

Introduction to Neural Scene Representation and Neural Rendering

Lingjie Liu



We Live in a World that is 3D and Contains Dynamics



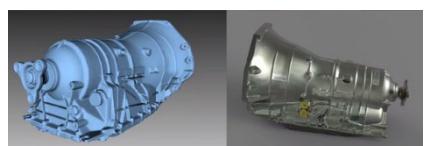
We Digitize Our World in 3D



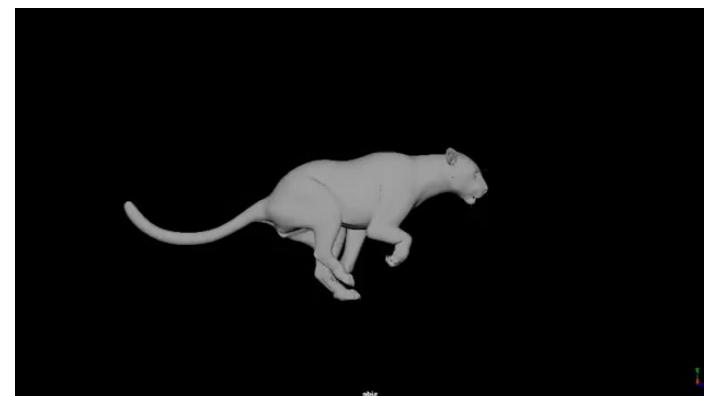
Future AI: Towards 3D Aware



3D Reconstruction of Real-world Scenes



Geometry
+ Appearance



Motion
+ Deformation

Photo-realistic Rendering

- **Image Synthesis** of Real-world Scenes with 3D Control.



Applications



AR / VR



Gaming / Movie



Healthcare



Autonomous Driving



Robot Grasping



Human-robot Interaction

Why are they challenging?

Problem formulation



Captured images →

Processing →

Rendering of real-world place



⋮



[Mildenhall et al., Neural Radiance Fields (NeRF), ECCV 2020]

[Wu et al., Scalable Neural Indoor Scene Rendering, SIGGRAPH 2022]

Classical Computer Graphics Pipeline

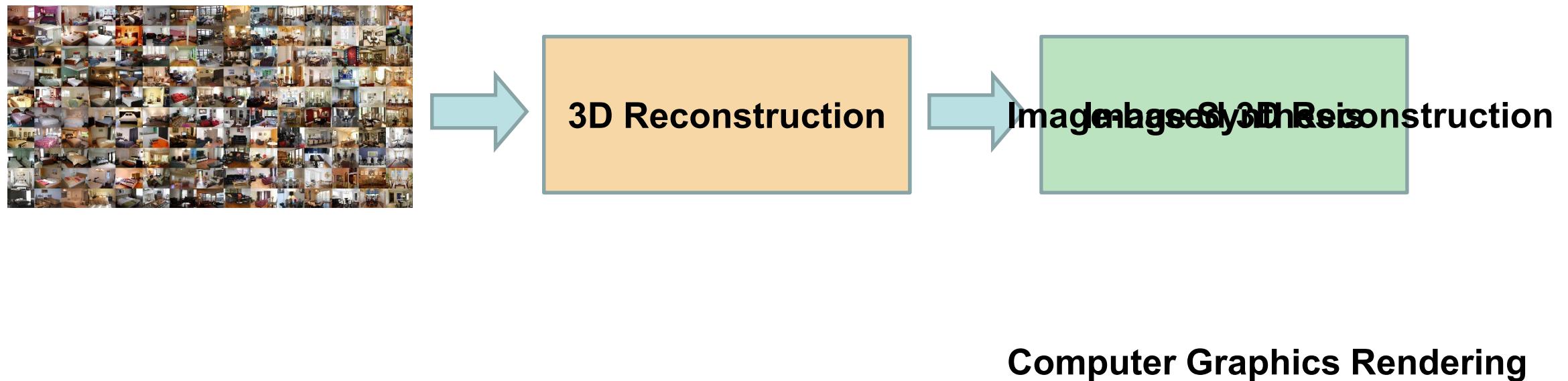
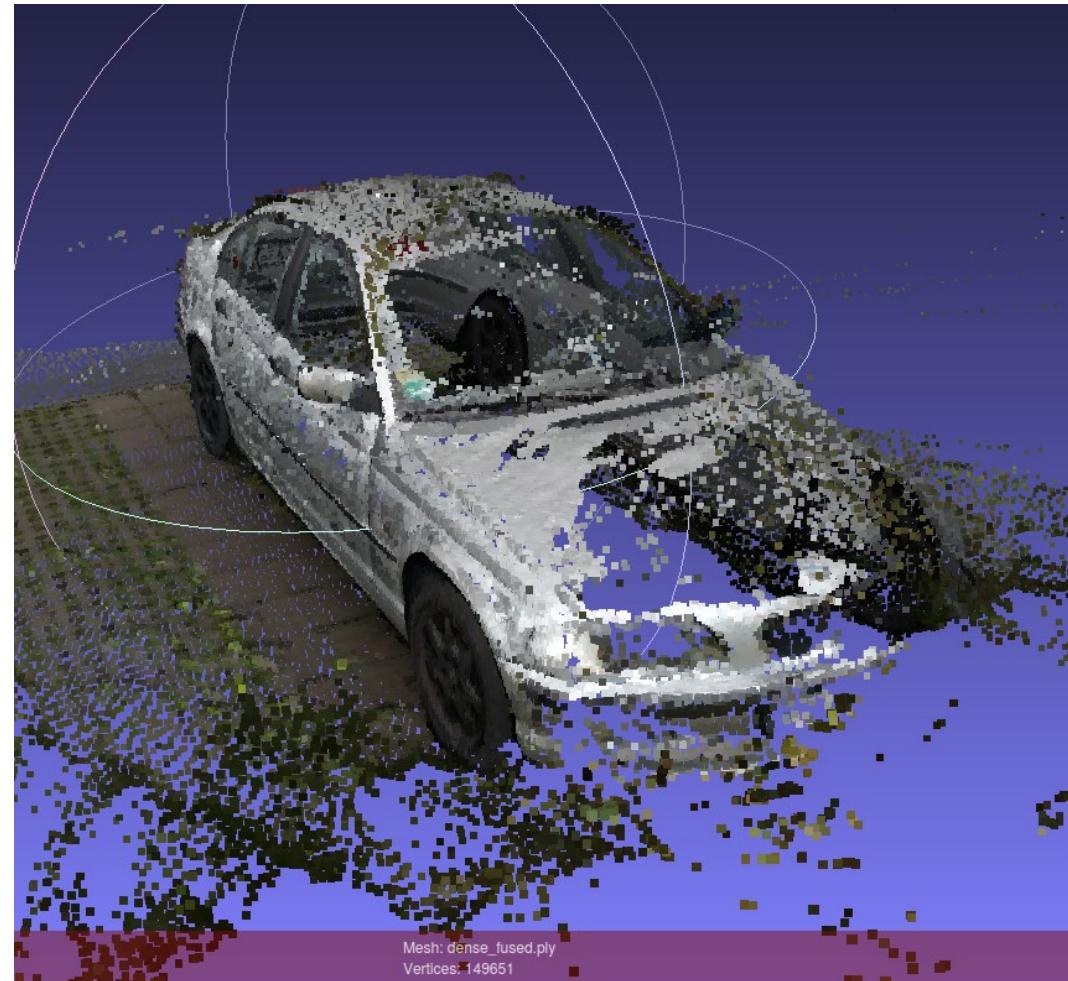
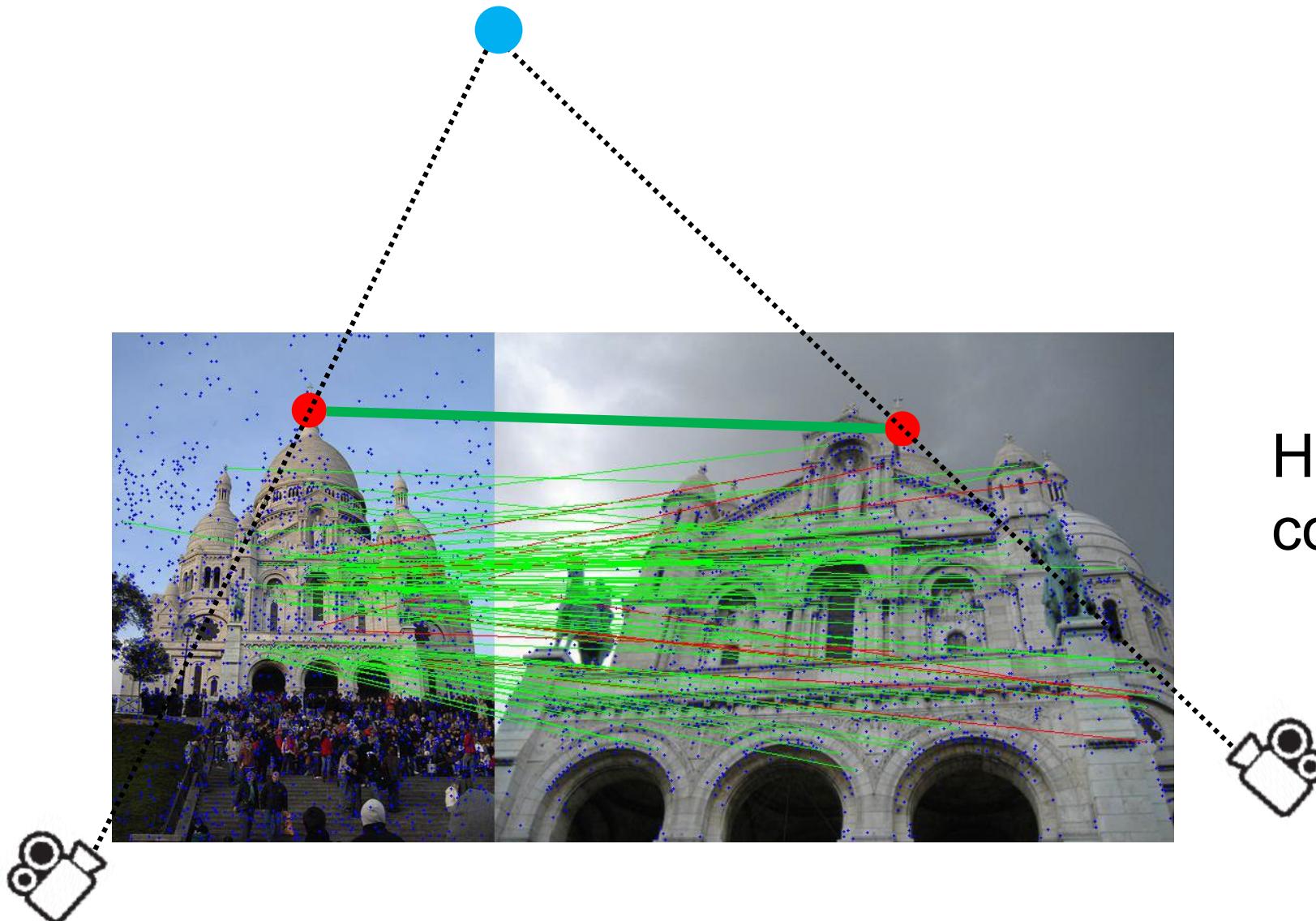


Image-based 3D Reconstruction



COLMAP [Johannes et al. 2016, Schoenberger et al. 2016]
(Input: 100 images)

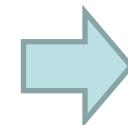
Challenges in Image-based Reconstruction

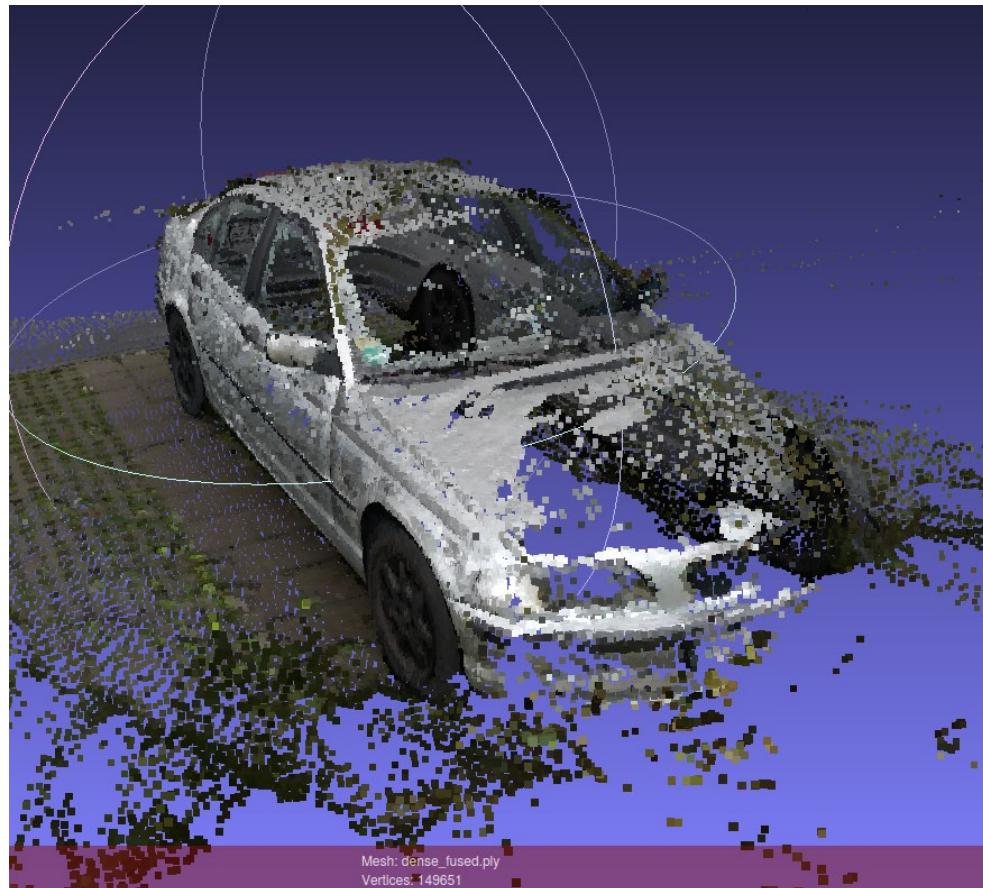


Hard to extract reliable correspondences!

Computer Graphics Rendering

Rendering requires very high-quality 3D models





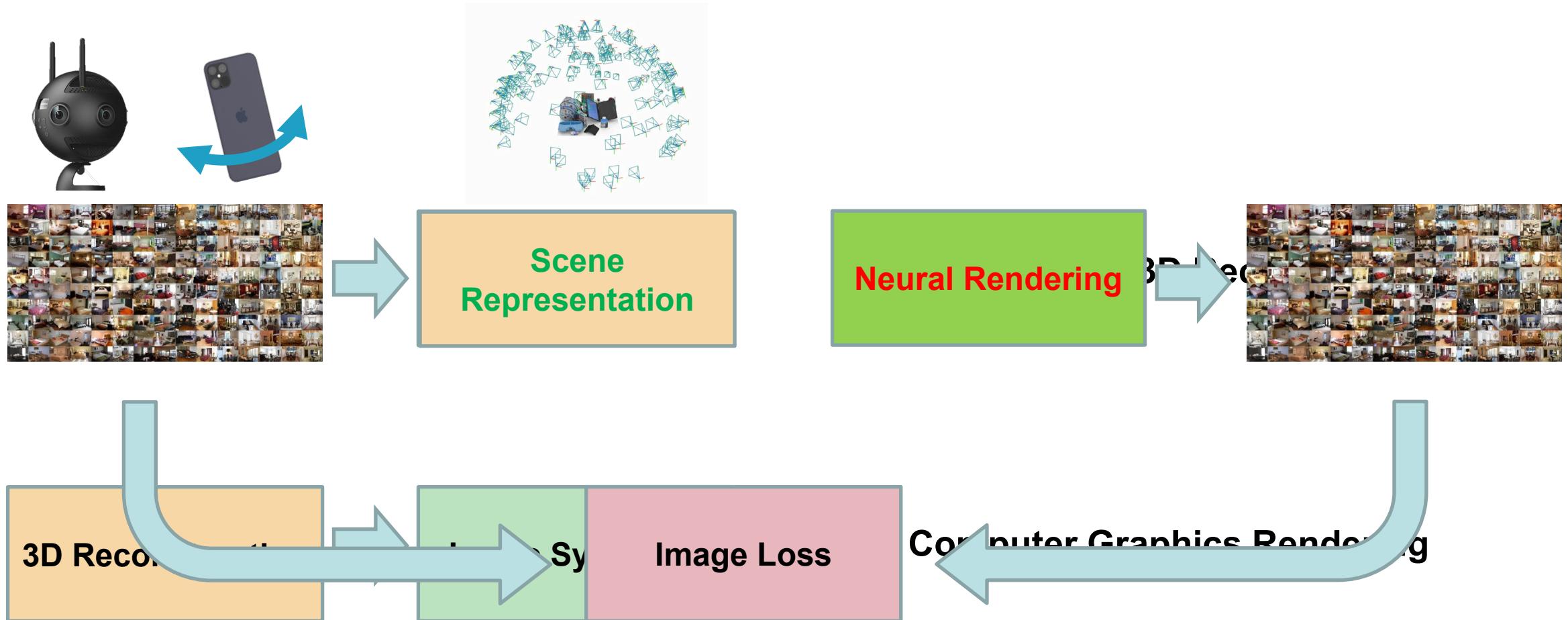
VS



Neural Scene Representation and Neural Rendering

To the rescue

Neural Scene Representation and Neural Rendering



Neural Scene Representation and Neural Rendering

Neural Rendering

DOI: 10.1111/cgf.14022
EUROGRAPHICS 2020
R. Manzak and V. Sundeck
(Guest Editors)

Volume 39 (2020), Number 2
STAR – State of The Art Report

State of the Art on Neural Rendering

A. Tewari^{1*}, O. Friedl^{2*}, J. Thies^{3*}, V. Sitzmann^{2*}, S. Lombardi⁴, K. Sunkavalli⁵, R. Martin-Brualla³, T. Simon⁴, J. Saragih⁶, M. Nießner⁷, R. Pandey⁸, S. Fanello⁹, G. Wetzstein², J.-Y. Zhu⁵, C. Theobalt¹, M. Agrawala², E. Shechtman¹⁰, D. B. Goldman⁶, M. Zollhöfer¹¹

¹MPI Informatics ²Stanford University ³Technical University of Munich ⁴Facebook Reality Labs ⁵Adobe Research ⁶Google Inc ⁷Equal contribution.

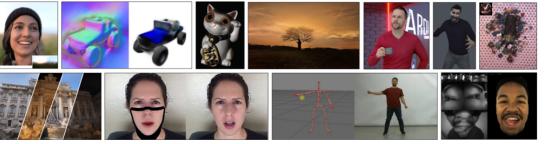


Figure 1: Neural renderings of a large variety of scenes. See Section 6 for more details on the various methods. Images from [SFT*19, SZW19, XBS*19, KHM17, GLD*19, MBP*18, XSHR18, MGK*19, FTZ*19, LZT*19, WSS*19].

