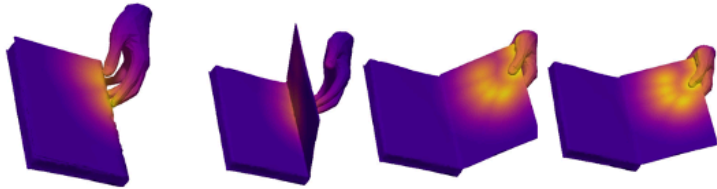# Recovering 3D Hands in Interaction from Images (<mark>NAVER+UvA Internship</mark>)

Description

Humans grasp and manipulate objects with their hands to perform everyday tasks. To enable computers and intelligent systems (such as robots) to understand our actions, we need algorithms for 3D visual perception, i.e., for reconstructing hands and objects in 3D from camera images.

However, this is a challenging task; 3D inference from 2D images is highly ill-posed, hands tend to be small and blurry in images, and the hand heavily occludes the object (and vice versa) causing missing visual information. Consequently, this is a hot research topic for both academia and the top-tech industry. For example, see the HANDS workshop (https://hands-workshop.org) in the upcoming top-tier ECCV conference. The following recent datasets provide excellent resources for progress in the field (in alphabetic order):
- GRAB dataset: https://grab.is.tue.mpg.de;
- HOGrasp dataset: https://hograspnet2024.github.io;
- SHOWMe dataset: https://europe.naverlabs.com/research/showme.

However, these are captured in controlled settings. The challenge is to train on these strong models, so that the models generalize out of distribution, i.e. to images taken in uncontrolled settings (such as internet images or images captured by autonomous robots).

Work Environment

The project is an internship at NAVER LABS Europe. The student will be mostly based in NAVER's campus in Grenoble, France. Regular weekly meetings will be scheduled for and discussing ideas and progress. Access to a GPU cluster and a work laptop will be provided.

You will be advised (in alphabetical order, QR-code links) by:
- Dr. Grégory Rogez (Naver);
- Dr. Dimitris Tzionas (UvA);
- Dr. Philippe Weinzaepfel (Naver).

All above people are international experts on the intersection of 3D Computer Vision, Graphics, and Machine Learning, with a strong publication record.

There is the possibility to refine the topic according to the candidate's prior experience, and the aligned interests of the advisors and the candidate (e.g. focus on video instead of single frames, focus on the body instead of hands, focus more on the object, etc). However, a connection to 3D humans and robotics is highly desired.

NAVER and the UvA are equal-opportunity institutes. For example, female students or students from any under-represented group are highly encouraged to apply.

Expectations

The project duration at NAVER is 6 months (a preparatory phase only at the UvA is possible), and requires a full-time engagement and hands-on coding skills. The internship will be financially supported; travel and accommodation are covered by NAVER.

As a minimum, at the end of this project the student should be able to: summarize the state of the art (SotA) w.r.t methods & datasets, explain how SotA methods work and formulate them mathematically, develop/reproduce working programs for these methods, analyze their performance and judge the strengths/weaknesses, suggest informed directions for future work, and write a thesis summarizing the above.

The goal is a publication at a top-tier conference, such as CVPR, ICCV, ECCV, SIGGRAPH, NeurIPS, etc. This naturally entails developing a novel method and evaluating it against SotA methods. In case the project is successful, there is the possibility to **continue as a fully-funded PhD** (in the same format as the internship, namely as part of NAVER with academic partnership with the UvA).

As the ideal candidate, you have the following profile:
- You have a background (e.g. by following related UvA courses) in at least one of the following: (3D) Computer Vision, (3D) Computer Graphics, and Machine/Deep Learning.
- You have strong programming skills, especially in Python and PyTorch (C/CUDA are welcome);
- You have strong communication, presentation and writing skills, and excellent command of English;
- You are self-motivated, value creative thinking, and are enthusiastic about working on the cutting edge;
- You work successfully both independently and in a team.

Optional: You have taken extra-curricular activities, such as participation in open-source projects, (inter-)national student competitions (e.g., robocup), etc. Experience on Natural Language Processing and Foundational Models are welcome – along with a strong desire to develop skills in these areas. You have a strong GitHub account.
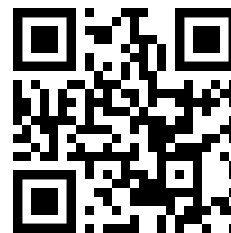

For **questions** & **expressing interest** (with your CV, transcript of records, and short motivation explanation) please contact:

Dimitris Tzionas
https://dtzionas.com
d.tzionas@uva.nl

Website   Email