nature communications

Article

https://doi.org/10.1038/s41467-023-42678-5

Polarization structured light 3D depth image sensor for scenes with reflective surfaces

Received: 9 March 2023

Accepted: 18 October 2023

Xuanlun Huang[1,2,4], Chenyang Wu[1,2,4], Xiaolan Xu[2], Baishun Wang[2], Sui Zhang[2], Chihchiang Shen[2], Chiennan Yu[2], Jiaxing Wang[2], Nan Chi[1,3], Shaohua Yu[1,3] & Connie J. Chang-Hasnain[1,2] ✉

Highly reflective surfaces are notorious in the field of depth sensing and three-dimensional (3D) imaging because they can cause severe errors in perception of the depth. Despite recent progress in addressing this challenge, there are still no robust and error-free solutions. Here, we devise a polarization structured light 3D sensor for solving these problems, in which high-contrast-grating (HCG) vertical-cavity surface-emitting lasers (VCSELs) are used to exploit the polarization property. We demonstrate accurate depth measurements of the reflective surfaces and objects behind them in various imaging situations. In addition, the absolute error and effective measurement range are measured to prove the applicability for a wide range of 3D applications. Our work innovatively combines polarization and depth information, opening the way for fully understanding and applying polarization properties in the 3D domain.

Highly reflective surfaces, such as glass, mirror, and water surface, are common scenes in 3D imaging. They often cause irrevocable errors in depth sensing and 3D imaging, such as mirages leading to wrong depth measurements or specular reflection blinding 3D sensors. However, the correct perception of the surrounding environment is critical for many applications, such as robotics, scene reconstruction, and virtual reality. Hence, it is extremely important to manage the problem of reflective surfaces. In recent decades, there are plenty of techniques managing to detect the reflective surfaces, including polarization imaging, active projectors and its fusion with other sensors, and deep learning. However, due to the effects of reflective or multipath noise, most of these techniques will result in erroneous measurements of the reflective surfaces.

Polarization imaging is able to obtain the 3D shape of an object by analyzing several images of reflected intensity at different polarized angles. Its first use to determine the 3D orientations of reflective surfaces dated back to the 1990s[1]. Since then, polarization imaging has been used to reconstruct the shape of transparent[2–4] and specular[5–8] objects. Furthermore, the integration with cues from various depth sensors[9–14] enables polarization imaging to obtain better shape recoveries of the targets. However, polarization imaging needs to capture multiple polarized images and can only estimate relative depths. In addition, most of them requires knowledge of the refractive index of the object, while some rely on other factors to ease this need, such as various light conditions[15,16] or multiwavelength[17]. In[11], Berger combined polarization and stereo vision and further extended the application to general environments with reflective surfaces. However, due to the inability to differentiate mirages from real objects, stereo vision will result in a virtual and wrong distance. Some methods use multiple polarized images[18–22] or local reflection cues[23,24] to separate the reflection. In some cases, an additional infrared sensor is used to provide depth information and to remove the mirage images[25,26]. One major drawback of polarization-stereo-vision methods is that they do not result in correct depth measurements of the reflective surfaces (e.g., the depth of the glass).

Active projectors, such as laser range finders and structured light projector, are widely used in the field of depth sensing. In the case of laser range finder, in order to deal with the reflective surfaces, some researchers use the reflected intensity profile to determine the glass area[27–29]. The fusion of laser range finder and sonar[30,31] is also developed for navigation in the glass environment. In addition, the method used in polarization imaging is applied to the laser range finder, in

---

[1]School of Information Science and Technology, Fudan University, 200433 Shanghai, China. [2]Berxel Photonics Co. Ltd., 518071 Shenzhen, China. [3]Peng Cheng Laboratory, 518055 Shenzhen, China. [4]These authors contributed equally: Xuanlun Huang, Chenyang Wu. ✉e-mail: connie.chang@berxel.com

which the degree of polarization is calculated and serves as a detection standard of glass areas[32]. However, these methods require scanning around to determine the depth and glass points, thus they are still in the stage of 2D route mapping. As for structured light projectors, a method[33] employs the fusion of Kinect and sonar to obtain the depth images of glass scenes. But, due to the sparse data and narrow angle of the sonar, it needs to sweep multiple times to obtain enough information. Besides, the fusion with sonar will make the system costly, huge and complex.

Deep learning is also a good candidate for the detection of reflective surfaces[34–37], where a network model trained on abundant images of reflective surfaces is employed to mark out the target area. But we find that even the recent works[36,37] might still misjudge the areas with frames or borders, such as an empty frame without glass, as the reflective surfaces.

Therefore, a compact and robust method is still needed for applications in environments full of highly reflective surfaces. Here, we develop a polarization structured light (PSL) 3D sensor[38] with polarization properties on both the transmitter (TX) and receiver (RX). In TX, high-contrast-grating (HCG) vertical-cavity surface-emitting lasers (VCSELs) are specially designed to provide structured light with a strong polarization selection ratio. In RX, a polarization-selection CMOS camera is designed to receive the signal selectively. According to Fresnel's theory[39], the specular reflection from a reflective surface maintains the same polarization as the incident polarized light. However, diffused reflection from objects without smooth surfaces does not exhibit any polarization even incident by a strongly polarized light. Hence, using a polarization-selection CMOS camera can differentiate reflection from a reflective surface. Thus, the depth information of a reflective surface or objects behind it can be obtained based on the choice of the polarizer direction. Here, we report three experiments to demonstrate how PSL 3D sensors can be used to see as well as to see through highly reflective surfaces.

Results

Polarization structured light 3D sensor

The PSL 3D sensor is illustrated in Fig. 1a. It consists of a transmitter TX, a receiver RX and an RGB camera. In TX, we have embedded an HCG-VCSEL array inside, whose working wavelength is 940 nm. VCSELs are critical sources for dot projectors used in structured light cameras[40]. Typical VCSEL uses distributed Bragg reflectors (DBRs) as its top and bottom mirrors. The DBRs are many 10s of layers of planar heteroepitaxial layers and do not provide or maintain a fixed polarization to the VCSELs. HCG is a thin-film subwavelength metastructure that is effective in providing high reflection with a fixed polarization. Hence, HCG VCSELs have exhibited a very high polarization selection ratio independent of operating temperature or drive conditions[41]. With this specially designed HCG-VCSEL array, the TX of PSL 3D sensor can project dot-array structured light (see Supplementary Fig. S1 and Methods) in either transverse electric (TE) or transverse magnetic (TM) polarization.

The polarization direction of the structured light emitted from TX is denoted by a double-head blue arrow in Fig. 1a and its projection range is $75.5° \times 65.7°$, which is indicated by a pink area bounded by

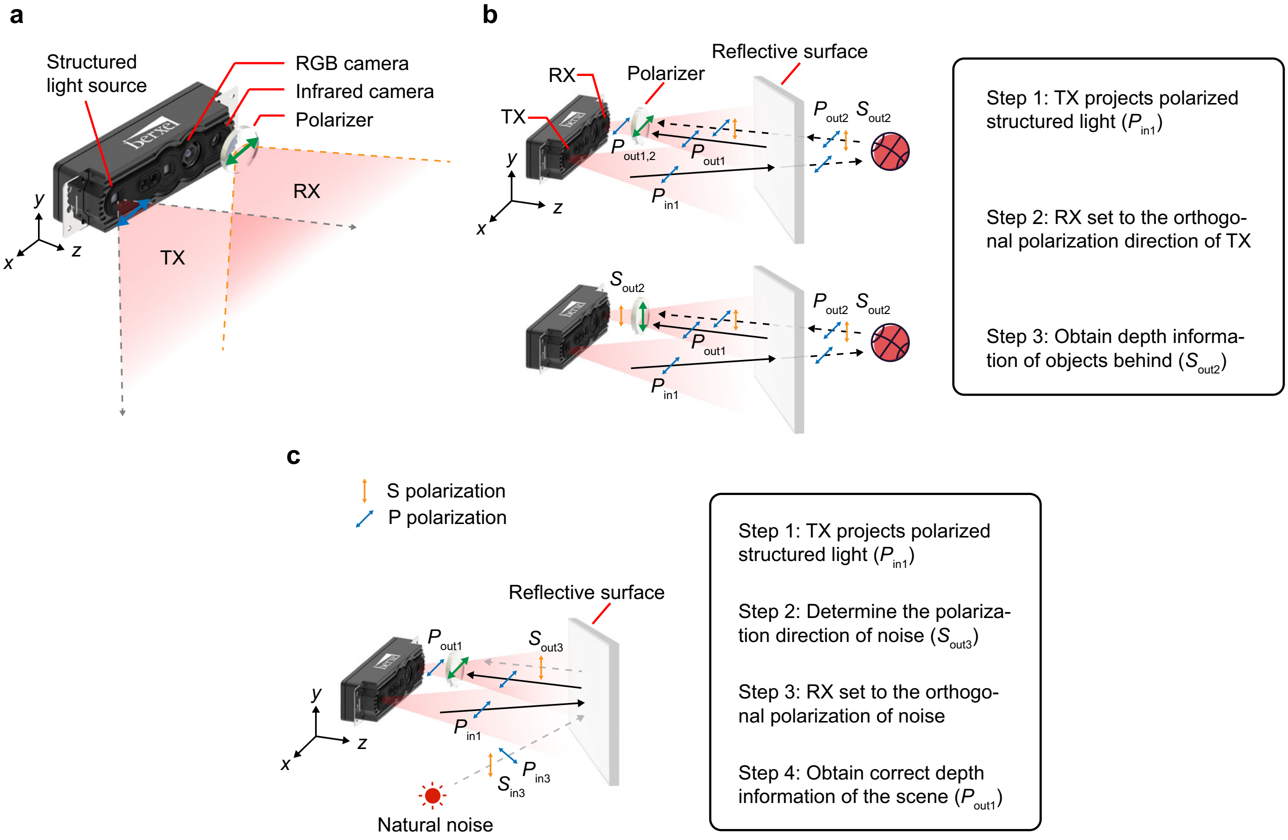

Fig. 1 | Principle of the polarization structured light 3D sensor. a The model of PSL 3D sensor. In TX, polarized structured light is produced by an HCG-VCSEL array inside. In RX, signals are received selectively by a rotatable polarizer mounted on the receiving infrared CMOS camera. In addition to TX and RX, PSL 3D sensor contains a RGB camera in the middle as well. **b** The scheme of seeing through the reflective surface. On the upper row, it shows the 3D sensor can obtain depth information of both the reflective surface and the objects behind. On the lower row, the 3D sensor can remove the depth of the reflective surface. The specific steps are shown in the flow diagram. **c** The scheme of seeing against reflective noise. On the left side, it shows the 3D sensor can obtain the correct depth of the scene by eliminating the effect of reflective noise. The flow of this mission is shown on the right. Note that the plane of the PSL 3D sensor is set to be *x*-*y* plane and the light propagation direction is in *z* direction.

gray dot arrow lines. Here, we set the plane of the PSL 3D sensor to be x-y plane and the light propagation direction to be z, then define the horizontal (x direction) polarization, as pointed by the blue arrow in Fig. 1a, to be TM polarization. In RX end, a rotatable polarizer is mounted on the infrared CMOS camera. Its polarization direction is denoted by a double-head green arrow and its receiving angle is $74^\circ \times 50.5^\circ$, which is indicated by a pink area bounded by orange dot arrow lines. The polarizer can be rotated to receive the dot patterns selectively. By calculating the spatial displacements of known dot patterns, depth information can be obtained for an entire image at high frame rates.

Seeing through the reflective surface

According to the polarization feature of the HCG-VCSEL-based PSL 3D sensor, we can fully exploit its merits for 3D imaging in scenarios with large reflective surfaces. Based on Fresnel equations[39], when S- or P-polarized light is incident on a highly reflective surface, the reflection will maintain the same polarized direction. As shown in the upper picture of Fig. 1b, TX of the PSL 3D sensor projects TM-polarization (P-polarization, P_{in1}) structured light onto the scene. The reflected light from the reflective surface (P_{out1}), which is denoted by the black solid line, will bounce back with the same polarization as the TX, while the light reflected from the object behind the reflective surface (S_{out2}, P_{out2}), whose light path is denoted by the black dot line, is polarized in various directions due to diffuse reflection. Then RX can receive the signals ($P_{\text{out1,2}}$) of both the reflective surface and the object behind with its polarizer set in the same polarized direction, i.e., x direction. As shown in the lower picture of Fig. 1b, if we change the direction of the polarizer to be orthogonal, i.e., y direction, the depth information of the reflective surface is eliminated, leaving that (S_{out2}) from the diffuse object behind.

Thus, for seeing through the reflective surface, the process can be divided into three steps, as illustrated in the flow of Fig. 1b. Step 1 is the projection of polarized structured light with a specific polarization direction (P_{in1}). Step 2 is the setting of RX polarization direction to be orthogonal to TX. Step 3 is obtaining the depth information of the scene (S_{out2}), where the part of the reflective surface (P_{out1}) is removed. In addition, the distinction between two polarization settings of RX can be further used to complete the reflective surface.

Seeing against the reflective noise

In Fig. 1c, we also show the principle of working against reflective noise. In this sketch, the influence of natural light (S_{in3}, P_{in3}) is indicated by the gray dash line, whose incident plane is in the x–z plane. Its reflected component is stronger in S-polarization (y direction, S_{out3}) according to Fresnel equations[39]. Hence it can be eliminated when a P-polarized TX and RX are used to enhance the signal-to-noise ratio (SNR), enabling clear 3D image acquisition of the whole scene where there is strong reflective noise from the natural light. The process for this mission can be concluded into four steps, as illustrated in the flow of Fig. 1c. Step 1 is the projection of polarized structured light with specific polarization direction (P_{in1}). Step 2 is the determination of polarization direction of noise (S_{out3}). Step 3 is the setting of RX polarization direction to be orthogonal to the reflective noise. Step 4 is obtaining the correct depth information of the scene (P_{out1}).

Completing the reflective surface

In addition to seeing through the reflective surface and seeing against the reflective noise, we further present the PSL 3D sensor's capability to detect and complete the reflective surfaces. The two-step working principle is illustrated in Fig. 2. In step 1, we extract the glass region by the collaboration of the depth and color channels. In the depth channel, we first obtain the depth image of the scene in polarization 0°. Polarization 0° means that both the TX and RX of the PSL 3D sensor are in the same polarization, here we set them in TM polarization. Next, we rotate the RX to the orthogonal polarization, obtaining the depth image in polarization 90°. As illustrated in the principle of seeing through the reflective surface, the glass part can be eliminated because its reflected polarization is orthogonal to the RX. Then, subtraction is applied between these two depth images to capture the area with sharp contrast. Meanwhile, in the color channel, we adopt the deep learning method[36] to get the glass boundary from the RGB image of the scene. For a glass pixel, it is predicted to be 1 (white) and for a non-glass pixel, it is 0 (black). Combining the subtraction result in the depth channel and the boundary result in the color channel, the depth points belonging to the glass can be determined and extracted, which is highlighted in red in the depth image. Note that, for scene without glass, there is no sharp contrast between two polarization settings in the depth channel (see Supplementary Fig. S3 and the results of completing the reflective surface).

After getting the glass region, we move forward to step 2. The extracted glass depth points $P_d = [u, v, z]^\top$ are transformed to the world coordinate of the infrared camera $P_{\text{ir}} = [x, y, z]^\top$ using the following equation:

$$\begin{array}{rcl} x & = & \frac{z(u-c_x)}{f_x} \\ y & = & \frac{z(v-c_y)}{f_y}, \end{array} \tag{1}$$

where $\{f_x, f_y, c_x, c_y\}$ are the internal parameters of the infrared camera. Then the glass is fitted to a large enough area in this world coordinate (Supplementary Fig. S4 explains why it needs to fit in the world coordinate). The fitted glass points are next transformed to the world coordinate of the RGB camera $P_{\text{rgb}} = [x', y', z']^\top$ with the external parameters, i.e. the rotational matrix R and translational matrix T:

$$P_{\text{rgb}} = RP_{\text{ir}} + T. \tag{2}$$

These points P_{rgb} are mapped to the color channel $P_c = [u', v', z']^\top$ using equation (1) but with the internal parameters of the RGB camera $\{f'_x, f'_y, c'_x, c'_y\}$. In the color channel, we can interpolate inside the glass boundary according to the fitted glass points P_c and finish the completion of the glass. At last, the completed point cloud can be converted back to the world coordinate. As shown in the right part of step 2, the original point cloud is plotted in the first row, where only the middle region of the glass is detected. After applying step 1 and 2, the whole glass area can be completed, which is highlighted in the second row with red color. With this proposed method, not only the glass can be seen and completed correctly, but also the objects behind the glass can be reconstructed, enabling seeing and seeing through the reflective surfaces.

