Penn Campus Building Multi-Platform Image Recognition

Mark Knichel | Matt Evans
Faculty Advisor: Jianbo Shi

Abstract:
- Detect with accuracy the campus building which is pictured in a submitted image

Background:
- **Image-driven Search** remains to be conquered by the search giants of the web. Web search engines can crawl the text-heavy web, but our world is a visual place with definite visual features. Very little has been done with genuine computer vision-driven image search.
- **Visual Navigation** most clearly follows the way humans navigate. We look for unique landmarks, recognizable patterns, or distinct colors in our attempts to navigate unfamiliar terrain. GPS technology provides macro-navigation, but we rely on visual cues for low-level identification.
- **Computer Vision** in current research focuses on face detection, single feature recognition, or motion techniques. Buildings are a difficult task, loaded with unique identifying features.

Approach:
- A persistent web service, with accurate computer vision algorithms, can offer an effective visual navigation tool

Challenge:
- How can we uniquely specify the visual fingerprint of a building? What image features are most accurate for comparisons?

**A MARRIAGE OF TECHNOLOGIES...**

**COMPUTER VISION TECHNIQUES...**

- **Edge detector**
  - strong, clear lines in an image

- **Corner detector**
  - sharp gradients across nearby pixels

- **Color Histogram**
  - pixel values into color bins

Results:

- Service database now contains over 4000 images of more than 50 campus buildings
- Database image features calculated offline
- Buildings tend to have strong edges and distinct corners
- Building color is an obvious feature of a photo
- These features are easy to compute and there has been much research on them
- Matches are calculated with a weighted rank of multiple matching algorithms
  - Uses combination SVM / brute force method to compare features
- The correct campus building is detected with 35% accuracy, and the system’s highest-scoring image is returned

**BUILDING: IRVINE AUDITORIUM**

**THE SYSTEM IN ACTION...**

- Mail server handles MMS flow
- OpenCV runs on query image
- MySQL DB of images polled for matches
- Best match selected, tag returned

**Conclusions:**
- Accuracy can be improved by taking more pictures, eroding and modifying images to replicate the angles and views humans observe
- Scale and angle invariance is difficult
- Building recognition is a difficult task with limited previous research
- Traditional image features do not work well with general building recognition

**SPECIAL THANKS TO PROF. KOSTAS DANIILIDIS and THE CIS DEPARTMENT**

Try it now!

Send a picture message to pisrmail@gmail.com