PHOTONICS Research

Addressable structured light system using metasurface optics and an individually addressable VCSEL array

CHENYANG WU,[1,2,†] (iD) XUANLUN HUANG,[1,2,†] (iD) YIPENG JI,[2] (iD) TINGYU CHENG,[2] JIAXING WANG,[2] NAN CHI,[1,3] (iD) SHAOHUA YU,[1,3] AND CONNIE J. CHANG-HASNAIN[1,2,4,*] (iD)

[1]School of Information Science and Technology, Fudan University, Shanghai 200433, China
[2]Berxel Photonics Co., Ltd., Shenzhen 518071, China
[3]Peng Cheng Laboratory, Shenzhen 518055, China
[4]Chinese University of Hong Kong, Shenzhen 518172, China
[†]These authors contributed equally to this work.
*Corresponding author: connie.chang@berxel.com

Structured-light (SL) based 3D sensors have been widely used in many fields. Speckle SL is the most widely deployed among all SL sensors due to its light weight, compact size, fast video rate, and low cost. The transmitter (known as the dot projector) consists of a randomly patterned vertical-cavity surface-emitting laser (VCSEL) array multiplicated by a diffractive optical element (DOE) with a fixed repeated pattern. Given that the separation of any two speckles is only one known and fixed number (albeit random), there are no other known scales to calibrate or average. Hence, typical SL sensors require extensive in-factory calibrations, and the depth resolution is limited to 1 mm at ∼60 cm distance. In this paper, to the best of our knowledge, we propose a novel dot projector and a new addressable SL (ASL) 3D sensor by using a regularly spaced, individually addressable VCSEL array, multiplicated by a metasurface-DOE (MDOE) into a random pattern of the array. Dynamically turning on or off the VCSELs in the array provides multiple known distances between neighboring speckles, which is used as a "built-in caliper" to achieve higher accuracy of depth. Serving as a precise "vernier caliper," the addressable VCSEL array enables fine control over speckle positions and high detection precision. We experimentally demonstrated that the proposed method can result in sub-hundred-micron level precision. This new concept opens new possibilities for applications such as 3D computation, facial recognition, and wearable devices.　© 2024 Chinese Laser Press

https://doi.org/10.1364/PRJ.516942

1. INTRODUCTION

Structured-light (SL) based 3D sensing technology has been widely used in reflective surface detection [1], machine vision [2], facial recognition [3,4], human-computer interaction [5], motion-sensing games [6], and biomimetic robotics [7] due to the video rate of high spatial resolution 3D images with a large field of view (FOV) as well as low power consumption and compact size. For speckle SL, the speckle pattern is generated by a dot projector consisting of a randomly patterned vertical-cavity surface-emitting laser (VCSEL) array multiplicated by a diffractive optical element (DOE) with a fixed repeated pattern [1–4]. With a CMOS receiver at a fixed known distance away from the dot projector, the speckle image at the object is recorded. The depth information is obtained from the changes of dot pattern positions based on the triangulation principle. The randomness of the VCSEL array position enables the

identification of a unique local area within one DOE order. However, this method is conventionally static and non-programmable. The spacing of speckles is the only known number used to calculate depth, resulting in low flexibility and limited depth resolution to ∼1 mm. In addition, a complex calibration procedure is required to cover a wide range of operation temperatures and observation distances.

In recent years, metasurface-DOEs (MDOEs) based on nanofabrication techniques are utilized to generate ultra-thin, miniaturized, and device-integrated SL [8–12]. Ni *et al.* designed a polarization-insensitive metasurface capable of generating SL speckle patterns over a 120° × 120° FOV [8]. Wang *et al.* integrated GaAs metasurfaces with standard VCSEL chips, enabling the generation of SL at an ultra-compact chip scale [9]. Kim *et al.* proposed a metasurface-enhanced SL depth-sensing system that scattered a high-density (∼10K) dot array over a 180° FOV [10]. Jing *et al.* presented a

single-shot 3D reconstruction method using metasurface-based SL point clouds for complex point cloud calculations [11]. The studies mentioned above have laid a solid foundation for the advancement of high-performance SL systems based on micro-nano optics. Nevertheless, following the fabrication of metasurfaces, the resulting speckle patterns still demonstrate static and non-programmable characteristics, and a complex calibration procedure is still required. Dynamic encoding of speckle patterns, including position variation and scaling modulation of speckles, remains an underexplored research area.

In this paper, a novel addressable SL (ASL) 3D sensor is proposed by using a regularly spaced, individually addressable VCSEL (IA-VCSEL) array, multiplied by an MDOE into a random pattern of the array. The designed IA-VCSEL array is composed of 8 × 8 VCSELs with 100 μm pitch, uniformly spaced. These VCSELs share the same cathode, but anodes are separated for on–off switching, allowing for the control of switch states of any individual or multiple VCSELs in the array through pre-encoding or real-time encoding, with response time in the ~10 ns range. The SOI reflective MDOE is characterized by a 3 μm thick intermediate SiO_2 layer and a 340 nm height top layer of single-crystal silicon for etching through elaborated design and selection of nanorod dimensions and rotations. The MDOE is designed to split each VCSEL into a random speckle pattern, which enables identification of a unique local area. By turning on or off the VCSELs in the array dynamically, distinct VCSELs within the array generate speckle patterns with dynamic position and scaling variation. Besides, the ASL 3D sensor provides fine control over speckle positions and multiple known distances between neighboring speckles, which is used as a "built-in caliper" for fine calibration to achieve higher accuracy of the depth. The incorporation of a "built-in caliper" potentially enables a departure from the intricate calibration processes traditionally required in SL 3D sensing. The following are the potential benefits.

(i) Dynamic calibration reference. The IA-VCSEL array acts as an inherent calibration reference. By dynamically controlling the activation and deactivation of VCSELs, the system establishes multiple known distances between speckles. This dynamic calibration reference adapts in real-time, reducing the need for predefined calibration setups for different distances.

(ii) Adaptive distance calibration. Traditional SL sensors necessitate extensive calibration efforts to accommodate varying observation distances, often resulting in reduced depth resolution. The "built-in caliper," through dynamic VCSEL control, allows for adaptive calibration based on the desired observation distance, mitigating the need for meticulous adjustments and ensuring accurate depth perception across different ranges.