Analysis of 3D imaging in scenes with reflective surfaces

Before showing the results of three experiments, we first analyse 3D imaging in scenes with reflective surfaces. Among the reflective surfaces, the indoor glass wall is a typically challenging situation (Fig. 3a, b). In this scene, we place the sensor in front of a glass wall, with a reflecting image of two persons and the sensor on the same side of the glass at about 1.2 m distance in front. In this situation, the depth image obtained with a stereo vision camera (Intel RealSense D455) is shown in Fig. 3c, which is unable to see the depth of the glass. In fact, the stereo vision camera provides erroneous depth measurements based upon the image of the persons reflected by the glass with a measured distance of about 2.4 m behind the glass wall, whereas the actual should be 1.2 m in front of the glass wall. In contrast, when TM polarized structured light patterns produced from the TX of PSL 3D sensor hit the glass, depth information of part of the glass can be reconstructed with TM polarization in RX, as shown by the area encircled by the dashed line in Fig. 3d. These measured values agree well with the actual

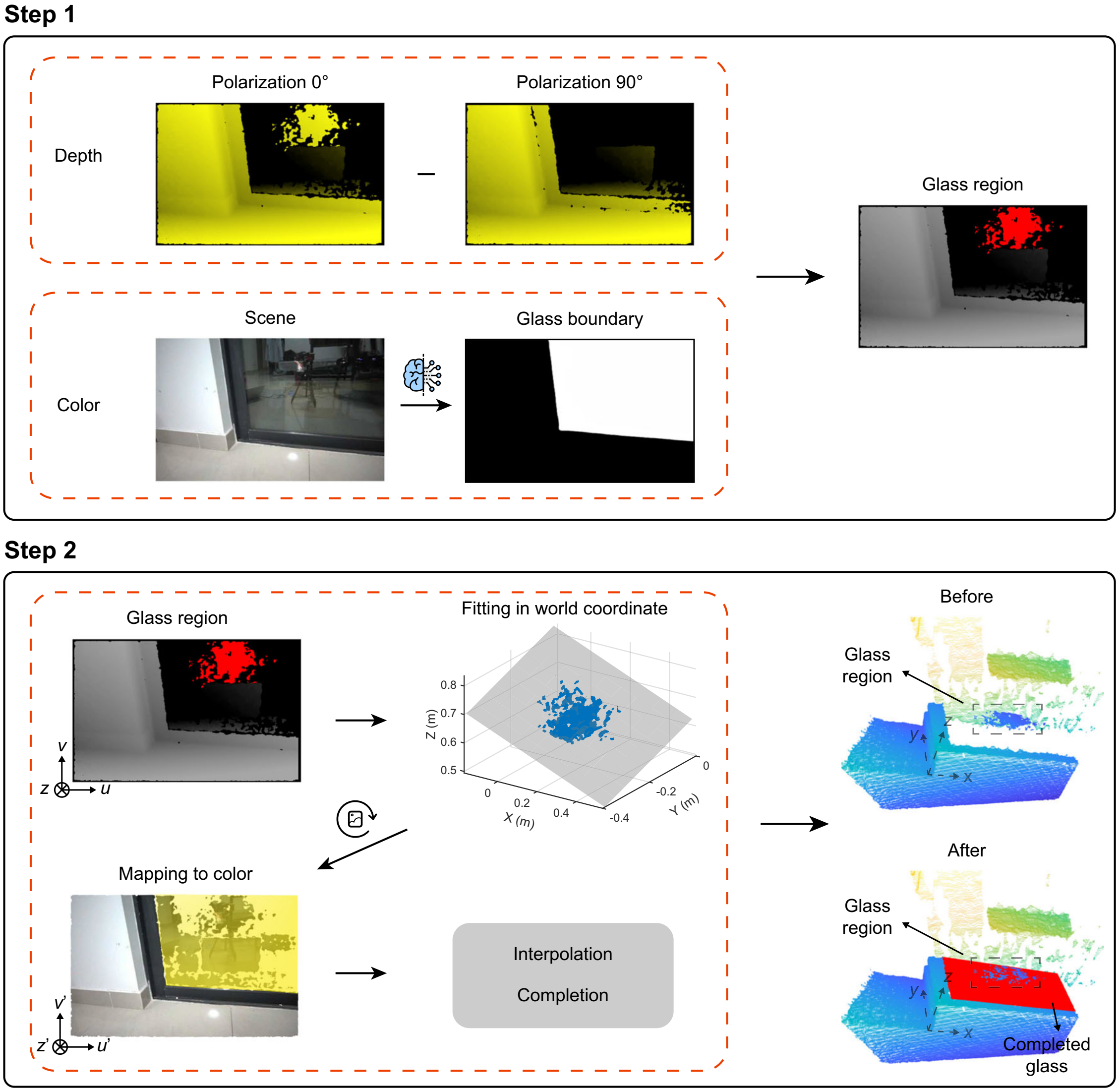

Fig. 2 | Principle of completing the reflective surface. In step 1, the glass region is extracted through a combination of depth and color channels, where the glass points are highlighted in red. In depth channel, two depth images are obtained in polarization 0° and 90° respectively. Then, subtraction is employed between these two depth images for extraction of the glass points. In color channel, the glass boundary is predicted from the color image of the scene using deep learning method. In step 2, the extracted glass points are transformed to the world coordinate for fitting. After fitting, the results are mapped to the color channel. In this channel, the reflective surface can be interpolated and completed inside the glass boundary. At last, the completed point cloud can be converted back to the world coordinate, as shown in the before and after comparison on the right side.

distance and this correct depth information can be further used to demonstrate and complete the glass wall.

Deep learning is an effective tool to detect glass[36,37], where a trained deep neural network is used to predict the glass part. However, such RGB-image-based deep learning method will cause misjudgement in some scenes. As shown in Fig. 3e, a cardboard box with its front surface removed is placed in front of the sensor. Inside the box, there is a box of gloves and a metal breadboard. The RGB image of this scene is sent to the network for prediction and its result is shown in Fig. 3f. The prediction result is the probability of being considered as glass, whose range is 0 to 1. For a glass pixel, it is evaluated as 1 (white) and for a non-glass pixel, it is 0 (black). From the result, we can see that almost the whole front surface area is considered to be the glass area. This is because the context of having box boundaries and objects inside will lead to errors in prediction. Moreover, we test another glass door scene. As shown in Fig. 3g, we place the sensor in front of a soundproof room with its door open. That is, there is glass on the left and no glass on the right. Inside the soundproof room, a cardboard box with a basketball on its top is placed on the left, and another box and a book are on the right. Again, the result in Fig. 3h shows that the network predicts both parts as glass areas.

Due to these disadvantages, a much more robust method for the detection of reflective surfaces is required to develop. On the one hand, it needs to obtain the correct 3D information on reflective surfaces. On the other hand, it should be able to determine whether the depth information belongs to them. In the following, we will illustrate how the PSL 3D sensor and its corresponding imaging methods facilitate in such a mission.

Results of seeing through the reflective surface

We start with a typical situation, i.e. glass door with objects behind, to illustrate PSL 3D sensor's ability to see through the reflective surface. We analyze different factors that can affect this mission, including the distance from the 3D sensor to the glass, the 3D sensor incident angle, the distance of objects behind glass and the density of objects behind glass. Their specific conditions are listed in Table 1.

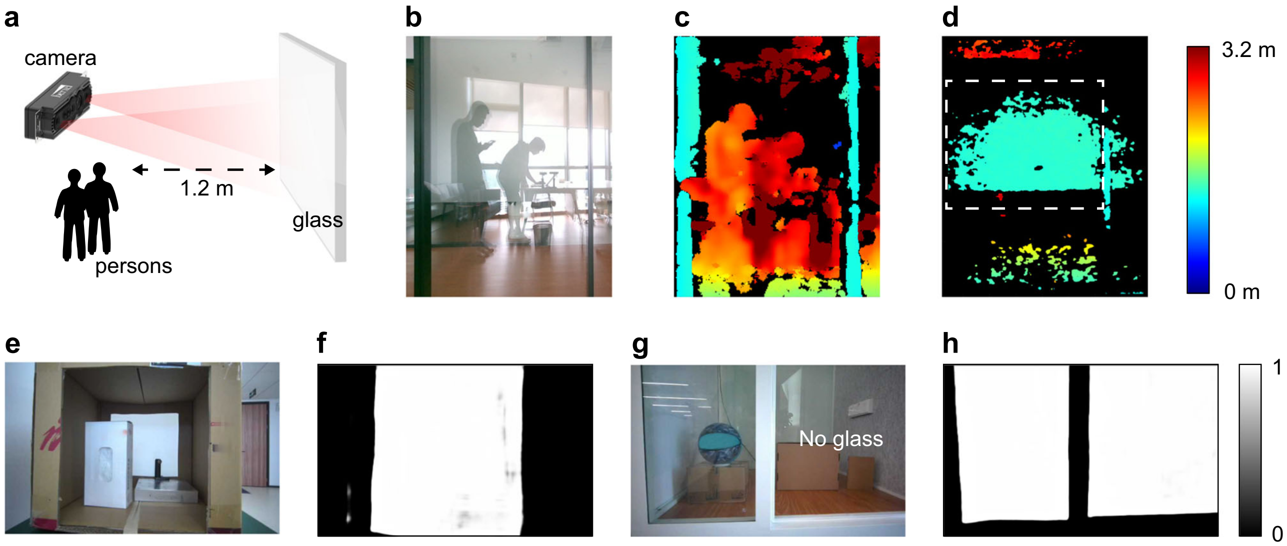

Fig. 3 | Analysis of 3D imaging in scenes with reflective surfaces. (**a**) and (**b**) are scheme and picture of the indoor glass wall scene. The glass wall is about 1.2 m in front of the sensor, with a reflecting image of two persons and the sensor. Note that two persons are on the same side as the sensor. **c** Depth image from the stereo vision camera (Intel RealSense D455). **d** Depth image from the PSL 3D sensor. The glass part is denoted by the dashed line. (**e**) and (**g**) are testing scenes of the deep learning method. Their corresponding prediction results are shown in (**f**) and (**h**). The jet color bar applies to (**c**) and (**d**), the gray color bar applies to (**f**) and (**h**).

In Fig. 4, we show the influence of the first factor. The color pictures of the scene are shown in the first column. In this situation, the PSL 3D sensor is facing the glass door with a distance changing from 0.4 m to 1.2 m. Behind the glass door, there is a movable cabinet with a basketball on the top. Their distances to the glass are fixed to 0.5 m. The TX of the PSL 3D sensor projects TM-polarized structured light to the scene. Then the structured light is reflected back to the RX, containing depth information of both the glass door and objects behind. As mentioned in the principle, the reflected light of the glass door maintains the same polarization as the TX, while those of the cabinet and basketball are in various polarized directions due to diffuse reflection.

Hence, when we set the RX to be TM polarization, depth information of the glass door, cabinet, and basketball can all be reconstructed, which is shown in Polarization 0° of Fig. 4a. In this depth map, bright yellow represents the near points and black represents the far or missing points. The middle bright yellow area accurately reveals the depth of the glass, while the farther cabinet and basketball are darker. Unlike Fig. 3d, the central part shows the depth of the basketball rather than the glass because reflection from the basketball is stronger. Next, if we rotate the RX to the orthogonal direction, i.e. TE polarization, the glass part can be filtered, leaving clear depth information of the cabinet and basketball (Polarization 90° in Fig. 4a). As we change the distance to 0.8 m, the middle ball area becomes smaller because the glass depth begins to dominate (dot line in Polarization 0° of Fig. 4b). As for 1.2 m (Fig. 4c), the middle is almost replaced by glass depth, while the whole proportion of glass depth is smaller as compared to that of 0.4 m. In Polarization 90° for both cases, the glass part can still be eliminated.

Table 1 | Influence factors for seeing through the reflective surfaces

Influence factor	Specific experimental condition		
Distance from 3D sensor to glass	0.4 m	0.8 m	1.2 m
3D sensor incident angle	−30°	0°	30°
Distance of objects behind glass	0.3 m	0.5 m	0.7 m
Density of objects behind glass	Low	Medium	High

More details about the incident angle of 3D sensor, distance and density of objects behind glass are displayed in Supplementary Fig. S5–S8. Regardless of different incidence angles, or different distances and object densities, the 3D sensor can eliminate the depth of the glass and achieve the function of seeing through the reflective surface. Supplementary Movie 1 clearly shows this phenomenon by rotating the RX polarizer 360 degrees. The high contrast between Polarization 0° and Polarization 90° can further serve as a determination method of the reflective surface, which will be demonstrated in the section after next.

Results of seeing against the reflective noise

We also set up outdoor and indoor scenes to demonstrate the advantage of detecting against noise (Fig. 5). We first show a case of outdoor glass scene (Fig. 5a). In this situation, a book is placed behind a tilted glass, where there is strong specular reflection from the natural light. We can see that the reflected noise incapacitates the stereo vision camera for obtaining the depth of the scene, leading to black-irregular-patch errors in the depth map (encircled by white and red dot lines while the book is indicated in red). However, as mentioned in the Fig. 1c, this noise has a larger component of S polarization. Note that in this setting, S polarization is in the horizontal direction (x direction) when facing the glass. Therefore, we can utilize P-polarized (i.e. y-direction polarization) TX and RX to filter this noise, enhancing the SNR. As shown in Fig. 5c, the book behind the glass (encircled by the red dot line) can be measured clearly with the PSL 3D sensor in this P-polarized setting.

Besides this sunlight noise, the shadow of objects may also severely blind the depth camera based on stereo vision. As illustrated before in Fig. 3c, the reflection noise of two persons makes the stereo vision camera provide erroneous measurement. Another outdoor scene with reflection overlapped with the detecting object is also illustrated in Supplementary Fig. S9, where the stereo vision camera is still affected by the reflection of outside objects while the PSL 3D sensor can get the depth information correctly.

In the next case, we analyze the ability to detect against noise in the situation of wall corner, which is a common scene for indoor serving robot. The results are displayed in Fig. 5d–l. As shown in Fig. 5d, the comparison is first done at a height of 0.24 m and a distance of 0.7

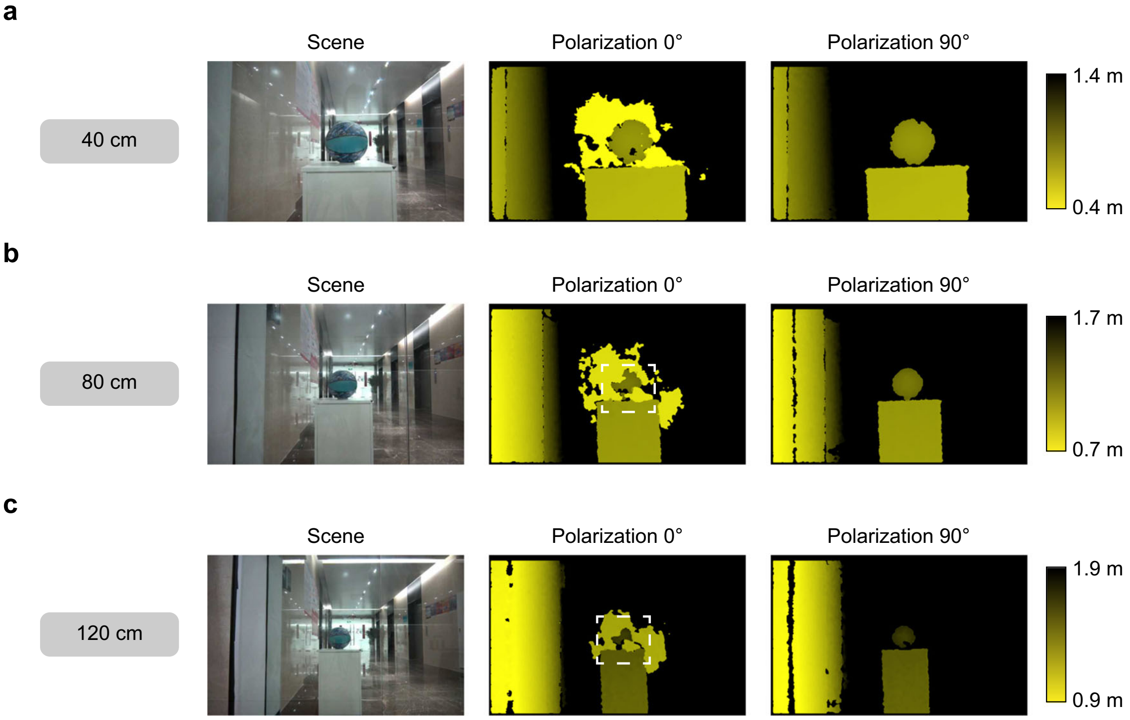

Fig. 4 | Seeing through the reflective surface. (**a**)–(**c**) are the results of different distances from 3D sensor to glass. Polarization 0° is the depth map when TX and RX are in the same polarization. Polarization 90° is the depth map when TX and RX are in the orthogonal polarization. The yellow-black color bars apply to both second and third columns.

m away from the corner, with the sensor 45° downward. In this situation, due to the influence of reflection and multipath noise, the stereo vision camera will make the wrong depth judgment. As indicated by the red dot line in Fig. 5e, the depth has gone deep below the floor plane. Only a small part (orange dot line) has accurate floor information. However, for the PSL 3D sensor (Fig. 5f), a complete and correct shape of the corner can be obtained by setting TX and RX to be P polarization (*y* direction). Then we change to the height of 0.4 m (Fig. 5g), with a distance 0.7 m and an incline angle 45° unchanged. It means that the incident angle decreases. In this condition, the correct floor area obtained from the stereo vision camera starts to increase (orange dot line in Fig. 5h), but the depth area encircled by the red dot line still penetrates into the floor. As we change to the height of 0.56 m (Fig. 5j), the floor area of the stereo vision camera goes back to the correct plane (Fig. 5k), but there are still some black speckles in the depth map because of noise. However, for the PSL 3D sensor, the intact depth map and point cloud are maintained in these two heights (Fig. 5i and l) with this P-polarized setting.

For these common scenes of reflective surface, PSL 3D sensor can remove noise through specific polarization combinations, improve the signal-to-noise ratio, and obtain complete depth information of the scene. More polarization combination experiments and comparison with time of flight (ToF) camera, presented in Supplementary Fig. S10–-S13, further demonstrate the superiority of PSL 3D sensor in these scenarios.

Results of completing the reflective surface

In the following, we prove the feasibility of completing the reflective surfaces. The results of eight common scenes are selected to illustrate in Fig. 6 and Supplementary Fig. S14. In Fig. 6, we display four of these scenes, including balcony glass, soundproof room glass, spherical glass and office door glass. In the balcony glass scene, a yellow square box is placed outside the glass door which has reflections of the interior furniture. With polarization 0° and polarization 90° settings, we obtain two depth images of the scene. In polarization 0°, partial of the glass and the box outside can both be reconstructed while the glass part is eliminated in polarization 90°. Compared to the stereo vision camera, which fails to detect the glass depth and misjudges the reflections as real objects, such a change in depth channel can serve as a reliable cue for determining the reflective surfaces. Thus, using the subtraction of two depth images and the glass boundary predicted from the RGB image of the scene, we are able to extract the glass region, which is highlighted in red in the fifth row. The extracted glass points are then used to fit and complete the reflective surface. The completed depth image is shown in the seventh row, where the partially empty glass has now been filled in. The final comparison between the original point cloud and the completed point cloud is shown in the last two rows. In the completed point cloud, we can see that the new glass plane, which is colored in red, matches well with the glass frame and the 3D information of the box outside the glass can be acquired simultaneously.

As for the soundproof room glass scene, we execute the same process. In this scene, the right door of the soundproof room is open, that is, only the left side has glass. Inside the room, a cardboard box with a basketball on its top is placed on the left, and another box and a book are on the right. As seen in the depth images, only the left side with glass has a large change while the right side shows no significant change. If simply relying on the predicted glass boundary from deep learning, we will mistakenly consider the right side as glass. However, our method can determine the regions that actually have glass and extract effective depth information belonging to the glass. Similarly, this accurately extracted depth information is used for the following fitting and completion. In the completed results, the left glass door and objects inside the room can be reconstructed, as if exploring the inside and outside of the soundproof room from God's perspective.

