

# Tom Yeh

430 UCB  
University of Colorado  
Boulder, CO, 80309

tom.yeh@colorado.edu  
<http://tomyeh.info>  
(303) 492-1762

## Education

- Massachusetts Institute of Technology**, Cambridge, MA June 2009  
Ph.D. in Computer Science  
Thesis: *Interacting with computers using images for search and automation*  
Advisor: Trevor Darrell
- Massachusetts Institute of Technology**, Cambridge, MA June 2004  
S.M. in Computer Science  
Thesis: *IDeixis : image-based deixis for recognizing locations*  
Advisor: Trevor Darrell
- Simon Fraser University**, Burnaby, BC, Canada December 2001  
B.Sc. in Computer Science, Honor

## Research Interests

Human-Computer Interaction, Computer Vision, Software Engineering, Information Retrieval, Mobile Computing

## Employment

- University of Colorado Boulder**, Boulder, CO 2011 - 2012  
*Assistant Professor*
- University of Maryland Institute for Advanced Computer Studies**, College Park, MD 2011 - 2012  
*Assistant Research Scientist*
- University of Maryland Institute for Advanced Computer Studies**, College Park, MD 2009 - 2011  
*Postdoctoral Research Associate*
- MIT Computer Science and Artificial Intelligence Laboratory**, Cambridge, MA 2009  
*Postdoctoral Research Associate*
- MIT Computer Science and Artificial Intelligence Laboratory**, Cambridge, MA 2003-2009  
*Research Assistant*

## Research Grants

- NSF: VOSS**: Crowdsourcing Interaction Design for Citizen Space Virtual Organizations 2012-2014
- DARPA: Active Authentication**: Screen Fingerprint as a New Modality for Active Authentication 2012-2013

## Awards

- Best Paper**, ACM Symposium on User Interface Software and Technology (UIST) 2010
- Best Student Paper**, ACM Symposium on User Interface Software and Technology (UIST) 2009

## Publications

### Conference papers

1. J. S. Pedro, **T. Yeh**, and N. Oliver. 2 Leveraging user comments for aesthetic aware image search reranking. In *Proceedings of the 21st international conference on World Wide Web (WWW '12)*. ACM, New York, NY, USA, 439-448.
2. B. Xie, **T. Yeh**, G. Walsh, I. Watkins, and M. Huang. Co-designing an e-health tutorial for older adults. In *Proceedings of the 2012 iConference (iConference '12)*. ACM, New York, NY, USA, 240-247.
3. **T. Yeh**, T. Chang, B. Xie, G. Walsh, K. Wongsuphasawat, I. Watkins, M. Huang, L. Davis, and B. Bederson. Creating Contextual Help for GUIs Using Screenshots. In *Proceedings of the 24th ACM Symposium on User Interface Software and Technology (UIST)*, Santa Barbara, CA, October 2011.
4. T. Chang, **T. Yeh** and R. Miller. Correlating the Visual Representation of User Interfaces with their Internal Structures and Metadata. In *Proceedings of the 24th ACM Symposium on User Interface Software and Technology (UIST)*, Santa Barbara, CA, October 2011.
5. B. Xie, **T. Yeh**, G. Walsh, I. Watkins, and M. Huang. Co-designing Contextual Tutorials for Older Adults on Searching Health Information on the Internet. In *Proceedings of the 74th Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, New Orleans, LA, October, 2011.
6. **T. Yeh**, B. White, J. San Pedro, B. Katz, and L. Davis. A Case for Query by Image and Text Content: Searching Computer Help Using Screenshots and Keywords. In *Proceedings of the 20th International Conference on World Wide Web (WWW)*, Hyderabad, India, March 2011.
7. **Best Paper Award** J. Biggam, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and **T. Yeh**. VizWiz: Nearly Real-time Answers to Visual Questions. In *Proceedings of the 23rd ACM Symposium on User Interface Software and Technology (UIST)*, New York, NY, USA, October 2010.
8. A. Kembhavi, **T. Yeh**, and L. Davis. Why Did the Person Cross the Road (There)? Scene Understanding using Probabilistic Logic Models and Common Sense Reasoning. In *Proceedings of the 11th European Conference on Computer Vision (ECCV)*, Crete, Greece, September 2010.
9. T. Chang, **T. Yeh**, and R. Miller. GUI Testing Using Computer Vision. In *Proceedings of the Conference on Human Factors in Computing System (CHI)*, Atlanta, GA, April 2010.
10. **Best Student Paper Award** **T. Yeh**, T. Chang, R. Miller. Sikuli: Using GUI Screenshots for Search and Automation. In *Proceedings of the 22nd ACM Symposium on User Interface Software and Technology (UIST)*, Victoria, BC, Canada, October 2009.
11. **T. Yeh**, J. Lee, T. and Darrell. Fast Concurrent Object Localization and Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, June 2009.
12. **T. Yeh**, J. Lee, T. and Darrell. Photo-based Question Answering. In *Proceedings of the ACM International Conference on Multimedia (MM)*, Vancouver, BC, October 2008.
13. **T. Yeh** and T. Darrell. Dynamic Visual Category Learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, AK, June 2008.
14. **T. Yeh** and T. Darrell. Multimodal Question Answering for Mobile Devices. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI)*, Canary Island, Spain, January 2008.

