



Robotic-arm-assisted flexible large field-of-view optical coherence tomography

YONG HUANG,^{1,2,3} XIAOCHEN LI,^{1,2} JUNYU LIU,¹ ZHENGYU QIAO,¹
JINGSI CHEN,¹ AND QUN HAO^{1,4}

¹*School of Optics and Photonics, Beijing Institute of Technology, No. 5 South Zhongguancun Street, Haidian, Beijing 100081, China*

²*Equal contributors*

³*huangyong2015@bit.edu.cn*

⁴*qhao@bit.edu.cn*

Abstract: Optical coherence tomography (OCT) is a three-dimensional non-invasive high-resolution imaging modality that has been widely used for applications ranging from medical diagnosis to industrial inspection. Common OCT systems are equipped with limited field-of-view (FOV) in both the axial depth direction (a few millimeters) and lateral direction (a few centimeters), prohibiting their applications for samples with large and irregular surface profiles. Image stitching techniques exist but are often limited to at most 3 degrees-of-freedom (DOF) scanning. In this work, we propose a robotic-arm-assisted OCT system with 7 DOF for flexible large FOV 3D imaging. The system consists of a depth camera, a robotic arm and a miniature OCT probe with an integrated RGB camera. The depth camera is used to get the spatial information of targeted sample at large scale while the RGB camera is used to obtain the exact position of target to align the image probe. Eventually, the real-time 3D OCT imaging is used to resolve the relative pose of the probe to the sample and as a feedback for imaging pose optimization when necessary. Flexible probe pose manipulation is enabled by the 7 DOF robotic arm. We demonstrate a prototype system and present experimental results with flexible tens of times enlarged FOV for plastic tube, phantom human finger, and letter stamps. It is expected that robotic-arm-assisted flexible large FOV OCT imaging will benefit a wide range of biomedical, industrial and other scientific applications.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Owing to its advantages of non-invasiveness, high sensitivity, and high resolution, optical coherence tomography (OCT) has been increasingly applied from biomedical research to industrial inspection [1–4]. Three-dimensional (3D) and high-resolution OCT images of biological tissue both *in vitro* and *in vivo* provide visual, realistic, and comprehensive information about tissue structure as well as disease characteristics that have been used in increasingly more medical research and diagnostic applications. Typical 3D OCT imaging applications include both biomedical research such as imaging of anterior segment of the eye, retina [5–9], small animal organs [10–12], and non-biomedical research such as courtroom science [13,14] and oil painting identification [15].

Although the advantage of 3D OCT imaging has been clearly demonstrated, it suffers from limited imaging FOV in both axial depth direction (a few millimeters) and lateral direction (a few centimeters) [5,16]. While the axial imaging range is limited by either short coherence length of swept source in SSOCT systems or finite spectral sampling resolution for SDOCT systems, lateral imaging range is dependent on the adopted scanning optics. Efforts to extend the imaging range have shown great value in peripheral retinal examination [17], whole-eye assessment [18], whole-brain vascular visualization in neuro-science [19], and skin imaging [20].

Currently, high resolution large volume OCT imaging is still very challenging. Methods proposed to extend the FOV of OCT systems can be divided into mainly two categories. The first is to use long coherence length light source combined with increased detection bandwidth for large volume OCT imaging. Recently, Wang et al. [21] performed large volume cubic meter imaging using tunable vertical cavity surface lasers and silicon photonic integrated circuit (PIC) system. Song et al. [22] achieved several hundred cubic centimeters FOV using all semiconductor programmable akinetic swept source and wide-angle lenses with long focal length. These methods mainly suffer from poor lateral resolution and imaging field distortion. At this stage, it is still challenging to achieve true high-resolution, large-volume OCT. The second category combines linear translation stages and image stitching techniques to extend the OCT range and achieve wide-field imaging [23–25]. It can maintain both good axial and lateral resolutions.

Current state-of-the-art OCT imaging technology utilizes actuation systems composed of XYZ translational stages with a maximum of 3 DOF to obtain extended 3D volume of the specimen tissue. These devices drive the probe to move in a strictly predetermined path, making scanning less flexible. Therefore, the probe attitude cannot be changed during the imaging process. Additional acquisitions from multiple positions and poses are required to create a complete 3D profile, which often requires changing the position and angle of the sample, limiting the application of 3D OCT to complex scenarios. Handheld OCT provides a possible solution, however, without a support framework, the inevitable jittering of the operator's hand, typically results in the shaking of the probe, which introduces position errors and even signal loss for the subsequent volume reconstruction [26].

Collaborative robotic arm has been combined with medical imaging and operation for a while due to its obvious advantages such as stability, high precision, repeatability, multi-degrees of freedom, mobility and remote control. Robotic surgery has become a reality in several surgical procedures. The DaVinci surgical robotic system is the most representative commercial surgical robotic system [27]. Draelos proposed a robotically-aligned OCT scanner capable of automatic eye imaging without chinrests [28]. F. Rossi has developed a vision-guided robotic platform for laser-assisted anterior eye surgery. [29]. Y. Sun proposed a method to plan grinding paths and velocities based on 3D medical images in the context of robot-assisted decompressive laminectomies, and the result suggested that robot-assisted decompressive laminectomies can be performed well [30]. Basov et al. combined a laser welding system with a robotic system and demonstrated its feasibility for skin incision closure in isolated mice [31]. Robnier et al. proposed the use of a 7-degree-of-freedom industrial robotic arm to enhance OCT scanning capabilities and discussed the feasibility of intraoperative OCT imaging of the cerebral cortex [32]. These examples illustrated the unique advantages and potential of the robotic arm.

