Nature Communications

Article in Press

Large-array sub-millimeter precision coherent flash three-dimensional imaging

Bin Wang, Junze Tian, Jianwei Wang, Shukang Xu, Shuangxiang Zhao, Jianhao Duan, Weifeng Zhang, Yangyang Liu, Tao Zeng & Erke Mao

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Large-array sub-millimeter precision coherent flash three-dimensional imaging

Bin Wang[1,2,3†], Junze Tian[1,2,3†], Jianwei Wang[4,5], Shukang Xu[1,2,3], Shuangxiang Zhao[1,2,3], Jianhao Duan[1,2,3], Weifeng Zhang[1,2,3*], Yangyang Liu[4,5*], Tao Zeng[1,2,3*], Erke Mao[1,2]

[1]Radar Technology Research Institute, School of Information and Electronics, Beijing Institute of Technology，Beijing 100081, China
[2]Chongqing Innovation Center, Beijing Institute of Technology, Chongqing 401120, China
[3]Chongqing Key Laboratory of Novel Civilian Radar, Chongqing 401120, China
[4]Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China
[5]University of Chinese Academy of Sciences, Beijing 100049, China

†These authors contributed equally: Bin Wang, Junze Tian.

*Corresponding author. Email: weifeng.zhang@bit.edu.cn, liuyy@aircas.ac.cn, zengtao@bit.edu.cn

Abstract

High-precision three-dimensional (3D) imaging is essential for accurately perceiving environments, providing critical depth and spatial awareness. Among various approaches, solid-state LiDAR systems have garnered significant attention. However, depth precision, detection range and pixel scalability remain key challenges for their widespread adoption. Here, we report a large-array coherent flash 3D imaging system that achieves a sub-millimeter range precision through stepped-frequency modulation and coherent detection with CCD sensors. A coherent image sensor is developed, and a prototype system is demonstrated, providing 3D imaging with a depth precision as high as 0.47 mm over a range of 30.50 m at an optical power of 15.86 mW and a maximum frame rate of 10 Hz. Our system features high range precision, exceptional sensitivity across long distances, and robust pixel scalability by directly leveraging well-established CCD sensors. This advancement introduces a scalable approach to long-range high-precision 3D imaging, with substantial implications for deformation monitoring, virtual reality, and cultural heritage preservation.

High-precision three-dimensional (3D) imaging is a fundamental technology for capturing, processing, and visualizing objects and environments in three dimensions[1], which can find numerous applications spanning deformation monitoring[2–4], building modelling[5–7], cultural heritage restoration[8, 9], virtual reality[10], to medical imaging[11, 12]. Light detection and ranging (LiDAR), a typical optical radar implementation, is envisioned as a promising approach to achieving high-precision 3D imaging[13-16].

Most commercially available LiDAR systems rely on mechanical scanning to direct the laser beam for data acquisition[17-19]. This dependence on bulky mechanical components often leads to a slow scanning speed and reduced reliability. In contrast, solid-state LiDAR systems offer a powerful alternative by leveraging solid-state components for both beam steering and detection, bringing benefits such as compact size, increased reliability, and improved scalability[20, 21]. The solid-state LiDAR systems are primarily implemented using optical phase arrays (OPAs) or focal plane arrays (FPAs). The OPAs employ an array of optical antennas to individually control the phase of light waves, enabling precise beam direction without mechanical movement[22]. However, the scaling has two limitations: exponentially increasing number of antennas elements creates active electrical control challenges[23], and for scanning over a field of view with a high spatial resolution, a long acquisition time is resulted[24].

Unlike the OPAs, the FPAs use an array of photosensitive elements or pixels to measure light intensity and captures the entire scene instantaneously. Currently, two main types of the FPAs are commonly employed: single-photon avalanche diodes (SPADs) and charge-coupled devices (CCDs). The SPADs are highly sensitive sensors capable of detecting individual photons, triggering an avalanche current for signal amplification[25]. However, scaling SPAD arrays to higher pixel counts presents considerable technical challenges, and the depth precision of the SPAD-based system is constrained by the time-to-digital converter[26-29]. Conversely, CCD technology, refined over five decades, supports integration of pixel arrays at scales reaching hundreds of megapixels, revolutionizing imaging and digital photography applications[30]. Recent advancements have also seen the development of flash LiDAR systems based on CCDs[31-35]. Using short-pulse illuminations, these systems incorporate additional active gated gain modulation or polarization modulation ahead of the CCD sensor, which are noncoherent and require a timing signal to drive the gate or modulator. This reliance on timing signals usually restricts their precision to several centimeters or more.

To fully exploit the potential of high-sensitivity CCD sensors, inspired by extensive use of coherent detection in long-haul fiber optical communications, a high-precision LiDAR system based on stepped-frequency modulation and coherent image sensor is proposed, as illustrated in Fig. 1a. At the transmitter, a stepped-frequency microwave signal, whose frequency increases in discrete steps over time, is fed into a Mach-Zehnder modulator (MZM). After the modulation, a stepped-frequency modulated optical signal is generated, which is subsequently split into two parts: one part flood-illuminates the scene, while the other serves as a local reference. At the receiver, a coherent image sensor is constructed using four CCD sensors and a free-space optical hybrid (see "Methods" section for details). Within each pixel of the CCD sensor, the local reference signal combines with the backscattered echo signal that incorporates an additional time delay due to the round-trip propagation. Through optical-to-electrical conversion, each pixel outputs a voltage signal proportional to accumulated electrical charges over a fixed exposure time. As the microwave signal sweeps through discrete frequencies, each pixel records a sequence of voltages across multiple frames, with the frame counts equal to the number of

microwave frequency steps. By employing a tailored coherent processing algorithm, depth information is extracted, enabling the reconstruction of a 3D image from the pixel array.

Figure 1b illustrates the principle of the coherent processing. A microwave signal with an instantaneous frequency f_m, increasing in discrete steps Δf over time, is generated. After the electro-optic modulation, the modulated optical signal, whose intensity temporal profile mirrors the varying microwave frequency, is split into two parts: one serves as the local signal A_l, while the other part is directed toward the target. The backscattered echo signal A_e returns with an additional time delay τ, corresponding to the target distance d. When the local and echo signals are mixed within a pixel of the CCD array, a multiplication component $A_l A_e$ is generated, directly associated with the target distance. Over a fixed exposure time T_e, the pixel accumulates charges, producing a voltage signal that contains an integrated term U_{sig} derived from the multiplication component. Through coherent processing across the multiple frames captured by the four CCDs, the term U_{sig} is extracted, revealing a temporal period related to the target distance. By identifying the fundamental frequency of this term, the depth information is recovered.

