Robot-assisted steady ultrasound imaging and servoing enabled by deep learning

CIS 2, Group 7: Tian Xie

Mentors: Drs. Mahya Shabazi, Emad Boctor & Russell Taylor

1. Objectives

Biopsy is a medical procedure for the diagnosis of suspicious masses. Sample cells or tissues are extracted from the lesion for examination under microscope or for chemical analysis to indicate the type of tumor (benign vs. malignant). During a biopsy procedure, ultrasound (US) imaging is used to visualize the lesion and navigate the needle to the target.

During a pathologist-in-room session, which is a well-established practice at most hospitals, it is required to repeat the biopsy procedures several times to ensure that the collected samples are informative enough for making an accurate diagnosis. At each time, the acquired sample cells are evaluated under a microscope by a pathologist in the room, before the next sample cell are acquired. During the multi-sample acquisition procedure, it is very essential to provide the radiologist with a consistent view of the lesion to enable them to acquire distinct samples from the lesion. Keeping a consistent view of the lesion, despite the patient's physiological movement, is a very tedious and cumbersome task for the sonographer, who has to keep the ultrasound probe in place for the whole duration of the biopsy session (which could even last more than half an hour).

Therefore, the goal of this project is to realize a robot-assisted visual servoing platform that can track a desired ultrasound view for the purpose of repeatable biopsy. The system will help make the repeatable biopsy procedure more efficient and convenient for the radiologists and the sonographers.

2. Background and significance

As a safe and cheap source of imaging, ultrasound is always implemented to guide needle biopsy procedures. The ultrasound probe is placed at a certain position to provide a visualization of the lesion. With the ultrasound image, the sonographer can then insert and advance the biopsy needle to the target lesion for acquiring sample cells or tissues. At the same time, the sample acquired must be evaluated by experienced pathology technicians to verify its effectiveness. If the sample is not diagnosable, the radiologist should repeat the procedures for the acquisition of more pathology samples, and the ultrasound probe should be placed to the certain slice again to view the lesion.

Currently, the ultrasound-guided biopsy is conducted all manually by sonographers. The sonographer has to hold the ultrasound transducer throughout several passes to maintain consistent visualization of the lesion or mass. It is tiring and almost impossible to hold the probe still and retain the slice for around ten to twenty minutes. Once moved, it will be time consuming to regain the view of the target. Therefore, it will be of great significance if the repeatable biopsy can be improved by the introduction of robotic systems for freehand locating and navigation.

In order to track the target image, it is essential to estimate both in-plane and out-of-plane motion of the probe. In existing literature, the in-plane motion can be estimated by conventional

2D image registration (Prager et al, 2007). And the out-of-plane motion can be estimated by speckle correlation. Because the focus of the ultrasound probe is poor along the elevational axis, the resolution cell elongates along this direction. The interference of the scatters in the cell form "noise" in the image, which is the speckle. And there is correlation among the speckles if two images cover the same resolution cells. Therefore, conventional methods (G.Treece et al, 2002) use the inter-patch correlation to obtain a plot of speckle decorrelation vs elevational distance at calibration stage, and then estimate the elevational distance by looking at this plot.

Today, with the development of deep learning, it provides the possibility to model the distance of the neighboring images in the robot frame via the neural network (NN). And then the NN can be augmented into the control loop, and finally let the robot locate and navigate to the pre-specified slice. In this project, we will first develop the testbed and acquire data on multiple phantoms and/or animal organs. The collected data will then be used to train the NN to model the motion based on correlation between two neighboring images. Finally, the NN will be added into the control loop for motion compensation, and the system will be evaluated on different organs.

3. Technical summary

Development of the (robotic) testbed

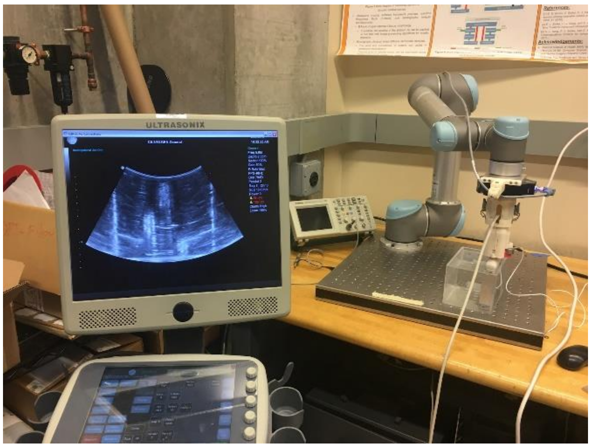


Figure 1 Testbed with UR5

Given the limited range (about 1mm) that speckle decorrelation is accurate and effective, the resolution of the acquired data should be around 25 microns. Therefore, a Cartesian stage with a dial gauge will be used to accurately adjust and measure the distance. As a backup, a UR5 robot arm can also be used to collect data, only the smallest step size is 0.2mm. The probe holder attached to the end effector of UR5 was developed in a previous CIS 2 project.

