Shih-Fu Chang

Morris A. and Alma Schapiro Professor
Dean**,** School of Engineering and Applied Science
Columbia University

Address:
500 W. 120th St, Mudd Building Rm 510, Columbia University, New York, NY 10027
URL: https://www.engineering.columbia.edu/faculty/shih-fu-chang
phone: 1-212-854-1123, email: shih.fu.chang@columbia.edu

Research Areas:
Computer vision, machine learning, multimedia analysis and retrieval.

Education:

Ph.D. in Electrical Engineering & Computer Sciences, University of California at Berkeley, 1993.
M.S. in Electrical Engineering & Computer Sciences, University of California at Berkeley, 1991.
B.S. in Electrical Engineering, National Taiwan University, 1985.

Professional Experience:

- Dean (2022-date), Interim Dean (2021-2022), Senior Executive Vice Dean (2015-2021), Senior Vice Dean (2012-2015), The Fu Foundation School of Engineering and Applied Science, Columbia University.

- Richard Dicker Chair Professor (2011-2022), started as Assistant Professor in 1993, Department of Electrical Engineering, Columbia University.

- Professor, Department of Computer Science, Columbia University, 2011-date.

- Amazon Scholar, 2020-2022.

- Director, Digital Video and Multimedia (DVMM) Laboratory, Columbia University, http://www.ee.columbia.edu/dvmm

- Director (2000-2003), Co-PI (1993-1999), *ADVENT University-Industry Research Consortium,* Columbia University (more than 25 sponsors in media and information technology areas)

- Co-Founder, Axon Image Inc., 2015-2018.

- Scientific Advisor, VidRovr Inc. (startup from the DVMM lab), 2016-2020.

- Advisory Board Member/Consultant, Eastman Kodak (1998), PictureTel (1998), iBeam (2000-2002), Sekani (2000), SOSi (2000-2002), Chonghwa Telecom (2003-06), InspireWork (2003), Motionbox (2007-2008), Institute of Information Industry (III) Taiwan (2006-2018), Academia Sinica (2014-date), and Industrial Technology Research Institute (ITRI) Taiwan (2016-2022).

- Visiting Faculty, IBM T.J. Watson Research Center, Hawthorn NY, 2004 –2006.

- Visiting Faculty, Microsoft Research Asia, 2006.

Honors and Awards:

- ***Member, National Academy of Engineering, 2023.***

- ***Member, National Academy of Inventors, 2022.***

- ***Academician, Academia Sinica, 2018.***

- ***ACM Fellow, 2017.***
  for contributions to large-scale multimedia content recognition and multimedia information retrieval.
  http://engineering.columbia.edu/news/shih-fu-chang-acm-fellow

- ***Most Influential Scholar in the Field of Multimedia, based on AMiner, 2016.***
  https://aminer.org/mostinfluentialscholar/mm

- ***Honorary Doctorate, University of Amsterdam, 384th Anniversary, 2016.***
  for "his pioneering contribution to our understanding of the digital universe, particularly in the areas of imagery, language, and sound."
  award ceremony video (presentation start time: 00:57:20)
  http://webcolleges.uva.nl/Mediasite/Play/8f1fbec0737f4755b705632624a448fc1d

- ***Great Teacher Award, The Society of Columbia Graduates, 2013.***

- ***Technical Achievement Award, IEEE Signal Processing Society, 2012.***
  for pioneering contributions to signal processing for multimedia content analysis and retrieval. IEEE SPS is the largest professional society for signal processing worldwide.

- ***Technical Achievement Award, ACM Special Interest Group in Multimedia, 2011.***
  for pioneering research and inspiring contributions in multimedia analysis and retrieval.

- ***Fellow, AAAS (the American Association for the Advancement of Science), 2010.***
  for contributions to multimedia content analysis and search.
  http://news.columbia.edu/oncampus/2275

- ***IEEE Kiyo Tomiyasu Award,** (IEEE-wide Award), 2009.*
  for pioneer contributions to image classification and search.

- ***IEEE Fellow, 2004.***
  for contributions to digital video and multimedia technologies.

- ***Meritorious Service Award, IEEE Signal Processing Society, 2011.***

- ***Recognition of Service Award**, Association for Computing Machinery (ACM), 2000.*

- ***IBM Faculty Award**, 1995, 1999.*

- ***Distinguished Lecturer**, IEEE Circuits and Systems Society, 2001-2002.*

- ***Young Investigator Award**, Office of Naval Research (ONR), 1998-2001.*

one of two ONR/YIA awards in the field of information, computer sciences, and mathematics.

- *CAREER Award*, National Science Foundation, 1995-98.

- **32 Issued Patents**, more than 10 patents licensed to industry.

- **Publication Impact:**
  68,000+ citations, h-index: 130 (Google Scholar link)

Research Summary and Impacts:

Chang's research is dedicated to multimedia information retrieval, computer vision, machine learning, and multimodal analysis. The primary focus of his work is on development of intelligent methods and systems for extracting information from visual and multimodal content that are prevalent in large archives and live sources. In the early 90's, his group developed some of the earliest and best-known content-based image search systems, VisualSEEk and VideoQ [1,2], which set the foundation of this vibrant area. During last two decades, he has made significant contributions to the field of multimedia retrieval by developing large scale classifiers and multimedia ontology construction [3-6]. These have strongly influenced design of the video search systems used in practice [7]. The image search system developed by his group has demonstrated the best performance in NIST image provenance and filtering evaluation three years consecutively [13]. He has developed influential compact hashing techniques [11] for efficient search over billion-scale image databases. His compact hashing work has enabled order of magnitude speedup and storage improvement in applications such as an online human trafficking crime fighting system that has been deployed in 200+ law enforcement agencies [12]. In addition, he has developed a series of fundamental methods of graph-based semi-supervised learning [8,10] that successfully address the challenge issues of training large-scale retrieval systems with noisy and sparse labels. These methods have been adopted to build the first commercial brain machine interface system [9] for rapid image retrieval using noisy brain signals. Recently, he has devoted efforts to research of multimodal knowledge graphs (with funding from several DARPA programs) that leverage multimodal information in extracting events, entities, and relations from multimodal content and use them in applications such as Question and Answer and Inferencing [14]. Chang is also active in cross-disciplinary research, including the recent development of a system in collaboration with social scientists in understanding of social media communication of Chicago youth members who are potentially involved in gang violence [15]