Therefore, we propose the use of a collaborative 7-DOF robotic arm for enhanced OCT scanning flexibility and FOV. This method uses a computer-controlled robotic arm to put the OCT probe to different viewpoints to achieve large volume scan of the sample. In this work, we demonstrate, for the first time to our knowledge, a flexible wide-field OCT system with a 7-DOF robotic arm and characterize its performance. A depth camera is adopted to capture the general position of the sample surface. A miniature OCT probe equipped with an RGB camera is mounted at the tool flange center of the robotic arm. Robotic arm base and tool, depth camera, OCT imaging coordinate systems are calibrated and registered first. Based on the depth camera information, target sample is located and the starting point of scan path is set. In addition, with the help of RGB camera image, OCT probe is finely aligned with the imaging target. The overall scan path can be either predetermined with fixed scanning point intervals or controlled manually using the RGB camera image. At each acquisition point, rapid 3D OCT imaging and reconstruction provides feedback on the relative position of the probe to the target as the input to optimize the probe position and attitude. Afterwards an OCT volume and the corresponding spatial location information are saved. When scanning is done, the system uses the collected

OCT volume data and corresponding spatial coordinates for 3D reconstruction, visualization and analysis.

The remainder of this paper is organized as follows. Details of our method are illustrated in Section 2. Experiment results of system for different samples are presented and discussed in detail in Section 3. Finally, the main conclusions are presented in Section 4.

2. Methods

2.1. System setup

Schematic system setup and photo of the prototype system are shown in Fig. 1(a) and Fig. 1(b), respectively. The system consisted of an OCT trolley enclosing a SDOCT imaging engine and a workstation (Dell Precision T3630, USA), a customized miniature OCT probe, a 7-DOF robotic arm (xArm7, UFACTORY, China) and a depth camera (RealsenseD435, Intel, USA). Customized system software was developed combing Qt (v5.14.2) and Microsoft Visual Studio 2019 in C++ for OCT and depth camera data acquisition, processing, display, storage, and system control. GPU accelerated OCT signal processing was implemented based on CUDA (v10.1). Communication between the workstation and the robotic arm controller is via the TCP/IP protocol using provided software development kits (xArm-C++-SDK-1.6.0). The robotic arm control software could read the configurations of the robotic arm in real time with a control latency of 4 ms, which include the pose of the robotic arm and the moving speed and acceleration.

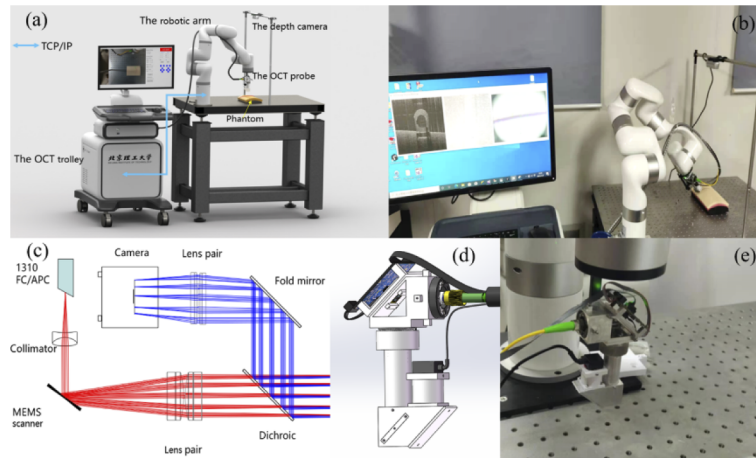


Fig. 1. Illustration of the prototype system. (a) overall system setup, (b) photo of the experimental system, (c) optical design of the miniature probe: red and blue rays describe the OCT optical path and the RGB camera optical path, respectively, (d) optomechanical design of the probe, and (e) fabricated probe fixed on the robotic arm.

A homebuilt $1.3\ \mu\text{m}$ SDOCT system was developed with a depth imaging range of 3.6 mm, axial resolution of $12\ \mu\text{m}$, lateral resolution of $31\ \mu\text{m}$, and maximal imaging speed of 76 kHz. During the experiment, each volume data consists of 256 B-scans with image size of 1024×1024 pixels. The system was running at imaging speed of 30 fps. A customized miniature OCT scanning probe with an integrated RGB camera was mounted on the robotic arm for the OCT volume data acquisition. The optical layout of the OCT probe is shown in Fig. 1(c). A two-axis MEMS micromirror with diameter of 3.6 mm was used for beam scanning. A dichroic mirror (DMLP950 T, $\varnothing = 12.7\ \text{mm}$, Thorlabs Inc.) with 950 nm cut-off wavelength was used to combine the OCT imaging path and the RGB camera imaging path. An off-the-shelf focal lens (AC127-050-C, $\varnothing = 12.7\ \text{mm}$, Thorlabs Inc.) with focal length of 50 mm was used as the

objective lens. The working distance of the probe is 23 mm. The RGB camera is a common miniature high-definition CMOS camera with integrated adjustable focal lens and an image size 1280×720 pixels. It fits well into the probe, allowing real-time imaging of the target area within a FOV of $12 \times 12 \text{ mm}^2$. Custom lens tubes and frames were designed to accommodate the closely spaced optics of the system and keep the dimensions small, as shown in Fig. 1(d). The internal skeleton and other structural components are made of aluminum. Photo of the fabricated probe is shown in Fig. 1(e). It weighs 330 g with a size of $4.3 \text{ cm} \times 3.5 \text{ cm} \times 11.8 \text{ cm}$ (L \times W \times H).

The D435 depth camera was fixed horizontally about 60 cm above the optical table where targeted phantoms were placed. It plays the role of sensing the external environment and locating the position of the target object relative to the robotic arm in our system. The D435 delivers depth and color images at resolution up to 1280×720 at 30 fps with working distance from 0.1 to 10 m and FOV of 85×58 degrees. The robotic arm can move the probe with 7-DOF and have a working radius of 691 mm and positioning accuracy of 0.1 mm for the probe actuation.

2.2. System calibration

To guide the probe, all components should be registered and calibrated into one unified coordinate system. As illustrated in Fig. 2, there are four coordinate systems to be unified including the robot base coordinate system R, tool coordinate system T, the depth camera coordinate system D, and OCT image coordinate system O. The robotic arm base coordinate system R is a coordinate system in which the central point of the robot base is the coordinate origin. The robotic arm tool coordinate system T is a coordinate system with the robotic arm tool center point (TCP) as the coordinate origin. Coordinate R and T are associated with the 7 DOF robotic arm. The origin of the depth camera coordinate system D is the center point of the left infrared lens. The origin of the 3D OCT image coordinate system O is the midpoint of the top line in the 128th image of 256 B-scan (one C-scan volume) images. Inset of Fig. 2 illustrates two vectors that determine the attitude of the probe. V_p denotes the unit vector parallel to the optical axis while V_e denotes the unit vector parallel to the B-scan direction.

