CS5330 Pattern Recognition and Computer Vision
Final Project Proposal

Team: Shi Zhang, Zhizhou Gu

Project Expectation

Our team aims to build an application that utilizes an image database for either 3D reconstruction from 2D images or real-time image retrieval from the image database itself. The rough idea could be letting users take pictures via a camera or upload pictures from different perspectives of landmarks. Then the application will reconstruct the landmark as a 3d point cloud scene and display it for the user.

If this approach is proved to be too hard to achieve, we will switch to a continuation of the image retrieval assignment, and try to build an application that takes real-time input and returns the name of the landmark by extracting and matching features from both the input image and database. This could be a combination of both image retrieval and deep network assignments.

Inspirations and Resources

The inspirations came from several Kaggle Competitions:
Google Image Matching Challenge 2023:
https://www.kaggle.com/competitions/image-matching-challenge-2023

Google Landmark Retrieval Challenge 2021:
https://www.kaggle.com/competitions/landmark-retrieval-2021

Google Landmark Recognition Challenge 2021:
https://www.kaggle.com/competitions/landmark-recognition-2021

For the image database we are planning to use:
Google Landmarks Dataset V2: https://storage.googleapis.com/gld-v2/web/index.html
This database contains over 5 million landmark images, with a total of over 200,000 landmarks included.

Project Plan

- Nov 27th: Finalize the project's primary goal: 3D reconstruction or image retrieval.
- Dec 4th: Data Acquisition and Preparation, with a progress update to Professor.
- Dec 5th - 10th: Design and Development: Implement the core functionality.
- Dec 11th - 13th: Testing and Optimization.
- Dec 14th: Documentation and Reporting.