15. **T. Yeh**, J. Lee and T. Darrell. Adaptive Vocabulary Forests for Dynamic Indexing and Category Learning. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brazil, October 2007.
16. **T. Yeh**, K. Grauman, K. Tollmar, and T. Darrell. A Picture is Worth a Thousand Keywords: Image-based Object Search on a Mobile Platform. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, Portland, OR, April 2005.
17. **T. Yeh** and T. Darrell. DoubleShot: An Interactive User-Aided Segmentation Tool. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI)*, San Diego, CA, January 2005.
18. K. Tollmar, **T. Yeh**, and T. Darrell. IDEixis: Image-based Deixis for Finding Location-based Information. In *Proceedings of the 6th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI)*, Glasgow, Scotland, September 2004.
19. **T. Yeh** and T. Darrell. Searching the Web with Mobile Images for Location Recognition, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, June 2004.
20. K. Voll, **T. Yeh** and V. Dahl. An Assumptive Logic Programming Methodology for Parsing, In *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, Vancouver, November 2000.

#### **Workshop papers**

21. D. Chen, M. Bilgic, L. Getoor, D. Jacobs, L. Mihalkova, and **T. Yeh**. Active Inference for Retrieval in Camera Networks. In *Proceedings of the IEEE Workshop on Person Oriented Vision*, January 2011.
22. B. White, **T. Yeh**, J. Lin, and L. Davis. Web-Scale Computer Vision using MapReduce for Multimedia Data Mining. In *Proceedings of the KDD Multimedia Data Mining Workshop*, Washington, DC, July 2010.
23. **T. Yeh**, J. Lee and T. Darrell. Scalable Classifiers for Internet Vision Tasks. In *Proceedings of the First IEEE Workshop on Internet Vision (IV)*, Anchorage, AK, June 2008.
24. J. M. Paluska, D. Saff, **T. Yeh** and K. Chen, Footloose: A Case for Physical Eventual Consistency and Selective Conflict Resolution. In *Proceedings of the 5th IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)*, Monterey, CA, October 2003.

#### **Posters and Demonstrations**

25. B. Xie, **T. Yeh**, G. Walsh, I. Watkins and M. Huang. Co-designing an integrated e-tutorial tool to improve older adults' e-health literacy. Presented at the *64th Annual Scientific Meeting of the Gerontological Society of America (GSA)*, Boston, MA, November 2011.
26. **T. Yeh** and B. Katz. Indexing and Searching Software Documentations using Text, OCR, and Image. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Poster, Boston, MA, July 2009.
27. **T. Yeh**, T. Chang and R. Miller. Obtaining Help on GUI Elements using Screenshots. In *Proceedings of the 21st ACM Symposium on User Interface Software and Technology (UIST)*, Poster, Monterey, CA, October 2008.
28. **T. Yeh**, K. Grauman and T. Darrell. Image-based Web Search on a Mobile Platform. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Demo, San Diego, CA, June 2005.
29. **T. Yeh**, K. Tollmar and T. Darrell. IDEixis - Image-based Deixis for Finding Location-Based Information, In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA)*, Demo, Vienna, Austria, April 2004.

#### **Journal Article**

30. K. Voll, **T. Yeh** and V. Dahl. An Assumptive Logic Programming Methodology for Parsing. *International Journal on Artificial Intelligence Tools*, Vol. 10 (4), pp. 573-588, December 2001.

## Software

**Sikuli Script/IDE** ([www.sikuli.org](http://www.sikuli.org))

200,000 downloads since it was open-sourced in January 2010.