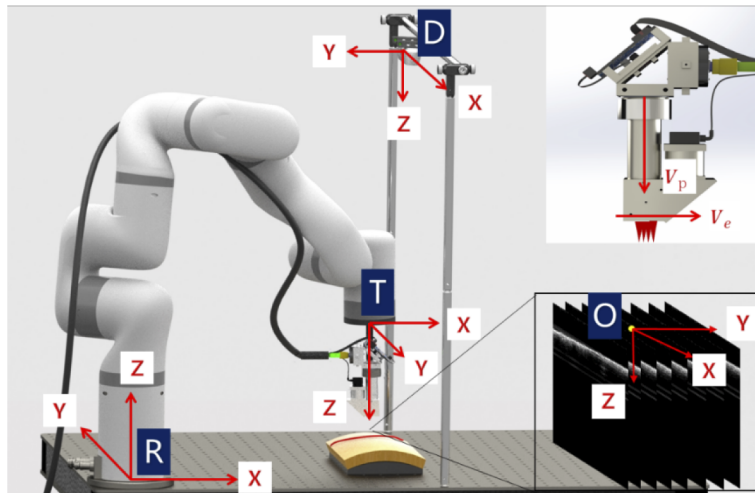


Fig. 2. Illustration of the robotic arm base coordinate system R, tool coordinate system T, the depth camera coordinate system D, OCT image coordinate system O, probe optical axis direction V_p and B-scan direction V_e .

First, spatial relationship between the coordinate systems D and R was calibrated. The calibration can be performed by obtaining a set of coordinate values of the same points in these

two separate coordinate systems [33]. The spatial calibration made use of a 9×12 chessboard board containing rectangles with a size of 1.5 cm×1.5 cm. A standard steel pin (0.15×41 mm, Diameter×Length) was fixed at the center point of the tool flange of the robotic arm. As shown in Fig. 3(a), the pin tip was driven to the intersection of the squares on the board and corresponding position parameters of the robotic arm were recorded. Since the length of the needle tip is fixed, the coordinates of the intersection point in the coordinate system R can be calculated by adding the length value to the robotic arm Z parameter. Then, coordinates of the corresponding intersection points in the coordinate system D can be obtained from the depth camera, as illustrated in Fig. 3(b).

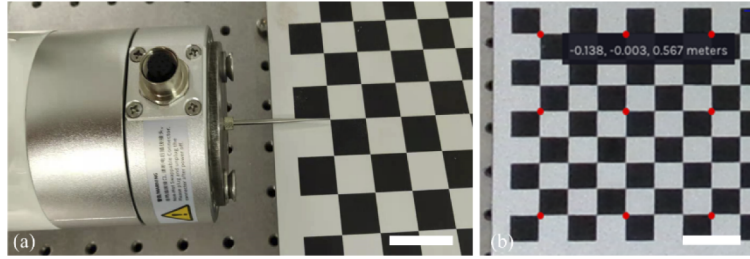


Fig. 3. Spatial calibration between the coordinate systems D and R. (a) placement of the steel needle tip to calibration points driven by the robotic arm, and (b) spatial coordinates of the corresponding calibration points (marked in red) obtained from the depth camera. Each square size is 15mm×15 mm. (scale bar: 30 mm).

Defining (X_D, Y_D, Z_D) as the spatial coordinate values of the cross intersections in D, and (X_R, Y_R, Z_R) as the spatial coordinate values of the cross intersections in R. The coordinate transformation of D to R is given by:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_R = R_0 \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_D + \begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \\ \Delta Z_0 \end{bmatrix}, \quad T_0 = \begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \\ \Delta Z_0 \end{bmatrix} \quad (1)$$

where R_0 represents the rotation matrix of the coordinate system D to the coordinate system R, and T_0 is the translation matrix. Then R_0 and T_0 were calculated by finding the least-squares fit of the two 3D point sets [33]. Each point set should contain more than three points. We chose 9 points for each set for the fitting. The above equation can be written as a nonlinear homogeneous Eq. (2).

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_R = {}^R M_D \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_D, \quad {}^R M_D = \begin{bmatrix} R_0 & T_0 \\ 0 & 1 \end{bmatrix} \quad (2)$$

where the form of ${}^R M_D$ represents the transformation matrix from coordinate system D to coordinate system R.

Second, spatial relationship between the coordinate systems O and T was calibrated. Registration of OCT images into coordinate system R is necessary for post image reconstruction and probe pose optimization. As shown in Eq. (3), it requires transforming sample position P_O in

OCT coordinate system O into position P_R in robotic base coordinate system R.

$$P_R = {}^R M_T^T M_O P_O, P_O = \begin{bmatrix} S_x x \\ S_y y \\ S_z z \\ 1 \end{bmatrix} \quad (3)$$

where P_O can be obtained by multiplying the pixel position index (x, y, z) in OCT image domain with the corresponding pixel domain to spatial domain conversion coefficients S_x , S_y , and S_z in three directions respectively. ${}^R M_T$ is the position transformation matrix from tool coordinate system T to robotic base coordinate system R, which can be calculated directly from the six control parameters of the robot arm. ${}^T M_O$ can be calibrated by the commonly used hand-eye calibration method, as shown in Fig. 4.

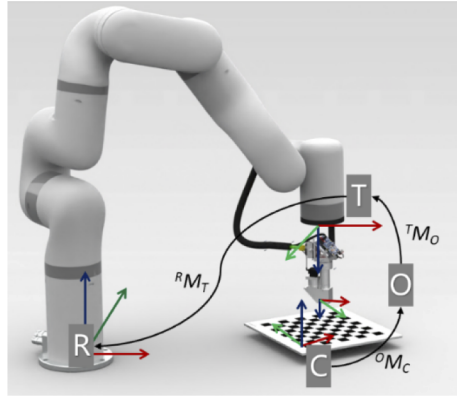


Fig. 4. Spatial calibration between the O and T coordinate systems.

