

CVPR 2026 Workshop

Visual General Intelligence: Vision Research Toward the AGI Era

Closing Remarks

Yoshihiro Fukuhara

[AIST](#) / [CADDi](#)

June 3, 2026

Colorado Convention Center - Room 703

CVPR
JUNE 3-7, 2026



DENVER
COLORADO



LIMIT.LAB

Gratitude to Our Keynote Speakers



Robert Geirhos
Google DeepMind



Aditi Raghunathan
Carnegie Mellon University



Matt Deitke
Meta Superintelligence Lab



Kristen Grauman
University of Texas, Austin / Meta



Yuki M. Asano
University of Technology Nuremberg



Andrea Vedaldi
University of Oxford / Meta



Alexei A. Efros
University of California, Berkeley



Alex Kendall
Wayve

Gratitude to Our Invited Poster Presenters

- **Theory of Space: Can Foundation Models Construct Spatial Beliefs through Active Exploration?**

Pingyue Zhang, Zihan Huang, Yue Wang, Jieyu Zhang, Letian Xue, Zihan Wang, Qineng Wang, Keshigeyan Chandrasegaran, Ruohan Zhang, Yejin Choi, Ranjay Krishna, Jiajun Wu, Li Fei-Fei, Manling Li

- **Unique Lives, Shared World: Learning from Single-Life Videos**

Tengda Han, Sayna Ebrahimi, Dilara Gokay, Li Yang Ku, Maks Ovsjanikov, Iva Babukova, Daniel Zoran, Viorica Patraucean, Joao Carreira, Andrew Zisserman, Dima Damen

- **Learning to See Through a Baby's Eyes: Early Visual Diets Enable Robust Visual Intelligence in Humans and Machines**

Yusen Cai, Qing Lin, Bhargava Satya Nunna, Mengmi Zhang

- **From None to All: Self-Supervised 3D Reconstruction via Novel View Synthesis**

Ranran Huang, Weixun Luo, Ye Mao, Krystian Mikolajczyk

- **Efficiently Reconstructing Dynamic Scenes One D4RT at a Time**

Chuhan Zhang, Guillaume Le Moing, Skanda Koppula, Ignacio Rocco, Liliane Momeni, Junyu Xie, Shuyang Sun, Rahul Sukthankar, Joëlle K. Barral, Raia Hadsell, Zoubin Ghahramani, Andrew Zisserman, Junlin Zhang, Mehdi S. M. Sajjadi

- **Seeing without Pixels: Perception from Camera Trajectories**

Zihui Xue, Kristen Grauman, Dima Damen, Andrew Zisserman, Tengda Han

- **Next-Embedding Prediction Makes Strong Vision Learners**

Sihan Xu, Ziqiao Ma, Wenhao Chai, Xuweiyi Chen, Weiyang Jin, Joyce Chai, Saining Xie, Stella X. Yu

- **4D-RGPT: Toward Region-level 4D Understanding via Perceptual Distillation**

Chiao-An Yang, Ryo Hachiuma, Sifei Liu, Subhashree Radhakrishnan, Raymond A. Yeh, Yu-Chiang Frank Wang, Min-Hung Chen

- **Steerable Visual Representations**

Jona Ruthardt, Manu Gaur, Deva Ramanan, Makarand Tapaswi, Yuki M. Asano

- **VGGT- Ω**

Jianyuan Wang, Minghao Chen, Shangzhan Zhang, Nikita Karaev, Johannes Schönberger, Patrick Labatut, Piotr Bojanowski, David Novotny, Andrea Vedaldi, Christian Rupprecht

- **A Frame is Worth One Token: Efficient Generative World Modeling with Delta Tokens**

Tommie Kerssies, Gabriele Berton, Ju He, Qihang Yu, Wufei Ma, Daan de Geus, Gijs Dubbelman, Liang-Chieh Chen

- **Visual General Intelligence: Vision Research Toward the AGI Era**

Hirokatsu Kataoka et al.

