Vision-Aided Through-Obstruction Handheld Imaging on 5G Smart Devices

Jacqueline M. Schellberg, *Student Member, IEEE*, and Sanjib Sur, *Member, IEEE*

***Abstract*—Traditional millimeter-wave imaging algorithms require millimeter-scale self-tracking, which handheld 5G smart devices cannot reliably achieve. We propose to generate accurate poses using the known relationship between the mmWave signals and the device antenna spacing. Since the reconstructed mmWave image may still appear degraded due to specularity or weak reflectivity, we implement a Conditional Generative Adversarial Network to generate human-perceptible 2D images.**

***Index Terms*—Synthetic Aperture Radar Imaging, millimeter-wave**

I. Introduction

Mobile 5G smart devices have expanded sensing possibilities for indoor localization and virtual reality, but these devices cannot sense beyond visual obstructions. 5G smart devices include millimeter-wave (mmWave) transceivers that could be used for handheld through-obstruction imaging, enabling several applications. (1) *Security*: A handheld imager could eliminate congestion in airports while maintaining robust security practices. (2) *Disaster relief*: The device could aid first responders to image through debris in hard-to-reach locations. (3) *Inventory management*: Through-obstruction imaging could account objects beyond packaging without contact, limiting packaging waste.

Traditional mmWave imaging systems, such as airport security scanners, use the Synthetic Aperture Radar (SAR) imaging technique to generate detailed through-obstruction images. SAR coherently combines signals transmitted and received at different spatial locations to reconstruct a through-obstruction image. Traditional SAR systems rely on expensive, bulky motion controllers to enable millimeter-scale tracking. However, applying the SAR concept to handheld 5G smart devices has been challenging, since 5G smart devices cannot generate accurate poses for SAR focusing. What's more, traditional SAR autofocus algorithms [1] cannot correct the pose error beyond $\frac{1}{4}^{th}$ of the system wavelength. For a mmWave device with a wavelength of 3.90 mm, the positioning error must be less than 0.98 mm, but 5G smart devices are only capable of centimeter scale self-tracking. The pose error would cause the mmWave reflections to sum destructively during image reconstruction, making the resultant output image imperceptible.

In this work, we design a system that corrects pose errors to enable improved handheld mmWave imaging. The system uses the mmWave signals themselves to provide a robust motion estimate and further improves image quality using machine learning.

J.M. Schellberg is a Computer Engineering B.Se student at the University of South Carolina, Columbia, SC 29208 U.S.A. (schellbj@email.sc.edu)

S. Sur is an Assistant Professor at the University of South Carolina, Columbia, SC 29208 U.S.A (sur@cse.sc.edu)

II. Design

A. Pose Correction

To generate a focused image, SAR requires precise motion tracking. Figure 3(b) shows a SAR image reconstructed of the optical image of the object in Figure 3(a) using the pose output from a vision-based self-tracking device. The pose error causes the mmWave signals to add destructively, and the output image appears heavily distorted. Thus, we must correct the poses to enable high-quality mmWave imaging. To obtain an improved initial pose estimate, we recover the device velocity using the antenna spacings on the mmWave device, similar to [2]. We obtain signal reflections from two Rx antennas sharing a transmit antenna, Rx1 and Rx2, separated by a distance H along the direction of motion. We cross-correlate each frame from Rx1 with surrounding received frames from Rx2. For a system capturing mmWave frames at framerate F with an antenna spacing H, the delay D is the number of frames at which Rx1 experiences the same response as Rx2. Then, the velocity V_i and the position X_i of the device at the ith frame is:

$$V_i = \frac{HF_i}{2D_i}; \quad X_{i+1} = X_i + \frac{V_i}{F} \tag{1}$$

However, since there is no depth-axis antenna spacing on the device that can be used for velocity estimation, the Z (depth) error accumulates and causes significant defocusing.

To overcome the depth-axis limitation, we divide each pose trajectory into several overlapping segments and apply a median filter to the Z-poses. From these segments, we reconstruct small 3D mmWave sub-images in the overlapping region. Since the poses within the overlapping region share the same mmWave reflections but have different poses due to drift, we register the sub-images to recover the offset between the separate pose segments. Since each mmWave signal is too sparse to register individual frames, we correct the poses following the method in our poster [3]. Now, the image can be reconstructed using the corrected poses to produce a coherent mmWave image, as shown in Figure 3 (c).

B. Deep-Learning Based Image Quality Improvement

The image may appear degraded due the fundamental limits of mmWave, such as specularity and weak reflectivity, that cause transmitted signals to never return to the receiver. We use a Conditional Generative Adversarial Network (cGAN) to recover high-frequency details that could not be obtained through pose correction, and implement it following [4].

C. Data Collection and Post-Processing

We implement a custom setup that mimics a realistic handheld setup integrating a vision-based self-tracking module and a mmWave device. We use the Intel RealSense T265 tracking camera [5] since the pose accuracy is similar to commodity self-localizing smartphones. We use the TI IWR1443 mmWave radar [6] operating from 77-81 GHz to collect the mmWave reflections with a depth resolution of 3.75 cm. The IWR1443 has 3 Tx and 4 Rx antennas, totaling 12 channels that can be used for velocity estimation. The T265 and the mmWave device cannot be triggered synchronously, and incorrect synchronization would cause the mmWave signals to combine destructively and produce a distorted image. To synchronize the data, we post-process the received data to associate each received mmWave signal to the correct pose following [3].

