2024 IEEE International Conference on Robotics and Automation (ICRA 2024)
May 13-17, 2024. Yokohama, Japan

Fast Photoacoustic Microscopy with Robot Controlled Microtrajectory Optimization

Yating Luo†, Yuxuan Liu†, Jiasheng Zhou, Sung-liang Chen, Yao Guo, Guang-Zhong Yang, *Fellow*, *IEEE*

***Abstract*— Photoacoustic Microscopy (PAM) is a relatively new imaging modality in biomedicine. However, point-by-point raster scanning in PAM suffers from low imaging speed. Sparse sampling has been studied in recent years and with the development of deep learning algorithms, extensive efforts have been devoted to sparse image reconstruction while little attention has been paid to sparse sampling trajectory design required for actual implementation. The use of real-time adaptive robotically controlled sampling with micro-scale accuracy with due consideration of physical constraints can pave the way for using PAM for robot-assisted microsurgery. This work proposes a fast PAM scheme with robot-controlled microtrajectory optimization. The proposed method is adaptive to imaging details of different regions of interest (ROI) and detailed experiments have been conducted on both simulation and in-vivo settings. Results show that our proposed method can achieve faster scanning speed than traditional raster scanning and improved image quality in ROI than the standard spiral trajectory, which demonstrates the effectiveness of our proposed method and its potential to be deployed in other point-by-point scanning systems.**

I. Introduction

Photoacoustic microscopy (PAM) is a non-invasive imaging technique that detects distinct acoustic signals generated via the photoacoustic effect [1], [2]. Since the optical absorption coefficient of blood vessels is relatively high compared to surrounding tissues, PAM has been widely used in vascular imaging, including the perception of vascular morphology, hemoglobin oxygen saturation dynamics, and time-dependent blood flow changes [3], [4]. Owing to the confocal imaging mechanism, PAM can achieve superior resolution and sensitivity compared to conventional optical imaging modalities. The predominant method for data acquisition in PAM involves the deployment of stepper motors for systematic point-by-point scanning on the 2-D plane over the target sample [5]. Hence, by adjusting the step size, there exists a trade-off between the imaging speed and resolution. When minimizing the step size to enhance the resolution, it often necessitates a long scanning time which makes it unsuitable for observing dynamic changes in the microstructure. Consequently, how to speed up the scanning process for high-resolution imaging is essential for PAM used for intervention tasks such as neuro-surgery.

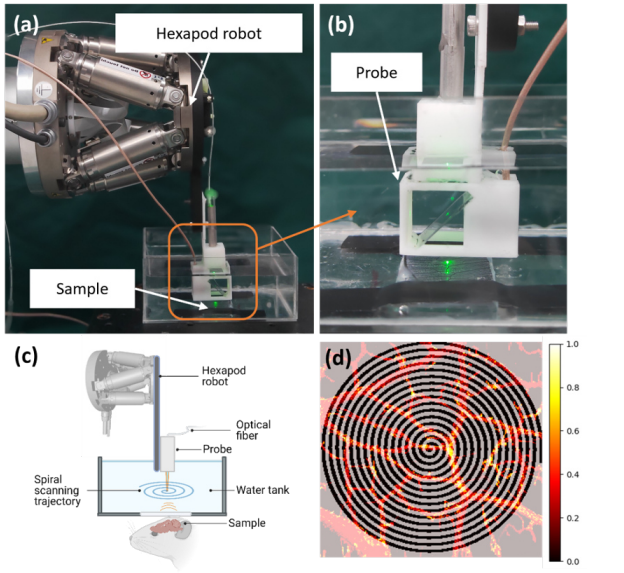

Fig. 1. (a) Overview of our robot-assisted PAM system; (b) a close-up view of the PAM probe; (c) schematic illustration of the fast PAM with a robot-controlled spiral trajectory; (d) an example of the sparse spiral sampling data.

To tackle the challenges of imaging speed, extensive efforts have been dedicated to the development of innovative scanning mechanisms [6]. Progresses included the exploration of alternative scanners like voice coil scanners [7], [8], galvanometer scanners [9], and water-immersible micro-electro-mechanical system (MEMS) scanners [10]–[13]. These advanced hardware solutions have shown success in promoting B-scan frame rates by surpassing the limitations of traditional stepper motor-based systems. However, these scanners are expensive and lack the flexibility to deploy across diverse scenarios [14].

An alternative approach to speed up imaging is to perform sparse sampling with a point-by-point scanner, and the low-rank assumption implies that it is possible to recover the original image from the sparse sampling data using mathematical algorithms like compressed sensing [15]. With recent advancements in deep learning techniques, several methods like convolution neural network (CNN) [14], [16]–[19] have been proposed for sparse data reconstruction since they are powerful and versatile in function modeling. These advanced algorithms provide new opportunities for fast yet high-quality PAM image acquisition.

Thus far, most aforementioned sparse sampling methods focused on the reconstruction algorithms while little attention

This work was supported by Shanghai Municipal Science and Technology Major Project 2021SHZDZX, and also in part supported by the Science and Technology Commission of Shanghai Municipality under Grant 20DZ2220400. †These authors contributed equally to this work.

