the 3LCD projector, the proposed technique can capture three phase-shifted fringe patterns at a real-time speed of 25.6 fps. In addition, a fully connected DNN machine learning model is introduced to directly extract the phase distributions with high accuracy from the captured images. Although the speed of the proposed system is slower than some of the existing fast-speed 3D imaging and shape measurement techniques, the system provides substantially higher accuracy thanks to the DNN machine learning scheme.

The errors introduced into the real-time 3D imaging system can be from a variety of sources. Such possible error sources include a mismatch between the input data and the ground-truth labels in the datasets, noise presented in the data, calibration uncertainty, environmental interference, projector imperfection, camera model, and sensor limitations. In spite of that, the errors are very small compared with the measured quantities. Experiments have shown that the system is capable of performing 3D imaging and shape measurement of multiple objects with complex shapes at real-time speed and relative accuracy of 0.012%.

In recent years, there have been high demands for the 3D imaging and shape measurement techniques to possess a number of advanced technical features, particularly fast speed and high accuracy. The system presented in this paper has demonstrated the capability of satisfying these demands. It also shows promise for using deep machine learning to develop future 3D imaging and shape measurement systems with unprecedented performance.
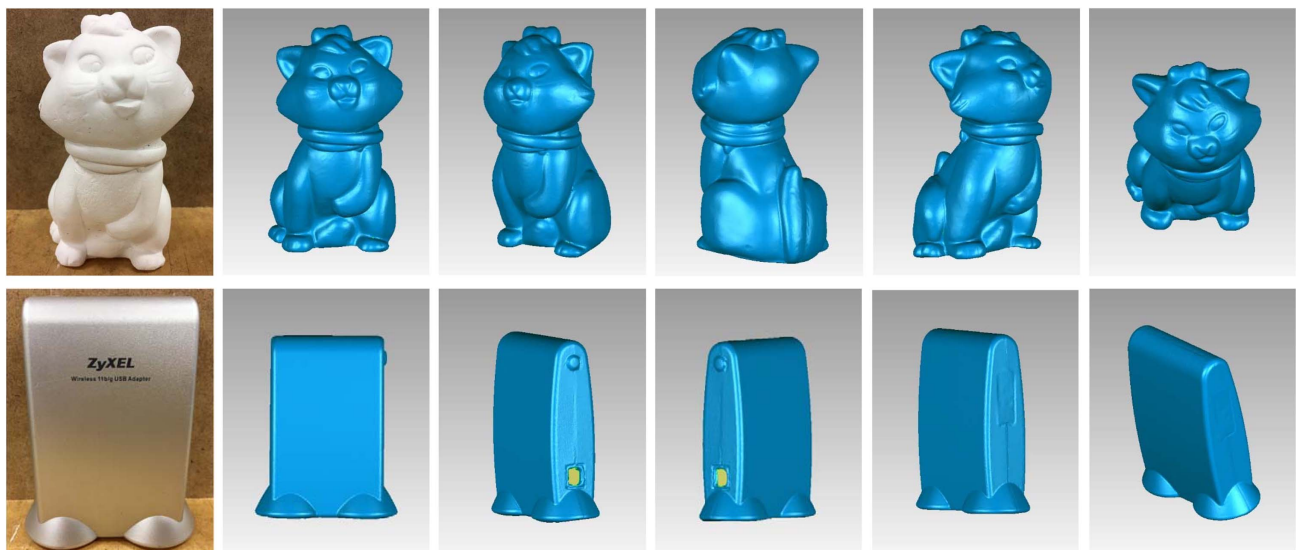
Fig. 12. 360° reconstructed models of two objects: a small cat sculpture and a Wi-Fi modem (see Visualization 3).

REFERENCES

1. S. Zhang, "High-speed 3D shape measurement with structured light methods: a review," Opt. Lasers Eng. **106**, 119–131 (2018).
2. R. Wu, Y. Chen, Y. Pan, Q. Wang, and D. Zhang, "Determination of three-dimensional movement for rotary blades using digital image correlation," Opt. Lasers Eng. **65**, 38–45 (2015).
3. Z. Wang, H. Kieu, H. Nguyen, and M. Le, "Digital image correlation in experimental mechanics and image registration in computer vision: similarities, differences and complements," Opt. Lasers Eng. **65**, 18–27 (2015).
