

Yang Cai

Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, USA

Phone: (412) 608-2152
Email: caiyang@cs.cmu.edu
Office: 4522 Newell Simon Hall
Homepage: <http://www.cs.cmu.edu/~caiyang>



Education

- **Master of Science in Computer Science** 9 / 2012 ~ Present (expected graduation: 5 / 2014)
 - Robotics Institute, School of Computer Science, Carnegie Mellon University, USA.
 - Advisor : Dr. Alexander Hauptmann (Principal Systems Scientist of CMU and Director of InforMedia Lab).
 - GPA : 4.0/4.0.
- **Bachelor of Engineering in Computer Science** 9 / 2007 ~ 6 / 2011
 - College of Computer Science and Technology, Zhejiang University, China.
 - Advisor : Prof. Qunsheng Peng (Director of State Key Lab of Computer-Aided Design & Computer Graphics).
 - GPA : 3.84/4.0.

Work Experience

- Graduate Research Assistant at Carnegie Mellon University 9 / 2012 ~ Present
- Visiting Research Scholar at Carnegie Mellon University 10 / 2011 ~ 8 / 2012
- Research Intern at Microsoft Research Asia 7 / 2010 ~ 9 / 2011

Project Experience

- **InforMedia Lab, Carnegie Mellon University**
 - **TRECVID¹ Multimedia Event Detection (MED) Task** 7 / 2012 ~ Present
 - * **Project description:** To detect videos with pre-specified events (defined by text descriptions and video examples) in a collection of videos.
 - * **Responsibilities:** Research and development of low-level visual feature encoding methods (e.g., Fisher Vector (FV) encoding).
 - * **Achievements:** (1) **Ranked 1st** in the pre-specified event detection at the TRECVID MED 2012; (2) FV encoding significantly improved the mean average precision (MAP) of event detection compared to traditional spatial bag-of-words encoding; (3) FV encoding greatly reduced detection time because only linear classifiers were needed.
 - **Prototype System for Multimedia Event Detection (MED)** 3 / 2013 ~ Present
 - * **Project description:** To develop an integrated web-based MED prototype system.
 - * **Responsibilities:** Design and development of the interface and most key functionalities of the prototype system.
 - * **Achievements:** Implemented the beta version of the system which offered useful features including: (1) the system was purely content-based, where videos were represented by automatically extracted semantic features (e.g., detected concepts, recognized text, etc.) and non-semantic features (e.g., low level visual features), (2) detections were done by leveraging both information retrieval techniques (e.g., text indexing/retrieval, content-based image retrieval) and machine learning techniques (e.g., classification using discriminative classifiers).
 - **TRECVID Surveillance Event Detection (SED) Task** 1 / 2012 ~ Present
 - * **Project description:** To detect pre-specified events (e.g. person runs) in a long surveillance video stream.
 - * **Responsibilities:** Low-level visual feature extraction, detection framework development and experimental evaluation.
 - * **Achievements:** (1) **Ranked 1st** in retrospective (automatic) SED and **ranked 2nd** in interactive SED at TRECVID 2012; (2) Built an interface which enabled effective detection results visualization and efficient user interaction for the interactive SED task; (3) Published two papers ([Cai-TRECVID12] and [Chen-MM13]).

¹TREC Video Retrieval Evaluation: an international competition with a series of tasks which are devoted to research in automatic segmentation, indexing, and content-based retrieval of digital video. TRECVID is sponsored by the National Institute of Standards and Technology (<http://trecvid.nist.gov/>).

- **Media Computing Group, Microsoft Research Asia**

- **Object Retrieval using Visual Query Context**

7 / 2010 ~ 11 / 2010

- * **Project description:** To more accurately estimate user's search intention from a user-provided bounding box in a query image for content-based object retrieval.
 - * **Responsibilities:** Algorithm implementation and experimental evaluation.
 - * **Achievements:** (1) Implemented a search intention estimation method based on Grab-Cut image segmentation algorithm; (2) Published one paper ([Yang-TMM11]).

- **Video-based Image Retrieval**

12 / 2010 ~ 3 / 2011

- * **Project description:** To develop a new image retrieval system where users are allowed to submit a short video clip as a query for image retrieval.
 - * **Responsibilities:** Approach design and implementation, system development and experimental evaluation.
 - * **Achievements:** (1) Developed a video-based image retrieval system which was significantly more reliable than retrieval using a single image as a query; (2) Published two papers ([Yang-MM11] and [Yang-IJMIR12]).