Abstract
Efficient rendering of photo-realistic virtual worlds is a long standing effort of computer graphics. Modern graphics techniques have succeeded in synthesizing photo-realistic images from hand-crafted scene representations. However, the automatic generation of shapes, materials, lighting, and other aspects of scenes remains a challenging problem that, if solved, would make photo-realistic computer graphics more widely accessible. Concurrently, progress in computer vision and machine learning have given rise to a new approach to image synthesis and editing: deeply generative models. Neural rendering is a new method for generating photo-realistic images and videos using deep generative models. Neural rendering draws on knowledge from computer graphics, e.g., by the integration of differentiable rendering into network training. With a plethora of applications in computer graphics and vision, neural rendering is poised to become a new area in the graphics community, yet no survey of this emerging field exists. This state-of-the-art report summarizes the recent trends and applications of neural rendering. We focus on approaches that combine classic computer graphics techniques with deep generative models to obtain controllable and photo-realistic outputs. Starting with an overview of the underlying computer graphics and machine learning concepts, we discuss critical aspects of neural rendering approaches. Specifically, our emphasis is on the type of control, i.e., how the control is provided to the neural rendering pipeline, and how it is used to generate novel images and videos for static scene synthesis. The second half of this state-of-the-art report is focused on the many important use cases for the described algorithms such as novel view synthesis, semantic photo manipulation, facial and body reenactment, relighting, free-viewpoint video, and the creation of photo-realistic avatars for virtual and augmented reality telepresence. Finally, we conclude with a discussion of the social implications of such technology and investigate open research problems.

1. Introduction
The creation of photo-realistic images or virtual worlds has been one of the primary driving forces for the development of sophisticated computer graphics techniques. Computer graphics approaches span the range from real-time rendering, which enables the latest generation of computer games, to sophisticated global illumination simulation for the creation of photo-realistic digital humans in feature films. In both cases, one of the main bottlenecks is content creation, i.e., that a vast amount of tedious and expensive manual work of skilled artists is required for the creation of the underlying scene representations in terms of surface geometry, appearance, material, light sources, and cameras. Computer neural rendering has recently emerged in the computer vision and machine learning communities. The seminal work on Generative Adversarial Neural Networks (GANs) by Goodfellow et al. [GPAM*14] has evolved in recent years into

© 2020 The Authors
Computer Graphics Forum © 2020 The Eurographics Association and John Wiley & Sons Ltd. Published by John Wiley & Sons Ltd.

[Tewari et al. 2020]

This state of the art report is accepted at EUROGRAPHICS 2022.

arXiv:2111.05849v2 [cs.GR] 30 Mar 2022

Advances in Neural Rendering

A. Tewari^{1,6*}, J. Thies^{3*}, B. Mildenhall^{1,8*}, P. Savoie^{3*}, E. Tretschk¹, W. Yifan^{4,8}, C. Lassner⁵, V. Sitzmann⁶, R. Martin-Brualla³, S. Lombardi⁴, T. Simon³, C. Theobalt¹, M. Nießner⁷, J. T. Barron⁸, G. Wetzstein², M. Zollhöfer¹, V. Golyamkin¹

¹MPI for Informatics ²MIT ³MPI for Intelligent Systems ⁴Google Research ⁵ETH Zurich ⁶Raility Labs Research ⁷Equal contribution.



Figure 1: This state-of-the-art report discusses a large variety of neural rendering methods which enable applications such as novel-view synthesis of static and dynamic scenes, generating models of objects, and scene relighting. See Section 4 for more details on the various methods. Images adapted from [MST*20, TY20, CMK*21, ZSD*21, BBJ*21, LSS*21, PSB*21, JXX*21, PDW*21] ©2021 IEEE.

Abstract
Synthesizing photo-realistic images and videos is at the heart of computer graphics and has been the focus of decades of research. Traditionally, synthetic images of a scene are generated using rendering algorithms such as rasterization or ray tracing, where the scene is represented by a collection of primitives (e.g., triangles) and the camera defines what parts of the scene to render, what the camera sees, and what is rendered. Neural rendering, on the other hand, does not define the actual scene and what is rendered, and are referred to as scene representations (where a scene consists of one or more objects). Example scene representations are triangle meshes (e.g., created by an artist), point clouds (e.g., from a depth sensor), volumetric grids (e.g., from a CT scan), or implicit surface functions (e.g., truncated signed distance fields). The reconstruction of such a scene representation from observations using differentiable rendering losses is known as inverse graphics or inverse rendering. Neural rendering is closely related, and combines ideas from classical computer graphics and machine learning to create algorithms for synthesizing images from real-world observations. Neural rendering has also moved beyond the goal of synthesis and now includes scene editing and composition. While most of these approaches are scene-specific, we also discuss techniques that generalize across object classes and can be used for generative tasks. In addition to reviewing these state-of-the-art methods, we provide an overview of fundamental concepts and definitions used in the current literature. We conclude with a discussion on open challenges and social implications.

1. Introduction
Synthesis of controllable and photo-realistic images and videos is one of the fundamental goals of computer graphics. During the last decades, methods and representations have been developed to mimic the image formation model of real cameras, including the handling of complex materials and global illumination. These methods are based on the laws of physics and simulate the light transport from light sources in the virtual camera for synthesis. To this end, all physical parameters of the scene have to be known for

[Tewari et al. 2021]

Lingjie Liu

Neural Rendering - Definition

- Definition:

*"Deep neural networks for **image or video generation** that enable **explicit or implicit control** of **scene properties**"*

1)

Generative networks that synthesis raw pixel output

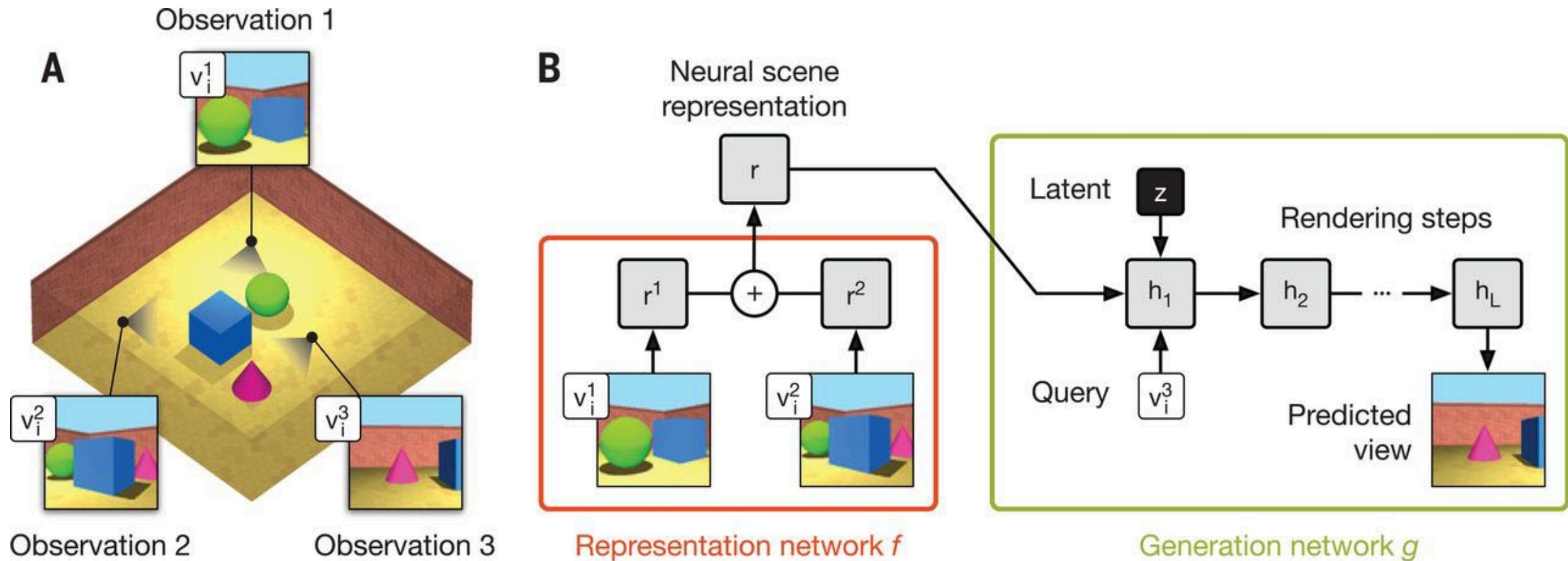
2)

Controllable by interpretable parameters or by video/audio input.

3)

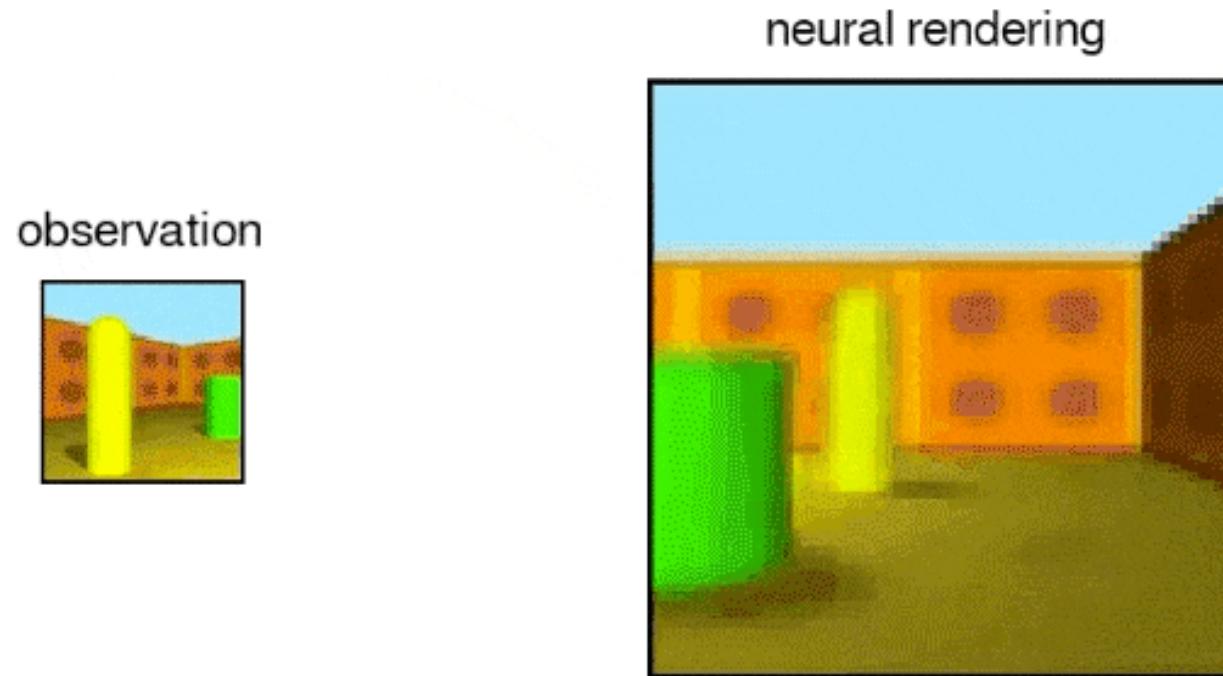
Illumination, camera, pose, geometry, appearance, or semantic structure controllable

Generative Query Network (GQN)



Neural scene representation and rendering, Eslami et al. 2018

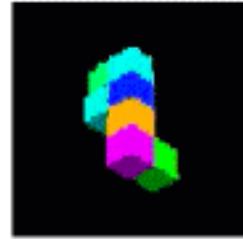
Generative Query Network (GQN)



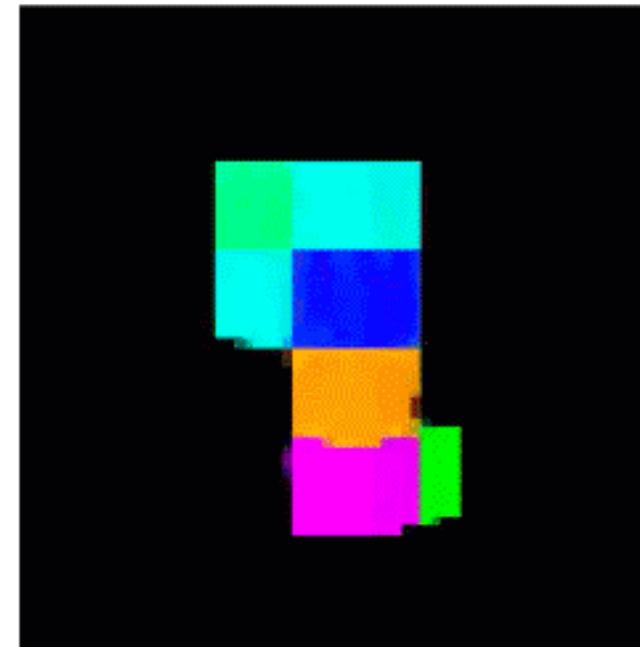
Neural scene representation and rendering, Eslami et al. 2018

Generative Query Network (GQN)

observation

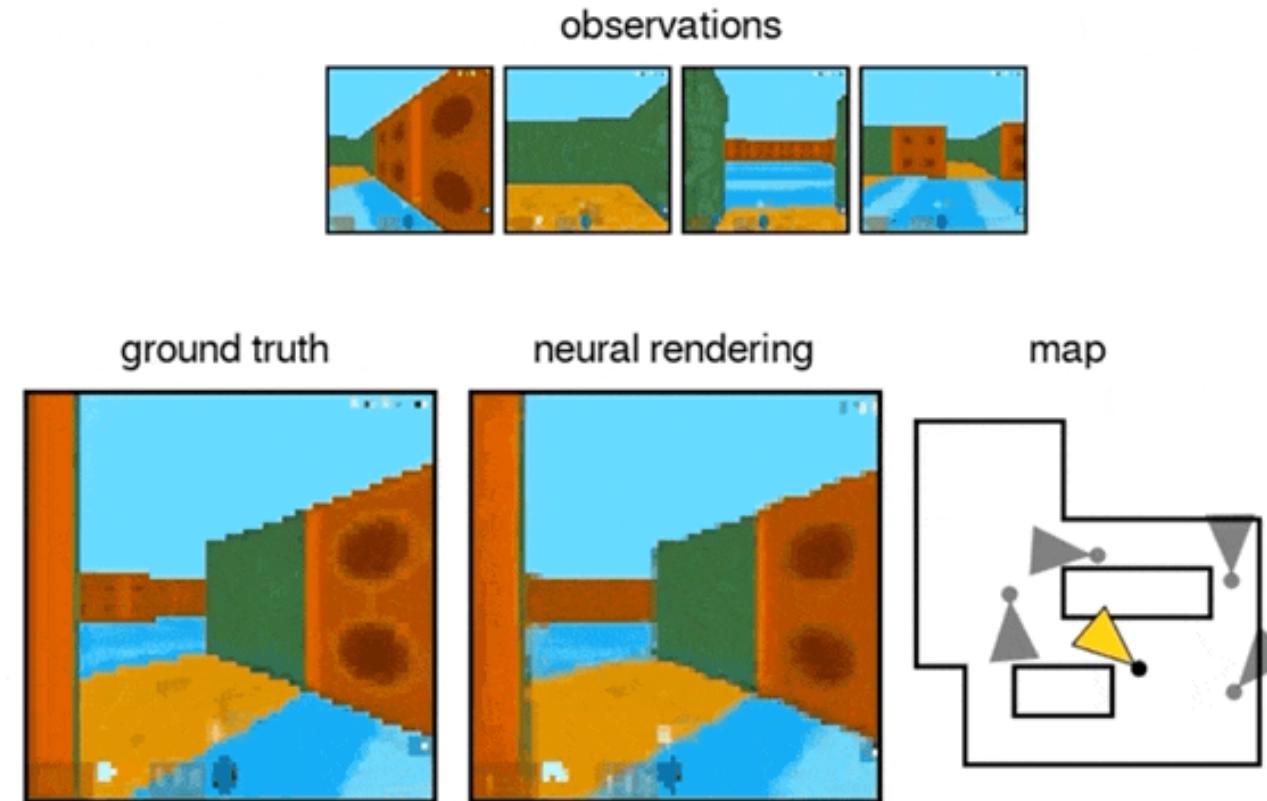


neural rendering



Neural scene representation and rendering, Eslami et al. 2018

Generative Query Network (GQN)



Neural scene representation and rendering, Eslami et al. 2018

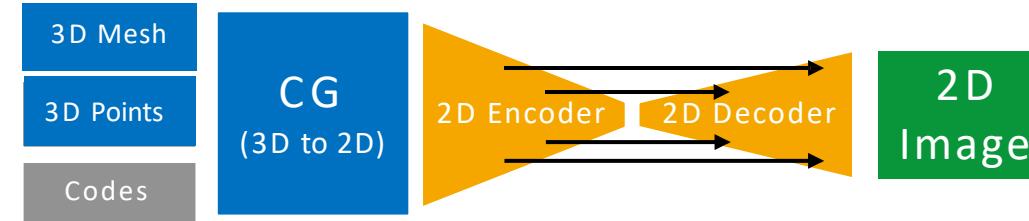
Neural Rendering Zoo

“Regress it”



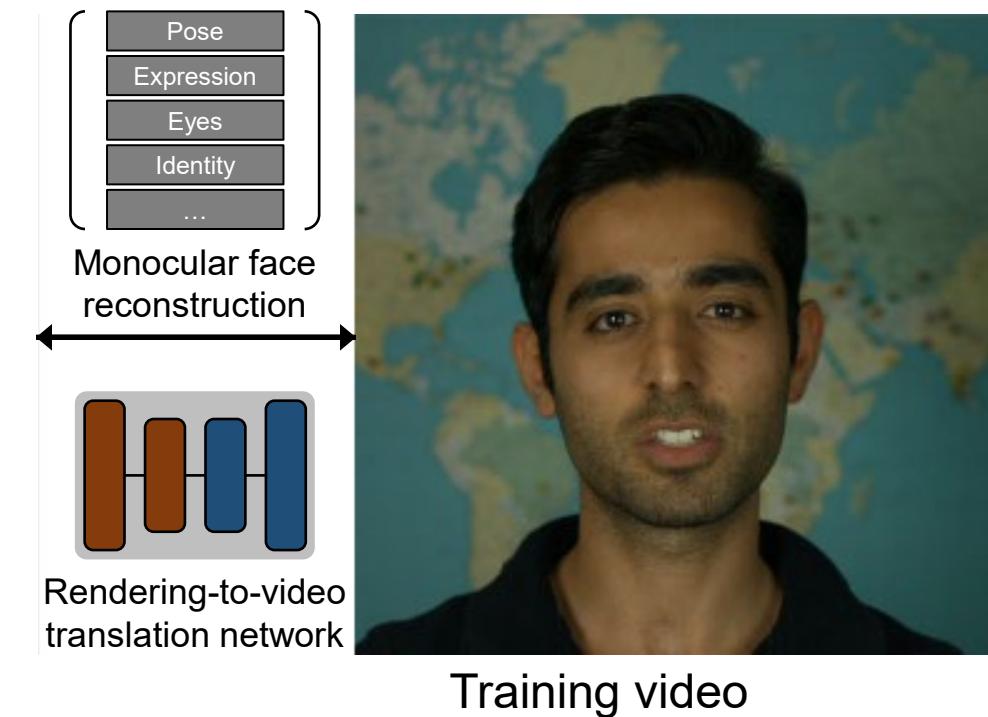
e.g., GQN

“Make it more real”



e.g., DVP or DNR

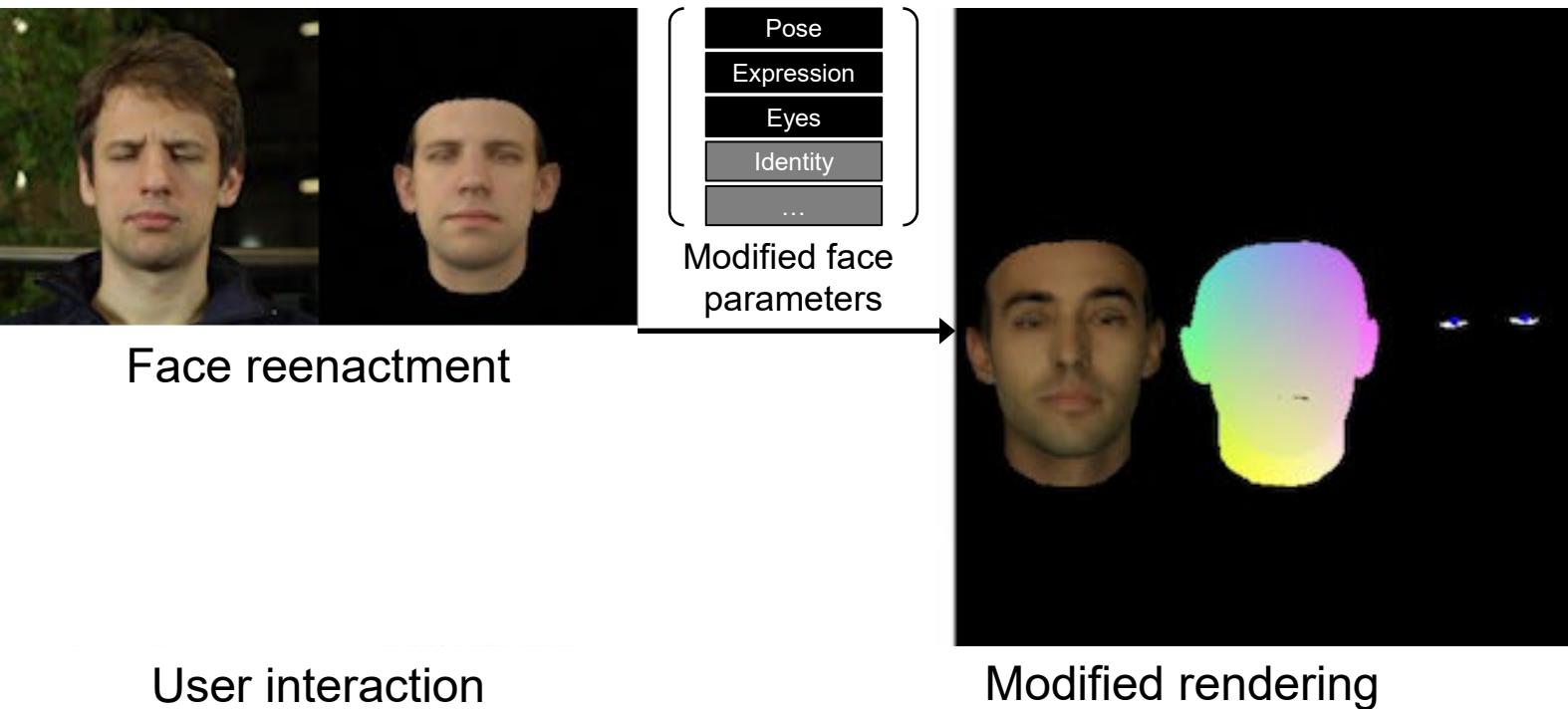
Deep Video Portraits (DVP)