- **3D sans 3D Scans: Scalable Pre-training from Video-Generated Point Clouds**

Ryosuke Yamada, Kohsuke Ide, Yoshihiro Fukuhara, Hirokatsu Kataoka, Gilles Puy, Andrei Bursuc, Yuki M. Asano

- **COG: Confidence-aware Optimal Geometric Correspondence for Unsupervised Single-reference Novel Object Pose Estimation**

Yuchen Che, Jingtu Wu, Hao Zheng, Asako Kanezaki

- **Towards Structured Recognition and Generation of Human-Object Interactions**

Takuma Yagi, Kohei Torimi, Jyun-Ting Song, Kris Kitani

- **BabyVLM-V2: Toward Developmentally Grounded Pretraining and Benchmarking of Vision Foundation Models**

Shengao Wang, Wenqi Wang, Zecheng Wang, Max Whetton, Michael Wakeham, Arjun Chandra, Joey Huang, Pengyue Zhu, Helen Chen, David Li, Jeffrey Li, Shawn Li, Andrew Zagula, Amy Zhao, Andrew Zhu, Sayaka Nakamura, Yuki Yamamoto, Jerry Jun Yokono, Aaron Mueller, Bryan A. Plummer, Kate Saenko, Venkatesh Saligrama, Boqing Gong

- **Planning in 8 Tokens: A Compact Discrete Tokenizer for Latent World Model**

Dongwon Kim, Gawon Seo, Jinsung Lee, Minsu Cho, Suha Kwak

- **E-RayZer: Self-supervised 3D Reconstruction as Spatial Visual Pre-training**

Qitao Zhao, Hao Tan, Qianqian Wang, Sai Bi, Kai Zhang, Kalyan Sunkavalli, Shubham Tulsiani, Hanwen Jiang

Gratitude to Our Sponsors and Supporters

 **SB Intuitions**

<https://www.sbintuitions.co.jp/en/>



ASPIRE

Where we are now →

<https://www.jst.go.jp/aspire/en/index.html>

Closing Remarks

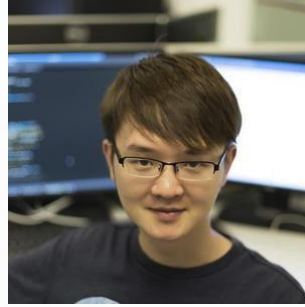
Special Thanks to the Organizing Team



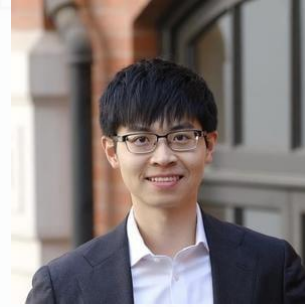
Hirokatsu Kataoka
AIST / Oxford VGG



Yoshihiro Fukuhara
AIST / CADDi



Yonglong Tian
OpenAI



Elliott / Shangzhe Wu
University of Cambridge



Oishi Deb
Oxford VGG



Ryousuke Yamada
AIST / UTN



Christian Rupprecht
Oxford VGG



Jianyuan Wang
Oxford VGG



Kohsuke Ide
University of Tsukuba



Xianzheng Ma
Oxford VGG



Koichi Namekata
Oxford VGG



Yiming Chen
Oxford VGG



Tetsu Furukawa
AIST Solutions



Akihiro Fuji
AIST Solutions



Shoko Ito
AIST Solutions



Shinji Yokoi
AIST Solutions

Upcoming: The VGI White Paper

[Tentative title] Visual General Intelligence: A White Paper

Hirokatsu Kataoka^{1,2} Yoshihiro Fukuhara¹ Yonglong Tian³ Shangzhe Wu⁴ Oishi Deb²
Ryousuke Yamada^{1,5} Christian Rupprecht² Jianyuan Wang² Kohsuke Ide^{1,6} Xianzheng Ma²
Koichi Namekata² Yiming Chen²

+Further Inviting Special Authors

¹ National Institute of Advanced Industrial Science and Technology (AIST)

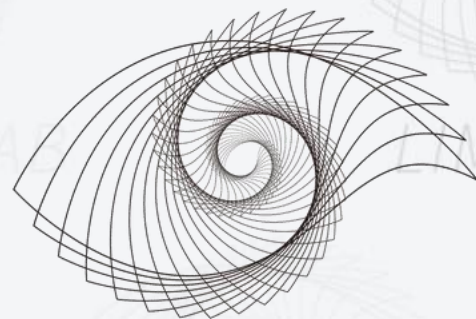
² Visual Geometry Group (VGG), University of Oxford ³ OpenAI ⁴ University of Cambridge

⁵ University of Technology Nuremberg ⁶ University of Tsukuba

The background of the slide features a repeating pattern of stylized, concentric spiral eyes in a light gray color. Each eye is composed of many thin, curved lines that form a spiral shape, resembling a nautilus shell or a stylized eye. The text "LIMIT.LAB" is repeated in a light gray, sans-serif font across the background, positioned below each eye.

Thank You All!

Enjoy the reset of CVPR 2026



LIMIT.LAB