Representative Papers and Impacts:

1. Smith, John R., and Shih-Fu Chang. "VisualSEEk: a fully automated content-based image query system." In ACM international conference on Multimedia, pp. 87-98. ACM, 1997.
2. Chang, Shih-Fu, William Chen, Horace J. Meng, Hari Sundaram, and Di Zhong. "A fully automated content-based video search engine supporting spatiotemporal queries." Circuits and Systems for Video Technology, IEEE Transactions on 8, no. 5 (1998): 602-615.
   [IEEE T-CSVT Best Paper Award]

The Web, as it started in the early 90's, introduced the unprecedented possibility for users to browse and search information from massive online sources. However, the breakthrough in information retrieval was initially only limited to text documents via textual query. The above two works were among the earliest and best known systems to break such limitation by developing novel query mechanisms that allow users to search images or videos by directly composing visual objects and their spatio-temporal attributes (such as color, texture, shape, layout, and motion trajectory). The content-based feature matching and the efficient inverted file indexing methodologies were broadly followed. These systems and subsequent works of Chang helped set the foundation of this vibrant field. Today, visual search has become a common feature in many products such as mobile product search from Google, Amazon, Pinterest, and many others.

3. With Naphade, Milind, et al., "Large-scale concept ontology for multimedia." MultiMedia, IEEE 13, no. 3 (2006): 86-91.
4. Borth, Damian, Rongrong Ji, Tao Chen, Thomas Breuel, and Shih-Fu Chang. "Large-scale visual sentiment ontology and detectors using adjective noun pairs." In Proceedings of ACM international conference on Multimedia, pp. 223-232. ACM, 2013.
5. Ye, Guangnan, Yitong Li, Hongliang Xu, Dong Liu, and Shih-Fu Chang. "Eventnet: A large scale structured concept library for complex event detection in video." In Proceedings of ACM international conference on Multimedia, pp. 471-480. ACM, 2015.
6. Li, Hongzhi, Joseph G. Ellis, Heng Ji, and Shih-Fu Chang. "Event specific multimodal pattern mining for knowledge base construction." In Proceedings of ACM on Multimedia Conference, pp. 821-830. ACM, 2016.

Following the early research efforts in content-based visual search, the series of works listed above further set the directions of multimedia research in the next decade. Breaking the limits of the existing practices relying on a small number of classes, our efforts aimed to develop systematic and sound strategies for constructing multimedia ontologies and developing automatic classifiers for detecting the vast number of concepts defined in the ontologies. The research aims to achieve the holistic optimization in using the ontology and classifiers to improve the overall search performance and experience – for example, how to answer free text queries over video databases by mapping the free-text query sentences to the defined multimedia concepts, how to support rapid intuitive browsing over the ontologies, and how to link the concepts in the ontologies to higher level targets such as emotions or events depicted in the multimedia content.

[3] presented the first multimedia ontology of about 1,000 concepts developed jointly with domain experts and librarians to support news video search. The ontology has been adopted for the international video search evaluation forum, TRECVID, organized by NIST starting 2005. [4] extended beyond the common semantic categories and developed the first ontology of 1,200 visual concepts (SentiBank) strongly related to sentiment and emotions expressed in the multimedia content. Such emotion related concepts were mined automatically from user generated content on online forums. SentiBank has been shown effective in predicting the visual sentiments and the popularity of visual content on social

media. [5] was the first to develop a Wikipedia-style ontology of 500 high-level events and 4,490 event-specific concepts to enhance browsing and searching of video events. Finally, [6] addressed the fundamental challenge of automatic construction of multimedia ontologies by multimodal pattern mining and naming based on data-driven approaches. It has shown very promising directions and has the potential to open another productive era of multimedia research in the coming years.

7. With Amir, Arnon, et al.,"IBM research TRECVID-2003 video retrieval system." NIST TRECVID Video Retrieval Evaluation (2003).

   Through close collaboration with industry collaborator, we made significant contributions to the development of a large video search system, called IMARS, commercialized by IBM. It incorporated a comprehensive set of new video search techniques, including large classifier libraries for detecting concepts defined in the multimedia ontology, ontology-based concept fusion, content-based image retrieval, mapping natural language queries to multimedia concepts, and interactive concept-based relevance feedback. The IBM system received the Wall Street Journal Technology Innovation Award in 2004. The ontology-assisted concept based search features have been adopted in many other video search systems today such as those of AT&T, BBN, and SRI.

8. Wang, Jun, Tony Jebara, and Shih-Fu Chang. "Semi-supervised learning using greedy max-cut." *Journal of Machine Learning Research* 14, no. Mar (2013): 771-800.
9. Wang, Jun, Eric Pohlmeyer, Barbara Hanna, Yu-Gang Jiang, Paul Sajda, and Shih-Fu Chang. "Brain state decoding for rapid image retrieval." In Proceedings of the 17th ACM international conference on Multimedia, pp. 945-954. ACM, 2009.
10. Wu, Xiao-Ming, Zhenguo Li, Anthony M. So, John Wright, and Shih-Fu Chang. "Learning with partially absorbing random walks." In *Advances in Neural Information Processing Systems (NIPS)*, pp. 3077-3085. 2012.