Deep Video Portraits, Kim et al. 2018

Lingjie Liu

Deep Video Portraits (DVP)



Deep Video Portraits, Kim et al. 2018

Lingjie Liu

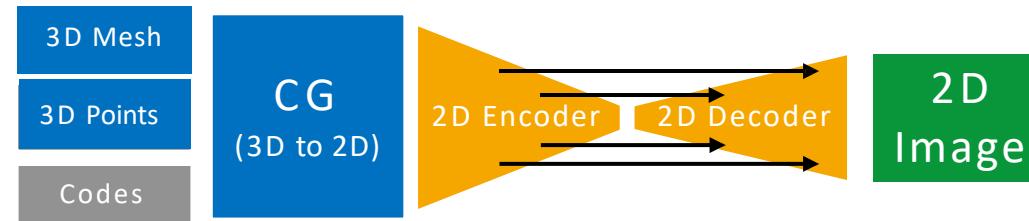
Neural Rendering Zoo

“Regress it”



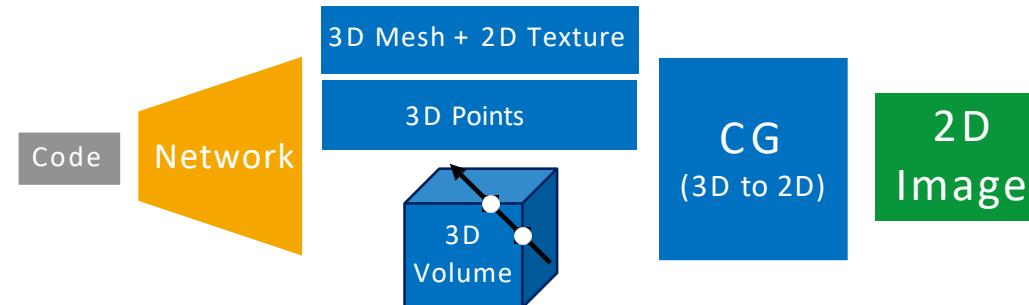
e.g., GQN

“Make it more real”



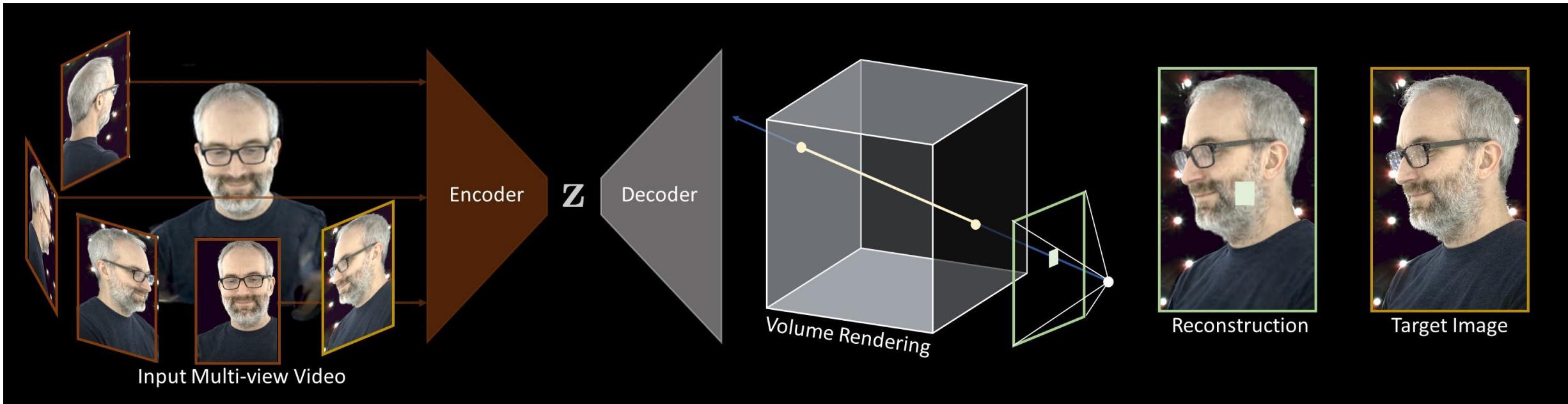
e.g., DVP or DNR

“Regress & render”



e.g., Neural Volumes

Neural Volumes



Neural Volumes: Learning Dynamic Renderable Volumes from Images, Lombardi et al. 2019

Neural Volumes



Neural Volumes



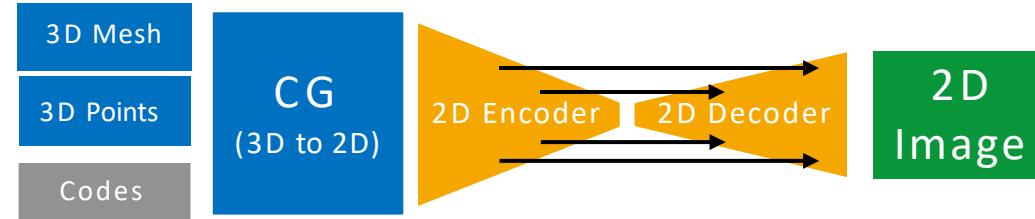
Neural Rendering Zoo

“Regress it”



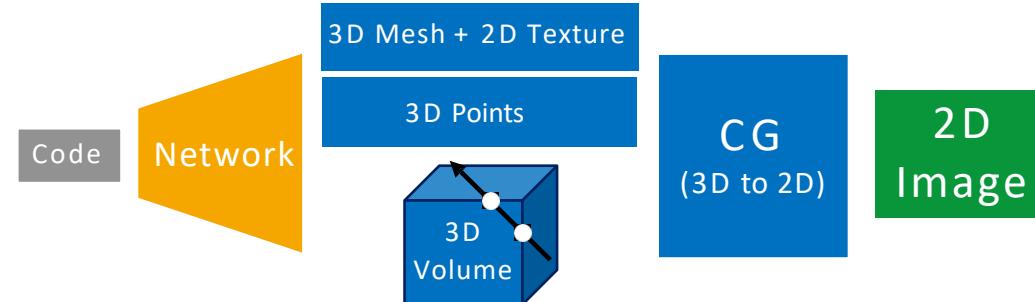
e.g., GQN

“Make it more real”



e.g., DVP or DNR

“Regress & render”



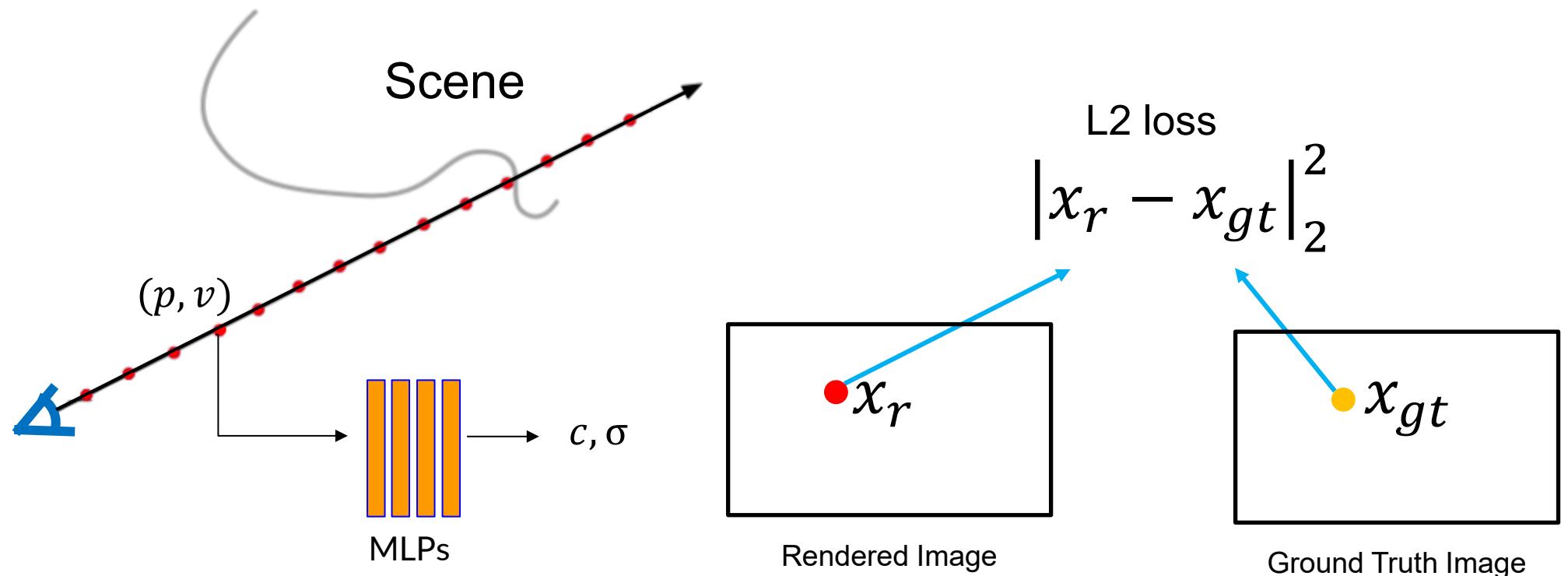
e.g., Neural Volumes

“Step, sample & blend”



e.g., NeRF

Neural Radiance Fields (NeRF)



[Mildenhall et al. 2020]

Neural Radiance Fields (NeRF)



[Mildenhall et al. 2020]

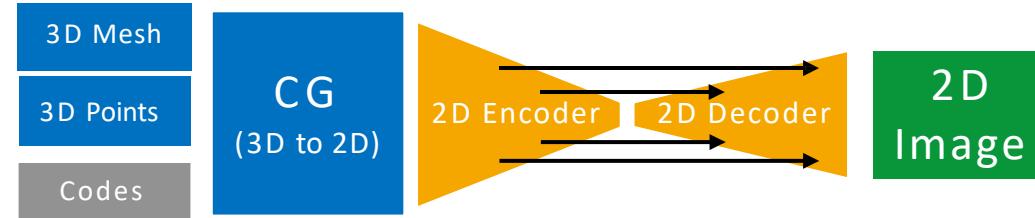
Neural Rendering Zoo

“Regress it”



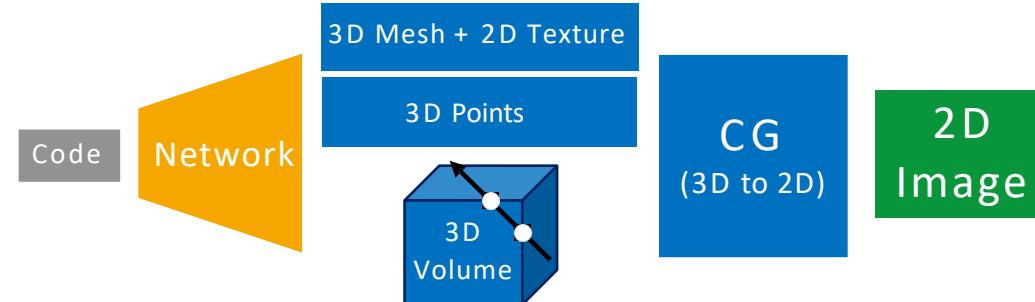
e.g., GQN

“Make it more real”



e.g., DVP or DNR

“Regress & render”



e.g., Neural Volumes

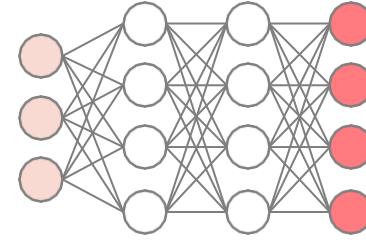
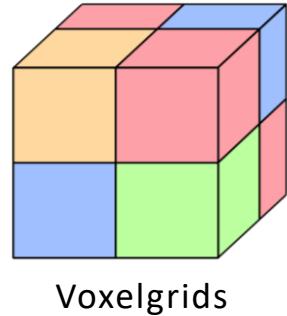
“Step, sample & blend”



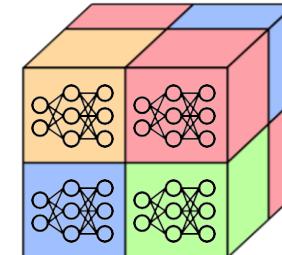
e.g., NeRF

Overview

Scene
Representation



Implicit Function



Hybrid
Implicit/Explicit

• • •

(We'll talk more about other
representations in our next class)

Renderer

Volumetric

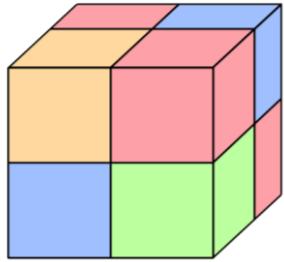
Sphere-Tracing
Volumetric

Volumetric

Both Scene Representation and Differentiable Renderer often
adapted from traditional computer graphics.

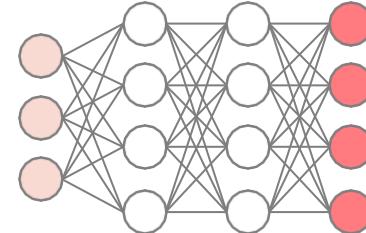
Requirements

Scene
Representation

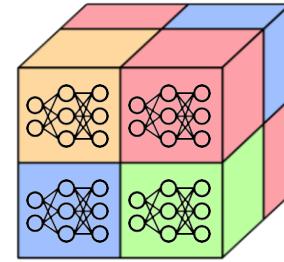


Voxelgrids

Renderer



Implicit Function



Hybrid
Implicit/Explicit

Sphere-Tracing
Volumetric

Volumetric

• • •

Pros

Cons

Voxel-based methods

DeepVoxels



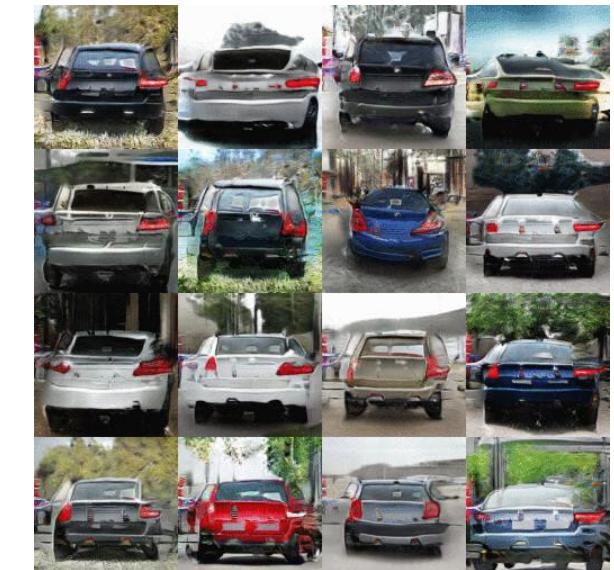
Sitzmann et al., CVPR 2018

Neural Volumes



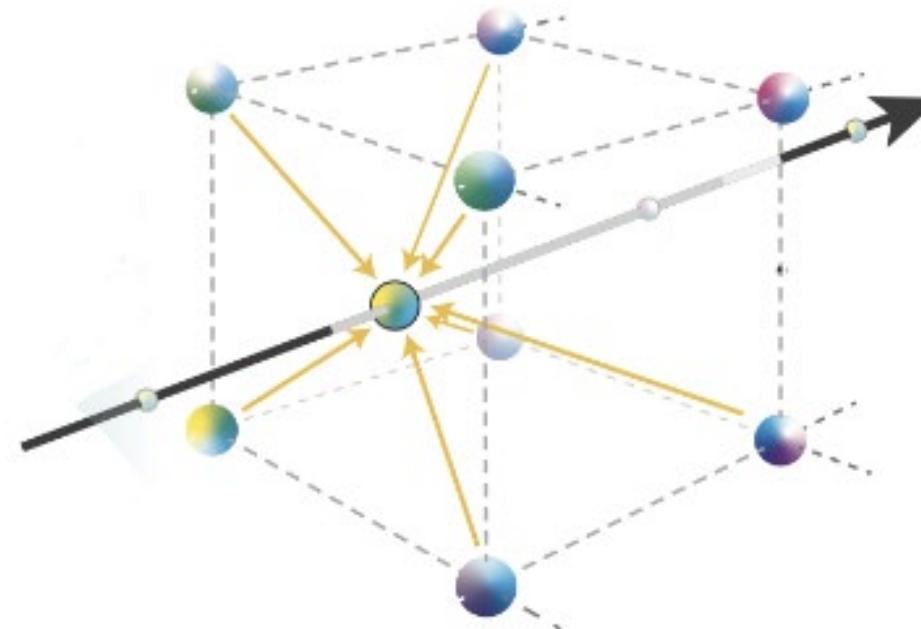
Lombardi et al., SIGGRAPH 2019

HoloGAN



Phuoc et al., ICCV 2019

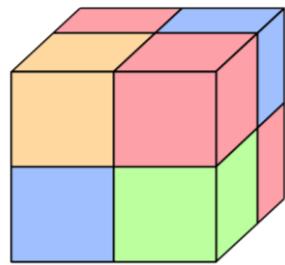
Voxel-based methods



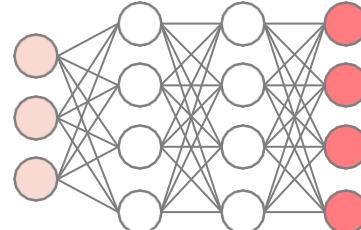
Trilinear Interpolation

Requirements

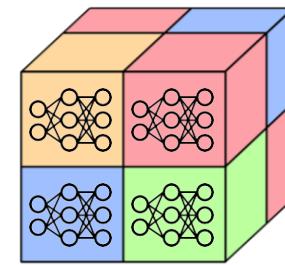
Scene
Representation



Voxelgrids



Implicit Function



Hybrid
Implicit/Explicit

Renderer

Volumetric

Volumetric

Pros

Fast rendering

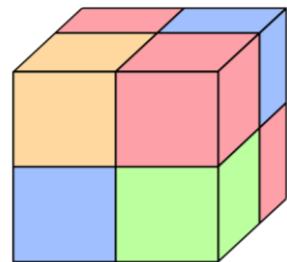
Cons

Memory $O(n^3)$
Limited spatial
resolution

• • •

Requirements

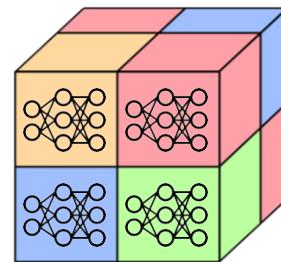
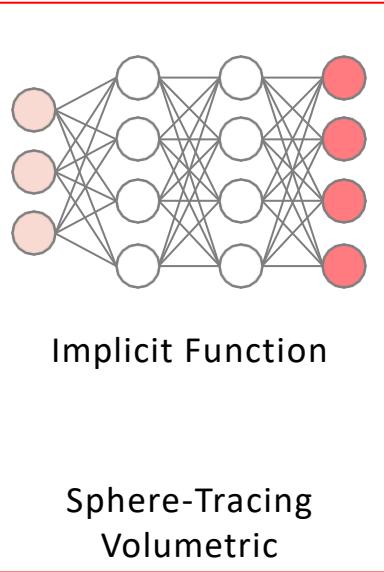
Scene
Representation