III. Results

Since obtaining sub-millimeter accuracy ground-truth with actual handheld motions is difficult, we use a mechanical controller to automatically move the device along a pre-defined 18×18 cm^2 grid (Figure 1).

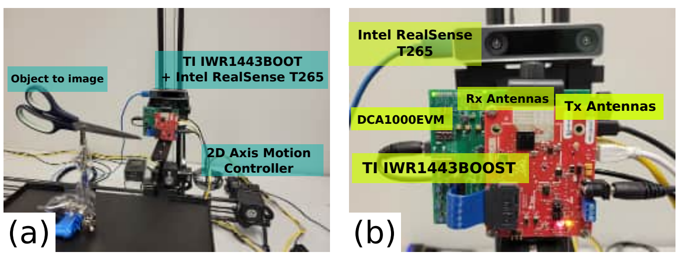

Figure 1. (a) 2D axis controller imaging setup; (b) T265 and TI IWR1443

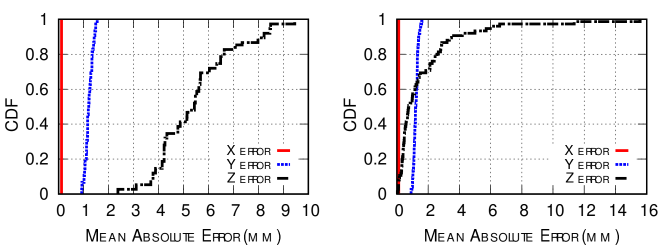

Figure 2. (a) MSE using raw T265 poses; (b) MSE using after pose correction

We have tested the accuracy of the pose correction by comparing the mean squared error (MSE) of the poses against the ground truth trajectory of the mechanical controller in Figure 2. Additionally, we show that the pose correction recovers coarse structure in Figure 3.

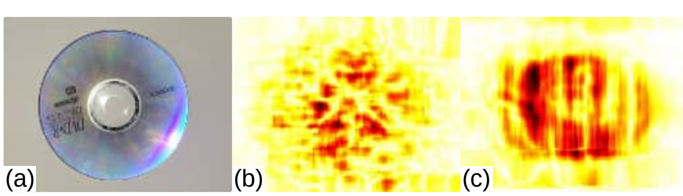

Figure 3. (a) Optical image of CD; (b) mmWave image generated using raw poses from self-tracking device; (c) mmWave image after performing pose-correction

Then, we test the cGAN by comparing the Structural Similarity Index Measure (SSIM) score [7], which measures the similarity between a ground truth image, with 0 being least similar and 1 being most similar. Our system improves the SSIM from .01 to .92, and the improvements can be seen in Figure 4

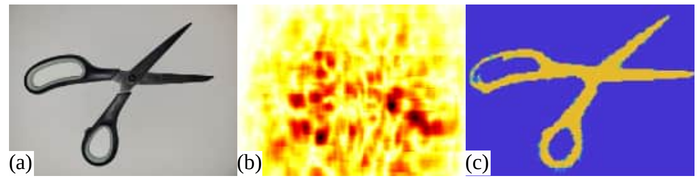

Figure 4. (a) Optical image of scissors; (b) mmWave image generated using raw poses from self-tracking device; (c) 2D image generated by cGAN

IV. Conclusion

In this work, we design and implement a system that enables handheld mmWave imaging on 5G smart devices. We use the antenna spacing and mmWave reflections to produce an accurate motion estimate. Then, we correct the global drift error by registering overlapping sub-images generated using the locally accurate poses. Finally, we use a generative machine learning framework to produce human-perceptible 2D images.

V. Acknowledgements

I am thankful to have been selected for the IEEE MTT-S Undergraduate/Pre-graduate scholarship, which has tremendously supported my efforts throughout this project. I had the opportunity to attend IMS2022 in Denver, Colorado, where I engaged with the RF community and learned from several interesting researchers. I am thankful to my supervisor, Dr. Sanjib Sur, for the guidance he provided me at the University of South Carolina. I will graduate with a bachelor's degree in computer engineering in May 2023, and I plan to attend graduate school to further pursue research related to wireless sensing.

References

[1] Morrison, Robert L and Do, Minh N and Munson, David C, "SAR image autofocus by sharpness optimization: A theoretical study," *IEEE Transactions on Image Processing*, vol. 16, no. 9, 2007.

[2] Qian, Kun and He, Zhaoyuan and Zhang, Xinyu, "3D Point Cloud Generation with Millimeter-Wave Radar," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 4, 2020.

[3] J. M. Schellberg and S. Sur, "Accurate Device Self-Tracking for Robust Millimeter-Wave Imaging on Handheld Smart Devices," in *ACM MobiSys*, 2022.

[4] Regmi, Hem and Saadat, Moh Sabbir and Sur, Sanjib and Nelakuditi, Srihari, "SquiggleMilli: Approximating SAR Imaging on Mobile Millimeter-Wave Devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 3, 2021.

[5] "Intel RealSense Tracking Camera T265." [Online]. Available: https://www.intelrealsense.com/tracking-camera-t265/

[6] Texas Instruments, "IWR1443 Single-Chip 76-to-81GHz mmWave Sensor Evaluation Module," 2018. [Online]. Available: http://www.ti.com/tool/IWR1443BOOST

[7] Wang, Zhou and Bovik, Alan C and Sheikh, Hamid R and Simoncelli, Eero P, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, 2004.