Mathematically, the term U_{sig} can be expressed in the frequency domain as (see Supplementary Note 1):

$$\left| \mathcal{U}_{sig}\left(n\right) \right| = \frac{4\pi\eta}{h\omega} S_V G T_e A_l A_e J_1^2\left(\beta\right) \left| \frac{\sin\left[\pi\left(n - \frac{2N\Delta f d}{c}\right) \right]}{\sin\left[\frac{\pi}{N}\left(n - \frac{2N\Delta f d}{c}\right) \right]} \right| \tag{1}$$

where η is the quantum efficiency, h is the Planck constant, ω is the angular frequency of the optical signal, S_V is the sensitivity of the CCD sense node, G is the gain of the CCD pixels and T_e is the exposure time per frame. A_l and A_e are the amplitudes of the local and echo optical signals, respectively. Δf is the frequency step size of the microwave signal, and N is the number of step points. J_1 denotes the first-order Bessel function of the first kind. β is the modulation index. d represents the target distance, and c is the speed of light in space.

Using Eqn. (1), the distance d can be determined by extracting the fundamental frequency:

$$d = \frac{c}{2N\Delta f} n_{funda} \tag{2}$$

Thanks to the wideband stepped-frequency modulation, the multiple frames are coherently synthesized, and the depth information can be extracted with a high precision. Moreover, the unambiguous range depends on the frequency step size:

$$d_{NAR} = \frac{c}{4\Delta f} \tag{3}$$

The smaller the frequency step size, the larger the unambiguous range.

Results

Coherent processing for 3D imaging

Figure 2 illustrates the 3D image reconstruction process based on coherent detection. A staircase target model is chosen and positioned 30.5 m away (see "Methods" section for details about the experimental set-up). Figure 2a gives a photograph of the staircase, composed of seven steps with heights increasing linearly from 5 mm to 30 mm in 5 mm increments. When the local and echo optical signal are mixed within the pixels, interference patterns are recorded simultaneously by each CCD. Figure 2b shows one frame raw image captured by the four CCDs, where no discernible target features are evident because of the weak echo power. A random pixel marked by the yellow dot is selected for further analysis to illustrate the coherent processing procedure. With the microwave frequency stepping, an output voltage sequence of the selected pixel is recorded. Figure 2c displays the recorded voltage sequences for the four CCDs, with each exhibiting a distinct time-varying profile encoding depth information. In the coherent image sensor, the free-space optical hybrid imposes a fixed phase relationship between the output voltage signals. Figure 2d provides a zoomed-in view of each output voltage signal around 15.6 s. It is clear to see that the voltage signal generated by CCD1 is out of phase of that generated by CCD4, and so is the voltage signals generated by CCD2 and CCD3, which verifies the effectiveness of the coherent image sensor. After the coherent processing, the integrated term derived from the multiplication component is extracted. Fig. 2e depicts the extracted signal, whose frequency components are resolved using a discrete Fourier transform (DFT). Figure 2f reveals two prominent peaks: a fundamental frequency at 56.20 Hz and a secondary peak at 37.40 Hz, attributed to high-order sidebands. Using Eqn. (2), the depth is resolved from the fundamental frequency. Figure 2g gives the recovered depth of 30469.10 mm at the selected pixel. Subsequently, a row of pixels along the red dashed line in Fig. 2a is selected and the same coherent processing is repeated for each pixel. Figure 2h shows the extracted fundamental frequencies across the row pixels. Different frequencies correspond to different depths. Figure 2i presents the recovered depth profile that resolves the staircase steps and background with high fidelity, consistent with the physical model. Benefiting from the large array of the CCD sensor, a complete 3D reconstruction is achieved. Figure 2j presents the final image, where the horizontal and vertical axis represent spatial coordinates, and the color encodes the depth. Notably, depth variations as small as 5 mm remain discernible. This reconstruction, enabled by stepped-frequency modulation and coherent detection, directly retrieves depth information from 2D data over multiple frames.

3D imaging of a traffic scene model

To evaluate the imaging performance of the prototype system, high-precision 3D imaging of a traffic scene model is conducted. Each CCD is set to have an exposure time of 4 ms at a frame rate of 150 Hz. By synchronizing the microwave source with the coherent image sensor, 3000 frames are collected over 20 s, in which the stepped-frequency microwave signal is set to have corresponding 3000 step points. Figure 3a depicts the traffic model, which features a pedestrian, vehicle, light pole, and guide board. Positioning the model at 30.5 m away and using stepped microwave frequencies ranging from 0.1 to 6.0 GHz, the CCDs captured multiple frames at a resolution of 320×256 pixels. The local optical power per pixel is calculated at 0.64 pW, while the echo power is at 0.26 fW, three orders of magnitude weaker than the local signal. Following coherent processing, Fig. 3b reveals a detailed 3D construction of the scene, with clear visualization of individual objects and their shadows. Figure 3c provides a side view of the 3D image, highlighting depth details and slope angles. A histogram in Fig. 3d quantifies the number of pixels contributing to the imaging of targets at different depths. For spatial resolution, Fig. 3e presents a front view of the 3D image, demonstrating a resolution of 1.2 mm both horizontally

and vertically. The scalability of CCD arrays offers further potential for enhancing resolution, underscoring the suitability of mature CCD technology for high-resolution 3D imaging. To provide a complementary perspective to the 3D image, a dedicated 2D reconstruction algorithm (see Supplementary Note 2) is also developed and the reconstructed 2D image is shown in Fig. 3f, where the object boundaries are clearly visible. The use of stepped-frequency modulation and coherent detection proves critical for detecting extremely weak echo signals and recovering depth information, enabling precise 3D imaging over long distances. Supplementary Movie S1 illustrates a comparative visualization of normal 2D imaging versus 3D reconstruction of a gladiolus bloom.

Range precision analysis and frame rate evaluation

The range precision is determined primarily by the step size and bandwidth of the microwave signal, while spatial resolution depends on pixel density and the field of view. These factors collectively enable precise, scalable, and robust imaging capabilities. To thoroughly evaluate the range precision, a planar target is leveraged. Figure 4a presents the 3D imaging result with the target placed at 30.5 m. The depth distribution histogram, shown in Fig. 4b, is analyzed by fitting a Gaussian curve based on the least-squares method. The standard deviation of the fitted curve yields a range precision as high as 0.47 mm. This high level of precision is attributed to three key factors: the broadband stepped-frequency modulation, enabling high range resolution and precise depth measurements; the use of coherent image sensor, which effectively suppresses noise and enhances signal fidelity; the extensive frame sampling, critical for accurate extraction of the fundamental frequency.

In addition to the experimental measurement of range precision, its theoretical evaluation is also conducted. By applying Cramér–Rao lower bound (CRLB) analysis[36], the theoretical range precision is derived (see Supplementary Note 3):

$$\delta d \geq \frac{c}{2\pi B} \sqrt{\frac{3}{2N} \cdot \frac{1}{SNR}} \tag{4}$$

where B is the bandwidth of the stepped-frequency microwave signal and SNR is the signal to noise ratio of the pixel output voltage signal. Generally, a broad bandwidth and large number of the step points lead to a high range precision.

