

**RESEARCH
INTERESTS**

- Computer vision, machine learning
- Multimodal perception for autonomous systems
- Adaptive intelligent human-computer interfaces

EDUCATION**Massachusetts Institute of Technology (MIT)****Cambridge, MA****Ph.D.** in Electrical Engineering and Computer Science

2009

Advisor: Prof. Trevor Darrell

Title: *"Image Sense Disambiguation: A Multimodal Approach"* (abstract attached).**M.Sc.** in Electrical Engineering and Computer Science

2004

Advisors: Prof. Jim Glass and Prof. Trevor Darrell

The University of British Columbia (UBC)**Vancouver, BC****B.Sc.** in Computer Science. **GPA 4.0**

2000

GRANTS**GRANTS**

- NSF: RI: Large: Collaborative Research: *Reconstructive Recognition: Advancing computer vision by uniting statistical scene understanding and physics-based visual reasoning*, PI. \$80,000/yr. 2012-1
- DARPA: Mind's Eye: *Recognizing Activities with Probabilistically Grounded Visual Intelligence Models*, PI. \$200,000/yr. 2010-1
- Toyota: *Visual Object Recognition for Practical Assistance Robots*, Co-PI. \$250,000/yr. 2010-1

**PROFESSIONAL
EXPERIENCE****University of Massachusetts, Lowell****Lowell, MA****Assistant Professor (Sept 2012 – present)**

- Tenure-track assistant professor in the Computer Science department

UC Berkeley CS Division**Berkeley, CA****Postdoctoral Researcher (Oct 2009 – Aug 2012)**

- Worked with Prof. Trevor Darrell on interactive object recognition in robots, transferring knowledge about objects across related domains; also working on action recognition in videos. Jointly appointed at the International Computer Science Institute (ICSI), where I served as PI on ICSI's DARPA Mind's Eye program subcontract and co-PI on Toyota's object recognition grant.

Harvard School of Engineering and Applied Sciences	Cambridge, MA
---	----------------------

Research Associate (Nov 2011 – present); Visiting Researcher (2010-11)

- Collaboration with Prof. Todd Zickler on modeling the color processing in consumer cameras with applications in computer vision.

MIT Computer Science and Artificial Intelligence Lab	Cambridge, MA
---	----------------------

Graduate Research Assistant (Sep 2002 – Aug 2009)

- Conducted research on modeling of audio-visual speech, image processing, object detection and recognition in images, and natural language processing.

Mitsubishi Electric Research Laboratories (MERL)	Cambridge, MA
---	----------------------

Research Intern (May 2005 – Aug 2005)

- Worked with Prof. Candy Sidner investigating human-robot engagement for hosting activities. Implemented algorithms for face detection and tracking on a mobile robot.

Advanced Telecommunications Research Intl.	Kyoto, Japan
---	---------------------

Student Researcher (Jun 1998 – Dec 1998)

- Conducted research in the area of speech synthesis. Improved speech databases for a concatenative speech synthesis system (CHATR).

UBC Computer Science Department	Vancouver, BC
--	----------------------

Undergraduate Research Assistant (Jan 1998 – May 1998)

- Worked on a research project in the field of performance prediction for object-oriented frameworks and libraries. Researched scientific literature and implemented algorithms in Java. Supervised by Prof. Gail Murphy.

**TEACHING
EXPERIENCE**

University of Massachusetts, Lowell	Lowell, MA
--	-------------------

Instructor

- Developed and taught a new graduate course, 91.550 Computer Vision, Fall 2012. The course covers introductory topics in computer vision and image processing through lectures, programming assignments and a final project.
- Developed and taught a new undergraduate and graduate course, 91.454 Machine Learning, Spring 2013. The course offers an introduction to machine learning through lectures, programming assignments, in-class quizzes, exams and a final project.

Science Club for Girls	Cambridge, MA
-------------------------------	----------------------

Volunteer Mentor Scientist (2010)

- Led a once-per-week after-school club for 5th grade girls in Cambridge public school, along with two other mentor scientists. Planned and conducted hand-on rocket science experiments to encourage the girls' interest in science careers.

MIT EECS Department	Cambridge, MA
----------------------------	----------------------

Graduate Teaching Assistant (Jan 2007 – May 2007)

- Helped develop a new multi-disciplinary graduate course on Intelligent

Multimodal Interfaces. Created and graded assignments, prepared and gave guest lectures.

UBC Mathematics Department

Vancouver, BC

Undergraduate Teaching Assistant (1999, 1997)

- Supervised and graded exams and homework assignments for first-year calculus courses. Provided one-on-one tutoring in calculus to the students.

**INDUSTRY
EXPERIENCE**

Redback Networks, Inc.

San Jose, CA

Software Developer, SmartEdge 800 (Jul 2000 – Aug 2002)

- Designed and implemented network protocols, system infrastructure, and software redundancy features in an embedded programming environment using C for NetBSD and VxWorks.

Motorola Inc.

Richmond, BC

Software Developer (May 1999 – Aug 1999)

- Developed test tools and designed test cases for GPRS wireless network protocol software in a team environment.

Newbridge Networks Corporation

Burnaby, BC

Software Engineer (Jan 1997 – Aug 1997)

- Designed, wrote and packaged software using Object Oriented design techniques in C++. Researched the latest in network security technology. Produced a security model for a distributed object system based on CORBA.