   One of the most promising retrieval paradigms for multimedia is search by examples with relevance feedback. In such settings, users select initial examples and use them as query seeds, which are used to trigger the information diffusion and label propagation processes over the entire dataset. The problem can be formulated as transductive or semi-supervised learning. We have developed a series of fundamental methods that use graph-based learning to derive optimal solutions for such diffusion processes. In [8], we addressed the problem of noisy initial labels that are frequently encountered in practice and may severally degrade the search results. We developed a multivariate optimization method to simultaneously diagnose and refine the initial noisy labels and diffuse the refined labels to the remaining samples on the graph. We developed an elegant solution equivalent to the well-known greedy max-cut with linear constraints. It showed superior performance in image retrieval and classification with sparse and noisy initial labels. Our solutions have broad applications and have been adopted to build the first image search system using brain machine interfaces (BCI) without any manual keyboard interactions [9]. Neural responses of the user to initial images displayed on the screen are processed to detect user

attention to the visual stimuli. Such attention labels of the seed images are then used to start the label diffusion process using our graph-based semi-supervised learning technique. The final diffusion results over the entire database are used to retrieve all the target images. The BCI image search system is rapid and effective as it intelligently combines the strengths of the human vision system and computer vision processes. The resulting search prototype, in collaboration with Prof. Paul Sajda's team, has been licensed to the company NeuroMatters and deployed in applications of aerial surveillance and security.

We have further extended the graph-based learning paradigm by developing a novel ranking process, called partially absorption random walks (PARW) [10]. We showed existing graph-based semi-supervised label propagation methods and many other models like PageRank and hitting/commute times can be elegantly unified under PARW. Furthermore, it has strong desirable local clustering properties in the sense that the final absorption probability of the random walk process will be mostly contained within a local cluster surrounding the initial starting point. Besides the proven superior performance for image retrieval, this method has recently been validated and deployed by Huawei in their large mobile app recommendation system, in which PARW is used to recommend apps from a pool of more than 500 million apps to more than 300 million users [H. Guo et al., *"Graph-based Push Service Platform,"* DASFAA'17].

11. Liu, Wei, Jun Wang, Rongrong Ji, Yu-Gang Jiang, and Shih-Fu Chang. "Supervised hashing with kernels." In *Computer Vision and Pattern Recognition (CVPR)*, IEEE Conference on, pp. 2074-2081, 2012.

Multimedia retrieval presents several unique challenges distinct from the traditional text or structured data search. First, there are no known discrete tokens like words or attributes that can be used to represent diverse multimedia content. Second, features and metrics used in multimedia matching are high dimensional and ambiguous. Such challenges call for new approaches to efficient multimedia representation based on compact binary codes that can reduce the search complexity and storage cost while preserving the original content similarity relations. We have developed a family of learning based hashing techniques to reduce the features to the order of 100 bits by utilizing the massive amount of image data and labels (if available) to learn the optimal hashing functions. We have developed supervised, semi-supervised, and unsupervised hashing functions for different learning scenarios. The above paper presented a novel learning-based technique, called supervised kernel hashing (KSH), which uses pair-wise label relations (same class or different class, or user click logs indicating relevance) to learn optimal hashing functions in the nonlinear kernel space. The optimization is done in a sequential way to minimize the redundancy among different hash bits. KSH and our other works on image hashing have been broadly used and shown effective in many large scale applications, including our implementation of image search over 400 million images in the application of fighting human trafficking crimes (see item 12 below).

12. With Pedro Szekely, et al., "Building and Using a Knowledge Graph to Combat Human Trafficking," In International Conference on Semantic Web (ICSW), Oct. 2015.
[Best Applied Paper Award]

> This paper describes the DIG system developed as part of the DARPA MEMEX project, in which we led the important task of scalable search over more than 400 million images. The initial application domain was human trafficking crime fighting. We leveraged our prior work of visual sentiment analysis and SentiBank visual concept classifiers to develop effective image representation, which showed excellent performance in matching image styles and appearances used in images in the human trafficking domain. To meet the needs of large-scale search, we applied our compact hashing techniques to achieve real-time search performance over 400 million images. Currently, the DIG system has been deployed in more than 200 law enforcement agencies and several NGOs that are using the system in various ways to fight human trafficking. Reports to date indicate that the system has already been successfully used to identify victims of human trafficking. Additional information can be found at
> http://www.ee.columbia.edu/dvmm/memex/ .

Translational Impacts:

Our research has resulted in 32 patents issued and several pending in the areas of brain machine interface for image retrieval, mobile visual search, video search, video editing, image authentication, MPEG-7, and MPEG-21 international standards. More than ten video technologies developed in our research have been licensed to companies, including BBN, Real Networks, Inspired Works, III, NeuroMatters, Digital Genesis, and VidRovr. Three startup companies applying technologies from our research have been created. Through the ADVENT consortium, we have conducted collaborative research sponsored by more than 25 companies. Results of our research are also used in a few large-scale government funded projects in collaboration with industry partners, such as Kitware, NeuroMatters, IBM, and BBN.

Entrepreneurship:

1. **VidRovr Inc. (Startup, 2016)**

VidRovr builds multimodal information extraction techniques and systems for intelligent video search. It provides novel techniques to link video across news video channels and social media sources. It also uses novel pattern discovery techniques to automatically discover semantic objects unique to each event and train machine learning classifiers, answering the long open challenge of manually defining ontologies for new application domains. VidRovr has won a prestigious startup prize Publicis90 in 2016 (ranked 12 out of 3,500+ submissions worldwide), a NSF SBIR grant for technology commercialization and selection into the Techstars cohort 2017. Its products are being used by major media companies, like Verizon, Forbe, AP, and major broadcast TV companies.

2. **Axon Image Inc. (Startup, 2015)**

The goal of Axon Image is to democratize the photo-to-market process and help non-professional photographers easily disseminate their personal photo collection to the stock photo marketplace. Traditional systems favor professional users who have rich resources for high-end photo processing and

distribution. Axon Image builds upon the intelligent techniques for automatically recommending high-quality photos, generating keyword tags and description titles, and connecting to a large number of stock photo destinations for quick dissemination. It addresses the dissemination disparity by providing one-stop end-to-end solutions helping massive users efficiently share their multimedia content.

3. **EyeStyle (Startup, 2017)**

Despite the major advances of image recognition in recent years, object search using visual query on light-weight mobile devices remains an open challenge. Besides the requirements for high search accuracy, there are additional requirements for reducing the computing/memory complexity as well as the communication cost. EyeStyle utilized the novel compact hashing and 3D visual retrieval techniques developed in the DVMM lab. It received funding from the COMBINE program (sponsored by NYCEDC Media Lab) and the Verizon grand challenge program.