Besides planar glass, our method can also be applied to curved glass, which is displayed in the third column. In this situation, we use a spherical exhibition stand with two 3D-printed models inside as the

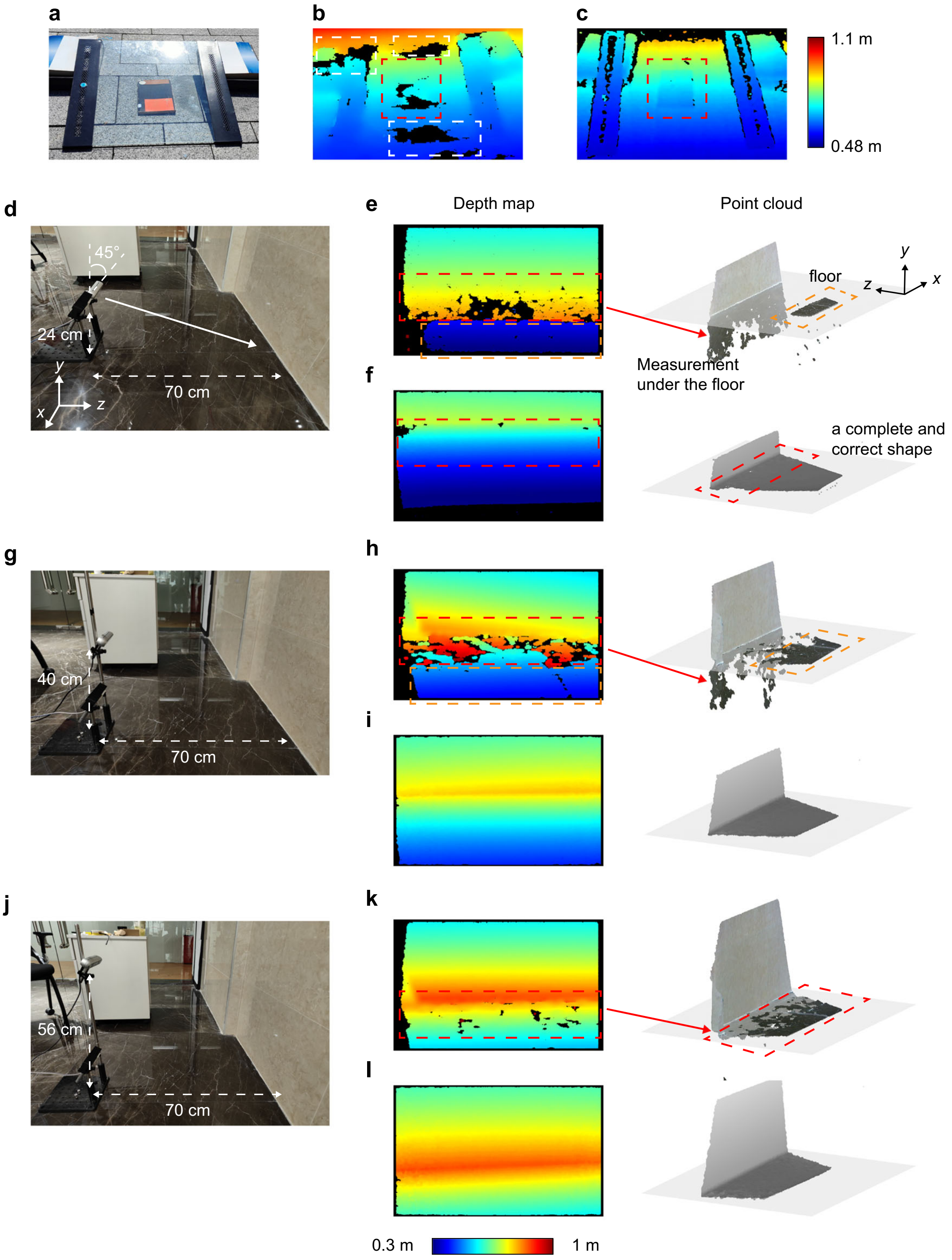

Fig. 5 | Seeing against the reflective noise. (**a**)–(**c**) are comparison results of an outdoor glass scene, where (**a**) is the picture of the scene, (**b**) is the depth map from stereo vision camera (Intel RealSense D455) and (**c**) is the depth map from PSL 3D sensor. (**d**)–(**l**) are comparison results of an indoor corner scene. The experiments are done in different height, whose schemes are shown in (**d**), (**g**), and (**j**). The corresponding depth map and point cloud from stereo vision camera are displayed in (**e**), (**h**) and (**k**) respectively, while those from PSL 3D sensor are in (**f**), (**i**), and (**l**). The jet color bar with range (0.48 m, 1.1 m) applies to (**b**), (**c**), and the one with range (0.3 m, 1 m) applies to the depth map column.

detecting object. Likewise, two depth images are obtained from two polarization directions. We can see that the front face of the glass ball is eliminated in polarization 90°. With the ball-shaped glass boundary, we acquire the glass region in the fifth row. These points are from the spherical glass so the fitting result is a quadric surface. Then we complete the glass inside the glass boundary and get the completion result. As seen from the point cloud in the last row, it completes well the front surface of the spherical glass. Note that, there are still some misjudged points in the predicted glass boundary, such as the unsmoothness and the false positive outliers at the ball edge. These minor flaws can be solved by optimizing the predicting network in future work.

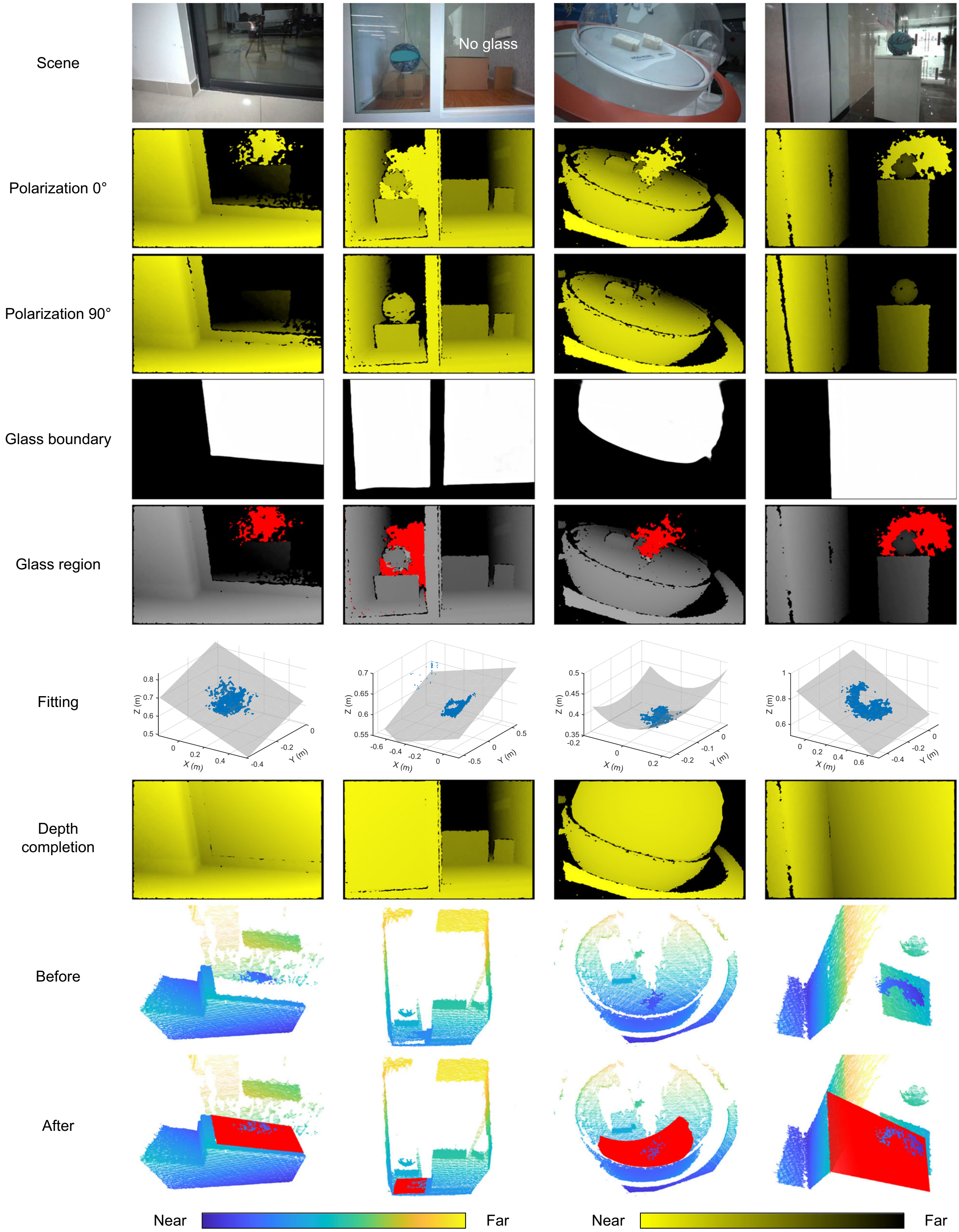

Fig. 6 | Detection and completion results. The results of four scenes are presented column-wise in this figure. The first row is the RGB image of each scene. The second and third rows are the depth map from polarization 0° and polarization 90° settings respectively. The fourth row is the predicted glass boundary using the deep learning method. The fifth row is the extracted glass region, which is obtained from the combination of depth map subtraction and glass boundary. The sixth row shows the fitting of the extracted glass points in the world coordinate. The seventh row is the completed depth map. The original point cloud and the completed point cloud are shown in the last two rows. In the completed point cloud, the completed glass is highlighted in red. The parula color bar applies to the last two rows and the yellow-black color bar applies to the rows of polarization 0°, polarization 90°, and depth completion.

At last, we show an office door glass scene with a larger tilted angle than the first column. In like manner, there is a cabinet with a ball on its top behind the glass. Due to different imaging angles and distances, the position of the glass depth is different in polarization 0°, whereas it will not affect the elimination of glass in polarization 90°. With the same subtraction method, the glass points are extracted and shown in the fifth row. Then the glass door is fitted and completed. From the completion results in the seventh row and the last row, we can see that the glass can be completed well and the box and ball behind the glass can also be reconstructed, reinforcing the idea of seeing and seeing through the reflective surfaces.

It can be seen from the above experiments that the position and proportion of effective glass information are variational for different experiment situations. This is because of the combined effect of the illumination angle and the cleanliness of the reflective surface. In Supplementary Fig. S15, we analyze the influence of different cleanliness of the reflective surface on the detection area. In fact, if the reflective surface perfectly follows the specular reflection model, only an area with the size of the emitter will be fully received. But in common scenes, this kind of situation rarely exists on the reflective surface. Actually, it will be affected by the randomly distributed particles on the surface, expanding the reflection range to a certain extent, which is the so-called off-specular reflection[42,43]. As our experiments show, the PSL 3D sensor can extract enough effective glass information in various scenes.

We also describe the effective measurement angles in Supplementary Fig. S16. We fix the position of the sensor facing the glass and then measure the working range by rotating the sensor horizontally and vertically to measure the proportion of the glass points. We use the number at 0° as a base and the ratio down to 2% as the limit, and the measured working range is ± 45° horizontally and ± 30° vertically. Moreover, the rest of the eight scenes are also illustrated in Supplementary Fig. S14, demonstrating the general applicability of our method.

As for the measurement accuracy, we use the absolute error to evaluate, and the results are shown in Fig. 7. Among them, Fig. 7a-h are the error maps of eight scenes, where Fig. 7a-d correspond to the four scenes in Fig. 6 and Fig. 7e-h correspond to the four scenes in Supplementary Fig. S14. Specifically, we cover the reflective surfaces with thin stickers and conduct depth detection on the scene again, the fitted depth of which is used as a benchmark. We then calculate the absolute error between the depth map of the completed reflective surface and this benchmark. Then we perform boxplot statistics on these eight error maps, and the results are shown in Fig. 7i. The green box includes the median and the upper and lower quartiles of each error map, the gray points are outliers exceeding 1.5 times the interquartile range, and the yellow circle dashed line is the mean value of each error map. From the error analysis, we can see that the errors are maintained at the level of 1mm, demonstrating the precise reconstruction of our method. Note that there are outliers in the error maps.

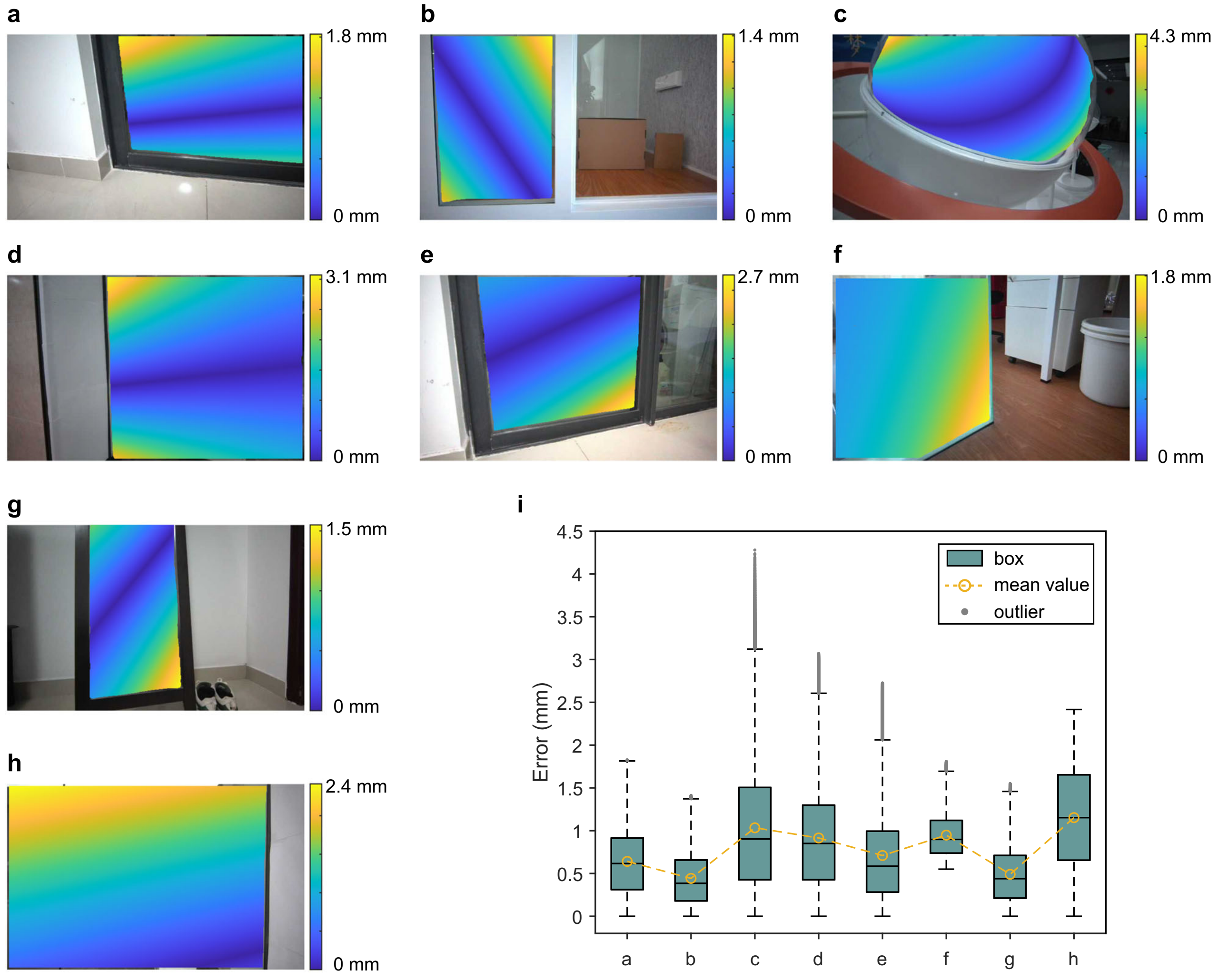

Fig. 7 | Error analysis. (**a**)–(**h**) are the error maps of eight scenes of reflective surfaces, which are the absolute errors between the completed reflective surface and the fitted benchmark. **i** The boxplot graph of these eight error maps. It includes the box, mean value and outliers of each map, showing their statistical properties.

These errors mainly come from the fluctuation of the data, such as the influence of the interreflection, which can be reduced to a much lower level if we include the model to suppress it in the future.

The above experiments of completion are from a single view. If we want to reconstruct the whole scene, we need to employ our method in different views. In Supplementary Fig. S17, we show the reconstruction of a fish tank by employing our method in four sides of the fish tank. In addition, a more complicated scene with a mirror at the back of glass is used to prove the feasibility of our method in Supplementary Fig. S18. The ToF camera is also compared in Supplementary Fig. S19, which once again demonstrates the distinction of the PSL 3D sensor.

Discussion

In summary, we have invented an HCG-VCSEL-based polarization structured light (PSL) 3D sensor and proposed the corresponding imaging methods for 3D reconstruction in a wide type of scenes with highly reflective surfaces.

We first chose a glass scene to demonstrate the ability of the PSL 3D sensor to obtain 3D information of objects behind reflective surfaces. Next, we compare with the stereo vision camera and ToF camera, and show that in the scene with strong reflective noise, through a specific polarization setting, our sensor can eliminate the reflective noise and accurately obtain the 3D image of the scene. Then, we select 8 different scenes to prove that our method can be robustly applied to the detection and completion of the reflective surfaces. Moreover, we show that the effective working range of incident angle is as large as 90° × 60°. In contrast, stereo vision cameras, ToF cameras, and deep learning methods will cause erroneous measurements in these scenarios. At last, we evaluate the absolute errors of our detecting results, and the mean error of each scene maintains at the level as low as 1mm, equivalent to 0.1%. Through these different experiments, we demonstrate that the PSL 3D sensor can realize seeing and seeing through the reflective surfaces.

Currently, the proposed methods are conducted in post-processing, but in fact they can be integrated into the module. We can also modify and optimize the design of the deep learning network, so that they can be transferred to the processing chip of the sensor for real-time detection and completion. On the receiving end, we can also design a unique polarizer array like that in[44], which can be coated on the CMOS camera and receive the signal from different polarizations simultaneously, making our system more compact.

Serving as the eyes of a robot, PSL 3D sensors will find many applications such as service robots and logistics robots. These robots will inevitably encounter the problem of reflective surfaces such as the glass door and floor corner shown in the main text. The PSL 3D sensor can make them handle these cases correctly. Besides, object or scene reconstruction of reflective surfaces, such as office and exhibition hall, will also benefit from this method. Thus, with the special polarization characteristic, the PSL 3D sensor can be further extended to a wide range of indoor and outdoor applications.

Methods

Generation of dot-array structured light

The polarized dot-array structured light is generated directly from the TX. Inside the TX, there is a 940 nm HCG-VCSEL array, a collimating lens with 4.8 mm focal length and a diffractive optical element (DOE). The HCG-VCSEL array and the DOE are mounted at the focal points of both sides of the lens. By designing the bandwidth and airgap of the HCG, the VCSEL sources produce polarized Gaussian beams. These beams propagate to the lens and are collimated to compensate the divergence. Then the collimated light travels to the DOE. Passing through the DOE, one beam can diffract to 11 × 9 orders uniformly. If we properly position the VCSEL on the array, then the far-field distribution of the laser array can be duplicated in 11 × 9 copies, enabling the production of dot-array structured light over 30,000 points. The far-field dot-array distribution is shown in Supplementary Fig. S2.

Postprocessing of the point cloud data and RGB images

The reading and drawing of point clouds are done by the Point Cloud Processing of Computer Vision Toolbox in MATLAB. The fitting uses the fit function of Curve Fitting Toolbox with its robust option set to bisquare. The other processing of the point clouds is also based on MATLAB. The prediction of the reflective surface boundary is based on PyTorch. In addition, the visualization of the point clouds in Fig. 5, Supplementary Fig. S10–S13 and Supplementary Fig. S17 is based on CloudCompare[45].

