Philipp Krähenbühl

EDUCATION

2009–2014 **PhD, Computer Science**, *Stanford University*, Stanford CA
2007–2008 **Student Exchange, Computer Science**, *University of Washington*, Seattle WA
2005–2009 **BSc, Computer Science**, *ETH Zurich*, Zurich, Switzerland

RESEARCH AND INDUSTRY EXPERIENCE

2022– **Associate Professor**, *UT Austin*, Department of Computer Science
2022– **Researcher**, *Apple*, AIML
2016–2022 **Assistant Professor**, *UT Austin*, Department of Computer Science
2014–2016 **Postdoctoral Researcher**, *UC Berkeley*, with Alexei Efros, Jitendra Malik, and Trevor Darrell
2016–2021 **Researcher (part-time)**, *Intel*
2009–2014 **Research Assistant**, *Stanford University*, with Vladlen Koltun
2012,2013 **Research Internship**, *Adobe Creative Technologies Lab*
2011 **Research Internship**, *Disney Research Zurich*
2008 **Research Assistant**, *University of Washington*, with Zoran Popovic
2006–2009 **Research Assistant**, *ETH Zurich*, with Markus Gross, Gaston Gonnet and Bertrand Meyer

SERVICE AND AFFILIATIONS

Area Chair
International Conference on Computer Vision (ICCV): 2017, 2025
Computer Vision and Pattern Recognition (CVPR): 2018, 2019, 2020, 2021, 2023
Computer Vision and Pattern Recognition (CVPR), Senior Area Chair: 2024
European Conference on Computer Vision (ECCV): 2016, 2018
International Conference on Learning Representations (ICLR): 2021, 2024, 2026
Conference on Neural Information Processing Systems (NeurIPS): 2021, 2022, 2023, 2025
Conference on Neural Information Processing Systems (NeurIPS), Evaluations and Datasets Track: 2026
International Conference on Machine Learning (ICML): 2025

COLLABORATORS & OTHER AFFILIATIONS

PhD Advisor: Koltun, Vladlen; Stanford

Postdoctoral Sponsors: Darrell, Trevor; UC-Berkeley
Efros, Alexei; UC-Berkeley
Malik, Jitendra; UC-Berkeley

HONORS AND AWARDS

GDC 4.824, 2317 Speedway – Austin, TX 78712 – USA
philkr@cs.utexas.edu • *www.philkr.net* • *Google Scholar*

2011 **NIPS Outstanding Student Paper Award**
Highest award at the conference, given to 3 out of 1400 submissions.

2010-2014 **Stanford Graduate Fellowship**
Pierre and Christine Lamond Fellow

GRANTS

2020-2025 **NSF AI Institute for Foundations of Machine Learning**
$\$20,000,000$; Senior Member (CO-PI)

2020-2023 **NSF SMALL: Towards inferring all properties of all things through time**
$\$420,000$; PI

2019-2026 **NSF CAREER: A Compression-Based Approach to Learning Video Representations**
$\$500,000$; PI

2016- **Industrial gifts: Intel, Berkeley Deep Drive**
$\$338,000$; PI

PUBLICATIONS

[1] Haran Raajesh, Kulin Shah, Adam Klivans, and Philipp Krähenbühl. Mask-aware policy gradients for diffusion language models. In *COLM*, 2026.

[2] Yue Zhao, Hanwen Jiang, Zhenlin Xu, Chutong Yang, Ehsan Adeli, and Philipp Krähenbühl. Spherical leech quantization for visual tokenization and generation. In *CVPR*, 2026.

[3] Shuhan Tan, Kashyap Chitta, Yuxiao Chen, Ran Tian, Yurong You, Yan Wang, Wenjie Luo, Yulong Cao, Philipp Krähenbühl, Marco Pavone, and Boris Ivanovic. Latent chain-of-thought world modeling for end-to-end driving. In *CVPR*, 2026.

[4] Aleksei Petrenko, Ben Lipkin, Kevin Chen, Erik Wijmans, Marco Cusumano-Towner, Raja Giryes, and Philipp Krähenbühl. Entropy-preserving reinforcement learning. In *ICLR*, 2026.