Figure 4c compares the theoretical (blue) and measured (red) range precision as a function of target distance, with the probe optical power fixed at 15.86 mW (see Supplementary Note 4). As the distance increases, the diminishing echo optical power results in a reduced range precision. The measured precision is notably worse than the theoretical precision, due to random speckle noise. At a distance of 30.5 m, the theoretical precision is calculated to be 0.079 mm, primarily constrained by quantum noise and dark current noise of the CCD sensors.

Further theoretical evaluation of range precision is performed. Assuming a fixed probe optical power, the theoretical range precision is calculated using Eqn. (4) under varying bandwidths and target distances. Figure 4d presents the results. For a given target distance, increasing the bandwidth improves the range precision. Conversely, for a fixed bandwidth, the range precision diminishes with greater distance due to the reduction in SNR over longer ranges.

In addition to range precision, frame rate is also a critical parameter, determining the number of the depth maps captured and processed per second. To quantitatively assess this performance metric in detail, 3D imaging of a BIT alphabet model are performed using varying frame counts

of 35, 88, 175, 350, 700, 1750 and 3500 at a fixed CCD frame rate of 350 Hz, corresponding to an acquisition time of 0.10s, 0.25s, 0.50s, 1.0s, 2.0s 5.0s and 10.0s. The reconstructed 3D images are shown in Figure 4e (I-V), where a maximum frame rate of 10Hz is achieved. In addition, increasing the acquisition time yields smoother appearances of the 3D images. Figure 4f summarizes the range precisions with data acquisition time. Specifically, with an acquisition time of 0.1s, a range precision of 1.33 mm is achieved. As the time spend increases, the range precision is highly improved to be 0.35 mm. This trade-off between acquisition time and measurement range precision provides operational flexibility for different application requirements.

Phase stability of the modulated optical signal is crucial for maintaining precision, particularly under environmental disturbances. To test robustness against air turbulence, a comparative experiment is designed and performed. A 2.18 m s^{-1} airflow generated by an electric fan is introduced across the free-space optical path during imaging. Remarkably, a clear 3D image is generated in Fig. 4e(VIII), and a range precision is measured to be 0.37 mm, just 0.02 mm degradation from static condition with the same acquisition time of 10s, shown in Fig. 4f, which confirms excellent resistance to airflow-induced phase perturbations. This resilience relies on that the fact that the optical signal is intensity-modulated by a microwave signal, and its intensity temporal profile is proportional to the microwave signal frequency at GHz, which makes the phase front inherently stable against air fluctuations.

Application: interactive 3D reconstruction

3D imaging plays a vital role in augmented reality and virtual reality, enabling the creation of realistic and immersive environments. Comprehensive panoramic data capture through 360° scanning ensures full angular coverage, leading to precise reconstructions and highly detailed 3D models. Figure 5a depicts a bust sculpture used as a test object for creating realistic digital representations. Positioning the bust sculpture at 30.3 m away, 3D imaging is performed after each rotation by 45°, and eight 3D images are collected (see Supplementary Note 5). Figure 5b shows one 3D image of the bust sculpture at the initial orientation, which exhibits the proper scale and perspective of the bust sculpture. By integrating data from all eight 3D images, a complete virtual model of the bust is created. Figure 5c presents the reconstructed bust from multiple perspectives, showcasing accurate depth information. Supplementary Movie S2 provides a 360-degree visualization of the 3D reconstructed bust sculpture. This virtual model enables enhanced interaction, allowing users to manipulate and explore the object in an immersive environment.

Discussion

Table 1 summarizes recent advancements in solid-state LiDAR 3D imaging techniques. As shown in the table, compared with other 3D imaging systems, our approach attains a range precision as high as 0.47 mm, representing an improvement of at least one order of magnitude. In addition, a frame rate of 10 Hz and a scan rate of up to 819 kpts s^{-1} are achieved. This reflects a competitive frame rate and a high scan rate among existing solid-state LiDAR systems, supported by the large pixel count of the CCD sensors. Our proposed method leverages stepped-frequency modulation and coherent detection with CCD arrays, achieving both high precision and extended range. Through a broadband stepped-frequency modulation, multi-frame coherent synthesis is enabled to extract high-precision depth information. The coherent image sensor simultaneously suppresses common-mode noise and amplifies coherent echo signal, providing exceptional sensitivity over extended distances. Furthermore, direct utilization of mature CCD

technology facilitates high spatial resolution with strong scalability to arrays with hundreds of megapixels, making this approach highly adaptable for future large-scale applications. Together, these factors deliver submillimeter range precision, extended detection range, and robust pixel scalability.

In the proposed approach, the unambiguous range is dependent on the step frequency size. Achieving a large unambiguous range requires a small step frequency size, while high range precision necessitates a large number of step points. Consequently, limited by the frame rate of CCD sensors, an extended measurement time is resulted, which poses a significant limitation for applications involving fast-changing scenes or velocity measurements. In our experimental demonstration, a maximum 3D frame rate is 10 Hz. To further accelerate the imaging speed, directly leveraging the high frame-rate mode of CCD sensors at a loss of spatial resolution is an effective solution. For example, the CCD sensors used in the demonstrations support the maximum frame rate as high as 10000 Hz with a spatial pixel of 32×4. In addition, sparse sampling offers a promising solution to decrease the number of step points required[37]. Moreover, replacing the CCD with a high-frame-rate CMOS sensor can also further reduce the data acquisition time[38].

In addition to CCD acquisition time, the real-time performance of our imaging system is influenced by data transfer, processing, and display. A control unit synchronizes the stepped-frequency microwave source with the four CCDs, which capture 2D frames comprising exposure and readout. After each frame is recorded, the CCDs immediately begin capturing the subsequent frame while the previously acquired data are transferred to the control unit memory. Upon completion of the final frame, a final transfer is executed. The data are then processed in parallel across all pixels, enabling reconstruction of the full 3D image, which is subsequently rendered and displayed. Among these four components, acquisition accounts for 98.78% of the total time (see Supplementary Note 6), demonstrating that the CCD sensor frame rate is the primary factor constraining overall system performance. Accordingly, the adoption of faster CCD sensors would directly improve real-time performance.

To assess the 3D imaging system's ability to capture transversely moving targets over long distances, additional experiments are conducted (see Supplementary Note 7). The results demonstrate that the proposed system can effectively reconstruct a moving target in 3D at speeds up to 300 mm s^{-1} and a distance of 30.60 m. This performance is achieved by reducing the exposure time, which helps minimize motion blur and enable clean depth reconstruction. Thus, increasing the CCD frame rate while shortening the exposure time proves to be an effective strategy for high-speed 3D imaging of moving targets.

