Yuta Oshima

The University of Tokyo | Tokyo, Japan

(+81) 90-3523-0449 | yuta.oshima@weblab.t.u-tokyo.ac.jp | shim0114.github.io | Google Scholar

Education

Apr. 2025 – Present	Ph.D. in Computer Science at **The University of Tokyo** *Advisor: Prof. Yutaka Matsuo*
Apr. 2023 – Mar. 2025	M.S. in Computer Science at **The University of Tokyo**
Apr. 2019 – Mar. 2023	B.S. in Engineering at **The University of Tokyo**

Selected Publications

[1] Yuta Oshima, Daiki Miyake, Kohsei Matsutani, Yusuke Iwasawa, Masahiro Suzuki, Yutaka Matsuo, Hiroki Furuta. *MultiBanana: A Challenging Benchmark for Multi-Reference Text-to-Image Generation.* **CVPR 2026 Main**. [arXiv] [Code] [HuggingFace]

[2] Yuta Oshima, Masahiro Suzuki, Yutaka Matsuo, Hiroki Furuta. *Inference-Time Text-to-Video Alignment with Diffusion Latent Beam Search.* **NeurIPS 2025**. [arXiv] [OpenReview] [Code] [Project]

[3] Shohei Taniguchi, Keno Harada, Gouki Minegishi, Yuta Oshima, Seong Cheol Jeong, Go Nagahara, Tomoshi Iiyama, Masahiro Suzuki, Yusuke Iwasawa, Yutaka Matsuo. *ADOPT: Modified Adam Can Converge with Any β_2 with the Optimal Rate.* **NeurIPS 2024**. [arXiv] [OpenReview] [Code]

[4] Yuta Oshima, Masahiro Suzuki, Yutaka Matsuo. *Enhancing Unimodal Latent Representations in Multimodal VAEs through Iterative Amortized Inference.* **New Generation Computing**. [DOI] [arXiv]

[5] Yuta Oshima, Yusuke Iwasawa, Masahiro Suzuki, Yutaka Matsuo, Hiroki Furuta. *WorldPack: Compressed Memory Improves Spatial Consistency in Video World Modeling.* **CVPR 2026 Workshop**. [arXiv]

[6] Riku Takahashi, Rongzhi Li, Yuta Oshima, Sho Kuno, Ryugo Morita, Issey Sukeda. *AKITalk: Audio-Implicit Keypoints for Identity-Preserving Talking-Head Video Synthesis.* **ACM MMAsia 2025 Workshop**. **3rd Place** in The ACM Multimedia Asia 2025 Grand Challenge. [DOI]

[7] Yuta Oshima, Shohei Taniguchi, Masahiro Suzuki, Yutaka Matsuo. *SSM Meets Video Diffusion Models: Efficient Video Generation with Structured State Spaces.* **ICLR 2024 Workshop**. [arXiv] [OpenReview] [Code]

[8] Tatsuya Kamijo, Tomoshi Iiyama, Yuta Oshima, Gentiane Venture, Tatsuya Matsushima, Yutaka Matsuo, Yusuke Iwasawa. *Tactile In-Hand Pose Estimation through Perceptual Inference.* **IROS 2023 Workshop**, **Spotlight**. [PDF]

Research Experience

Google DeepMind (Collaboration) Sep. 2024 – Present

Research Collaborator, working with Dr. Hiroki Furuta (Senior Research Scientist, Google DeepMind). Formally supported through the UTokyo–Google AI Symbiotic Future Society Program (Sep. 2025 – Aug. 2026).

Matsuo–Iwasawa Lab, The University of Tokyo Dec. 2023 – Present

Ph.D. / M.S. Student Researcher, advised by Prof. Yutaka Matsuo. Funded through scholarship programs for Ph.D. students.

Industry Experience

AI Researcher, EQUES Inc. Jul. 2025 – May 2026

Vision generation for anime production within the GENIAC project, supported by Japan's Ministry of Economy, Trade and Industry (METI).

Academic Honors

Awards

2025 3rd Place, The ACMMM Asia 2025 Grand Challenge: Multimodal Multiethnic Talking-Head Video Generation

2024 Outstanding Presentation Award, JSAI Annual Conference

2023 Spotlight Presentation, IROS 2023 Workshop on World Models and Predictive Coding in Cognitive Robotics

Grants

2026 Overseas Travel Grant, The Telecommunications Advancement Foundation (TAF)

2025 Research Funding, Leadership Development Program for Ph.D. (LDPP), The University of Tokyo

Academic Service

Conference Reviewer: NeurIPS 2026, ICML 2026, ICLR 2026, ECCV 2026, CVPR 2026