Voxelgrids

Renderer

Volumetric



Hybrid
Implicit/Explicit

Volumetric

• • •

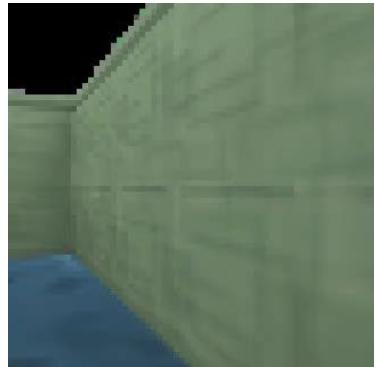
Pros

Fast rendering

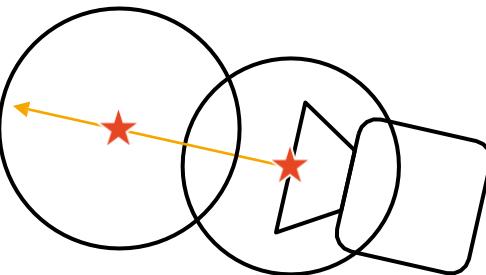
Cons

Memory $O(n^3)$
Limited spatial
resolution

Neural Implicit Approaches



Scene Representation Networks
Generalizes across scenes
Sitzmann et al., NeurIPS 2019



Sphere tracing



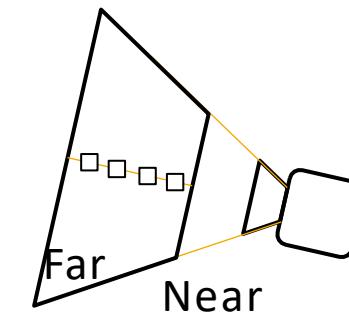
Differentiable Volumetric Rendering
Generalizes across scenes
Niemeyer et al., CVPR 2020



NeRF
Single-scene
Mildenhall et al., ECCV 2020



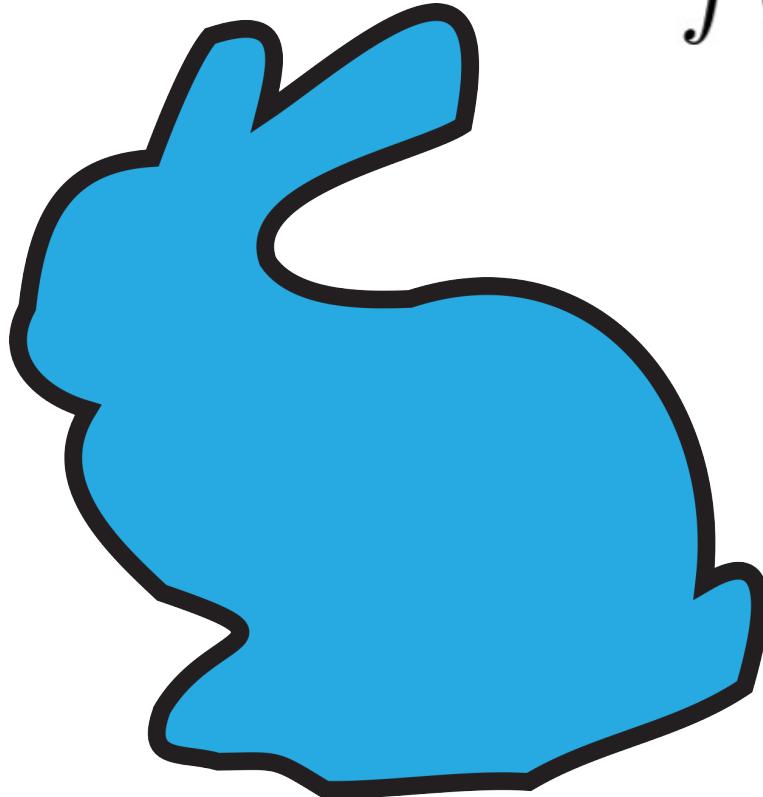
Implicit Differentiable Renderer
Single-scene
Yariv et al., NeurIPS 2020



