Proceedings of the Workshop on 3D Geometry Generation for Scientific Computing

Marissa Ramirez de Chanlatte*
UC Berkeley
marissachanlatte@berkeley.edu

Phil Colella*
Lawrence Berkeley National Lab
pcolella@lbl.gov

Trevor Darrell*
UC Berkeley
trevordarrell@berkeley.edu

Alexandra Katherine Carlson
University of Michigan - Ann Arbor

Peter H.N. de With
Eindhoven University of Technology

Huayu Deng
Shanghai Jiao Tong University

Shanyan Guan
Shanghai Jiao Tong University

James Hays
Georgia Institute of Technology

Tim Houben
Eindhoven University of Technology

Thomas Huisman
ASML

Nikita Jaipuria
Ford Motor Company

Hans Johansen
Lawrence Berkeley National Lab

Shuja Khalid
University of Toronto

Akshay Krishnan
Georgia Institute of Technology

Chuming Li
University of Sydney

Maxim Pisarenco
ASML

Amit Raj
Google

Frank Rudzicz
Dalhousie University

Tim J. Schoonbeek
Eindhoven University of Technology

Sandhya Sridhar
Georgia Institute of Technology

Nathan Tseng
Ford Motor Company

Fons van der Sommen
Eindhoven University of Technology

Chen Wang
University of Pennsylvania

Yunbo Wang
Shanghai Jiao Tong University

Tong Wu
Amazon

Xiaokang Yang
Shanghai Jiao Tong University

Jiawei Yao
University of Washington

Derek Young
UC Davis

Xianling Zhang
Ford Greenfield Labs

Abstract

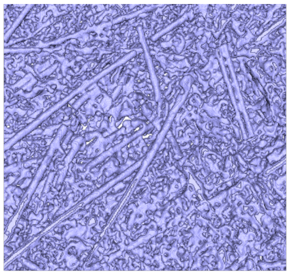
High-fidelity 3D geometries of the natural and built world around us are an essential part of answering some of the most pressing scientific questions of our day. Through advances in deep learning, computer vision, and artificial intelligence more broadly, much progress has been made in reconstructing real geometries from images and/or sparse data, but these methods are just beginning to be applied to scientific problems. On January 7th, 2024 we bring together researchers from computer vision, applied mathematics, and several scientific disciplines, to discuss the state of the art in 3D geometry generation and how it can be applied to open problems in science. This paper summarizes a selection of the talks and papers that have been accepted.

1. Introduction

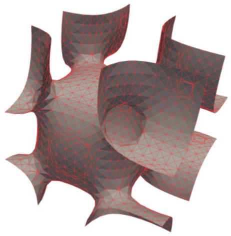
Answering many of the important scientific questions of our time - "How are ice caps going to melt?"; "How will extreme weather events affect our cities?"; and "How does a forest regrow after fire?" - all require a geometric understanding of the world around us. There have long been efforts to map the world, but current methods in AI & ML allow us to make 3D geometries of outdoor scenes with an ease and fidelity that has never been seen before. Rapid progress is being made in high quality, photo-realistic reconstruction of scenes from few images. Neural approaches to shape completion allows us to obtain dense point cloud representations even if only sparse measurement data is available. However, despite this amazing progress, state-of-the-art methods are not yet fully embraced by the scientific community. While Computer Vision & AI conferences are filled with impressive work on 3D geometry generation, what can be done with that work, particularly in regards to scientific progress, is under-explored. While almost all computational scientists use geometries in their work, the applied scientists who study how to actually build these specific geometries are scattered across disciplines and often isolated from the much larger general geometry generation community that exists in computer vision. This workshop brings together researchers from across AI, computer vision, scientific disciplines, and applied mathematics to discuss the state of the art in 3D geometry generation and the plethora of real-world problems and applications that can use these geometries. Below we include a selection of the abstracts and talks presented in the workshop.