(iii) Temperature-independent calibration. The addressable VCSEL array offers temperature robustness by providing a continuous range of known distances for calibration. This adaptability enables the system to maintain accuracy under different temperature conditions without requiring recalibration. The inherent flexibility compensates for temperature-induced variations, contributing to a more reliable and stable 3D sensing solution.

(iv) Reduced manual calibration efforts. Unlike conventional SL sensors that demand extensive manual calibration procedures, the built-in caliper streamlines the process. The dynamic control over VCSELs automates aspects of calibration, saving time and effort, and reducing the dependency on meticulous manual adjustments.

In essence, the integration of a "built-in caliper" not only simplifies the calibration process but also offers adaptability, temperature robustness, and precision control, addressing the challenges associated with traditional SL sensor calibration over varying distances and temperatures. We experimentally demonstrated that the proposed method can get sub-hundred-micron level precision, which is much better than the conventional method. Moreover, the prototype model for AR/VR glass application is proposed, potentially enabling eye movement tracking. The novel concept opens up new possibilities for applications such as 3D computation, facial recognition, and wearable devices, and promotes the development of consumer electronic products based on 3D imaging technologies.

2. METHOD AND DESIGN

A. Principle of ASL

To realize the capability of ASL, we initiate the process by collimating the light emitted from the IA-VCSEL array with a lens. Subsequently, the MDOE is positioned along the path of the collimated light, and the separation distance between the MDOE and the VCSEL array light source is adjusted to project the desired encoded light sources onto the coverage area of the MDOE. The MDOE is designed to split each VCSEL into a random speckle pattern, which involves the encoded information. Figure 1(c) illustrates the schematic optical paths of the ASL based on the reflective MDOE with the IA-VCSEL array, alongside the corresponding far-field encoded SL speckle patterns. Here, red and blue circles of the VCSEL array symbolize the on and off states, respectively. In Fig. 1(a) from top to bottom, source patterns are created by 1, 2, and 24 activated VCSEL apertures from the VCSEL array, serving the purpose of position encoding. Through the control of switch states at different positions within the IA-VCSEL array, we can effectively encode VCSEL light spot positions projected onto the MDOE, resulting in the far-field projection of encoded SL speckle patterns. It is noteworthy that the position encoding of VCSEL array apertures directly corresponds to the encoding of individual spots within the SL speckle pattern, including replication and expansion. This facilitates the achievement of a much higher resolution with a fixed VCSEL pitch for fine
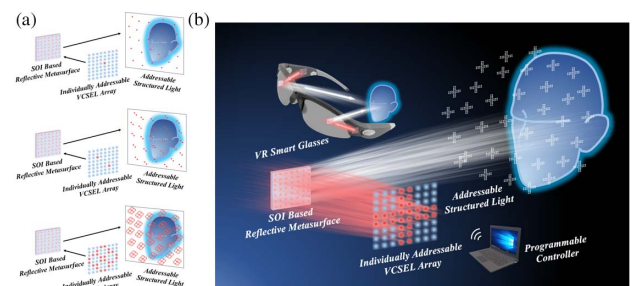
Fig. 1. (a) Schematic diagram of the ASL system based on reflective MDOE with the IA-VCSEL array; (b) prototype model for AR/VR smart glasses for eye movement tracking based on the proposed method.

calibration to attain depth information using the triangulation method. There are many interesting applications including AR/VR glasses for eye movement tracking, as illustrated in Fig. 1(b), where the IA-VCSEL array and SOI reflective MDOE are mounted on the frame and lens of the AR/VR glasses, respectively.

B. Design of SOI Reflective MDOE

The SOI reflective MDOE is composed of a top layer of amorphous silicon cuboid nanorods with a height of 340 nm, a 3 μm thick intermediate SiO_2 layer, and a bottom layer of silicon substrate. Due to the strong coupling effect between the amorphous silicon nanorods and the optical wave, each nanorod can be regarded as a waveguide with truncated ends. The optical field is mainly confined within the high refractive index nanorods. Therefore, the optical influence of each individual nanorod is principally determined by its geometric parameters, specifically the length and width, while factors related to the dimensions and orientations of neighboring nanorods can be feasibly disregarded [13,14]. Consequently, each unit cell within the lattice can be treated as an individual pixel for independent design. By elaborate manipulation of the geometrical parameters of the nanorods, the phase distribution of the incident light wavefront can be modulated, achieving a comprehensive phase coverage spanning from 0 to 2π. We employ two approaches, namely, the propagation phase (by changing the length and width of the nanorods) and Pancharatnam–Berry (PB) phase (by changing the rotation angle of the nanorods) to design the SOI reflective MDOE. A schematic diagram of the proposed SOI reflective MDOE unit cell based on propagation phase and PB phase is shown in Figs. 2(a) and 2(b), respectively, where t_g and t_m denote the height of the nanorods and the thickness of the SiO_2 middle layer, respectively; P and θ represent the lattice constant of the unit cell and the rotation angle of the nanorods, respectively. The Gerchberg–Saxton (GS) algorithm [15] is employed to calculate the phase distribution, enabling the generation of the designed random/pseudo-random SL speckle patterns in the far-field.

For design of the propagation phase, 2D scanning of the length and width parameters of the SOI reflective MDOE unit cell is conducted using the finite-difference time-domain

(FDTD) simulation software (Lumerical Inc.). Dimensions of P, t_g, and t_m are set to 500 nm, 340 nm, and 3 μm, respectively. 940 nm plane wave is used as the incident light to perform 2D scanning of the length and width of a single nanorod unit cell, ranging from 100 nm to 450 nm. The reflection coefficient and reflection phase distributions of the unit cell are shown in Figs. 2(c) and 2(d), respectively, with the reflection coefficient ranging from 0 to 1 and the reflection phase ranging from $-\pi$ to π in radians. The results demonstrate that by varying the length and width parameters of the unit cell, the comprehensive coverage of reflection phase from $-\pi$ to π can be achieved, encompassing 2π range. The reflection coefficients and discretized phase distributions of the selected 16 kinds of unit cells are presented in Fig. 2(e). The curves formed by the black circular points and red square points represent the corresponding reflection coefficients and discretized phase distributions, respectively. It is evident that the majority of the chosen unit cells exhibit reflection coefficients exceeding 0.95.