Volumetric

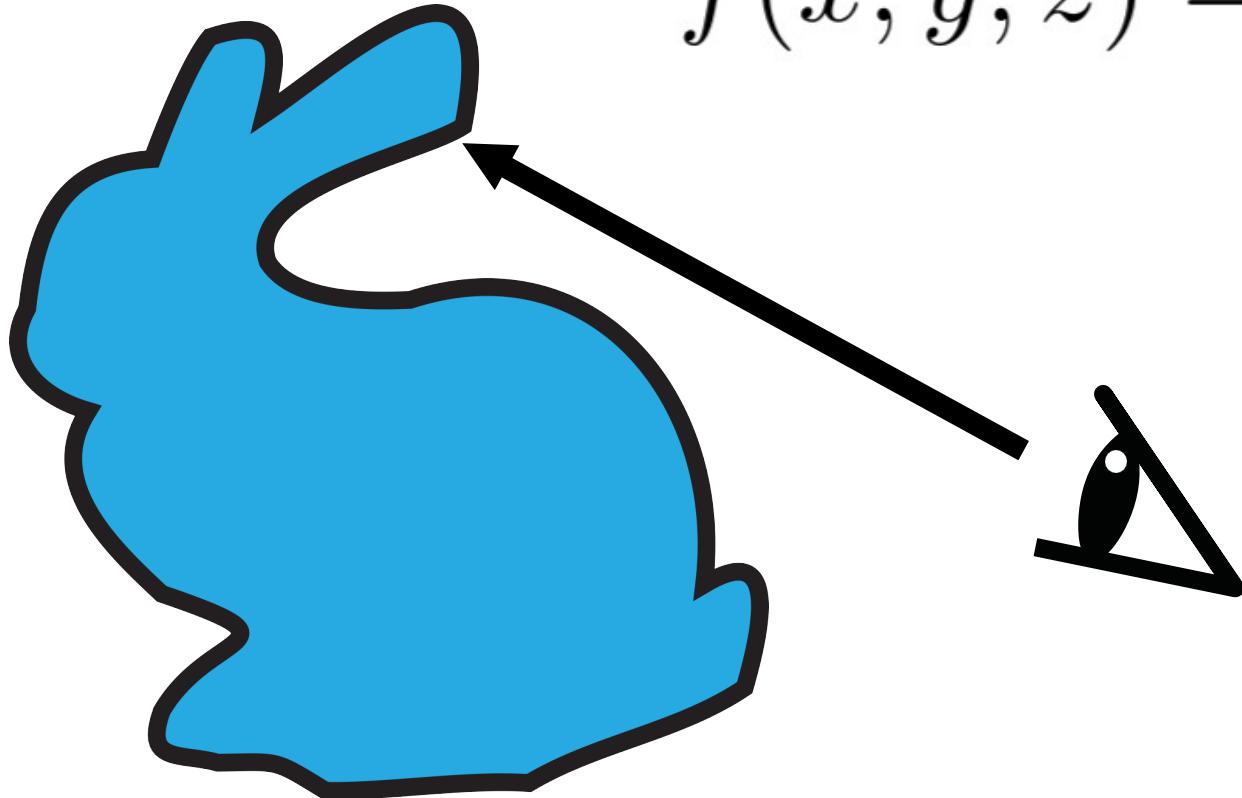
Sphere Tracing

$$f(x, y, z) = d$$



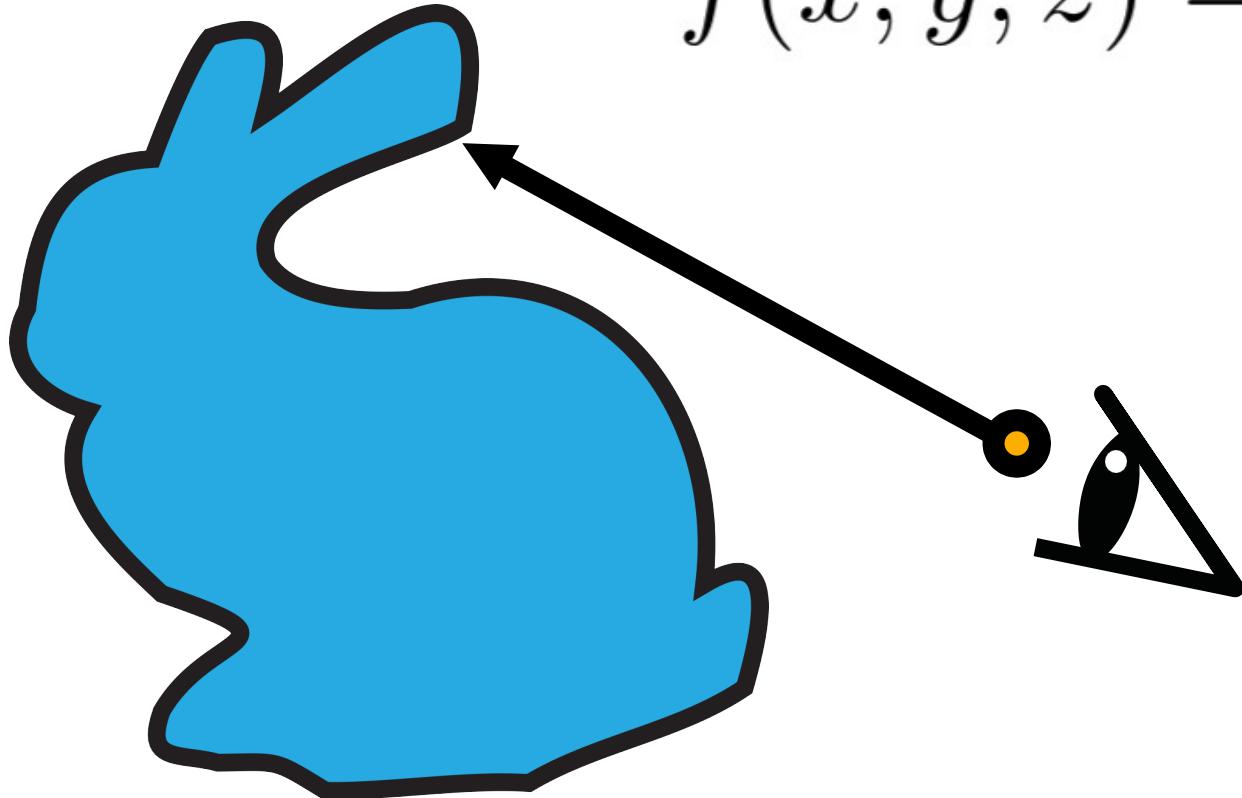
Sphere Tracing

$$f(x, y, z) = d$$



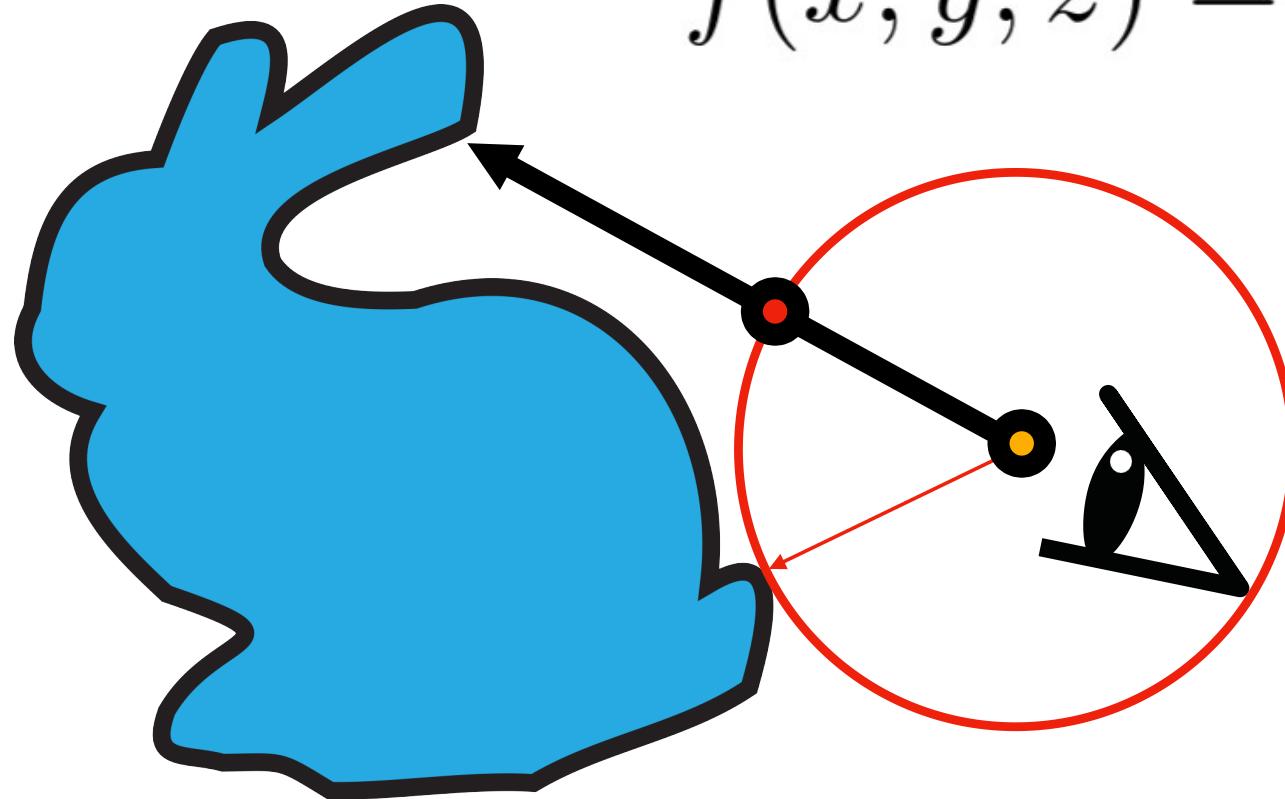
Sphere Tracing

$$f(x, y, z) = d$$

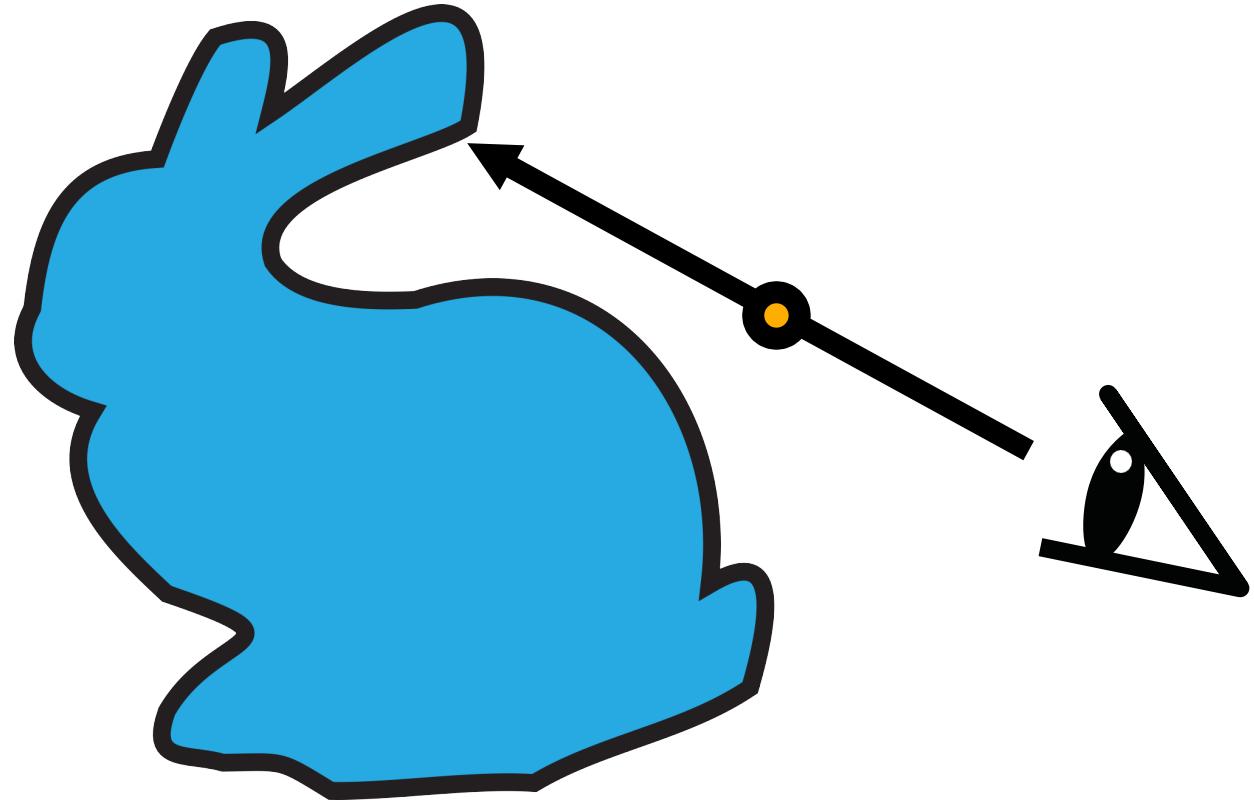


Sphere Tracing

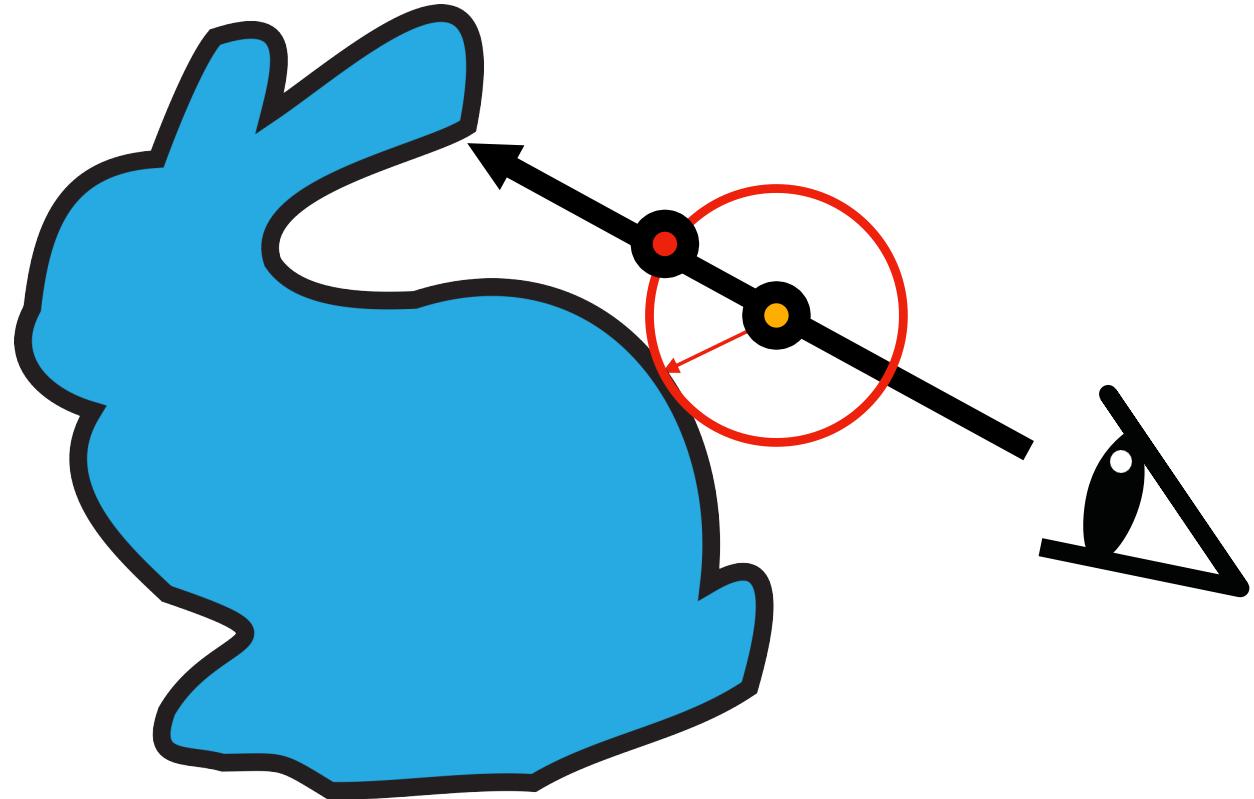
$$f(x, y, z) = d$$



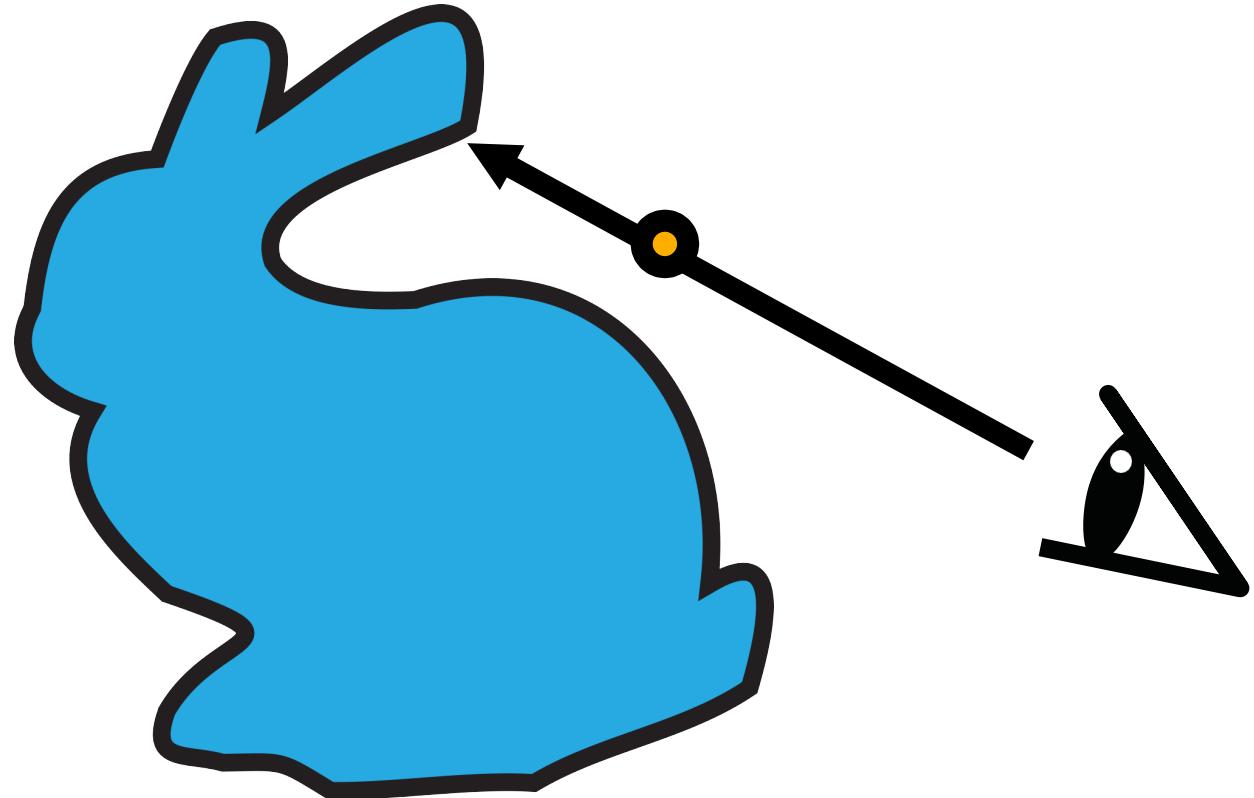
Sphere Tracing



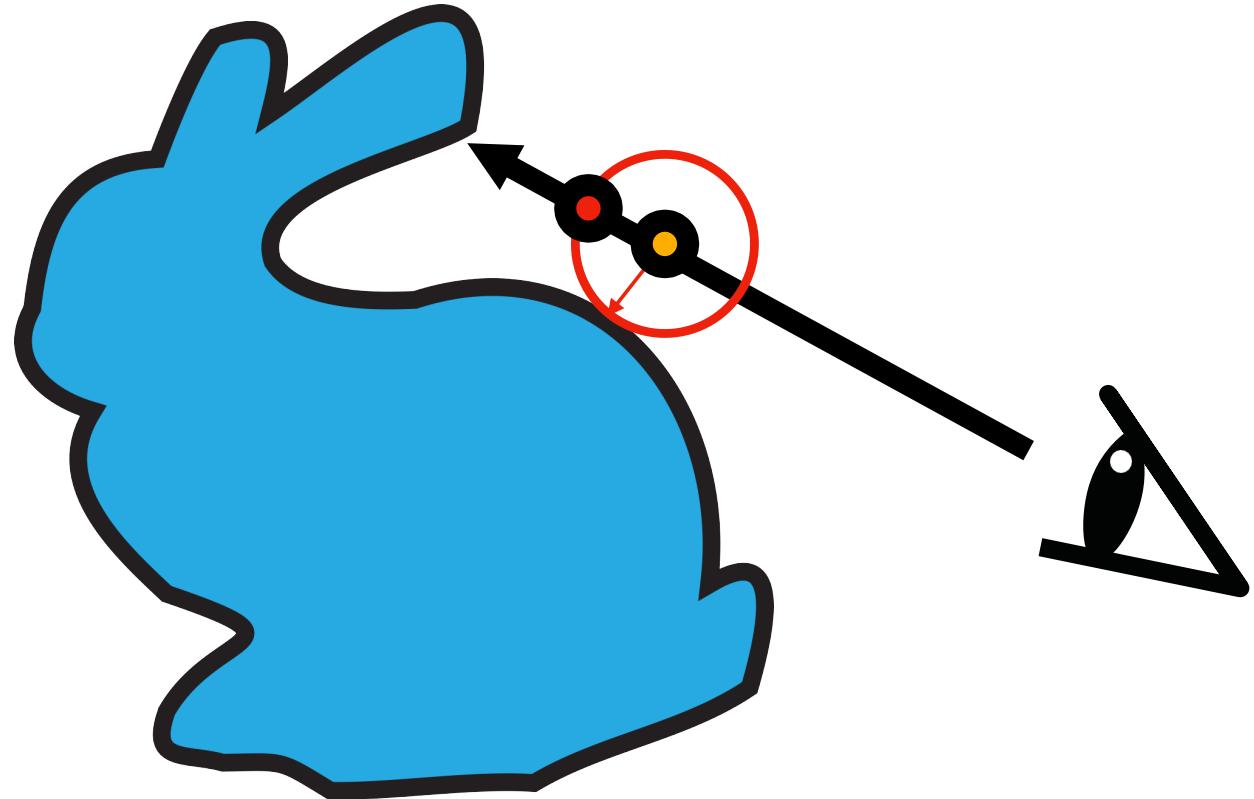
Sphere Tracing



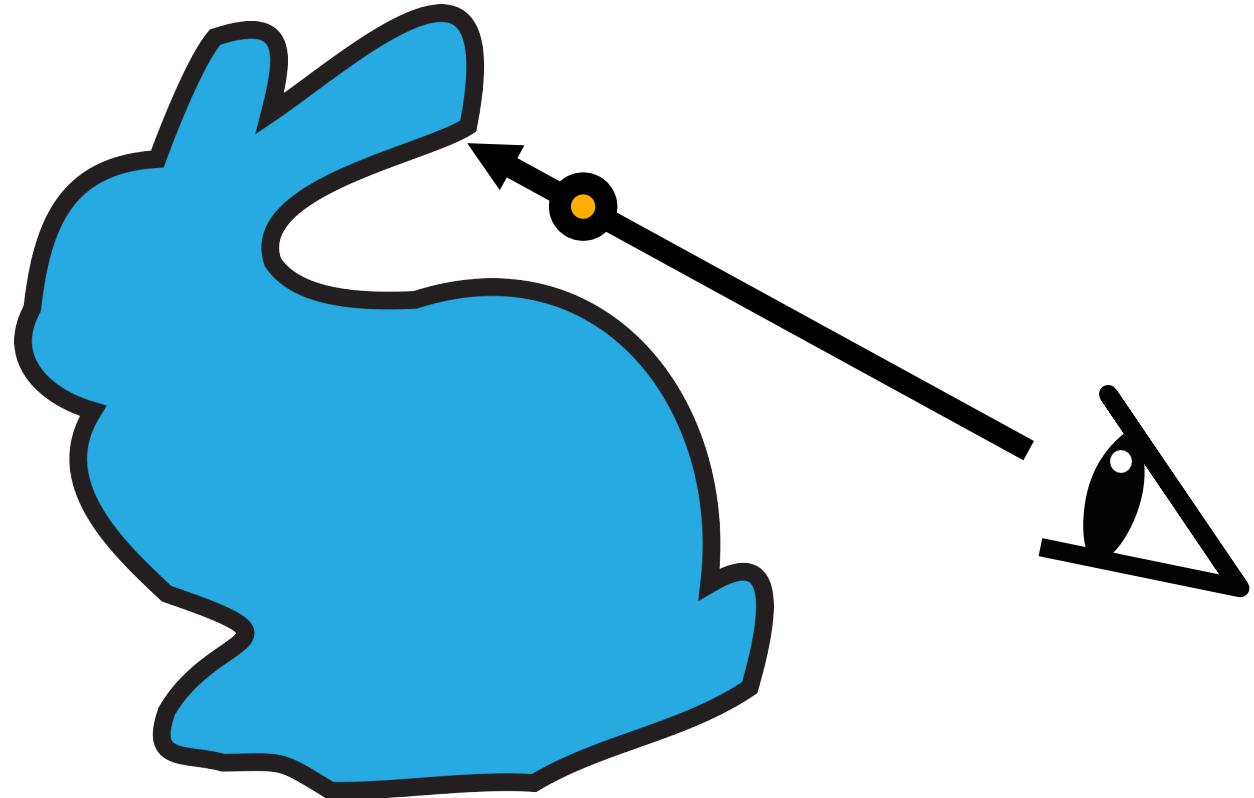
Sphere Tracing



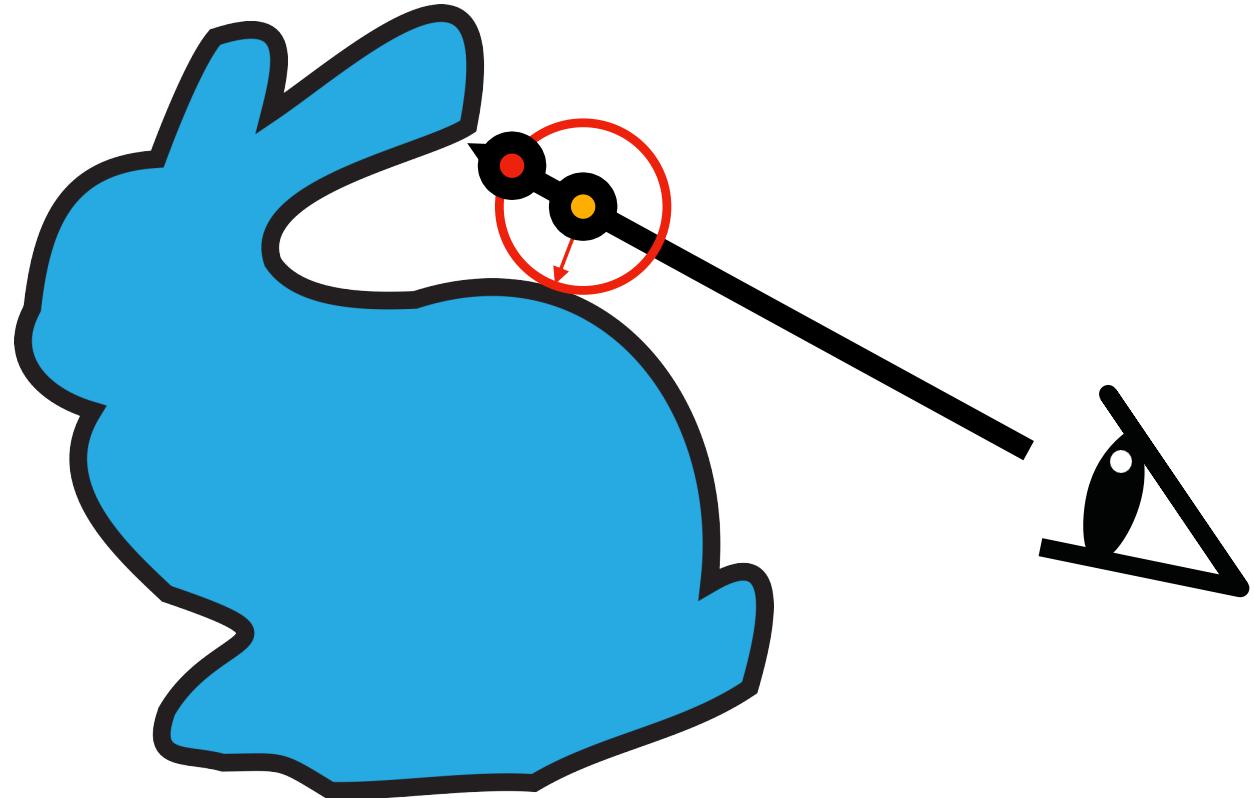
Sphere Tracing



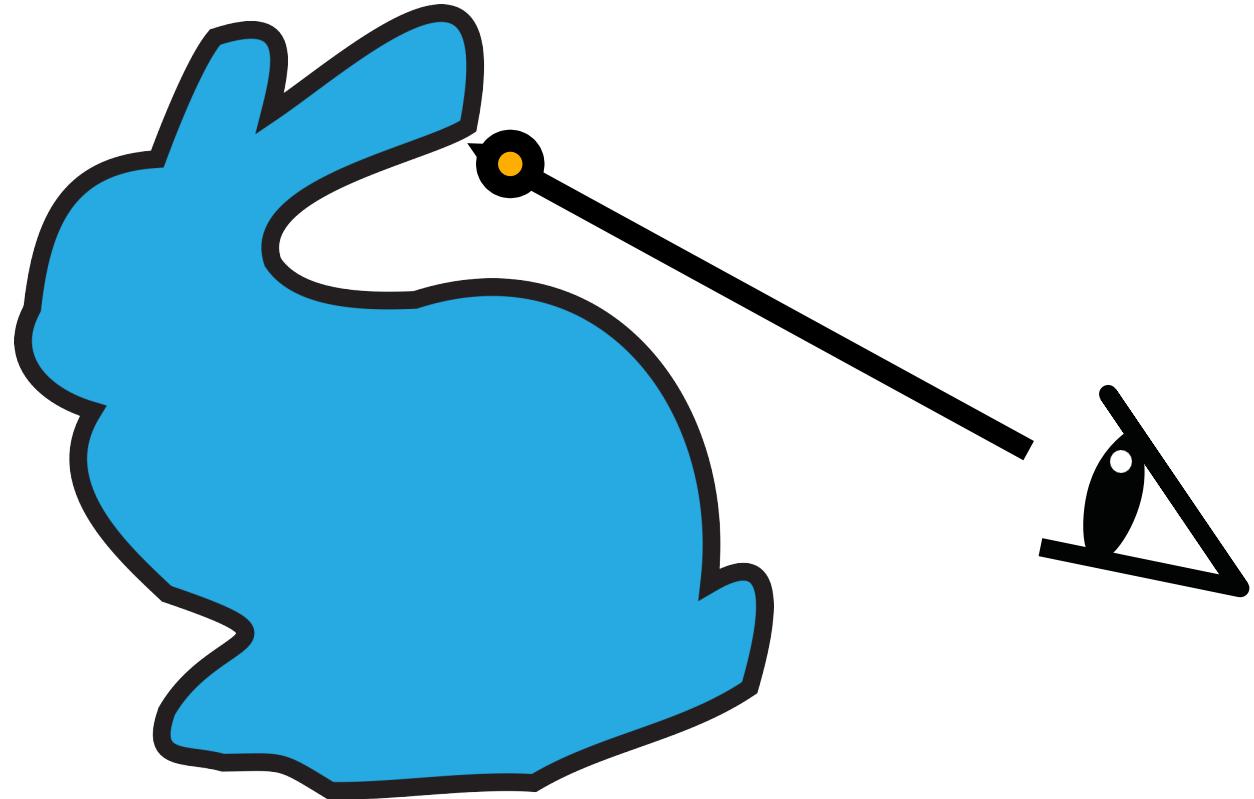
Sphere Tracing



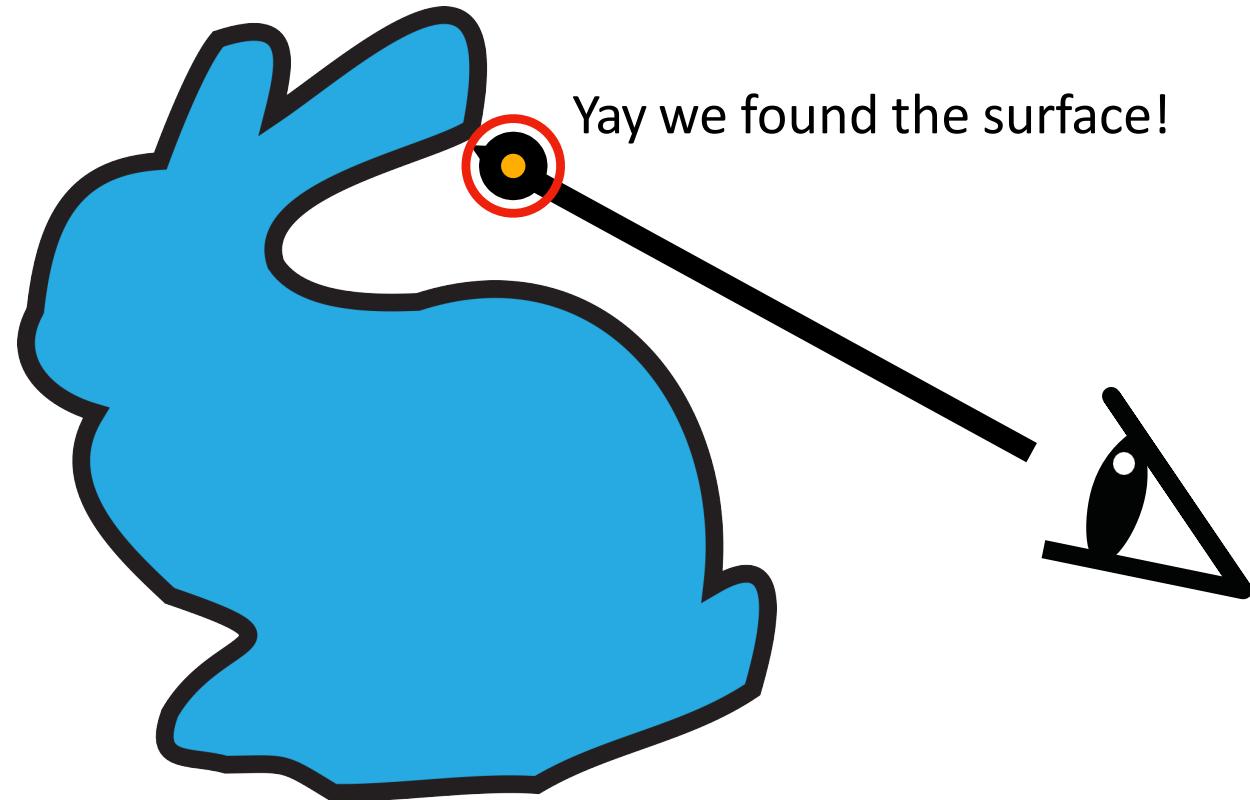
Sphere Tracing



Sphere Tracing

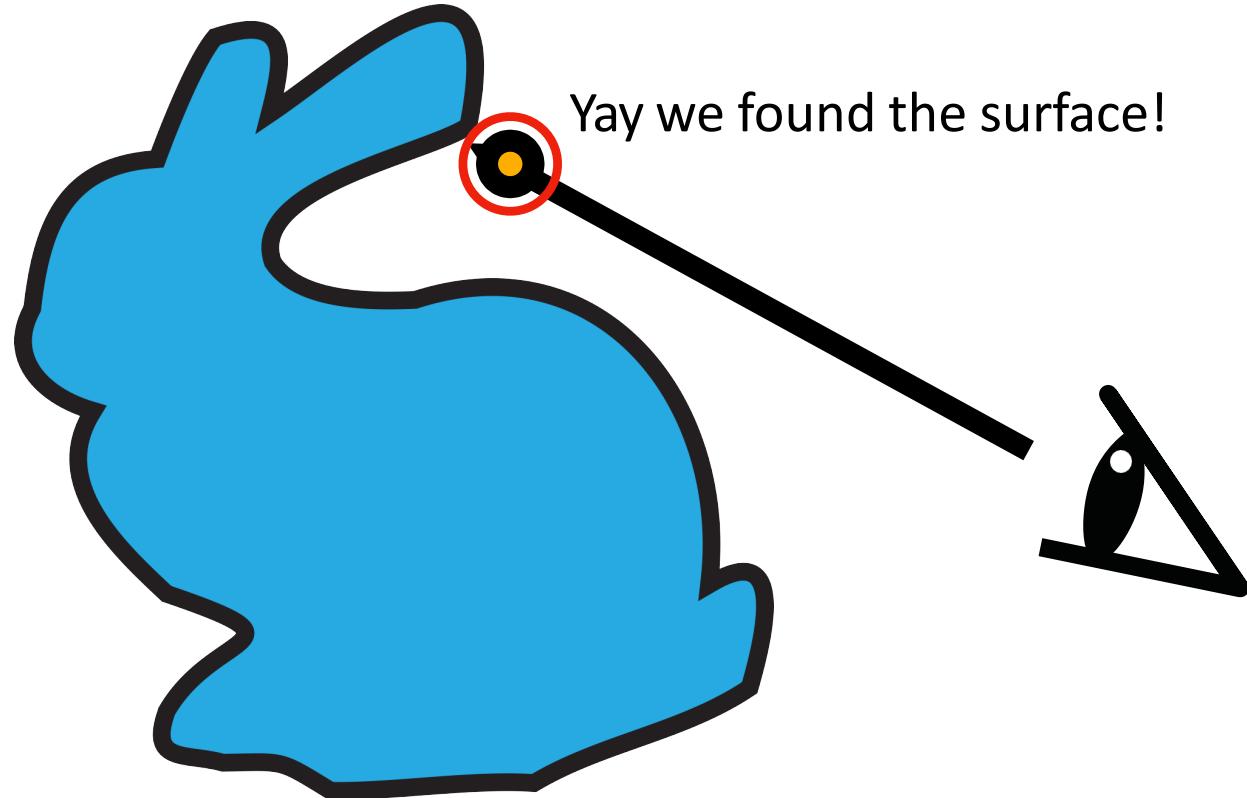


Sphere Tracing

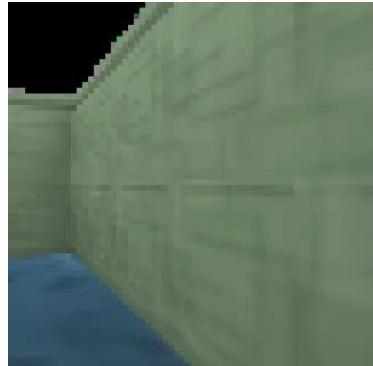


Sphere Tracing

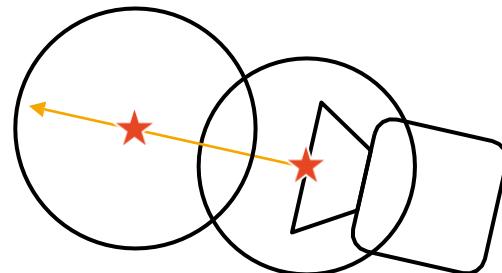
$$f(x, y, z) = d$$



Neural Implicit Approaches



Scene Representation Networks
Generalizes across scenes
Sitzmann et al., NeurIPS 2019



Sphere tracing

- Faster
- Fewer network evaluations
- Convergence more difficult



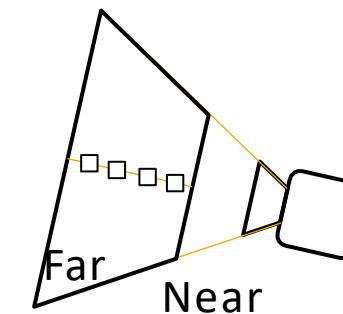
Differentiable Volumetric Rendering
Generalizes across scenes
Niemeyer et al., CVPR 2020



NeRF
Single-scene
Mildenhall et al., ECCV 2020



Implicit Differentiable Renderer
Single-scene
Yariv et al., NeurIPS 2020

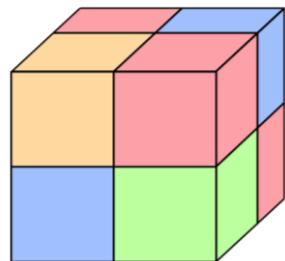


Volumetric

- Higher Quality
- Easy convergence
- Very expensive

Requirements

Scene
Representation



Voxelgrids

Renderer

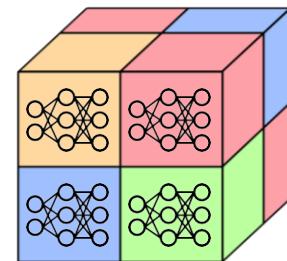
