Winston Hsu (徐宏民)

Professor, Dept. of Computer Science and Information Eng., National Taiwan University, Taipei
whsu@ntu.edu.tw | http://winstonhsu.info

Summary of Experiences

Advancing machine intelligence and deep learning over large-scale multimodal data streams including image, video, text, 3D point clouds, etc., for intelligent perception and decision. Receiving technical awards in multimedia and computer vision research communities. Realizing advanced research towards business deliverables via collaborating with leading industry partners and co-founding startups.

Education

- **Ph.D. in Electrical Engineering** — Columbia University, New York, USA (01/2007)
  Advisor: Prof. Shih-Fu Chang, Department of Electrical Engineering
- **M.S. in Computer Science** — National Taiwan University, Taipei, Taiwan (07/1995)
  Advisor: Prof. Jau Huang, Department of Computer Science and Information Engineering

Professional Experience

- **National Taiwan University**, Taipei, Taiwan
  Professor (2015 — present), *Associate Professor* (2011 — 2015); *Assistant Professo*r (2007 — 2011)
  Co-leads Communication and Multimedia Lab (CMLab), with 130 members and 7 faculties
- **thingnario Inc. (**慧景科技**),** Taipei, Taiwan
  Co-founder; a fast-growing AI/IoT startup for optimizing green energy and power infrastructure, with subscription-based business model, and serving global energy industry (4/2017 — present)
- **NVIDIA AI Lab,** National Taiwan University, Taipei, Taiwan
  Founding Director; awards with 4-year gift money and supercomputer DGX-1 from NVIDIA CEO Jensen Huang; the first in Asia and the 5th in the world; enabling neural-networks-based multimedia comprehension (10/2016 — 12/2020)
- **IBM TJ Watson Research Center**, Yorktown Heights, New York, USA
  Visiting Scientist - Cognitive Computing (07/2016 — 07/2017)
  Computer Vision Department, Watson AI Division; contributing the first AI trailer
- **Microsoft Research**, Redmond, Washington, USA
  Visiting Scientist (06/2014 — 09/2014)
  leveraging web-scale images for training deep convolutional neural networks for image recognition
- **CyberLink Corp. (**訊連科技**)**, Taipei, Taiwan
  Founding Engineer and R&D Manager (06/1997 — 06/2001)
  The 8th employee for the startup, now a public multimedia software company (5203.TW). Leading 15 engineers, growing the team, and reporting to CEO (Alice Chang) directly.

Awards and Honors (Selected)

- FIRST PLACE in IARPA Disguised Faces in the Wild Competition 2018 (along with CVPR 2018)
- IBM Research Pat Goldberg Memorial Best Paper Award 2018 [one of the 6 most influential projects among global IBM research and engineering teams in 2017]
- 2016 IBM Research Image Award. [The technology, industry, and media impacts for the first AI trailer, Morgan. Jan. 2017]
- Microsoft Research Award 2009/2012/2014-2018 in machine intelligence for image/video
- Best Brave New Idea Paper Award in ACM Multimedia 2017. [Top multimedia conference]
- ACM Multimedia 2014 Grand Challenge Multimodal Award
- FIRST PLACE in MSR-Bing Image Retrieval Challenge 2013 [Hosted by Microsoft Research Redmond and Bing. Award presented by Microsoft CTO Dr. Harry Shum in Seattle, WA. Oct. 2013]
- 2019 年科技部「AI 創新研究中心投資潛力獎第一名」
- 2013 National Outstanding IT Elite Award (傑出資訊人才獎) [National awards for recognizing advanced research and industry collaborations. Dec. 2013]
- 2012 NTU EECS Research Contribution Award [6 winners only, top 3% among 180 faculties in EECS]
- 2011 National Science Council (NSC) Ta-You Wu Memorial Award (吳大猷先生紀念獎) [A national and prestigious recognition for young researchers. August 2011]

Leadership

- Co-leading Communication and Multimedia Lab (CMLab), NTU, with more than 120 students (MS & PhD), 3 postdocs, 7 research assistants, and 7 faculties
- Technical Committee for Science Park, Ministry of Science and Technology (MOST), Taiwan
- Director/Board Member for Columbia Alumni Association of Taiwan
- Columnist in AI & Machine Learning for DIGITIMES (電子時報)
- Opinion and technical leader in machine intelligence for Taiwan-based IT industries. Strong industry-academia collaborations in machine intelligence. Rich media interviews regarding the technical and industrial trends.