Although our experimental setup employs an infrared laser source and four high-end short-wave infrared (SWIR) CCD sensors, the underlying principle is equally applicable to visible wavelengths. This adaptation could provide a cost-effective alternative and, more importantly, enable direct fusion of visual and LiDAR data. Unlike LiDAR alone, images offer rich texture and color information. Using the same sensors for both modalities allows seamless alignment of LiDAR and image data, leveraging their combined strengths to improve object detection, recognition, localization, and 3D reconstruction. Furthermore, the coherent detection architecture, which integrates noise cancellation and signal enhancement (see Supplementary Note 8), ensures stable operation even under strong background illumination or intentional jamming. For convenience, all experimental conditions and performance metrics are summarized in Supplementary Note 9.

Leveraging the high pixel density of well-established CCD sensors, our 3D imaging camera achieves both high resolution and sub-millimeter precision. This enables critical applications including: (1) high-precision deformation monitoring; (2) digital preservation of cultural artifacts, monuments, and archaeological sites; (3) damage assessment and restoration of heritage structures; and (4) high-fidelity virtual reality content creation. Compared to conventional point-by-point or line-by-line laser scanning systems, our approach offers two key advantages: a fully solid-state architecture without moving parts and instantaneous full-scene capture with a superior spatial resolution.

In conclusion, a large-array coherent flash LiDAR system that achieves a sub-millimeter range precision was reported. A coherent image sensor was developed, and a prototype system was demonstrated, providing 320×256 3D imaging with a depth precision as high as 0.47 mm over a range of 30.50 m at an optical power of 15.86 mW and a maximum frame rate of 10 Hz. The proposed system features a high range precision, a high sensitivity over extended distances, and a strong pixel scalability with mature CCDs. This advancement enhances the accessibility and versatility of accurate 3D imaging, delivering exceptional depth sensing capability.

Methods

Coherent image sensor

The coherent image sensor is a crucial component in the proposed 3D imaging system. It is constructed using four CCDs and a free-space optical hybrid. Figure S1 illustrates the setup of the coherent image sensor. The optical hybrid is comprised of two polarizers (P), a quarter-wave plate (QWP), a non-polarizing beam splitter (NPBS), and two polarizing beam splitters (PBS).

By adjusting polarizer P1 and the QWP, the polarization state of the incident local signal is set to left-handed circular polarization. After passing through the NPBS, the local signal is divided into reflected and transmitted beams, both preserving the original polarization state. Relative to the reflected beam, a phase shift of $\pi/2$ is introduced to the transmitted beam. The transmitted beam is then directed to the polarizing beam splitter PBS1, where it is further split into two linearly polarized beams. Due to the $\pi/2$ phase difference between the orthogonal electric field components of the circularly polarized light, these two beams at the output of PBS1 have the same phase of π. The reflected beam is sent to the PBS2, where the two linearly polarized beams have the same phase of $\pi/2$ at its output.

The backscattered echo signal is collected using a convex lens and directed to the polarizer P2 for linear polarization. After passing through the NPBS, the signal is divided into the reflected and transmitted beams, which retain the original polarization state. A phase shift of $\pi/2$ is introduced to the transmitted beam relative to the reflected beam. This transmitted beam is then directed to the polarizing beam splitter PBS2, where it is split into two linearly polarized beams. At the output of the PBS2, these beams have a phase difference of $\pi/2$. The reflected beam is subsequently sent to the PBS1, where the output also consists of two linearly polarized beams with a phase difference of $\pi/2$.

Thanks to the free-space optical hybrid, the local and echo signals meet the phase requirements for coherent detection before injecting into the four identical CCDs. The calibration of the coherent image sensor is critical for high-quality 3D image capture. Zhang's camera calibration method is employed to achieve precise alignment, determining the rotation matrices and

translation vectors for each CCD[39]. Figure S2 illustrates the alignment performance. Figure S2a shows a chessboard image captured by each CCD. Binarization and calibration are then performed using the determined rotation matrices and translation vectors, as shown in Figure S2b. An XOR operation is applied to the pairs of the binarized images to detect misaligned pixels, with the results presented in Fig. S2c. The white pixels in the XOR results indicate the misalignments. The misalignment ratios are calculated to be 1.99%, 1.58%, and 1.54%. These small ratios indicate a good alignment.

Experimental set-up

In the experiment, a target to be imaged is placed at a distance as far as 30.5 m. A narrow-linewidth fiber laser (NKT Photonics Koheras BASIK E15) is used to generate a CW optical signal at a wavelength of 1550 nm. The optical signal is modulated at a Mach-Zehnder modulator (MZM, Fujitsu FTM7937EZ) driven by a stepped-frequency microwave signal from a microwave source (Keysight N5182B). The stepped-frequency microwave signal has a frequency ranging from 0.1 to 6 GHz with a frequency step size as small as 1.97 MHz. Each frequency has a temporal duration of 6.7 ms. After power amplification using an erbium-doped fiber amplifier (EDFA, Amonics AEDFA-PA-35-B-FA), the modulated optical signal is split into two paths: the probe signal and reference signal. The probe signal has a power of 15.86 mW and the reference signal has a power of 14.12 µW.

At the receiver, a lens with a focal length of 500 mm (LBTEK MCX10621-C) is used to collect the echo signal, corresponding to a field of view (FOV) of $0.72° \times 0.58°$. The coherent image sensor is leveraged to capture the image, which consists of a free-space optical hybrid and four short-wave infrared CCD cameras with a pixel of 320×256 (Xenics Bobcat-320-GigE-400Hz).

Data availability: The authors declare that the data supporting the findings of this study are available within the paper and its Supplementary Information. The experimental data generated in this study are deposited in Figshare at https://doi.org/10.6084/m9.figshare.30948620.

References

1. Schwarz, B. Mapping the world in 3D, *Nat. Photonics* **4**, 429–430 (2010).
2. Wulder, M. A. *et al.* Lidar sampling for large-area forest characterization: a review. *Remote Sens. Environ.* **121**, 196–209 (2012).
3. Jaboyedoff, M. *et al.* Use of LIDAR in landslide investigations: a review. *Nat. Hazards* **61**, 5–28 (2012).
4. Torres, Y. *et al.* Integration of LiDAR and multispectral images for rapid exposure and earthquake vulnerability estimation. Application in Lorca, Spain. *Int. J. Appl. Earth Obs. Geoinf.* **81**, 161–175 (2019).
5. Trocha, P. *et al.* Ultrafast optical ranging using microresonator soliton frequency combs. *Science* **359**, 887–891 (2018).
6. Riemensberger, J. *et al.* Massively parallel coherent laser ranging using a soliton microcomb. *Nature* **581**, 164–170 (2020).
7. Kuse, N. & Fermann, M. E. Frequency-modulated comb LIDAR. *APL Photonics* **4**, 106105 (2019).