Calculation of the depth of reflective surfaces

By calculating the spatial displacements of known dot patterns, depth information of reflective surfaces can be obtained according to triangulation. The received dot patterns on the reflective surfaces are shown in Supplementary Fig. S20 and a model of triangulation is also shown to illustrate how depth is calculated.

Data availability

All the data and codes that support the findings of this study are available in Supplementary files and Figshare under accession code https://doi.org/10.6084/m9.figshare.23722416[46].

Code availability

All custom codes are uploaded along with the data, as stated in Data Availability.

References

1. Wolff, L. B. Surface orientation from two camera stereo with polarizers. In: *Optics, Illumination, and Image Sensing for Machine Vision IV* Vol. 1194 (ed. Svetkoff, D. J.) 287–297 (SPIE, 1990). https://doi.org/10.1117/12.969861.
2. Miyazaki, D., Kagesawa, M. & Ikeuchi, K. Transparent surface modeling from a pair of polarization images. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 73–82 (2004).
3. Miyazaki, D. & Ikeuchi, K. Shape estimation of transparent objects by using inverse polarization ray tracing. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 2018–2030 (2007).
4. Xu, X., Qiao, Y. & Qiu, B. Reconstructing the surface of transparent objects by polarized light measurements. *Opt. Express* **25**, 26296–26309 (2017).
5. Rahmann, S. & Canterakis, N. Reconstruction of specular surfaces using polarization imaging. In: *Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* Vol. **1**, I149–I155 (2001).
6. Morel, O., Meriaudeau, F., Stolz, C. & Gorria, P. Polarization imaging applied to 3Dreconstruction of specular metallic surfaces. In: *Machine Vision Applications in Industrial Inspection XIII* Vol. 5679 (eds. Price, J. R. & Meriaudeau, F.) 178–186 (SPIE, 2005). https://doi.org/10.1117/12.586815.
7. Morel, O., Stolz, C., Meriaudeau, F. & Gorria, P. Active lighting applied to three-dimensional reconstruction of specular metallic surfaces by polarization imaging. *Appl. Opt.* **45**, 4062–4068 (2006).
8. Lu, J., Ji, Y., Yu, J. & Ye, J. Mirror surface reconstruction using polarization field. In: *2019 IEEE International Conference on Computational Photography (ICCP)* 1–9 (2019). https://doi.org/10.1109/ICCPHOT.2019.8747335.
9. Kadambi, A., Taamazyan, V., Shi, B. & Raskar, R. Polarized 3d: high-quality depth sensing with polarization cues. In: *2015 IEEE International Conference on Computer Vision (ICCV)* 3370–3378 (2015). https://doi.org/10.1109/ICCV.2015.385.
10. Cui, Z., Gu, J., Shi, B., Tan, P. & Kautz, J. Polarimetric multi-view stereo. In: *2017 IEEE Conference on Computer Vision and Pattern*

Recognition (CVPR)*, 369–378 (2017). https://doi.org/10.1109/CVPR.2017.47.

11. Berger, K., Voorhies, R. & Matthies, L. H. Depth from stereo polarization in specular scenes for urban robotics. In: *2017 IEEE International Conference on Robotics and Automation (ICRA)* 1966–1973 (2017). https://doi.org/10.1109/ICRA.2017.7989227.
12. Zhu, D. & Smith, W. A. P. Depth from a polarisation + rgb stereo pair. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 7578–7587 (2019). https://doi.org/10.1109/CVPR.2019.00777.
13. Tian, X., Liu, R., Wang, Z. & Ma, J. High quality 3d reconstruction based on fusion of polarization imaging and binocular stereo vision. *Inf. Fusion* **77**, 19–28 (2022).
14. Yuan, L., Gao, J., Wang, X. & Cui, H. Research on 3d reconstruction technology based on the fusion of polarization imaging and light field depth information. In: *2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP)* 1792–1797 (2022). https://doi.org/10.1109/ICSP54964.2022.9778534.
15. Thilak, V., Voelz, D. G. & Creusere, C. D. Polarization-based index of refraction and reflection angle estimation for remote sensing applications. *Appl. Opt.* **46**, 7527–7536 (2007).
16. Ngo, T.-T., Nagahara, H. & Taniguchi, R.-I. Shape and light directions from shading and polarization. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2310–2318 (2015). https://doi.org/10.1109/CVPR.2015.7298844.
17. Huynh, C. P., Robles-Kelly, A. & Hancock, E. Shape and refractive index recovery from single-view polarisation images. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1229–1236 (2010). https://doi.org/10.1109/CVPR.2010.5539828.
18. Kong, N., Tai, Y.-W. & Shin, S. Y. High-quality reflection separation using polarized images. *IEEE Trans. Image Process.* **20**, 3393–3405 (2011).
19. Li, Y. & Brown, M. S. Exploiting reflection change for automatic reflection removal. In: *2013 IEEE International Conference on Computer Vision*, 2432–2439 (2013). https://doi.org/10.1109/ICCV.2013.302.
20. Kong, N., Tai, Y.-W. & Shin, S. Y. A physically-based approach to reflection separation: from physical modeling to constrained optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 209–221 (2014).
21. Lei, C. et al. Polarized reflection removal with perfect alignment in the wild. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 1747–1755 (2020). https://doi.org/10.1109/CVPR42600.2020.00182.
22. Wen, S., Zheng, Y. & Lu, F. Polarization guided specular reflection separation. *IEEE Trans. Image Process.* **30**, 7280–7291 (2021).
23. Levin, A., Zomet, A. & Weiss, Y. Separating reflections from a single image using local features. In: *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* I–I (2004). https://doi.org/10.1109/CVPR.2004.1315047.
24. Shih, Y., Krishnan, D., Durand, F. & Freeman, W. T. Reflection removal using ghosting cues. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 3193–3201 (2015). https://doi.org/10.1109/CVPR.2015.7298939.
25. Sun, J., Chang, Y., Jung, C. & Feng, J. Multi-modal reflection removal using convolutional neural networks. *IEEE Signal Process. Lett.* **26**, 1011–1015 (2019).
26. Chang, Y., Jung, C. & Sun, J. Joint reflection removal and depth estimation from a single image. *IEEE Trans. Cybern.* **51**, 5836–5849 (2021).
27. Foster, P., Sun, Z., Park, J. J. & Kuipers, B. Visagge: visible angle grid for glass environments. In: *2013 IEEE International Conference on Robotics and Automation*, 2213–2220 (2013). https://doi.org/10.1109/ICRA.2013.6630875.
28. Wang, X. & Wang, J. Detecting glass in simultaneous localisation and mapping. *Robot. Autonomous Syst.* **88**, 97–103 (2017).
29. Shiina, T. & Wang, Z. An indoor navigation algorithm incorporating representation of quasi-static environmental object and glass surface detection using lrf sensor. In: *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)* 2508–2514 (2017). https://doi.org/10.1109/ROBIO.2017.8324797.
30. Yang, S.-W. & Wang, C.-C. Dealing with laser scanner failure: mirrors and windows. In: *2008 IEEE International Conference on Robotics and Automation*, 3009–3015 (2008). https://doi.org/10.1109/ROBOT.2008.4543667.
31. Wei, H., Li, X., Shi, Y., You, B. & Xu, Y. Multi-sensor fusion glass detection for robot navigation and mapping. In: *2018 WRC Symposium on Advanced Robotics and Automation (WRC SARA)* 184–188 (2018). https://doi.org/10.1109/WRC-SARA.2018.8584213.
32. Yamaguchi, E., Higuchi, H., Yamashita, A. & Asama, H. Glass detection using polarization camera and lrf for slam in environment with glass. In: *2020 21st International Conference on Research and Education in Mechatronics (REM)* 1–6 (2020). https://doi.org/10.1109/REM49740.2020.9313933.
33. Zhang, Y., Ye, M., Manocha, D. & Yang, R. 3D reconstruction in the presence of glass and mirrors by acoustic and visual fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 1785–1798 (2018).
34. Yang, X. et al. Where is my mirror? In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* 8808–8817 (2019). https://doi.org/10.1109/ICCV.2019.00890.
35. Tan, J., Lin, W., Chang, A. X. & Savva, M. Mirror3d: depth refinement for mirror surfaces. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 15985–15994 (2021). https://doi.org/10.1109/CVPR46437.2021.01573.
36. Mei, H. et al. Don't hit me! glass detection in real-world scenes. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 3684–3693 (2020). https://doi.org/10.1109/CVPR42600.2020.00374.
37. He, H. et al. Enhanced boundary learning for glass-like object segmentation. In: *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* 15839–15848 (2021). https://doi.org/10.1109/ICCV48922.2021.01556.
38. Huang, X. et al. Seeing and seeing through the reflection surface: 3d imagingusing polarization structured light camera. In: *Frontiers in Optics/Laser Science* JW5B.49 (Optica Publishing Group, 2022). https://doi.org/10.1364/FIO.2022.JW5B.49.
39. Hecht, E. & Zajac, A. *Optics*. (Addison-Wesley Publishing Company, 1974).
40. Huang, M. C., Ye, Z. & Chang-Hasnain, C. J. A surface-emitting laser incorporating a high-index-contrast subwavelength grating. *Nat. Photonics* **1**, 119–122 (2007).
41. Huang, M. C., Ye, Z. & Chang-Hasnain, C. J. Polarization mode control in high contrast subwavelength grating vcsel. In: *2008 Conference on Lasers and Electro-Optics and 2008 Conference on Quantum Electronics and Laser Science*, 1–2 (2008). https://doi.org/10.1109/CLEO.2008.4551186.
42. Torrance, K. E. & Sparrow, E. M. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.* **57**, 1105–1114 (1967).
43. Wolff, L. B., Nayar, S. K. & Oren, M. Improved diffuse reflection models for computer vision. *Int. J. Comput. Vis.* **30**, 55–71 (1998).
44. Millerd, J. et al. Pixelated phase-mask dynamic interferometers. In: *Fringe 2005* (ed. Osten, W.) 640–647 (Springer, 2006). https://doi.org/10.1007/3-540-29303-5_86.
45. CloudCompare (version 2.10.alpha) [GPL software] (2023). Retrieved from http://www.cloudcompare.org/.
46. Huang, X. et al. PSL_3DSensor. Figshare (2023). https://doi.org/10.6084/m9.figshare.23722416 (2023).

Acknowledgements

This research was partially funded by the Shenzhen Science and Technology Programs No. KQTD20200820113053102 and No. ZDSYS20220325163600001; the Guangdong Major Talent Introduction Project (No. 2021ZT09X328); the National Key Research and Development Program of China (2022YFB2802803); the Natural Science Foundation of China Project (No. 61925104, No. 62031011) and Major Key Project of PCL.

Author contributions

C.J.C.-H. conceived and supervised the project. X.H. wrote the processing codes. X.H. and C.W. performed the experiments and analyzed the data. X.X., B.W., S.Z., and C.Y. designed, assembled, and tested the PSL 3D sensor. C.S. and J.W. designed and managed the fabrication of HCG VCSELs. N.C. and S.Y. provided guidance and advice on experiments. X.H., C.W., and C.J.C.-H. interpreted the data and wrote the paper, with input from all the other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-023-42678-5.

Correspondence and requests for materials should be addressed to Connie J. Chang-Hasnain.

Peer review information *Nature Communications* thanks Yongqiang Zhao and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Supplementary Information: Polarization Structured Light 3D Depth Image Sensor for Scenes with Reflective Surfaces

Xuanlun Huang[a,b,#], Chenyang Wu[a,b,#], Xiaolan Xu[b], Baishun Wang[b], Sui Zhang[b], Chihchiang Shen[b], Chiennan Yu[b], Jiaxing Wang[b], Nan Chi[a,c], Shaohua Yu[a,c], Connie J. Chang-Hasnain[a,b,*]

[a]*School of Information Science and Technology, Fudan University, Shanghai 200433, China*
[b]*Berxel Photonics Co. Ltd., Shenzhen 518071, China*
[c]*Peng Cheng Laboratory, Shenzhen 518055, China*

---

[*] Corresponding author

Email address: `connie.chang@berxel.com` (Connie J. Chang-Hasnain)

[#] These authors contributed equally to this work

I. Polarization characteristics of the HCG-VCSEL array.

As mentioned in the main text, HCG VCSELs have exhibited a very high polarization selection ratio independent of operating temperature or drive conditions, which can be even up to 30 dB [1, 2] with precise fabrication. Polarization selection ratio is defined by the ratio of power in TE and TM polarization. It can also be characterized by the orthogonal polarization suppression ratio (OPSR) curve, where

$$\text{OPSR} = 10\log(\frac{P_{\text{TE}}}{P_{\text{TM}}}) \tag{1}$$

Here, we compare the OPSR results of the HCG VCSEL array and the single HCG VCSEL in the same production patch in order to show that the VCSEL array possesses similar polarization characteristics as the single VCSEL. In Fig. S1a, we show the OPSR and power curve as the function of working current, where solid lines are the results of HCG-VCSEL array and dot lines are the results of DBR-VCSEL array. We can see that as the current increases, the OPSR of HCG-VCSEL array reaches the plateau at around 17 dB, meaning the output light of the HCG-VCSEL array possesses polarization. However, as for the DBR-VCSEL array, it keeps around 0 dB, which has no polarization. From the power curves of two types of arrays, we can see that the top HCG does not affect the output power much, making the HCG-VCSEL array suitable for the TX source. In Fig. S1b, we also show the OPSR curve and power curve of a single HCG VCSEL in the same fabrication patch. We can see that the OPSRs of the VCSEL array and the single VCSEL are close and the power obeys the multiple relation, which means the array almost has no impact on the original characteristics. In this patch, because of the fabrication error, the OPSR is not as high as that of [1, 2]. But this polarization ratio is sufficient to prove its superiority in the application of 3D imaging, as will be demonstrated in the main text. In addition, in order to control variables, we set the current of the PSL 3D sensor to be 1.5 A at each experiment, which means the output power is around 2.2 W. The driving condition is pulsed current with 33 milliseconds in period and 0.03 to 0.13 in duty cycle, namely 1 millisecond to 4.3 milliseconds in pulse width.

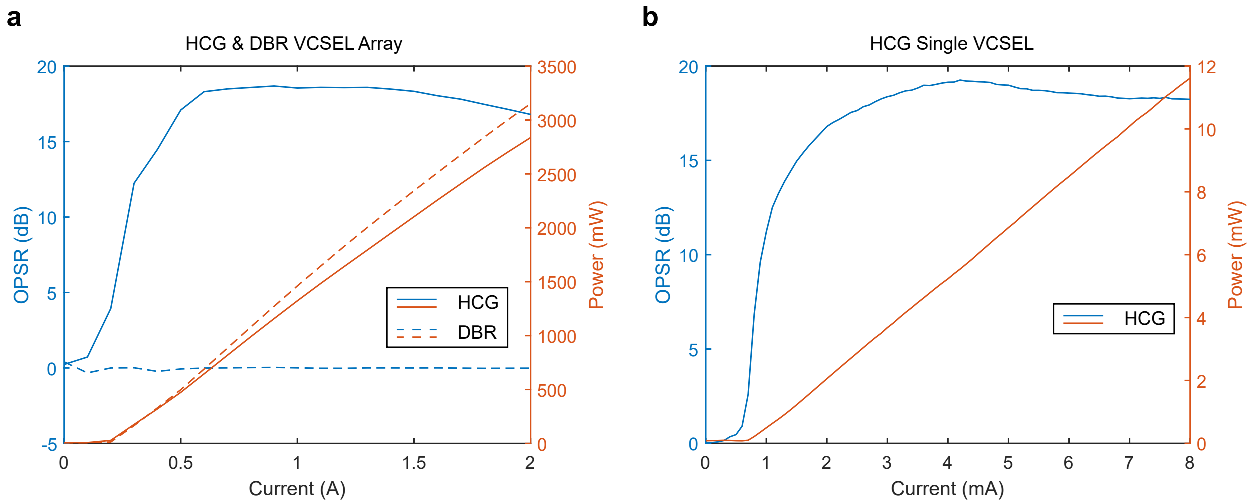

Figure S1: **OPSR and power curve of VCSELs.** **(a)** OPSR and power curve of HCG and DBR VCSEL arrays. **(b)** OPSR and power curve of a single HCG VCSEL.

II. Far-field of the dot-array structured light

As described in the Methods of the main text, the PSL 3D sensor produces dot-array structured light with over $30,000$ points, whose far-field distribution is shown in Fig. S2. This is achieved by the duplication of the VCSEL source array. As seen from Fig. S2, each block is a far-field projection of a VCSEL array, containing 364 randomly distributed points. Then this projection is duplicated to 11×9 diffraction orders uniformly by the diffraction optical element, enabling the generation of over $30,000$ dots in total. Note that the dot-array structured light possesses high polarization selection ratio. It means that these far-field points maintain the same polarized direction as the VCSEL source. If we put a polarizer whose polarized direction is orthogonal to the VCSEL on top of the TX, the far-field dots can be eliminated completely.

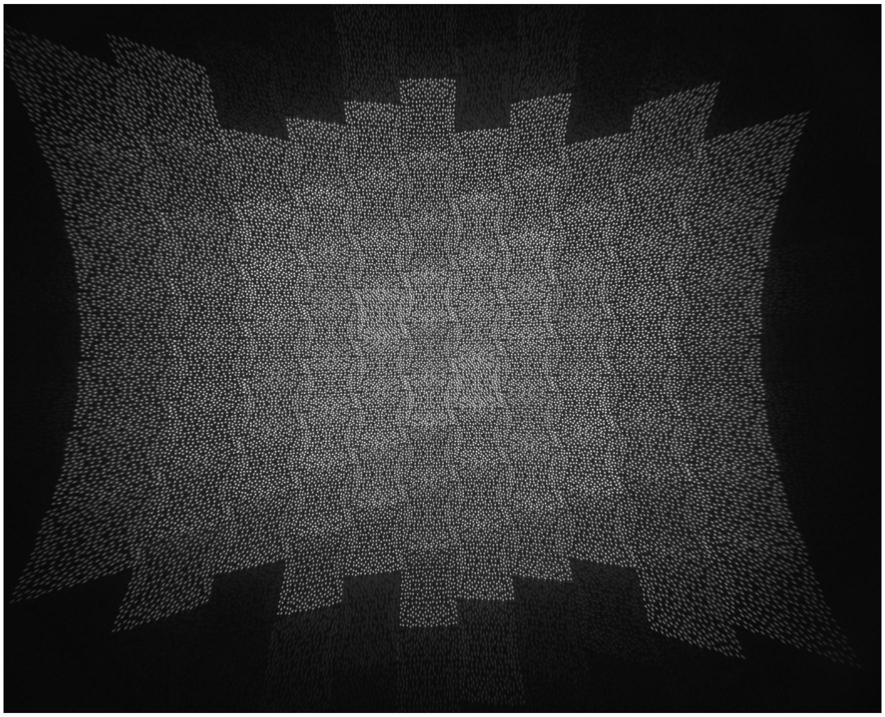

Figure S2: **Dot-array structured light.** This far-field distribution is captured with an industrial infrared camera at 40 cm.