For design of PB phase, according to the fundamental principles of PB phase [16], the geometric phase $\varphi(x, y)$ can be modulated by employing cuboid nanorods with identical dimensions but varying rotation angles $\theta(x, y)$. The interrelationship between these two parameters is $\theta(x, y) = (1/2) \cdot \varphi(x, y)$, implying that the rotation angle $\theta(x, y)$ of the nanorods within the incident plane, spanning the range of 0 to π, facilitates the transmission of the geometric phase $\varphi(x, y)$ across the reflection plane within the range of 0 to 2π. Similar to the design for propagation phase, we set the lattice constant P of the nanorods to 500 nm and t_g to 340 nm. Under incident light with wavelength of 940 nm, we select a set of geometric dimensions which yield high polarization conversion efficiency, specifically $L = 355$ nm and $W = 165$ nm. Subsequently, the geometric parameters of the meta-structures are derived through the conversion of the designed phase into rotation angles, and the MDOE is fabricated using conventional standard electron beam lithography (EBL) in conjunction with the etching process.
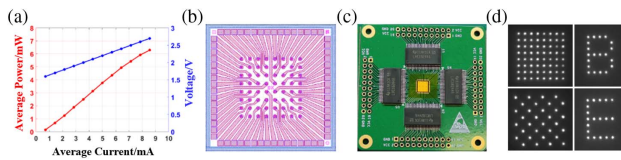
C. Design of the IA-VCSEL Array

The designed IA-VCSEL array is composed of 8×8 VCSELs with 100 μm pitch. These VCSELs share a common cathode, but the anodes are independently isolated for precise on–off switching. The designed VCSEL array chip facilitates individual or multiple control of VCSELs within the 8×8 VCSEL array through a programmable controller, enabling control over the pattern encoding of VCSEL. The IA-VCSEL array mainly consists of the IA-VCSEL array driver board and the pattern generation FPGA board. The IA-VCSEL array driver board is designed to power and control the 8×8 IA-VCSEL chip, enabling independent activation of 64 apertures separately. The operating voltage range is about 1.6 V to 3.6 V, ensuring compatibility and stability across various power conditions; the pattern encoding FPGA board is designed to generate customized light-up patterns. The FPGA board supports 8×8 graphical mapping encoding, allowing for complex and flexible control over the laser patterns. It can produce different pulse widths and repetition frequencies as needed, with a minimum pulse width of 10 ns. The light output-current-voltage (LIV) characteristic curve of the single VCSEL aperture is measured
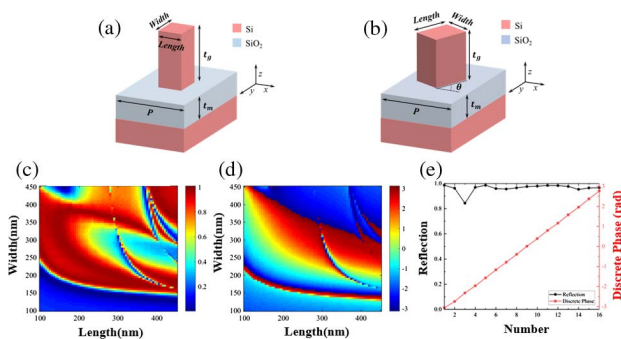
Fig. 2. Schematic diagram of the proposed SOI reflective MDOE unit-cell based on (a) propagation phase and (b) PB phase; distributions of (c) reflection coefficient and (d) reflection phase of the unit cell based on propagation phase; (e) reflection coefficients and discretized phase distributions of the selected unit cells.

Fig. 3. (a) LIV curve of single VCSEL aperture of the IA-VCSEL array; (b) wire bonding of the IA-VCSEL array; (c) assembled PCB of the IA-VCSEL array; (d) infrared images of laser spot pattern encoding using IA-VCSEL arrays.

Fig. 4. (a) Far-field distribution simulations of the designed random/ pseudo-random speckle patterns. SEM images of fabricated MDOE: (b) side view of the nanorods; (c) and (d) top view of the MDOE designed by propagation phase and PB phase, respectively; (e)–(g) 3D perspective of (b) to (d), respectively.

as shown in Fig. 3(a). The left and right vertical axes denote the average output optical power (in red) and voltage (in blue), respectively. Notably, the threshold voltage for a single aperture is measured at approximately 1.5 V, accompanied by the threshold current of 0.7 mA. The approximate resistance of a single aperture is measured to be approximately 70 Ω. The VCSEL array is interconnected and integrated onto an external printed circuit board (PCB) using a wire bonding technique. The wire bonding schematic is shown in Fig. 3(b). It is clear that all the VCSEL chips within the 8×8 array are individually linked to the surrounding 16×4 wire bonding pads without any cross-interference or mutual influence. This interconnection of the wire bonding pads with the external driving circuitry enables independent and addressable control of the VCSEL array. The ultimate assembled PCB of the IA-VCSEL array is presented in Fig. 3(c).

The manufactured IA-VCSEL array can be effectively employed for laser spot pattern encoding. Following collimation with a lens, the generated far-field patterns are captured employing an infrared (IR) camera, as illustrated in Fig. 3(d). Proceeding from left to right, these images correspond to laser-encoded patterns created utilizing 64, 32, 20, and 19 VCSEL apertures sourced from the 8×8 VCSEL array. Notably, the designed VCSEL array enables precise control over the activation and deactivation of individual or multiple VCSELs, attainable through either pre-encoding or real-time encoding methods. This capability ensures precise modulation over both the quantity and spatial positioning of VCSEL apertures projected onto the metasurface, thereby facilitating the encoding functionality of the SL generated by the metasurface.

3. EXPERIMENTAL RESULT

A. Fabrication of the SOI MDOE

Three distinct categories of random and pseudo-random SL speckle patterns have been designed through the utilization of the GS algorithm. According to the Huygens-Fresnel principle, the far-field spatial distributions for these three patterns are calculated via Fresnel integral simulations. Figure 4(a) illustrates the far-field distributions of the simulated speckle patterns, where the range of the x and y axes corresponds to an FOV spanning $45° \times 45°$.