8. Godin, G. *et al.* Active optical 3D imaging for heritage applications. *IEEE Comput. Graph. Appl.* **22**, 24–35 (2002).

9. Luo, L. *et al.* Airborne and spaceborne remote sensing for archaeological and cultural heritage applications: a review of the century (1907–2017). *Remote Sens. Environ.* **232**, 111280 (2019).

10. Xiong, J., Hsiang, E. L., He, Z., Zhan, T. & Wu, S. T. Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light Sci. Appl.* **10**, 216 (2021).

11. Bouazizi, M., Ye, C. & Ohtsuki, T. 2-D LIDAR-based approach for activity identification and fall detection. *IEEE Internet Things J.* **9**, 10872–10890 (2022).

12. Fang, Z. *et al.* A review of emerging electromagnetic-acoustic sensing techniques for healthcare monitoring. *IEEE Trans. Biomed. Circuits Syst.* **16**, 1075–1094 (2022).

13. Jiang, Y., Karpf, S. & Jalali, B. Time-stretch LiDAR as a spectrally scanned time-of-flight ranging camera. *Nat. Photonics* **14**, 14–18 (2020).

14. Stellinga, D. B. *et al.* Time-of-flight 3D imaging through multimode optical fibers. *Science* **374**, 1395–1399 (2021).

15. Zhang, X., Kwon, K., Henriksson, J., Luo, J. & Wu, M. C. A large-scale microelectromechanical-systems-based silicon photonics LiDAR. *Nature* **603**, 253–258 (2022).

16. Chen, R. *et al.* Breaking the temporal and frequency congestion of LiDAR by parallel chaos. *Nat. Photonics* **17**, 306–314 (2023).

17. Rogers, C. *et al.* A universal 3D imaging sensor on a silicon photonics platform. *Nature* **590**, 256–261 (2021).

18. Li, Z. *et al.* Towards an ultrafast 3D imaging scanning LiDAR system: a review. *Photon. Res.* **12**, 1709–1729 (2024).

19. John, D. D. *et al.* Wideband electrically pumped 1050-nm MEMS-tunable VCSEL for ophthalmic imaging. *J. Light. Technol.* **33**, 3461–3468 (2015).

20. Park, J. *et al.* All-solid-state spatial light modulator with independent phase and amplitude control for three-dimensional LiDAR applications. *Nat. Nanotechnol.* **16**, 69–76 (2021).

21. Li, N. *et al.* A progress review on solid-state LiDAR and nanophotonics-based lidar sensors. *Laser Photonics Rev.* **16**, 2100511 (2022).

22. Sun, J., Timurdogan, E., Yaacobi, A., Hosseini, E. S. & Watts, M. R. Large-scale nanophotonic phased array. *Nature* **493**, 195–199 (2013).

23. Poulton, C. V. *et al.* Coherent LiDAR with an 8,192-element optical phased array and driving laser. *IEEE J. Sel. Top. Quantum Electron.* **28**, 1–8 (2022).

24. Chen, B. *et al.* SiN-on-SOI optical phased array LiDAR for ultra-wide field of view and 4D sensing. *Laser Photonics Rev.* **18**, 2301360 (2024).

25. Bruschini, C. *et al.* Single-photon avalanche diode imagers in biophotonics: review and outlook. *Light Sci. Appl.* **8**, 87 (2019).

26. Beer, M., Haase, J. F., Ruskowski, J. & Kokozinski, R. Background light rejection in SPAD-based LiDAR sensors by adaptive photon coincidence detection. *Sensors* **18**, 4338 (2018).

27. Zhang, C. *et al.* A 30-frames/s, 252 × 144 SPAD flash LiDAR with 1728 dual-clock 48.8-ps TDCs, and pixel-wise integrated histogramming. *IEEE J. Solid-State Circuits* **54**, 1137–1151 (2019).

28. Padmanabhan, P. *et al.* A 256×128 3D-stacked (45 nm) SPAD flash LiDAR with 7-level coincidence detection and progressive gating for 100 m range and 10 klux background light. In *2021 IEEE International Solid-State Circuits Conference (ISSCC)*, 111–113 (IEEE, 2021).

29. Hutchings, S. W. *et al.* A reconfigurable 3-D-stacked SPAD imager with in-pixel histogramming for flash LIDAR or high-speed time-of-flight imaging. *IEEE J. Solid-State Circuits* **54**, 2947–2956 (2019).

30. Smith, G. E. Nobel Lecture: The invention and early history of the CCD. *Rev. Mod. Phys.* **82**, 2307–2312 (2010).

31. Xiuda, Z., Huimin, Y. & Yanbing, J. Pulse-shape-free method for long-range three-dimensional active imaging with high linear accuracy. *Opt. Lett.* **33**, 1219–1221 (2008).

32. Jin, C., Sun, X., Zhao, Y., Zhang, Y. & Liu, L. Gain-modulated three-dimensional active imaging with depth-independent depth accuracy. *Opt. Lett.* **34**, 3550–3552 (2009).

33. Jo, S. *et al.* High resolution three-dimensional flash LIDAR system using a polarization modulating Pockels cell and a micro-polarizer CCD camera. *Opt. Express* **24**, A1580–A1592 (2016).

34. Zhang, P., Du, X., Zhao, J., Song, Y. & Chen, H. High resolution flash three-dimensional LIDAR systems based on polarization modulation. *Appl. Opt.* **56**, 3889–3895 (2017).

35. Song, Y. *et al.* Potassium tantalite niobite (KTN) crystal-based polarization-modulated 3D lidar with a large field of view. *Opt. Lett.* **45**, 5319–5322 (2020).

36. Gu, Z. *et al.* Theoretical range precision obtained by maximum likelihood estimation in laser radar compared with the Cramer-Rao bound. *Appl. Opt.* **57**, 9951–9956 (2018).

37. Blu, T., Dragotti, P. L., Vetterli, M., Marziliano, P. & Coulot, L. Sparse sampling of signal innovations. *IEEE Signal Process. Mag.* **25**, 31–40 (2008).

38. Fossum, E. R. CMOS image sensors: electronic camera-on-a-chip. *IEEE Trans. Electron Devices* **44**, 1689–1698 (1997).

39. Chen, X., Chen, Y., Song, X., Liang, W. & Wang, Y. Calibration of stereo cameras with a marked-crossed fringe pattern. *Opt. Lasers Eng.* **147**, 106733 (2021).

Author contributions: B.W., J.Z.T., W.F.Z. and Y.Y.L. conceived of and designed the project. J.Z.T., S.X.Z, J.H.D., and S.K.X. performed most experiments; J.Z.T., J.W.W., T.Z., and E.K.M. analyzed data and organized figures; and all authors provided intellectual input and contributed to the text.

Competing interests: The authors declare no competing interests.

