Carlos Esteves

Research Scientist at Google

Education

University of Pennsylvania, *Philadelphia, PA, United States of America*.

2020 **Ph.D.**, *Computer and Information Science*, GRASP Laboratory.
Thesis: Learning Equivariant Representations.
Advisor: Prof. Kostas Daniilidis.
Committee: Prof. Jean Gallier, Prof. Jianbo Shi, Prof. Alejandro Ribeiro, Dr. Ameesh Makadia (Google Research).

Instituto Tecnológico de Aeronáutica (ITA) [Aeronautics Institute of Technology], *São José dos Campos, SP, Brazil*.

2010 **M.S.**, *Electronic and Computer Engineering*.
Thesis (transl.): *"Modeling, Control, and Flight Simulation of a Visually-Guided Autonomous Glider"*.
Advisor: Prof. Elder Moreira Hemerly.

2007 **B.S.**, *Computer Engineering*.
Admission rate: 1.4% (8,287 candidates for 120 positions).
All ITA undergraduate students are granted full scholarships, room and board.
Thesis (transl.): *"Autonomous Navigation and Obstacle Avoidance with Visual Feedback"*.
Advisors: Prof. Jackson Paul Matsuura and Dr. Christian Giorgio Roberto Taranti.

Selected publications

ECCV'18 IJCV'19 **Learning SO(3) Equivariant Representations with Spherical CNNs**, *C. Esteves, C. Allen-Blanchette, A. Makadia, K. Daniilidis*, European Conference on Computer Vision, ECCV'18 (oral; acceptance rate 2.4%). An extended version appeared in the IJCV "Best of ECCV'18" special edition.
Introduced spherical convolutional neural networks that learn rotation-equivariant feature maps by design via spherical convolutions evaluated in the spectral domain. Achieved superior results on 3D shape analysis tasks with inputs in arbitrary orientation.

machc@google.com • machc.github.io • machc
_machc • machc

ICCV'19 **Equivariant Multi-View Networks**, *C. Esteves, Y. Xu, C. Allen-Blanchette, K. Daniilidis*, International Conference on Computer Vision, ICCV'19 (oral; acceptance rate 4.6%).
Introduced an icosahedral group CNN that allows jointly reasoning about multiple views of an object or scene in an equivariant fashion. Achieved state of the art in several shape retrieval benchmarks and one panoramic scene classification benchmark.

CVPR'22 **Light Field Neural Rendering**, *M. Suhail, C. Esteves, L. Sigal, A. Makadia*, Conference on Computer Vision and Pattern Recognition, CVPR'22 (oral).
Conceptualized a model that leverages light field rendering ideas and geometric constraints to accurately synthesize novel views with view-dependent effects and sharp details. Achieved state-of-the-art performance on multiple popular datasets.

ICLR'18 **Polar Transformer Networks**, *C. Esteves, C. Allen-Blanchette, X. Zhou, K. Daniilidis*, International Conference on Learning Representations, ICLR'18 (top 3% review scores).
Proposed a differentiable polar transformer layer with learned origin, endowing CNNs with translation-invariance and continuous rotation/dilation equivariance.

NeurIPS'20 **Spin-Weighted Spherical CNNs**, *C. Esteves, A. Makadia, K. Daniilidis*, Conference on Neural Information Processing Systems, NeurIPS'20.
Defined convolution between spin-weighted spherical functions and built a CNN based on it. The model outperformed competing methods in several tasks and was also the first to handle spherical vector fields in an equivariant way.

ICML'21 **Implicit-PDF: Non-Parametric Representation of Probability Distributions on the Rotation Manifold**, *K. Murphy*, C. Esteves*, V. Jampani, S. Ramalingam, A. Makadia.*, International Conference on Machine Learning, ICML'21.
Introduced a general representation for object pose as a non-parametric probability distribution computed implicitly by a neural network. The method allows representation of intricate uncertainty and symmetry patterns that were not possible before.

ICML'19 **Cross-domain 3D Equivariant Image Embeddings**, *C. Esteves, A. Sud, Z. Luo, K. Daniilidis, A. Makadia*, International Conference on Machine Learning, ICML'19.
Introduced a map from images to a spherical feature space resulting in 3D equivariant embeddings with applications to relative pose estimation and novel view synthesis.