Keynote/Invited Speeches:

(recent list and slides at http://www.ee.columbia.edu/~sfchang/talks.html)

- Keynote Talk, IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Multimodal Video Analysis. June 2019.
  Multimodal Knowledge Graphs: Automatic Extraction & Applications

- Keynote Talk, IEEE Conference on Computer Vision and Pattern Recognition, Workshop on Media Forensics. June 2019.
  GAN Deepfake Image Detection: Up-Sampling Artifact & GAN Pipeline Emulator

- Keynote Talk, ECCV Workshop on Compact and Efficient Feature Representation and Learning in Computer Vision, September 2018
  "Robust and Compact Deep Learning Features for Image Matching and Retrieval".

- Keynote Talk, CVPR Workshop on ActivityNet, June 2018
  "Video Event Localization, Supervision, and Online Detection".

- Keynote Talk, ACM Turing 50th Celebration Conference, Symposium on Multimedia, Shanghai, May 2017.

- Invited Talk, Symposium on Machines and Media, NYC Media Lab, April, 2017.
  "AI for Image/Video Content Search and Understanding".

- Invited Talk, Joint ACM Multimedia Conference and European Conference on Computer Vision, Oct. 2016.
  "Video Event Detection, Localization and Cross-Media Linking"

- Keynote Talk, IEEE Multimedia Signal Processing (MMSP), September, 2016.

- Keynote Talk, ACM International Conference on Multimedia Retrieval (ICMR), June 2016.
  "New Frontiers of Large Scale Multimedia Information Retrieval"

- Invited talk, University of Amsterdam, Jan. 2016.
  "Evolution of Multimedia Content Understanding - from Categorization to Emotion and Knowledge"

- Keynote Talk, ACM Multimedia Conference, Workshop on Crowdsourcing Multimedia, Brisbane, Australia, Oct. 2015.
- Distinguished Lecture, Department of Electrical Engineering, UCLA, Jan. 2015.
  "Recent Advances of Compact Hashing for Large-Scale Search"
- Invited Talk, DARPA Workshop on Multimodal Information Analysis, March 2015.
- Invited Talk, European Conference on Computer Vision (ECCV), Workshop on Parts and Attributes, Sept. 2014.
  "Concept-Based Framework for Detecting High-Level Events in Video"
- Invited Talk, European Conference on Computer Vision (ECCV), Workshop on Storytelling with Images and Videos, Sept. 2014.
  "Assistive Image Comment Robot – Teaching Machines to Suggest Social Comments Relevant to Image Content"
- Keynote Speech, IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP 2014), August 2014.
- Distinguished Lecture, Florida Atlantic University College of Engineering and Computer Science, April, 2014.
- Keynote Speech, SPIE/Electronic Imaging, Human Vision and Electronic Imaging (HVEI) conference, Feb. 2014.
  "Images Shared in Social Media: a Window into Human Sentiment and Emotion"
- Keynote Speech, Bay Area Multimedia Forum, Feb. 2014.
- Invited Talk, NSF/FBI/DARPA sponsored workshop on Frontiers in Image and Video Analysis, Jan. 2014.
- Invited Talk, Army Research Office Visual Search Workshop, December 2013.
- Plenary Talk, IEEE Computer Vision and Pattern Recognition (CVPR) Conference, Big Data Computer Vision Workshop, Portland, OR, June 2013.
  "Large-Scale Concepts and Classifiers for Describing Visual Sentiment on Social Multimedia"
- Invited Talk, Annual Conference on Neural Information Processing Symposium (NIPS), Workshop on BigVision, Lake Tahoe, NV, Dec. 2012.
  "Recent Advances of Compact Hashing for Large Scale Visual Search"
- Invited Talk, European Conference on Computer Vision (ECCV) Workshop on Web-Scale Vision and Social Media, Florence, Italy, Oct. 2012.
  "Recent Advances of Compact Hashing for Large Scale Visual Search"
- Keynote Speech, ACM SIGMM Technical Achievement Award, ACM Multimedia, Nov. 2011.
  "Content Based Multimedia Retrieval: Lessons Learned from Two Decades of Research"
- Distinguished Lecture, Department of EECS, Northwestern University, Feb. 2011.
  "Large Scale Mobile Visual Search"
- Keynote Speech, IBM Emerging Leader Workshop on Multimedia, IBM, Oct. 2010.

"Brain State Decoding for Rapid Image Retrieval"

List of Best Paper/Demo Awards:

1. **Best Demo Paper Award,** the 58th Annual Meeting of the Association for Computational Linguistics (ACL) July, 2020.

   Manling Li*, Alireza Zareian*, Ying Lin, Xiaoman Pan, Spencer Whitehead, Brian Chen, Bo Wu, Heng Ji, Shih-Fu Chang, Clare Voss, Daniel Napierski and Marjorie Freedman (*equal contributions).
   "GAIA: A Fine-grained Multimedia Knowledge Extraction System."

2. **Best Poster Paper Award**, ACM on International Conference on Multimedia Retrieval (ICMR), 2018.

   Li, Hongzhi, Joseph G. Ellis, Lei Zhang, and Shih-Fu Chang.
   "Patternnet: Visual pattern mining with deep neural network."

3. **Best Multimodal Paper Award**, ACM International Conference on Multimedia Retrieval (ICMR), 2016.

   Nikolaos Pappas, Miriam Redi, Mercan Topkara, Brendan Jou, Hongyi Liu, Tao Chen, Shih-Fu Chang. "Multilingual Visual Sentiment Concept Matching."
   In ACM International Conference on Multimedia Retrieval (ICMR), 2016.

   This paper presented the first study on how online users of 12 different languages describe emotion-related concepts in images shared on the social media. We applied word embedding of visual concepts over 7.4 million images of 16K multilingual concepts into a common representation space, allowing multilingual visual concepts to be efficiently matched and clustered. The results offered novel insights about the commonality and difference of affect expression through online images across languages.

4. **Best Paper Award Honorable Mention**, IEEE Transactions on Multimedia, 2016.