Delivered Systems

- *Photon* (by thingnario Inc.), the fast-growing PV monitoring and optimization SaaS (AI-enabled prediction), and serving more than 1700 global PV sites, ranging from 100KW to 256MW.
- Technical consultancy (advising or founding AI/ML teams) for Taipei-based leading ICT companies for delivering AI-enabled products. The companies include *CyberLink*/*Perfect Corp.* (image/face recognition), *Synology* (photo management), *QNAP* (system, retail AI), *iCatch* (low-power IC, edge AI), *Genesys Logic* (neural ISP, edge AI), *HTC VIVE* (AR/VR, 3D vision), etc.
- 3D vision for autonomous driving and ADAS, collaborating with FIH (a Foxconn company), 4 patents filed and 1 CVPR publication; multimodal perception for LiDAR, RADAR, and stereo-cameras.

- *GDN*, the best 6-DoF grasp detection networks, operating on novel objects and low-cost 3D cameras; awarded MOST FutureTech Award 2020 (未來科技獎) and published in CoRL 2020
- The first AI trailer by IBM Watson. Contributing CNN-based object and place recognition and mining effective visual aspects for automatic trailer generation (Sep. 2016)
- Top-3 performer for brain tumor segmentation in BRATS benchmark and published in CVPR (2017)
- Contributing the image to text neural networks for novel "Image to Poetry" (看圖造詩) technical core in Microsoft chatbot "xiaoice"（小冰) (2017)
- 3-year industry-academia collaborative project for efficient convolutional neural networks model (compression) for augmenting MediaTek Deep Learning Platform (2014-2017)

Professional Services

- Associate Editor for IEEE Transaction on Multimedia (TMM)
- Associate Editor for IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)
- Editorial Board for IEEE Multimedia Magazine (2010-2017)

Scholarly Addresses

Keynote Addresses and Invited Presentations

- Keynote speaker for IEEE International Conference on Identity, Security and Behavior Analysis (ISBA) 2019, Hyderabad, India
- Invited speaker for Microsoft Academic Day, Beijing, China, November 2019
- Invited speaker for Computing in the 21st Century Conference & Asia Faculty Summit on MSRA's 20th Anniversary, Beijing, China, November 2018
- Keynote speaker for AI Session, Computex Forum, Taipei, June 2018
- Keynote speaker for Dell 2017 Tech Forum, Taipei, August 2017
- Keynote speaker for the 1st International Workshop on Understanding Situations Through Multimodal Sensing (uSitu'16), Singapore, January 2016

Top Conference Tutorials

- *Learning from 3D (Point Cloud) Data* in ACM Multimedia 2019, Nice, France. Highly attended and appreciated by the attendees and 140+ page slides shared (http://bit.ly/2MG2PGf).
- *Face Recognition - from Scientific Research to Commercial Products* in NVIDIA GPU Technology Conference (GTC) 2019, San Jose, USA. (with strong ratings including Speaker Rating: 4.47/5 & Content Rating: 4.58/5)
- *Investigating Data Augmentation Strategies for Advancing Deep Learning Training* in NVIDIA GTC 2018 Taipei, full lecture room with 500+ attendees
- I*nvestigating Data Augmentation Strategies for Advancing Deep Learning Training*, in NVIDIA GPU Technology Conference (GTC) 2018, San Jose, USA. (full lecture room, more than 180 attendees)

- *Learning Large-Scale Multimodal Data Streams: Ranking, Mining, and Machine Comprehension*, in NVIDIA GTC 2017 (strong ratings including Speaker Rating: 4.5/5 & Content Rating: 4.36/5)

Publications

[Google citations: 6730+, H-index: 37, i10-index: 86, more than 180 publications, as of March 2021]

Top Journal (Selected)

