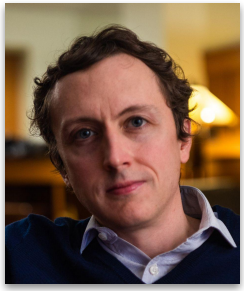
Welcome to CVPR 2024

IEEE COMPUTER SOCIETY

CVPR SEATTLE, WA JUNE 17-21, 2024

CVF

General Chairs



Walter Scheirer
Notre Dame

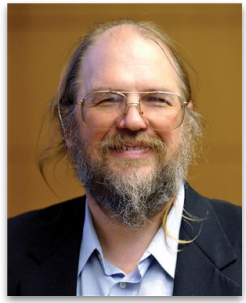
Octavia Camps
Northeastern

Ramin Zabih
Cornell / Google

Rita Cucchiara
U. Modena

Sudeep Sarkar
USF

Advisor to the PCs



David Forsyth
UIUC

Program Chairs

Zeynep Akata
U. Tübingen

Ali Farhadi
U. Washington / Allen Institute for AI

David Crandall
Indiana U.

Jianxin Wu
Nanjing U.

Robert Pless
George Washington U.

Imari Sato
National Institute of Informatics

Technical Chair

Yoshitomo Matsubara

Finance Chair

Gerard Medioni

Workshop Chairs

Luisa Verdoliva

Andrew Owens

Abhinav Shrivastava

Antitza Dantcheva

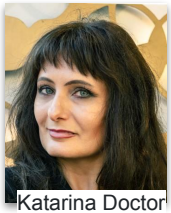
Demo Chairs

Shu Kong

Sathya Aakur

Tutorial Chairs

Vito Struc

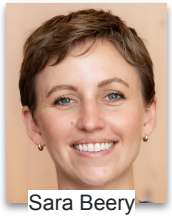


Katarina Doctor

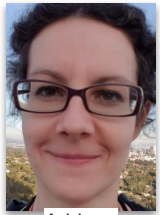
Diversity, Equity and Inclusion Chairs

Shuran Song

CJ Taylor



Sara Beery



Adriana Kovashka

Tamara Berg

Roni Sengupta

Mike King

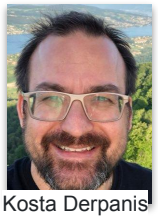
AI Art Curator

Luba Elliott

Accessibility Chair

Danna Gurari

Publicity Chairs



Kosta Derpanis

Abby Stylianou

Boqing Gong

Shenghua Gao

Jia-Bin Huang

Publications Chair

Eric Mortensen

Doctoral Consortium Chairs


Aparna Bharati

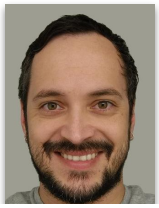

Nathan Jacobs

Corporate Relations Chairs


Brian Clipp


Victor Fragoso

Website Chair


Mauricio Pamplona Segundo

Local Chairs


Ranjay Krishna


Ira Kemelmacher-Shlizerman

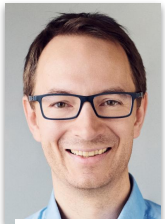
Social Activities Chairs


Vitor Albiero


Yale Song


Giovanni Maria Farinella

Virtual Platform Chair


Andreas Geiger

Workflow Chair


Sherry Xue

Conference Ombuds


Derek Hoiem


Angjoo Kanazawa

Web Developer


Lee Campbell

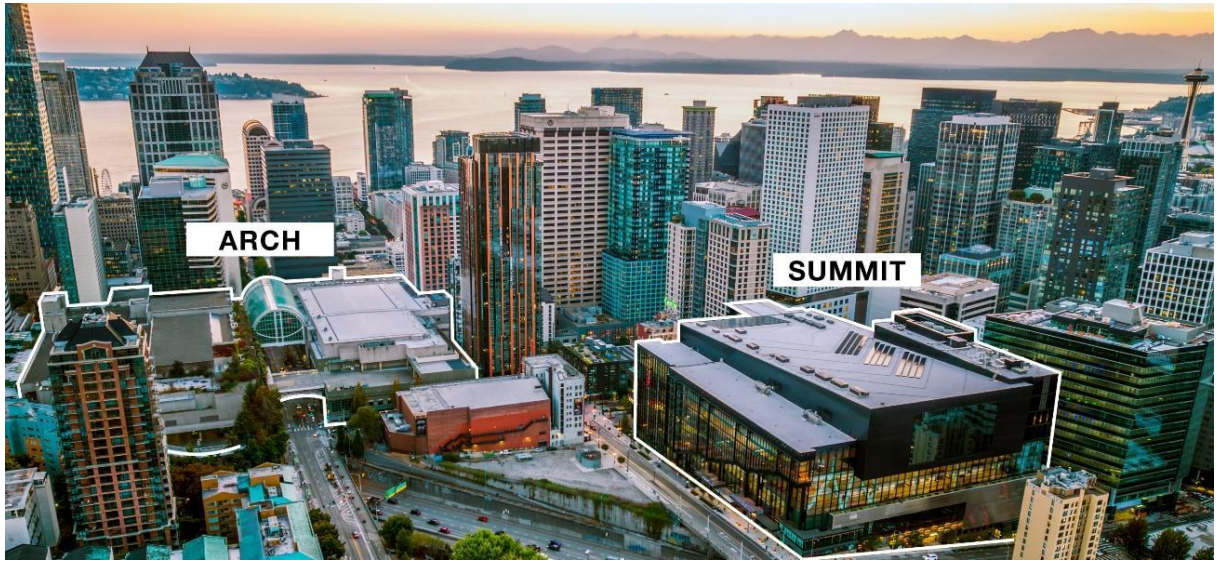
Conference Producer

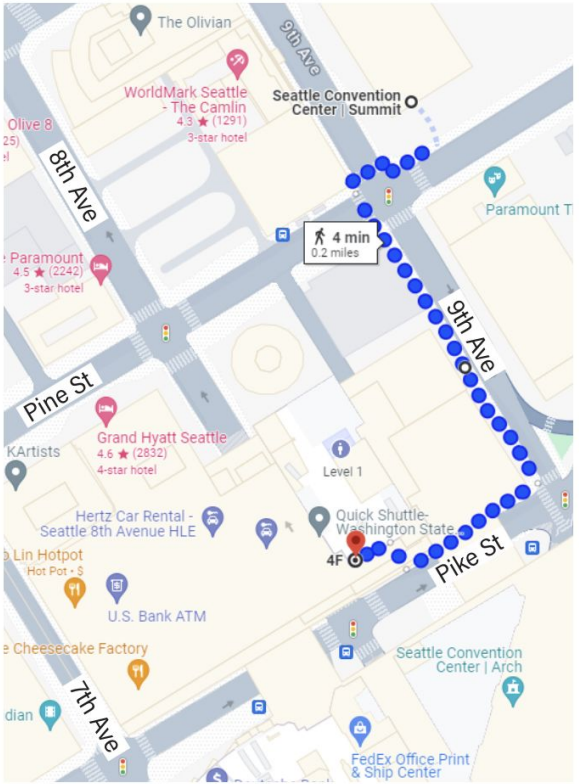

Nicole Finn

Event Program Manager


Carmen Saliba

Seattle Convention Center





See the maps in the back of your program guide and follow the stickers:

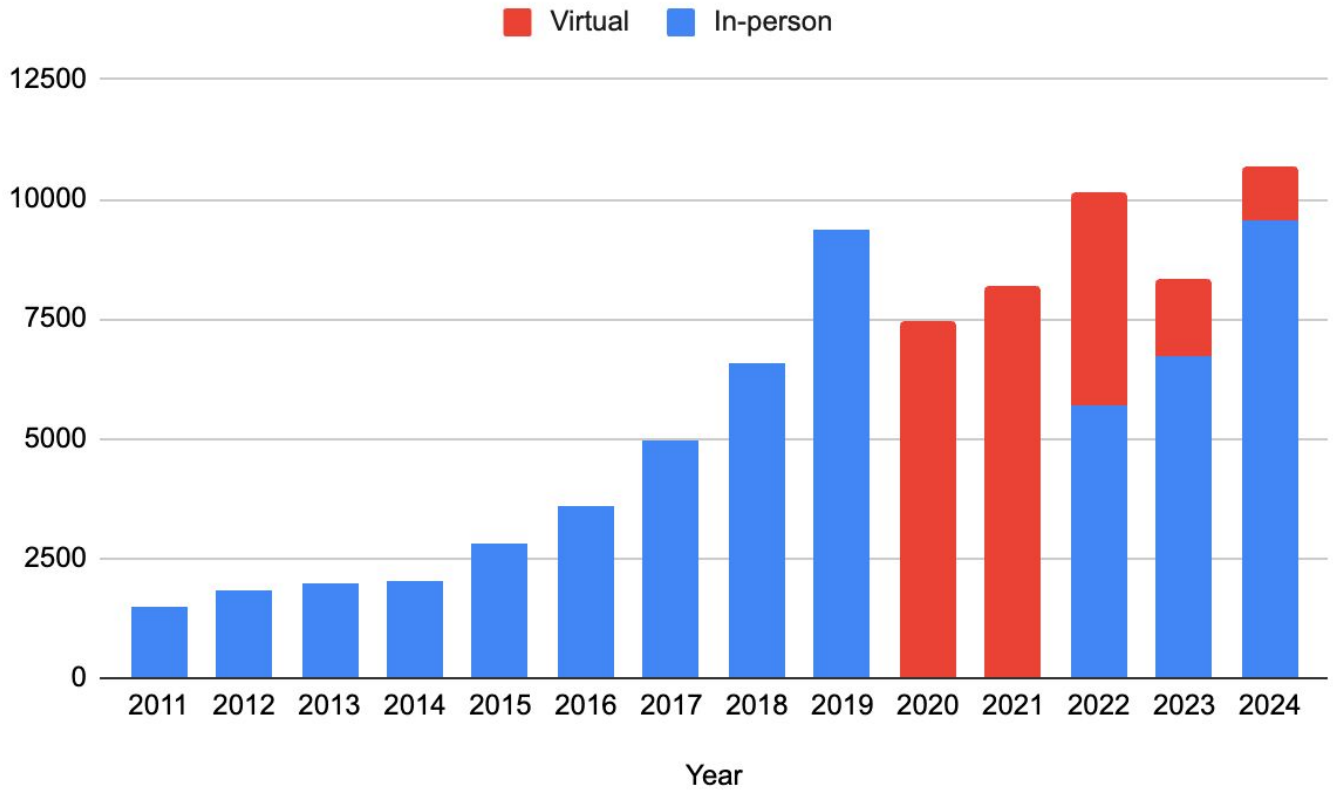
Sponsors and Expo

Thank you for your support!