   Yu-Gang Jiang, Jun Wang, Xiangyang Xue, Shih-Fu Chang.
   "Query-Adaptive Image Search with Hash Codes." IEEE Transactions on Multimedia, 15(2):442-453, 2013.

   This paper proposed a novel method changing how existing binary hashing codes were used to compute distances between images and how such distances were used to search large image databases. Instead of treating all hash bits equally, the importance of each hash bit is adapted optimally on the fly during the query time based on the semantic category of the query target. Experiments over large image datasets confirmed clear performance gains of the proposed method.

5. **Best Applied Paper Award**, the 14th International Semantic Web Conference (ISWC), Oct. 2015.

P. Szekely, C. Knoblock, J. Slepicka, A. Philpot, A. Singh, C. Yin, D Kapoor, P. Natarajan, D. Marcu, K. Knight, D. Stallard, S. S. Karunamoorthy, R Bojanapalli, S. Minton, B. Amanatullah, T. Hughes, M. Tamayo, D. Flynt, R. Artiss, S.-F. Chang, T. Chen, G. Hiebel and L. Ferreira, "Building and Using a Knowledge Graph to Combat Human Trafficking," In International Conference on Semantic Web (ICSW), Pennsylvania, Oct. 2015

This paper described an integrated system, DIG, for crawling, ingesting, and searching large-scale heterogeneous data from the Web. The fast image similarity search tool was developed by my group using novel compact hashing techniques to significantly reduce storage and computation complexity, and sentiment-aware image features to increase search accuracy over a very large database of 400 million images. The system has been used by more than 200 law enforcement agencies and several non-governmental organizations for countering online human trafficking crimes.

6. **Best Paper Award**, ACM International Conference on Multimedia Retrieval, 2014.

   Subhabrata Bhattacharya, Felix X. Yu, Shih-Fu Chang.
   "Minimally Needed Evidence for Complex Event Recognition in Unconstrained Videos." In ACM International Conference on Multimedia Retrieval (ICMR), April 2014.

   We developed a novel framework to understand needed evidences for human judgements of event occurrence in videos. The finding that one microshot (2 seconds long video) is sufficient in many cases for correct identification or rejection of events is surprising and important. It influences how researchers approach the challenging problem of event representation and event localization in untrimmed videos encountered in real world applications.

7. **Best Demo Award**, NYC Media Lab Annual Summit, Sept. 2014.
   First Prize, ACM Multimedia Conference Grand Challenge, 2013.
   Best Demo Award, the 3rd Greater New York Area Workshop on Multimedia and Vision, June 2013.
   New York Tech MeetUp, Nov. 2013.

   Brendan Jou, Hongzhi Li, Joseph G. Ellis, Daniel Morozoff, and Shih-Fu Chang, "News Rover: Exploring Topical Structures and Serendipity in Heterogeneous Multimedia News".

   News Rover is an intelligent news video aggregation and search system that integrates various multimodal information extraction techniques, including video story linking across broadcast channels and social media sources, multimedia named entities linking and labeling, and interactive dashboard for accessing who&what key information in massive archives. The system has won several awards listed above and formed the technical core of the new startup company VidRovr.

8. **Best Paper Award**, ACM Multimedia Conference, Nov. 2011.

   F. X. Yu, R.-R. Ji, and S.-F. Chang, "Active Query Sensing for Mobile Location Search," ACM Multimedia Conference, Scottsdale, AZ, Nov. 2011.

Presented a novel platform and solution for determining the best viewing angle for capturing mobile image queries for recognizing the geo-locations based on automatic image matching. The analytical formulation was designed to maximize the uncertainty reduction by using subsequent mobile visual query. An end-to-end system was developed to demonstrate the complete capabilities on both the server (searching over 300,000 street view images) and the client apps on the mobile devices.

9. **Most Cited Paper of the Decade Award**, the Journal of Visual Communication and Image Representation, 2010.
Yong Rui, Thomas Huang, and Shih-Fu Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," Journal of Visual Communication and Image Representation 10 (1999), 39–62.

10. **Best Paper Award**, ACM Conference on Image and Video Retrieval (CIVR), 2008.
E. Zavesky, S.-F. Chang, C.-C. Yang, "Visual Islands: Intuitive Browsing of Visual Search Results," ACM International Conference on Image and Video Retrieval, Niagara Falls, Canada, 2008.

This paper described novel algorithms and interfaces for summarizing large visual search results according to dynamically discovered dominant concepts contained in the search result sets.

11. **Best Paper Finalist**, ACM Multimedia Conference, Oct. 2009.
W. Jiang, C. Cotton, S.-F. Chang, D. Ellis, and A. Loui, "Short-Term Audio-Visual Atoms for Generic Video Concept Classification," ACM Multimedia Conference, Beijing, Oct. 2009.

This paper described a novel framework for extracting short-term temporal regions from the video track, and fusing such visual cues with the atomic signals extracted from the audio. Such joint audio-video representations at the atomic level facilitate robust detection of multimedia concepts (such as horseback riding and music performance) and importantly provide intuitive explanations for automatic detection models of multimodal concepts.

12. **Best Paper Finalist**, ACM Multimedia Conference, Oct. 2006.
Winston Hsu, Lyndon Kennedy, Shih-Fu Chang, "Video Search Reranking via Information Bottleneck Principle," In ACM Multimedia, Santa Barbara, CA, USA, 2006.

This paper described a content-based information theoretic approach to reranking the noisy search results returned by initial keyword based search. It offers a general and powerful way to combine visual analysis with conventional text based image search engines.

13. **Best Student Paper Award,** ACM Multimedia Conference 2005.
Tian-Tsong Ng, Shih-Fu Chang, Jessie Hsu, Lexing Xie, Mao-Pei Tsui, "Physics-Motivated Features for Distinguishing Photographic Images and Computer Graphics," ACM Multimedia Conference 2005.

Given the rapid advances of computer graphics (CG) capabilities, it became challenging for human eyes to distinguish natural photos from photo-realistic CG images. This paper was among the first to investigate fundamental image characteristics (physical, geometric, and statistical) that can help discriminate photos from CG images. Besides demonstrating large performance gains, we also released the first benchmark dataset for research in this area.

