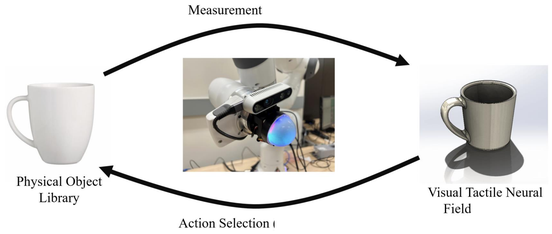
Visual Tactile Neural Fields for Active Digital Twin Generation

Lead PI: Philip Dames, Temple University
PI: Monroe Kennedy, Stanford University
PI: Kostas Daniilidis, University of Pennsylvania

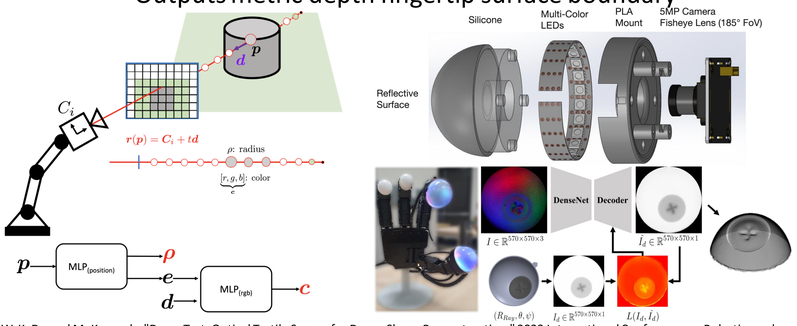
Motivation and Broader Impact

- Goals
  - Model visual and physical characteristics of individual objects
  - Predict expected interaction between an object and a robot/sensors
- Uses
  - Create digital twins, e.g., archiving, tailored interaction
  - Downstream tasks, e.g., dexterous manipulation



Neural Fields and DenseTact Design

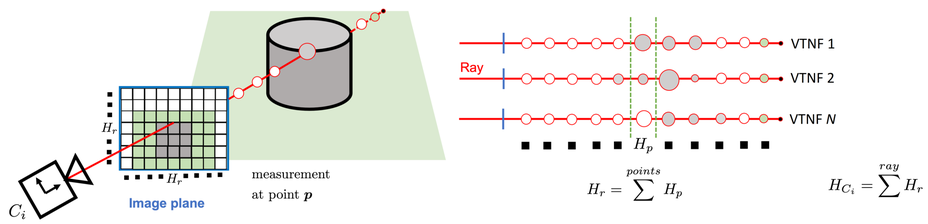
- Neural Fields
  - Use neural network to characterize fields in volumetric space (e.g., density, color)
  - State-of-the-art method in vision-based scene reconstruction
  - We create the Visual Tactile Neural Field (VTNF)
    - Unified representation that can be updated using vision and touch
- DenseTact
  - Optical-tactile sensor
  - Outputs metric depth fingertip surface boundary



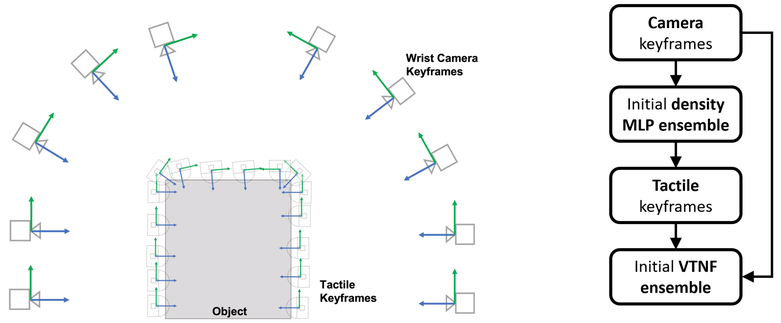
W. K. Do and M. Kennedy, "DenseTact: Optical Tactile Sensor for Dense Shape Reconstruction," 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 2022, pp. 6188-6194, doi: 10.1109/ICRA46639.2022.9811966.

Proposed Methodology

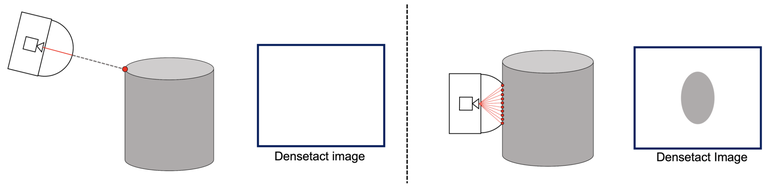
- Thrust 1
  - Generate a VTNF model from sensor data
    - Mathematical models
    - Neural network architecture
    - Algorithmic pipeline
  - Sensor package
    - DenseTact
      - calibrated using library of known surfaces
      - Use surface measurements (over time) to extract shape, texture, friction, etc.
    - RGB camera
    - Combine DenseTact with color imagery to extract other material properties
- Thrust 2
  - Optimize VTNF creation
    - Minimize number of measurements
    - Select sequence of senses and sensing locations
  - Characterize uncertainty in current model
    - Use ensemble of VTNFs to characterize entropy
    - Use entropy to select next best sensing sequence for vision and touch



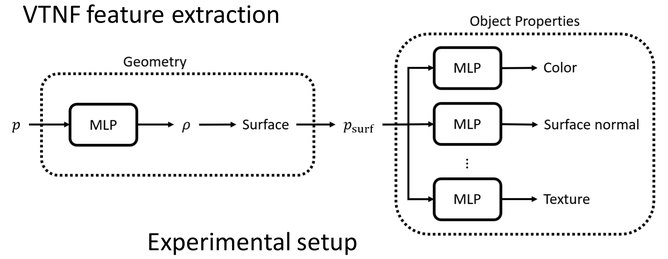
VTNF ensemble for entropy guided sensing



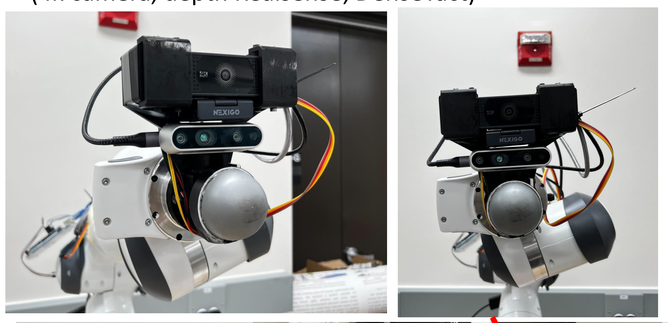
DenseTact depth sensing example



VTNF feature extraction



Experimental setup
(4k camera, depth Realsense, DenseTact)



Award ID#: 2220866 (Temple), 2220867 (Stanford), 2220868 (Penn)