The designed SOI reflective MDOE is fabricated utilizing standard EBL in conjunction with the lift-off process. Figures 4(b)–4(g) present scanning electron microscope (SEM) images of the resulting MDOE. Figure 4(b) displays a side view of the nanorods, with a height measuring 334 nm, closely aligning with the intended height of 340 nm. Figures 4(c) and 4(d)
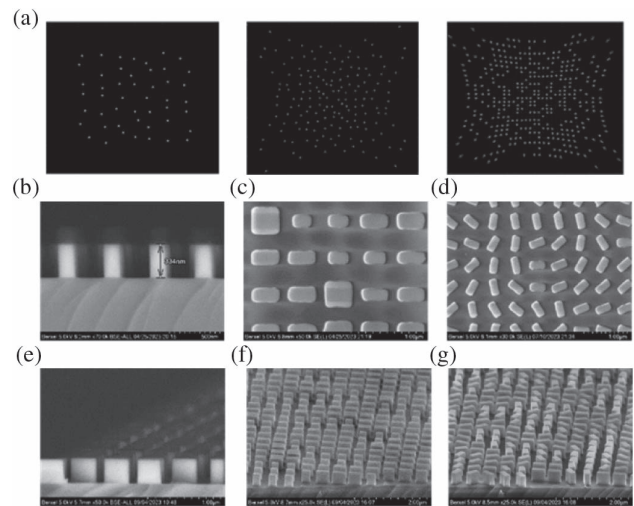
provide a top view of the MDOE design based on propagation phase and PB phase, respectively. Figures 4(e)–4(g) offer corresponding 3D perspectives of Figs. 4(b)–4(d), respectively. These SEM images of the MDOE verify the precision maintained during the EBL processing. For the assessment of the fabricated MDOE sample, a single commercial collimated 940 nm VCSEL source is employed, and the far-field diffraction patterns are captured using an infrared camera, which exhibits consistency with the results derived from Fresnel integral simulations. The pronounced presence of energy in the 0th diffraction order is primarily attributed to fabrication accuracy. Moreover, further enhancement can be achieved through improvement of fabrication precision and post-processing methods involving the utilization of the 0th diffraction order with image processing techniques, among other potential refinements.

B. Verification of ASL

Next, we experimentally verify the effectiveness and superiority of the ASL in achieving speckle pattern encoding and speckle density control. The experimental setup is illustrated in Fig. 5(a). The IA-VCSEL array is connected to a programmable controller, which generates pre-encoded or real-time encoded laser beams. The encoded laser beams are collimated and projected onto the SOI reflective MDOE at a slight incident angle. The diffraction speckle pattern reflected by the MDOE is received on a screen (acting as a reference plane) and is dynamically detected and captured with an IR camera.

We performed experimental validation utilizing the 2×2 and 3×3 configurations of the VCSEL array situated within the central region of the 8×8 IA-VCSEL array. By finely tuning the experimental setup, the collimated beams originating from the VCSEL array are precisely directed onto the MDOE, ensuring comprehensive coverage of the entire MDOE. Real-time encoding control of the laser array is executed through the programmable controller. The resulting far-field diffraction
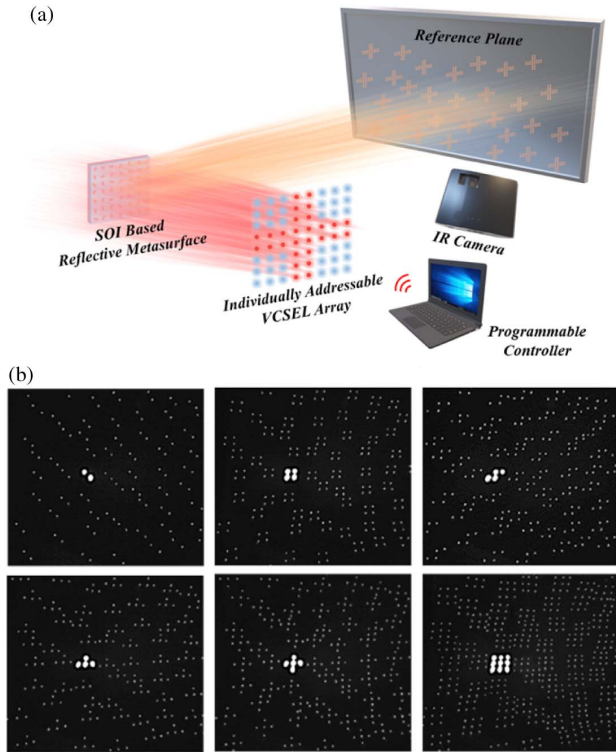
Fig. 5. (a) Experimental setup schematic for ASL system verification. (b) Measured far-field infrared images of ASL.

Fig. 6. Calibrated unique distances of SL generated by (a) a single VCSEL and (b) and (c) 2×3 and 3×3 IA-VCSEL arrays; and (d) schematic diagram of calibrated unique distances of $N \times M$ IA-VCSEL array.

speckle patterns under diverse laser array encoding configurations are exhibited in Fig. 5(b). The 0th diffraction order, characterized by its strong intensity, is centrally positioned within the programmable speckle patterns, aligning precisely with the laser array encoding configurations facilitated by the independently addressable VCSEL array.
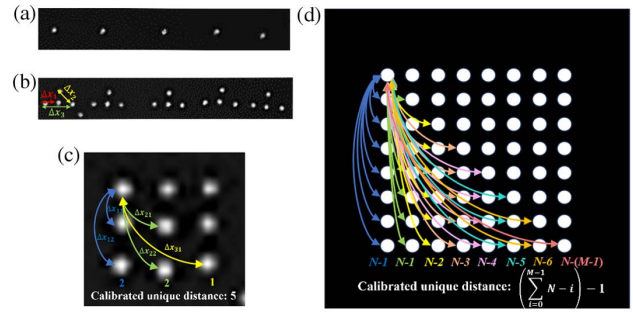
The experimental results evidently validate the system's effectiveness in encoding diffraction speckle patterns and regulating speckle density. As the number of involved VCSEL arrays increases, there is a corresponding increase in the diversity of encoded far-field diffraction speckle patterns. Given the predetermined maximum dimensions of the MDOE structure at 1 mm by 1 mm, the imaging performance of the encoded far-field reflection patterns, when utilizing VCSEL arrays sized up to and including 3×3, proves to be satisfactory. For beam encoding with larger VCSEL arrays, such as 8×8 VCSEL array or larger configurations, a correspondingly expanded larger MDOE area becomes a prerequisite to ensure complete coverage of the light source.