14. **Best Student Paper Award,** IEEE International Conference on Image Processing, 2004.
Lexing Xie, L. Kennedy, S.-F. Chang, A. Divakaran, H. Sun and C.-Y. Lin at, "Discovering Meaningful Multimedia Patterns with Audio-Visual Concepts and Associated Text," IEEE International Conference on Image Processing (ICIP) Oct. 2004, Singapore.

15. **Outstanding Young Author Paper Award**, IEEE Circuits and Systems Society 2003.
Ching-Yung Lin and Shih-Fu Chang, "A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation," IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 2, pp. 153-168, February 2001,

16. **Best Paper Award (2nd prize)**, IEEE Sarnoff Symposium, March 2002.
Raj Kumar, Mihaela van der Schaar, and Shih-Fu Chang, "FGS+: Optimizing the Joint SNR-Temporal Video Quality in MPEG-4 Fine Grained Scalable Coding," IEEE Sarnoff Symposium, March 2002.

17. **Best Student Paper Award**, ACM Multimedia Conference 2002.
Hari Sundaram, Lexing Xie and Shih-Fu Chang, "A Utility Framework for the Automatic Generation of Audio-Visual Skims", ACM Multimedia Conference, France, Oct. 2002.

We presented a unified theoretical framework based on film theories and user perception model for generating optimal summaries of video programs.

18. **Best Paper Award**, IEEE Transactions on Circuits and Systems for Video Technology, 2000.
S.-F. Chang, W. Chen, H.J. Meng, H. Sundaram, and D. Zhong, "A Fully Automatic Content-Based Video Search Engine Supporting Multi-Object Spatio-Temporal Queries," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, No. 5, pp. 602-615, Sept. 1998.

This paper described the first video search engine that supports multi-modality multi-feature searching of automatically extracted video objects. One of the important features is searching video objects based on automatically extracted motion trajectories over time. Such capabilities are useful for finding information in security or sports videos.

19. **Best Paper Award**, SPIE International Symposium on Visual Communications and Image Processing, May 1995.

Shih-Fu Chang and John R. Smith, "Extracting Multi-Dimensional Signal Features for Content-Based Visual Query," SPIE Visual Communications and Image Processing, May 1995.

This paper presents content-based image query methods based on multiple local spatio-visual features.

20. **Best Student Paper Award**, the 1st ACM Multimedia Conference, Anaheim, CA, Aug. 1993.
S.-F. Chang and D. G. Messerschmitt, "Transform Coding of Arbitrarily-Shaped Image Segments," ACM 1st Multimedia Conference, Anaheim, CA, August 1993.

This paper explored the key issues involved in object-based visual representation and presented an efficient solution for coding arbitrarily shaped image regions.

Patents Issued (32):

P1. Chang, Shih-Fu, David Messerschmitt, and Wen-Lung Chen. "Method and apparatus for compositing compressed video data." U.S. Patent 5,408,274, issued April 18, 1995.

P2. Chang, Shih-Fu, Justin Che-I. Chuang, Gustavo De Los Reyes, and Amy Ruth Reibman. "Error resilient transcoding for video over wireless channels." U.S. Patent 6,339,450, issued January 15, 2002.

P3. Ebadollahi, Shahram, Shih-Fu Chang, and Henry Wu. "Method and apparatus for processing echocardiogram video images." U.S. Patent 6,514,207, issued February 4, 2003.

P4. Chang, Shih-Fu, and Ching-Yung Lin. "Method and apparatus for image authentication." U.S. Patent 6,532,541, issued March 11, 2003.

P5. Chang, Shih-Fu, and Horace J. Meng. "Methods and architecture for indexing and editing compressed video over the world wide web." U.S. Patent 6,735,253, issued May 11, 2004.

P6. Chang, Shih-Fu, William Chen, Horace J. Meng, Hari Sundaram, and Di Zhong. "Algorithms and system for object-oriented content-based video search." U.S. Patent 6,741,655, issued May 25, 2004.

P7. Divakaran, Ajay, Anthony Vetro, Huifang Sun, Peng Xu, and Shih-Fu Chang. "Extraction of high-level features from low-level features of multimedia content." U.S. Patent 6,763,069, issued July 13, 2004.

P8. Xu, Peng, Shih-Fu Chang, and Ajay Divakaran. "Method and system for high-level structure analysis and event detection in domain specific videos." U.S. Patent 6,813,313, issued November 2, 2004.

P9. Chang, Shin-Fu, Di Zhong, Raj Kumar, and Alejandro Jaimes. "Method and system for indexing and content-based adaptive streaming of digital video content." U.S. Patent 20,040,125,877, issued July 1, 2004.

P10. Sun, Qibin, Shih-fu Chang, Di Zhong, and Desai Nayasimhalu. "Methods and systems for generating multimedia signature." U.S. Patent 20,040,128,511, issued July 1, 2004.

P11. Benitez, Ana B., Alejandro Jaimes, Shih-Fu Chang, John R. Smith, and Chung-Sheng Li. "Fundamental entity-relationship models for the generic audio visual data signal description." U.S. Patent 6,847,980, issued January 25, 2005.

P12. Xie, Lexing, Shih-Fu Chang, Ajay Divakaran, and Huifang Sun. "Structural analysis of videos with hidden markov models and dynamic programming." U.S. Patent 6,865,226, issued March 8, 2005.

P13. Lin, Ching-Yung, and Shih-Fu Chang. "Method and apparatus for watermarking images." U.S. Patent 6,879,703, issued April 12, 2005.

P14. Benitez, Ana B., Alejandro Jaimes, Paek Seungyup, Shih-Fu Chang, Chung-Sheng Li, and John R. Smith. "Multimedia archive description scheme." U.S. Patent 6,941,325, issued September 6, 2005.

P15. Maeno, Kurato, Qibin Sun, Shih-Fu Chang, and Masayuki Suto. "Method and system for watermarking an electrically depicted image." U.S. Patent 20,050,129,268, issued June 16, 2005.