Used by many companies and organizations such as Google, Intel, Symantec, Spotify, and Office of Budget and Management

## Patent

**Photo-based Mobile Deixis System and Related Techniques** – US 7,872,669

## Service

<i>Chair</i>	MIT CSAIL Student Workshop	2007
<i>Program</i>	ACM Symposium on User Interface Software and Technology (UIST)	2012
<i>Committee</i>	IEEE Workshop on Applications of Computer Vision (WACV)	2011
<i>Conference</i>	Conference on Human Factors in Computing System (CHI)	
<i>Reviewer</i>	International Conference on Intelligent User Interface (IUI)	
	ACM Symposium on User Interface Software and Technology (UIST)	
	IEEE Conference on Computer Vision and Pattern Recognition (CVPR)	
	IEEE Conference on Computer Vision and Pattern Recognition (ICCV)	
	ACM International Conference on Multimodal Interaction (ICMI)	
<i>Journal</i>	IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)	
<i>Reviewer</i>	Machine Vision and Application (MVA)	

## Talks

- 2011 Sikuli: GUI Testing using Computer Vision  
*Office of Management and Budget, Washington, DC*
- Sikuli: GUI Testing using Computer Vision  
*Northern Virginia Test Automation Interest Group, Herndon, Virginia*
- 2010 Online GUI Help using Computer Vision  
*HCIL, University of Maryland, College Park, MD*
- 2009 Interacting with Computer using Images  
*University of British Columbia, Vancouver, BC*
- Interacting with Computer using Images  
*Harvard University, Cambridge, MA*
- Interacting with Computer using Images  
*Columbia University, New York, NY*
- Interacting with Computer using Images  
*University of Maryland, College Park, MD*
- Interacting with Computer using Images  
*Media Lab, MIT, Cambridge, MA*

## Press

- 2011 The UID group: Sikuli, picture-driven computing  
*MIT EECS News Letter*

- 2010 Screen-Capture Programming: What You See is What You Script  
*IEEE Computing Now*
- Picture-driven Computing  
*MIT Press*
- Sikuli Rethinks Programming  
*MIT CSAIL Spotlights*
- Programming Visually with Sikuli  
*Computing Community Consortium (CCC)*
- MIT Offers Picture-Centric Programming To the Masses With Sikuli  
*Slashdot*
- Sikuli Automates Almost Anything with Screenshot Ease  
*LifeHacker*
- Sikuli: Scripting with Screenshots  
*PC Pro*
- MIT Creates Picture-Driven Programming for the Masses  
*CIO*
- Sikuli Can Automate Any GUI by Taking Screenshots  
*TUAW: The Unofficial Apple Weblog*
- MIT Project Lets You Author Code with Screenshots, Pictures  
*DigitalBeat*
- Sikuli: Create “Smart Macros” based on Screenshots  
*dotTech*
- Sikuli: the Coolest Python Project I Have yet Seen  
*Python411 Podcast*
- 2009 Goggle Image Search Coming to Android  
*LinuxDevices*
- 2004 When Databases Think  
*Smart Device Central*
- Picture This  
*MIT Technology Review*

## Teaching

<i>Lecturer</i>	<b>Introduction to Human Computer Interaction (CMSC 434)</b> , UMD Taught a class of 45 students and supervised 10 team projects.	2010
<i>Guest Lecturer</i>	<b>Developing Interfaces for Rehabilitation (IS 698)</b> , UMBC Taught the use of computer vision and Kinect in user interfaces	2011
	<b>Introduction to Human Computer Interaction (CMSC 434)</b> , UMD Taught principles of visual design	2011
	<b>Multimodal User Interfaces (6.870)</b> , MIT Taught multimodal user interfaces on mobile devices	2007, 2008
<i>Teaching Assistant</i>	<b>Structure and Interpretation of Computer Program (6.001)</b> , MIT Led 7 sections every week, developed custom exercises and quizzes.	2006

## Advising

<i>PhD Thesis Committee</i>	Tsung-Hsiang Chang, MIT ( <i>Google Fellowship in HCI</i> )	2009-2011
<i>PhD</i>	Brandyn White, UMD	2009-2011
<i>Research</i>	Ejaz Ahmed, UMD	2011
<i>Mentor</i>	Aniruddha Kembhavi, UMD	2009-2010
	Jingchen Liu, PSU	2010-2011
<i>Undergrad</i>	Eric Kuang, UMD	2010-2011
<i>Research</i>	Leyla Nornooz, UMD	2011
<i>Supervisor</i>	Noor Siddiqi, UMD	2010
	John Lee, MIT ( <i>M. Eng Thesis Prize, 1st place</i> )	2008-2009
	Allan Deckelbaum, MIT	2006
	Yunus Samaz, MIT	2006

## References

**Trevor Darrell**

Professor  
University of California, Berkeley  
trevor@eecs.berkeley.edu

**Rob Miller**

Associate Professor  
Massachusetts Institute of Technology  
rcm@mit.edu

**Boris Katz**

Principal Research Scientist  
Massachusetts Institute of Technology  
boris@mit.edu

**Larry Davis**

Professor  
University of Maryland, College Park  
lsd@cs.umd.edu

**Ben Bederson**

Professor  
University of Maryland, College Park  
bederson@cs.umd.edu