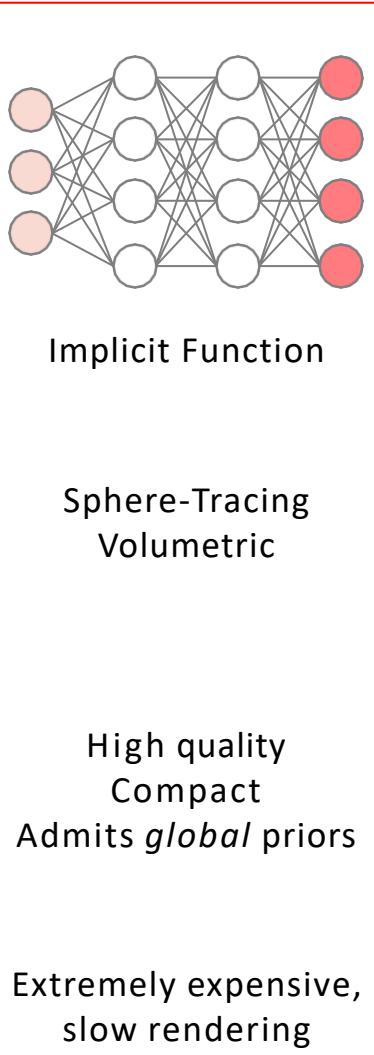
Volumetric

Pros

Fast rendering

Cons

Memory $O(n^3)$
Limited spatial
resolution



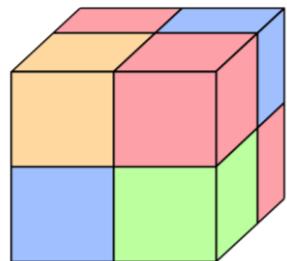
Hybrid
Implicit/Explicit

Volumetric

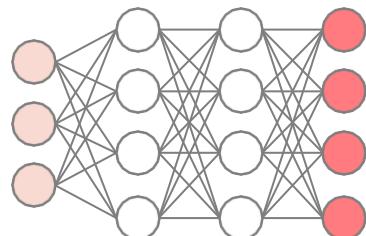
• • •

Requirements

Scene
Representation



Voxelgrids

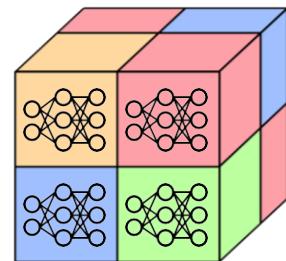


Implicit Function

Renderer

Volumetric

Sphere-Tracing
Volumetric



Hybrid
Implicit/Explicit

Volumetric

• • •

Pros

Fast rendering

High quality
Compact
Admits *global* priors

Cons

Memory $O(n^3)$
Limited spatial
resolution

Extremely expensive,
slow rendering

Hybrid Implicit / Explicit

Wine Holder



SRN (Sitzmann et al. 2019)
(Rendering speed: 1.10 s/frame)

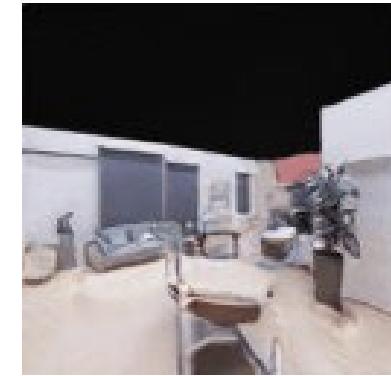


NSVF
(Rendering speed: 1.68 s/frame)

Neural Sparse Voxel Fields,
Liu et. al., NeurIPS 2020



PiFU, Saito et al., ICCV 2019
GRF, Trevithick et al., arXiv 2020
pixelNeRF, Yu et. al., CVPR 2021
MVSNeRF, Chen et al., arXiv 2021
Learn *local* (image patch-based) priors



Unconstrained Scene Generation with
Locally Conditioned Radiance Fields,
DeVries et al., arXiv 2021

Neural Sparse Voxel Fields (NSVF)

- Avoid sampling points in empty space as much as possible.

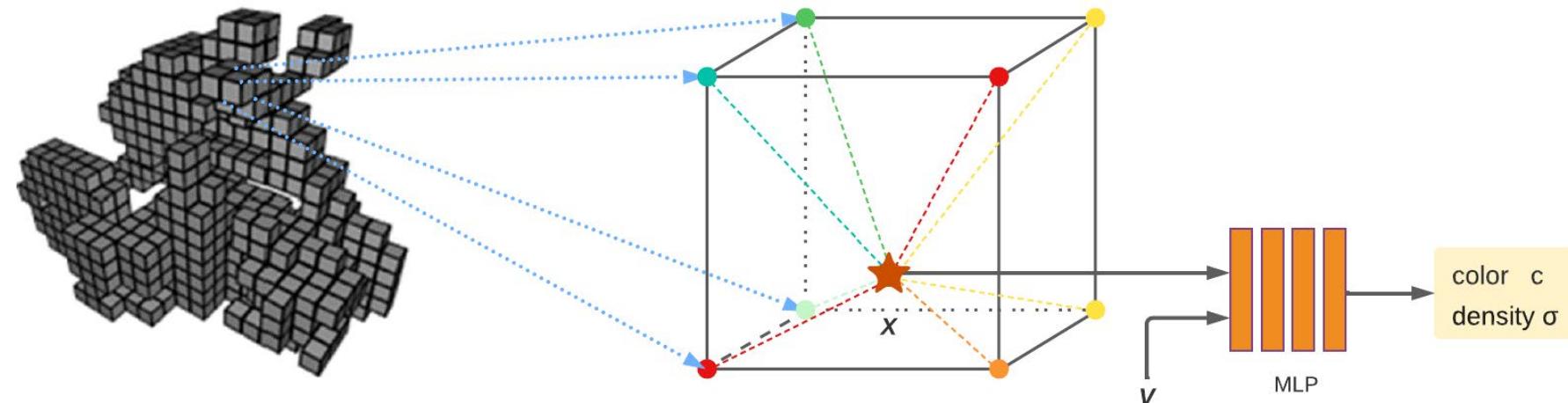


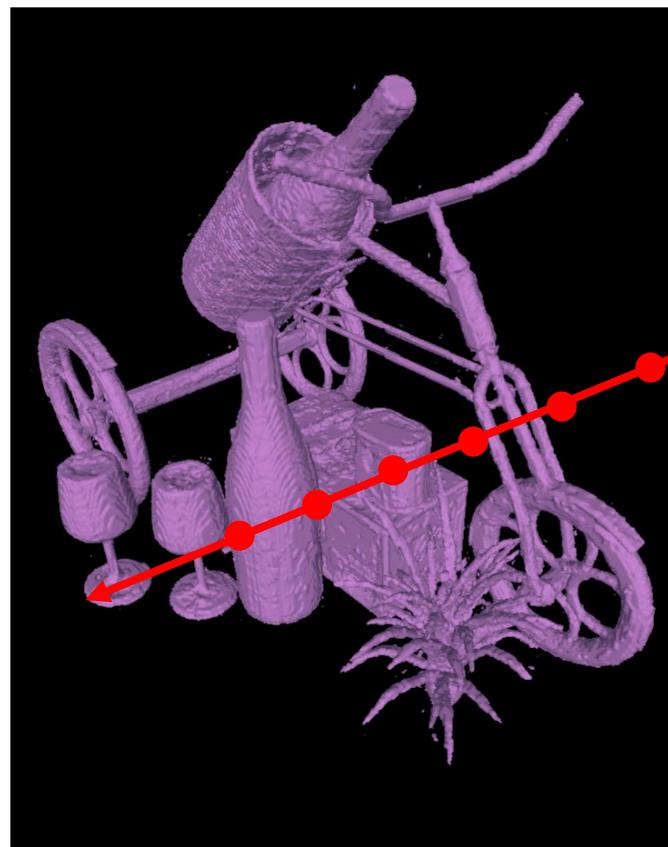
Illustration of Sparse Voxels

Illustration of a voxel-bounded neural field

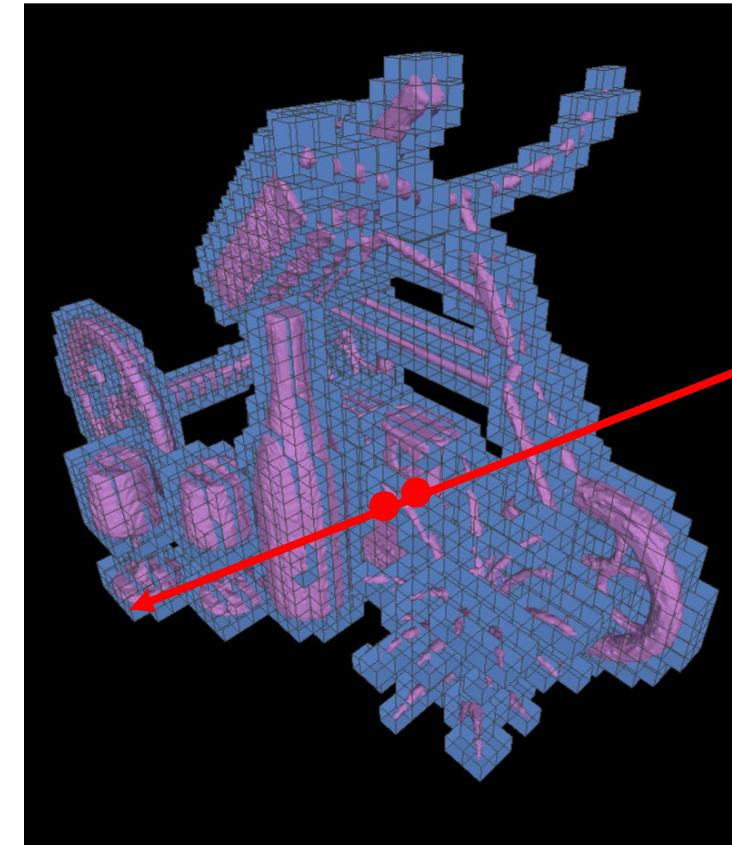
Neural Sparse Voxel Fields, Liu et al. 2020

Neural Sparse Voxel Fields (NSVF)

- Avoid sampling points in empty space as much as possible.



