Learning Material Property from Video-Tactus Association

Chenyang Ma (Honors Capstone), Fengyu Yang, and Andrew Owens
Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, United States

ENGINEERING HONORS PROGRAM
UNIVERSITY OF MICHIGAN

Introduction

- In this research project, we try to retrieve touching information using “GelSight”, a specially designed tactile sensor to learn the material properties of objects and combine them with visual information to perform interesting tasks.

- More specifically, we present a method for manipulating the appearance of an object to match its material property, a problem we term **tactile-driven image stylization**.

- Our model learns to restyle images through self-supervised learning. Given a dataset of paired visual and tactile data, the model learns to manipulate the image such that, after manipulation, the object in the image transforms its material property (e.g., make an object harder, let grass grows on a stone).

Contributions

1. We collected **the first** egocentric dataset of video + tactile data. Our dataset contains 2 hours of video, 20 objects, 1000 touches with the tactile sensor. We are still expanding our dataset.

2. We applied our dataset to the novel problem of visual-tactile stylization. We developed a deep learning model based on CUT [1] which is suitable to this task.

3. To prove the robustness of our dataset and model, we trained them to conduct stylization on RGB images. We empirically found that changing the tactile signal and mixture results in predictable visual changes. We are designing an evaluation metric to quantitively measure the accuracy of our style transfer.

4. We are also working on a novel multisensory video prediction problem. Given a image + tactile signal, we want to predict the next tactile signal (e.g., if you are pressing a soft object, then the tactile signal should show more deformation than a hard object), or predict the next visual image (e.g., if you stop pressing, your hand is likely to move off the object).

Selected References

[1] Taesung Park, Alexei A. Efros, Richard Zhang, and Jun-Yan Zhu. Contrastive Learning for Unpaired Image-to-Image Translation. *Proceedings of European Conference on Computer Vision*, 2020.

Acknowledgment

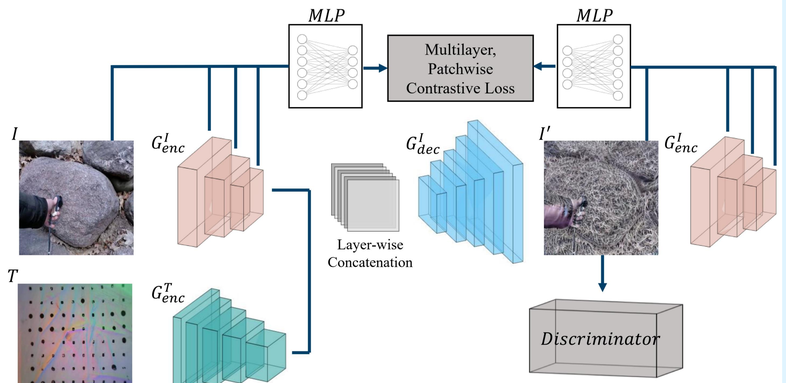
Special thanks to Prof. Wenzhen Yuan from Carnegie Mellon University for providing GelSight for this research.

Methodology

Sensors

Data Collection Set-up

Deep Neural Network Architecture / Pipeline



Results & Visualizations