NeurIPS'20 **An Analysis of SVD for Deep Rotation Estimation**, *J. Levinson, C. Esteves, K. Chen, N. Snavely, A. Kanazawa, A. Rostamizadeh, A. Makadia*, Conference on Neural Information Processing Systems, NeurIPS'20.
Contributed in proofs about error expectations of SVD and Gram-Schmidt orthonormalization when estimating rotation matrices using deep regression.

Academic Service

2018- **Reviewer**, *Computer vision and machine learning conferences and journals*, NeurIPS, CVPR, ICML, ICLR, ICCV, ECCV, TPAMI, JMLR, AAAI.
ICLR 2022 Highlighted Reviewer.
CVPR 2019 Outstanding Reviewer.
CVPR 2021 Outstanding Reviewer.
ICML 2021 Expert Reviewer.

Outreach

2021- **Mentor**, *Google CS Research Mentorship Program*.
Mentored students from historically marginalized groups to support their pursuit of computing research pathways.

2021- **Mentor**, *LXAI Mentorship Program*.
Mentored early career professionals and students working on AI research.

Skills

Python, JAX, TensorFlow, Git, Linux, LaTeX, Emacs, Frequent use.

PyTorch, Theano, C, MATLAB, Subversion, Past frequent use.

Emacs Lisp, OpenCV, Bash, Occasional use.

C++, VHDL, Past occasional use.

Experience

2020– **Research Scientist**, *Google Research*, New York, NY.
- Conducting research on machine learning, computer vision, 3D vision and applications to quantum chemistry, drug design and climate science.

2019 **Research Intern**, *Facebook AI Research (FAIR)*, Menlo Park, CA.
- Designed model to reconstruct 3D meshes from a single occluded natural image.
- Introduced hypergraph convolutions and a self-intersection loss for high-quality mesh generation.
- Introduced multi-task feature sharing and adversarial training to improve robustness to occlusion and self-occlusion.

2018 **Student Researcher**, *Google Research/Machine Intelligence*, New York, NY.

- Improved and scaled spherical CNNs to allow applications to semantic segmentation (ECCV-W paper: "Labeling Panoramas with Spherical Hourglass Networks").
- Introduced spherical embeddings from single views of 3D objects for novel view synthesis and relative pose estimation (ICML paper: "Cross-Domain 3D Equivariant Image Embeddings").
- Designed a method for wide baseline cross-modality registration, with applications to camera pose estimation, scene completion, and unsupervised depth and surface normals from RGB.

2016–2020 **Teaching Assistant**, *University of Pennsylvania*, Philadelphia, PA.

- **Math for Machine Learning, MCIT-515** – graduate level. Foundations of linear algebra and optimization, covering both theoretical and algorithmic aspects.
- **Machine Perception, CIS-580 (2x)** – graduate level. Signal processing, Fourier analysis, projective geometry, structure from motion. Attributions: lectured, designed homeworks and projects.
- **Advanced Robotics, MEAM-620** – graduate level. Dynamics, control, planning, and vision. Students use the Crazyflie quadcopter with Vicon tracking feedback. Attributions: lectured, designed homeworks and exams, led lab sessions.
- **Robotics: Planning and Perception, CIS-390** – undergraduate level. Planning, estimation and perception; students develop for an iRobot Create controlled by a Raspberry Pi with April tag tracking. Attributions: lectured, designed and led lab sessions, designed homeworks and exams.

2008–2015 **Development Engineer/Researcher/Instructor/Advisor**, *Aeronautics and Space Institute – IAE, Aeronautics Institute of Technology – ITA (transl.)*, São José dos Campos, SP, Brazil.

- Designed and implemented simulation models and control systems, maintained embedded sofware as a development engineer on a joint program with South Africa.
- Designed cryptographic system; implemented aircraft bus simulator.
- Co-advised four undergraduate theses from the Computer, Electronic and Aeronautical Engineering departments. Two won best thesis awards from their class.
- Taught two courses in a full-time extension program offered to engineers.
- Published five articles in national symposiums.