Y. Luo, Y. Liu, J. Zhou, Y. Guo, and G.-Z. Yang are with the Institute of Medical Robotics, School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China. ({yating.luo, 20000905lyx, jiasheng_zhou, yao.guo, gzyang}@sjtu.edu.cn). S. Chen is with the University of Michigan-Shanghai Jiao Tong University Joint Institute, Shanghai Jiao Tong University, Shanghai, China. (sungliang.chen@sjtu.edu.cn).

Corresponding authors: *Yao Guo, Guang-Zhong Yang*.

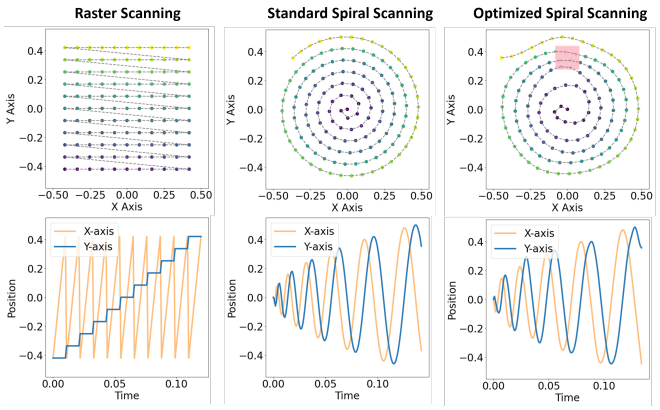

Fig. 2. Comparison of the motion trajectories among different scanning strategies. Each trajectory is shown in cartesian space in the first row and as a function over time in the second row. The user-selected ROI region is masked as a red rectangle.

has been paid to sparse data acquisition which has a dedicated impact on imaging speed. Previous works mainly used a uniformly downsampling strategy to obtain sparse data by post-processing full-scanned images, while they didn't consider whether the sampling trajectory formed by those downsampled points is physically feasible for practical use under hardware constraints. It is therefore necessary to investigate methods that can attain a feasible yet robot-friendly sparse sampling trajectory under velocity and acceleration constraints while maintaining comparable image reconstruction performance. To achieve a fast but smooth sampling trajectory, several works have been introduced in other point-by-point scanning tasks [20], [21] which demonstrated that the spiral trajectory outperforms the raster scanning in terms of continuous velocity and acceleration resulting in improved imaging speed and less image distortion. An illustration of using robots to perform the standard spiral sampling is demonstrated in Fig. 1. However, the standard spiral trajectory uniformly samples over the tissue can lead to low sampling densities in some important regions, e.g., vascular structures, resulting in the poor performance of image reconstruction. Therefore, to improve sampling density in regions of interest (ROI) without increasing sampling points for fast imaging, a non-uniform sampling strategy should be used. How to derive a smooth and feasible trajectory with high sampling densities in those important areas and use those points to recover the original image remains a challenge.

To improve the PAM image reconstruction quality with limited sampling points, in this paper, we propose a neural network-based trajectory optimization algorithm to output a feasible and robot-friendly downsampling trajectory with dense sampling in ROI which provides opportunities for robot-assisted micro-surgery [22]. Following the idea of [23], the sampling trajectory is modeled as an ordinary differential equation (ODE) where a neural network is used to represent the derivative for optimizing [24]. In order to increase the number of sampling points in the user-defined ROI, we design a distance loss function to minimize the distance between the ROI and its neighboring sampling points. Physical constraints are added to the loss term using the soft shrinkage function. An illustration of our optimized spiral trajectory and other sampling trajectories is presented in Fig. 2, where the optimized trajectory is required to sample more points in the ROI which is represented as a red patch. Meanwhile, we implement a deep convolution neural network to reconstruct high-quality images from the generated sampling trajectories and evaluate the performance of the entire system in both simulation and in-vivo settings.

The main contributions of this paper include:

- A neural ODE-based trajectory optimization method for fast PAM is proposed to output a physically applicable sampling trajectory with more sampling points lying in user-defined ROI.
- A reconstruction algorithm is proposed to reconstruct the PAM images based on sparse sampling trajectories.
- Detailed experiments have been conducted in both simulation and in-vivo settings. Quantitative and qualitative results demonstrate that our method achieves comparable sampling speed to standard spiral trajectories but with good imaging quality in the ROI which emphasizes its potential to be deployed in the other point-by-point scanning systems.