- **Content-based Million-scale Near-duplicate Video Retrieval**

4 / 2011 ~ 9 / 2011

- * **Project description:** To develop a large scale near-duplicate video retrieval system which serves a database comprised of one million web videos.
 - * **Responsibilities:** Efficient retrieval method design, system development and experimental evaluation.
 - * **Achievements:** (1) Implemented a system that can respond to a query in 41ms with 78.4% mean average precision (MAP) on average using the highly efficient retrieval approach; (2) Built a web-based near-duplicate video retrieval system based on the proposed techniques and showed a demo of the system at the ACM Multimedia Conference 2011; (3) Published two papers ([Cai-MM11] and [Cai-IEEEMM12]).

Skills

- Specialties: Computer Vision, Machine Learning, Information Retrieval and their applications in Multimedia (e.g. video, image, text, etc.) analytic, understanding, indexing and retrieval.
- Programming Languages: C, C++, C#, Python, Matlab, HTML, PHP, Javascript and jQuery.
- Languages: English (fluent) and Mandarin (native).

Awards and Honors

- Research Assistantship, Carnegie Mellon University, 2012 ~ 2014.
- Student Travel Grants, ACM Multimedia, 2013.
- Intern Excellence Award, Microsoft Research Asia, 2011.

Publications

- [Cai-MM11]: **Yang Cai**, Linjun Yang, Wei Ping, Fwei Wang, Tao Mei, Xian-Sheng Hua, Shipeng Li, "Million-scale Near-duplicate Video Retrieval System", ACM Multimedia (MM), Arizona, USA, 2011.
- [Yang-MM11]: Linjun Yang, **Yang Cai**, Alan Hanjalic, Xian-Sheng Hua, Shipeng Li, "Video-based Image Retrieval", ACM Multimedia (MM), Arizona, USA, 2011.
- [Yang-TMM11]: Linjun Yang, Bo Geng, **Yang Cai**, Alan Hanjalic, Xian-Sheng Hua, "Object Retrieval using Visual Query Context", IEEE Transactions on Multimedia, 2011.
- [Cai-ICMR12]: **Yang Cai**, Wei Tong, Linjun Yang, Alexander Hauptmann, "Constrained Keypoint Quantization: Towards Better Bag-of-Words Model for Large-scale Multimedia Retrieval", ACM International Conference on Multimedia Retrieval (ICMR), Hong Kong, China, 2012.
- [Ma-MM12]: Zhigang Ma, Yi Yang, **Yang Cai**, Nicu Sebe, Alexander Hauptmann, "Knowledge Adaptation for Ad Hoc Multimedia Event Detection with Few Exemplars", ACM Multimedia (MM), Nara, Japan, 2012.
- [Cai-TRECVID12]: **Yang Cai**, Qiang Chen, Lisa Brown, Ankur Datta, Quanfu Fan, Rogerio Feris, Shuicheng Yan, Alexander Hauptmann, Sharath Pankanti, "CMU-IBM-NUS@TRECVID 2012: Surveillance Event Detection", TRECVID Video Retrieval Evaluation Workshop, NIST, Gaithersburg, MD, USA, 2012.
- [Cai-IEEEMM12]: **Yang Cai**, Linjun Yang, "Large-scale Near-duplicate Web Video Retrieval – Challenges and Approaches", IEEE Multimedia, 2012.
- [Yang-IJMIR12]: Linjun Yang, **Yang Cai**, Alan Hanjalic, Xian-Sheng Hua, Shipeng Li, "Searching for images by video", International Journal of Multimedia Information Retrieval, 2012.
- [Chen-MM13]: Qiang Chen, **Yang Cai**, Lisa Brown, Ankur Datta, Quanfu Fan, Rogerio Feris, Shuicheng Yan, Alex Hauptmann and Sharathchandra Pankanti, "Spatio-Temporal Fisher Vector Coding for Surveillance Event Detection", ACM Multimedia (MM), Barcelona, Spain, 2013.
- [Cai-MM13]: **Yang Cai**, Yi Yang, Alexander Hauptmann, Howard Wactlar, "A Cognitive Assistive System for Monitoring the Use of Home Medical Devices", ACM Multimedia Workshop (MM MIIRH), Barcelona, Spain, 2013.

Talks

- "Multimedia Event Detection Prototype System", IARPA Project Site Visit, Carnegie Mellon University, USA, 2013.
- "CMU-IBM-NUS@TRECVID 2012: Surveillance Event Detection", TRECVID Workshop, National Institute of Standards and Technology (NIST), USA, 2012.