C. Applications of ASL
In this section, we will list several potential applications of ASL, including but not limited to the precise "built-in caliper," adaptable speckle density modulation, enhancement of diversity in feature pattern, and improved speckle image matching accuracy.

1. Precise "Built-In Caliper"
Serving as precise "built-in caliper," the separation between speckles of ASL can be further precisely subdivided.

The distribution of ASL provides much more variable scales of speckle spacing, corresponding to the encoding of the IA-VCSEL array, which is fundamentally different from conventional fixed speckles with no scale. The measured SL generated by a single VCSEL and $2 \times 3 / 3 \times 3$ IA-VCSEL array is shown in Figs. 6(a)–6(c), where Δx_i represents the variable calibrated unique distance (CUD), and Fig. 6(c) is enlarged for clarity. There are three and five kinds of CUDs in Figs. 6(b) and 6(c), respectively. As the size of VCSEL array increases, the variety of CUD also increases. Figure 6(d) demonstrates the schematic diagram of CUD generated by $N \times M$ IA-VCSEL array ($N \geq M$), which is calculated by

$$\mathrm{CUD} = \left(\sum_{i=0}^{M-1} N - i \right) - 1, \tag{1}$$

where N and M indicate the number of rows and columns of the VCSEL array, respectively. Therefore, there are 35 kinds of CUDs for the 8×8 IA-VCSEL array. According to the triangulation principles, the depth value z of each pixel can be calculated using the following equation:

$$z = \frac{z_0}{1 + (z_0 / f \cdot b) \cdot D}, \tag{2}$$

where b, f, and z_0 are built-in parameters of the baseline, focal length, and distance to the reference plane, respectively, and D denotes the disparity. The process of obtaining D involves projecting calibrated speckle patterns onto a plane at a specific distance, represented in a reference image. Subsequently, when these speckle patterns are projected onto the surface of an object and reflected back, they create a deformation map. The computation of disparity entails selecting a speckle point as the center and choosing a corresponding speckle image block. Through a search on the reference image by the specific algorithm, a matching block is identified, and by calculating the offset between these two image blocks, D of the speckle point is determined.

The addressable VCSEL array potentially offers temperature robustness by providing a continuous range of known distances of speckles for calibration. This inherent flexibility potentially compensates for temperature-induced variations, enabling the system to maintain accuracy under different temperature conditions with a joint reconstruction algorithm. Next, we

establish a mathematical model for temperature drift correction for depth information reconstruction. We can introduce temperature-related terms to correct depth measurement biases caused by temperature.

The depth information calculation model without considering temperature drift is represented as follows:

$$Y = \Phi \cdot X + N_0, \tag{3}$$

where Y represents the measured depth information, X denotes the true disparity information to be reconstructed, Φ is the system observation matrix capturing the geometric relationship between depth and disparity, and N_0 encompasses system noise and depth errors. The reconstruction of disparities X is achieved by solving the L_1 regularization problem:

$$\hat{X} = \underset{X}{\mathrm{argmin}}\{\|Y - \Phi X\|_F^2 + \beta\|X\|_1\}, \tag{4}$$

where $\hat{X}$ is the reconstructed true disparity, $\|\cdot\|_F^2$ is the Frobenius norm of the matrix, and β is the regularization parameter.

The expanded depth information calculation model considering temperature drift is represented as follows:

$$Y = \Phi \cdot X + N_0 + N_T, \tag{5}$$

where N_T represents the temperature-dependent deviation, reflecting the depth measurement bias caused by temperature variations.

Next, we need to establish a model to describe the relationship between the temperature drift term N_T and the temperature T. Assuming this relationship can be represented by a simple linear model (more complex modeling can be performed based on actual conditions):

$$N_T = k \cdot (T - T_0), \tag{6}$$

where k is the temperature sensitivity coefficient, representing the amount of depth information bias per unit temperature change. T_0 is the reference temperature.

Temperature correction and depth information reconstruction. Combining the above two models, we obtain the depth information calculation model including temperature correction:

$$Y = \Phi \cdot X + N_0 + k \cdot (T - T_0). \tag{7}$$

To reconstruct the true disparity information $\hat{X}$ from the measured values with temperature bias Y, we need to modify the original L_1 regularization problem to include the temperature correction term. The modified L_1 regularization problem becomes

$$\hat{X} = \underset{X}{\mathrm{argmin}}\{\|Y - \Phi X - k(T - T_0)\|_F^2 + \beta\|X\|_1\}. \tag{8}$$

By solving this modified problem, we can obtain the reconstructed disparity information $\hat{X}$ with temperature drift correction. Considering the calibrated unique distance (CUD) of Δx_i and the corresponding depth information Y_{IMi} calculated from the image matching algorithm, and incorporating the temperature drift term N_{Ti}, the ground truth depth information Y_{GTi} is expressed as

$$\begin{cases} Y_{IM1} = Y_{GT1} + N_1 + N_{T1} \\ Y_{IM2} = Y_{GT2} + N_2 + N_{T2} \\ \quad\vdots \\ Y_{IM1} = Y_{GTi} + N_i + N_{Ti} \end{cases}, \tag{9}$$

where N_i represents the differences between measured and ground truth depth information, encompassing noise, system errors, and calibration errors. The temperature drift term N_{Ti} is introduced to account for the bias in depth measurement due to temperature variations. The joint reconstruction model incorporating temperature drift correction is formulated as

$$\begin{bmatrix} Y_{IM1} \\ Y_{IM2} \\ \vdots \\ Y_{IMi} \end{bmatrix} = \begin{bmatrix} Y_{GT1} \\ Y_{GT2} \\ \vdots \\ Y_{GTi} \end{bmatrix} + \begin{bmatrix} N_1 + N_{T1} \\ N_2 + N_{T2} \\ \vdots \\ N_i + N_{Ti} \end{bmatrix}, \tag{10}$$

