CS 6384 Computer Vision Project Proposal Description

Professor Yu Xiang

February 6, 2023

1 Introduction

For the computer vision course project, students can choose a topic related to computer vision, and explore the topic in one of the three different ways:

- **Research-oriented.** In this direction, students are going to propose a new idea that has not been explored before in the literature, then implement the new idea and conduct experiments to verify it.

- **Application-oriented.** In this direction, students can apply an existing computer vision algorithm or method to a new problem or a new application. For example, if a method is proposed for domain A, the project can explore applying the method to domain B where different data are collected.

- **Implementation-oriented.** In this direction, students can select an existing computer vision algorithm or method, and then implement it and conduct experiments to verify the implementation. Since most computer vision methods are open-source these days, for implementation-oriented projects, students cannot just use an open-source code and run experiments with it.

For project evaluation, all three categories will be considered equally. A project will be evaluated according its quality in terms of implementation, experiments, presentation and writing, regardless of its category. However, students are encouraged to consider research-oriented projects and application-oriented projects. Even if the introduced novelty is incremental, it is still exploring new things researchers have not been tried before or applying an approach to new applications. Moreover, collecting real-world data for testing is highly encouraged.

2 Proposal Format

The project proposal should be prepared using the CVPR latex template. A useful online LaTex tool is Overleaf [https://www.overleaf.com/](https://www.overleaf.com/). We have the CVPR latex template accessible here via overleaf: [https://www.overleaf.com/read/gpjssbtrr5qm](https://www.overleaf.com/read/gpjssbtrr5qm). You can download a copy of the template or make a copy in overleaf for your own project, and then edit it.

The project proposal should be a 1-page PDF using the latex template with the following items:
• **Title.** Let’s give a name to your project.

• **Team Members.** List the names of the team members as the authors in the proposal. We expect you to work in groups of 2-4 students for the course projects.

• **Problem Statement.** Describe what is the problem you are trying to solve in this project.

• **Approach.** Describe what is your idea to solve the problem. It is fine if some details have not been figured out in the project proposal. But students should have rough ideas on how to proceed.

First, explicitly state that which category the project is in: research-oriented, application-oriented or implementation-oriented. Second, for research-oriented projects, describe the proposed idea and the novelty of the idea. For application-oriented projects, describe which approach is going to be used and how to apply this approach to a new application. For implementation-oriented projects, describe which approach is going to be implemented and the plan for the implementation.

• **Data.** Describe what dataset the project is going to use. Students can use existing datasets for experiments, or collect your own datasets, or even test the method with real-time data stream from a camera.

• **Evaluation.** Describe how to evaluate the success of the project. For example, what evaluation metrics will be used to evaluate the performance of the method?

• **References.** Cite related works in the proposal.

3 **Suggested Topics**

Based on the materials we cover in the lectures, we suggest the follow topics for the course project. However, the scope of the project is not limited to the mentioned topics below. Students can explore other topics in computer vision as well. Also, the references in the suggested topics are recent representative works. Students can explore methods beyond these references and propose new ideas for different topics.

• **Neural 3D Representations and Neural Rendering.** Neural networks can be utilized to learn 3D representations of objects or scenes and render these 3D representations into images \[ 26, 14, 38, 36, 54 \].

• **Feature Detection and Matching.** For this topic, students can explore methods for detecting keypoint features \[ 10 \], edges \[ 63, 39 \], lines \[ 65 \] or contours \[ 48 \] as well matching these features \[ 45, 53 \].

• **Stereo Depth Estimation.** This topic is about using stereo images for depth estimation \[ 29, 1, 28 \].

• **Structure from Motion (SfM) and SLAM.** SfM and SLAM are very actively research areas with large numbers of references \[ 46, 2, 3 \].
• **3D Object Reconstruction.** This topic is about reconstructing a 3D model of an object from a single image or multiple images [11, 15, 21, 30, 52].

• **Optical Flow.** This topic is about estimating the optical flow between two images [23, 40, 22].

• **Object Detection.** This topic is about detecting objects on images using bounding boxes [42, 41, 5].

• **Semantic Segmentation.** This topic is about labeling pixels in an image into semantic classes [32, 37, 43, 51].

• **Object Pose Estimation.** This topic is about estimating the pose of objects from images or videos [62, 56, 59, 8].

• **Object Tracking.** There are two types of object tracking problems: multi-object tracking [47, 35, 60] and tracking of a specific object in an input video [19, 27, 9].

• **Camera Pose Estimation.** This topic is about estimating the pose of cameras from images [34, 13, 18].

• **Human and Hand Pose Estimation.** This topic is about estimating the pose of humans or hands in either 2D or 3D from images or videos [49, 55, 6, 4, 17, 58, 64].

• **Human Activity Recognition.** The topic is about recognizing human activities from videos [24, 16, 12].

• **Images and Languages.** This topic is related to research on linking images and languages such as object grounding [7, 44, 25, 31] and visual query answering [61, 50, 57].

• **Visual Navigation.** This topic is about mobile navigation using vision [20, 66, 33].

4 **Deep Learning Resources**

Most recent vision methods leverage deep learning to train neural networks to tackle various problems in computer vision. If your project requires training of deep neural networks, you may need to have GPUs for training. **Google Colab** is a great free resource for small amounts of GPU resources: [https://colab.research.google.com/](https://colab.research.google.com/) Two widely-used deep learning frameworks:

- PyTorch [https://pytorch.org/](https://pytorch.org/)
- TensorFlow [https://www.tensorflow.org/](https://www.tensorflow.org/)

4 References


[7] Lei Chen, Mengyao Zhai, Jiawei He, and Greg Mori. Object grounding via iterative context reasoning. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, pages 0–0, 2019.


