Song Cao

Department of Computer Science
Cornell University
345 Gates Hall
Ithaca, NY 14850

Phone: (607) 220-6745
Email: *caosong@cs.cornell.edu*
Homepage: *http://www.cs.cornell.edu/~caosong/*

Current Position

5th Year PhD Candidate (Research Assistant), Department of Computer Science, Cornell University

GPA: 3.73 Expected graduation date: December 2014

Previous Education

B.E. Computer Software, Tsinghua University, Beijing, China, 2008.

GPA: 3.7 (93/100)

Undergraduate Thesis: Multiple Motion Trajectory-Based Video Retrieval

Employment

Research Scientist Intern, Amazon.com LLC, May 2014 - August 2014

Research Interests

Computer Vision and Applied Machine Learning (Advisor: Noah Snavely)

Publications

Song Cao, Noah Snavely: Minimal Scene Descriptions from Structure from Motion Models. CVPR 2014.

(Also presented at Scene Understanding Workshop - CVPR 2014, and won Best Poster Award.)

Song Cao, Noah Snavely: Graph-Based Discriminative Learning for Location Recognition. CVPR 2013. IJCV 2014.

Song Cao, Noah Snavely: Learning to Match Images in Large-Scale Collections. Workshop on Web-scale Vision and Social Media, ECCV 2012.

Yijun Yang, Song Cao, Junhai Yong, Hui Zhang, Jean-Claude Paul, Jiaguang Sun: Approximate computation of curves on B-spline surfaces. Computer-Aided Design, Volume 40, Issue 2, Pages 223-234, February 2008.

Selected Projects

Minimal Scene Descriptions from Structure from Motion Models

January 2013 - November 2013

How much data do we need to describe a location? We explore this question in the context of 3D scene reconstructions created from running structure from motion on large Internet photo collections, where reconstructions can contain many millions of 3D points. We consider several methods for computing much more compact representations of such reconstructions for the task of location recognition, with the goal of maintaining good performance with very small models. In

particular, we introduce a new method for computing compact models that takes into account both image-point relationships, as well as feature distinctiveness, and show that this method produces small models that yield better recognition performance than previous model reduction techniques.

Graph-Based Discriminative Learning for Location Recognition

January 2012 - November 2012

Recognizing the location of a query image by matching it to a database is an important problem in computer vision, and one for which the *representation* of the database is a key issue. We explore new ways for exploiting the structure of a database by representing it as a graph, and show how the rich information embedded in a graph can improve a bag-of-words-based location recognition method. In particular, starting from a graph on a set of images based on visual connectivity, we propose a method for selecting a set of subgraphs and learning a local distance function for each using discriminative techniques. For a query image, each database image is ranked according to these local distance functions in order to place the image in the right part of the graph. In addition, we propose a probabilistic method for increasing the diversity of these ranked database images, again based on the structure of the image graph. We demonstrate that our methods improve performance over standard bag-of-words methods on several existing location recognition datasets.

Learning to Match Images in Large-Scale Collections

February 2011 - November 2011

Many computer vision applications require computing structure and feature correspondence across a large, unorganized collection of images. This is generally a difficult, computationally expensive process, because the set of matching image pairs is unknown in advance, and so good methods for quickly predicting which images match are critical. Image comparison method such as bag-of-words models or global features, are often used to predict similar pairs, but can be very noisy. In this paper, we propose a new image matching approach that uses discriminative learning techniques—applied to training data gathered automatically during the image matching process—to gradually compute a better similarity measure for predicting whether two images overlap. By using such a learned similarity measure, our algorithm can select image pairs that are more likely to match for performing further matching and geometric consistency checks, improving overall efficiency in the matching process. Our approach processes a set of images in an iterative manner, alternately performing pairwise feature matching and learning an improved similarity measure. Our experiments show that our learned measures can significantly improve match prediction over both the standard *tf-idf* weighted similarity measure and more recent unsupervised techniques even with small amounts of training data, and can improve the overall speed of the image matching process by more than a factor of two.

Teaching

CS 7670: Special Topics in Computer Vision, Co-instructor, 2014 Fall, Cornell University.

CS 6670: Computer Vision (PhD-level course), Teaching Assistant, 2011 Spring, Cornell University.

Skills

Python (Proficient), C/C++ (Proficient), Java (Proficient), Unix Tools (Familiar), C#(Familiar), MATLAB (Familiar) and Mathematica (Familiar)

Familiar with image retrieval & recognition, discriminative learning and graphical models

Honors & Awards

Best Poster, Scene Understanding Workshop, CVPR 2014

Excellent Graduate, Tsinghua University, China, 2008

Tsinghua Academic Excellence Scholarship, 1st Prize(Mitsubishi UFJ Scholarship), 2007

Tsinghua Comprehensive Excellence Scholarship, 1st Prize (Toyota Scholarship), 2006

1st Prize of National College Physics Competition, 2005