[5] Jeffrey Ouyang-Zhang, Pranav Murugan, Daniel J. Diaz, Gianluca Scarpellini, Richard Strong Bowen, Nate Gruver, Adam Klivans, Philipp Krähenbühl, Aleksandra Faust, and Maruan Al-Shedivat. Triangle multiplication is all you need for biomolecular structure representations. In *ICLR*, 2026.

[6] Jang Hyun Cho, Andrea Madotto, Effrosyni Mavroudi, Triantafyllos Afouras, Tushar Nagarajan, Muhammad Maaz, Yale Song, Tengyu Ma, Shuming Hu, Suyog Jain, Miguel Martin, Huiyu Wang, Hanoona Rasheed, Peize Sun, Po-Yao Huang, Daniel Bolya, Nikhila Ravi, Shashank Jain, Tammy Stark, Shane Moon, Babak Damavandi, Vivian Lee, Andrew Westbury, Salman Khan, Philipp Krähenbühl, Piotr Dollár, Lorenzo Torresani, Kristen Grauman, and Christoph Feichtenhofer. Perceptionlm: Open-access data and models for detailed visual understanding. In *NeurIPS*, 2025.

[7] Xiuyu Yang, Shuhan Tan, and Philipp Krähenbühl. Long-term traffic simulation with interleaved autoregressive motion and scenario generation. In *ICCV*, 2025.

[8] Marco Cusumano-Towner, David Hafner, Alex Hertzberg, Brody Huval, Aleksei Petrenko, Eugene Vinitsky, Erik Wijmans, Taylor Killian, Stuart Bowers, Ozan Sener, Philipp Krähenbühl, and Vladlen Koltun. Robust autonomy emerges from self-play. In *ICML*, 2025.

[9] Erik Wijmans, Brody Huval, Alexander Hertzberg, Vladlen Koltun, and Philipp Krähenbühl. Cut your losses in large-vocabulary language models. In *ICLR*, 2025.

[10] Jeffrey Ouyang-Zhang, Chengyue Gong, Yue Zhao, Philipp Krähenbühl, Adam R Klivans, and Daniel J Diaz. Distilling structural representations into protein sequence models. In *ICLR*, 2025.

[11] Santhosh Kumar Ramakrishnan, Erik Wijmans, Philipp Kraehenbuehl, and Vladlen Koltun. Does spatial cognition emerge in frontier models? In *ICLR*, 2025.

[12] Yue Zhao, Yuanjun Xiong, and Philipp Krähenbühl. Image and video tokenization with binary spherical quantization. In *ICLR*, 2025.

[13] Jang Hyun Cho, Boris Ivanovic, Yulong Cao, Edward Schmerling, Yue Wang, Xinshuo Weng, Boyi Li, Yurong You, Philipp Krähenbühl, Yan Wang, and Marco Pavone. Language-image models with 3d understanding. In *ICLR*, 2025.

[14] Shuhan Tan, Boris Ivanovic, Yuxiao Chen, Boyi Li, Xinshuo Weng, Yulong Cao, Philipp Krähenbühl, and Marco Pavone. Promptable closed-loop traffic simulation. In *CoRL*, 2024.

[15] Yue Zhao, Long Zhao, Xingyi Zhou, Jialin Wu, Chun-Te Chu, Hui Miao, Florian Schroff, Hartwig Adam, Ting Liu, Boqing Gong, Philipp Krähenbühl, and Liangzhe Yuan. Distilling vision-language models on millions of videos. In *CVPR*, 2024.

[16] Jang Hyun Cho, Philipp Krähenbühl, and Vignesh Ramanathan. Partdistillation: Learning parts from instance segmentation. In *CVPR*, 2023.

[17] Yue Zhao, Ishan Misra, Philipp Krähenbühl, and Rohit Girdhar. Learning video representations from large language models. In *CVPR*, 2023.

[18] Jeffrey Ouyang-Zhang, Daniel J. Diaz, Adam R. Klivans, and Philipp Krähenbühl. Predicting a protein's stability under a million mutations. In *NeurIPS*, 2023.