Sample in the whole space



Only sample inside the sparse-voxels

Comparison



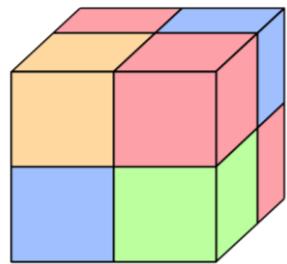
NeRF (Mildenhall et al. 2020)
(Rendering speed: 100 s/frame)



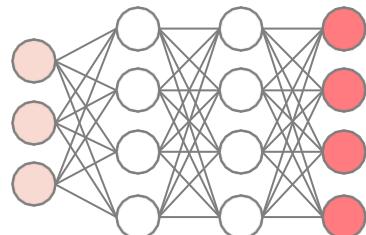
Ours (NSVF)
(Rendering speed: 2.62 s/frame)

Requirements

Scene
Representation



Voxelgrids



Implicit Function

Renderer

Volumetric

Sphere-Tracing
Volumetric

Pros

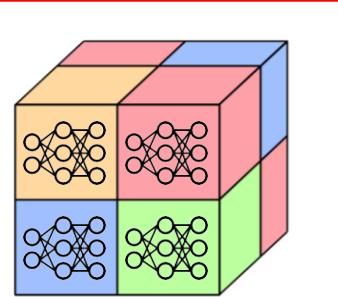
Fast rendering

High quality
Compact
Admits *global* priors

Cons

Memory $O(n^3)$
Limited spatial
resolution

Extremely expensive,
slow rendering



Hybrid
Implicit/Explicit

Volumetric

Significant Speedup
Admits *local* priors

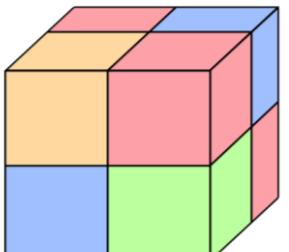
No compact
representation
No *global* priors

• • •

Neural Scene Representation and Neural Rendering

Neural Fields

Scene
Representation



Voxelgrids

Renderer

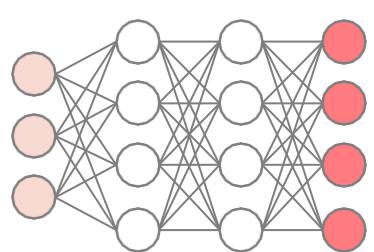
Volumetric

Pros

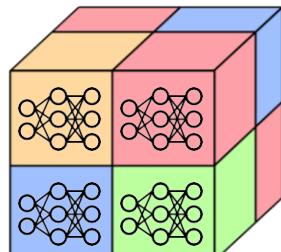
Fast rendering

Cons

Memory $O(n^3)$
Limited spatial
resolution



Implicit Function



Hybrid
Implicit/Explicit

Volumetric

Sphere-Tracing
Volumetric

High quality
Compact
Admits *global* priors

Extremely expensive,
slow rendering

Significant Speedup
Admits *local* priors

No compact
representation
No *global* priors

EUROGRAPHICS 2022
D. Meneva and G. Patané
(Guest Editors)

Volume 41 (2022), Number 2
STAR – State of The Art Report

Neural Fields in Visual Computing and Beyond

Yiheng Xie^{1,2} Towaki Takikawa^{3,4} Shunsuke Saito⁵ Or Litany⁶ Shiqin Yan⁷ Numair Khan¹ Federico Tombari^{6,7}
James Tompkin⁸ Vincent Sitzmann^{9†} Srinath Sridhar¹⁰

¹Brown University ²Unity Technologies ³University of Toronto ⁴NVIDIA ⁵Meta Reality Labs Research ⁶Google ⁷Technical University of Munich
⁸Massachusetts Institute of Technology ⁹Equal advising

<https://neuralfields.cs.brown.edu/>

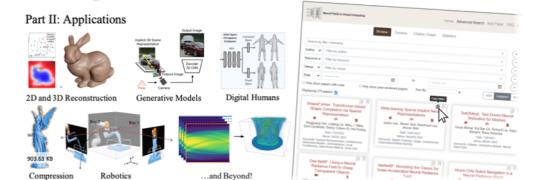
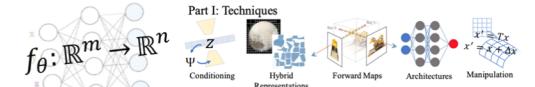


Figure 1: Contribution of this report. Following a survey of over 250 papers, we provide a review of (Part I) techniques in neural fields such as prior learning and conditioning, representations, forward maps, architectures, and manipulation, and of (Part II) applications in visual computing including 2D image processing, 3D scene reconstruction, generative modeling, digital humans, compression, robotics, and beyond. This report is complemented by a [community-driven website](#) with search, filtering, bibliographic, and visualization features.

Abstract

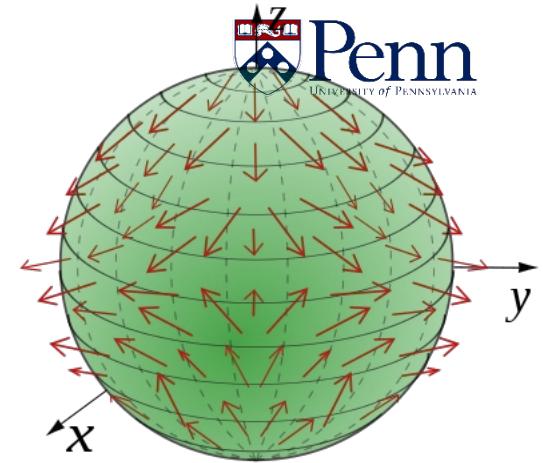
Recent advances in machine learning have led to increased interest in solving visual computing problems using methods that employ coordinate-based neural networks. These methods, which we call *neural fields*, parameterize physical properties of scenes or objects across space and time. They have seen widespread success in problems such as 3D shape and image synthesis, animation of human bodies, 3D reconstruction, and pose estimation. Rapid progress has led to numerous papers, but a consolidation of the discovered knowledge has not yet emerged. We provide context, mathematical grounding, and a review of over 250 papers in the literature on neural fields. In Part I, we focus on neural field techniques by identifying common components of neural field methods, including different conditioning, representation, forward map, architecture, and manipulation methods. In Part II, we focus on applications of neural fields to different problems in visual computing, and beyond (e.g., robotics, audio). Our review shows the breadth of topics already covered in visual computing, both historical and in current incarnations, and highlights the improved quality, flexibility, and capability brought by neural field methods. Finally, we present a [companion website](#) that acts as a living database that can be continually updated by the community.

CCS Concepts

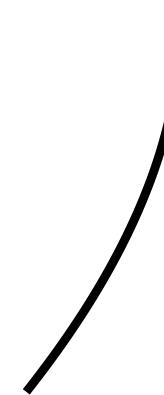
• Computing methodologies → Machine Learning; Artificial Intelligence;

A *field* is a quantity defined for all spatial and / or temporal coordinates.

Examples of Fields

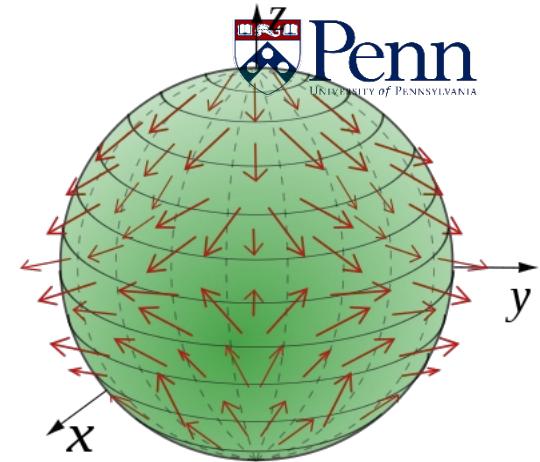
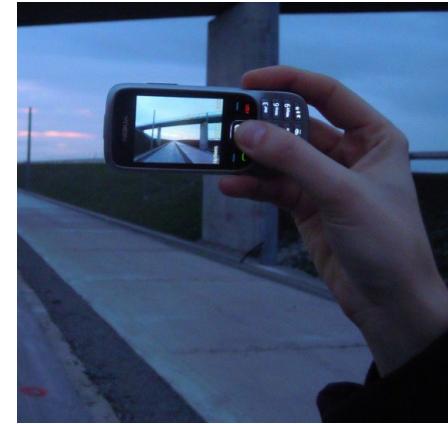


Vector Field



Fields

Examples of Fields

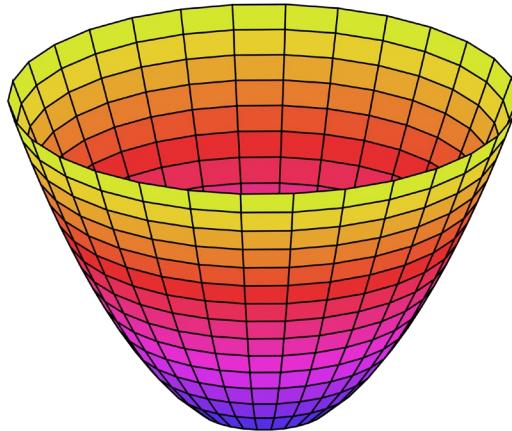


Image

Vector Field

Fields

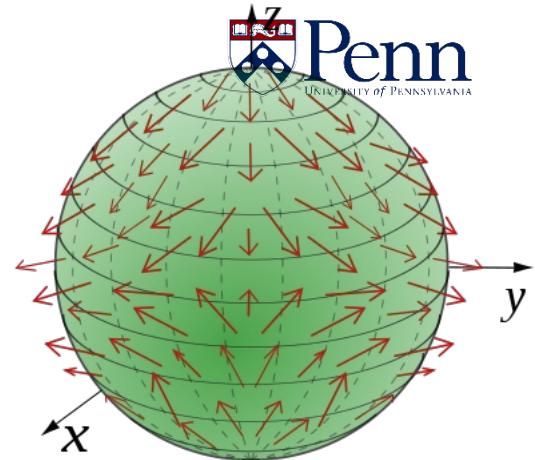
Examples of Fields



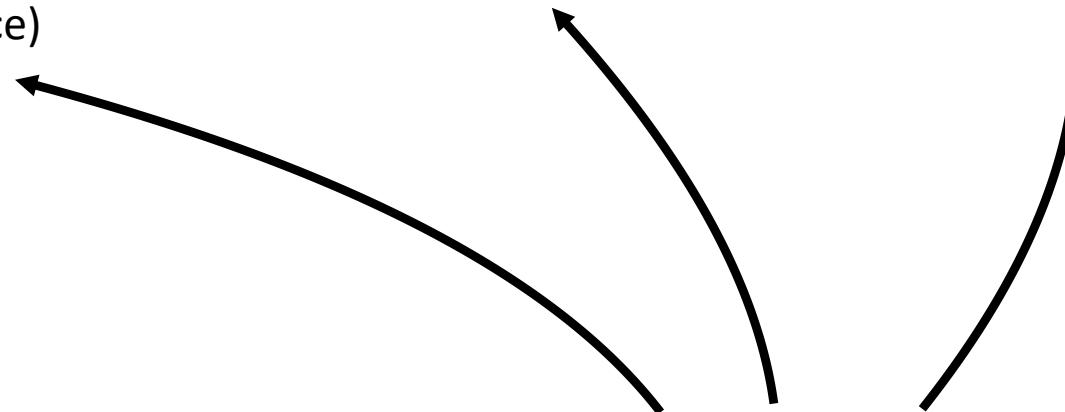
3D Parabola
(Explicit Surface)



Image

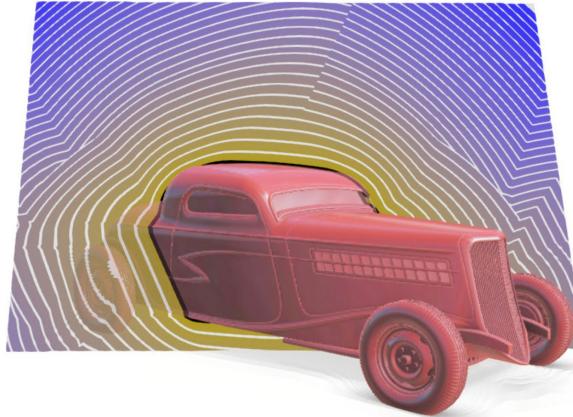


Vector Field

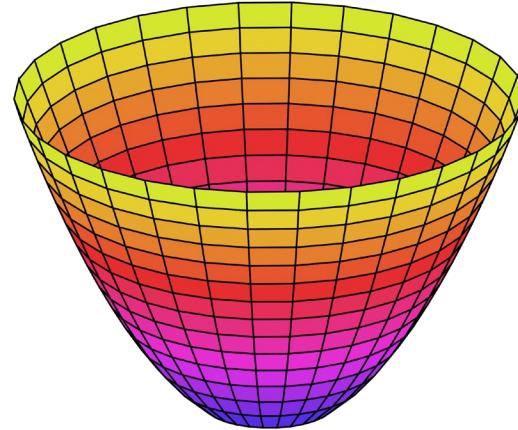


Fields

Examples of Fields



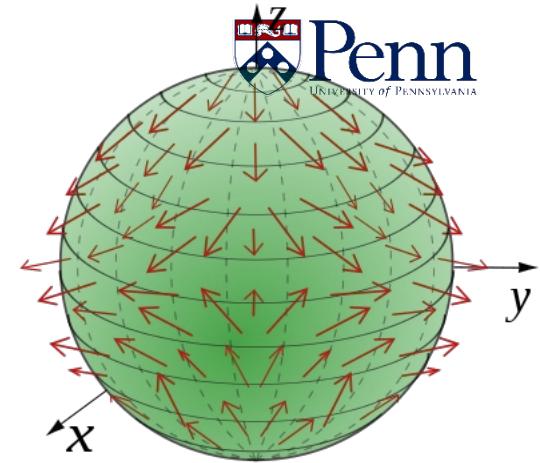
3D Signed Distance Fields
(Implicit Surface)



3D Parabola
(Explicit Surface)



Image



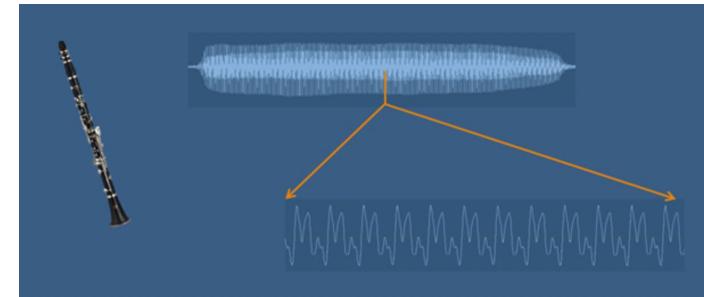
Vector Field

Fields

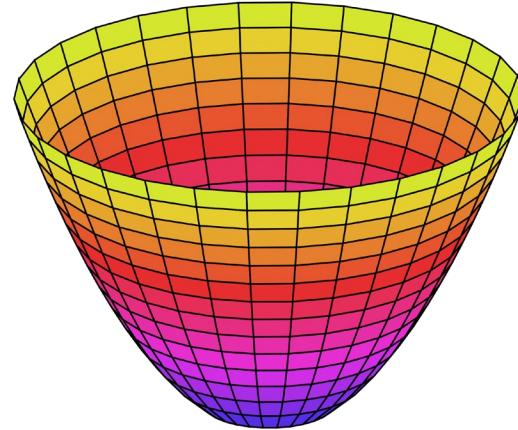
Examples of Fields



3D Signed Distance Fields
(Implicit Surface)



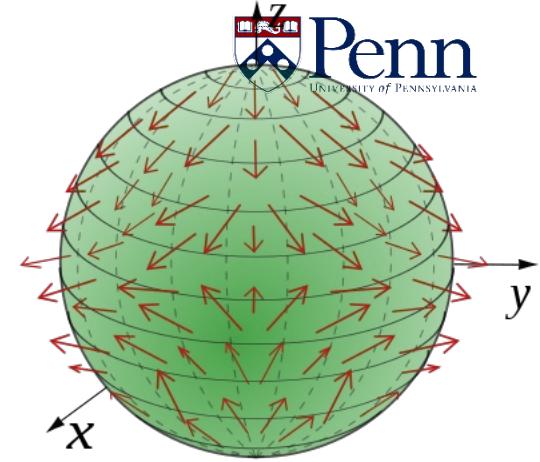
Audio



3D Parabola
(Explicit Surface)



Image

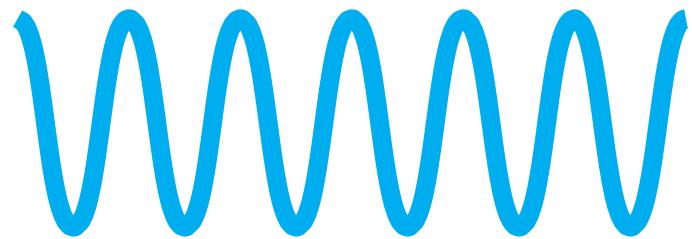


Vector Field

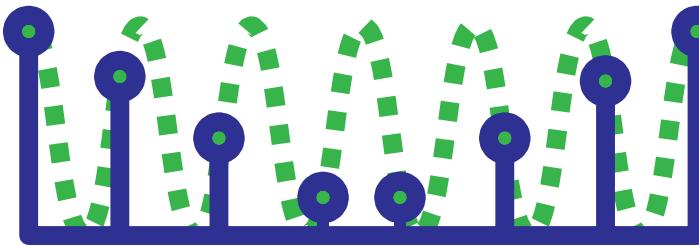
Fields

What are neural fields?

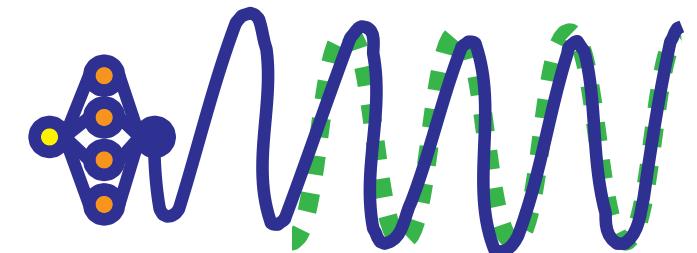
Fields / signals can be represented in many ways.



Continuous

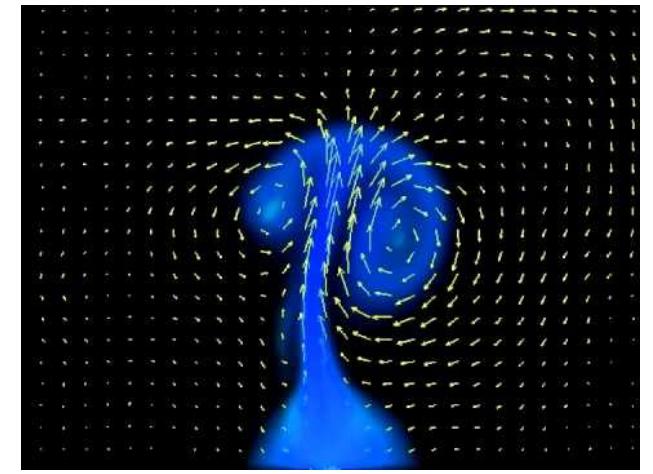
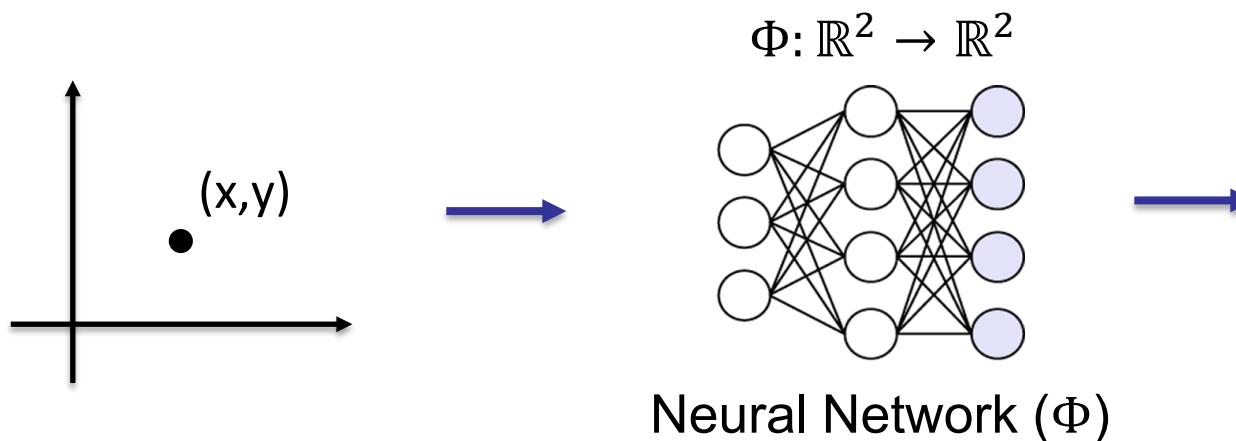
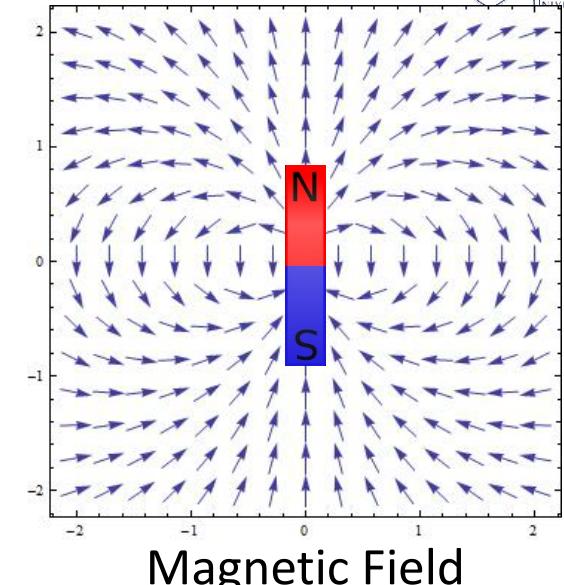
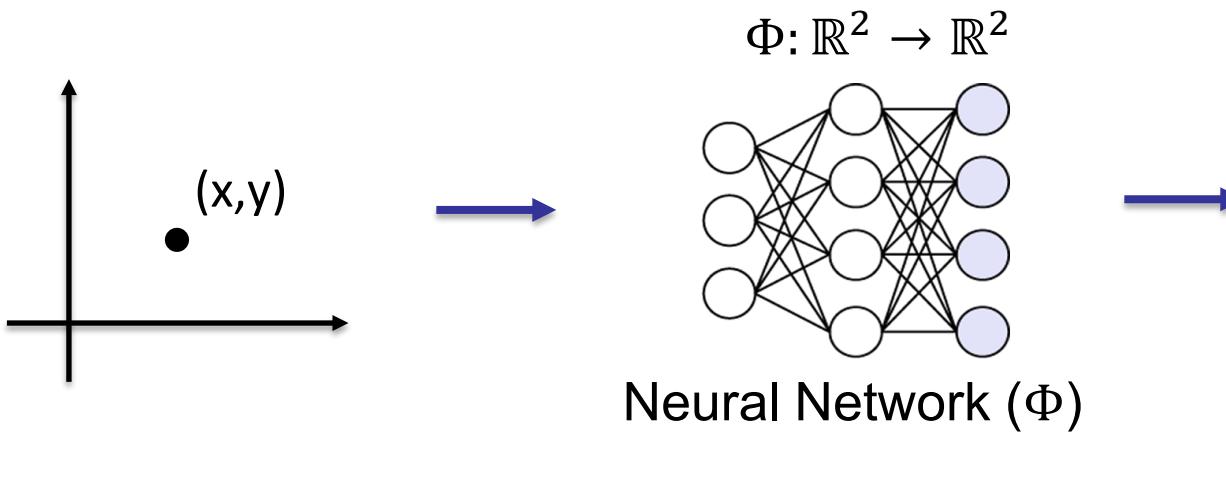


Discrete



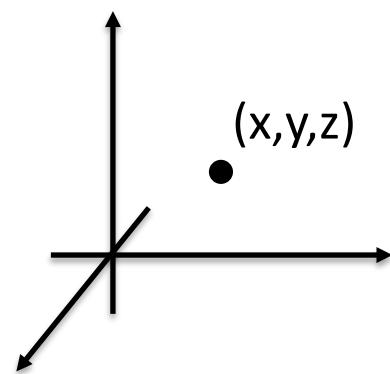
Neural

What are neural fields?