The robotic arm drives the probe to image the chessboard with different poses. Since the position of the chessboard is fixed with respect to the base of the robotic arm, for any two different imaging poses, Eq. (4) can be obtained:

$$P_R = {}^R M_{T1}^T M_O {}^O M_{C1} P_C = {}^R M_{T2}^T M_O {}^O M_{C2} P_C, P_O = {}^O M_C P_C \quad (4)$$

where ${}^O M_C$ represents the transformation matrix from chessboard coordinate system C to OCT coordinate system O, which can be deduced from 3D OCT images of the chessboard for every imaging pose. Since the imaging pose are different, two transformation matrix ${}^R M_{T1}$ and ${}^R M_{T2}$ can be calculated from two sets of robotic arm control parameters.

Then Eq. (4) can be deduced into Eq. (5), which is in the form of $AX = XB$. ${}^T M_O$ can be solved using Tsai-Lenz method described in [34]. To get one unique solution, at least three imaging poses are required.

$$({}^R M_{T2})^{-1} ({}^R M_{T1})^T M_O = {}^T M_O ({}^O M_{C2}) ({}^O M_{C1})^{-1} \quad (5)$$

2.3. Determination of the probe position and attitude

Depth camera was used to guide the probe to the initial imaging pose after system calibration. Depth camera is only needed for the initial alignment for a stationary sample. Figure 5(a) shows a representative image of the targeted vessel phantom (plastic tube filled with red dyes) with

depth information overlaid. In Fig. 5(a), three points marked in green to form a tiny plane area of interest, two points marked in blue along the vessel to form the small region scan path vector and one scanning start point marked in red were selected manually. Figure 5(b) shows the initial area of interest with 3D surface rendering corresponding to normal camera image of the sample in Fig. 5(c). Yellow arrow represents the surface normal vector. Ideally, V_p should be parallel to the tissue surface normal vector and V_e should be perpendicular to the scan path vector, and the distance between the probe and the tissue plane is a fixed value, as shown in Fig. 5(d). Figure 5(e) and Fig. 5(f) show the probe at the original position and the specific pose under the guidance of the D435, respectively.

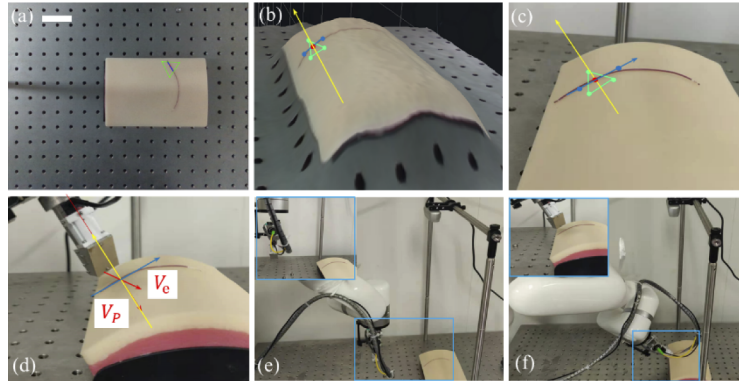


Fig. 5. Depth camera based initial probe placement. (a) image of target area with depth information overlaid, (b) 3D surface rendering of target area (c) normal camera image of target area, (d) zoomed view of the relative position between probe and target sample, (e) probe at original position, and (f) probe at the specific pose. (blue vector indicates the scan path and the yellow vector indicates the surface normal vector.) (scale bar: 50 mm).

The RGB camera integrated in the probe was used for fine alignment of the probe to the sample due to the reason that the positioning accuracy of the depth camera D435 is only about 1 mm, only rough alignment can be achieved. As shown in Fig. 6(a), after initial positioning, the center of the scanning area of the probe (green cross) was not aligned well with the phantom tube. Manual fine adjustment of the probe position was performed based on the RGB camera image, as shown in Fig. 6(b).

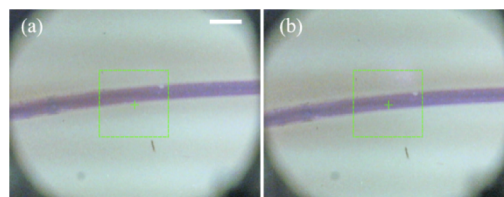


Fig. 6. RGB camera based fine probe placement. (a) RGB camera image of the targeted area after coarse placement of the probe guided by depth camera, (b) RGB camera image of the targeted area after fine placement of the probe guided by the RGB camera. (scale bar is 2 mm. green rectangular shows the lateral FOV of OCT imaging.)

2.4. Probe pose optimization by OCT volume feedback

After initial and fine adjustment, the imaging probe might not be at the optimal pose relative to special samples that are of relatively complex morphology due to the lack of positioning

accuracy from depth camera and lack of orientation adjustment capability of the RGB camera. OCT volume images can be further used as the feedback to optimize the imaging probe when necessary as they carry spatial information of the probe relative to the target sample.

It worth mentioning first that probe optimization strategy and corresponding processing algorithms are application dependent. Here we give a representative example of curved plastic tube which mimics future intraoperative imaging application for blood vessel inspection after anastomosis operation. At each acquisition point, within one C-scan volume, Hough transform linear detection method was used to fit the underlying skin phantom surface contour line of each B-scan, shown in red line in Fig. 7(a). Thus, 3D skin surface contour can be obtained. For the plastic tube target, edge detection method was used to extract the edge vertex points shown in Fig. 7(b). The system then randomly selects three points on the skin surface shown as green dots in Fig. 7(c) and two points on the upper edge of the tube shown as blue dots in Fig. 7(e) to calculate the surface normal vector and scan path vector for pose optimization.