CVPR 2024 Expo (Exhibit Hall Summit)

- 121 leading organizations
- 22,100 sq. ft.
- Companies from across the globe
- Organized by Hall-Erickson, Inc.

Attendance



*As of 6/12

CVPR 2024: ~12,000 Attendees from 76 Countries / Regions

United States 5074
Canada 352
Mexico 16
Puerto Rico 1

Brazil 27
Colombia 11
Ecuador 7
Peru 6
Argentina 4
Chile 2
Costa Rica 1
Jamaica 1

Germany 377
United Kingdom 330
Switzerland 171
France 141
Italy 118
Netherlands 64
Spain 60
Sweden 56
Belgium 34
Denmark 33
Austria 23
Czech Republic 22
Turkey 18
Norway 13
Ireland 12
Greece 11

Finland 16
Poland 15
Hungary 9
Portugal 9
Slovenia 9
Luxembourg 8
Portugal 9
Serbia 4
Croatia 3
Estonia 3
Romania 3
Albania 1
Slovakia 1

Ethiopia 69
Nigeria 9
Senegal 3
Rwanda 2
Burkina Faso 1

Burundi 1
Cameroon 1
Ghana 1
Mauritius 1

Australia 125
New Zealand 8

China 1511
South Korea 775
Japan 347
Singapore 171
Hong Kong 134
Israel 117
India 110
Taiwan 83
United Arab Emirates 40
Saudi Arabia 35
Viet Nam 22
Qatar 6
Thailand 6
Armenia 3
Bangladesh 3
Afghanistan 2
Iran 2
Macau 2
Kazakhstan 1
Malaysia 1
Pakistan 1

*As of 6/12

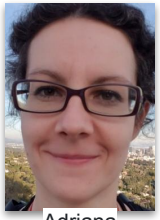
DEI Initiatives

Diversity, Equity and Inclusion Chairs


Shuran Song


CJ Taylor


Sara Beery


Adriana Kovashka


Tamara Berg


Roni Sengupta


Mike King

Travel Support for Attendees

- $200K allocated by the IEEE Computer Society and CVF for DEI Initiative
- 807 Applicants
- 181 Travel Scholarships
- 296 Registration Waivers

Website Accessibility

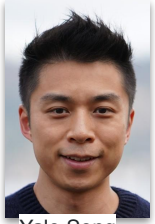
- Ongoing audit of CVPR website
- Improvements to support future CVPRs and related ML conferences

Accessibility Chair


Danna Gurari

Social Activities Chairs


Vitor Albiero


Yale Song


Giovanni Maria Farinella

Social Program

Socialization a key part of the conference experience

- Community driven socials
  - Challenges / Opportunities for ECRs in Fast Paced AI
  - How to Balance Research Interests and Academic Tasks
  - Diversity and Inclusion for Everyone
  - How to Know Your True Market Value as an AI Researcher
  - CV Entrepreneurship — Founders, Freelancers & Friends

  All attendees were invited to register for these events but they are limited to those who received confirmation (show this at the door)

- Musical performances at conference reception (June 20 7-9pm, Summit Ballroom)

Student Activities

Doctoral Consortium (Thurs. 11:30am)

- Opportunity for recent grads/close-to-complete PhD students to interact with experienced researchers
- Two-to-one matching of students to mentors
- Discuss career plans and research
- 36 students

Doctoral Consortium Chairs


Aparna Bharati


Nathan Jacobs

Speed Mentoring (Thurs. 12pm)

- Open call was issued to all students
- Informal small-group discussions with senior faculty and industry leaders
- 108 students

Social Activities Chairs


Vitor Albiero

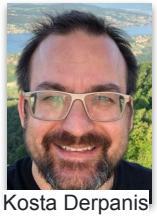

Yale Song

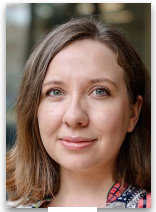

Giovanni Maria Farinella

Social Media

- Regular posts to Twitter/X (@CVPR) and Weibo (https://weibo.com/u/7778156717)
- News of Twitter's death greatly exaggerated…
  - Lots of community dialogue!

Publicity Chairs

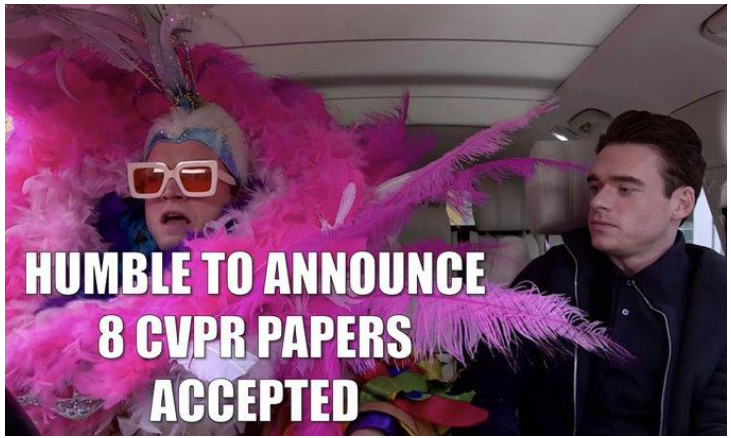

Kosta Derpanis


Abby Stylianou


Boqing Gong


Shenghua Gao


Jia-Bin Huang

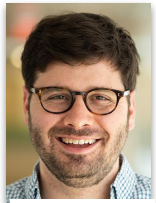


Workshops and Tutorials

- 123 workshops
  - **29 Thematic Tracks**
- 24 Tutorials

Workshop Chairs

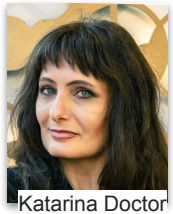

Luisa Verdoliva


Andrew Owens


Abhinav Shrivastava


Antitza Dantcheva

Tutorial Chairs


Katarina Doctor

CVPR 2024 Workshops

Track on 3D Scene Understanding

Multimodalities for 3D Scenes	Changan Chen	06/17	PM	Arch 2B
OpenSUN3D: 2nd Workshop on Open-Vocabulary 3D Scene Understanding	Francis Engelmann	06/18	PM	Arch 211
ScanNet++ Novel View Synthesis and 3D Semantic Understanding Challenge	Angela Dai	06/18	AM	Arch 211

Track on 3D Vision

2nd Workshop on Compositional 3D Vision	Habib Slim	06/18	AM+PM	Summit 327
(3rd) Monocular Depth Estimation Challenge	Matteo Poggi	06/18	PM	Summit 331

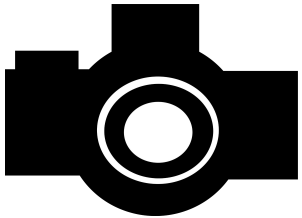
Demos

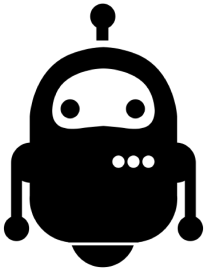
Demo Chairs


Shu Kong


Sathya Aakur

- 52 Demos (Arch 4CDE)
  - Session 1: Weds. 10:30am - 6:45pm
  - Session 2: Thurs. 10:30am - 6:45pm
  - Session 3: Fri. 10:30am - 6:45pm





Art Gallery

AI Art Curator

Luba Elliott

The gallery will be on view between 10.30am and 6.45pm between Weds. and Fri. in Arch 4CDE

Immersive Art Experience in 4F when expo talks are not scheduled

Check out the CVPR Virtual Gallery: https://thecvf-art.com/

Credit: Valdemar Danry x Cenk Güzelis

Special Art Sessions:

Weds.

- **11am** Gallery Tour with Curator and Artists
- **5pm** Gallery Tour with Curator and Artists

Thurs.

- **11am** Gallery Tour with Curator and Artists
- **1:30pm** Panel Discussion with Tom White, Varvara Guljajeva, Valdemar Danry, Avijit Ghosh and Luba Elliott
- **5pm** Gallery Tour with Curator and Artists

Fri.

- **11am** Gallery Tour with Curator and Artists

Credit: Jessica Tucker

CVPR 2024 Conference T-Shirt

Remember to pick your shirt up in the Expo Hall!

Our AI artists for this year's shirt:


Aaron Park


Daniel Geng


Andrew Owens

"Factorized Diffusion: Perceptual Illusions by Noise Decomposition"

https://dangeng.github.io/factorized_diffusion

Housekeeping Notes

- Posters can be picked up at the POSTER DISTRIBUTION CENTER in the back of 4E.
- Please hold on to your badge! We are not able to reprint them under any circumstances.
- Guests are not allowed to attend any event. All attendees must have on a badge to attend any part of CVPR.
- Accessibility questions? Check out our FAQ: https://cvpr.thecvf.com/Conferences/2024/Accessibility
- **Be kind to the event staff and each other**

IEEE COMPUTER SOCIETY
Assist

Committee to Assist in Reporting Ethics and Conduct Violations

David Forsyth
UIUC

Linda Shapiro
UW (Onsite)

Derek Hoiem
UIUC (Onsite)

Angjoo Kanazawa
Berkeley

- IEEE Computer Society and the Computer Vision Foundation are committed to a positive environment for all members of our diverse CVPR community. We do not tolerate bullying, harassment, or other misconduct.
- The **IEEE CS Committee to Assist in Reporting Ethics & Conduct Violations** (CS Assist) is on-site to help anyone who needs it report any issues to the IEEE Ethics Reporting Line or other appropriate authority.