Table 1. Performance comparison of recent 3D solid-state LiDAR imaging systems

Reference	Method	Sensor	Precision (mm)	Power (mW)	Range (m)	Pixels	Frame rate (Hz)	Scan rate (kpts s⁻¹)
2022[23]	FMCW	OPA	NA	400	35	154×20	10	33
2024[24]	FMCW	OPA	5.5	NA	3.5	400×18	1	10
2018[26]	ToF	SPAD	47	11.25	6.5	192×2	25	10
2019[27]	ToF	SPAD	50	2	50	126×144	30	544
2021[28]	ToF	SPAD	NA	5	10	128×128	NA	NA
2017[34]	PM	CCD	<4	3.5	17.3	512×512	NA	NA
2020[35]	PM	CCD	44	50	15	640×480	2	614
This work	SF-CD	CCD	0.47	15.86	30.5	320×256	10	819

FMCW: frequency-modulated continuous wave. ToF: time of flight. PD: photodetector. BPD: balanced photodetector. OPA: optical phased array. SPAD: single-photon avalanche diode. PM: polarization modulation. CCD: charge-coupled device. SF-CD: stepped-frequency + coherent detection.

Fig. 1. Flash coherent CCD 3D imaging system. a, Schematic of the 3D imaging system. MZM: Mach-Zehnder modulator. L: lens. P: polarizer. QWP: quarter-wave plate. NPBS: non-polarized beam splitter. PBS: polarized beam splitter. CCD: charge-coupled device. **b,** Principle of the imaging system. f_m: frequency of the microwave signal. A_l and A_e: amplitude of the local and echo optical signal. $A_l A_e$: multiplication component. U_{sig}: integrated term derived from the multiplication component. T_e: exposure time of each frame.