which can be compactly expressed as

$$Y_{IM} = Y_{GT} + N + N_T. \tag{11}$$

To reconstruct the ground truth depth information Y_{GT} jointly, considering temperature drift, we modify the original L_1 regularization optimization problem as follows:

$$\begin{aligned} \hat{Y}_{GT} = \underset{Y_{GT}}{\mathrm{argmin}}\Big\{ \sum_i \big\| Y_{IMi} - (Y_{GTi} + N_{Ti}) \big\|_2^2 \\ + \beta \big\| Y_{GT1}, Y_{GT2}, \cdots, Y_{GTi} \big\|_{2,1} \Big\}, \end{aligned} \tag{12}$$

where

$$\begin{aligned} &\| Y_{GT1}, Y_{GT2} \cdots, Y_{GTi} \|_{2,1} \\ &= \sum_j (|(Y_{GT1})_j|^2 + |(Y_{GT2})_j|^2 + \cdots + |(Y_{GTi})_j|^2)^{1/2} \end{aligned} \tag{13}$$

and $(\cdot)_j$ represents the jth column. The optimization problem can be solved using regular regularization algorithms such as the iterative threshold algorithm (ITA) and complex approximated message passing (CAMP) algorithm [17,18]. As the relationship between temperature drift and the resulting speckle distance variation exhibits a well-behaved function, we can potentially install a thermometer sensor module at the TX end to detect temperature variations. The data can then be fed back into the calibration model, enabling temperature-independent calibration.

The "built-in caliper" effect potentially enables further fine control over speckle positions and enhancement of depth precision.

(i) Increased constraints for improved disparity accuracy. The expansion of each speckle pointing into an array with computable CUD provides additional depth reference information for the stereo disparity calculation. This diversification of information allows the system to leverage more constraints during the calculation of disparity, mitigating the cumulative effects of errors and thereby enhancing measurement accuracy.

(ii) Flexibility in disparity calculation with multiple reference distances. With each speckle point having computable CUD to other points, the system gains the flexibility to choose the most suitable distance during the calculation of disparity,
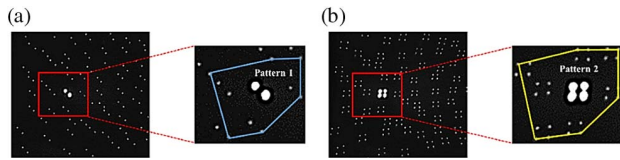
Fig. 7. Feature pattern extraction corresponding to scenarios with (a) $N = 6$, $k = 2$; (b) $N = 6$, $k = 4$.

adapting to various depth ranges. This flexibility ensures more accurate disparity calculations, unaffected by changes in depth.

(iii) Improved calibration accuracy. Calibration is a crucial step in ensuring measurement precision in stereo disparity calculations. The availability of multiple known distances enables the system to perform more accurate calibration and reduce systematic errors.

(iv) Enhanced adaptability for complex scene processing. The multiple CUD selection empowers the system to adapt to diverse scenes and surface characteristics, particularly when dealing with complex object surfaces. This capability is pivotal for handling intricate topologies and surface shapes, contributing to the maintenance of high-precision stereo disparity calculations.

In summary, the "built-in caliper" effect provides richer and more flexible information for stereo disparity calculations, significantly improving the accuracy of depth measurements.

2. Adaptable Speckle Density Modulation

Speckle density can be adaptively adjusted, proportional to the number of IA-VCSELs in the array, considering detection requirements based on factors such as scene size, detection distance, and accuracy requirements, as illustrated in Fig. 5(b). This adaptive control effectively improves detection precision, flexibility, and resource utilization.

3. Enhanced Diversity in Feature Pattern

The use of IA-VCSEL arrays effectively increases the diversity of speckle point feature patterns, thereby improving depth accuracy through methods such as averaging and other complex computations in a single measurement. Setting the number of feature pattern vertexes as N and the repetition count of single-point encoding for the IA-VCSEL array as k, the maximum numbers of feature patterns that can be generated for a single VCSEL and an IA-VCSEL array are 1 and k^N, respectively. The values corresponding to scenarios with $N = 6$, $k = 2$ and $N = 6$, $k = 4$ are illustrated in Figs. 7(a) and 7(b), respectively.

4. Improved Accuracy of Speckle Image Matching

The principle of spatial 3D reconstruction using SL algorithms is based on matching deformed speckle images reflected from the detected target with a known speckle reference image, considering the known internal parameters. By computing the displacement between two image blocks, the corresponding optical parallax is determined, enabling the subsequent calculation of depth values. The core of the algorithm lies in image matching technology, directly influencing the accuracy of the detection. Below, we employ the classical normalized cross-correlation (NCC) image matching algorithm for validation.
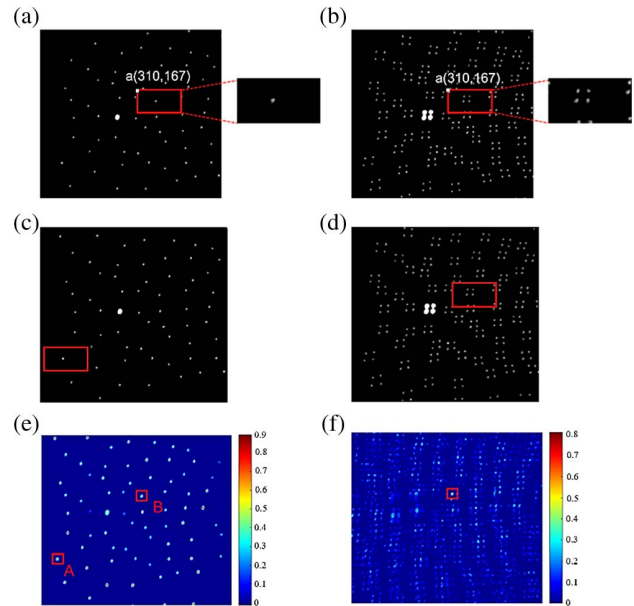
Fig. 8. Selected speckle for image matching generated by the (a) single VCSEL and (b) IA-VCSEL array; speckle image matching results by the (c) single VCSEL and (d) IA-VCSEL array; (e) and (f) corresponding correlation coefficient distributions of (c) and (d), respectively.