III. Depth images comparison in a scene without glass

According to the Fresnel equation, the light reflected from the reflective surfaces maintains the same polarized direction as the incident structured light, while those from the diffuse surfaces have omnidirectional directions. This property can be best used to differentiate two types of reflecting surfaces. In Fig. 2, 4 and 6 of the main text, we have demonstrated that for a scene with the reflective surfaces, there is a sharp contrast between the depth channels from two polarization settings, which is an important cue for the determination of the reflective surfaces. Here, we provide the depth images in a scene without glass to support this method. In this scene (Fig. S3a), a cardboard box with its front surface removed is placed in front of the PSL 3D sensor. Inside the box, there is a box of gloves and a metal breadboard. We obtain two depth images (Fig. S3b and S3c) from two polarization settings, i.e., polarization 0° and polarization 90°. In Fig. S3b, we show the depth image in polarization 0°, which is obtained by setting both the TX and RX of the PSL 3D sensor in TM polarization. As for Fig. S3c, this is the depth image in polarization 90°, which is obtained by setting the TX in TM polarization and the RX in TE polarization. From the results, we can see that for a scene without glass, there is no sharp contrast between two polarization settings in the depth channel.

Figure S3: **Depth images comparison in a scene without glass.** **(a)** Scene of a cardboard box with its front surface removed. Inside the box, there is a box of gloves and a metal breadboard. **(b)** Depth image from the PSL 3D sensor, obtained by setting both the TX and RX in TM polarization. **(c)** Depth image from the PSL 3D sensor, obtained by setting the TX in TM and the RX in TE. The yellow-black color bar applies to both **(b)** and **(c)**.

IV. Fitting in the world coordinate

As shown in the step 2 of Fig. 2 in the main text, the extracted glass points are transformed to the world coordinate for fitting. Here we explain why it needs to fit in the world coordinate. We take the balcony glass #2 scene in the first column of Fig. S14 as an example. In Fig. S4a, we show the depth point cloud in x-z view. Because the depth point cloud originates from the depth image, it uses pixel as x unit and millimeter as z unit. In this subfigure, the glass points, which are extracted by the subtraction method as illustrated in step 1 of Fig. 2, are highlighted in red and the rest points are shown in grey. We can see that the glass frame (encircled by the white dashed line) is curved in the depth coordinate. This is because the internal parameters $\{f_x, f_y, c_x, c_x\}$ of the infrared camera will transform the point cloud coordinate to make it appear properly in the depth image. If we fit the glass plane based on the extracted glass points directly in the depth coordinate, it will deviate seriously from the glass frame, which is shown in Fig. S4b. Thus, the extracted glass points need to transform to the world coordinate. In Fig. S4c, we show the transformed point cloud in the world coordinate, where the glass points are also highlighted in red. In the world coordinate, both the x and z axis are in unit of meter. We can see that the glass frame is straight in this subfigure, which accords with the physical intuition. Then, the fitting can be employed under this coordinate and the result in Fig. S4d reveals that the red fitted plane matches well with the straight glass frame. At last, in Fig. S4e, the fitted and completed point cloud data is converted back to the depth coordinate, showing the correct result of our method.

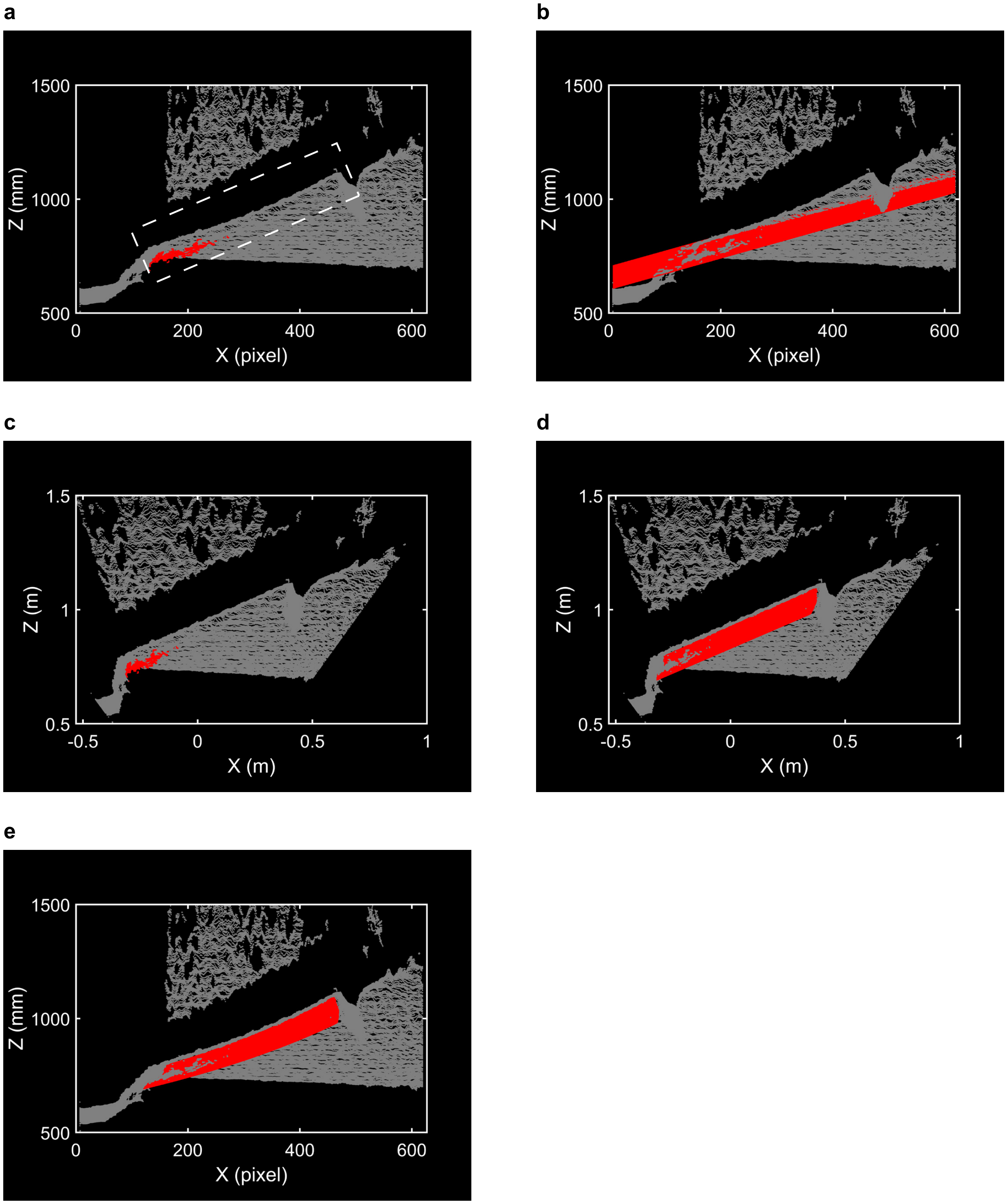

Figure S4: **Point cloud results in different coordinates.** **(a)** The original point cloud in the depth coordinate. The straight glass frame is curved in this coordinate. **(b)** Fitting directly in the depth coordinate. The fitted glass plane is shown in red. **(c)** Original point cloud in the world coordinate. **(d)** Fitting in the world coordinate, where the red fitted plane accords well to the straight glass frame. **(e)** The fitted point cloud that is transformed back to the depth coordinate. Note that in the depth coordinate, x axis is in pixel unit and z axis is in millimeter unit while x and z axes are both in meter unit in the world coordinate.

V. Analysis on the influence factors of seeing through the reflective surface

In this section, we analyze the other three influence factors of seeing through the reflective surface. In Fig. S5, we show the results of PSL 3D sensor incident angle. In this experiment, we test three detecting angles, i.e., −30°, 0°, 30° and their pictures of scene are shown in the first column of Fig. S5a, S5b, S5c. For each angle, the objects behind the glass are fixed to 0.5 m away and the distance from the PSL 3D sensor to the middle of glass is fixed to 0.8 m. Due to different incident angles, the position of the detected glass depth is also different. At −30°, the glass appears on the left of the depth map and does not overlap with the ball and cabinet, as shown in polarization 0° (Pol 0°) of Fig. S5a. When the sensor is set to polarization 90° (Pol 90°), the glass depth on the left can be eliminated, leaving only the ball and cabinet. When the detection angle is changed to 0°, the glass depth appears around the ball at polarization 0°, and similarly this glass information is eliminated at polarization 90°. At 30°, the detection and elimination of glass appears on the right side of the depth map. We can see that the different incident angle does not affect the function of seeing through the glass.

In Fig. S6, we show the results of different distances of objects behind glass. In this experiment, the distance from the sensor to the glass is set to 0.8 m and the detecting angle is 0°. Then we change the distances of objects behind glass from 0.3 m to 0.7 m, and their results are shown in Fig. S6a, S6b, S6c, respectively. When the ball is close to the glass, more depth information about the ball is detected in the middle (white dot line in polarization 0° of Fig. S6a). As the distance increases to 0.5 m, the area of the ball decreases, and instead the glass in the middle is detected (white dot line in Fig. S6b). At 0.7 m, the middle is almost all the glass information (white dot line in Fig. S6c). At these three different distances, we can see that at polarization 90°, the glass depth all can be eliminated. In the next experiment, we also discuss the effect of the density of objects behind glass.

Fig. S7 is the result of different densities of objects behind the glass. In this experiment, the distance of the PSL 3D sensor is 0.8 m, the distance of the rear object is 0.5 m, and the detecting angle is 0°. When the density is low (Fig. S7a), the glass depth occupies most of the middle in polarization 0°. When the density is in medium level (Fig. S7b), because the reflection of the rear object is stronger, part of the glass depth is replaced by the rear ball and cabinet. When the density becomes high (Fig. S7c), only the glass between the ball and the box can be detected. For all these three different levels, the part of glass can still be removed in polarization 90°, leaving a clear depth map of objects behind.

Although the density of objects behind affects the glass depth that we can detect, we can get enough depth information of glass by varying the distance and angle as demonstrated in the experiments of first three influence factors. Moreover, different experiments have shown that the glass depth can be eliminated at polarization 90°. This benefits from our specially designed HCG-VCSEL array, which allows the depth information belonging to the reflective surface to retain its original polarization characteristics.

Meanwhile, we also compare the sensor based on DBR VCSEL in Fig. S8. It can be seen that the glass depth area in polarization 0° and polarization 90° is similar, and the function of elimination cannot be achieved in polarization 90°. In addition, in the Supplementary Movie 1, we show the effect of 360 degrees rotation of the RX polarizer, and we can see the glass depth varies accordingly, which once again proves the ability of the PSL 3D sensor.

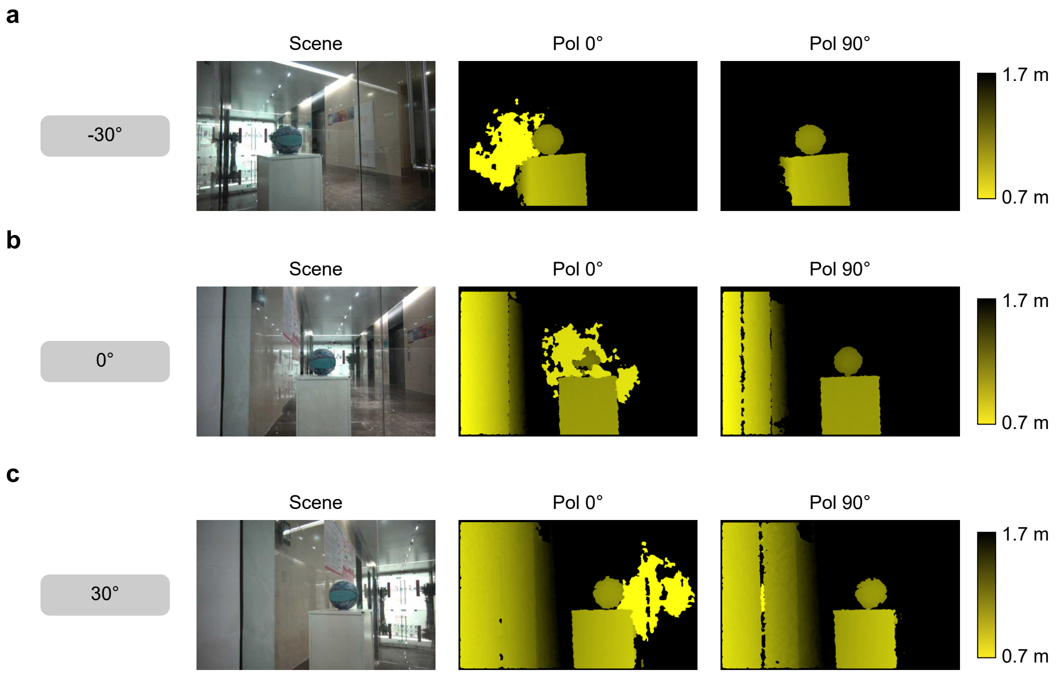

Figure S5: **Results of different incident angles of the PSL 3D sensor.** **(a)**-**(c)** are results for angles of −30°, 0°, 30°, respectively. The pictures of scenes are shown in the first column. The column of Pol 0° is the depth map when TX and RX are in the same polarization. The column of Pol 90° is the depth map when TX and RX are in the orthogonal polarization. The yellow-black colorbars apply to both second and third columns.

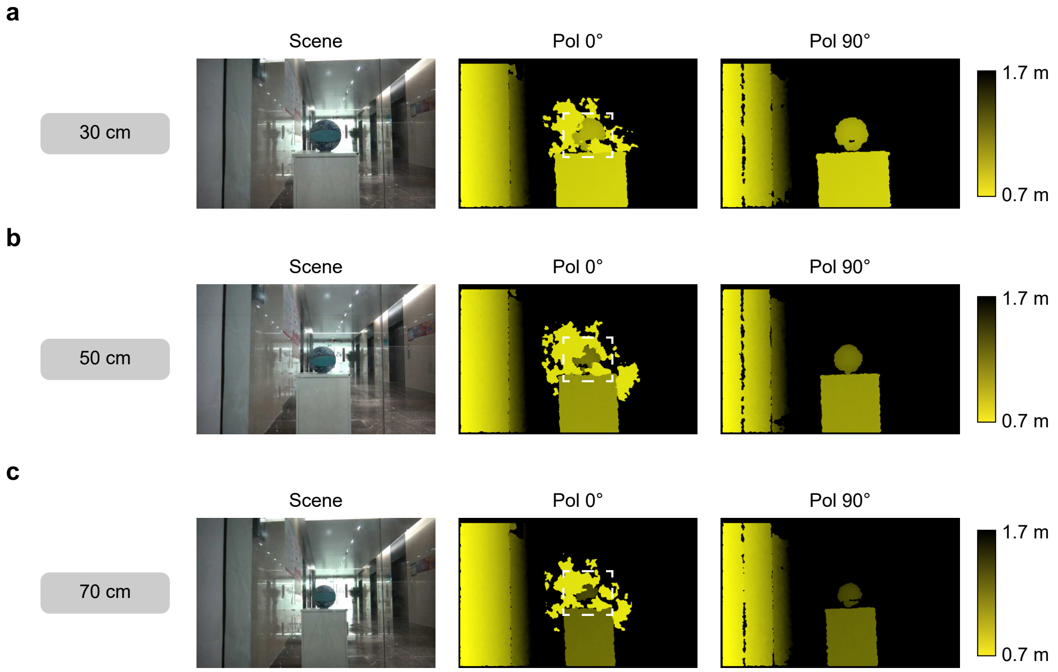

Figure S6: **Results of different distances of object behind glass.** **(a)**-**(c)** are results for distances of 30 cm, 50 cm, 70 cm, respectively. The pictures of scenes are shown in the first column. The column of Pol 0° is the depth map when TX and RX are in the same polarization. The column of Pol 90° is the depth map when TX and RX are in the orthogonal polarization. The yellow-black colorbars apply to both second and third columns.

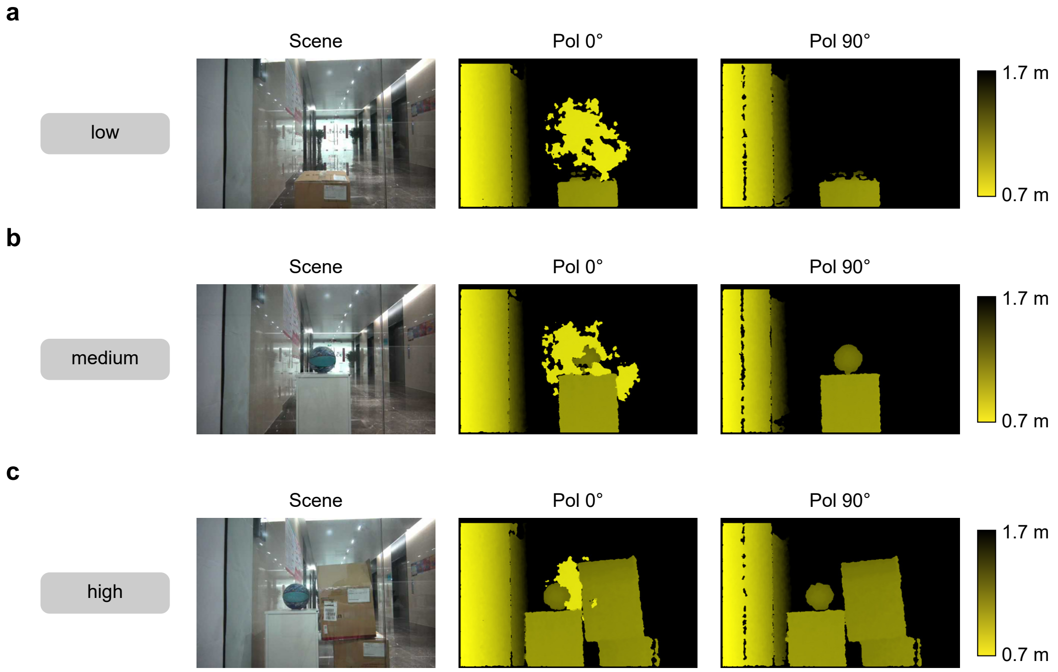

Figure S7: **Results of different densities of object behind glass.** **(a)**-**(c)** are results for density levels of small, medium and large, respectively. The pictures of scenes are shown in the first column. The column of Pol 0° is the depth map when TX and RX are in the same polarization. The column of Pol 90° is the depth map when TX and RX are in the orthogonal polarization. The yellow-black colorbars apply to both second and third columns.