II. Methodology

A. Sampling Trajectory Parameterization

A 2-D sampling trajectory parameterized by time index can be described as $\mathbf{s}(t), t \in [1, N]$, where each point $\mathbf{s}(t)$ in the trajectory is a 2-D vector (s_x, s_y) and N is the number of sampling points. Motion velocity and acceleration can be obtained from the first-order and second-order derivatives of the trajectory as $\dot{\mathbf{s}}$ and $\ddot{\mathbf{s}}$. The constant velocity spiral trajectory is described as

$$s_x = A\sqrt{t}\cos(\omega\sqrt{t}), \quad s_y = A\sqrt{t}\sin(\omega\sqrt{t}), \tag{1}$$

where A is the amplitude and ω is the scanning frequency. These two parameters can be adjusted to control the step size and circle spacing based on the resolution requirements like the fast axis step size and slow axis step size in the raster scanning framework. Denote the step size as d_s and circle spacing as d_c, the parameters A and ω can be derived as

$$A = \sqrt{\frac{d_s \cdot d_c}{dt \cdot \pi}}, \quad \omega = \frac{2 \cdot A \cdot \pi}{d_c}, \tag{2}$$

where dt represents the sampling interval.

In the robot-assisted sparse sampling setting, the step size is limited by the motor speed and can not be increased indefinitely, so we will adjust the circle spacing to control the downsampling ratio. However, in the standard constant velocity spiral trajectory, the spacing between adjacent circles is a constant value, resulting in uniformly distributed sampling points in space. In order to improve the sampling efficiency, the circle spacing at different regions should be adjusted adaptively according to the selected ROI.

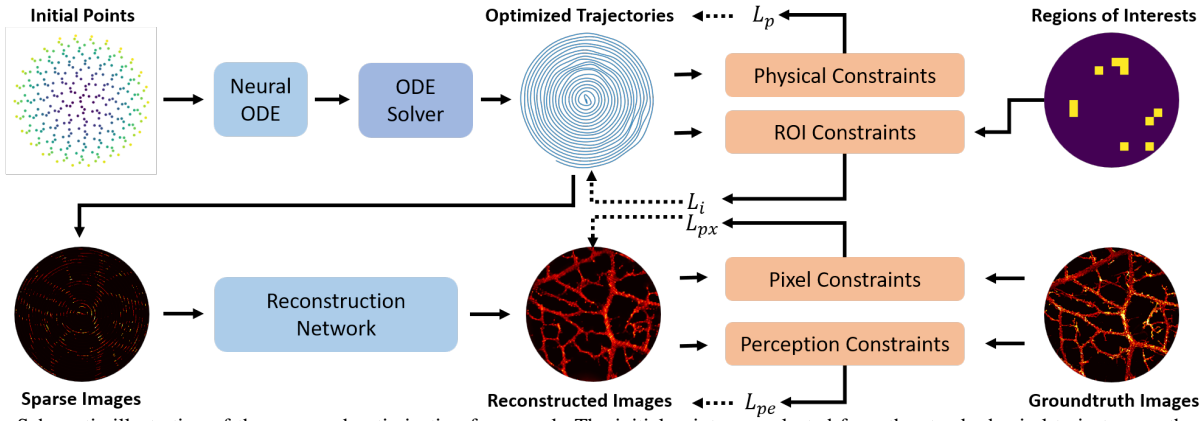

Fig. 3. Schematic illustration of the proposed optimization framework. The initial points are selected from the standard spiral trajectory as the trajectory control points and the derived trajectory is optimized under physical constraints and ROI distance loss. The Optimized trajectory is used to attain sparse data for reconstruction.

B. Neural ODE-based Trajectory Optimization

To solve the issue of varying circle spacing, we utilize a neural ODE model to approximate the constant velocity sampling trajectory and optimize it with both importance loss and physical constraint loss as shown in Fig. 3. As described above, the 2-D sampling trajectory $\mathbf{s}(t)$ can be represented as a matrix of $[N, 2]$. This matrix can be optimized directly as variables or outputted by a fully connected neural network. However, these methods pose a challenge in maintaining trajectory continuity and smoothness. Therefore, in this work, we formulate it into the ODE form to ensure it's differential:

$$\frac{\delta \mathbf{s}(t)}{\delta t} = g(\mathbf{s}(t), t), \tag{3}$$

where $g(\cdot)$ represents the dynamic change of the sampling trajectory which is unknown in our work and needs to be estimated and optimized. With the estimation of the derivative, the sampling trajectory is derived as

$$\mathbf{s}(t) = \mathbf{s}(0) + \int_0^t g(\mathbf{s}(\tau), \tau)\ d\tau, \tag{4}$$

where $\mathbf{s}(0)$ is the initial point of the sampling trajectory and this equation can be solved using conventional ODE solvers like Runge-Kutta.

Inspired by [24], the derivative function g is modeled by a neural network by parameterizing as g_θ since the neural network can be seen as a nonlinear model to approximate any function theoretically. In our work, we aim to optimize g_θ to sample more points in those important regions while the derived trajectory is not only physically feasible but also robot-friendly. The task function is defined as

$$g_{\theta^*} = \arg\min_\theta \mathcal{L}_{traj}(\mathbf{s}(t)), \tag{5}$$

where $\mathbf{s}(t) = ODESolve(\mathbf{s}(0), g_\theta, t)$ and the loss function $\mathcal{L}_{traj}(\cdot)$ is designed for dense sampling in ROI and avoiding exceeding physical constraints.