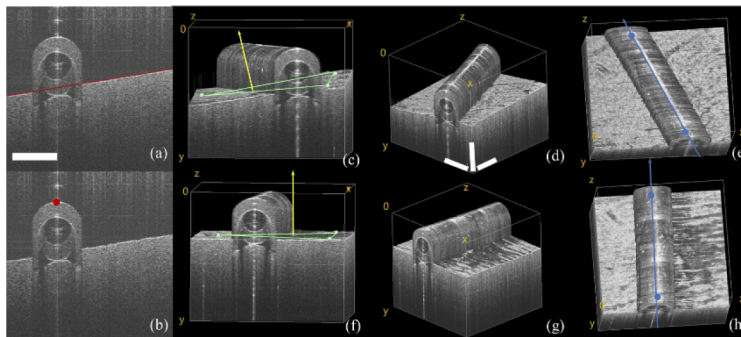


Fig. 7. Optimization of the probe pose by OCT volume feedback. (a) edge detection of underlying skin surface; (b) tube edge vertex point detection; 3D reconstruction of the plastic tube before probe pose optimization (c) front view, (d) isotropic view, and (e) top view; after probe pose optimization (f) front view, (g) isotropic view, and (h) top view. (scale bar: 1 mm)

From Fig. 7(c) to Fig. 7(e), we can see that before imaging pose optimization the tissue plane and sample tube are in a tilted state, indicating that the probe is not perpendicular to the skin surface plane and the OCT B-Scan direction is not perpendicular to the long axis of phantom tube. After probe pose optimization, 3D OCT images were acquired and reconstructed again shown in Fig. 7(f) to Fig. 7(h), where the phantom skin surface plane is basically at the same imaging depth and the OCT B-scan direction is basically perpendicular to the long axis of plastic tube segment.

3. Results and discussions

3.1. Robotic arm positioning accuracy test

The positioning accuracy and stability of the robotic arm directly affect the final OCT 3D reconstruction results. To verify the positioning accuracy of the xArm7 robotic arm used in the system, repeated positioning accuracy and absolute positioning accuracy test experiments were conducted. The control system first moved the robotic arm to the same designated position ten times from different initial positions. Then a chessboard sample was imaged at that position. Figure 8(a) shows the *enface* projection images of the intersection area with the intersection point marked by yellow dots. Spatial distribution of the position coordinates of the intersection point is shown in Fig. 8(b). A minimum enclosing sphere with radius of 0.112 mm can be found,

which means the repetitive positioning accuracy of the robot arm is ± 0.112 mm. It is close to the official claimed ± 0.1 mm.

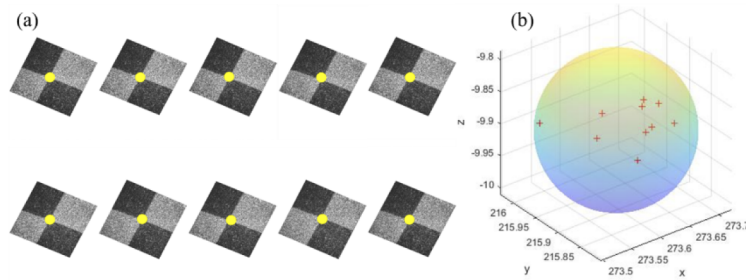


Fig. 8. Repetitive positioning accuracy test result of the robotic arm. (a) *enface* OCT images of the same chessboard intersection area. (b) spatial distribution of the position coordinates of intersection point for 10 consecutive repetitions.

Absolute positioning accuracy test results are shown in Fig. 9. The robotic arm drove the probe to the chessboard intersection positions and the coordinate parameters of the robotic arm at each point were recorded. 24 intersections of the chessboard marked with red dots in Fig. 9(a) were imaged. RGB camera image shown in Fig. 9(b) was used as the reference for the positioning. Figure 9(c) shows the position deviation of robotic arm when imaging intersection points with top left point as the distance origin. The average positioning deviation for the robotic arm is 0.21 ± 0.09 mm, which means the absolute positioning accuracy of the robot arm is about 0.21 mm. We set a redundancy of 0.3 mm between two adjacent scanning points during the scanning process to ensure that there is an overlap between two adjacent scanning areas.

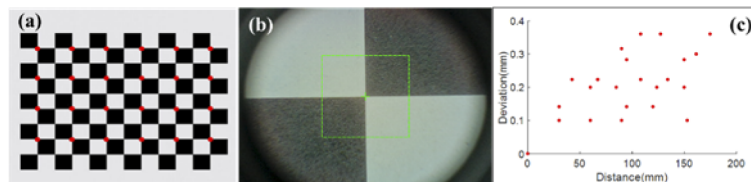


Fig. 9. Robotic arm positioning variation test results. (a) the calibration chessboard with imaged intersection points marked with red dots, (b) RGB camera image for positioning alignment, (c) position deviation of the robotic arm with respect to the distance from the top-left origin position.

3.2. Flexible large field of view imaging

To evaluate the performance of the system in different tissues and the feasibility of large-scale scanning and 3D reconstruction, we used it to image a fingertip phantom, a letter stamp, and a curved plastic tube filled with red dye. The single imaging FOV size is $4.1 \text{ mm} \times 4 \text{ mm} \times 3.6 \text{ mm}$ ($X \times Y \times Z$). It took about 10 s to record an OCT volume of 256 B-scans at each scan point. For future different OCT applications, it is important to mention that OCT probe can be customized on demand and the size of the FOV can be optimized depending on the application. For example, clinical intraoperative vascular imaging requires miniature probe as was used in this manuscript due to tight space limitation [35]. For large-scale nondestructive material evaluation such as artwork inspection, an OCT probe with larger FOV is preferred. Meanwhile, design of lateral FOV size needs to be matched with axial FOV of the imaging system. Too large lateral FOV with

limited axial FOV will cause signal loss at edges of the lateral FOV if sample surface profile varies heavily.

Figure 10(a) shows the image of fingertip phantom. The scanning range is outlined with yellow box. To cover the whole scan range, 12 scanning points (green dots) were programmed evenly along the scanning path marked with the green line. At each scanning point, the corresponding robotic arm position and pose parameters were saved. Pixel coordinates of each pixel point in the OCT volume can be converted to the robotic arm base coordinate system, and each pixel point is projected to the corresponding position in real space to generate point cloud data. For the overlapping area of two adjacent volumes, we cover the overlapping part of the previous volume with the latter volume based on scanning order. A redundancy of 0.3 mm was set for the scanning region to ensure that there is enough overlap between the two adjacent volumes so that no gaps are created.