P16. Paek, Seungyup, Ana Benitez, Shih-Fu Chang, Atul Puri, Chung-Sheng Li, John R. Smith, and Lawrence Bergman. "Video description system and method." U.S. Patent 7,143,434, issued November 28, 2006.

P17. Chang, Shih-Fu, and Jianhao Meng. "Watermarking of digital image data." U.S. Patent 7,154,560, issued December 26, 2006.

P18. Benitez, Ana Belen, Shih-Fu Chang, Qian Huang, Seungyup Paek, and Atul Puri. "Multimedia integration description scheme, method and system for MPEG-7." U.S. Patent 7,185,049, issued February 27, 2007.

P19. Paek, Seungup, Ana Benitez, Shih-Fu Chang, Chung-Sheng Li, John R. Smith, Lawrence D. Bergman, Atul Puri, and Qian Huang. "Image description system and method." U.S. Patent 7,254,285, issued August 7, 2007.

P20. Xie, Lexing, Ajay Divakaran, and Shih-Fu Chang. "Unsupervised learning of video structures in videos using hierarchical statistical models to detect events." U.S. Patent 7,313,269, issued December 25, 2007.

P21. Chang, Shih-Fu, and Dongqing Zhang. "System and method for extracting text captions from video and generating video summaries." U.S. Patent 7,339,992, issued March 4, 2008.

P22. Wang, Yong, and Shih-Fu Chang. "System and method for motion estimation and mode decision for low-complexity h. 264 decoder." U.S. Patent 20,080,181,308, issued July 31, 2008.

P23. Paek, Seungup, Ana Benitez, and Shih-Fu Chang. "Systems and methods for interoperable multimedia content descriptions." U.S. Patent 7,653,635, issued January 26, 2010.

P24. Chang, Shih-Fu, Wei Jiang, and Alexander C. Loui. "Active context-based concept fusion." U.S. Patent 7,720,851, issued May 18, 2010.

P25. Chang, Shih-fu, Jun Wang, and Tony Jebara. "System And Method For Annotating And Searching Media." U.S. Patent 20,110,314,367, issued December 22, 2011.

P26. Jiang, Wei, Courtenay Cotton, Shih-Fu Chang, Daniel P. Ellis, and Alexander C. Loui. "Video concept classification using audio-visual atoms." U.S. Patent 8,135,221, issued March 13, 2012.

P27. Kim, Jae-Gon, Yong Wang, Shih-Fu Chang, Kyeongok Kang, and Jinwoong Kim. "Method and system for optimal video transcoding based on utility function descriptors." U.S. Patent 8,218,617, issued July 10, 2012.

P28. Chang, Shih-Fu, and Eric Zavesky. "System and method for dynamically and interactively searching media data." U.S. Patent 8,364,673, issued January 29, 2013.

P29. Chang, Shih-Fu, Jun Wang, Paul Sajda, Eric Pohlmeyer, Barbara Hanna, and David Jangraw. "Rapid image annotation via brain state decoding and visual pattern mining." U.S. Patent 8,671,069, issued March 11, 2014.

P30. Kennedy, Lyndon, and Shih-Fu Chang. "Systems and methods for image archaeology." U.S. Patent 8,849,058, issued September 30, 2014.

P31. He, Junfeng, Shih-Fu Chang, and Tai-Hsu Lin. "Systems and methods for mobile search using Bag of Hash Bits and boundary reranking." U.S. Patent 9,009,149, issued April 14, 2015.

P32. Chang, Shih-fu, Yan-ying Chen, and Tao Chen. "Systems and Methods for Visual Sentiment Analysis." U.S. Patent 20,170,046,601, issued February 16, 2017.

Contributions to International Standards and Evaluation Forums:

Through close collaboration with industrial partners (IBM, AT&T, Mitsubishi, SONY, Kodak, Philips, ETRI), we have made significant contributions to the development of international standards: 21 MPEG-7 contributions (1998-2001), one MPEG-4 contribution (2002), and several MPEG-21 contributions (2002). Several of them have been accepted and become parts of the standards. Two members of our team also served as editors for sections focusing on multimedia description schemes in MPEG-7 and MPEG-21.

We actively participated in the international TRECVID video retrieval evaluation forum and demonstrated the best performance in high-level feature detection in 2008 (among 160+ submissions), and multimedia event detection in 2010.

In the 2011 DARPA Grand Challenge, *Shredder*, we applied automatic image matching techniques and interactive search tools to reassemble thousands of shredded pieces back to original documents, which allow extraction of information needed for solving puzzles. Our system was ranked the 5th place among near 9000 team submissions from around the world. (see http://www.ee.columbia.edu/ln/dvmm/shredder/).

Professional Activities and Services:

Professional Leadership and Organization (recent):

- Chair, ACM Special Interest Group on Multimedia, 2013-2017.
  Significantly grew the local community activities, expanded talent recognition/development programs, and strengthened the financial resources of the society.
- Co-Organizer, NSF Workshop on Multimedia Challenges and Research Directions for the Next 10 Years, March 2017.
- Area Chair, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, 2018.
- General Co-Chair, ACM Multimedia Conference, Florence, Italy, 2010. (the flagship conference of ACM SIG Multimedia).
- Co-Chair, NSF Workshop on Hybrid Neuro-Computer Vision Systems, April 2010, (http://www.columbia.edu/cu/hybridvision/)
- Chair, ARO/DARPA Workshop on Interactive Query Refinement, 2011.

- General Co-Chair, ACM Multimedia Conference, November, 2000.
- General Co-Chair, IEEE International Multimedia Conference and Exhibition, 2004. (joint conference sponsored by four IEEE Societies)
- Elected member, Information Forensics and Security (TIFS) TC, IEEE Society of Signal Processing, 2010-2013.
- Elected member, Image and Multidimensional Digital Signal Processing (IMDSP) TC, IEEE Society of Signal Processing, 2003-2008.