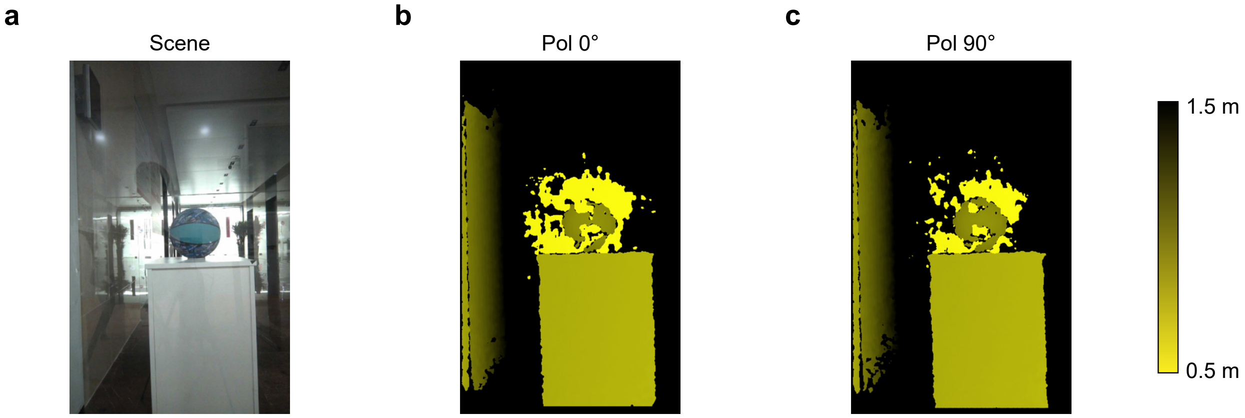

Figure S8: **Results of DBR-VCSEL-based sensor.** **(a)** The picture of the scene. **(b)** Depth map from Pol 0°. **(c)** Depth map from Pol 90°. The yellow-black colorbar apply to both **(b)** and **(c)**.

VI. Analysis on seeing against reflective noise.

In Fig. S9, we show another comparative experiment for detecting objects inside glass from the outside. In this scene, different from the normal incidence in Fig. 3b of the main text, we detect the glass scene in an angle (Fig. S9a). On the glass, there are reflections of external trees and hand. Inside the glass, there is a person to be detected. We can see that the reflective noise on the glass overlaps with the position of the person to be detected, which causes the stereo vision camera to fail to obtain the correct depth of the person. As we can see from Fig. S9b, the depth of the external tree and hand is incorrectly displayed on the depth map and blocks the depth information of the person inside, leaving only the leg in the depth map. But for the PSL 3D sensor, the correct depth information can be obtained through the polarization combination of TX and RX (Fig. S9c).

In the second case of seeing against reflective noise experiment in the main text, we compare the stereo vision camera and PSL 3D sensor in the scene of wall corner. In this case, although there are effects of reflection and multipath noise, PSL 3D sensor can obtain complete and correct corner depth and point cloud through the combination of TX and RX P polarization. In the following experiments, we continue to analyze the effects of TX and RX polarization combination in the same scenario.

In Fig. S10, we show the experimental results at the height of 24 cm, where Fig. S10a, S10c and S10e are the results of TX y polarization (pol y), and Fig. S10b, S10d and S10f are the results of TX x polarization (pol x). In Fig. S10a, we see that when RX has no polarization, the depth map will have holes in the corner due to the influence of reflective noise. Compared with the results of stereo vision camera in the main text, although there are holes, the angle between the floor and the wall remains at a right angle in the point cloud. When we use TX pol x (Fig. S10b), since both signal and noise are x-polarized, the holes in the red dot line are larger than in Fig. S10a. Since the reflective noise is in the x direction, when we use the combination of TX pol y and RX pol y, as shown in Fig. S10c, we can complete the missing part in Fig. S10a and obtain the complete and correct depth information. Similarly, even if the TX and RX directions are orthogonal (Fig. S10d), most of the noise can be filtered out when we use RX pol y, making the missing part in Fig. S10b complete. But when RX is in the x direction, it allows a lot of noise to go through, so the depth map at the corner is missing again (Fig. S10e). These holes are more severe when both TX and RX are in the x direction (Fig. S10f). Through the above comparisons, we can see that in the case of reflection and multipath noise, the combination of TX pol y and RX pol y can filter the noise and enable a clear 3D information acquisition of the scene. Although TX is not always orthogonal to the noise, we can still determine the direction of noise and use RX to increase the SNR like demonstrated in Fig. S10d. The polarization characteristic of PSL 3D sensor make it superior in the case of detecting against the reflective noise.

Similar to Fig. S10, we show the results at the height of 40 cm in Fig. S11. Compared with those of Fig. S10, the impact of reflective noise is not so big at 40 cm, as shown in the dashed box in Fig. S11a. However, for TX pol x, there is still noise in the middle of the depth map (red dot line in Fig. S11b). Therefore, the combination of TX pol y and RX pol y in Fig. S11c still provides the best depth map information. Similarly, even if TX and RX are orthogonal, the RX pol y can still filter the noise in the middle (Fig. S11d). For Fig. S11e and S11f, because the polarization direction of RX and noise are both x, the depth information is missing and Fig. S11f is the worst situation because TX and RX are both pol x.

In Fig. S12, we show the results at the height of 56 cm. At this height, the reflective noise has little effect. As shown in Fig. S12a, with only TX pol y, the depth in the red dot line is already complete. For TX pol x, there will still be a small effect of reflective noise in the middle (blue dot line in Fig. S12b). As for the combination of TX pol y and RX pol y, it can get a complete depth map and point cloud as Fig. S12a, and will not be affected by the halving of the polarization power at the RX. In the setting of Fig. S12d, although the noise in the middle of the blue box is eliminated, there is some missing in the edge due to the orthogonal polarization (shown in the red box). Similarly, Fig. S12e will have this problem. As for Fig. S12f, it has the worst depth result because both TX and RX are set in the same direction as the noise. Therefore, in this scenario, the polarization combinations in Fig. S12a-S12c all work for depth detection. In future work, we can design a RX that receives both pol x and pol y at the same time, solving the problems in Fig. S12d and S12e.

In addition, we also test the ToF sensor in this scenario. As shown in Fig. S13, ToF sensor is unable to obtain the correct depth information at all heights. In detail, we can see that the floor part in the pointcloud (pointed by red dot box and solid arrow) penetrates into the plane because of the impact of noise while the PSL 3D sensor works well at all heights with its polarization settings.

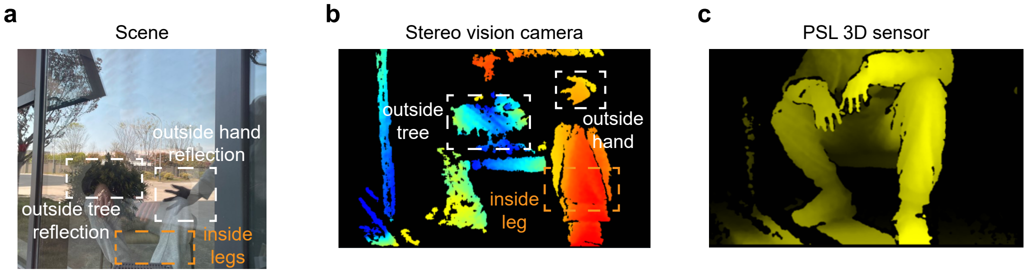

Figure S9: **Comparison on an outdoor glass scene.** **(a)**Scene of the experiment. **(b)** Depth map from the stereo vision camera. **(c)** Depth map from the PSL 3D sensor.

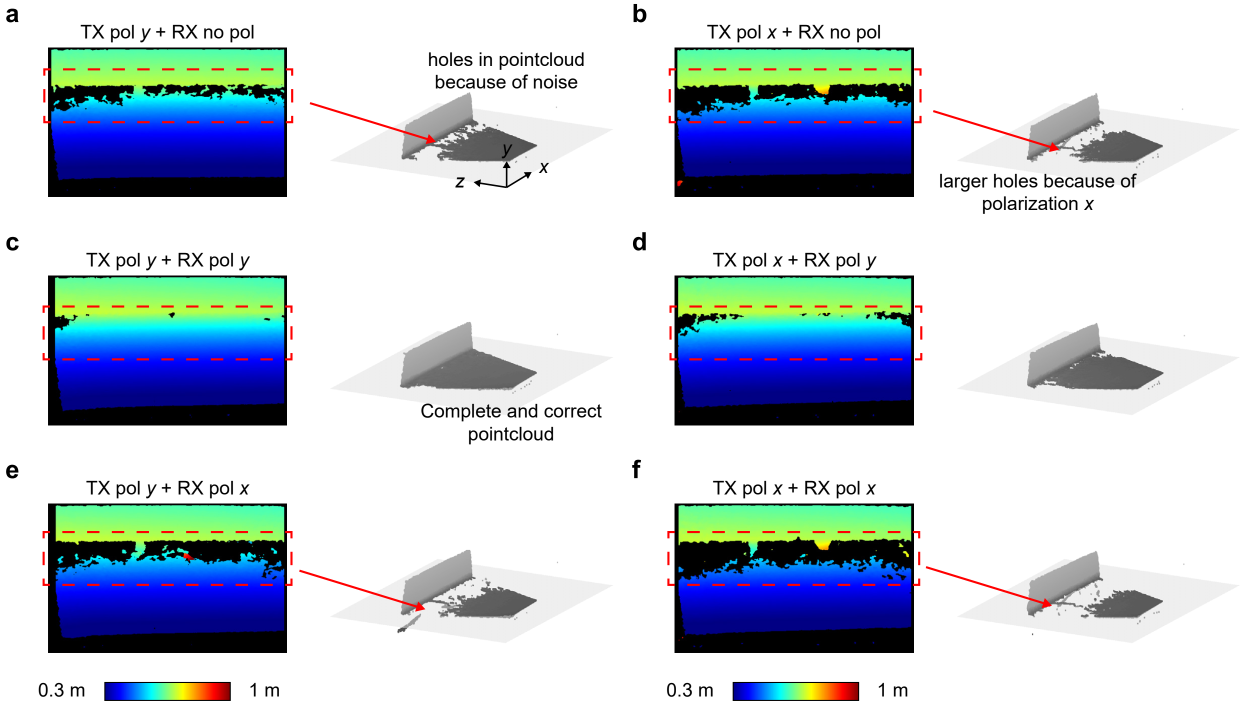

Figure S10: **Comparison at the height of 24 cm.** **(a)**-**(c)** are results of TX pol y and **(b)**-**(f)** are results of TX pol x. RX has no polarization in **(a)** and **(b)**, RX is pol y in **(c)** and **(d)**. RX is pol x in **(e)** and **(f)**.

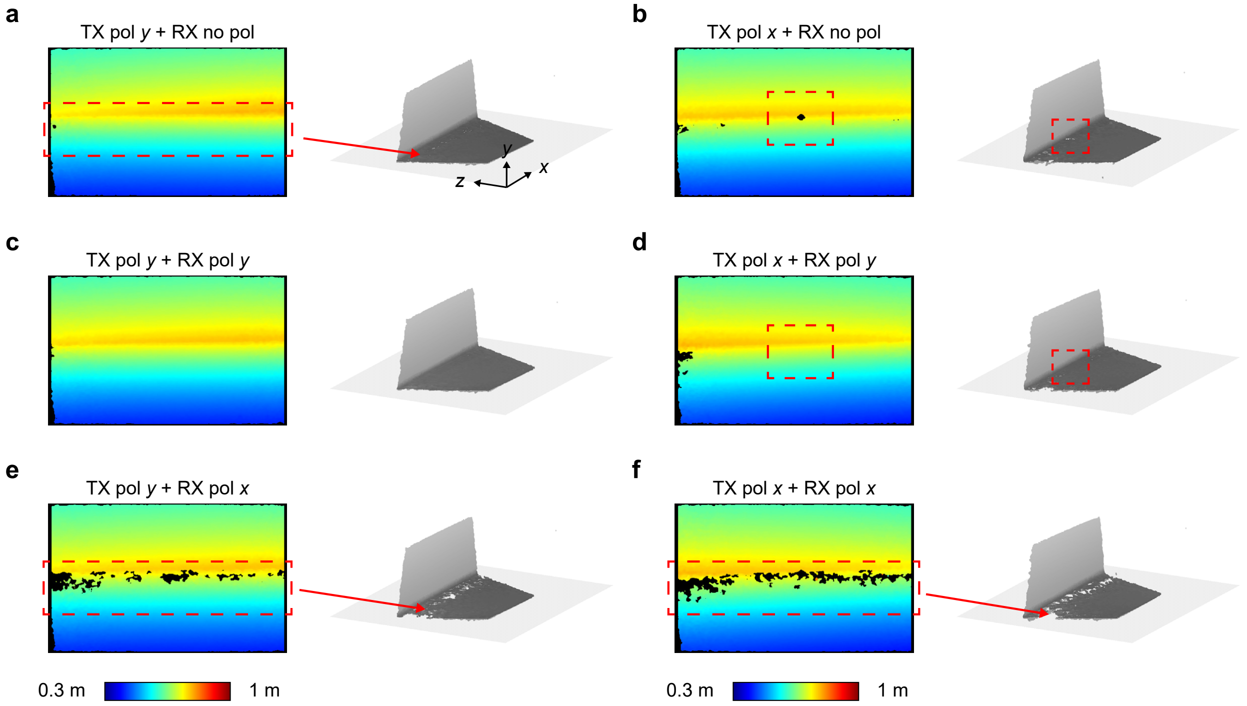

Figure S11: **Comparison at the height of 40 cm.** **(a)**-**(c)** are results of TX pol *y* and **(b)**-**(f)** are results of TX pol *x*. RX has no polarization in **(a)** and **(b)**, RX is pol *y* in **(c)** and **(d)**. RX is pol *x* in **(e)** and **(f)**.

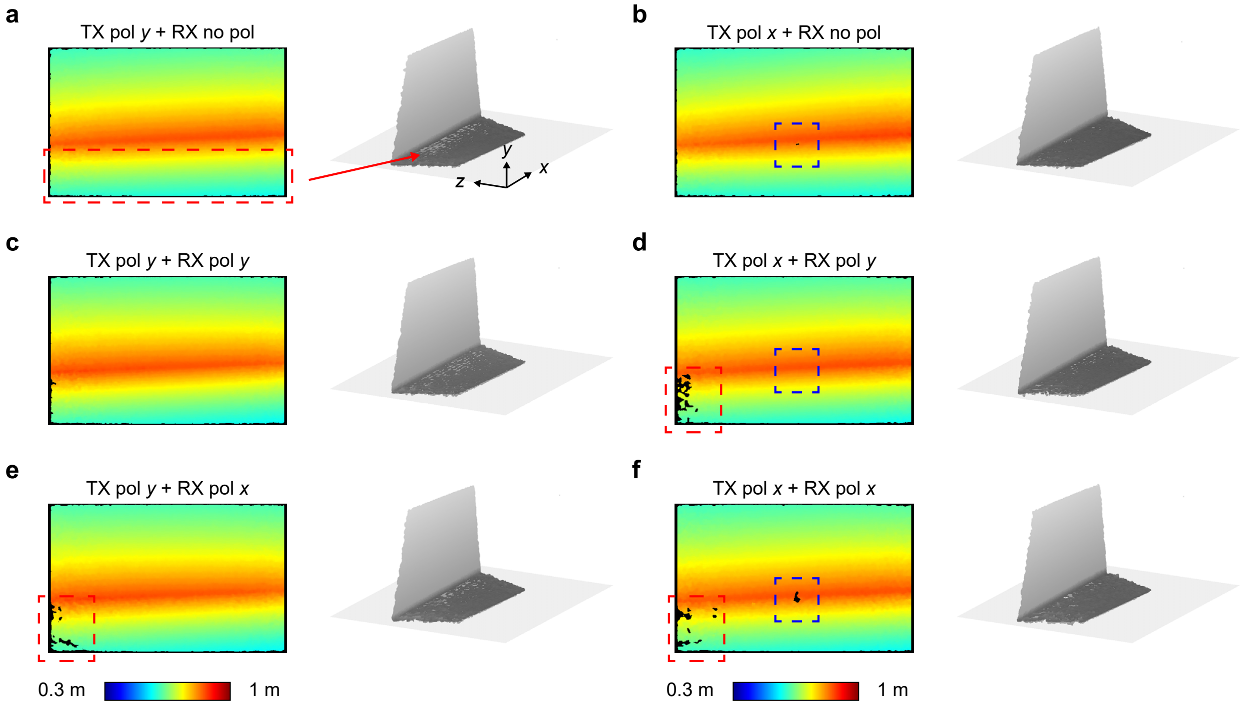

Figure S12: **Comparison at the height of 56 cm.** **(a)**-**(c)** are results of TX pol *y* and **(b)**-**(f)** are results of TX pol *x*.

RX has no polarization in **(a)** and **(b)**, RX is pol *y* in **(c)** and **(d)**. RX is pol *x* in **(e)** and **(f)**.

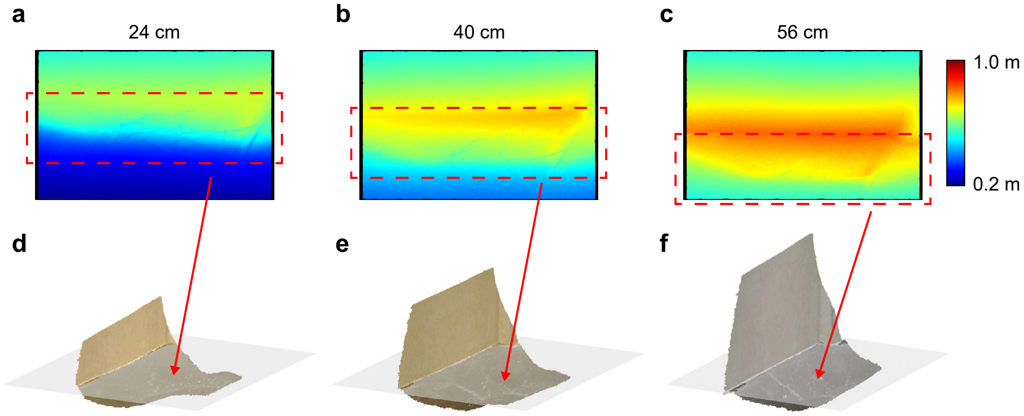

Figure S13: **Detecting results from ToF sensor.** **(a)**-**(c)** are depth maps at different heights. **(d)**-**(f)** are their corresponding point clouds. The jet colorbar applies to **(a)**-**(c)**.