Questions or Need Assistance?

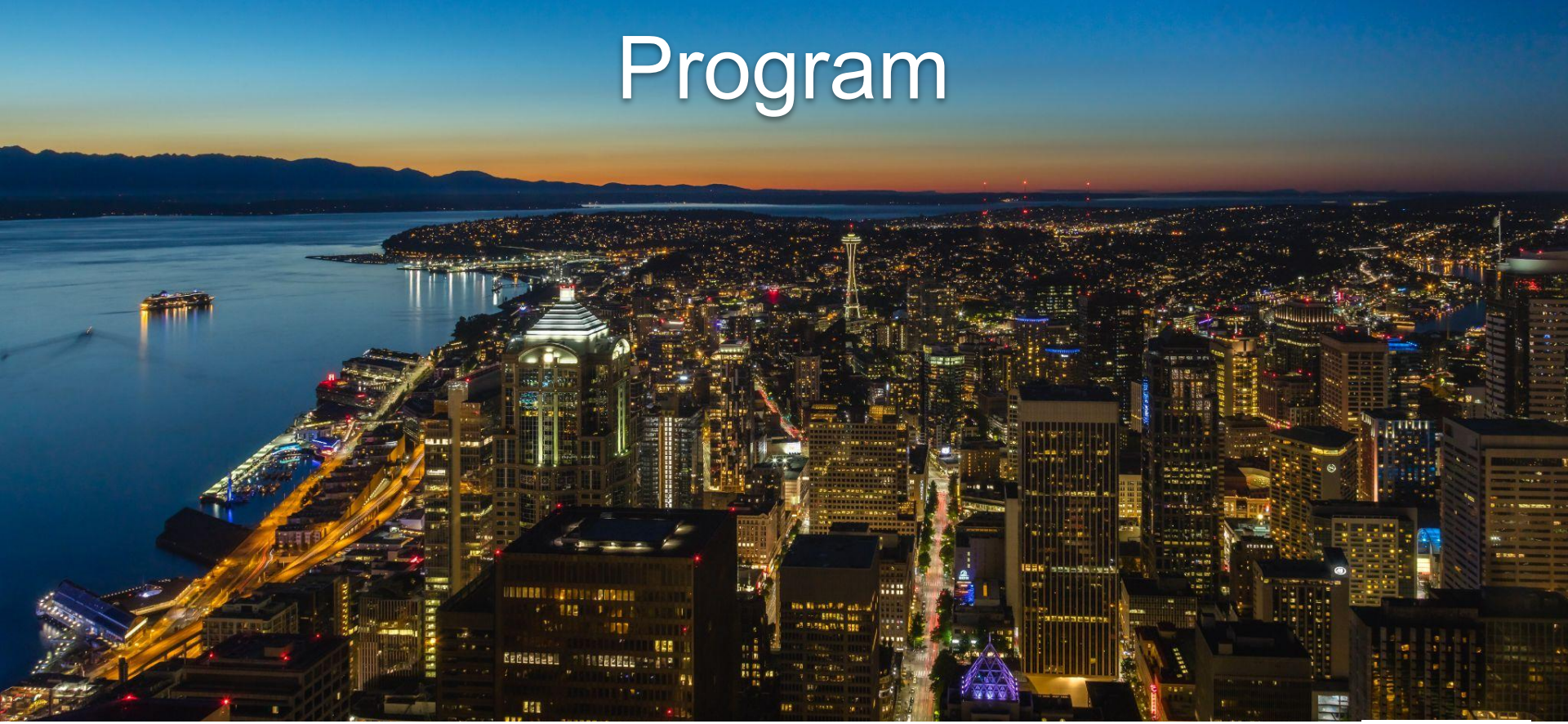
Assist Committee Members are available to support you!