[19] Shuhan Tan, Boris Ivanovic, Xinshuo Weng, Marco Pavone, and Philipp Kraehenbuehl. Language conditioned traffic generation. In *CoRL*, 2023.

[20] Jang Hyun Cho and Philipp Krähenbühl. Language-conditioned detection transformer. In *CVPR*, 2023.

[21] Yue Zhao and Philipp Krähenbühl. Real-time online video detection with temporal smoothing transformers. In *ECCV*, 2022.

[22] Jang Hyun Cho and Philipp Krähenbühl. Long-tail detection with effective class-margins. In *ECCV*, 2022.

[23] Xingyi Zhou, Rohit Girdhar, Armand Joulin, Philipp Krähenbühl, and Ishan Misra. Detecting twenty-thousand classes using image-level supervision. In *ECCV*, 2022.

[24] Dian Chen and Philipp Krähenbühl. Learning from all vehicles. In *CVPR*, 2022.

[25] Brady Zhou and Philipp Krähenbühl. Cross-view transformers for real-time map-view semantic segmentation. In *CVPR*, 2022.

[26] Xingyi Zhou, Tianwei Yin, Vladlen Koltun, and Philipp Krähenbühl. Global tracking transformers. In *CVPR*, 2022.

[27] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Simple multi-dataset detection. In *CVPR*, 2022.

[28] Tianwei Yin, Xingyi Zhou, and Philipp Krähenbühl. Multimodal virtual point 3d detection. In *NeurIPS*, 2021.

[29] Dian Chen, Vladlen Koltun, and Philipp Krähenbühl. Learning to drive from a world on rails. In *ICCV*, 2021.

[30] Chao-Yuan Wu and Philipp Krähenbühl. Towards long-form video understanding. In *CVPR*, 2021.

[31] Tianwei Yin, Xingyi Zhou, and Philipp Krähenbühl. Center-based 3d object detection and tracking. In *CVPR*, 2021.

[32] Aashaka Shah, Chao-Yuan Wu, Jayashree Mohan, Vijay Chidambaram, and Philipp Krähenbühl. Memory optimization for deep networks. In *ICLR*, 2021.

[33] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Probabilistic two-stage detection. In *arXiv*, 2021.

[34] Brady Zhou, Nimit Kalra, and Philipp Krähenbühl. Domain adaptation through task distillation. In *ECCV*, 2020.

[35] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Tracking objects as points. In *ECCV*, 2020.

[36] Chao-Yuan Wu, Ross Girshick, Kaiming He, Christoph Feichtenhofer, and Philipp Krähenbühl. A multigrid method for efficiently training video models. In *CVPR*, 2020.

[37] Dian Chen, Brady Zhou, Vladlen Koltun, and Philipp Krähenbühl. Learning by cheating. In *CORL*, 2019.

[38] Dequan Wang, Coline Devin, Qi-Zhi Cai, Philipp Krähenbühl, and Trevor Darrell. Monocular plan view networks for autonomous driving. In *IROS*, 2019.

[39] Xingyi Zhou, Dequan Wang, and Philipp Krähenbühl. Objects as points. In *arXiv preprint arXiv:1904.07850*, 2019.

[40] Chao-Yuan Wu, Christoph Feichtenhofer, Haoqi Fan, Kaiming He, Philipp Krähenbühl, and Ross Girshick. Long-term feature banks for detailed video understanding. In *CVPR*, 2019.

[41] Xingyi Zhou, Jiacheng Zhuo, and Philipp Krähenbühl. Bottom-up object detection by grouping extreme and center points. In *CVPR*, 2019.

[42] Hou-Ning Hu, Qi-Zhi Cai, Dequan Wang, Ji Lin, Min Sun, Philipp Krähenbühl, Trevor Darrell, and Fisher Yu. Joint monocular 3d vehicle detection and tracking. In *ICCV*, 2019.

[43] Brady Zhou, Philipp Krähenbühl, and Vladlen Koltun. Does computer vision matter for action? In *Science Robotics*, 2019.

