Introducing AI to Undergraduate Students via Computer Vision Projects

Kaiman Zeng, Yancheng Li, Yida Xu, Di Wu, Nansong Wu
Arkansas Tech University, Russellville, Arkansas 72801
{kzeng, yli, yxu1, dwu2, nwu}@atu.edu

Abstract

Computer vision, as a subfield in the general artificial intelligence (AI), is a technology can be visualized and easily found in a large number of state-of-art applications. In this project, undergraduate students performed research on a landmark recognition task using computer vision techniques. The project focused on analyzing, designing, configuring, and testing the two core components in landmark recognition: feature detection and description. The project modeled the landmark recognition system as a tour guide for visitors to the campus and evaluated the performance in the real world circumstances. By analyzing real-world data and solving problems, student's cognitive skills and critical thinking skills were sharpened. Their knowledge and understanding in mathematical modeling and data processing were also enhanced.

Introduction

It has been studied that through research experiences students become better learners, more engaged in their education programs, and more inclined to pursue graduate studies. It is also reported that students' problem-solving skills are improved, communication and interpersonal skills are improved, and self-confidence increased. (Eagan Jr et al. 2013) When considering research projects for undergraduate students, it is of great benefit to involving students in projects that are closely related to their life and even from ideas that are generated from themselves.

Computer vision, as a subfield in the general AI, has drawn great attention among students. They are curious about how Facebook automatically tags their friends in a photo. They are also fascinated by augmented reality, turned on by Pok'mon GO. This research is originated from freshman students' experience, the frustration to locate buildings on campus. At Arkansas Tech University, more than 50% students are reported as first-generation college students. Even a minor setback could be a major roadblock at the beginning of their college life. When the students know the techniques in computer vision can automatically recognize the buildings by image, they are excited and motivated by this landmark recognition project. In this paper, we will describe our experience in mentoring undergraduate students in a computer vision project and report our current progress of the project.

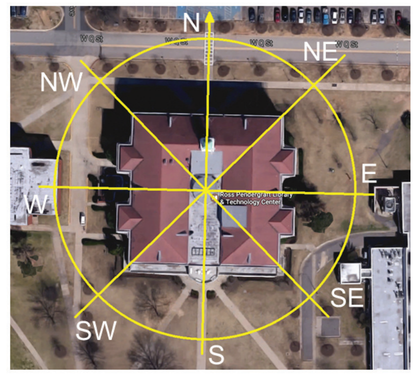

Figure 1: Eight Direction Imaging Positions

Problem Analysis and Data Collection

A goal of this project is to provide the students an experience of analyzing and solving the real-world problems. The landmark recognition system was modeled as a mobile tour guide for visitors to the campus, where the visitor can ask the mobile device for information about the building in his/her view. The students were guided to research the layout and history of Arkansas Tech University to obtain the basic information of landmarks on campus. They conducted a survey to gather the data of popular or hard-to-find locations among students and visitors. Then, an imaging plan was initialized for acquiring reference images of the landmarks. Each building was designed to be imaged from eight directions, as shown in Figure 1. If space allows, images were obtained from three positions in each direction. And the reference images were taken at different lighting conditions, i.e. in the morning and afternoon of a day. The longitude and latitude coordinates were also recorded with the images. Figure 2 shows the library on campus layered with the imaging positions. Currently, the campus landmarks dataset contains 3725 images for 23 buildings.

After collecting the reference images, the students analyzed the scene of the reference images, discussed the challenges of recognition task, and explored the modern algo-

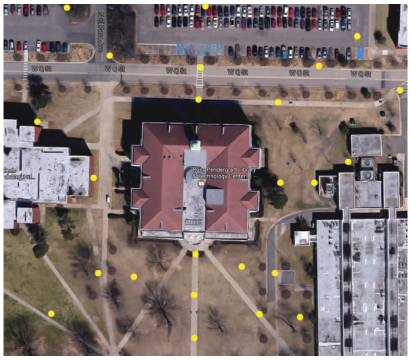

Figure 2: An Example of Imaging Positions

Figure 3: Examples of Matching SURF Features between Two Images

rithms through case study and literature review. Compared to conventional object recognition or location recognition, mobile landmark recognition differs from the following aspects: (1) the context information such as GPS and direction is available since the GPS sensor and gyro sensor are becoming a standard in current mobile devices; and (2) despite accuracy mobile users evaluate high on quick response. As a result, students focused on investigating the Speeded Up Robust Feature (SURF) (Bay et al. 2008) extraction and tackling the landmark recognition from the perspectives of features size and computation efficiency.

Design of the Landmark Recognition System

There have been interests in the problem of landmark recognition in the computer vision community for decades. The development of landmark recognition on the mobile platform has been relatively recent. The advent of commercial large-scale applications like Google's Goggles, and Microsoft's HoloLens have spurred further interest in this area.

Most current landmark recognition systems are based on client-server architecture and heavily rely on server-side support and processing. (Chen et al. 2009) The reference images of locations (typically at city scale) at the server, and landmark recognition are accomplished via sending the query image to the server, extracting and matching the visual features at the server side.

A key shortcoming of this type of system is the users' concern about the privacy. They disclose their locations, route of travel, and access to the image album to the remote server. Besides, the response time highly depends on the data speed of the wireless network. As a result, the on-device landmark recognition has drawn increasing research in recent years. For example, (Guan et al. 2015) can compress each image descriptor into several bytes; however, the user experience is less than satisfactory because recognition accuracy is reduced due to the significant compression of image descriptors.

In this project, the SURF feature was applied to detect and describe the local salience. The experiments of image matching were performed between the query image and relevant images, as well as irrelevant images. Figure 3 shows examples of matching SURF features between two images. It demonstrates the effectiveness of the proposed method in distinguishing relevant from irrelevant buildings.

Conclusion and Future Work

Through this project, students not only acquire intellectual technique skills, but also grow in personal and professional aspects. Students reflected that the feature extraction was too abstract to understand in the first place, but a better understanding was formed through image processing and programming. It is an exciting time that AI becomes real and empowering the students with AI technology becomes essential to the current generation of engineering students. In the future, we will develop more computer vision related projects such as self-driving cars to foster students' interests in AI, and promote AI in undergraduate education.

Acknowledgements

This project is funded by Arkansas Tech University Undergraduate Research Grant. We thank Dr. Mostafa Hemmati and undergraduate research committee for invaluable comments and suggestions.

References

Bay, H.; Ess, A.; Tuytelaars, T.; and Van Gool, L. 2008. Speeded-up robust features (surf). *Computer vision and image understanding* 110(3):346–359.

Chen, T.; Wu, K.; Yap, K.-H.; Li, Z.; and Tsai, F. S. 2009. A survey on mobile landmark recognition for information retrieval. In *Mobile Data Management: Systems, Services and Middleware, 2009. Tenth International Conference on*, 625–630. IEEE.

Eagan Jr, M. K.; Hurtado, S.; Chang, M. J.; Garcia, G. A.; Herrera, F. A.; and Garibay, J. C. 2013. Making a difference in science education: the impact of undergraduate research programs. *American educational research journal* 50(4):683–713.

Guan, T.; Wang, Y.; Duan, L.; and Ji, R. 2015. On-device mobile landmark recognition using binarized descriptor with multifeature fusion. *ACM Transactions on Intelligent Systems and Technology* 7(1):12.