Email: assist-cvpr@computer.org

IEEE Ethics Reporting Line: www.ieee-ethics-reporting.org | +1 888-359-6323

CVF IEEE COMPUTER SOCIETY IEEE

Program

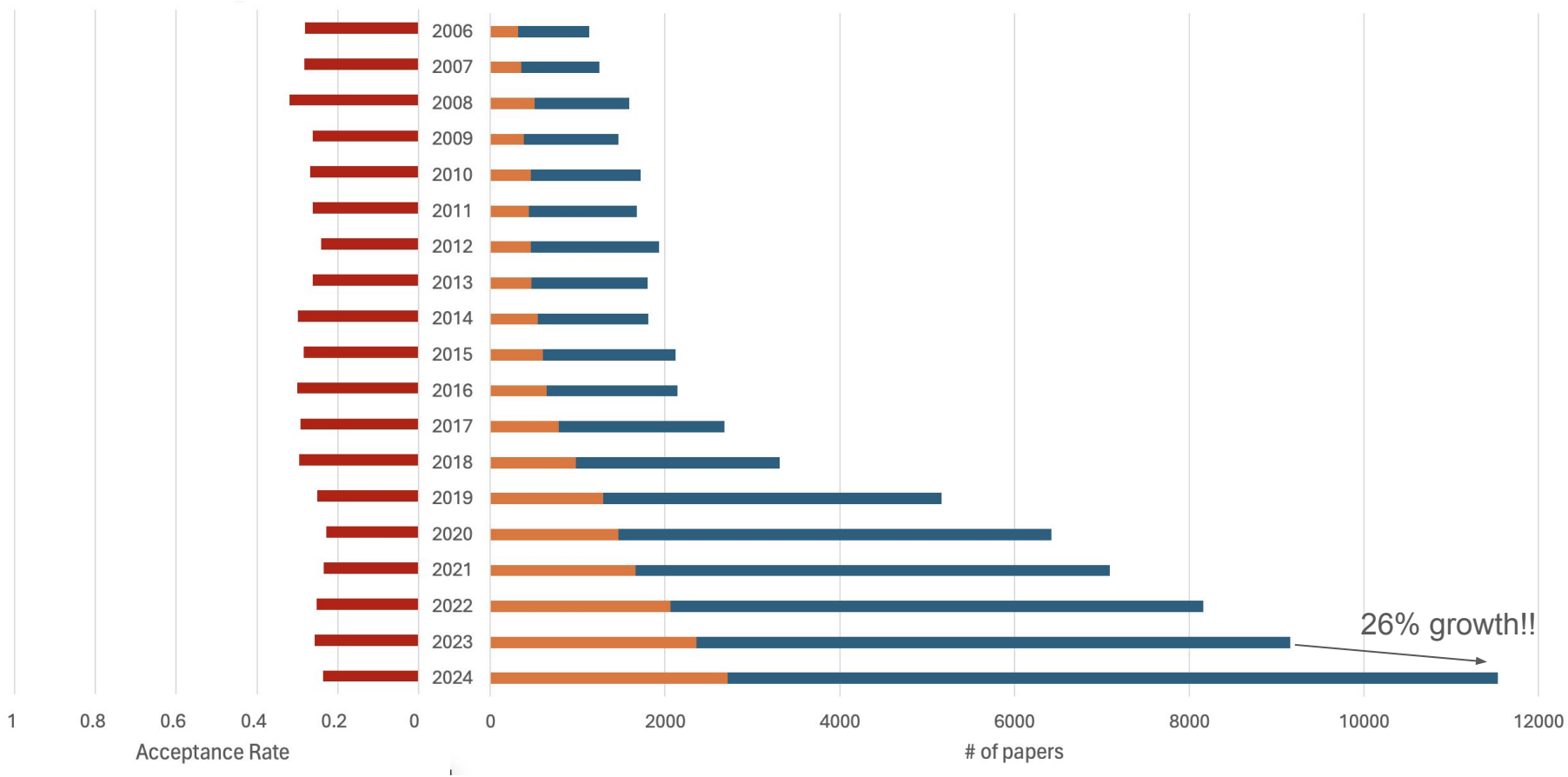


IEEE COMPUTER SOCIETY

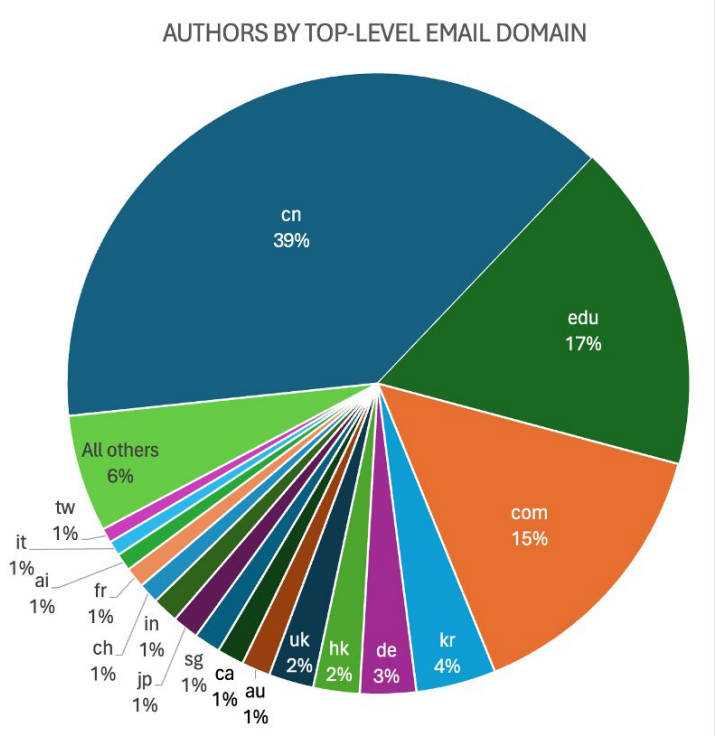
CVPR SEATTLE, WA JUNE 17-21, 2024

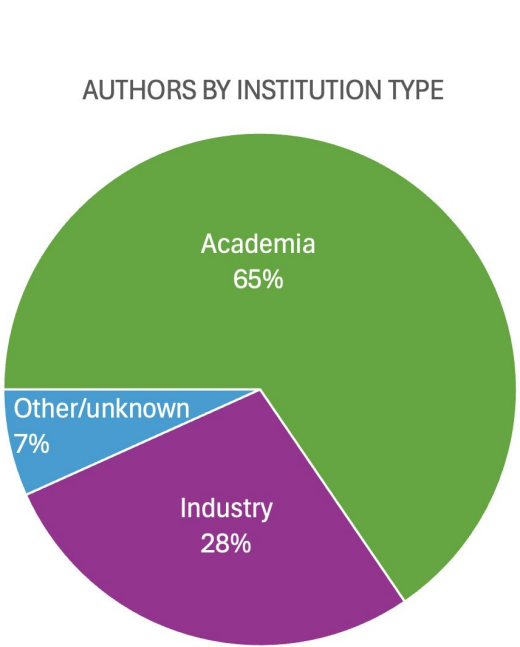
CVF

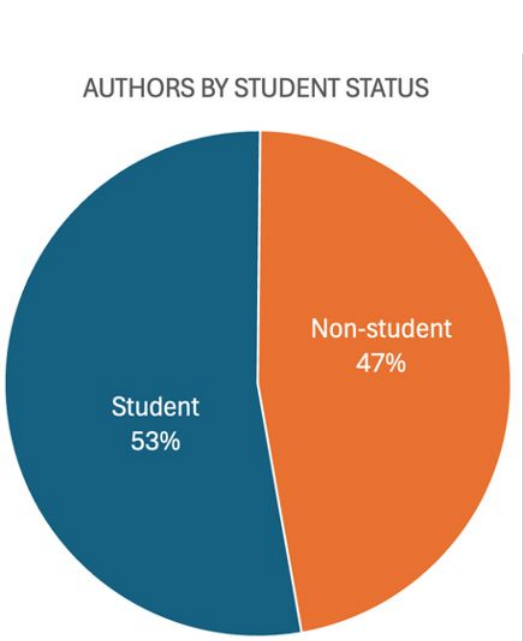
New records: 11,532 submissions, 2,719 accepted papers!



35,691 registered authors!







Program Chairs

Zeynep Akata
U. Tübingen

Ali Farhadi
U. Washington / Allen Institute for AI

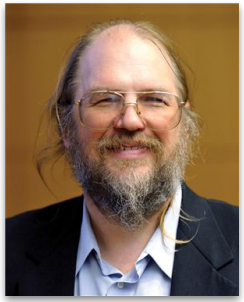
David Crandall
Indiana U.

Jianxin Wu
Nanjing U.

Robert Pless
George Washington U.

Imari Sato
National Institute of Informatics

Senior Advisor to the PCs



David Forsyth
UIUC

Technical Chair

Yoshitomo Matsubara
Spiffy AI

Senior Area Chairs

Alex Schwing
Alexei A Efros
Aniruddha Kembhavi
Anthony Hoogs
Aude Oliva
Bernard Ghanem
Bharath Hariharan
Bohyung Han
Fernando De la Torre
Jingdong Wang
Jingyi Yu
Juan Carlos Niebles
Kyoung Mu Lee
Nathan Jacobs
Pascal Fua
Philipp Kraehenbuehl
Phillip Isola
Richard Souvenir
Roozbeh Mottaghi
Ruigang Yang
Sing Bing Kang
Srinivasa Narasimhan
Tal Hassner
Yoichi Sato

Area Chairs (n=477)

Aayush Bansal, Abby Stylianou, Abhinav Shrivastava, Abhishek Gupta, Adam Czajka, Adam Kortylewski, Adriana Kovashka, Aishwarya Agrawal, Akihiro Sugimoto, Alejandro F. Frangi, Alexander C. Berg, Alexander T Toshev, Alexandre Alahi, Alimoor Reza, Alireza Fathi, Aljosa Osep, Ameesh Makadia, Amir Zamir, Andre Araujo, Andrea Tagliasacchi, Andrew Owens, Angel X Chang, Angela Yao, Anh Tuan Tran, Animesh Garg, Anna Rohrbach, Anpei Chen, Anurag Arnab, Arsha Nagrani, Arun Mallya, Asako Kanezaki, Ashok Veeraraghavan, Aswin C. Sankaranarayanan, Baoyuan Wu, Bastian Leibe, Ben Mildenhall, Bin Fan, Bing Su, Binh-Son Hua, Björn Ommer, Bo Chen, Bolei Zhou, Boqing Gong, Boxin Shi, Boyi Li, Brendan Tran Morris, Bruce Allen Maxwell, Bryan A. Plummer, Bryan Morse, Bumsub Ham, Carl Doersch, Carl Vondrick, Chang Xu, Chao Ma, Chao-Yuan Wu, Chaowei Xiao, Charless Fowlkes, Chen Change Loy, Chen Sun, Chen Wang, Chengjiang Long, Chenliang Xu, Chenyou Fan, Chetan Arora, Chong Luo, Christian Rupprecht, Christopher Clark, Christopher Funk, Chuang Gan, Chuhang Zou, Chunyu Wang, Cihang Xie, Cornelia Fermuller, Cristian Canton Ferrer, Cuiling Lan, Cuong V. Nguyen, Dahun Kim, Dan Xu, Danna Gurari, David B. Lindell, David Fouhey, David Novotny, Deng-Ping Fan, Devis Tuia, Di He, Di Huang, Diane Larlus, Dim Papadopoulos, Dima Damen, Dimosthenis Karatzas, Dinesh Jayaraman, Dingwen Zhang, Dong Chen, Dong Xu, Dongfang Liu, Eddy Ilg, Elisa Ricci, Emily Morgan Hand, Enrique Dunn, Eshed Ohn-Bar, Evan Shelhamer, Evangelos Kalogerakis, Fahad Khan, Fartash Faghri, Fatma Guney, Federico Tombari, Feng Lu, Francesc Moreno-Noguer, Gang Yu, Gedas Bertasius, Georgia Gkioxari, Georgios Pavlakos, Georgios Tzimiropoulos, Gim Hee Lee, Giovanni Maria Farinella, Gordon Wetzstein, Guanbin Li, Guansong Pang, Gül Varol, Guosheng Lin, Guoyu Lu, Hadi Pouransari, Hae-Gon Jeon, Hajime Nagahara, Hamid Rezatofighi, Han-Jia Ye, Hanbyul Joo, Hao Chen, Hao Dong, Henghui Ding, Hengshuang Zhao, Hideki Nakayama, Hilde Kuehne, Hirokatsu Kataoka, Hiroshi Kawasaki, Hong Chang, Hong Xuan, Hongyang Li, Hossein Rahmani, Hsin-Ying Lee, Hu Han, Huaizu Jiang, Hui Ji, Huijuan Xu, Huiwen Chang, Hung-Yu Tseng, Hyun Soo Park, Hyunjung Shim, Hyunwoo J. Kim, Iacopo Masi, Ioannis Gkioulekas, Ishan Misra, Ishani Chakraborty, Jaesik Park, James Hays, Jason J Corso, Jasper Uijlings, Jean-Francois Lalonde, Jeany Son, Jiahuan Zhou, Jiajun Wu, Jian Wang, Jianfei Cai, Jiankang Deng, Jianwen Xie, Jimei Yang, Jingya Wang, Jinwei Gu, Joao P Barreto, Jonathan T. Barron, Jong Chul Ye, Jonghyun Choi, Jordi Pont-Tuset, Joseph Tighe, Judy Hoffman, Jufeng Yang, Jun Liu, Jun-Yan Zhu, Jungseock Joo, Junseok Kwon, Junsong Yuan, Kai Han, Kaiyang Zhou, Kannappan Palaniappan, Karteek Alahari, Katerina Fragkiadaki, Keiji Yanai, Ken Sakurada, Kevin J Liang, Khoa Luu, Kiana