The ROI can be selected manually by operators or determined automatically according to other sensing modalities. For instance, such as the long-term observation of the dynamics of blood vessels, we can also select regions of poorer imaging quality based on the first frame of the image. The selected ROI is represented by a weight map $\mathcal{W}$ of the same size as the fully sampled image where the values are scaled between 0 and 1 with higher values indicating more importance. During training, we aim to distribute more points to regions with higher weights. Therefore, we propose a weighted importance loss by calculating the distance between each pixel on the plane with a weight greater than 0 and the surrounding K-nearest sampling points. Multiplying the weight of this pixel can force the sampling points to move closer to the pixel with high weights. Denote the K-nearest sampling points around each weight pixel q as $\mathcal{P}(q)$ and the cartesian indices of the weight map as $\mathcal{Q}$. Then the weighted importance loss can be expressed as

$$\mathcal{L}_i = \sum_{p\in\mathcal{P}(q), q\in\mathcal{Q}} w(q) D(p,q) \delta(w(q) > 0), \tag{6}$$

where $w(q)$ represents the weight at pixel q and $D(p, q)$ is the euclidean distance between p and q.

To make the generated trajectories physically applicable, the physical constraints of the robot, such as velocity and acceleration limits, also need to be taken into account when optimizing the trajectories. Given the sampling trajectory $\mathbf{s}(t)$ obtained by the neural ODE, the velocity and acceleration limits of the trajectory can be derived as

$$v_{min} \le \dot{\mathbf{s}} \le v_{max}, \quad a_{min} \le \ddot{\mathbf{s}} \le a_{max}, \tag{7}$$

where the bound of velocity and acceleration is typically symmetric that $v_{max} = -v_{min}$ and $a_{max} = -a_{min}$. Then the physical constraints loss is defined as

$$\mathcal{L}_p = \lambda_1 ||S(\dot{\mathbf{s}}, v_{max})||_2^2 + \lambda_2 ||S(\ddot{\mathbf{s}}, a_{max})||_2^2, \tag{8}$$

where $S(\cdot)$ is the element-wise soft shrinkage function that $S(x, y) = \text{sgn}(x) \cdot \max(0, |x| - y)$.

Then the loss function for sampling trajectory optimization is composed of these two loss functions $\mathcal{L}_i$ and $\mathcal{L}_p$ as

$$\mathcal{L}_{traj} = \mathcal{L}_i + \mathcal{L}_p, \tag{9}$$

To train the derivative estimator g_θ with the loss function $\mathcal{L}_{traj}$, we need to calculate the gradients with respect to θ. However, it's difficult to optimize g_θ using the standard

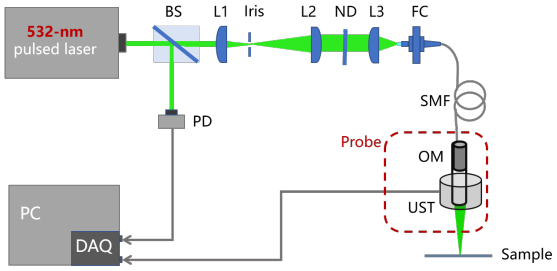

Fig. 4. Schematic illustration of our PAM system. BS, beam splitter; PD, photodiode; ND, neutral density filter; L1, lens #1; L2, lens #2; L3, lens #3; FC, fiber coupler; SMF, single mode fiber; OM, optical module; UST, ultrasound transduser; DAQ, data acquisition card; PC, personal computer.

error back-propagation due to the integration operation in the ODE solver introducing a high memory cost and large numerical error. Hence, we use the adjoint sensitivity method presented in [24] to solve this issue since researchers found it's effective to calculate gradients by solving another ODE equation resulting in less memory costs and numerical error.

C. Deep Networks for Image Restoration

After we obtain the sparse sampling trajectory, we aim to reconstruct the high-quality PAM images from the undersampled images. We define the input undersampled images as $\mathbf{I}_u$, the groundtruth fully sampled images as $\mathbf{I}_{gt}$ and the reconstructed images as $\mathbf{I}_r$. Our designed deep network $\mathcal{F}(\cdot)$ purposes to transform the undersampled images to high-quality images. The task can be defined as:

$$\mathcal{F}_{\theta^*} = \arg\min_{\theta} \mathcal{L}_{img}(\mathcal{F}(\mathbf{I}_u), \mathbf{I}_{gt}), \tag{10}$$

where $\mathcal{L}_{img}(\cdot)$ refers to the training constraints and θ represents the model parameters.

To train our designed network, we generate the image pairs for supervision. Since we have discussed neural ODE-based trajectory optimization, we can generate the sampling trajectories based on the given images and ROI. We transform the sampling trajectories $\mathbf{s}(t)$ to image masks denoted as $\mathbf{M}$ based on the coordinate system transformation, where 1 represents the pixels that fall into the sampling trajectories and 0 represents not. Consequently, we obtain the undersampled images as the input of our designed reconstruction network:

$$\mathbf{I}_u = \mathbf{M} \odot \mathbf{I}_{gt}. \tag{11}$$

The input of our network is the undersampled images $\mathbf{I}_u$ and the output is reconstructed images $\mathbf{I}_r$. The basic blocks in our network are built upon residual convolutional neural blocks with a combination of two group norm layers, two sigmoid layers, and two convolution layers. Inspired by [25], we design the deep networks containing the feature extraction and feature reconstruction processes. The feature extraction part contains four blocks and each block is composed of two residual convolution units. The downsample layer is implemented to reduce the feature maps by half. The feature extraction part extracts the shallow feature maps of the undersampled images. The feature reconstruction part contains the same four blocks as the feature extraction part which aims to reconstruct high-quality images. The upsample layer is implemented to enlarge the feature map. To reduce the information loss during the feature extraction, we implement skip-connection blocks to concatenate different feature maps extracted from different layers.