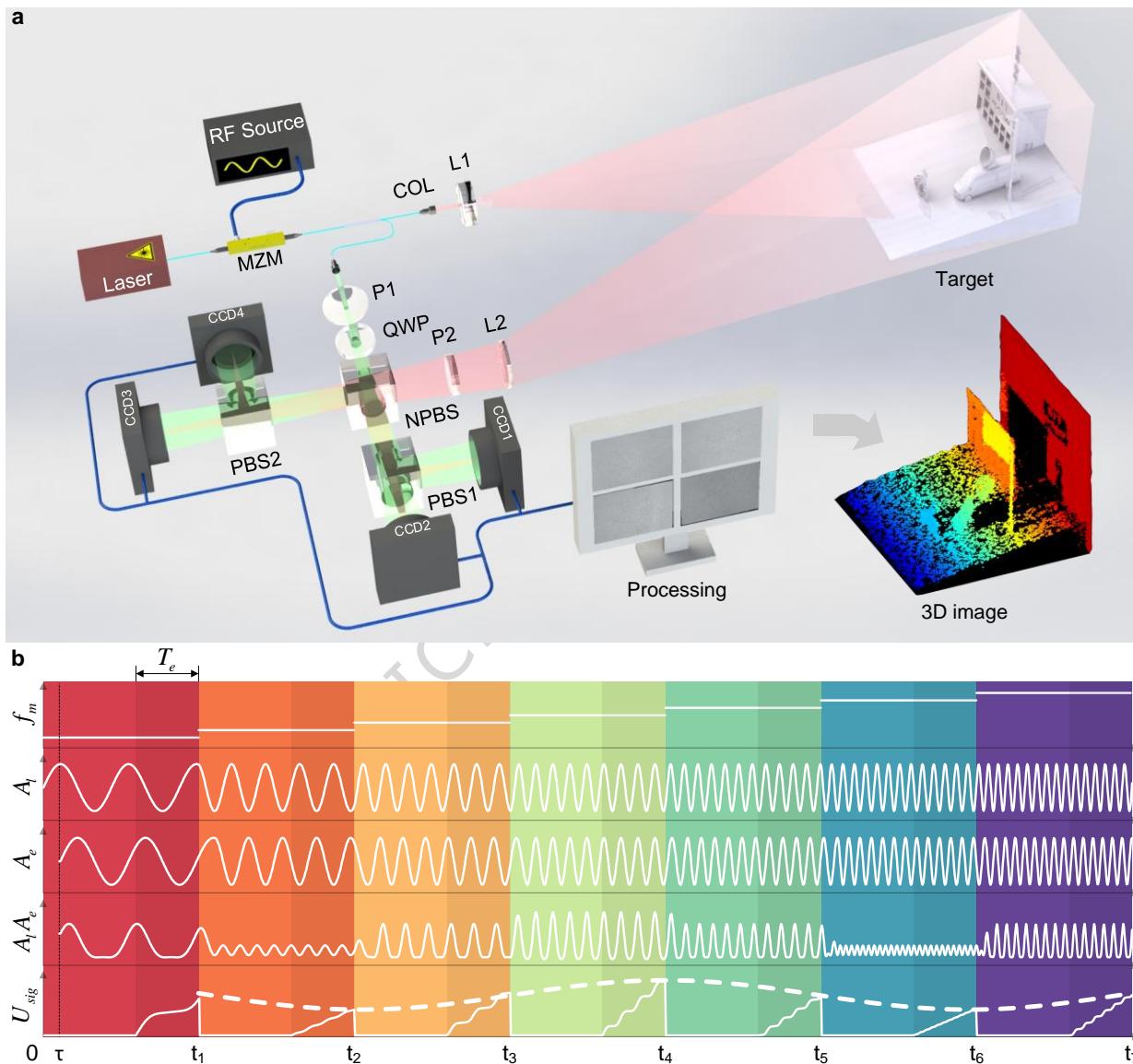
Fig. 2. Coherent processing for 3D imaging. a, Photograph of a staircase target model. **b**, Raw image captured by each CCD. **c**, Recorded output voltages of the selected pixel. **d**, Zoom-in view of the recorded output voltages. **e**, Temporal profile of the integrated term after coherent processing. **f**, Calculated frequency components. **g**, Recovered the depth information at the selected pixel. **h**, Calculated fundamental frequencies across a row of the pixels along the dashed red line shown in **a**. **i**, Recovered depth information. **j**, Reconstructed 3D image.
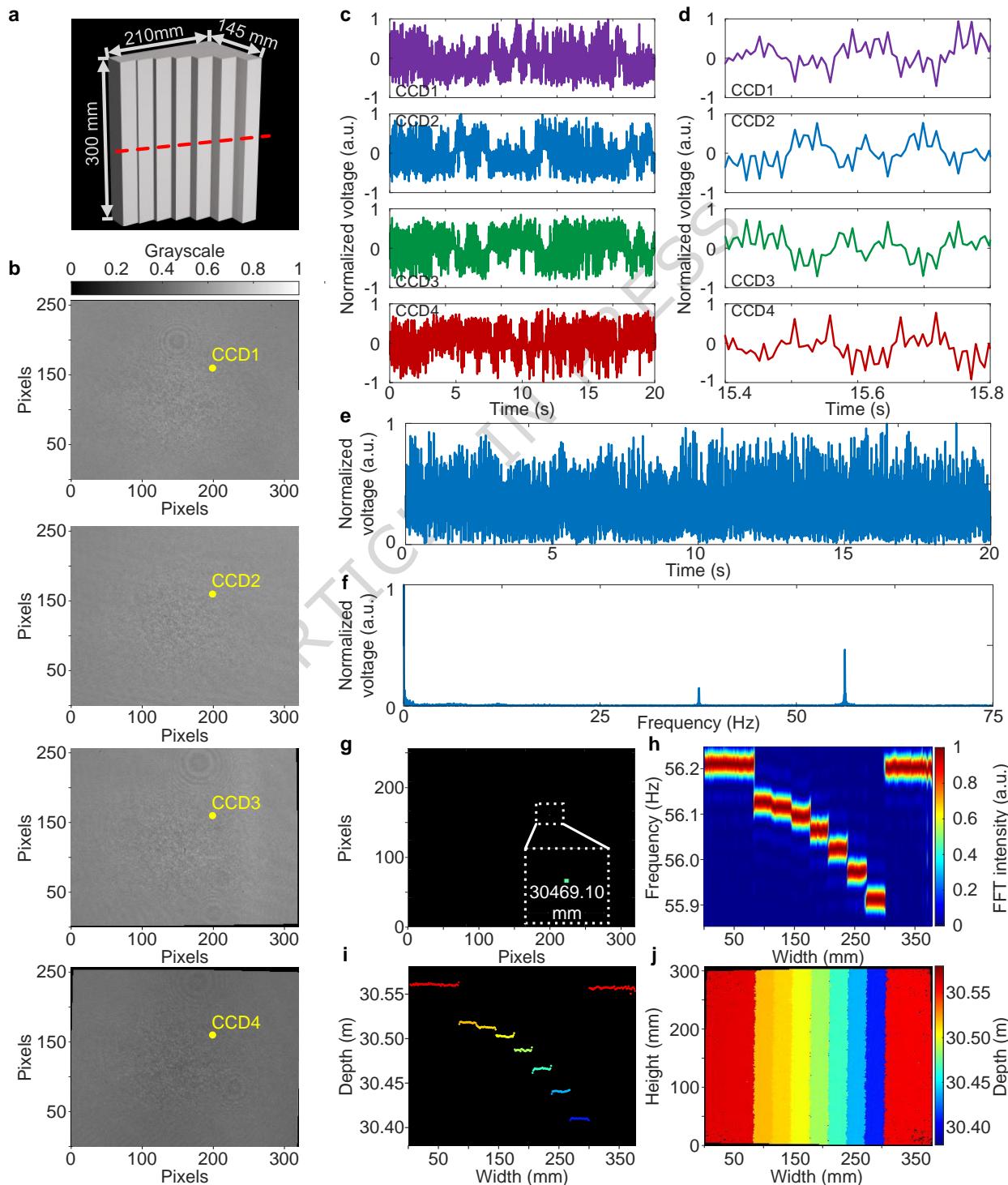
Fig. 3. 3D imaging results of a traffic scene model. a, Photograph of the traffic scene model. **b**, Reconstructed 3D image of the model. **c**, Side view of the 3D image. **d**, Occurrences at various depths. **e**, Front view of the 3D image. **f**, Recovered 2D image of the model using a 2D reconstruction algorithm.
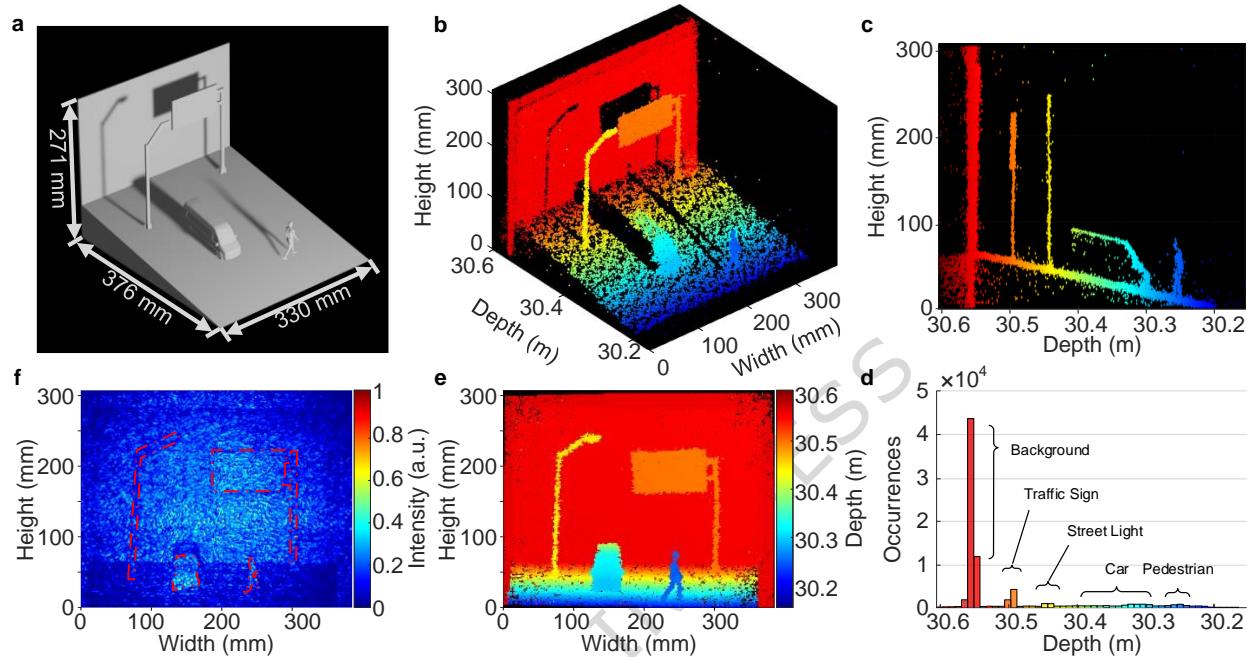