2. 3D Implicit Functions and the Embedded Boundary Method

An Invited Talk by Hans Johansen



(a) An example "noisy" surface constructed from scans of a carbon-fiber matrix in hydrogen fuel cell applications, where surface representation is greatly facilitated by 3D computer-vision approaches.



(b) An example "water-tight" surface representation built of triangles for use in an embedded boundary scheme.

For science-focused computer simulations, there is an open debate as to what accuracy is required for machine learning-augmented simulations to be "correct." The question of minimum accuracy certainly depends on the context: as a preconditioner or initial guess, one might be able to get away with almost anything. However, as a partial differential equation (PDE) solver, if some important physical property is lacking, like conservation of mass in chemical reactions, results could be disastrously wrong. For example, high-resolution finite volume simulations require accurate, "water-tight" implicit-function definitions of complex geometries; these are a good fit for computer-vision based 3D reconstructions, as long as the required properties are enforced. We will present some of these sensitivities using an embedded boundary discretization and examples of "noisy" data that demonstrate the challenges and requirements for 3D AI-generated reconstructions.

3. Aerial imagery, photogrammetry, and multi-view computer vision for forest mapping at the individual tree scale

An Invited Talk by Derek Young: Forest ecology research and forest management planning often require detailed forest inventory data at the individual tree level. Using traditional ground-based manual survey methods, such data are time-consuming and costly to collect. Rapidly improving technologies including drones, photogrammetry, lidar, and computer vision are creating great potential for automated and low-cost individual tree mapping, and this potential could be enhanced by stronger collaborations between forest ecologists and computer vision scientists. This talk, from the perspective of a forest ecologist, describes some recent and ongoing work to map individual trees from

---

*Workshop organizers

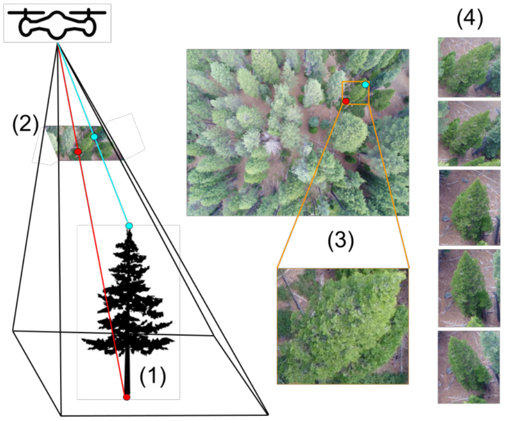


Figure 2. Workflow for programmatically locating individual detected trees within raw drone images for computer vision-based species classification. (1) Tree is detected using CHM, orthomosaic, and/or point cloud. (2-3) Tree location is projected from 3D space onto the drone image. (4) The process is carried out for every drone image in which the focal tree appears, enabling multi-view species classification.

aerial imagery, using the work of the Open Forest Observatory (www.openforestobservatory.org) as a case study. This work includes (a) tree detection from photogrammetric canopy height models using traditional algorithmic approaches; (b) tree detection from orthoimagery using computer vision object detection; and (c) tree species assignment using computer vision-based classification, including multi-view approaches that take advantage of multiple views (images) of each tree. The talk concludes by highlighting opportunities to improve forest mapping via collaborations between ecologists and computer vision scientists, with a key example being 3D computer vision (a) for detecting and characterizing trees from photogrammetrically derived structure models and (b) as an alternative to some key components of traditional photogrammetry.