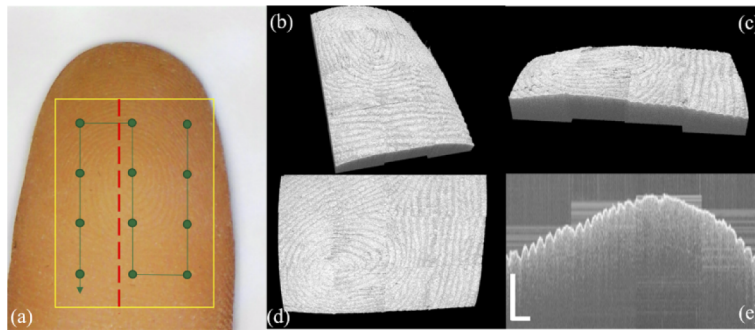


Fig. 10. (a) Photo of the fingertip phantom. scanning range is outlined with yellow box, scan path is shown in green line with scan points marked; 3D reconstruction of the fingertip phantom in (b) isotropic view, (c) front view, (d) top view and (e) stitched cross sectional image along the red dashed line. (scale bar: 1 mm)

During the imaging process, the robotic arm drove the probe just as conventional 3D translation stages. 3D reconstruction of the final stitched fingertip from isotropic view, front view and top view are shown in Fig. 10(b-d) respectively. We can see clearly that both surficial fingerprint pattern and natural surface contour were depicted. Since at current stage all the sub-volumes were manually registered and fused using the simultaneously recorded probe pose information, boundary outlines were noticeable. Future automatic image registration and fusion study is necessary. With one large volume data, cross-sectional inspection of the sample with large FOV can be achieved. Figure 10(e) shows the sample profile along the red dashed line in Fig. 10(a).

To further demonstrate the capability of the system to scan and reconstruct the surface of sample over a large area, thus the potential for applications such as industrial defect detection and artwork identification, we imaged a letter stamp with 32 single volumes stitched. Figure 11(a) shows the image of the letter stamp, the scanning range outlined by yellow box and scan path in yellow line with scan points marked. Figures 11(b)–11(d) presents the final reconstructed stamp surface from different viewpoints. We can see the reconstruction of the elevation of “OCT” letters. Figure 11(e) shows the cross-sectional image of the red dashed line in Fig. 11(a).

To validate that the stitched volume after reconstruction can capture the physical property of the target, a height of 12.55 mm for the letter “O” was measured after reconstruction of the whole stamp from the software shown in Fig. 12(a), which is very close to the measured value of 12.5 mm from the caliper.

To demonstrate the flexible large FOV imaging capability of the system with pose optimization, a curved plastic tube sample filled with red dye was placed on an arc shaped skin phantom. Figure 13 shows the relative pose between the probe and tube during the scanning process with

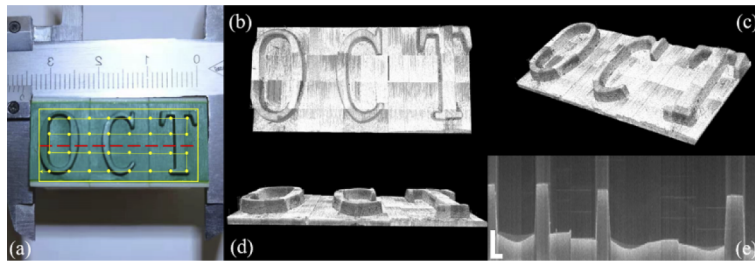


Fig. 11. (a) Picture of the letter stamp. scanning range is outlined with yellow box, scan path is shown in yellow line with scan points marked; 3D reconstruction of the fingertip phantom in (b) top view, (c) isotropic view, (d) front view and (e) stitched cross sectional image along the red dashed line. (scale bar: 1 mm)

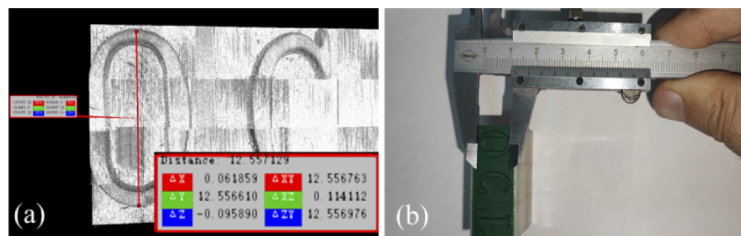


Fig. 12. Measurement of the height of letter “O”: (a) after reconstruction from software, (b) from Vernier caliper.

the scanning path length of approximate 68 mm consisting of 18 scanning points. The scan path points were manually selected by mouse-clicking from the probe RGB images to ensure that two adjacent imaging areas overlapped. From Fig. 13 we can see that the probe moved along the curved blood vessels on the complex skin surface and at each scanning point the probe position and attitude was optimized with the feedback from 3D OCT images, ensuring high-quality image acquisition. Depending on the specific target such as blood vessels after anastomosis, certain pattern recognition algorithms need to be developed in the future to enable automatic scanning.

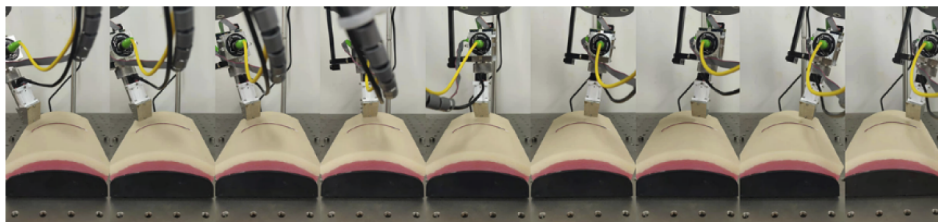


Fig. 13. Scanning process of a curved plastic tube on a skin surface phantom.

