

Akio Hayakawa

Tokyo, Japan | Email: akio080314 [at] gmail.com, akio.hayakawa [at] sony.com
Website: ahykw.github.io | Github: github.com/ahykw | [Google Scholar](#)

Education

M.S. in Information Science and Technology | The University of Tokyo

Apr 2016 – Mar 2018

B.S. in Engineering | The University of Tokyo

Apr 2012 – Mar 2016

Selected Publications

Linear Fusion MultiDiffusion for Fast Training-Free Spherical Panorama Generation

ECCV 2026

Akio Hayakawa, Yusuke Mukuta, Tatsuya Harada

First author. Training-free panorama generation formulated as regularized least squares with a Krylov iterative solver; reported 15.36x faster inference than the strongest training-free baseline.

Step-by-Step Video-to-Audio Synthesis via Negative Audio Guidance

ECCV 2026

Akio Hayakawa, Masato Ishii, Takashi Shibuya, Yuki Mitsufuji

First author. A step-by-step video-to-audio framework that suppresses previously generated sounds through negative audio guidance.

Echoes Over Time: Unlocking Length Generalization in Video-to-Audio Generation Models

CVPR 2026

Christian Simon, Masato Ishii, Wei-Yao Wang, Koichi Saito, **Akio Hayakawa**, et al.

Research on length generalization for video-to-audio generation.

TITAN-Guide: Taming Inference-Time Alignment for Guided Text-to-Video Diffusion Models

ICCV 2025

Christian Simon, Masato Ishii, **Akio Hayakawa**, et al.

Inference-time alignment for guided text-to-video diffusion models.

Taming Multimodal Joint Training for High-Quality Video-to-Audio Synthesis

CVPR 2025

Ho Kei Cheng, Masato Ishii, **Akio Hayakawa**, et al.

Multimodal joint training for high-quality video-to-audio synthesis.

MMDisCo: Multi-Modal Discriminator-Guided Cooperative Diffusion for Joint Audio and Video Generation

ICLR 2025

Akio Hayakawa, Masato Ishii, Takashi Shibuya, Yuki Mitsufuji

First author. Cooperative diffusion framework for joint audio and video generation using multimodal discriminator guidance.

Additional Publications

- [A Simple but Strong Baseline for Sounding Video Generation: Effective Adaptation of Audio and Video Diffusion Models for Joint Generation](#) IJCNN 2025
- [Reference-based Video Colorization with Spatiotemporal Correspondence](#) MIRU 2021
- [Out-of-core Training for Extremely Large-scale Neural Networks with Adaptive Window-based Scheduling](#) MIRU 2021
- [Fine-grained Image Editing by Pixel-wise Guidance Using Diffusion Models](#) CVPRW 2023
- [Instruct 3D-to-3D: Text Instruction Guided 3D-to-3D Conversion](#) MIRU 2023
- [AutoRefiner: Improving Autoregressive Video Diffusion Models via Reflective Refinement Over the Stochastic Sampling Path](#) arXiv 2025
- [Coherent Audio-Visual Editing via Conditional Audio Generation Following Video Edits](#) arXiv 2025
- [SoundReactor: Frame-level Online Video-to-Audio Generation](#) arXiv 2025
- [Neural Network Libraries: A Deep Learning Framework Designed from Engineers' Perspectives](#) arXiv 2021

Work Experience

Research Scientist | Sony Research Inc.

Apr 2023 – Present

- Conduct research on multimodal generative AI, including audio-video generation, video-to-audio synthesis, video diffusion, and controllable generation.
- Publish and collaborate on research presented at ICLR, CVPR, ICCV, and ECCV.

Research Engineer | R&D Center, Sony Corporation

Apr 2018 – Mar 2023

- Worked on machine learning and generative modeling research and development.
- Contributed to research on diffusion-based image editing, large-scale neural network training, and computer vision.
- Served as a core contributor to [Neural Network Libraries \(NNabla\)](#), with contributions spanning the Python API, C++

execution backend, CUDA kernel and runtime optimization, CPU/GPU memory management, and the implementation of distributed training infrastructure for large-scale neural networks.

Honors

- Selected as a Mitou Creator by the Information-technology Promotion Agency (IPA), Japan, 2017.

Invited talks

- “Tutorial on Diffusion Models,” The 29th Symposium on Sensing via Image Information (SSII), 2023; presentation received more than 15,000 views.
- “Tutorial on Diffusion Models and Recent Trends,” The 37th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI), 2023.

Professional Activities

- Conference reviewer for CVPR, ECCV, ICLR, and NeurIPS.
- Recognized as a Top 200 Reviewer at ICLR 2026.
- Recognized as an [Outstanding Reviewer at ECCV 2026](#).
- Executive co-chair for the 29th Meeting on Image Recognition and Understanding (MIRU2026).

Patent

- “Bias adjustment device, information processing device, information processing method, and information processing program,” US Patent Application 17/771,051, 2022. Co-inventor.

Links

Website: <https://ahykw.github.io/>

Google Scholar: <https://scholar.google.com/citations?user=sXAJHFIAAAAJ>

GitHub: <https://github.com/ahykw>

Sony GitHub: <https://github.com/AkioHayakawa-sony>