1. Chih-Hung Liang, Yu-An Chen, Yueh-Cheng Liu, and <u>Winston H. Hsu</u>. Raw Image Deblurring. In *IEEE Transactions on Multimedia*, 2021.
2. Hung-Ting Su, Chen-Hsi Chang, Po-Wei Shen, Yu-Siang Wang, Ya-Liang Chang, Yu-Cheng Chang, Pu-Jen Cheng, <u>Winston H. Hsu</u>. End-to-End Video Question-Answer Generation with Generator-Pretester Network. In *IEEE Transactions on Circuits and Systems for Video Technology*, 2021
3. Sebastian Agethen, Winston H. Hsu. Deep Multi-Kernel Convolutional LSTM Networks and an Attention-Based Mechanism for Videos. In *IEEE Transactions on Multimedia*, March 2020.
4. Wen-Yu Lee, Winston H. Hsu, Shin'ichi Satoh. Learning from Cross-Domain Media Streams for Event-of-Interest Discovery. In *IEEE Transactions on Multimedia*, January 2018.
5. Wei-Tse Sun, Ting-Hsuan Chao, Yin-Hsi Kuo, <u>Winston H. Hsu</u>. Photo Filter Recommendation by Category-Aware Aesthetic Learning. In *IEEE Transactions on Multimedia*, 2017.
6. Wen-Yu Lee, <u>Winston H. Hsu</u>. Learning from Cross-Domain Media Streams for Event-of-Interest Discovery. In *IEEE Transactions on Multimedia*, 2017. [Accepted]
7. Yin-His Kuo, <u>Winston H. Hsu</u>. De-Hashing: Server-Side Context-Aware Feature Reconstruction for Mobile Visual Search. In *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, Volume 27, No 1, Pages 139-148, 2017.
8. Bor-Chun Chen, Yan-Ying Chen, Yin-Hsi Kuo, Thanh Duc Ngo, Duy-Dinh Le, Shin'ichi Satoh, <u>Winston H. Hsu</u>. Scalable Face Track Retrieval in Video Archives using Bag-of-Faces Sparse Representation. In *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, Volume 27, No 7, Pages 1595 – 1603, July 2017.
9. Bor-Chun Chen, Chu-Song Chen, <u>Winston H. Hsu</u>. Face Recognition and Retrieval Using Cross-Age Reference Coding with Cross-Age Celebrity Dataset. In *IEEE Transactions on Multimedia*, Volume 17, No 6, Pages 804-815, 2015.
10. Yu-Chuan Su, Tzu-Hsuan Chiu, Yin-Hsi Kuo, Chun-Yen Yeh, <u>Winston H. Hsu</u>. Scalable Mobile Visual Classification by Kernel Preserving Projection Over High-Dimensional Features. In *IEEE Transactions on Multimedia*, Volume 16, No 6, Pages 1645-1653, October 2014.
11. Yan-Ying Chen, <u>Winston H. Hsu</u>, Hong-Yuan Mark Liao, Automatic Training Image Acquisition and Effective Feature Selection from Community-Contributed Photos for Facial Attribute Detection. In *IEEE Transactions on Multimedia*, Pages 1388-1399, October 2013.
12. Yan-Ying Chen, An-Jung Cheng, <u>Winston H. Hsu</u>. Travel Recommendation by Mining People Attributes and Travel Group Types from Community-Contributed Photos. In *IEEE Transactions on Multimedia*, Volume 15, No 6, Pages 1283-1295, October 2013.

13. Bor-Chun Chen, Yan-Ying Chen, Yin-Hsi Kuo, <u>Winston H. Hsu</u>. Scalable Face Image Retrieval Using Attribute-Enhanced Sparse Codewords. In *IEEE Transactions on Multimedia*, Volume 15, No 5, Pages 1163-1173, August 2013.

14. Yin-Hsi Kuo, Wen-Huang Cheng, Hsuan-Tien Lin, <u>Winston H. Hsu</u>, Unsupervised Semantic Feature Discovery for Image Object Retrieval and Tag Refinement. In *IEEE Transactions on Multimedia*, Volume 14, No 4, Pages 1079-1090, August 2012.

15. Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, <u>Winston H. Hsu</u>, Shao-Yi Chien. Preference-Aware View Recommendation System for Scenic Photos Based on Bag-of Aesthetics-Preserving Features. In *IEEE Transactions on Multimedia*, Volume 14, No 3, Page 833-843, June 2012.

16. Yen-Liang Lin, Ming-Kuang Tsai, <u>Winston H. Hsu</u>, Chih-Wei Chen. Investigating 3-D Model and Part Information for Improving Content-based Vehicle Retrieval. In *IEEE Transactions on Circuits and Systems for Video Technology*, Volume 23, No 3, Pages 401-413, March 2013.

17. Guan-Long Wu, Yin-Hsi Kuo, Tzu-Hsuan Chiu, <u>Winston H. Hsu</u>, and Lexing Xie. Scalable Mobile Video Retrieval with Sparse Projection Learning and Pseudo Label Mining. In *IEEE Multimedia Magazine*, Volume 20, No 3, Pages 47-57, July-Sept. 2013.