During the training phase, since we have generated the image pairs $\{\mathbf{I}_u, \mathbf{I}_{gt}\}$, we train the network with two constraints from both pixel and perception level in a supervised manner. The pixel-level loss function $\mathcal{L}_{px}$ refers to the well-known Mean Square Error (MSE) function, which calculates the L2 distance between the predicted images and the ground truth images in image space:

$$\mathcal{L}_{px} = \frac{1}{N}\sum_{n=1}^{N}\sum_{u=1}\sum_{v=1}||I_{gt}^n(u,v) - I_r^n(u,v)||_2^2. \tag{12}$$

However, existing works [14], [17] found that, for PAM image reconstruction, the model trained only on pixel-level loss function tends to be over-smoothing and lacks visual-perceptual quality. Consequently, we introduce the perception-level loss term $\mathcal{L}_{pe}$. We implement the pre-trained VGG model $\mathcal{H}(\cdot)$ to extract latent feature maps of the reconstructed images and the ground truth images, and compose the L2 distance to measure the similarity between these two feature maps in feature space:

$$\mathcal{L}_{pe} = \frac{1}{N}\sum_{n=1}^{N}\sum_{u=1}\sum_{v=1}||\mathcal{H}^n(I_{gt})(u,v) - \mathcal{H}^n(I_r)(u,v)||_2^2, \tag{13}$$

Finally, the combined loss function is the weighted sum of these two terms:

$$\mathcal{L}_{img} = \beta_1\mathcal{L}_{px} + \beta_2\mathcal{L}_{pe}. \tag{14}$$

III. Experiment

A. System Design

In this paper, we utilize a 532-nm pulsed laser (SPOT-10-200-532, Elfolight, UK) with a pulse repetition frequency (PRF) of 1 kHz to excite PA signals which are detected by a custom-made ultrasound transducer (central frequency: 10 MHz; relative bandwidth: 88%). The schematic diagram is illustrated in Fig. 4. The photoacoustic probe is mounted on a 6-axis hexapod robot (H-811.I2 6-Axis Miniature Hexapod, Physik Instrumente) shown in Fig. 1. The maximum velocity of the hexapod robot is 10 mm/s and the maximum acceleration is 116 mm/s^2. The robot's minimum step size is 0.1 μm and its repeatability positioning accuracy is 0.06 μm.

B. Experimental Setting

To evaluate the performance of our proposed methods, we compare our method with raster and standard spiral trajectory in two experimental settings: simulation and real-world scenarios.

1) Simulation scene: For the simulation case, downsampled images are obtained by post-processing fully sampled images, and two public photoacoustic datasets are used here.

- Leaf vein dataset (I): The leaf vein dataset is proposed in [14] which contains 268 PAM images of synthetic

leaf veins in total. We split the datasets into the train, validation, and test sub-datasets following the ratios 7:2:1. There are 187 images for training, 54 images for validation, and 27 images for testing.

- Mouse ear vessel dataset (II): The mouse ear vessel dataset is proposed in [19]. Similar to the first dataset, there are 71 images for training, 20 images for validation, and 10 images for testing.

2) Real-world scene: To further verify that our optimized trajectory is physically executable and is superior to raster scanning in terms of sampling speed, we performed experiments on the mouse brain using our designed PAM system. Based on operator-defined ROI, the optimized trajectory and the standard constant velocity spiral trajectory are also executed to evaluate the sampling efficiency. Three sampling trajectories are designed to scan a $2.048mm \times 2.048mm$ area with nearly the same number of sampling points. The ground truth image is acquired with 25800 sampling points using the standard spiral trajectory(d_s = 8 μm, d_c = 16 μm).

C. Implementation Details

1) Trajectory optimization: The constant velocity spiral trajectory is divided into m segments with equal length L, and the start point of each segment is used as the initial point input to the neural ODE model. Then all the segments generated by the neural ODE model are concatenated for trajectory optimization. To ensure a smoother transition between segments, we utilize the cubic spline interpolation to connect adjacent segments. The proposed model is implemented using PyTorch based on the torchdiffeq [24] package. In our implementation, the segment length is selected as 30 and the nearest neighbor number for importance loss calculation is set as 40. The Adam optimizer is used for error backpropagation and the initial learning rate is 0.012. λ_1 = 0.05 and λ_2 = 0.0005 are set for trajectory optimization.