NCC demonstrates robustness to noise, is unaffected by changes in brightness, and offers high matching accuracy. The algorithm's steps involve selecting a speckle region of a certain size for matching. For all possible grayscale arrays within the reference speckle image, a pixel-wise comparison is performed using a similarity metric, and the results are normalized to a range of [0,1]. Values closer to 1 indicate higher correlation and greater similarity between the two images, while lower values imply lower similarity or mismatches. We chose speckle images generated by a single VCSEL and an IA-VCSEL array, selecting the same region at the same position for matching. As shown in Figs. 8(a) and 8(b), the selected region's top-left reference pixel coordinates are (310,167).

Using the NCC algorithm, we calculate the maximum correlation coefficient to identify the position of the matching speckle image in the reference image. The matching results are depicted by the red rectangles in Figs. 8(c) and 8(d), and the corresponding correlation coefficient distributions are shown in Figs. 8(e) and 8(f). The identified reference pixel coordinates for the matching results are (51,338) and (310,167), respectively, with maximum correlation coefficients of 0.8972 and 0.8108. It is evident that the use of a single VCSEL speckle image results in a mismatch, whereas the IA-VCSEL array speckle image leads to a correct match. In Fig. 8(e), there are multiple reference pixel coordinate positions with high correlation coefficients, such as pixels A and B. Pixel A represents an incorrect result calculated by the algorithm, while point B indicates the correct reference point location, with correlation coefficients of 0.8972 and 0.8926, respectively. This demonstrates that there are multiple similar image blocks in the reference image, which can easily lead to matching errors. In contrast, Fig. 8(f) has only one position with a significantly

higher correlation coefficient, leading to a noticeable improvement in matching accuracy.

D. Depth Detection Experiment

The real-time ASL has the capacity for complex encoding and decoding computations, thus enhancing the precision of 3D reconstruction depth perception. Subsequently, we elucidate a specific application through a straightforward depth detection experiment. It is imperative to underscore that this experiment does not encompass intricate reconstruction algorithms; rather, it serves to validate the feasibility of depth detection based on geometric relationships and projection transformations. It is evident that the proposed ASL system holds significant potential for delving into a broader array of applications in the realm of precision depth detection and perception by leveraging various well-established 3D reconstruction algorithms.

The experiment objective is to measure the depth difference between the test plane (object surface) and the reference plane, as the schematic diagram of the experimental setup shown in Fig. 9(a). d_1 and d_2 represent depth differences at a distinct position. By fixing the position of the reference plane and shifting the test plane in the depth direction (z-direction), infrared images of the far-field diffraction patterns generated by the 3×2 IA-VCSEL array at different detection distances are captured using a fixed-position infrared camera. The range of depth differences between the surface of two planes is from 5 mm to 33 mm, with a step size of 2 mm. A total of 15 positions of the pattern array are captured. During the shifting process, the first two columns of the expanded 3×2 array remain on the test plane, while the third column remains on the reference plane. Far-field infrared images measured at depth differences of 5 mm and 25 mm are shown in Figs. 9(b) and 9(c), respectively. It can be seen that as the depth difference increases, the distances between the two parts of diffraction patterns also increase.

To quantitatively measure the numerical relationship of triangular projection transformation between the depth difference and the variation of the speckle pattern images, we extract feature points of the second column of the 2×2 diffraction speckle pattern on the test plane and the 1×2 diffraction speckle pattern on the reference plane. These four feature points are used to calculate the depth information. By connecting the center of four feature points, we obtain two feature patterns: one based on the average length ($\bar{L}$) of L_1 and L_2, and the other based on the enclosed quadrilateral area (S), as shown in Figs. 9(d) and 9(e), respectively. We measure the actual value
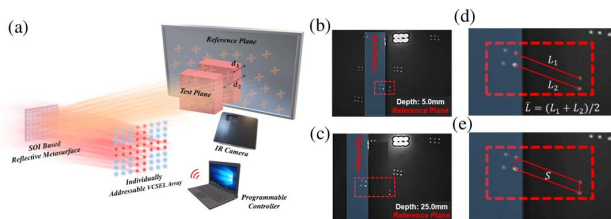
Fig. 9. (a) Experimental setup schematic for object depth detection; far-field infrared images measured at depth differences of (b) 5 mm and (c) 25 mm; exacted (d) average length $\bar{L}$ of feature points and (e) area S of feature patterns.
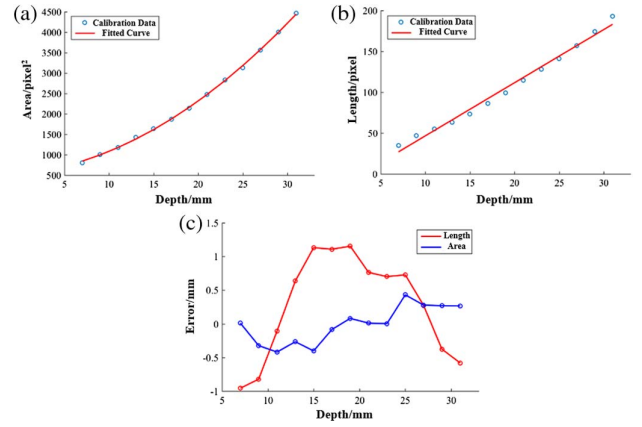
Fig. 10. Measured and fitted curves of (a) $\bar{L}$ and (b) S as a function of depth difference; (c) error curves of $\bar{L}$ and S as a function of depth difference.