Fig. 4. Range precision analysis and frame rate evaluation. a, 3D Imaging result with a planar target placed at 30.5 m. **b**, Depth distribution histogram. **c**, Theoretical (blue) and measured (red) precisions comparison as the planar target distance increases. **d**, Theoretical precision when the bandwidth of the probe signal is increased from 6 GHz to 30 GHz, and the target is located at the distance from 30 m to 240 m. The probe optical power is set to be 15.86 mW. **e**, 3D images of a BIT alphabet model are generated at a data acquisition time of 0.10s, 0.25s, 0.50s, 1.0s, 2.0s, 5.0s, 10.0s and 10.0s with air turbulences. **f**, Measured precision with different frame counts.
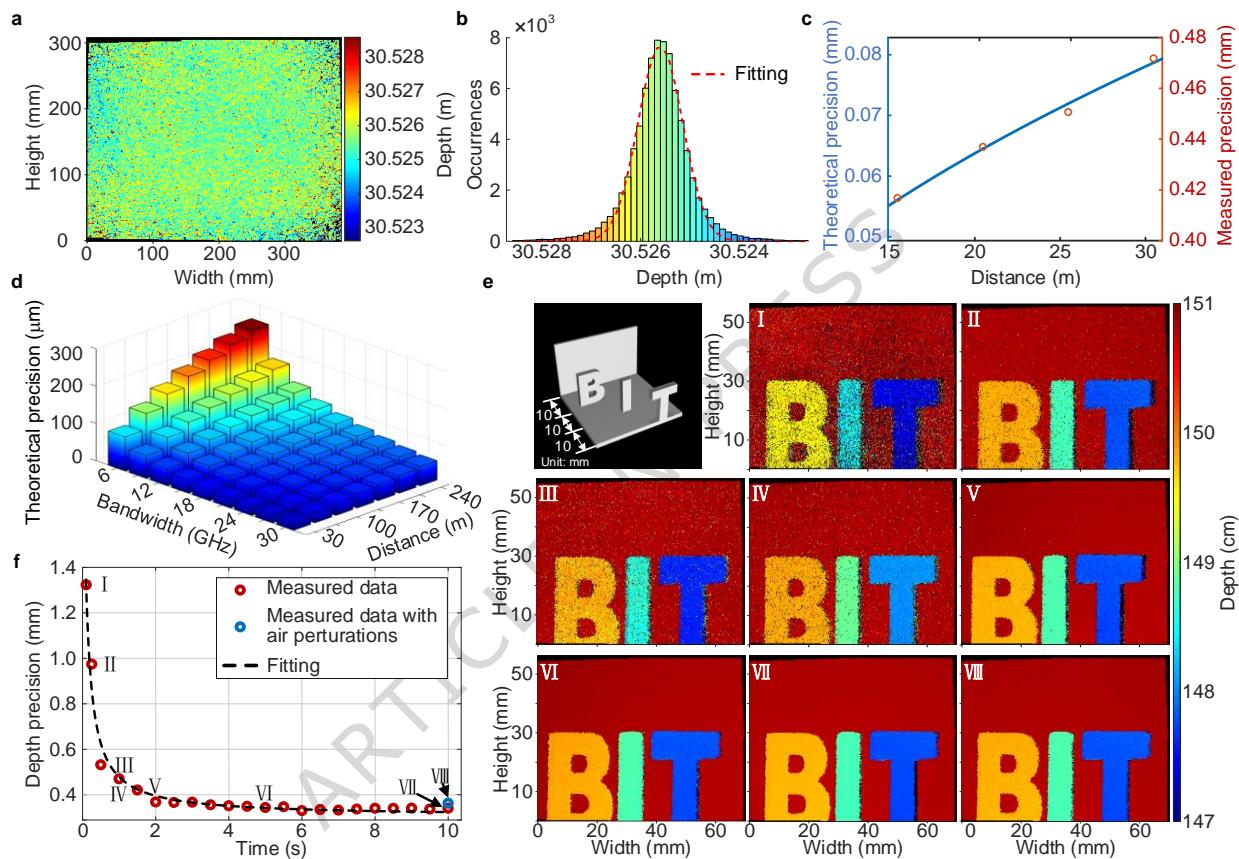
Fig. 5. Interactive 3D reconstruction of a bust sculpture. a, Photograph of a bust sculpture. **b**, 3D image of the bust sculpture at the initial orientation. **c**, Virtual bust sculpture from various perspectives.
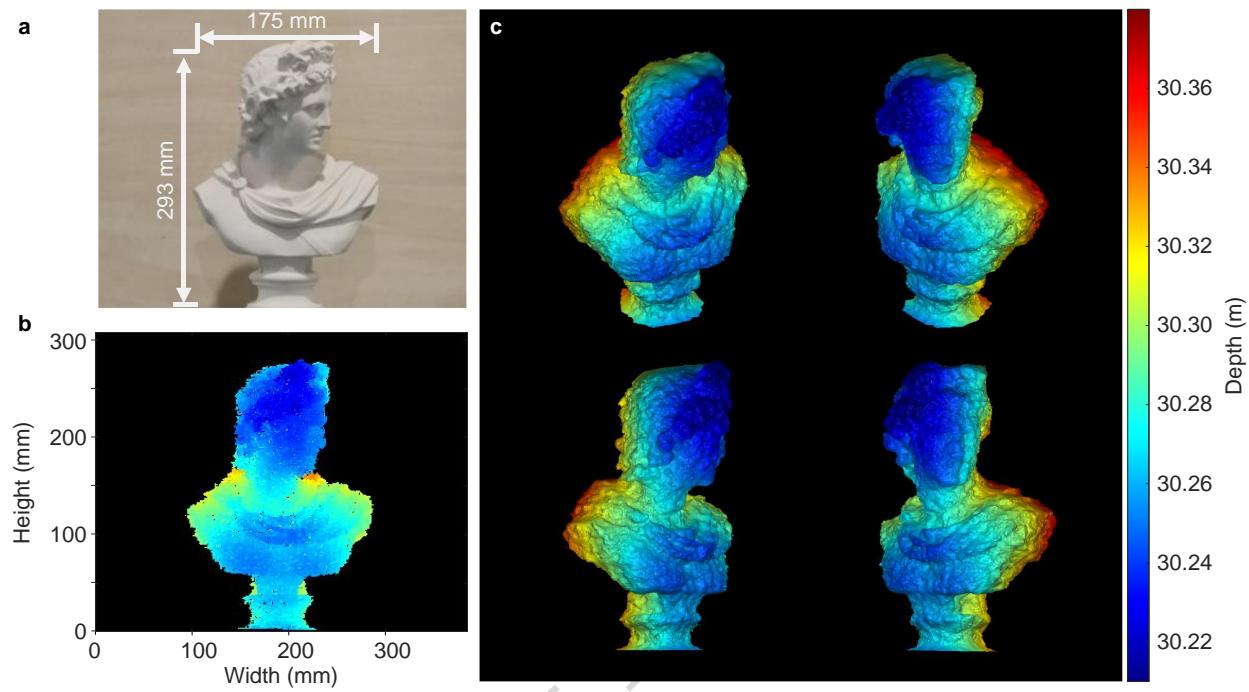
Editorial Summary:
A solid-state flash 3D imaging system is demonstrated using stepped-frequency modulation and CCD-based coherent detection, achieving submillimeter precision, long-range detection, and scalable pixel performance.

Peer review information: *Nature Communications* thanks Xingjun Wang and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is availablae.