2) Image reconstruction: The proposed deep neural networks are implemented by PyTorch and trained on an NVIDIA TITAN RTX 3090 GPU. We use Adam optimizer and implement the strategy of learning rate decay. The learning rate is selected as 1e-4 and the batch size is 32. For better performance, we pre-train the image reconstruction network based on the randomly generated trajectories with the set β_1=1.0 and β_2=0.05. Since we aim to make the method generalize well to different sampling strategies, we generate about 100 different sampling trajectories by randomly selecting the important regions, and we pre-train the deep reconstruction network based on these trajectories.

D. Evaluation Metrics

The effectiveness of our optimized trajectories is validated in two aspects. One is the sampling speed evaluated in the in-vivo setting. The other is sampling efficiency which is evaluated by the sampling points number in the selected ROI and reconstruction quality compared to the uniformly constant velocity spiral trajectory. To evaluate the quality of generated images, we implement three metrics for both pixel-level and perception-level comparisons, i.e., Peak Signal Noise Ratio (PSNR↑), Structural Similarity Index Measure (SSIM↑), and Learned Perceptual Image Patch Similarity (LPIPS↓). We report these three metrics for the selected regions of each image for better comparisons of the improvement.

IV. Results And Analysis

A. Results on the Sampling Speed

TABLE I
COMPARISON RESULTS OF SAMPLING SPEED

Sampling Trajectory	Sampling Point Num	Time Cost (s)	Time Cost Per Point (ms)
Raster Scanning	9472	20.414	2.155
Standard Spiral Trajectory	9900	9.082	0.917
Optimized Spiral Trajectory	9600	8.809	0.918

The detailed sampling points number and time costs in the in-vivo setting are reported in Table. I. We found that the time cost per point in the spiral-type trajectory is about 0.918 ms but that is 2.155 ms in the traditional raster scanning trajectory which means the scanning speed has improved by 135% owing to the trajectory smoothness. Additionally, the optimized spiral trajectory improves ROI sampling efficiency while maintaining a similar sampling speed to the standard spiral trajectory.

B. Results on the Image Reconstruction

TABLE II
COMPARISON RESULTS OF PAM IMAGE RECONSTRUCTION

Dataset	Approaches	PSNR↑	SSIM↑	LPIPS↓
I	Optimized Spiral	**15.328**±1.806	**0.536**±0.149	**0.060**±0.027
	Standard Spiral	15.157±1.745	0.493±0.165	0.062±0.028
II	Optimized Spiral	**17.768**±1.914	**0.630**±0.143	**0.071**±0.037
	Standard Spiral	17.610±2.085	0.606±0.156	0.077±0.041

1) Simulation scene: The results of our method on dataset I and II are reported in Table II. We compare our method with the baseline method which refers to the basic spiral scanning trajectories. Note that the sampling points used for different trajectories are the same. As described before, we manually downsample the groundtruth images with the masks we obtain according to different trajectories. It can be found that our method has achieved better performance compared with the spiral scanning method, which demonstrates that by introducing ROI the optimized trajectories can scan more useful information for image reconstruction. We also provide some qualitative results in Fig. 5. It can be found that our optimized trajectories can help the reconstruction network to better focus on the detailed textures.

2) Real-world scene: The results of our method on our collected in-vivo data are visualized in Fig. 6. As shown in the figure, compared to the standard spiral trajectory, the optimized trajectory demonstrates better reconstruction continuity of blood vessels in the ROI.

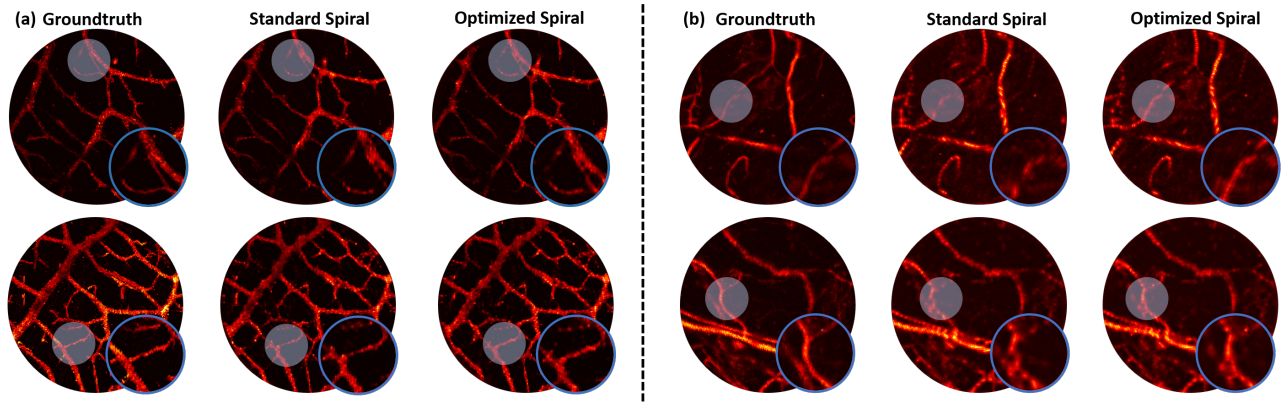

Fig. 5. Results for the simulation images. The left part refers to the representative results in dataset I (a) and the right part refers to the result in dataset II (b). The zoomed images of the blue shaded area are visualized in the lower right part.