4. H. Nguyen, H. Kieu, Z. Wang, and H. N. D. Le, "Three-dimensional facial digitization using advanced digital image correlation," Appl. Opt. **57**, 2188–2196 (2018).
5. X. Su and Q. Zhang, "Dynamic 3-D shape measurement method: a review," Opt. Lasers Engr. **48**, 191–204 (2010).
6. H. Nguyen, Z. Wang, P. Jones, and B. Zhao, "3D shape, deformation, and vibration measurements using infrared Kinect sensors and digital image correlation," Appl. Opt. **56**, 9030–9037 (2017).
7. L. Keselman, J. Woodfill, A. Jepsen, and A. Bhowmik, "Intel real-sense stereoscopic depth cameras," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2017), pp. 1–10.
8. Apple, "About Face ID advanced technology," https://support.apple.com/en-us/HT208108.
9. Zeiss, "3D scanning with blue LED fringe projection ZEISS COMET," https://www.zeiss.com/metrology/products/systems/optical-systems/3d-scanning/zeiss-comet.html.
10. GOM, "Industrial 3D scanning technology," https://www.gom.com/metrology-systems/atos.html.
11. S. Heist, P. Lutzke, I. Schmidt, P. Dietrich, P. Kuhmstedt, A. Tunnermann, and G. Notni, "High-speed three-dimensional shape measurement using GOBO projection," Opt. Lasers Eng. **87**, 90–96 (2016).
12. T. Weise, B. Leibe, and L. Gool, "Fast 3D scanning with automatic motion compensation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2007), pp. 1–8.
13. J. Hyun, G. Chiu, and S. Zhang, "High-speed and high-accuracy 3D surface measurement using a mechanical projector," Opt. Express **26**, 1474–1487 (2018).
14. C. Zuo, T. Tao, S. Feng, L. Huang, A. Asundi, and Q. Chen, "Micro Fourier transform profilometry (μFTP): 3D shape measurement at 10,000 frames per second," Opt. Lasers Eng. **102**, 70–91 (2018).
15. S. Rusinkiewicz, O. Hall-Holt, and M. Levoy, "Real-time 3D model acquisition," ACM Trans. Graph. **21**, 438–446 (2002).
16. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," Opt. Lasers Eng. **48**, 149–158 (2010).
17. K. Liu, Y. Wang, D. Lau, Q. Hao, and L. Hassebrook, "Dual-frequency pattern scheme for high-speed 3-D shape measurement," Opt. Express **18**, 5229–5244 (2010).
18. C. Zuo, Q. Chen, G. Gu, S. Feng, and F. Feng, "High-speed three-dimensional profilometry for multiple objects with complex shapes," Opt. Express **20**, 19493–19510 (2012).
19. 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. **34**, 416–418 (2009).
20. K. Yatabe, K. Ishikawa, and Y. Oikawa, "Compensation of fringe distortion for phase-shifting three-dimensional shape measurement by inverse map estimation," Appl. Opt. **55**, 6017–6024 (2016).
21. C. Mao, R. Lu, and Z. Liu, "A multi-frequency inverse-phase error compensation method for projector nonlinear in 3D shape measurement," Opt. Commun. **419**, 75–82 (2018).
22. Z. Lei, C. Wang, and C. Zhou, "Multi-frequency inverse-phase fringe projection profilometry for nonlinear phase error compensation," Opt. Lasers Eng. **66**, 249–257 (2015).
23. S. Gai, F. Da, and C. Liu, "Multiple-gamma-value based phase error compensation method for phase measuring profilometry," Appl. Opt. **57**, 10290–10299 (2018).
24. S. Ma, C. Quan, R. Zhu, and C. J. Tay, "Investigation of phase error correction for digital sinusoidal phase-shifting fringe projection profilometry," Opt. Lasers Eng. **50**, 1107–1118 (2012).
25. J. Yao, C. Xiong, Y. Zhou, H. Miao, and J. Chen, "Phase error elimination considering gamma nonlinearity, system vibration and noise for fringe projection profilometry," Opt. Eng. **53**, 094102 (2014).
26. F. Lu, S. Xing, and H. Guo, "Self-correction of projector nonlinearity in phase-shifting fringe projection profilometry," Appl. Opt. **56**, 7204–7216 (2017).
27. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, "Fringe pattern analysis using deep learning," Adv. Photon. **1**, 1 (2019).
28. K. Yan, Y. Yu, C. Huang, L. Sui, K. Qian, and A. Asundi, "Fringe pattern denoising based on deep learning," Opt. Commun. **437**, 148–152 (2019).