[44] Brady Zhou and Philipp Krähenbühl. Don't let your discriminator be fooled. In *ICLR*, 2019.

[45] Chao-Yuan Wu, Nayan Singhal, and Philipp Krähenbühl. Video compression through image interpolation. In *ECCV*, 2018.

[46] Haoshuo Huang, Qixing Huang, and Philipp Krähenbühl. Domain transfer through deep activation matching. In *ECCV*, 2018.

[47] Chao-Yuan Wu, Manzil Zaheer, Hexiang Hu, R. Manmatha, Alexander J. Smola, and Philipp Krähenbühl. Compressed video action recognition. In *CVPR*, 2018.

[48] Philipp Krähenbühl. Free supervision from video games. In *CVPR*, 2018.

[49] Charles Packer, Katelyn Gao, Jernej Kos, Philipp Krähenbühl, Vladlen Koltun, and Dawn Song. Assessing generalization in deep reinforcement learning. In *arXiv*, 2018.

[50] Chao-Yuan Wu, R. Manmatha, Alexander J. Smola, and Philipp Krähenbühl. Sampling matters in deep embedding learning. In *ICCV*, 2017.

[51] Jeff Donahue, Philipp Krähenbühl, and Trevor Darrell. Adversarial feature learning. In *ICLR*, 2017.

[52] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alexei A. Efros. Generative visual manipulation on the natural image manifold. In *ECCV*, 2016.

[53] Deepak Pathak, Philipp Krähenbühl, Jeff Donahue, Trevor Darrell, and Alyosha Efros. Context encoders: Feature learning by inpainting. In *CVPR*, 2016.

[54] Tinghui Zhou, Philipp Krähenbühl, Mathieu Aubry, Qixing Huang, and Alyosha Efros. Learning dense correspondence via 3d-guided cycle consistency. In *CVPR*, 2016.

[55] Philipp Krähenbühl, Carl Doersch, Jeff Donahue, and Trevor Darrell. Data-dependent initializations of convolutional neural networks. In *ICLR*, 2016.

[56] Jun-Yan Zhu, Philipp Krähenbühl, Eli Shechtman, and Alyosha Efros. Learning a discriminative model for the perception of realism in composite images. In *ICCV*, 2015.

[57] Tinghui Zhou, Philipp Krähenbühl, and Alyosha Efros. Learning data-driven reflectance priors for intrinsic image decomposition. In *ICCV*, 2015.

[58] Deepak Pathak, Philipp Krähenbühl, and Trevor Darrell. Constrained convolutional neural networks for weakly supervised segmentation. In *ICCV*, 2015.

[59] Philipp Krähenbühl and Vladlen Koltun. Learning to propose objects. In *CVPR*, 2015.

[60] Philipp Krähenbühl and Vladlen Koltun. Geodesic object proposals. In *ECCV*, 2014.

[61] Philipp Krähenbühl. Dense random fields. In *Phd Thesis*, 2014.

[62] Philipp Krähenbühl and Vladlen Koltun. Parameter learning and convergent inference for dense random fields. In *ICML*, 2013.

[63] Philipp Krähenbühl and Vladlen Koltun. Efficient nonlocal regularization for optical flow. In *ECCV*, 2012.

[64] Federico Perazzi, Philipp Krähenbühl, Yael Pritch, and Alexander Hornung. Saliency filters: Contrast based filtering for salient region detection. In *CVPR*, 2012.

[65] Philipp Krähenbühl and Vladlen Koltun. Efficient inference in fully connected crfs with gaussian edge potentials. In *NIPS*, 2011.

[66] Sergey Levine, Philipp Krähenbühl, Sebastian Thrun, and Vladlen Koltun. Gesture controllers. In *SIGGRAPH*, 2010.

[67] Philipp Krähenbühl, Manuel Lang, Alexander Hornung, and Markus Gross. A system for retargeting of streaming video. In *SIGGRAPH Asia*, 2009.

[68] Philipp Krähenbühl. Art directable retargeting for streaming video. In *Bachelor Thesis, ETH Zurich*, 2009.

Complete list available at `www.philkr.net`.

Last CV update: September, 2026