18. Tao Mei, <u>Winston H. Hsu</u>, Jiebo Luo. Knowledge Discovery from Community-Contributed Multimedia. In *IEEE Multimedia Magazine*, Pages 16-17, October-December 2010.

19. Yi-Hsuan Yang, <u>Winston H. Hsu</u>, Homer H. Chen. Online Reranking via Ordinal Informative Concepts for Context Fusion in Concept Detection and Video Search. In *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, Pages 1880-1890, December 2009.

20. <u>Winston H. Hsu</u>, Lyndon Kennedy, Shih-Fu Chang. Reranking Methods for Visual Search. In *IEEE Multimedia Magazine*, Volume 14, Issue 3, Pages 14-22, July-September 2007.

21. Milind Naphade, John R. Smith, Jelena Tesic, Shih-Fu Chang, <u>Winston Hsu</u>, Lyndon Kennedy, Alexander Hauptmann, and Jon Curtis. Large-Scale Concept Ontology for Multimedia. In *IEEE Multimedia Magazine*, Vol. 13, No.3, July-September 2006.

Top Conferences (Selected)

1. Yu-Kai Huang, Yueh-Cheng Liu, Tsung-Han Wu, Hung-Ting Su, Yu-Cheng Chang, Tsung-Lin Tsou, Yu-An Wang, <u>Winston H. Hsu</u>. S3: Learnable Sparse Signal Superdensity for Guided Depth Estimation. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2021

2. Ke-Jyun Wang, Yun-Hsuan Liu, Hung-Ting Su, Jen-Wei Wang, Yu-Siang Wang, <u>Winston H. Hsu</u>, Wen-chin Chen. OCID-Ref: A 3D Robotic Dataset with Embodied Language for Clutter Scene Grounding. *NAACL-HLT* 2021

3. Chen-Hsi Chang, Hung-Ting Su, Juiheng Hsu, Yu-Siang Wang, Yu-Cheng Chang, Zhe Yu Liu, Ya-Liang Chang, Wen-Feng Cheng, Ke-Jyun Wang, <u>Winston H. Hsu</u>. Situation and Behavior Understanding by Trope Detection on Films. *WWW '21: The Web Conference* 2021.

4. Kuang-Yu Jeng, Yueh-Cheng Liu, Zhe Yu Liu, Jen-Wei Wang, Ya-Liang Chang, Hung-Ting Su, <u>Winston H. Hsu</u>. GDN: A Coarse-To-Fine (C2F) Representation for End-To-End 6-DoF Grasp Detection. In the Conference on Robot Learning (CoRL). November 2020.

5. Jhih-Yuan Lin, Yu-Cheng Chang, <u>Winston H. Hsu</u>. Efficient and Phase-Aware Video Super-Resolution for Cardiac MRI. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2020.

6. Ya-Liang Chang, Zhe Yu Liu, Kuan-Ying Lee, <u>Winston H. Hsu</u>. Free-form Video Inpainting with 3D Gated Convolution and Temporal PatchGAN. In *International Conference on Computer Vision (ICCV)*, November 2019.

7. Kaipeng Zhang, Zhanpeng Zhang, Chia-Wen Cheng, <u>Winston H. Hsu</u>, Yu Qiao, Wei Liu, Zhifeng Li. Super-Identity Convolutional Neural Network for Face Hallucination. In *European Conference on Computer Vision (ECCV)*, September 2018.

8. Wen Hua Lin, Kuan-Ting Chen, HungYueh Chiang, <u>Winston H. Hsu</u>. Netizen-Style Commenting on Fashion Photos: Dataset and Diversity Measures. In *WWW*, April 2018.

9. John R. Smith, Dhiraj Joshi, Benoit Huet, <u>Winston H. Hsu</u>, and Zef Cota. Harnessing A.I. for Augmenting Creativity: Application to Movie Trailer Creation. In *ACM Multimedia*, October 2017. [Best Brand New Idea Paper Award]

10. Meng-Ru Hsieh, Yen-Liang Lin, <u>Winston H. Hsu</u>. Drone-based Car Counting by Spatially Regularized Regional Proposal Network. In *International Conference on Computer Vision (ICCV)* 2017.

11. Kuan-Lun Tseng, Yen-Liang Lin, <u>Winston H. Hsu</u>, Chung-Yang Huang. Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2017.