Editorial Board (recent):

- Associate Editor, ACM Book Series, 2013-date
- Associate Editor, International Journal of Multimedia Information Retrieval, 2011-
- Associate Editor, Editorial Board for Overview Papers for IEEE Signal Processing Society Transactions, 2009-12.
- Editor in Chief, Signal Processing Magazine, IEEE Signal Processing Society, 2006-8.
  The flagship publication of IEEE Signal Processing Society distributed to about 16,000 members. Ranked by ISI in 2004 and 2009 as the top publication with the highest citation impact among all 200+ IEEE journals in the electric and electronic area.
- Associate Editor, Journal of Statistical Analysis and Data Mining, John Wiley & Sons, Inc., 2006-2009
- Associate Editor, IEEE Transactions on Multimedia (T-MM), 2000-2003.
- Associate Editor, EURASIP Journal on Applied Signal Processing, 2002-2003.
- Associate Editor, ACM Transactions on Multimedia Computing, Communications, and Applications, (TOMCAPP) 2004-2005.

Review Panels and Others:

- International Scientific Advisor Board, DATAIA Institute of Data Science, France, 2018.

- Review Panel, NIH/NLM Special Emphasis Panel, Feb. 2014.

- International Review Panel, the Swiss National Science Foundation (SNSF), "National Centres of Competence in Research" (NCCR), for the IM2 National Centre on multi-modal information interaction research, 2002-2013.

- Review Panel, National Centre for Science and Engineering Technology (CSET), CLARITY, Science Foundation Ireland, 2010-11.

- Review Panel, Academia Sinica, Division of Mathematics and Physical Sciences, 2009.

- International Advisory Board, European Union 7th Framework Project, Network of Excellence, PetaMedia, 2008-2011.

- Advisory Board, European Union Integrated Project Research Consortium (WeKnowIt), 2008-2011.

- Review Committee, University of Illinois at Urbana-Champaign, Beckman Institute, HCII, 2010.

- Area Study Expert (for the area of Multimedia Information Retrieval), Computer and Information Science and Engineering Directorate (CISE), National Science Foundation, ten-year perspective review, 2005.
- Program Review Panelist, National Science Foundation, 1996, 2000, 2001, 2003, 2005, 2008, 2009, 2010.

Teaching/Education:

- PI of the multi-year ASCENT Postdoc Best Practice program (2014-2018) sponsored by Computing Research Association (CRA) including collaboration among 4 universities in the New York region. It aims to develop best practices and offer programs for mentoring and training postdocs in the computer science field.

- Co-led the NSF Integrative Graduate Education and Research Traineeship (IGERT) Program (2012-2018) focusing on the theme "From Data to Solutions". The IGERT program is intended to establish new models for graduate education and training in a fertile environment for collaborative research that transcends traditional disciplinary boundaries.

- Supervised 40 Ph.D. students, 13 postdocs, and more than 45 MS students, many of whom became active leaders in academia or industry, or entrepreneurs. (see list below)

- Lead Columbia-Amazon summer undergraduate research program SURE [link] including 70+ students from HBCU, MSIs, and underrepresented groups since 2021.

Ph.D. Students and Postdocs: (as advisor and sponsor)

Ph.D. Students Completed
1. John R. Smith (IBM, IBM Fellow), 1997.
2. Gustavo Reyes (AT&T, Executive Director of Technology), 1999.
3. Paul Bocheck (BarterQuest, Founder and CEO), 1999.
4. Seungyup Paek (Jubilant Technology, Founder and CEO), 2000.
5. Ching-Yung Lin (Graphen, Founder and CEO), 2000.
6. Di Zhong (Citibank), 2001.
7. Hari Sundaram (University of Illinois Urbana-Champaign, Professor), 2002.
8. Alejandro Jaimes (Dataminer, Chief Scientist & SVP), 2003.
9. Ana Benitez (Google), 2004.
10. Shahram Ebadollahi (Norvatis, Global Head of Data Science and AI), 2005.
11. Yong Wang (Mugeda, Co-Founder and CTO), 2005.
12. Lexing Xie (Australian National University, Professor), 2005.
13. Dong-Qing Zhang (Huawei Research), 2005.
14. Winston Hsu (National Taiwan University, Professor), 2006.

15. Tian-Tsong Ng (Institute for Information Research, Singapore, Project Lead), 2007.
16. Lyndon Kennedy (Apple), 2008.
17. Jessie Hsu (Industrial Technology Research Institute), 2009.
18. Eric Zavesky (AT&T, Director Inventive Science), 2010.
19. Wei Jiang (Huawei Research), 2010.
20. Jun Wang (Alibaba), 2011.
21. Wei Liu (Tencent), 2012.
22. Mandis Beigi (IBM T.J. Watson Research), 2013.
23. Junfeng He (Facebook), 2013.
24. Yan Wang (Microsoft), 2015.
25. Felix Xinnan Yu (Google Research), 2015.
26. Guangnan Ye (IBM T.J. Watson Research), 2015.
27. Xiao-Ming Jennifer Wu (Hong Kong Polytechnic University, Associate Professor), 2016.
28. Hongzhi Li (Microsoft Research), 2016.
29. Brendan Jou (Google), 2016.
30. Jie Feng (Amazon), 2017.
31. Zheng Shou (National University of Singapore, Assistant Professor), 2019.
32. Alireza Zareian (Snap Research), 2021.
33. Brian Chen (Meta), 2023.

Postdoc Completed
1. Qibin Sun (Cisco), 2000-2001.
2. Dong Xu (Sydney University), 2007-2008.
3. Yu-Gang Jiang (Fudan University), 2009-2012.
4. Rong-Rong Ji (Xiaman University), 2010-2013.
5. Zhengou Li (Noah's Ark Research Lab), 2009-2013.
6. Yadong Mu (Peking University), 2011-2013.
7. Subhabrata Bhattacharya (NetFlix Research), 2013-2014.
8. Dong Liu (RetailNext), 2011-2015.
9. Tao Chen (500px), 2013-15.
10. Hanwang Zhang (Nanyang Technology University, Singapore), 2017.
11. Svebor Karaman (DataMiner), 2015-2020.
12. Xu Zhang (Amazon), 2016-2019.
13. Bo Wu (IBM Research), 2018-2020.
14. Chris Thomas (Virginia Tech), 2020-2022.