curves of $\bar{L}$ and S with respect to depth difference, as shown in blue dots in Figs. 10(a) and 10(b), respectively. We exclude the first and last data sets to reduce the influence of experimental errors introduced by initial and final manual translation operation with minor distance. By using the least squares polynomial fitting method, measured values of $\bar{L}$ and S are fitted to the depth difference, resulting in the fitted curves shown in red lines in Figs. 10(a) and 10(b), respectively. The linear polynomial fitted curve of $\bar{L}$ (y_1) with respect to depth difference (x) is calculated as

$$y_1 = 6.5015x - 18.3873, \tag{14}$$

while the quadratic polynomial fitted curve of S (y_2) with respect to depth difference (x) is calculated as

$$y_2 = 3.3415x^2 + 23.0323x + 518.818. \tag{15}$$

The measured data are used to validate the deviation of fitting curves and error curves of $\bar{L}$ and S as a function of depth difference, as calculated in Fig. 10(c). The measurement mean error (ME) is calculated by $\varepsilon_1 = \sum(x_a - x_f)/n$, where x_a and x_f represent the actual and fitted values of the depth difference at distinct positions, respectively, while the mean absolute error (MAE) is calculated by $\varepsilon_2 = \sum|x_a - x_f|/n$. The error statistics of ε_1 and ε_2 are shown in Table 1. It can be observed that the average error obtained using the pattern area S is smaller than that using the average length $\bar{L}$. This can be attributed to the fact that as the depth distance changes, the variation of the pattern area is more significant due to the quadratic polynomial relationship, resulting in higher accuracy of the fitted curve. The calculated average measurement error ε_1 using average

Table 1.   Error Statistics in the Depth Detection Experiment

Calculation Method	MAE/mm	ME/mm
Average length ($\bar{L}$)	0.718	0.281
Pattern area (S)	0.219	0.011

length $\bar{L}$ and pattern area S is as low as 0.281 mm and 0.011 mm, respectively. The experimental results demonstrate the depth precision of up to sub-hundred microns. To further improve the measurement accuracy and reduce errors, measures such as improving the precision of the translation experiment, reducing the translation step size, and conducting more measurements for averaging can be taken.

4. CONCLUSION

We demonstrated a novel ASL system for 3D depth sensing by integrating the SOI MDOE and IA-VCSEL chip array. By coding the VCSEL on–off states in the VCSEL array, ASL patterns with dynamically variable positions and scales of speckles spacing are achieved. It is noteworthy that the "vernier caliper" effect promisingly enables fine control over speckle positions and further precision enhancement of depth detection. The method is experimentally validated and achieves depth detection precision of up to sub-hundred microns, which is much better than the conventional method. The focus of this study primarily lies in introducing a programmable structured light system with speckle pattern encoding capabilities and providing preliminary conclusions through fundamental depth detection experiments. Since the main emphasis of this paper is on presenting the implementation principles of this novel programmable structured light system, the intricacies of complex algorithms may not fall within the scope of this study and have not been delved into extensively. Looking ahead, we anticipate that with the integration of more complex and effective imaging algorithms and machine learning techniques, the fusion of metasurface optics and computer vision technology will lead to a growing number of interesting and practically valuable applications based on this system. The advancement provides fresh perspectives and technical support for the progress of diverse domains, including 3D computation, facial recognition, wearable devices, and machine vision, as well as consumer electronic products based on 3D imaging technologies.

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. X. Huang, C. Wu, X. Xu, *et al.*, "Polarization structured light 3D depth image sensor for scenes with reflective surfaces," Nat. Commun. **14**, 6855 (2023).
2. Z. Ren, F. Fang, N. Yan, *et al.*, "State of the art in defect detection based on machine vision," Int. J. Precis. Eng. Manuf. Green Technol. **9**, 661–691 (2022).
3. H. Lin, L. Nie, and Z. Song, "A single-shot structured light means by encoding both color and geometrical features," Pattern Recogn. **54**, 178–189 (2016).
4. Z. Song, S. Tang, F. Gu, *et al.*, "DOE-based structured-light method for accurate 3D sensing," Opt. Laser Eng. **120**, 21–30 (2019).
5. Z. Lv, Y. Xu, G. Li, *et al.*, "A new finger touch detection algorithm and prototype system architecture for pervasive bare-hand human computer interaction," in *IEEE International Symposium on Circuits and Systems (ISCAS)* (2013), pp. 725–728.
6. F. A. Zuberi, S. Khatri, and K. N. Junejo, "Dynamic gesture recognition using machine learning techniques and factors affecting its accuracy," in *6th International Conference on Innovative Computing Technology* (INTECH, 2016), pp. 310–313.
7. S. Palagi, A. Mark, S. Reigh, *et al.*, "Structured light enables biomimetic swimming and versatile locomotion of photoresponsive soft microrobots," Nat. Mater. **15**, 647–653 (2016).
8. Y. Ni, S. Chen, Y. Wang, *et al.*, "Metasurface for structured light projection over 120° field of view," Nano Lett. **20**, 6719–6724 (2020).
9. Q. Wang, P. Ni, Y. Y. Xie, *et al.*, "On chip generation of structured light based on metasurface optoelectronic integration," Laser Photon. Rev. **15**, 2000385 (2021).
10. G. Kim, Y. Kim, J. Yun, *et al.*, "Metasurface-driven full-space structured light for three-dimensional imaging," Nat. Commun. **13**, 5920 (2022).
11. X. Jing, R. Zhao, X. Li, *et al.*, "Single-shot 3D imaging with point cloud projection based on metadevice," Nat. Commun. **13**, 7842 (2022).
12. P. Qiao, W. Yang, and C. Chang-Hasnain, "Recent advances in high-contrast metastructures, metasurfaces, and photonic crystals," Adv. Opt. Photon. **10**, 180–245 (2018).
13. N. Yu, P. Genevet, M. A. Kats, *et al.*, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," Science **334**, 333–337 (2011).
14. A. Arbabi, Y. Horie, M. Bagheri, *et al.*, "Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission," Nat. Nanotechnol. **10**, 937–943 (2015).
15. R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," Optik **35**, 237–246 (1972).
16. M. V. Berry, "The adiabatic phase and Pancharatnam's phase for polarized light," J. Mod. Opt. **34**, 1401–1407 (1987).
17. I. Daubechies, M. Defriese, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," Commun. Pure Appl. Math. **57**, 1413–1457 (2004).
18. A. Maleki, L. Anitori, Z. Yang, *et al.*, "Asymptotic analysis of complex LASSO via complex approximate message passing (CAMP)," IEEE Trans. Inf. Theory **59**, 4290–4380 (2013).