12. Ting-Hsuan Chao, Yen-liang Lin, Yin-Hsi Kuo, <u>Winston H. Hsu</u>. Scalable Object Detection by Filter Compression with Regularized Sparse Coding. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, June 2015.

13. Yan-Ying Chen, Yin-Hsi Kuo, Chun-Che Wu, and <u>Winston H. Hsu</u>. Visually Interpreting Names as Demographic Attributes by Exploiting Click-Through Data. In *AAAI Conference on Artificial Intelligence*, January 2015.

14. Kuan-Ting Chen, Kezhen Chen, Peizhong Cong, <u>Winston H. Hsu</u>, Jiebo Luo. Who are the Devils Wearing Prada in New York City? In *ACM Multimedia* 2015.

15. Hsin-Fu Huang, Wei-Hung Weng, <u>Winston H. Hsu</u>, Chi-Kuang Sun. Automated Detection of Noninvasive Imaging of Basal Cell Carcinoma by Convolutional Neural Network. In *NIPS 2015 Workshop on Machine Learning in Healthcare*.

16. Yu-Chuan Su, Tzu-Hsuan Chiu, Yan-Ying Chen, Chun-Yen Yeh, <u>Winston H. Hsu</u>. Enabling Low Bitrate Mobile Visual Recognition – A Performance versus Bandwidth Evaluation. In *ACM Multimedia*, October 2013.

17. Chia-Hung Lin, Yan-Ying Chen, Bor-Chun Chen, Yu-Lin Hou, <u>Winston H. Hsu</u>. Facial Attribute Space Compression by Latent Human Topic Discovery. In *ACM Multimedia,* November 2014.

18. Yen-Liang Lin, Vlad I. Morariu, <u>Winston H. Hsu</u>, Larry S. Davis. Jointly Optimizing 3D Model Fitting and Fine-Grained Classification. In *European Conference on Computer Vision (ECCV)*, September 2014.

19. Che-Chun Lee, Yin-Hsi Kuo, <u>Winston H. Hsu</u>, Shin'ichi Satoh, Sebastian Agethen. Efficient Cross-Domain Image Retrieval by Multi-Level Matching and Spatial Verification for Structural Similarity. In *ACM Multimedia*, November 2014.

20. Yan-Ying Chen, <u>Winston H. Hsu</u>, Hong-Yuan Mark Liao. Discovering Informative Social Subgraphs and Predicting Pairwise Relationships from Group Photos. In *ACM Multimedia*, October 2012.

21. Yan-Ying Chen, <u>Winston H. Hsu</u>, Hong-Yuan Mark Liao. Discovering Informative Social Subgraphs and Predicting Pairwise Relationships from Group Photos. In *ACM Multimedia*, October 2012.

22. Yen-Liang Lin, Cheng-Yu Huang, Hao-Jeng Wang, <u>Winston H. Hsu</u>. 3D Sub-Query Expansion for Improving Sketch-based Multi-View Image Retrieval. In *International Conference on Computer Vision (ICCV)*, December 2013.

23. Bor-Chun Chen, Chu-Song Chen, <u>Winston H. Hsu</u>. Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval. In *European Conference on Computer Vision (ECCV)*, September 2014.

24. Chiang-Yu Tsai, Yin-Hsi Kuo and <u>Winston H. Hsu</u>. Approximating Weighted Hamming Distance by Probabilistic Selection for Multiple Hash Tables. In *European Conference on Information Retrieval (ECIR)*, March 2015.

25. Chun-Che Wu, Kuan-Yu Chu, Yin-Hsi Kuo, Yan-Ying Chen, Wen-Yu Lee, <u>Winston H. Hsu.</u> Search-Based Relevance Association with Auxiliary Contextual Cues. In *ACM Multimedia*, 2013.

26. Yu-Heng Lei, Yan-Ying Chen, Bor-Chun Chen, Lime Iida, <u>Winston H. Hsu</u>, Where Is Who: Large-Scale Photo Retrieval by Facial Attributes and Canvas Layout. In *ACM SIGIR*, August 2012.

27. Yin-Hsi Kuo, Hsuan-Tien Lin, Wen-Huang Cheng, Yi-Hsuan Yang, <u>Winston H. Hsu.</u> Unsupervised Auxiliary Visual Words Discovery for Large-Scale Image Object Retrieval. In *IEEE Computer Vision and Pattern Recognition* (CVPR), June 2011.