29. B. Lin, S. Fu, C. Zhang, F. Wang, S. Xie, Z. Zhao, and Y. Li, "Optical fringe patterns filtering based on multi-stage convolution neural network," arXiv: 1901.00361v1 (2019).
30. A. Figueroa and M. Rivera, "Deep neural network for fringe pattern filtering and normalization," arXiv: 1901.00361v1 (2019).
31. F. Hao, C. Tang, M. Xu, and Z. Lei, "Batch denoising of ESPI fringe patterns based on convolutional neural network," Appl. Opt. **58**, 3338–3346 (2019).
32. S. Jeught and J. Dirckx, "Deep neural networks for single shot structured light profilometry," Opt. Express **27**, 17091–17101 (2019).
33. A. Babaei, M. Saadatseresht, and J. Kofman, "Exponential fringe pattern projection approach to gamma-independent phase computation without calibration for gamma nonlinearity in 3D optical metrology," Opt. Express **25**, 24927–24938 (2017).
34. A. Kamagara, X. Wang, and S. Li, "Nonlinear gamma correction via normed bicoherence minimization in optical fringe projection metrology," Opt. Eng. **57**, 034107 (2018).
35. Z. Cai, X. Liu, X. Peng, Z. Zhang, H. Jiang, Y. Yin, and S. Huang, "Phase error compensation methods for high-accuracy profile measurement," Meas. Sci. Technol. **27**, 045201 (2016).
36. T. Hoang, B. Pan, D. Nguyen, and Z. Wang, "Generic gamma correction for accuracy enhancement in fringe-projection profilometry," Opt. Lett. **35**, 1992–1994 (2010).
37. C. Quan, W. Chen, and C. J. Tay, "Phase-retrieval techniques in fringe-projection profilometry," Opt. Lasers Eng. **48**, 235–243 (2010).
38. C. Zhang, H. Zhao, F. Gu, and Y. Ma, "Phase unwrapping algorithm based on multi-frequency fringe projection and fringe background for fringe projection profilometry," Meas. Sci. Technol. **26**, 045203 (2015).
39. M. Servin, J. Padilla, A. Gonzalez, and G. Garnica, "Temporal phase-unwrapping of static surfaces with 2-sensitivity fringe-patterns," Opt. Express **23**, 15806–15815 (2015).
40. H. Le, H. Nguyen, Z. Wang, J. Opfermann, S. Leonard, A. Krieger, and J. Kang, "Demonstration of a laparoscopic structured-illumination three-dimensional imaging system for guiding reconstructive bowel anastomosis," J. Biomed. Opt. **23**, 1–10 (2018).
41. S. Xing and H. Guo, "Temporal phase unwrapping for fringe projection profilometry aided by recursion of Chebyshev polynomials," Appl. Opt. **56**, 1591–1602 (2017).
42. H. Du and Z. Wang, "Three-dimensional shape measurement with an arbitrarily arranged fringe projection profilometry system," Opt. Lett. **32**, 2438–2440 (2007).
43. M. Vo, Z. Wang, B. Pan, and T. Pan, "Hyper-accurate flexible calibration technique for fringe-projection-based three-dimensional imaging," Opt. Express **20**, 16926–16941 (2012).
44. O. Packer, L. Diller, J. Verweij, B. Lee, J. Pokorny, D. Williams, D. Dacey, and D. Brainard, "Characterization and use of a digital light projector for vision research," Vis. Res. **41**, 427–439 (2001).
45. L. Hornbeck, "Digital Light Processing for high-brightness high-resolution applications," Proc. SPIE **3013**, 27–40 (1997).
46. H. Nguyen, D. Nguyen, Z. Wang, H. Kieu, and M. Le, "Real-time, high-accuracy 3D imaging and shape measurement," Appl. Opt. **54**, A9–A17 (2015).
47. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
48. A. Hoerl and R. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," Technometrics **12**, 55–67 (1970).
49. G. Golub, P. Hansen, and D. O'Leary, "Tikhonov regularization and total least squares," SIAM J. Matrix Anal. Appl. **21**, 185–194 (1999).
50. R. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Statist. Soc. B **58**, 267–288 (1996).
51. H. Zuo and T. Hastie, "Regularization and variable selection via the elastic net," J. R. Statist. Soc. B **67**, 301–320 (2005).