Figure 14(a) shows the imaging area with the plastic tube outlined by yellow box. Based on the spatial position information of each scanned point, a large-scale 3D image was obtained after registration. 3D reconstruction of the plastic tube with front view, top view and back view are shown in Fig. 14(b), Fig. 14(c) and Fig. 14(d), respectively. Reconstructed plastic tube dimension can reach a total length of approximate 67.8 mm measured from the software while the lateral FOV is 2.8 mm×4 mm. Figure 14(e) shows the virtual cut of the acquired data presenting the radial profile along the tube. From Fig. 14 we can see that flexible large FOV imaging of a

long-curved sample with surface contour following capability was enabled with robotic arm assistance, which would be difficult for conventional translation stage-based image stitching.

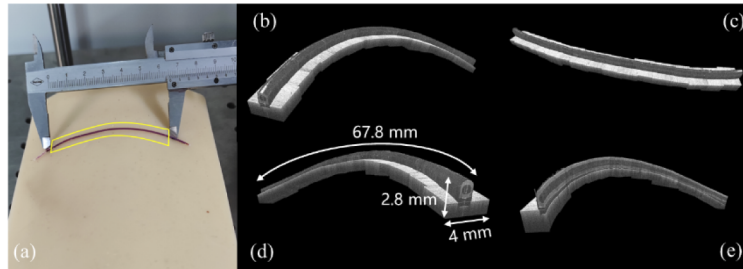


Fig. 14. (a) Photo of the plastic tube. scanning range is outlined by the yellow box; 3D reconstruction of the tube (b): front view, (c) top view, (d) back view and (e) virtual cut profile.

Currently, it took about 30 s for probe pose optimization, which includes initial volume acquisition, automatic optimal pose calculation, volume acquisition after pose adjustment and data saving. Image acquisition of 18 volumes took about 11 minutes including pose optimization and manual scanning point selection operation. Admittedly, relative long data acquisition time is the limitation for our prototype system. We would expect increasing the imaging speed can help reduce the data acquisition time with the tradeoff of image quality. In this study, all the samples we imaged are stationary, which can be satisfied for applications such as industrial inspection and artwork identification. However, when it comes to *in vivo* biomedical imaging, sample motion is inevitable. To address the issues caused by sample motion, both software method such as image pattern recognition and tracking and hardware method by adopting more sensors to compensate the motion effect are necessary in the future.

4. Conclusion

In summary, we demonstrated a prototype robotic-arm-assisted OCT system with flexible scanning capabilities and ultra-large FOV taking the advantages of the stability, precision and flexibility of the robotic arm. To achieve fast targeting, we adopted the depth camera D435 to acquire the point cloud of the sample. The RGB camera integrated in the probe was used to view the target area and perform fine alignment. During the scanning process, a 7-DOF robotic arm was used to drive the probe to follow a manually set or automatically planned scanning path. At each scan point, the probe position can be optimized by feedback from the 3D OCT images when necessary. The main advantage of the system is that there is enough flexibility to obtain the required OCT images for large scale 3D reconstruction without sacrificing resolution. Our robotic-arm-assisted OCT probe produced encouraging results in all tests. Notably, to the best of our knowledge, this is the first demonstration for flexible large FOV OCT imaging enabled by robotic arm. We believe it will open up new opportunities for OCT imaging applications which requires flexible large FOV, remote control and automatic imaging capability.

Funding. National Natural Science Foundation of China (61505006); Beijing Institute of Technology (2018CX01018); Overseas Expertise Introduction Project for Discipline Innovation (B18005); CAST Innovation Foundation (2018QNRC001).

Disclosures. The authors declare no conflicts of interest.

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

References

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafito, "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).