VII. Detection and completion results

In this section, we display the other four scenes (Fig. S14). They include the balcony glass #2, the fish tank glass, the mirror and the front door glass. As for the balcony glass #2 scene, it is also a glass door which has reflections of the interior furniture and the glass inclines to the right. With polarization 0° and polarization 90° settings, we obtain two depth images of the scene. In polarization 0°, partial of the glass and the objects around it can be reconstructed. In polarization 90°, the glass part is eliminated, leaving those from diffuse reflection. Thus, in the same way, using the subtraction of two depth images and the glass boundary predicted from the RGB image of the scene, we are able to extract the glass region, which is highlighted in red in the fifth row. The extracted glass points are then used to fit and complete the reflective surface. The completed depth images are shown in the seventh row. Because the glass on the right is not on the same plane as that on the left, we only select the left glass for completion manually. The final comparison between the original point cloud and the completed point cloud is shown in the last two rows, where the new completed glass is colored in red. Similar to the balcony glass #1 scene in the main text, the new completed glass also matches well to the glass frame.

In the fish tank scene, the glass tank with water inside is placed at the left of the field of view, while on the right side there is a bucket and a cabinet. In polarization 0°, we can see that the upper part of glass is detected. In polarization 90°, the upper glass depth is eliminated due to the orthogonal polarization. Again, we combine with the glass boundary in the fourth row to extract the glass region, which is shown in the fifth row. Similarly, this extracted depth information is used for the following fitting and completion. As shown in the row of depth completion and the last row, the front face of the fish tank can be reconstructed precisely.

We also test our method in the mirror scene. In this situation, the PSL 3D sensor is placed in front of the mirror. On the mirror, there are reflections of the sensor and other indoor objects. Likewise, two depth images are obtained from two polarization settings. We can see that the depth of the mirror surface is also eliminated in polarization 90°. With the subtraction from two depth images and the mirror boundary, we acquire the mirror region in the fifth row. Then we complete the mirror inside the boundary and get the completion result. Because we cannot see through a mirror physically, the measured depth information which is caused by the reflection of the mirror is removed, as shown in the comparison in the last two rows. For a surface with high reflectance like mirror, the structured light will first reflect to the objects in front of this surface and next propagate back to the RX in the same way, leading to the acquisition of the mirror images. But this will not affect the elimination of the mirror between two depth images. Therefore, our method is still able to detect and complete the mirror, making the correct reconstruction of the scene. Furthermore, it is different from the situation encountered by the stereo vision camera. For a stereo vision camera, on the one hand, it is unable to obtain and determine the depth of the reflective surface. On the other hand, whenever there is a reflection in either a mirror or glass, it will mistakenly consider the reflected image as a real object behind the surface.

In the last column, we show the front door glass with reflecting image of two persons, confirming the advantage of PSL 3D sensor in the glass scene once again. In this situation, two persons and the tripod are reflected images while the vase is a real object behind the glass. In polarization 0°, only the vase and partial glass are obtained. When it changes to polarization 90°, the glass part is eliminated, leaving the depth of the vase. Then subtraction is employed between two depth images to extract the glass region, combined with the glass boundary. Next, the glass region is fitted and used for completion. From the final result, we can see that the glass can be completed well and the vase behind the glass can also be reconstructed.

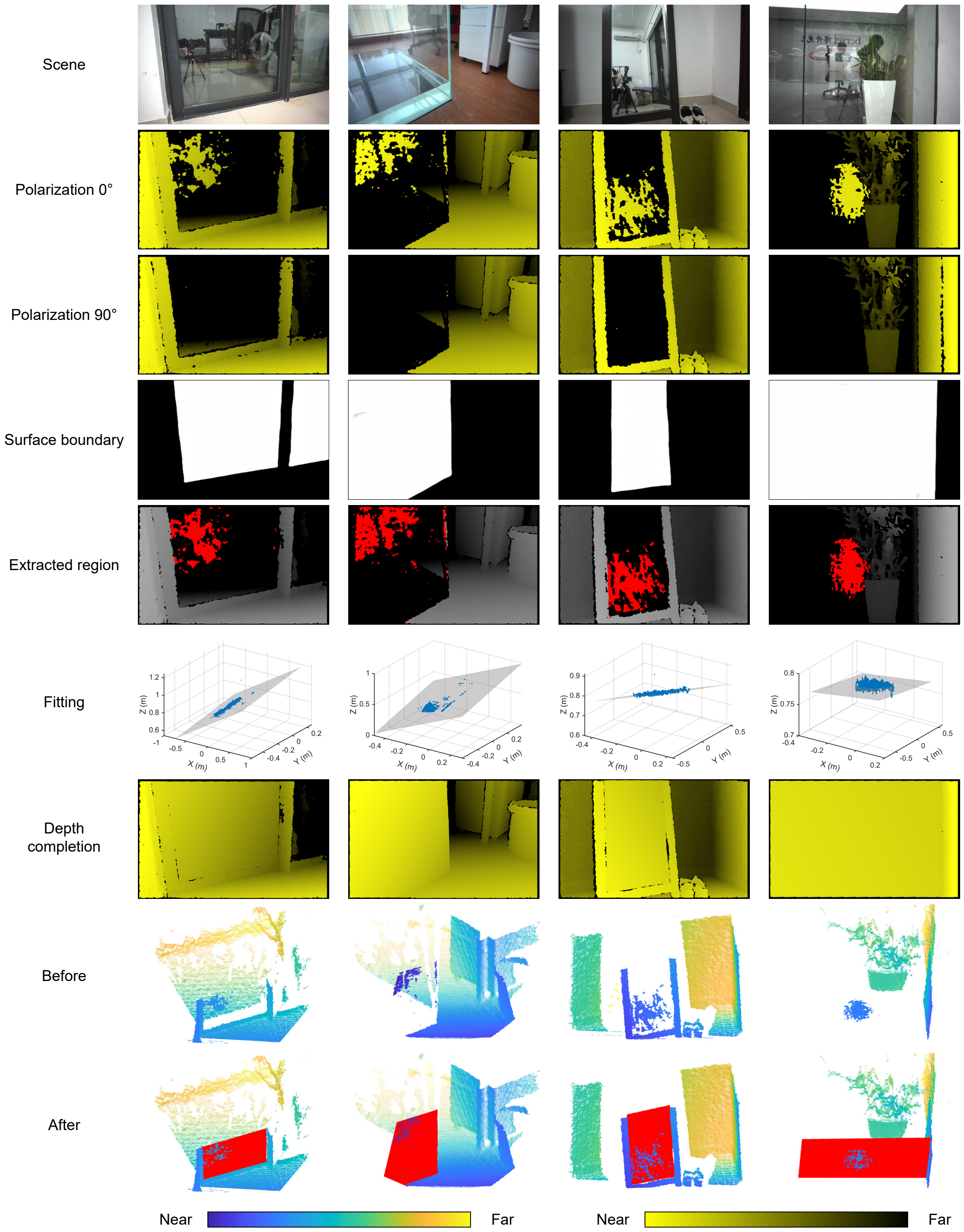

Figure S14: **Detection and completion results.** The results of the other four scenes are presented column-wise in this figure. The first row is the RGB image of each scene. The second and third row are the depth map from polarization 0°

and polarization 90° settings respectively. The fourth row is the predicted boundary using deep learning method. The fifth row is the extracted reflective surface region, which is obtained from the combination of depth image subtraction and predicted boundary. The sixth row shows the fitting of the extracted points in the world coordinate. The seventh row is the completed depth image. The original point cloud and the completed point cloud are shown in the last two rows. In the completed point cloud, the completed reflective surface is highlighted in red. The parula colorbar applies to the last two rows and the yellow-black colorbar applies to the rows of polarization 0°, polarization 90° and depth completion.

VIII. Detection area analysis

In the experiment, we find that the area of the reflective surface detected by the PSL 3D sensor will be affected by the cleanliness of the surface. In this section, we analyse the relationship between the level of cleanliness on the surface and the amount of depth information that can be extracted. As shown in Fig. S15a, we place the PSL 3D sensor at a distance of 0.8m directly facing the mirror. There are different kinds of natural dirt on the mirror surface, such as dust, water stain or fingerprints. We first set the PSL 3D sensor in pol 0°, that is, TX and RX are both in TM polarization, and then obtain the depth map of this mirror scene, the result of which is shown in the first row of Fig. S15b. We can see that due to the influence of surface particles, the specular reflection is redirected, and most of the reflected light can be received by the RX, thus obtaining a large area of mirror depth. This kind of reflection still maintains the original polarization direction, so when we set the RX in pol 90°, as shown in the second row of Fig. S15b, the mirror part is eliminated, while the reflection from the wall which has different polarized directions due to diffuse reflection is reserved in the depth map. Then we wipe the mirror surface with alcohol to gradually improve the cleanliness, and their depth results are shown in Fig. S15c, S15d and S15e. We can see that in pol 0°, the depth area of the mirror surface gradually decreases as the cleanliness increases and in pol 90°, the depth information belonging to the mirror surface will be eliminated. Fig. S15e is the result of the cleanest situation, which means there is only complete specular reflection and not off-specular reflection [3]. Hence in pol 0°, the mirror area that can be detected is very small, and its corresponding theoretical model is illustrated in Fig. S15f. As shown in Fig. S15f, the baseline distance between the TX and RX of PSL 3D sensor is b (b = 4cm in the PSL 3D sensor), and the distance from the PSL 3D sensor to the reflective surface is d (d = 80cm in this section). In the situation where the PSL 3D sensor faces the reflective surface, the imaging light path of complete specular reflection is shown by the blue solid line. According to the reflection law, the incident angle θ_{i1} is equal to the reflection angle θ_{r1}, so only an area with the size of s_1 can be received by the RX. In the situation of oblique illumination where the reflective surface has an inclined angle α, the imaging light path is shown by the grey solid line. Similarly, the incident angle θ_{i2} is equal to the reflection angle θ_{r2}, and only the area with the size of s_2 can be received by the RX. The size of s_1 and s_2 is basically equal to the area of TX. When d/b is large, s_1 or s_2 will appear relatively small, as shown in the first row of Fig. S15e.

However, in common scenes, these situations where there is only specular reflection are rarely seen. Instead, a more suitable reflection model is shown in Fig. S15g. In this model, the projection range of TX is indicated by grey dot arrow lines and the receiving range of RX is indicated by orange dot arrow lines. As pointed by the blue solid arrow lines, the TX projects structured light onto the reflective surface and the reflected light will be affected by the reorientation of surface particles. In this situation, off-specular reflection dominates, which broadens the reflection lobe. Thus, in the receiving range, a larger area s can be detected by the RX. Furthermore, this reflection is not diffuse reflection and will still maintain the original polarized direction [3, 4], so when TX and RX have orthogonal polarized directions, the information belonging to the reflective surface will be completely eliminated. Benefitting from the off-specular reflection, for daily application scenarios, our PSL 3D sensor can extract enough depth information of the reflective surface and use it for the detection and completion applications, as demonstrated in Fig. 6 of the main text and Fig. S14 of the Supplementary Information.

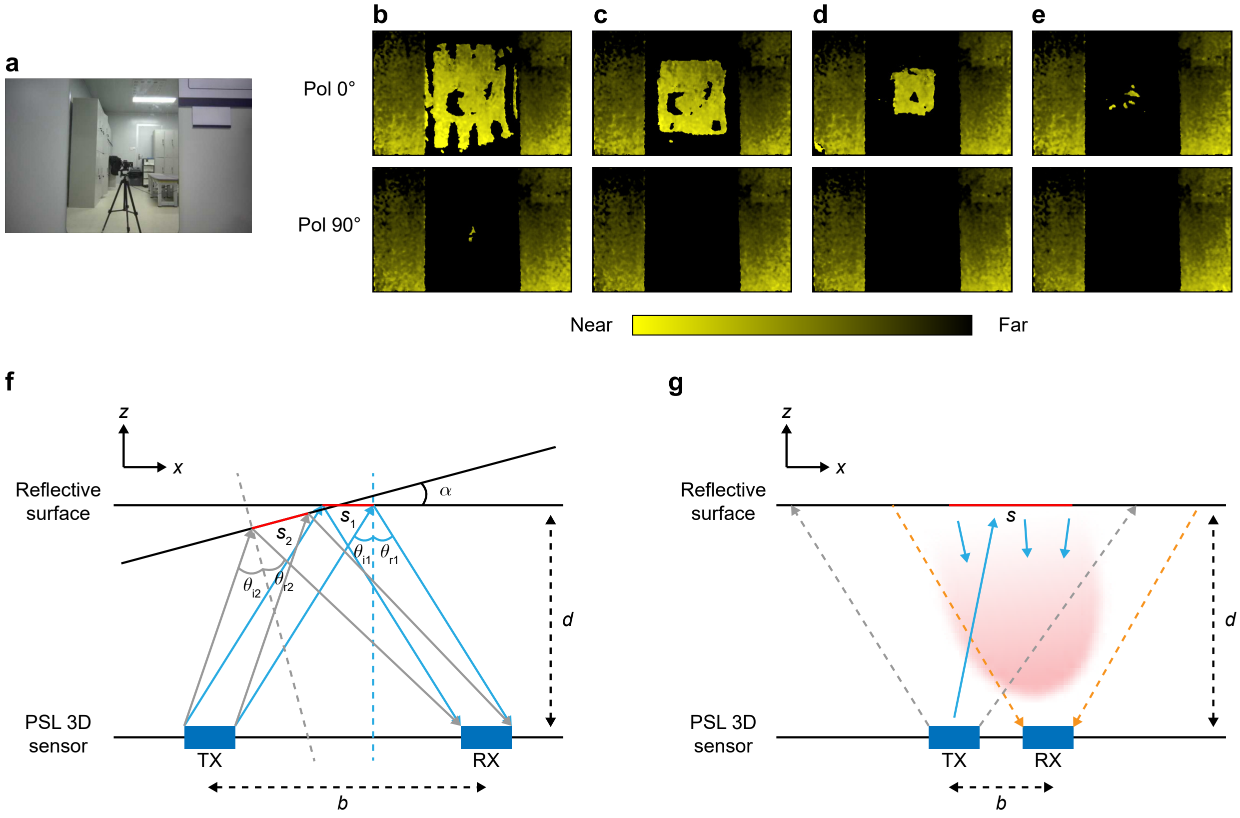

Figure S15: **Detection area analysis.** **(a)** The experimental scene for detection area analysis. **(b)-(e)** are the depth maps of the scenes with different degrees of cleanliness. The first row is the result of pol 0°. The second row is the result of pol 90°. **(f)** Light path model of complete specular reflection. **(g)** The reflection model for common scenarios.

IX. Working range of the PSL 3D sensor

In actual 3D applications, the glass scene is very complicated and cannot face our PSL 3D sensor every time. Therefore, we give the effective measurement range here. We fix the distance from the sensor to the flat glass door as 1 meter, and then measure the working range of the sensor by rotating it in the horizontal and vertical directions (Fig. S16a). Specifically, we take the number of depth points that belong to the glass when the sensor is at angle 0° as a benchmark, and then calculate the proportion of glass points relative to this benchmark at different angles. We take the angle at which the scale drops to 2% as the imaging limit.

In the horizontal direction, we measure 5 angles, $\{0°, 15°, 30°, 37.5°, 45°\}$. The depth maps of each angle are shown in the first row of Fig. S16b and their corresponding extracted glass regions are highlighted in the second row. Thus, we take the glass region of angle 0° as a benchmark and calculate the proportion of glass regions in other angles with respect to this benchmark. The proportion is shown in Fig. S16d, where the percentage drops to 2.66% at angle 45°. In the vertical direction, we also measure 5 angles, $\{0°, 7.5°, 15°, 22.5°, 30°\}$. With the same method, we show the results in Fig. S16c and Fig. S16e, where the percentage drops to 2.49% at angle 30°. Note that we only measure in the positive range, whereas it can expand to the negative range. Thus, the effective working range of our PSL 3D sensor is ±45° horizontally and ±30° vertically.

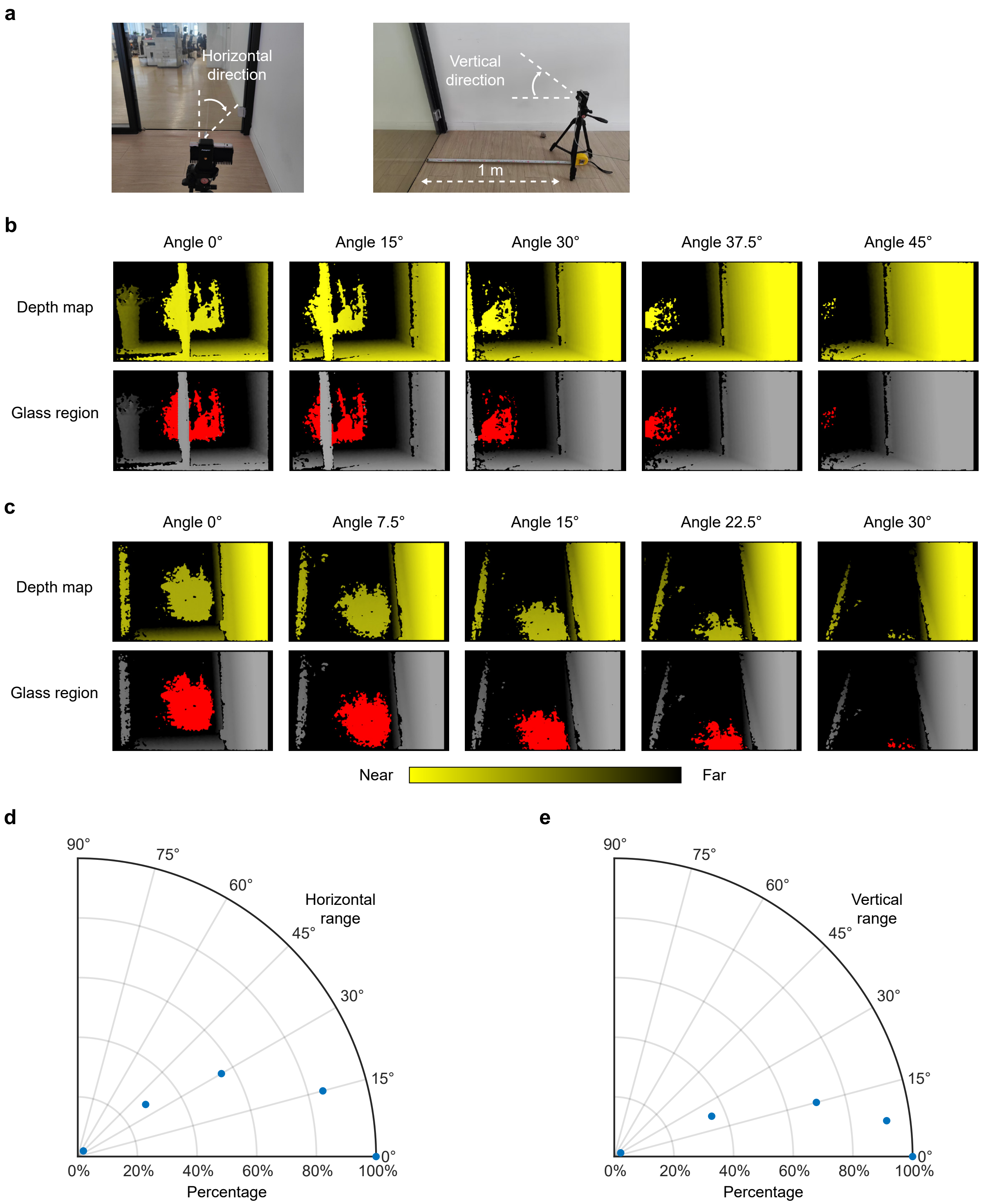

Figure S16: **Horizontal and vertical working range. (a)** Angle measurement setups of horizontal and vertical direction. **(b)** Measurement results of the horizontal range. **(c)** Measurement results of the vertical range. **(d)** The proportion of glass points in the horizontal direction. **(e)** The proportion of glass points in the vertical direction.