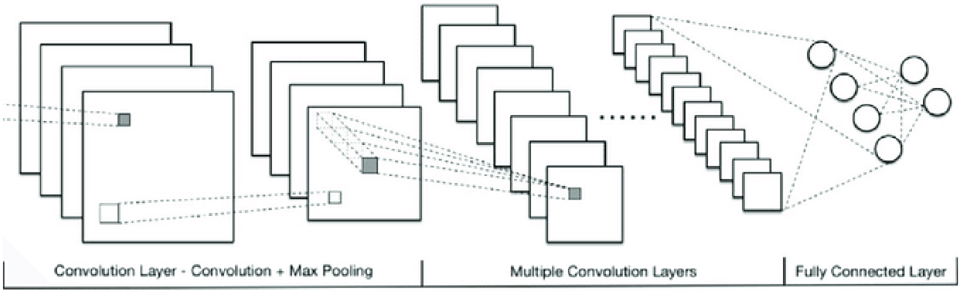
Ultrasonix ultrasound system and a linear array transducer is used to get B-mode images. The image depth is set to 4.0cm and the sampling frequency is 10MHz. The gain will be adjusted accordingly and then kept unchanged to obtain a clear view.

In order to make use of speckle decorrelation, the images acquired will have a resolution of 0.02mm along the elevational direction. The step size could be larger for axial and lateral translation. The range of translational motion will be about 1 cm. For rotations, the resolution will be 0.2 degrees and the range will be 2 degree. Data collection will start with 1DoF and ultimately upgrade to 6DoF.

<u>CNN to find the correlation in two neighboring images</u>

Convolutional neural network (CNN) is widely used in image classification, image and video recognition and medical image analysis. The complexity of CNN may help better estimate the elevational distance. In this project, the initial architecture will be based on a previous work (Prevost et al, 2017). It is a standard CNN with two input channels (each for one image), four convolutional (Conv2D) layers and one fully connected (FC) layer. The output is the six parameters used to represent rigid body transformation. However, their architecture may not match our objectives. Therefore, tuning the hyperparameters and changing the architecture both worth trying.

Mean absolute error and mean absolute percentage are the two parameters used to evaluate the effectiveness of the neural network. Comparison with some current state-of-art methods will also be helpful for evaluation.



https://www.google.com/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwi0ntuxgvHgAhVFNd8KHbbTD7AQjRx6BAgBEAU&url=https%3A%2F%2Fwww.

Figure 2 Possible CNN architecture

<u>Visual servoing</u>

After the estimation of both in-plane and out-of-plane motion, a hybrid control scheme can be developed for visual servoing. The paper (Krupa et al, 2009) will be a good reference for visual servoing. And the detail of this part will be settled after the development of NN.

4. List of dependencies and alternatives

Dependency	Solution	Alternatives	Due Date	If not met	Status
UR5	Access to the lab	N/A	Feb 14	Delay of the project	Resolved
Phantoms/ animal organs	In the lab	N/A	Feb 20	/	Resolved
Simulated Data	Provide by Dr. Marius	N/A			
Ultrasonix Ultrasound system and software	In the lab. And software downloaded from Baichuan.	N/A	Feb 14	Delay of the project	Resolved
Computation	GPU or Google	Training	Feb 26	/	Resolved