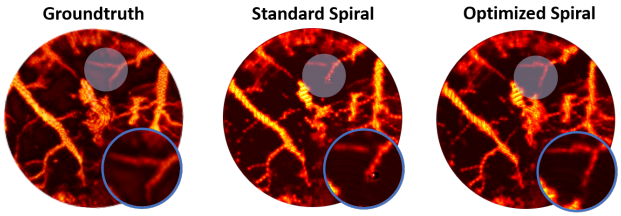

Fig. 6. Results for the real-world in-vivo images. The zoomed images of the blue shaded area are visualized in the lower right part.

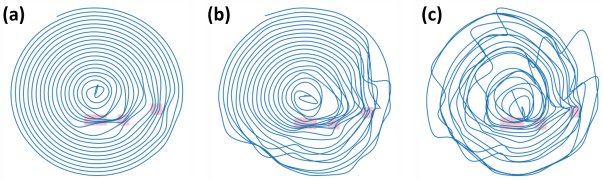

Fig. 8. Results on the influence of physical constraints. (a) the trajectory optimized with both velocity and acceleration constraints. (b) the trajectory optimized with the velocity constraints. (c) the trajectory optimized without any physical constraints. The red rectangles refer to the selected ROI.

C. Results on Sampling Point Number in the Selected ROI

The amounts of valid sampling points in the ROI under different ROI selections are calculated to evaluate the sampling efficiency of the optimized trajectory in important regions. The results are shown in Fig. 7. Valid sampling points refer to points with color in the sampled data since PAM data is sparse with the need for more sampling data in the foreground regions. According to the statistical analysis, it is observed that the optimized trajectory samples more points in the ROI compared to the standard spiral trajectory with a fixed circle spacing, which provides an opportunity for subsequent high-quality reconstruction.

D. Ablation Study of Physical Constraints

The influence of the physical constraints loss in our trajectory optimization is illustrated in Fig. 8 where the optimized trajectory optimized with physical constraints is compared to the other two optimized trajectories optimized without acceleration constraints or with only importance loss. The optimized trajectory without acceleration constraints tends to move with abrupt turns which is not robot-friendly in physical practice. If only the importance loss is used to constrain the trajectory, the overall shape of the generated trajectory is uncontrollable. From the comparison results, we could conclude that physical constraints can effectively force the optimized trajectory to be physically applicable.

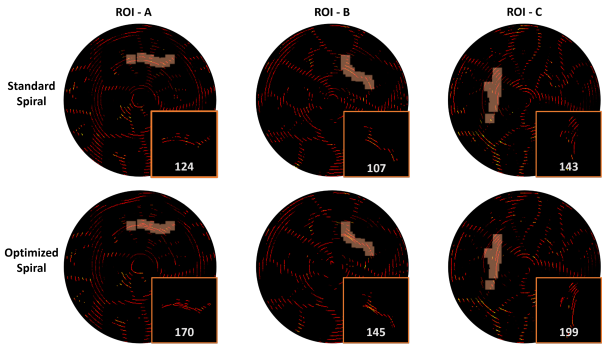

Fig. 7. Results of sampling point number in selected ROIs(orange boxes). The sampling data in the ROI are shown in the lower right corner of each image. The number indicated the quantity of valid samples in the ROI.

V. Conclusions

This work proposes a robot-controlled PAM scanning system for fast photoacoustic microscopy imaging, consisting of a neural ODE-based sampling trajectory optimization module and a downsampled image reconstruction module. Physical constraints and ROIs are introduced to make the sampling trajectories robot-friendly and adaptive to the image details. The reconstruction network then attempts to reconstruct high-quality images based on the optimized trajectories. The current study is from a robot-assisted and real-time perspective, which is different from existing work by retrospectively sparse sampling from densely acquired datasets without considering practical implementation constraints. To fully evaluate our proposed system, we conduct experiments on both simulation cases and real-world cases. Our results on simulation experiments indicate improved performance of our method compared with the baseline method on image reconstruction, and results on the real-world in-vivo experiments demonstrate that our proposed system is robot-friendly and physically applicable. In the future, we will implement our system in more diverse scenarios and improve the reconstruction performance.

REFERENCES

[1] J. Yao and L. V. Wang, "Photoacoustic microscopy," *Laser & photonics reviews*, vol. 7, no. 5, pp. 758–778, 2013.

[2] L. V. Wang, "Multiscale photoacoustic microscopy and computed tomography," *Nature photonics*, vol. 3, no. 9, pp. 503–509, 2009.

[3] S. Hu, K. Maslov, and L. V. Wang, "Second-generation optical-resolution photoacoustic microscopy with improved sensitivity and speed," *Optics letters*, vol. 36, no. 7, pp. 1134–1136, 2011.