Ehsani, Kihyuk Sohn, Kris M. Kitani, Kristin Dana, Kwang In Kim, Kwang Moo Yi, Lam M. Nguyen, Lamberto Ballan, Laura Leal-Taixé, Laurent Kneip, Le Lu, Lei Wang, Leif Kobbelt, Li Erran Li, Li Fuxin, Li Yi, Li Zhang, Liang Lin, Liang Zheng, Liang-Chieh Chen, Limin Wang, Lingjie Liu, Lingxi Xie, Linjie Yang, Liwei Wang, Long Chen, Lourdes Agapito, Lu Jiang, Lu Sheng, Lu Yuan, Luca Weihs, M. Salman Asif, Mahdi S. Hosseini, Mang Ye, Manohar Paluri, Manolis Savva, Marcello Pelillo, Marco Lorenzi, Marcus Rohrbach, Margret Keuper, Maria Vakalopoulou, Mark Yatskar, Massimiliano Mancini, Matej Kristan, Matt Feiszli, Matthew Brown, Matthew O'Toole, Matthieu Cord, Mehrdad Farajtabar, Miaomiao Liu, Michael Maire, Michael Wray, Mike Zheng Shou, Min H. Kim, Min Sun, Mingkui Tan, Mingsheng Long, Mingze Xu, Minh Hoai, Minjie Cai, Minsu Cho, Mohamed Elhoseiny, Mohit Gupta, Mohsen Ali, Muhammad Haris Khan, Negar Rostamzadeh, Neill D. F. Campbell, Nianyi Li, Niloy Mitra, Ning Yu, Noah Snavely, Norimichi Ukita, Nuno Vasconcelos, Oana-Maria Camburu, Oisin Mac Aodha, Oliver Wang, Oncel Tuzel, Pablo Arbelaez, Paolo Favaro, Pascal Mettes, Paul Hongsuck Seo, Pedro Morgado, PENG WANG, Peter Hedman, Peter Vincent Gehler, Peter Wonka, Peyman Milanfar, Philippos Mordohai, Ping Hu, Ping Luo, Piotr Bojanowski, Piotr Koniusz, Pratul P. Srinivasan, Praveen Krishnan, Qi Dai, Qi Wu, Qian Yu, Qifeng Chen, Qilong Wang, Qixing Huang, Rafal Mantiuk, Raffay Hamid, Rahaf Aljundi, Ran He, Rana Hanocka, Ranjay Krishna, Ravi Ramamoorthi, Raymond A. Yeh, Rei Kawakami, Richard Zhang, Robby T. Tan, Rohit Girdhar, Roman Pflugfelder, Ryoma Bise, Sai-Kit Yeung, Salman Khan, Sanjeev Jagannatha Koppal, Sara Beery, Sarah Adel Bargal, Sareh Rowlands, Sathyanarayanan N. Aakur, Satoshi Tsutsui, Saurabh Gupta, Saurabh Singh, Scott McCloskey, Scott Workman, Seong Joon Oh, Ser-Nam Lim, Sergey Tulyakov, Seunghoon Hong, Seungryong Kim, Seungyong Lee, Shalini De Mello, Shang-Hong Lai, Shanshan Zhang, Shaohui Lin, Sharon X Huang, Shayok Chakraborty, Shenghua Gao, Shenlong Wang, Shijian Lu, Shin'ichi Satoh, Shizhe Chen, Shohei Nobuhara, Shubham Tulsiani, Shuicheng YAN, Siavash Arjomand Bigdeli, Siyu Tang, Song Bai, Srinath Sridhar, Stefan Lee, Stefano Mattoccia, Stella X. Yu, Stephan Alaniz, Stephan Richter, Stephen James, Stratis Gavves, Subhashini Venugopalan, Suha Kwak, Tae Hyun Kim, Tae-Hyun Oh, Tae-Kyun Kim, Taesung Park, Taesup Moon, Takayuki Okatani, Tanmay Gupta, Tat-Jen Cham, Tat-Jun Chin, Tatsuya Harada, Thibaut Durand, Thomas Kipf, Tian Han, Tianfan Xue, Ting Yao, Ting Zhang, Tinne Tuytelaars, Todd Zickler, Tolga Birdal, Tony Tung, Toshihiko Yamasaki, Tsung-Yi Lin, Tushar Nagarajan, Ulugbek S. Kamilov, Unnat Jain, Varun Jampani, Venkatesh Babu Radhakrishnan, Viktoriia Sharmanska, Vincent Lepetit, Vincent Sitzmann, Vishal M. Patel, Vítor Albiero, Vittorio Murino, Vlad I Morariu, Wangmeng Zuo, Waqas Sultani, Wei Liu, Wei Yang, Wei-Chih Hung, Wei-Chiu Ma, Weihong Deng, Weisong Shi, Wen Li, Wenguan Wang, Wenqi Ren, William Robson Schwartz, Xi Peng, Xiangyu Xu, Xiangyu Yue, Xiaoguang Han, Xiaojuan Qi, Xiaoming Liu, Xiaowei Zhou, Xiaoyu Wang, Xihui Liu, Xin Wang, Xin Yang, Xin Yu, Xingang Pan, Xinggang Wang, Xinlei Chen, Xiu-Shen Wei, Xueting Li, Xuming He, Xun Cao, Yale Song, Yan Huang, Yan Zhang, Yanchao Yang, Yang Song, Yang Wang, Yang You, Yannis Kalantidis, Yaoyao Liu, Yasushi Makihara, Yebin Liu, Yedid Hoshen, Yen-Yu Lin, Yezhou Yang, Yi Wu, Yi-Hsuan Tsai, Yi-Ting Chen, Yibing Song, Yifei Huang, Yijun Li, Yin Li, Ying Fu, Ying Wu, Yinghuan Shi, Yingli Tian, Yinqiang Zheng, Yisen Wang, Yiyi Liao, Yonatan Bisk, Yong Jae Lee, Yosi Keller, Yu Cheng, Yu Kong, Yu Li, Yu Wu, Yu-Chuan Su, Yu-Xiong Wang, Yuandong Tian, Yuchao Dai, Yuki M Asano, Yulun Zhang, Yumin Suh, Yunchao Wei, Yung-Yu Chuang, Yunzhu Li, Yusuke Sugano, Yuyin Zhou, Zaid Harchaoui, Zhangyang Wang, Zhaowei Cai, Zhaoxiang Zhang, Zhe Gan, Zheng Zhang, Zhengming Ding, Zhengqi Li, Zhirong Wu, Zhizhong Han, Zhun Zhong, Ziwei Liu, Zongwei Zhou, Zsolt Kira, Zuxuan Wu, Zuzana Kukelova