4. RefiNeRF: Modelling dynamic neural radiance fields with inconsistent or missing camera parameters

A Poster by Shuja Khalid and Frank Rudzicz Novel view synthesis (NVS) is a challenging task in computer vision that involves synthesizing new views of a scene from a limited set of input images. Neural Radiance Fields (NeRF) have emerged as a powerful approach to address this problem, but they require accurate knowledge of camera *intrinsic* and *extrinsic* parameters. Traditionally, structure-from-motion (SfM) and multi-view stereo (MVS) approaches have been used to extract camera parameters, but these methods can be unreliable and may fail in certain cases. In this paper, we propose a novel technique that leverages unposed images from dynamic datasets, such as the NVIDIA dynamic scenes dataset, to learn camera parameters directly from data. Our approach is highly extensible and can be integrated into existing NeRF architectures with minimal modifications. We demonstrate the effectiveness of our method on a variety of static and dynamic scenes and show that it outperforms traditional SfM and SVM approaches. Our approach offers a promising new direction for improving the accuracy and robustness of NVS using NeRF, and we anticipate that it will be a valuable tool for a wide range of applications in computer vision and graphics.

5. Geometry-Guided Ray Augmentation for Neural Surface Reconstruction with Sparse Views

A Poster by Jiawei Yao, Chen Wang, Tong Wu, and Chuming Li

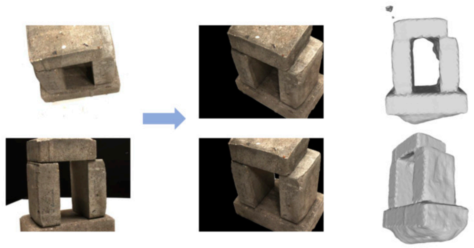
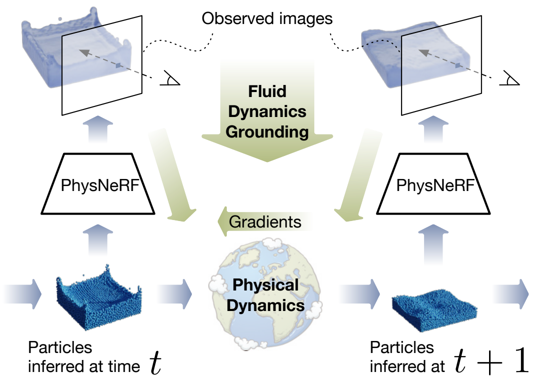


Figure 3. With a limited number of posed images (left), our proposed RayAug could produce high-quality novel view synthesis and accurate surface reconstruction when training on clean images (right). We reconstruct the meshes using the marching cubes algorithm.

In this paper, we propose a novel method for 3D scene and object reconstruction from sparse multi-view images. Different from previous methods that leverage extra information such as depth or generalizable features across scenes, our approach leverages the scene properties embedded in the multi-view inputs to create precise pseudo-labels for optimization without any prior training. Specifically, we introduce a geometry-guided approach that improves surface reconstruction accuracy from sparse views by leveraging spherical harmonics to predict the novel radiance while holistically considering all color observations for a point in the scene. Also, our pipeline exploits proxy geometry and correctly handles the occlusion in generating the pseudo-labels of radiance, which previous image-warping methods fail to avoid. Our method, dubbed Ray Augmentation (RayAug), achieves superior results on DTU and Blender datasets without requiring prior training, demonstrating its effectiveness in addressing the problem of sparse view reconstruction. Our pipeline is flexible and can be integrated into other implicit neural reconstruction methods for sparse views.

6. NeuroFluid: Fluid Dynamics Grounding with Particle-Driven Neural Radiance Fields

A Poster by Shanyan Guan, Huayu Deng, Yunbo Wang, and Xiaokang Yang



Deep learning has shown great potential for modeling the physical dynamics of complex particle systems such as fluids. Existing approaches, however, require the supervision of consecutive particle properties, including positions and velocities. We consider a partially observable scenario known as fluid dynamics grounding, that is, inferring the state transitions and interactions within the fluid particle systems from sequential visual observations of the fluid surface. We propose a differentiable two-stage network named NeuroFluid. Our approach consists of (i) a particle-driven neural renderer, which involves fluid physical properties in the volume rendering function, and (ii) a particle transition model optimized to reduce the differences between the rendered and the observed images. NeuroFluid provides the first solution to unsupervised learning of particle-based fluid dynamics by training these two models jointly. It is shown to reasonably estimate the underlying physics of fluids with different initial shapes, viscosity, and densities.