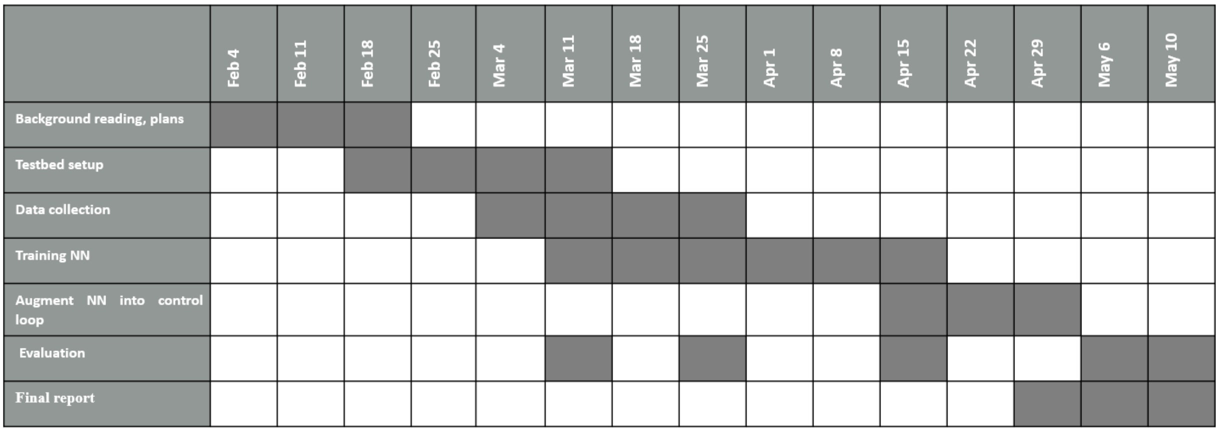
power	Cloud Engine	with CPU			
Optical tracker	The rest can be borrowed.	/	Resolve this only when it is needed.	/	
Motorized Linear stage	Borrow one	Manual linear stage	Feb 28	Delay of data collection	Not yet
Dial Gauge	Buy one (Dr. Boctor/ Dr. Taylor); Borrow?		Mar 8	Delay of data collection	Not yet

5. **Deliverables**

- Minimum: development of robotic testbed and initial data acquisition on multiple phantoms
- Expected: development and evaluation of the NN to accurately model in-plane and out-of-plane motions based on correlations between neighboring images
- Maximum: augmenting the NN into the control loop of the robot for motion compensation and evaluating the system on different types of organs
- Follow-on: finding the location of and navigating to a pre-specified slice, given some nearby slices

6. **Schedule**

Detailed Gantt chart is provided in an additional file on the web page. See Gantt Chart.xlsx.



7. **Management Plan**

1. Weekly meeting with mentors (every Wednesday and Friday);
2. Codes will be stored in GitHub;
3. Ultrasound images will be uploaded to JH Box.

8. **List of readings**

<u>Speckle decorrelation</u>

R. Prevost, M. Salehi, J. Sprung, R. Bauer, & W. Wein. (2017). Deep Learning for Sensorless 3D Freehand Ultrasound Imaging. Medical Image Computing and Computer-Assisted Intervention − MICCAI 2017. MICCAI 2017.

C. Laporte, & T. Arbel. (2009). Learning a tissue invariant ultrasound speckle decorrelation model. Paper presented at the 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 995-998. doi:10.1109/ISBI.2009.5193222

Krupa, Alexandre & Fichtinger, Gabor & Hager, Gregory. (2009). Real-time Motion Stabilization with B-mode Ultrasound Using Image Speckle Information and Visual Servoing. I. J. Robotic Res.. 28. 1334-1354. 10.1177/0278364909104066.

H. Rivaz, R. Zellars, G. Hager, G. Fichtinger, & E. Boctor. (2007). Beam steering approach for speckle characterization and out-of-plane motion estimation in real tissue. Paper presented at the 2007 IEEE Ultrasonics Symposium Proceedings, 781-784. doi:10.1109/ULTSYM.2007.200

R. J. Housden, A. H. Gee, G. M. Treece, & R. W. Prager. (2007). Sensorless reconstruction of unconstrained freehand 3D ultrasound datadoi://doi.org/10.1016/j.ultrasmedbio.2006.09.015

G. M. Treece, R. W. Prager, A. H. Gee, & L. Berman. (2002). Correction of probe pressure artifacts in freehand 3D ultrasound doi://doi-org.proxy1.library.jhu.edu/10.1016/S1361-8415(02)00080-4

Deep learning

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & Lecun, Y. (2014). Overfeat: Integrated recognition, localization and detection using convolutional networks. In International Conference on Learning Representations (ICLR2014), CBLS, April 2014 [http://openreview.net/document/d332e77d-459a-4af8-b3ed-55ba, http://arxiv.org/abs/1312.6229]

N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, & J. Liang. (2016). Convolutional neural networks for medical image analysis: Full training or fine tuning?IEEE Transactions on Medical Imaging, 35(5), 1299-1312. doi:10.1109/TMI.2016.2535302

Visual Servoing

W. J. Wilson, C. C. Williams Hulls, & G. S. Bell. (1996). Relative end-effector control using cartesian position based visual servoing. IEEE Transactions on Robotics and Automation, 12(5), 684-696. doi:10.1109/70.538974

H. Rivaz, R. Zellars, G. Hager, G. Fichtinger, & E. Boctor. (2007). Beam steering approach for speckle characterization and out-of-plane motion estimation in real tissue. Paper presented at the 2007 IEEE Ultrasonics Symposium Proceedings, 781-784. doi:10.1109/ULTSYM.2007.200