Reviewers (n=9,872)

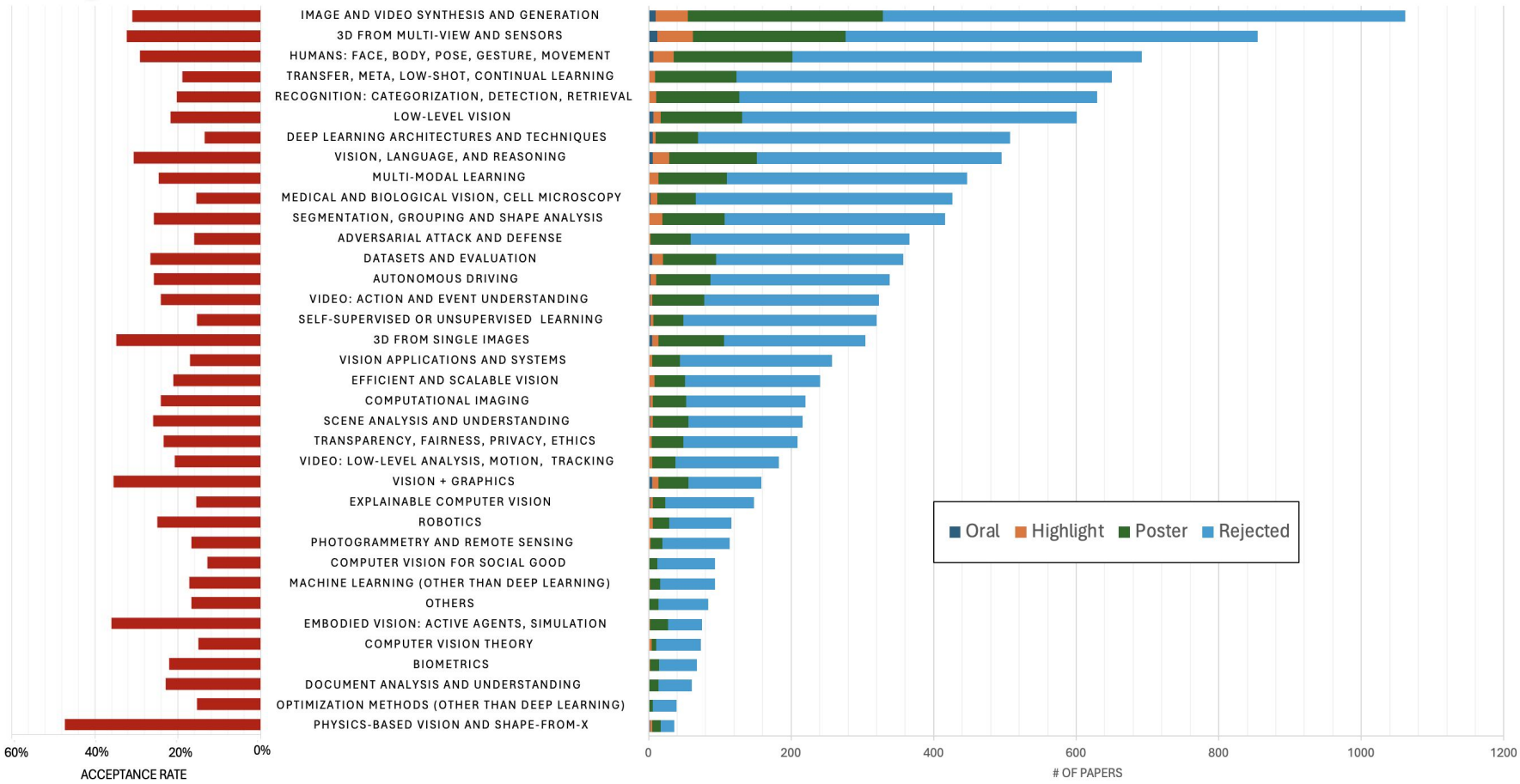
Outstanding Reviewers

Aadarsh Sahoo, Aaron Walsman, Abhinav Shukla, Adnan Qayyum, Adrian Lopez, Adrian Penate-Sanchez, Ahmed Taha, Alessio Xompero, Alexander Hermans, Alexandros Haliassos, Alvaro Budria, Amine Bourki, Ana C Murillo, Anastasia Antsiferova, Andrea Conti, Andrew Brown, Anh-Quan Cao, Anil Usumezbas, Anna Kukleva, Arjun Karpur, Armin Hadzic, Ashish Ramayee Asokan, Assia Benbihi, Ayush Jain, Ayush Saraf, Benjamin Busam, Bicheng Xu, Bingliang Jiao, Bumsoo Kim, Carlo Masone, Carlos Rodriguez-Pardo, Changqing Zhang, Chenglin Yang, Chengze Miaomiao Li, Cheston Tan, Chiao An Yang, Chi Xu, Chunghyun Park, Cristian Rodriguez-Opazo, Dahyun Kang, Daqi Liu, Davide Cozzolino, David Fan, Da-Wei Zhou, Despoina Paschalidou, Di Hu, Dominique Ginhac, Donald G. Dansereau, Dongbo Min, Dong-Jun Han, Effrosyni Mavroudi, Elisavet Konstantina, Stathopoulou, Evangelos Kazakos, Fabio Tosi, Faisal Z Qureshi, Forrester Cole, Francis Engelmann, Francois Rameau, Fuyong Xing, Ganesh Subramanian Iyer, Gilles Puy, Ginger Delmas, Goutam Bhat, Guoli Jia, Gurumurthy Swaminathan, Hakan Bilen, Hanchen Xie, Hannah Kerner, Haomin Liu, Hao Zhu, Hawook Jeong, Hazel Doughty, Hermann Blum, Hongwei Bran Li, Huan Wang, Hui Liang, Inhwan Bae, Ioannis Papoutsis, Iro Laina, Ishan Rajendrakumar Dave, Jaeyoo Park, Jan Eric Lenssen, Jangwon Lee, Jaskirat Singh, Jens Behley, Jiangpeng He, Jianyi Wang, Jieji Ren, Jihyong Oh, Jingkang Yang, Jingyuan Yang, Joe Mathai, Joerg Stueckler, Jonghee Kim, Juan Camilo Perez, Julian Straub, Junhwa Hur, Jun Xiao, Junyao Hu, Kevin Qinghong Lin, Khang Truong Giang, Kiru Park, Konda Reddy Mopuri, Kongming Liang, Konstantinos Batsos, Kumar Ashutosh, Lin Geng Foo, Lin Zhu, Luisa Verdoliva, M. Saquib Sarfraz, M. Usman Rafique, Mahmoud Afifi, Mannat Singh, Martin R. Oswald, Matteo Poggi, Matthew Joseph Leotta, Mattia Savardi, Mayu Otani, Michaël Gharbi, Michael S. Brown, Michael Weinmann, Michel Antunes, Ming Jiang, Mohamed E. Hussein, Mohammad Reza, Hosseinzadeh Taher, Monica Hernandez, Mo Zhou, Nan Pu, Nicholas Kolkin, Paola Cascante-Bonilla, Patrick Ebel, Paul-Edouard Sarlin, Philipp Lindenberger, Qiang Xu, Qian Zheng, Qize Yang, Renato Martins, Reuben Tan, Riccardo Volpi, Richard Newcombe, Ruicong Liu, Ruihan Gao, Saihui Hou, Samaneh Azadi, Satwik Kottur, Satya Krishna Gorti, Shuhong Zheng, Shu Zhang, Shyamal Buch, Siamul Karim Khan, Sima Behpour, Simon Stepputtis, Siyan Dong, Sourav Garg, Spencer Whitehead, Stefan Leutenegger, Stefano Berretti, Sunghoon Im, Swathikiran Sudhakaran, Terrance E. Boult, Thomas Hummel, Toby Perrett, Torsten Sattler, Touqeer Ahmad, Valerio Cambareri, Vishaal Udandarao, Volker Rodehorst, Williem, Wojciech Zielonka, Wonwoong Cho, Xiang An, Xiang Li, Xiangyu He, Xiaohan Zhang, Xingyi Zhou, Xiongwei Wu, Yanbin Liu, Yankai Jiang, Yannick Hold-Geoffroy, Yi Dong, Yihua Zhang, Yinan Zhao, Yongzhen Huang, Yuanhao Cai, Yuesong Wang, Yu-Lun Liu, Yunqi Miao, Yunze Man, Yurui Zhu, Yusuke Matsui, Yu-Wei Chao, Yuyu Guo, Zan Gojcic, Zerong Zheng, Zheheng Jiang, Zhekun Luo, Zheng Chen, Zhibin Liao, Zhicheng Zhang, Zhihe Lu, Zhou Yu

Paper decisions

- Each paper received 3 reviews and a meta-review from an Area Chair.
- Decisions made within triplets of ACs.
- We decided to bring back orals, but also keep the CVPR 2023 innovation of “highlighted” posters
- Orals and highlight candidates collected from the ACs
- Overall acceptance rate: 23.6%
  - 90 (3.3%) of papers are Orals+posters
  - 324 (11.9%) of papers are “highlights” posters, with special annotation in the program
  - 2,305 additional posters

Submissions and decisions, by primary topic



	Wednesday	Thursday	Friday
8:30 AM	**Opening**		
9:00 AM	**Orals** 1A Low-level vision 1B Vision and Graphics 1C Humans: Face, body, pose, gesture, movement	**Orals** 3A 3D from single view 3B Vision, Language, and Reasoning 3C Medical and Physics-based vision	**Orals** 5A Datasets and evaluation 5B 3D from multiview and sensors 5C Low-shot, self-supervised, semi-supervised learning
10:30 AM	**Posters**	**Posters**	**Posters**
12:00 PM	Lunch	Lunch	Lunch
1:00 PM	**Orals** 2A Image & Video Synthesis 2B Deep learning architectures & techniques 2C 3D from multiview and sensors	**Orals** 4A Autonomous navigation & egocentric vision 4B 3D Vision 4C Action and motion	**Orals** 6A Low-level vision and remote sensing 6B Image & Video Synthesis 6C Multi-modal learning
2:45 PM	**Keynote** Joshua Bongard	**Keynote** David Baker	**Keynote** Sofia Crespo
4:00 PM	**Panel** Societal opportunities and challenges of AI	**PAMI-TC Meeting**	**Panel** CVPR: Past, Present, and Future
5:15 PM	**Posters**	**Posters**	**Posters**
7:00 PM		**Reception and Music**	

Keynotes

Wednesday 2:45pm

Joshua Bongard

Veinott Professor of Computer Science, University of Vermont

“The Tip and the Iceberg: Deep Learning and Embodiment”

Thursday 2:45pm

David Baker

Professor of Biochemistry, University of Washington

“Design of New Protein Functions Using Deep Learning”

Friday 2:45pm

Sofia Crespo

Artist

“Entanglements, Exploring Artificial Biodiversity”

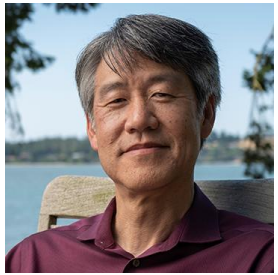
Panels

Societal opportunities and challenges of AI

Wednesday 4:00pm

Nicole Decari
Director of AI & Society
Allen Institute for AI

Oren Etzioni
Founder
Truemedia.org



Peter Lee
President
Microsoft Research

Fei-Fei Li
Sequoia Professor
Stanford University

Matt McIlwain
Managing Director
Madrona Ventures

Hadi Partovi
CEO
Code.org

CVPR: Past, Present, and Future

Friday 4:00pm

Kiana Ehsani
Senior Research Scientis
Allen Institute for AI

Dima Damen
Professor, Univ of Bristol and
Google DeepMind

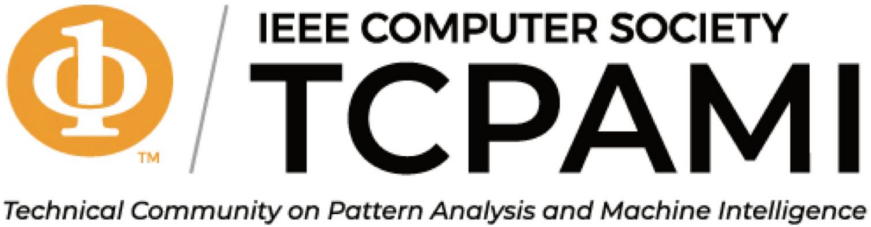
Cordelia Schmid
Research Director, INRIA
and Google

Ranjay Krishna
Assistant Professor
Univ of Washington

PAMI TC Meeting

4pm Thursday in Summit Flex Hall ABC

Motions will be discussed!



Expo Track Keynotes

Wednesday 1:15pm

Swami Sivasubramanian
Vice President of AI & Data, AWS
“Computer vision at scale: Driving customer innovation and industry adoption”

Thursday 10:30am

Andrea Gagliano
Senior Director of AI/ML, Getty Images
“Today’s Pictures, Tomorrow’s Training Data: The Synergy Between Human Creativity and AI”

Friday 9:00am

Ece Kamar
Managing Director, Microsoft Research AI Frontiers Lab
“Phase Transition in AI: Opportunities and Gaps Towards Making AI Real”

Awards

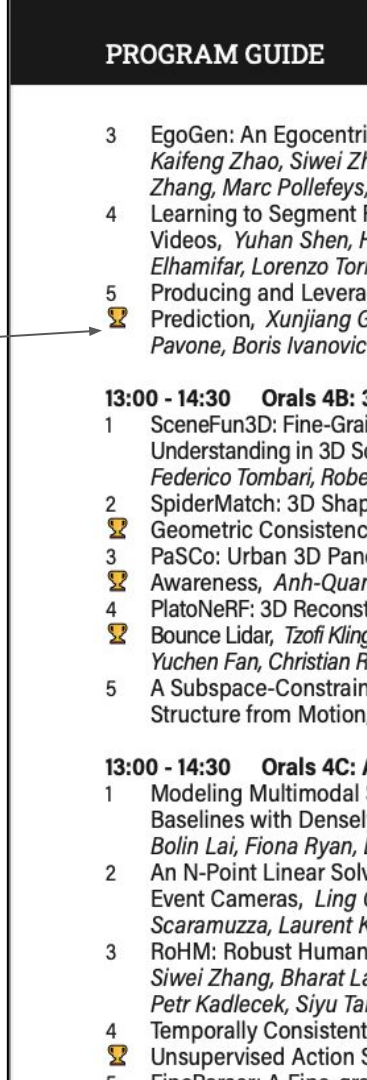
IEEE COMPUTER SOCIETY

CVPR SEATTLE, WA JUNE 17-21, 2024

CVF

Best Paper Awards

- 24 best paper award candidates were nominated by the Area Chairs, marked in the program
- From the candidates, 10 papers were selected for an award by the Program Chairs
  - 4 Best Student Paper Runners-Up
  - 2 Best Student Papers
  - 2 Best Paper Runners-Up
  - 2 Best Papers
- We'll present certificates to the award winners at the PAMI-TC meeting on Thursday