7. Surface Reconstruction of Semiconductor Devices from Scanning Electron Microscope Images using Simulated Geometries and Domain Adaptation

A Poster by Tim Houben, Thomas Huisman, Maxim Pisarenco, Tim J. Schoonbeek, Fons van der Sommen and Peter de With

Accurate metrology techniques for semiconductor devices are indispensable for controlling the manufacturing process. For instance, the dimensions of a transistor's current channel (displayed in Figure 3) are an important indicator of the device's performance regarding switching voltages and parasitic capacities. This work [1] expands upon traditional 2D analysis by utilizing computer vision techniques for 3D surface reconstruction. We propose a data-driven approach that predicts the surface map of a semiconductor device. Because of the absence of ground truth data we heavily rely on the creation of synthetic data, which is custom constructed for the specific application. Additionally, we introduce an unsupervised domain adaptation step to overcome the domain gap between simulated and experimental data. Our model is further fine-tuned [2] with a height measurement from a second scatterometry sensor and optimized through a tailored training scheme for optimal performance. During operation, the proposed method solely requires experimental images from a scanning electron microscope (SEM) of the patterns concerned and results in accurate depth predictions. Additionally, qualitative results indicate that detailed surface features, such as rounded corners, are realistically predicted, indicating that using the information learned from simulated data is of high value. This study shows that, when an algorithm is pretrained and calibrated adequately, accurate 3D metrology techniques can be viable for high-volume manufacturing.

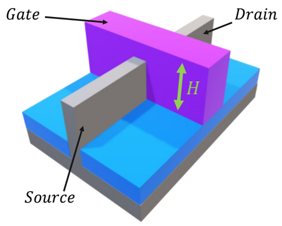


Figure 4. Example drawing of FinFET transistor on a wafer, indicating the main dimension of interest (H) in this work.

8. LANe: Lighting Aware Neural Fields for Compositional Scene Synthesis

A Poster by Akshay Krishnan, Amit Raj, Xianling Zhang, Alexandra Katherine Carlson, Nathan Tseng, Sandhya Sridhar, Nikita Jaipuria and James Hays

Recent advances in neural fields have enabled powerful representations and rendering of 3D scenes. However, most state-of-the-art implicit models capture static or dynamic scenes as a single entity, with limited ability to compose novel variations. We present Lighting-Aware Neural Fields (LANe), which enables compositional synthesis of real-world driving scenes in a physically consistent manner. Specifically, we disentangle the background and transient elements into separate world and object neural radi-

Figure 5. Examples of unshaded (left) and LANe shaded (right) object models composed onto the original positions of the car. The shader accounts for global changes in the scene lighting, darkening the cars in cloudy scenes and brightening them (with specularity changes) in sunny scenes. In scenes where the shading changes are subtle, regions on the car with most change have been highlighted.

ance fields. This allows for compositional rendering of multiple objects with consistent scene lighting. This is achieved by constructing a light field of the scene, which modulates object appearance, via a learned shader. We demonstrate LANe's advantages on synthetic and real-world driving datasets with varying lighting conditions. By learning object radiance fields from one scene and composing them into new backgrounds with distinct lighting, LANe outperforms prior work in challenging compositional synthesis tasks while maintaining lighting consistency. For additional results, project website can be found at: `https://lanerf.github.io/`.

References

[1] Tim Houben, Thomas Huisman, Maxim Pisarenco, Fons van der Sommen, and Peter de With. Training procedure for scanning electron microscope 3d surface reconstruction using unsupervised domain adaptation with simulated data. *Journal of Micro/Nanopatterning, Materials, and Metrology*, 22(3):031208–031208, 2023. 4

[2] Tim Houben, Thomas Huisman, Maxim Pisarenco, Fons van der Sommen, and Peter HN de With. Depth estimation from a single sem image using pixel-wise fine-tuning with multimodal data. *Machine Vision and Applications*, 33(4):56, 2022. 4