[4] L. V. Wang and S. Hu, "Photoacoustic tomography: in vivo imaging from organelles to organs," *science*, vol. 335, no. 6075, pp. 1458–1462, 2012.

[5] S. Jeon, J. Kim, D. Lee, J. W. Baik, and C. Kim, "Review on practical photoacoustic microscopy," *Photoacoustics*, vol. 15, p. 100141, 2019.

[6] K. Wang, C. Li, R. Chen, and J. Shi, "Recent advances in high-speed photoacoustic microscopy," *Photoacoustics*, vol. 24, p. 100294, 2021.

[7] L. Wang, K. Maslov, J. Yao, B. Rao, and L. V. Wang, "Fast voice-coil scanning optical-resolution photoacoustic microscopy," *Optics letters*, vol. 36, no. 2, pp. 139–141, 2011.

[8] L. Wang, K. Maslov, W. Xing, A. Garcia-Uribe, and L. V. Wang, "Video-rate functional photoacoustic microscopy at depths," *Journal of biomedical optics*, vol. 17, no. 10, pp. 106 007–106 007, 2012.

[9] J. Yao, L. Wang, J.-M. Yang, K. I. Maslov, T. T. Wong, L. Li, C.-H. Huang, J. Zou, and L. V. Wang, "High-speed label-free functional photoacoustic microscopy of mouse brain in action," *Nature methods*, vol. 12, no. 5, pp. 407–410, 2015.

[10] J. Yao, C.-H. Huang, L. Wang, J.-M. Yang, L. Gao, K. I. Maslov, J. Zou, and L. V. Wang, "Wide-field fast-scanning photoacoustic microscopy based on a water-immersible mems scanning mirror," *Journal of biomedical optics*, vol. 17, no. 8, pp. 080 505–080 505, 2012.

[11] J. Y. Kim, C. Lee, K. Park, G. Lim, and C. Kim, "Fast optical-resolution photoacoustic microscopy using a 2-axis water-proofing mems scanner," *Scientific reports*, vol. 5, no. 1, p. 7932, 2015.

[12] J. W. Baik, J. Y. Kim, S. Cho, S. Choi, J. Kim, and C. Kim, "Super wide-field photoacoustic microscopy of animals and humans in vivo," *IEEE transactions on medical imaging*, vol. 39, no. 4, pp. 975–984, 2019.

[13] L. Lin, P. Zhang, S. Xu, J. Shi, L. Li, J. Yao, L. Wang, J. Zou, and L. V. Wang, "Handheld optical-resolution photoacoustic microscopy," *Journal of biomedical optics*, vol. 22, no. 4, pp. 041 002–041 002, 2017.

[14] J. Zhou, D. He, X. Shang, Z. Guo, S.-L. Chen, and J. Luo, "Photoacoustic microscopy with sparse data by convolutional neural networks," *Photoacoustics*, vol. 22, p. 100242, 2021.

[15] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[16] A. DiSpirito, D. Li, T. Vu, M. Chen, D. Zhang, J. Luo, R. Horstmeyer, and J. Yao, "Reconstructing undersampled photoacoustic microscopy images using deep learning," *IEEE transactions on medical imaging*, vol. 40, no. 2, pp. 562–570, 2020.

[17] Y. Liu, J. Zhou, Y. Luo, J. Li, S.-L. Chen, Y. Guo, and G.-Z. Yang, "Upamnet: A unified network with deep knowledge priors for photoacoustic microscopy," 2023.

[18] H. Zhao, Z. Ke, F. Yang, K. Li, N. Chen, L. Song, C. Zheng, D. Liang, and C. Liu, "Deep learning enables superior photoacoustic imaging at ultralow laser dosages," *Advanced Science*, vol. 8, no. 3, p. 2003097, 2021.

[19] D. He, J. Zhou, X. Shang, X. Tang, J. Luo, and S.-L. Chen, "Denoising of photoacoustic microscopy images by attentive generative adversarial network," *IEEE Transactions on Medical Imaging*, 2022.

[20] X. Sang, A. R. Lupini, R. R. Unocic, M. Chi, A. Y. Borisevich, S. V. Kalinin, E. Endeve, R. K. Archibald, and S. Jesse, "Dynamic scan control in stem: Spiral scans," *Advanced Structural and Chemical Imaging*, vol. 2, pp. 1–8, 2016.

[21] K. P. Kelley, M. Ziatdinov, L. Collins, M. A. Susner, R. K. Vasudevan, N. Balke, S. V. Kalinin, and S. Jesse, "Fast scanning probe microscopy via machine learning: non-rectangular scans with compressed sensing and gaussian process optimization," *Small*, vol. 16, no. 37, p. 2002878, 2020.

[22] Y. Guo, W. Chen, J. Zhao, and G.-Z. Yang, "Medical robotics: opportunities in china," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 361–383, 2022.

[23] W. Peng, L. Feng, G. Zhao, and F. Liu, "Learning optimal k-space acquisition and reconstruction using physics-informed neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20 794–20 803.

[24] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud, "Neural ordinary differential equations," *Advances in neural information processing systems*, vol. 31, 2018.

[25] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.