Best Student Paper Runner-Ups

SpiderMatch: 3D Shape Matching with Global Optimality and Geometric Consistency

Paul Roetzer[1] Florian Bernard[1]
University of Bonn[1]

Idea: Match Cyclic Path $\mathcal{C}$ on $\mathcal{X}$ to $\mathcal{Y}$ | Dense Geometrically Consistent Shape Matching | Scalability Analysis

Figure 1. (**Left**) Schematic illustration of our method: we extract a cyclic path $\mathcal{C}$ on the surface of 3D shape $\mathcal{X}$, and match $\mathcal{C}$ to the 3D target shape $\mathcal{Y}$. (**Middle**) If $\mathcal{C}$ covers the whole source shape $\mathcal{X}$ (not shown for clarity, see Sec. 4.1 for details), our method can be used to compute globally optimal results for dense non-rigid and geometrically consistent 3D shape matching. (**Right**) We show runtimes w.r.t. shape resolution and compare to Roetzer et al. [55] (which is the only existing method that achieves geometric consistency while having a global flavour and being initialization-free) and Cao et al. [14] (which does not provide any geometric consistency guarantees).

Best Student Paper Runner-Ups

Image Processing GNN: Breaking Rigidity in Super-Resolution

Yuchuan Tian[1], Hanting Chen[2], Chao Xu[1], Yunhe Wang[2*]

[1] National Key Lab of General AI, School of Intelligence Science and Technology, Peking University. [2] Huawei Noah's Ark Lab.

tianyc@stu.pku.edu.cn, yunhe.wang@huawei.com

Abstract

Super-Resolution (SR) reconstructs high-resolution images from low-resolution ones. CNNs and window-attention methods are two major categories of canonical SR models. However, these measures are rigid: in both operations, each pixel gathers the same number of neighboring pixels, hindering their effectiveness in SR tasks. Alternatively, we leverage the flexibility of graphs and propose the Image Processing GNN (IPG) model to break the rigidity that dominates previous SR methods. Firstly, SR is unbalanced in that most reconstruction efforts are concentrated to a small proportion of detail-rich image parts. Hence, we leverage degree flexibility by assigning higher node degrees to detail-rich image nodes. Then in order to construct graphs for SR-effective aggregation, we treat images as pixel node sets rather than patch nodes. Lastly, we

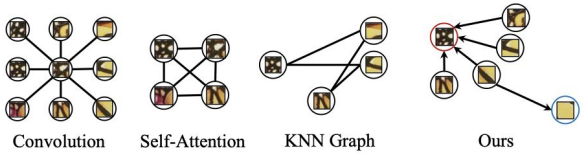


Figure 1. Convolution (left), Self-Attention (middle-left), KNN Graph Aggregation (middle-right), and Graph Aggregation in IPG (right). Compared with other methods, IPG graph aggregation considers the unbalanced nature of SR: detail-rich, high-frequency image nodes (red) have higher node degrees; while flat, low-frequency image nodes (blue) require fewer aggregations.

mote sensing [15], et cetera.

Despite various measures, it is taken for granted that mainstream SR models treat all pixels in a fairly rigid manner. For instance, as shown in Figure 1, in convolution layers of the CNN-based SR model VDSR [16], the

Best Student Paper Runner-Ups

Objects as volumes: A stochastic geometry view of opaque solids

Bailey Miller, Hanyu Chen, Alice Lai, Ioannis Gkioulekas
Carnegie Mellon University

Abstract

We develop a theory for the representation of opaque solids as volumes. Starting from a stochastic representation of opaque solids as random indicator functions, we prove the conditions under which such solids can be modeled using exponential volumetric transport. We also derive expressions for the volumetric attenuation coefficient as a functional of the probability distributions of the underlying indicator functions. We generalize our theory to account for isotropic and anisotropic scattering at different parts of the solid, and for representations of opaque solids as stochastic implicit surfaces. We derive our volumetric representation from first principles, which ensures that it satisfies physical constraints such as reciprocity and reversibility. We use our theory to explain, compare, and correct previous volumetric representations, as well as propose meaningful extensions that lead to improved performance in 3D reconstruction tasks.

1. Introduction

swers to queries such as "are two points mutually visible" (a visibility query) and "what is the distance to first intersection along a ray" (a ray-casting query), when the geometry occluding visibility and terminating rays is stochastic.

Volumetric representations for translucent and participating media are stochastic abstractions of their microscopic structure: Such media comprise numerous microscopic particles that reflect and occlude light rays. Modeling explicit microparticle configurations, and rendering light transport through them, is prohibitively expensive. As a compromise for efficiency, volumetric representations allow to simulate light transport in such media *on average* [6]. These representations replace explicit with *statistical* descriptions of microparticle configurations (e.g., average particle location, size, shape, and orientation), analogously to statistical BRDF models for surface microgeometry [13, 14, 24, 46, 60]. Computer graphics has developed volumetric representations for microparticle media that account for details such as varying particle size and material [16, 32], shape and orienta-

Best Student Paper Runner-Ups

Comparing the Decision-Making Mechanisms by Transformers and CNNs via Explanation Methods

Mingqi Jiang, Saeed Khorram, Li Fuxin
Collaborative Robotics and Intelligent Systems (CoRIS) Institute
Oregon State University
{jiangmi, khorrams, lif}@oregonstate.edu

Abstract

In order to gain insights about the decision-making of different visual recognition backbones, we propose two methodologies, sub-explanation counting and cross-testing, that systematically applies deep explanation algorithms on a dataset-wide basis, and compares the statistics generated from the amount and nature of the explanations. These methodologies reveal the difference among networks in terms of two properties called compositionality and disjunctivism. Transformers and ConvNeXt are found to be more compositional, in the sense that they jointly consider multiple parts of the image in building their decisions, whereas traditional CNNs and distilled transformers are less compositional and more disjunctive, which means that they use

tant contributing factor is not the attention itself but those design principles? If so, which specific design principles particularly affect the decision-making of networks? Better answers to those questions would help us to gain more insights into those deep and complicated black-box networks.

In this paper, we propose a novel methodology to examine these questions through *deep explanation algorithms*. Explanation algorithms have improved significantly in recent years and can generate accurate explanations that can be verified through *intervention experiments* on images [21, 26]. Recent search-based explanation algorithms can find a comprehensive set of *minimally sufficient explanations* (MSEs) [30], defined as the minimal set of patches that, if shown to the network, lead to predictions that are

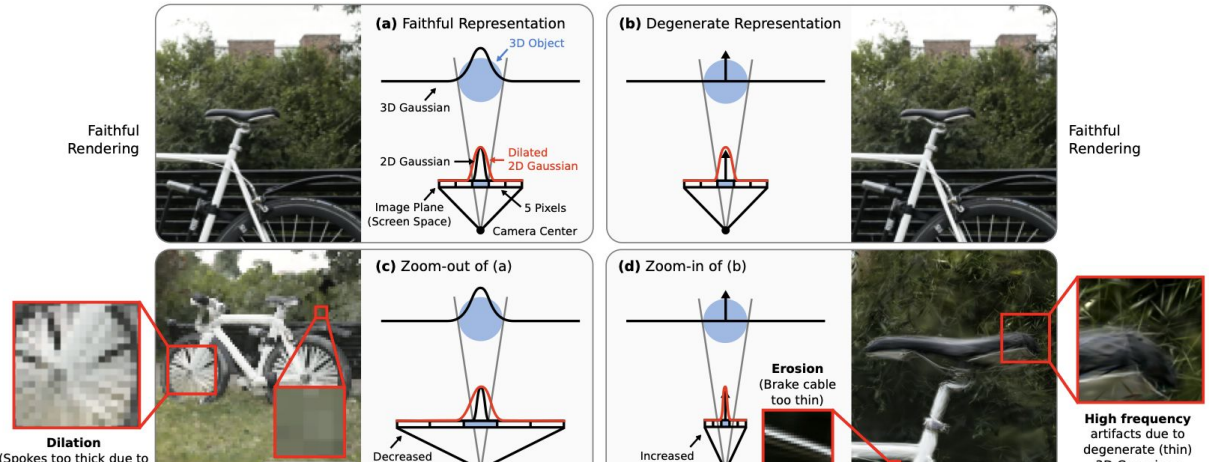
Best Student Papers

Mip-Splatting: Alias-free 3D Gaussian Splatting

Zehao Yu[1,2] Anpei Chen[†,1,2] Binbin Huang[3] Torsten Sattler[4] Andreas Geiger[1,2]

[1]University of Tübingen [2]Tübingen AI Center [3]ShanghaiTech University
[4]Czech Technical University in Prague

https://niujinshuchong.github.io/mip-splatting



Best Student Papers

BioCLIP: A Vision Foundation Model for the Tree of Life

Samuel Stevens[1*†], Jiaman Wu[1*], Matthew J Thompson[1], Elizabeth G Campolongo[1], Chan Hee Song[1], David Edward Carlyn[1], Li Dong[2], Wasila M Dahdul[3], Charles Stewart[4], Tanya Berger-Wolf[1], Wei-Lun Chao[1], and Yu Su[1†]

[1]The Ohio State University, [2]Microsoft Research, [3]University of California, Irvine, [4]Rensselaer Polytechnic Institute

Abstract

Images of the natural world, collected by a variety of cameras, from drones to individual phones, are increasingly abundant sources of biological information. There is an explosion of computational methods and tools, particularly computer vision, for extracting biologically relevant information from images for science and conservation. Yet most of these are bespoke approaches designed for a specific task and are not easily adaptable or extendable to new questions, contexts, and datasets. A vision model for general organismal biology questions on images is of timely need. *To*

79–81, 87, 88] into actionable information (e.g., species classification, individual identification, and trait detection) has accelerated and enabled new advances in tasks such as species delineation [32], understanding mechanisms of adaptation [23, 39], abundance and population structure estimation [3, 40, 58, 82], and biodiversity monitoring and conservation [83].

