COMP 9517 Computer Vision

S2, 2011

Group Project - Draft

Preliminaries

This is two-person project. Pick a project team member, and register your group ONLINE by Thurs week 6.

Project Synopsis

The project consists of two parts, with checkpoints built in to enable you to manage the project and the deliverables. Read the project specification through carefully, and discuss with potential team member before registering your group!

Part 1

This part is common to all groups.

Feature detection and matching form the foundation of many successful applications of computer vision. The objective of this task is for you to become familiar with these techniques and their implementation in OpenCV.

Given a collection of photographs automatically arrange them in a panograph (Szeliski, 6.12). See the flickr link for some artistic examples.

http://www.flickr.com/photos/tags/panograph/

In addition to the final panograph you should be able to display any pair of images side by side with lines showing potential matches. You can base your solution on the following OpenCV sample programs:

```
opencv\samples\c\find_obj.cpp
opencv\samples\cpp\video_homography.cpp
opencv\samples\cpp\generic_descriptor_match.cpp
```

There will be two bonus marks for implementing additional features such as automatically detecting images are to be included in the collage, exposure correction or alpha blending.

Testing of Part 1

Collect several sets of images to verify the performance your algorithm e.g. photos of the UNSW campus. A collection of eight test images of an outdoor scene will be released on the day of the demo. You should also provide a two page report giving a high level overview of your approach.

Part 2

In the second part, your team should select ONE direction of further work, do the necessary literature survey on techniques and algorithms, implement them and write a detailed report. You should also find suitable datasets to work with. The week 9 presentation (see checkpoints below) will provide an opportunity to get feedback from lecturers on your chosen direction.

Here are some possible lines of further work. Your team is encouraged to do its own research and decide on a line of further development. The direction you choose may also be different from the suggestions made in this section- use the week 9 presentation or ask us prior to get feedback.

Possible directions

Datasets:

"Computer vision papers online" has a comprehensive list of public datasets used in recent computer vision research. You can use these in your projects.

http://www.cvpapers.com/datasets.html

Easier projects:

ICMLA 2011 Street View challenge

> The objective of this competition is to recognise business signage and certain types of objects in images collected for Google StreetView. It is not required that you actually enter a submission into the competition – only class presention and your report will be assessed. You will need to contact the organisers to obtain the data.
>
> http://www.icmla-conference.org/icmla11/challenge.htm

Efficient video search

> Build a system that efficiently finds all (or at least some) occurrences of a particular backdrop, prop, or another object in a video. Only retrieval needs to be fast, you can do any amount of preprocessing beforehand.
>
> J. Sivic, A. Zisserman Video Google: A Text Retrieval Approach to Object Matching in Videos
> http://www.robots.ox.ac.uk/~vgg/publications/papers/sivic03.pdf

High Quality Panaromas

Develop an automatic panorama generation system allowing for a richer class of transformations than a simple panograph. There is a discussion in chapter 9 of Szeliski.

Harder projects:

3-D Reconstruction from video / augmented reality

Reconstruct 3D geometry (e.g. a point cloud) from a video captured by a handheld camera. Optionally, can you use this information to insert a rendered object into the video?

Iryna Gordon and David G. Lowe, "Scene modelling, recognition and tracking with invariant image features"
http://www.cs.ubc.ca/~lowe/papers/gordon04.pdf

Stephen Se, David G. Lowe and Jim Little, "Global localization using distinctive visual features
http://www.cs.ubc.ca/~lowe/papers/iros02b.pdf

Photo Tourism

Construct a 3D model of an architectural landmark from a collection of photographs collected from the internet. You can use the following paper as the starting point and search for others that cite it:

N Snavely, SM Seitz, R Szeliski, Photo Tourism: Exploring photo collections in 3D, 2006
http://research.microsoft.com/apps/pubs/default.aspx?id=75618

Checkpoints

Week 8 Demo on Part 1 (10 marks)

Week 9 Presentation + Report (10 marks)

Week 13 Final Demo + Report (20 + 20 marks)

Total marks 60

Reports

Both preliminary and final reports for the second part of the project should be in two column IEEE format (http://www.ieee.org/web/publications/authors/transjnl/index.html), maximum of 4 and 8 pages respectively. Marks will be deducted for poor formatting.

Code Submission

Even though it is not assessed directly, you will need to submit the code for the project by the time of the final demo.