X. Panoramic reconstruction of a fish tank

All the detection and completion experiments in the main text are captured in the single view. If we want to reconstruct a complete 3D shape of an object, we need to employ our method in various angle. In this section, we have achieved the 360° panoramic reconstruction of a fish tank scene using multi-angle fusion. Specifically, we capture four sides of the fish tank using the PSL 3D sensor, as shown in the first row of Fig. S17a. By setting the polarization of the PSL 3D sensor to 0° and 90°, we obtain depth images at two different angles, as illustrated in the second and third rows of Fig. S17a. Consistent with previous experiments, when polarization is 0°, the PSL 3D sensor can reconstruct the glass part of the fish tank and the surrounding objects. When polarization is 90°, it can eliminate the glass part and retain only the surrounding diffuse components. Combined with the glass boundary in the fourth row, we then extract, fit and complete the glass, as shown in the fifth, sixth and seventh rows of Fig. S17a, respectively. Finally, by combining the reconstructions of the four sides of the fish tank with the ground as the reference plane, we use CloudCompare [5] to fuse the multi-angle data and achieve the 360° panoramic reconstruction of the fish tank scene, as shown in the 3D color and grayscale schematics in Fig. S17b and S17c, respectively.

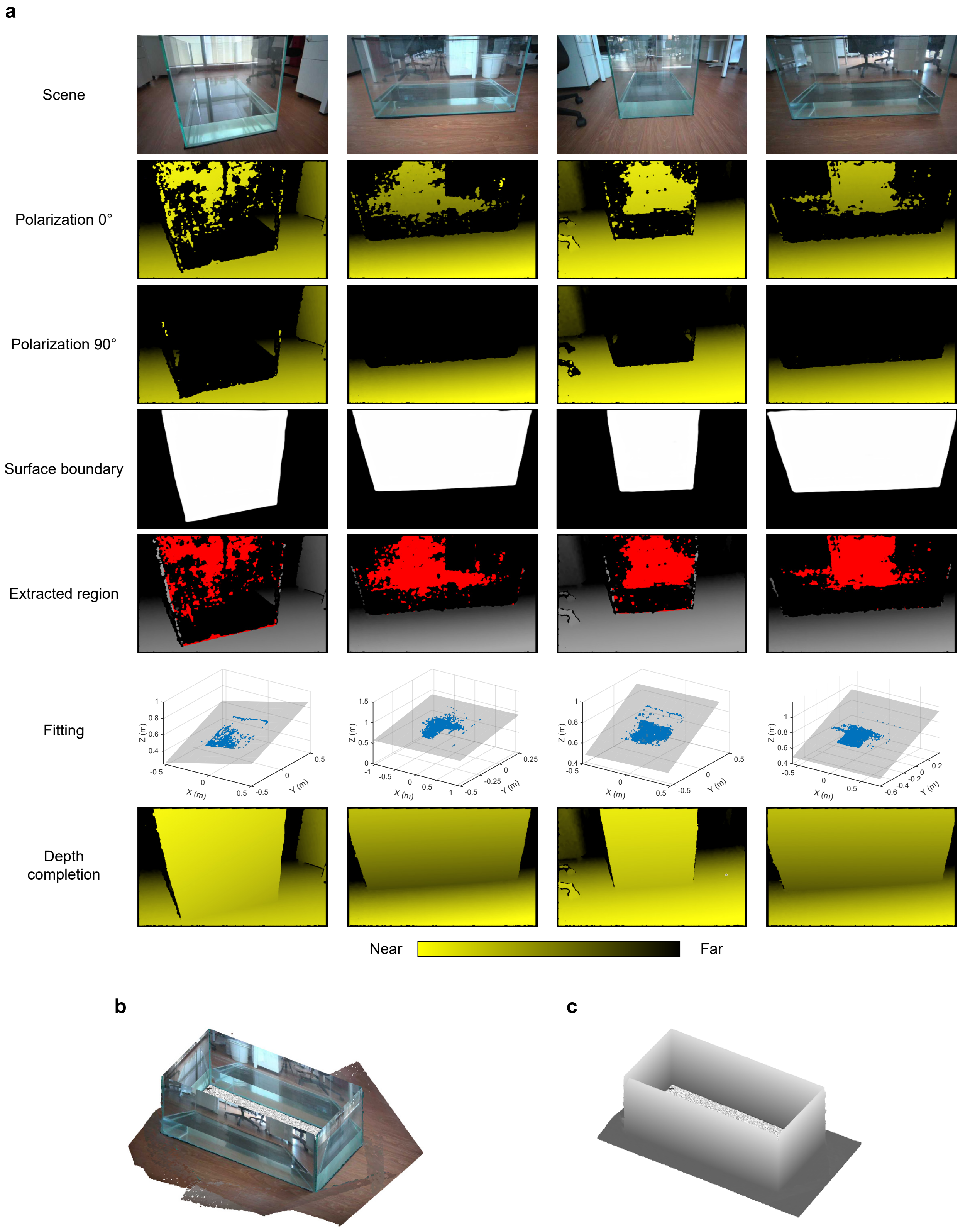

Figure S17: **Panoramic reconstruction of a fish tank.** **(a)** Detection and completion results of four sides of the fish tank. **(b)** and **(c)** are the panoramic reconstruction of the fish tank, rendered with color and in grayscale, respectively.

XI. Detection and completion result of scene with glass and mirror

In this section, we experimentally verify the reconstruction ability of the PSL 3D sensor when there is another strongly reflective object (such as a mirror) behind the glass. In this scene, the PSL 3D sensor is placed 1 m in front of the glass door, and a rectangular mirror and a cabinet are placed 0.5 m behind the glass door, as shown in Fig. S18a and S18b. We use the PSL 3D sensor to obtain two depth images at polarization 0° and 90°, as shown in Fig. S18c and S18d. At polarization 0°, we are able to obtain the depth of the front glass (red dot line) and the back mirror (white dot line). At polarization 90°, these two parts are eliminated simultaneously and only the surrounding diffuse components are remaining. After applying subtraction, the depths of the glass and mirror are both retained, as shown in S18f. But our main purpose is to identify the front obstacle. Hence, we can filter out the reflection from the back, making it successful to determine the front glass area, which is demonstrated in Fig. S18g. The final completed result is shown in Fig. S18i. As seen from the error map in Fig. S18j, our method can achieve reconstruction with high precision of 2 mm on the average.

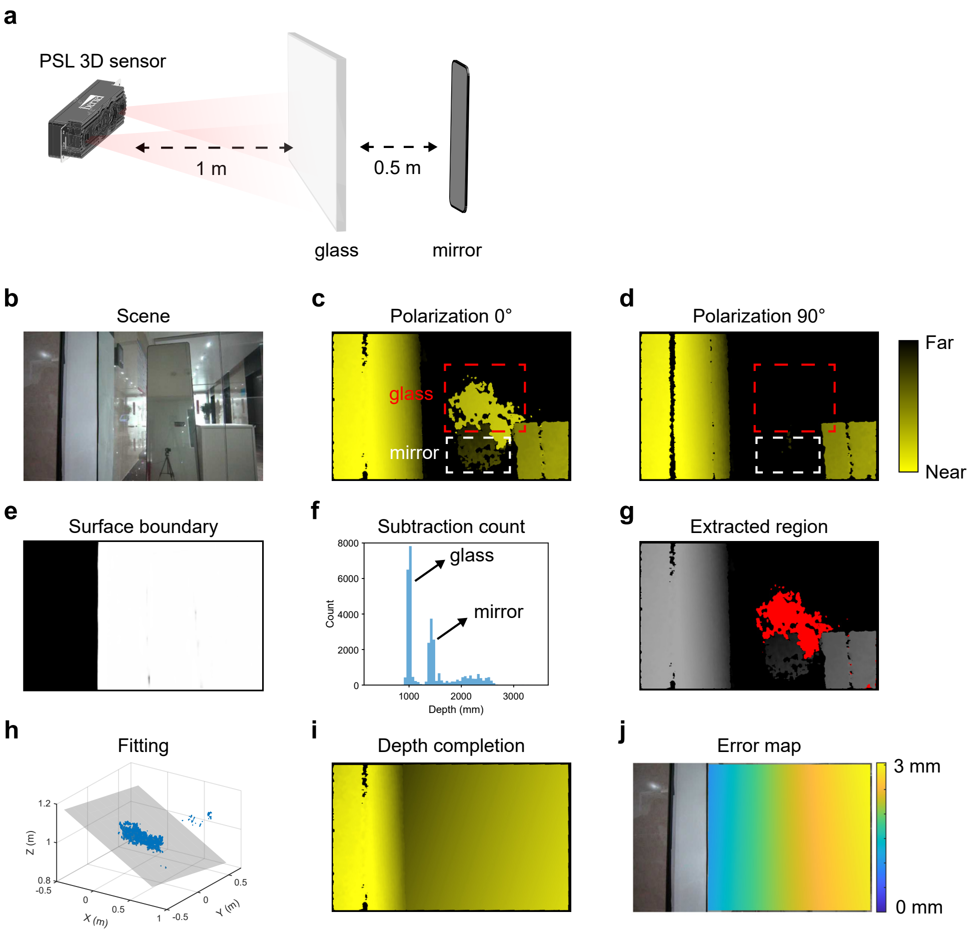

Figure S18: **Detection and completion result in scene with glass and mirror.** **(a)** Scheme of the experiment. **(b)** RGB image of the scene. **(c)** Depth map at polarization 0°. **(d)** Depth map at polarization 90°. **(e)** The predicted boundary. **(f)** The subtraction count. **(g)** The extracted reflective surface region. **(h)** Fitting of the extracted points in the world coordinate. **(i)** The completed depth image. **(j)** The calculated error map.

XII. Comparision with ToF sensor on detection and completion of reflective surfaces

In this section, we show the shortcoming of ToF sensor in the task of detecting and completing glass. In this comparison, we place the sensor 51 cm directly in front of the glass (Fig. S19a). Since the purchased ToF sensor (Orbbec Femto W) has been integrated together, we cannot replace the source with the HCG VCSEL, so we add polarizers to its TX and RX for testing. As shown in Fig. S19b, TX is set to pol 0° and RX is set to two directions. We also test the PSL 3D sensor in the same scene. The RGB pictures of the scene obtained by the ToF sensor and PSL 3D sensor are displayed in Fig. S19c and Fig. S19d respectively, where the diffuse sticker on the glass is denoted by orange dot line.

In this experiment, the results of ToF sensor are shown in Fig. S19e to S19h. As seen from Fig. S19e, the area of glass that can be detected is very small (denoted by white dot line), even though neither TX nor RX has a polarizer added. In addition, the depth of the sticker on the glass is denoted by orange dot line. When we add the polarizer of pol 0° to TX, the glass and sticker can be measured (Fig. S19f). When we add a polarizer of 0° to the RX, both the glass and sticker can still be detected because of the same polarization of TX and RX (Fig. S19g). Although the tree and wall behind are absent because of the occlusion, these are irrelevant factors, and we only need to focus on the parts of glass and sticker. When we place the polarizer of RX in pol 90°, the glass part is eliminated in the depth map because its reflection maintains the same polarization, while the diffuse sticker is detected in the depth map (Fig. S19h).

As for the PSL 3D sensor, as shown in Fig. S19i, we get a wide range of glass depth when we set it to polarization 0°. If we change to polarization 90° (Fig. S19j), the glass part can be removed, while the sticker on the glass is retained because of the diffuse reflection.

Although the ToF sensor can also add polarization characteristics to detect and eliminate the glass depth, it cannot obtain enough glass depth information because the depth calculation method is different from that of structured-light-based method. Therefore, PSL 3D sensor remains a powerful tool for the task of detecting and completing reflective surfaces.

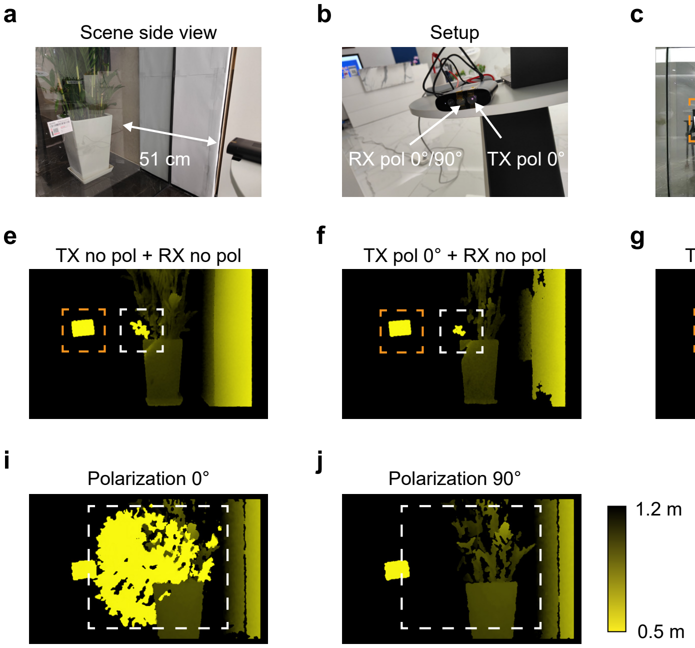

Figure S19: **Comparision with ToF sensor.** **(a)** Side view of the scene. **(b)** Setup of the ToF sensor. **(c)** Color image of the scene from the ToF sensor. **(d)** Color image of the scene from the PSL 3D sensor. **(e)**-**(h)** are depth maps from the ToF sensor, where their experimental settings are shown on the top of each figure. **(i)**-**(j)** are the results from the PSL 3D sensor.

XIII. Calculating the depth of the reflective surface

In this section, we use the scene in Fig. S5b as an example to show how the glass depth is calculated. We extract the infrared images received at the RX and present them at Fig. S20b and S20c. As we can see from Fig. S20b, at polarization 0°, we can obtain dot pattern reflected from the glass (denoted by the white dot line). Then we can calculate the depth of the glass according to the offset relationship between these received speckles and the original reference pattern, and the depth result is shown in Fig. S20d. Because these reflected speckles remain the original polarization, when we set to polarization 90°, the glass speckles are removed in the infrared image (Fig. S20c), and also the information that belongs to the glass in the depth image (Fig. S20e).

In addition, we illustrate the common triangulation method in Fig. S20f. In this scheme, the distance to the reference plane L, the baseline b between TX and RX, the imaging focal length f are known parameters. The point C is the detecting point, whose depth is our target. Due to the displacement effect of point C, the original reference point F on the imaging plane is now moved to G. If we calculate the displacement distance of FG, we can get the depth based on similar triangle relation,

$$\text{Depth} = \frac{L \times b \times f}{f \times b + \overline{FG} \times L} \tag{2}$$

Thus, the depth of the glass can be calculated with this method based on the received speckles.

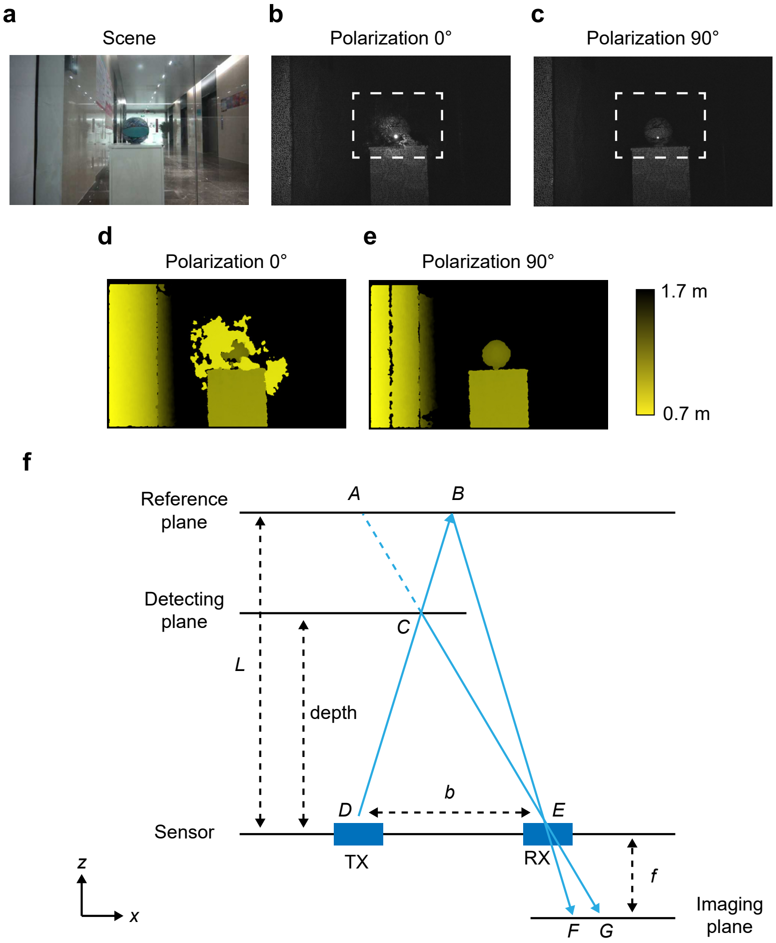

Figure S20: **The method of calculating depth.** **(a)** Picture of the scene. **(b)** Infrared image at polarization 0°. **(c)** Infrared image at polarization 90°. **(d)** and **(e)** are the corresponding depth maps. **(f)** Principle of the triangulation method.

Supplementary References

[1] Huang, M. C., Ye, Z. & Chang-Hasnain, C. J. A surface-emitting laser incorporating a high-index-contrast subwavelength grating. *Nature Photonics* **1**, 119–122 (2007).

[2] Huang, M. C., Ye, Z. & Chang-Hasnain, C. J. Polarization mode control in high contrast subwavelength grating vcsel. In: *2008 Conference on Lasers and Electro-Optics and 2008 Conference on Quantum Electronics and Laser Science*, 1–2 (2008).

[3] Torrance, K. E. & Sparrow, E. M. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.* **57**, 1105-1114 (1967).

[4] Wolff, L. B., Nayar, S. K. & Oren, M. Improved diffuse reflection models for computer vision. *International Journal of Computer Vision* **30**, 55–71 (1998).

[5] CloudCompare (version 2.10.alpha) [GPL software]. (2023). Retrieved from http://www.cloudcompare.org/