However, applying computer vision to answer a biological question is still a laborious task requiring substantial machine learning expertise and effort—biologists must manually label sufficient data for the specific taxa and task of interest, and find and train a suitable model for the task. Meanwhile, foundation models [12] such as CLIP [69] and

Best Paper Runners-Up

EventPS: Real-Time Photometric Stereo Using an Event Camera

Bohan Yu[1,2] Jieji Ren[3] Jin Han[4,5] Feishi Wang[1,2] Jinxiu Liang[1,2] Boxin Shi[1,2*]

[1] National Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

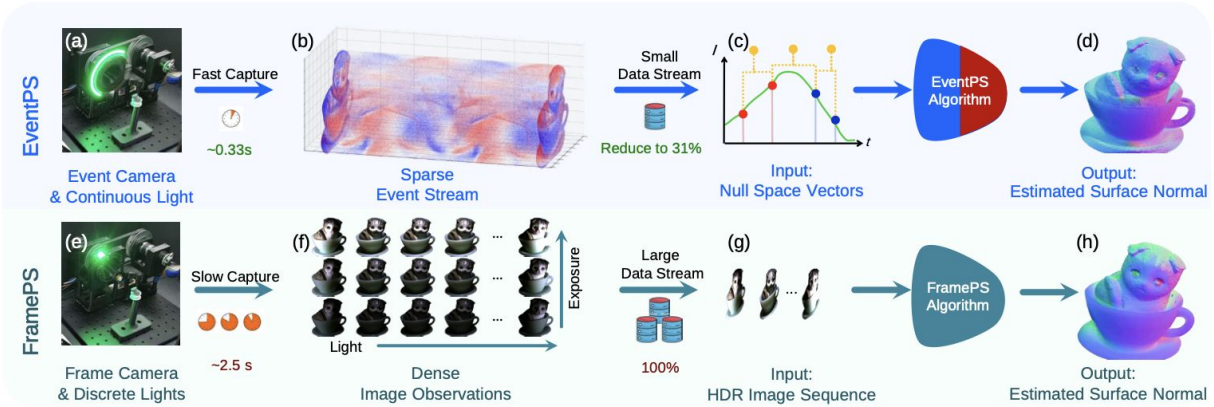
[2] National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

[3] School of Mechanical Engineering, Shanghai Jiao Tong University

[4] Graduate School of Information Science and Technology, The University of Tokyo [5] National Institute of Informatics

{ybh1998, wangfeishi, cssherryliang, shiboxin}@pku.edu.cn

jiejiren@sjtu.edu.cn, jinhan@nii.ac.jp



Best Paper Runners-Up

pixelSplat: 3D Gaussian Splats from Image Pairs for Scalable Generalizable 3D Reconstruction

David Charatan[1] Sizhe Lester Li[1] Andrea Tagliasacchi[2] Vincent Sitzmann[1]
[1]Massachusetts Institute of Technology [2]Simon Fraser University, University of Toronto
{charatan, sizheli, sitzmann}@mit.edu andrea.tagliasacchi@sfu.ca

Abstract

We introduce pixelSplat, a feed-forward model that learns to reconstruct 3D radiance fields parameterized by 3D Gaussian primitives from pairs of images. Our model features real-time and memory-efficient rendering for scalable training as well as fast 3D reconstruction at inference time. To overcome local minima inherent to sparse and locally supported representations, we predict a dense probability distribution over 3D and sample Gaussian means from that probability distribution. We make this sampling operation differentiable via a reparameterization trick, allowing us to back-propagate gradients through the Gaussian splatting representation. We benchmark our method on wide-baseline novel view synthesis on the real-world RealEstate10k and ACID datasets, where we outperform state-of-the-art light

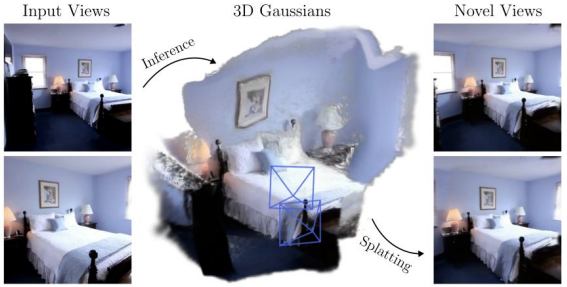


Figure 1. **Overview.** Given a pair of input images, pixelSplat reconstructs a 3D radiance field parameterized via 3D Gaussian primitives. This yields an explicit 3D representation that is renderable in real time, remains editable, and is cheap to train.

Best Papers

Generative Image Dynamics

Zhengqi Li Richard Tucker Noah Snavely Aleksander Holynski
Google Research

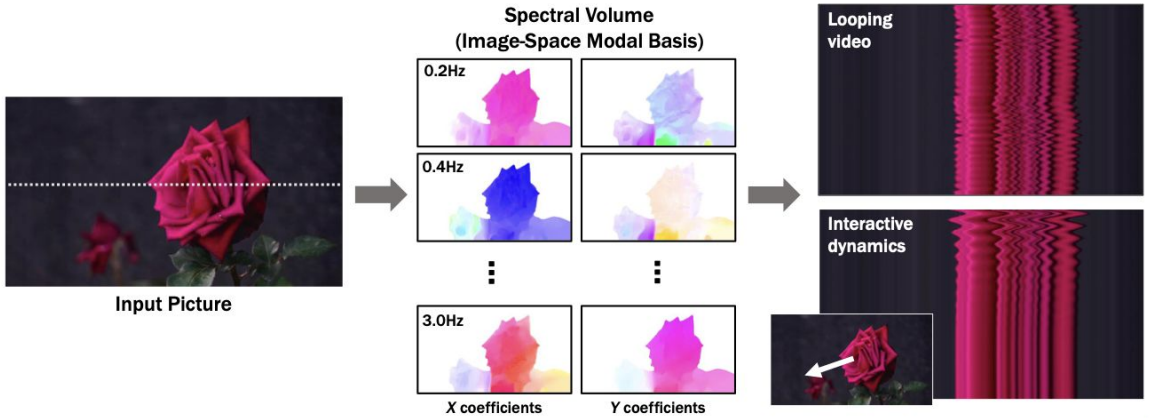


Figure 1. We model a generative image-space prior on scene motion: from a single RGB image, our method generates a *spectral volume* [23], a motion representation that models dense, long-term pixel trajectories in the Fourier domain. Our learned motion priors can be used to turn a single picture into a seamlessly looping video, or into an interactive simulation of dynamics that responds to user inputs like dragging and

Best Papers

Rich Human Feedback for Text-to-Image Generation

Youwei Liang[*†1], Junfeng He[*‡2], Gang Li[*‡2], Peizhao Li[†5], Arseniy Klimovskiy[2], Nicholas Carolan[2], Jiao Sun[†§3], Jordi Pont-Tuset[2], Sarah Young[2], Feng Yang[2], Junjie Ke[2], Krishnamurthy Dj Dvijotham[2], Katherine M. Collins[†4], Yiwen Luo[2], Yang Li[2], Kai J Kohlhoff[2], Deepak Ramachandran[2], and Vidhya Navalpakkam[2]

[1]University of California, San Diego
[2]Google Research
[3]University of Southern California
[4]University of Cambridge
[5]Brandeis University

Abstract

Recent Text-to-Image (T2I) generation models such as Stable Diffusion and Imagen have made significant progress in generating high-resolution images based on text descriptions. However, many generated images still suffer from issues such as artifacts/implausibility, misalignment with text descriptions, and low aesthetic quality. Inspired by the

in our GitHub repository: https://github.com/google-research/google-research/tree/master/richhf_18k.

1. Introduction

Text-to-image (T2I) generation models [12, 17, 41, 42, 56, 58, 59] are rapidly becoming a key to content creation in

PAMI TC Award Process

- PAMITC Awards committee: R. Zabih & T. Tuytelaars (chairs), D. Forsyth, W. Freeman, R. Hartley, T. Kanade, J. Malik, G. Medioni, C. Schmid, A. Zisserman
- Conflicts of interest policy adapted from NSF guidelines
- Voting system allows mild conflicts not to count positively or negatively

2024 PAMI TC Awards Committees

Longuet-Higgins Prize

R. Hartley (chair), P. Perona, J. Shi

Young Researcher Award

R. Zabih & T. Tuytelaars (chairs), K. Kutulakos, C. Liu, N. Paragios, M. Pollefeys

Thomas Huang Memorial Prize

G. Medioni (chair), M. Black, R. Chellappa, A. Fitzgibbon, D. Fleet, E. Grimson, R. Hartley, K.M. Lee, C. Schmid, H. Shi

Longuet-Higgins Prize

Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation

Ross Girshick, Jeff Donahue, Trevor Darrell and Jitendra Malik

Young Researcher Award (Honorable Mention)

Katie Bouman

Young Researcher Award

Angjoo Kanazawa

Carl Vondrick

Thomas Huang Memorial Prize

Andrea Vedaldi

IEEE COMPUTER SOCIETY

CVPR SEATTLE, WA JUNE 17-21, 2024

CVF