2. C.-L. Chen and R. K. Wang, "Optical coherence tomography based angiography [Invited]," *Biomed. Opt. Express* **8**(2), 1056–1082 (2017).
3. D. Yang, M. Hu, M. Zhang, and Y. Liang, "High-resolution polarization-sensitive optical coherence tomography for zebrafish muscle imaging," *Biomed. Opt. Express* **11**(10), 5618–5632 (2020).
4. F. Romano, S. Parrulli, M. Battaglia Parodi, M. Lupidi, M. Cereda, G. Staurenghi, and A. Invernizzi, "Optical coherence tomography features of the repair tissue following RPE tear and their correlation with visual outcomes," *Sci. Rep.* **11**(1), 5962 (2021).
5. C. Kerbage, H. Lim, W. Sun, M. Mujat, and J. F. de Boer, "Large depth-high resolution full 3D imaging of the anterior segments of the eye using high speed optical frequency domain imaging," *Opt. Express* **15**(12), 7117–7125 (2007).
6. F. LaRocca, D. Nankivil, T. DuBose, C. A. Toth, S. Farsiu, and J. A. Izatt, "In vivo cellular-resolution retinal imaging in infants and children using an ultracompact handheld probe," *Nat. Photonics* **10**(9), 580–584 (2016).
7. E. Götzinger, M. Pircher, B. Baumann, C. Ahlers, W. Geitzenauer, U. Schmidt-Erfurth, and C. K. Hitzenberger, "Three-dimensional polarization sensitive OCT imaging and interactive display of the human retina," *Opt. Express* **17**(5), 4151–4165 (2009).
8. C. Ahlers and U. Schmidt-Erfurth, "Three-dimensional high resolution OCT imaging of macular pathology," *Opt. Express* **17**(5), 4037–4045 (2009).
9. L. An and R. K. Wang, "In vivo volumetric imaging of vascular perfusion within human retina and choroids with optical micro-angiography," *Opt. Express* **16**(15), 11438–11452 (2008).
10. J. C. Burton, S. Wang, C. A. Stewart, R. R. Behringer, and I. V. Larina, "High-resolution three-dimensional in vivo imaging of mouse oviduct using optical coherence tomography," *Biomed. Opt. Express* **6**(7), 2713–2723 (2015).
11. R. Huber, M. Wojtkowski, J. G. Fujimoto, J. Y. Jiang, and A. E. Cable, "Three-dimensional and C-mode OCT imaging with a compact, frequency swept laser source at 1300 nm," *Opt. Express* **13**(26), 10523–10538 (2005).
12. C.-Y. Tsai, C.-H. Shih, H.-S. Chu, Y.-T. Hsieh, S.-L. Huang, and W.-L. Chen, "Submicron spatial resolution optical coherence tomography for visualising the 3D structures of cells cultivated in complex culture systems," *Sci. Rep.* **11**(1), 3492 (2021).
13. G. Liu and Z. Chen, "Capturing the vital vascular fingerprint with optical coherence tomography," *Appl. Opt.* **52**(22), 5473–5477 (2013).
14. N. Laan, R. H. Bremmer, M. C. G. Aalders, and K. G. de Bruin, "Volume determination of fresh and dried bloodstains by means of optical coherence tomography," *J. Forensic Sci.* **59**(1), 34–41 (2014).
15. P. Targowski, M. Iwanicka, L. Tyminska-Widmer, M. Sylwestrzak, and E. A. Kwiatkowska, "Structural examination of easel paintings with optical coherence tomography," *Acc. Chem. Res.* **43**(6), 826–836 (2010).
16. I. Grulkowski, J. J. Liu, B. Potsaid, V. Jayaraman, J. Jiang, J. G. Fujimoto, and A. E. Cable, "High-precision, high-accuracy ultralong-range swept-source optical coherence tomography using vertical cavity surface emitting laser light source," *Opt. Lett.* **38**(5), 673–675 (2013).
17. J. P. Kolb, T. Klein, C. L. Kufner, W. Wieser, A. S. Neubauer, and R. Huber, "Ultra-widefield retinal MHz-OCT imaging with up to 100 degrees viewing angle," *Biomed. Opt. Express* **6**(5), 1534–1552 (2015).
18. I. Grulkowski, J. J. Liu, B. Potsaid, V. Jayaraman, C. D. Lu, J. Jiang, A. E. Cable, J. S. Duker, and J. G. Fujimoto, "Retinal, anterior segment and full eye imaging using ultrahigh speed swept source OCT with vertical-cavity surface emitting lasers," *Biomed. Opt. Express* **3**(11), 2733–2751 (2012).
19. Y. Jia and R. K. Wang, "Label-free in vivo optical imaging of functional microcirculations within meninges and cortex in mice," *J. Neurosci. Methods* **194**(1), 108–115 (2010).
20. J. Xu, W. Wei, S. Song, X. Qi, and R. K. Wang, "Scalable wide-field optical coherence tomography-based angiography for in vivo imaging applications," *Biomed. Opt. Express* **7**(5), 1905–1919 (2016).
21. Z. Wang, B. Potsaid, L. Chen, C. Doerr, H.-C. Lee, T. Nielson, V. Jayaraman, A. E. Cable, E. Swanson, and J. G. Fujimoto, "Cubic meter volume optical coherence tomography," *Optica* **3**(12), 1496–1503 (2016).
22. S. Song, J. Xu, and R. K. Wang, "Long-range and wide field of view optical coherence tomography for in vivo 3D imaging of large volume object based on akinetic programmable swept source," *Biomed. Opt. Express* **7**(11), 4734–4748 (2016).
23. E. Min, J. Lee, A. Vavilin, S. Jung, S. Shin, J. Kim, and W. Jung, "Wide-field optical coherence microscopy of the mouse brain slice," *Opt. Lett.* **40**(19), 4420–4423 (2015).
24. T. Callewaert, J. Guo, G. Hartevelde, A. Vandivere, E. Eisemann, J. Dik, and J. Kalkman, "Multi-scale optical coherence tomography imaging and visualization of Vermeer's Girl with a Pearl Earring," *Opt. Express* **28**(18), 26239–26256 (2020).
25. V. Mazlin, K. Irsch, M. Paques, J.-A. Sahel, M. Fink, and C. A. Boccara, "Curved-field optical coherence tomography: large-field imaging of human corneal cells and nerves," *Optica* **7**(8), 872–880 (2020).
26. B. Krajancich, A. Curatolo, Q. Fang, R. Zilkens, B. F. Dessauvage, C. M. Saunders, and B. F. Kennedy, "Handheld optical palpation of turbid tissue with motion-artifact correction," *Biomed. Opt. Express* **10**(1), 226–241 (2019).
27. C. Freschi, V. Ferrari, F. Melfi, M. Ferrari, F. Mosca, and A. Cuschieri, "Technical review of the da Vinci surgical telemanipulator," *Int J Med Robotics Comput Assist Surg* **9**(4), 396–406 (2013).
28. M. Draelos, P. Ortiz, R. Qian, B. Keller, K. Hauser, A. Kuo, and J. Izatt, "Automatic optical coherence tomography imaging of stationary and moving eyes with a robotically-aligned scanner," in *2019 International Conference on Robotics and Automation (ICRA)*, 8897–8903(2019).

29. F. Rossi, F. Micheletti, G. Magni, R. Pini, L. Menabuoni, F. Leoni, and B. Magnani, "A robotic platform for laser welding of corneal tissue," *Proc.SPIE* 10413, Novel Biophotonics Techniques and Applications IV, 104130B (2017).
30. Y. Sun, Z. Jiang, X. Qi, Y. Hu, B. Li, and J. Zhang, "Robot-assisted decompressive laminectomy planning based on 3D medical image," *IEEE Access* **6**, 22557–22569 (2018).
31. S. Basov, A. Milstein, E. Sulimani, M. Platkov, E. Peretz, M. Rattunde, J. Wagner, U. Netz, A. Katzir, and I. Nisky, "Robot-assisted laser tissue soldering system," *Biomed. Opt. Express* **9**(11), 5635–5644 (2018).
32. P. Robnier Reyes, J. Jamil, and X. D. Y. M. D. Victor, "Intraoperative optical coherence tomography of the cerebral cortex using a 7 degree-of freedom robotic arm," *Proc.SPIE* 10050, Clinical and Translational Neurophotonics, 100500 V (2017).
33. K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**(5), 698–700 (1987).
34. R. Y. Tsai and R. K. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *IEEE Trans. Robot. Automat.* **5**(3), 345–358 (1989).
35. Y. Huang, G. J. Furtmuller, D. Tong, S. Zhu, and W. P. A. Lee, "MEMS-based handheld fourier domain doppler optical coherence tomography for intraoperative microvascular anastomosis imaging," *PLoS One* **9**(12), e114215 (2014).