AWARDS

SELECTED AWARDS

- **Best Paper Prize** at the Domain Adaptation Workshop: Theory and Application, Sierra Nevada, Spain 2011
- Science Club for Girls **Mentor Award** 2010
- **Best Student Paper Award** at the International Conference on Multimodal Interfaces 2006
- Canadian Natural Sciences and Engineering Research Council (NSERC) **Postgraduate Scholarship** 2000
- **Canadian Premier's Award** for Young Women in Science 1999
- Honorable Mention in the CRA's **Outstanding Undergraduate Award** Competition 1998

PUBLICATIONS

Publications

See my Google Scholar profile for full list of publications
<http://scholar.google.com/citations?user=9xDADY4AAAAJ>

Selected Publications

[Efficient Learning of Domain-invariant Image Representations](#)

J Hoffman, E Rodner, J Donahue, K Saenko, T Darrell
arXiv preprint arXiv:1301.3224 2013

[Mid-level Features Improve Recognition of Interactive Activities](#)

K Saenko, B Packer, CY Chen, S Bandla, Y Lee, Y Jia, JC Niebles, D Koller, L ...
Technical report, 2012. <http://www.eecs.berkeley.edu/Pubs/TechRpts/>. 7, 8 2012

[A combined pose, object, and feature model for action understanding](#)

B Packer, K Saenko, D Koller
Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on ... 2012

[From pixels to physics: Probabilistic color de-rendering](#)

Y Xiong, K Saenko, T Darrell, T Zickler **Oral Presentation: 3.5% Acceptance Rate**
Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 358-365 2012

[Modeling the Uncertainty in Inverse Radiometric Calibration](#)

Y Xiong, K Saenko, T Zickler, T Darrell 2011

[Practical 3-D object detection using category and instance-level appearance models](#)

K Saenko, S Karayev, Y Jia, A Shyr, A Janoch, J Long, M Fritz, T Darrell
Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International ... 2011

[Learning object color models from multi-view constraints](#)

T Owens, K Saenko, A Chakrabarti, Y Xiong, T Zickler, T Darrell
Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, 169-176 2011

[What you saw is not what you get: Domain adaptation using asymmetric kernel transforms](#)

B Kulis, K Saenko, T Darrell **Oral Presentation: 3.5% Acceptance Rate**
Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on ... 2011

[The Ratio Method for Multi-view Color Constancy](#)

T Owens, K Saenko, A Chakrabarti, Y Xiong, T Zickler, T Darrell
UC Berkeley Tech. Report 2011

[Domain adaptation with multiple latent domains](#)

J Hoffman, K Saenko, B Kulis, T Darrell
NIPS Domain Adaptation Workshop 2011

[The NBNN kernel](#)

T Tuytelaars, M Fritz, K Saenko, T Darrell, ICCV 2011

PROFESSIONAL SERVICE

- Co-organizer of the [tutorial at CVPR 2012](#) on Domain Transfer Learning for Vision Applications
- Reviewer for CVPR 2013
- Co-organizer of the [Workshop on Integrating Language and Vision](#), December 2011, Granada, Spain.
- Co-organizer of the 6th [Workshop for Women in Machine Learning](#), December 2010, Vancouver, Canada.
- Volunteer mentor for [Science Club for Girls](#) in Cambridge, MA.
- Reviewer for *Transactions on Pattern Analysis and Machine Intelligence*, *IEEE International Conference on Computer Vision*, *IEEE Transactions on Multimedia*.
- Session chair for the [3rd CSAIL Student Workshop, 2007](#), Gloucester, MA.
- Co-organizer of the Johns Hopkins 2006 Summer Workshop on articulatory feature-based speech recognition ([final report](#)).

OTHER

- Fluent in Russian, working knowledge of German and Japanese.
- Canadian citizen, permanent resident of the U.S.

Ph. D. Thesis Abstract**Image Sense Disambiguation: A Multimodal Approach**

If a picture is worth a thousand words, can a thousand words be worth a training image? Most successful object recognition algorithms require manually annotated images of objects to be collected for training. The amount of human effort required to collect training data has limited most approaches to the several hundred object categories available in the labeled datasets. While human-annotated image data is scarce, additional sources of information can be used as weak labels, reducing the need for human supervision. In this thesis, we use three types of information to learn models of object categories: speech, text and dictionaries. We demonstrate that our use of non-traditional information sources facilitates automatic acquisition of visual object models for arbitrary words without requiring any labeled image examples. Spoken object references occur in many scenarios: interaction with an assistant robot, voice-tagging of photos, etc. Existing reference resolution methods are uni-modal, relying either only on image features, or only on speech recognition. We propose a method that uses both the image of the object and the speech segment referring to it to disambiguate the underlying object label. We show that even noisy speech input helps visual recognition, and vice versa. We also explore two sources of linguistic sense information: the words surrounding images on web pages, and dictionary entries for nouns that refer to objects. Keywords that index images on the web have been used as weak object labels, but these tend to produce noisy datasets with many unrelated images. We use unlabeled text, dictionary definitions, and semantic relations between concepts to learn a refined model of image sense. Our model can work with as little supervision as a single English word. We apply this model to a dataset of web images indexed by polysemous keywords, and show that it improves both retrieval of specific senses, and the resulting object classifiers.