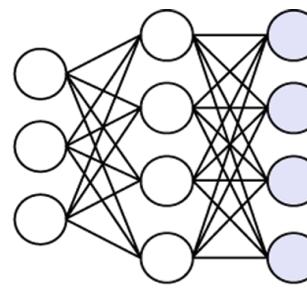


Eulerian Flow Field of a Fluid
[Koldora CC]

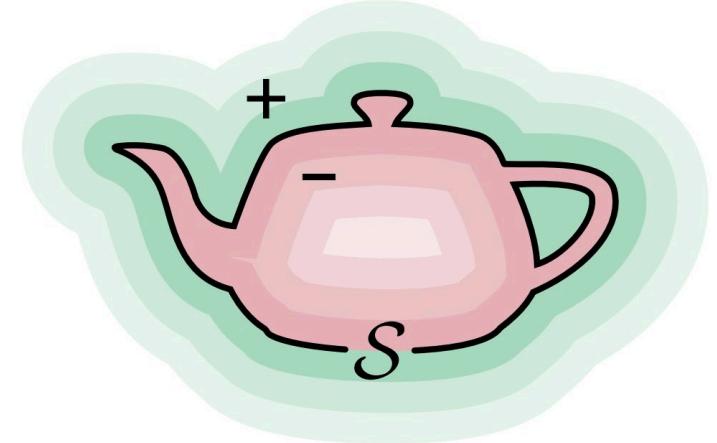
What are neural fields?



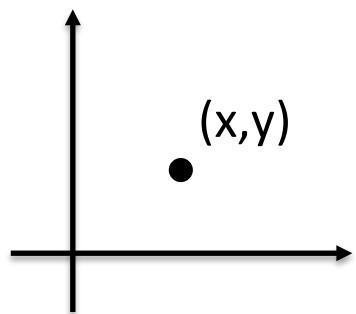
$$\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$$



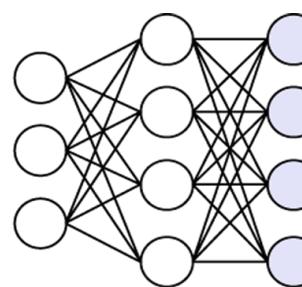
Neural Network (Φ)



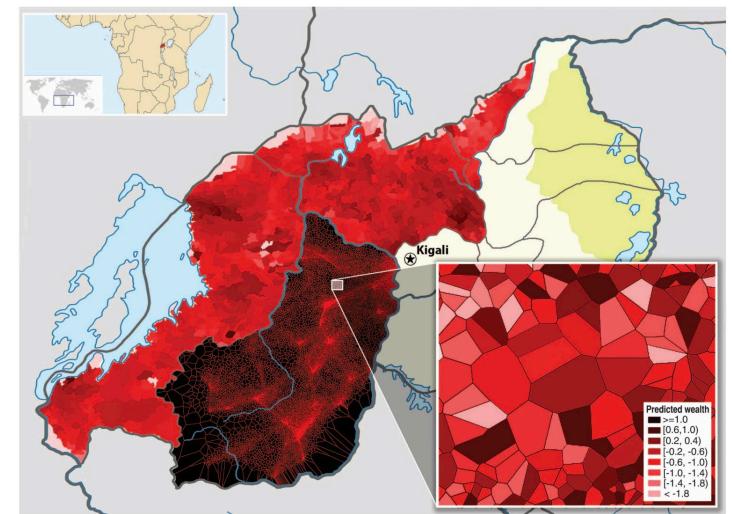
Signed Distance Function (SDF)



$$\Phi: \mathbb{R}^2 \rightarrow \mathbb{R}^n$$



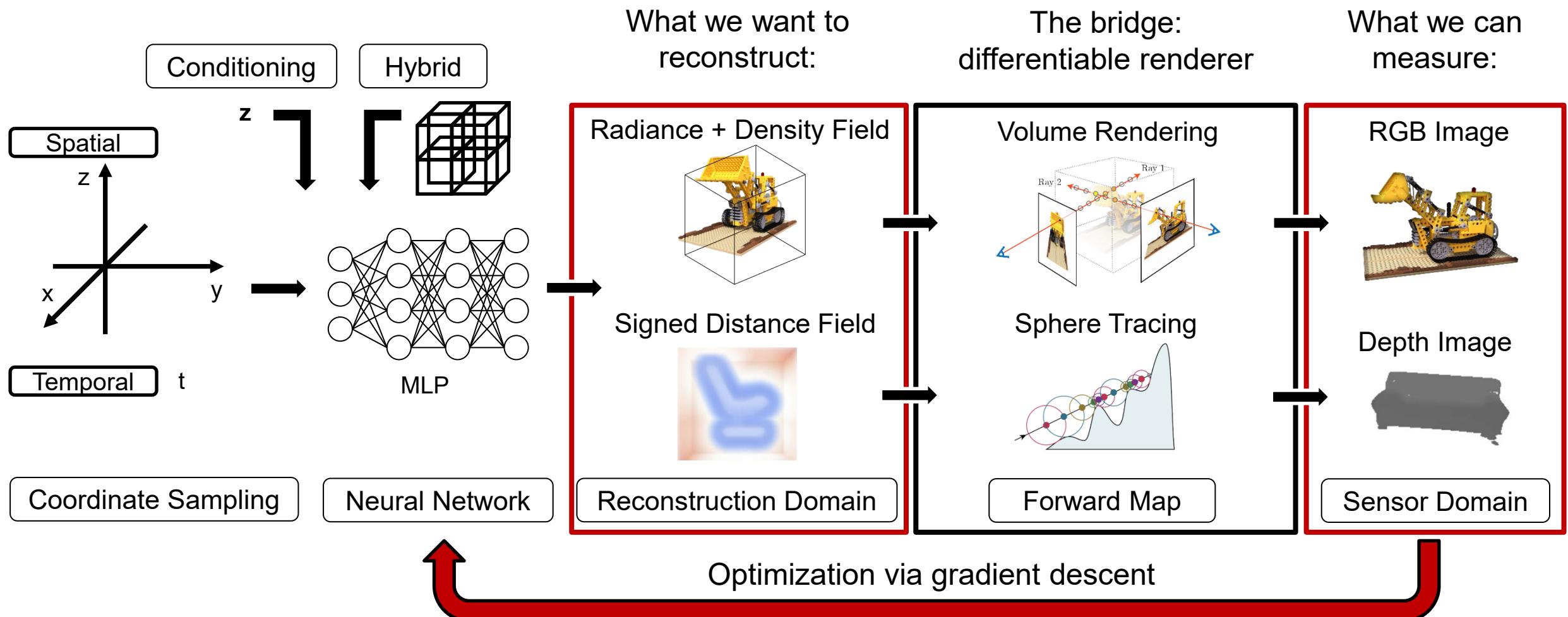
Neural Network (Φ)



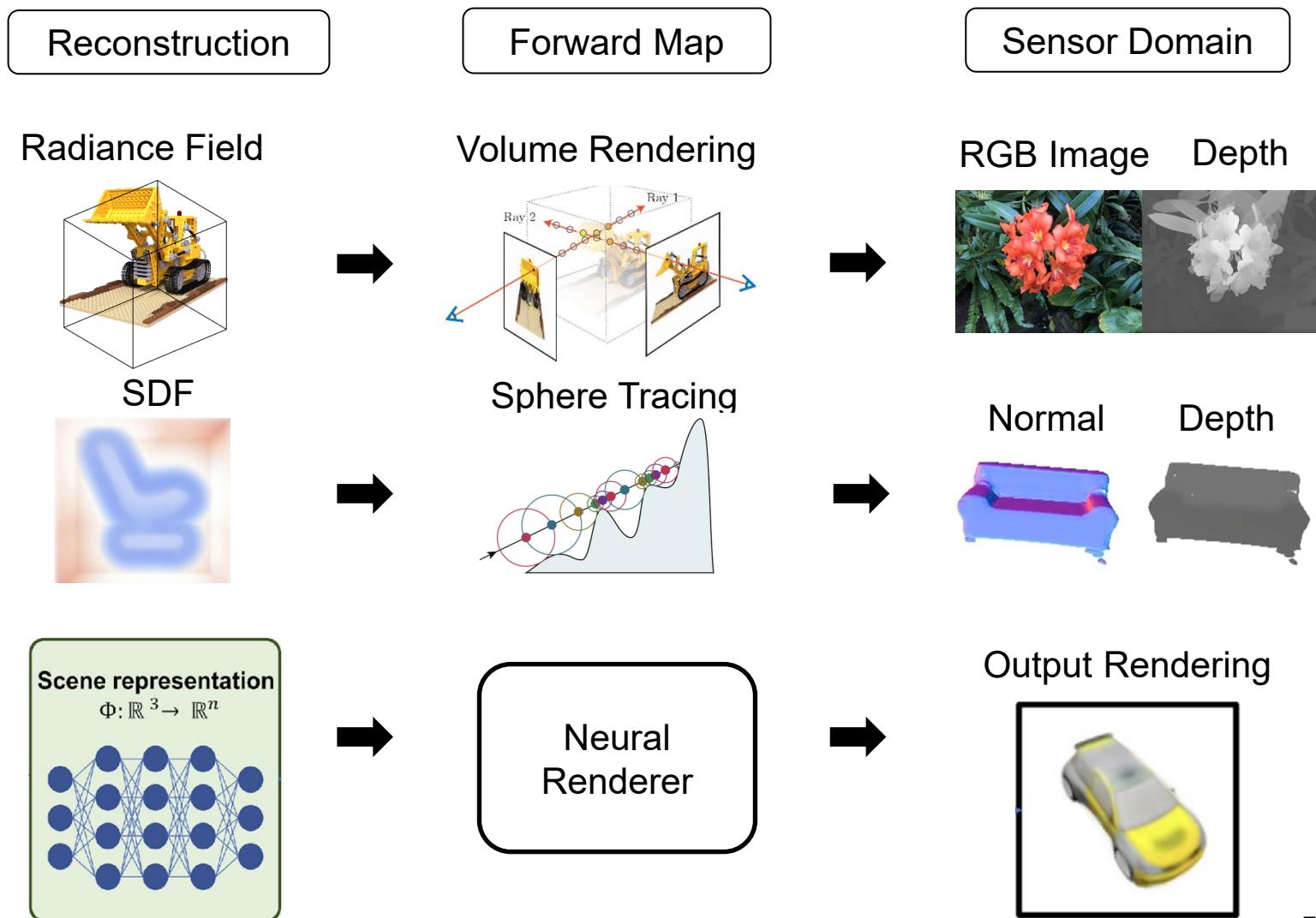
Geospatial Data
[Blumenstock et al. 2015]

Lingjie Liu

Neural Fields General Framework

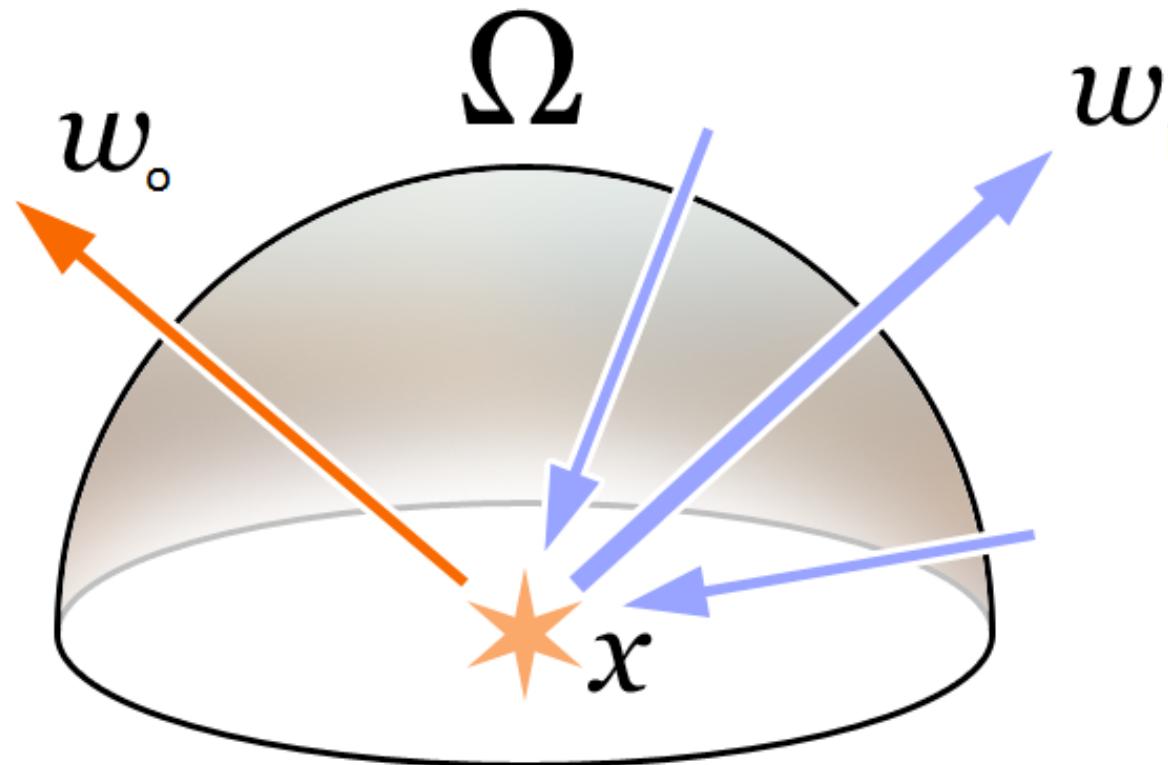


Differentiable Rendering



Figures adapted from:
Mildenhall et al. 2020 (NeRF)
Sitzmann et al. 2019 (SRN)
Lingjie Liu

BRDF Shading



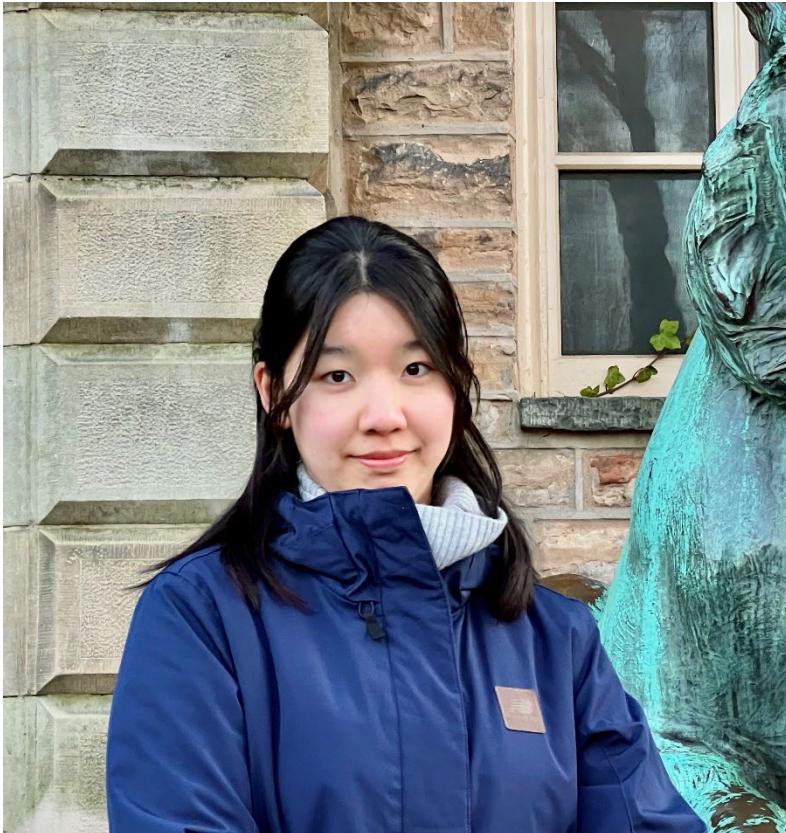
$$L(\mathbf{x}, \vec{\omega}_o) = L_e(\mathbf{x}, \vec{\omega}_o) + \int_s f_r(\mathbf{x}, \vec{\omega}_i \rightarrow \vec{\omega}_o) L(\mathbf{x}', \vec{\omega}_i) G(\mathbf{x}, \mathbf{x}') V(\mathbf{x}, \mathbf{x}') d\omega_i$$

Course Link:

<https://neural-representation-2024.github.io/topics.html>



TAs



Xuyi Meng
Email: mengxuyi@seas.upenn.edu



Chuhao Chen
Email: morphling233@gmail.com

Preliminary Syllabus

No.	Date	Content
1	Aug 28 (Wed)	Intro
2	Sept 4 (Wed)	Intro 2
3 - 12	Sept 9 (Mon) – Oct 9 (Wed)	Paper Presentations (round 1)
13 - 16	Oct 14 (Mon) – Oct 23 (Wed)	Guest Talks
17 – 26	Oct 28 (Mon) – Nov 27 (Wed)	Paper Presentations (round 2)
27	Dec 2 (Mon)	Practice lecture (e.g., NerfStudio)
28	Dec 4 (Wed)	Discussion on your favorite papers in Neural Representation and Neural Rendering (5 mins per person)
29	Dec 9 (Mon)	Summary + Brainstorming new ideas

Next Class

1. Present some pioneering works in this field, e.g., NeRF, SRN, Neural Volumes, ...
2. Fundamentals of Classical 3D Representations and Rendering in Computer Graphics

Topic and Papers

Fast Inference

BakedSDF: Meshing Neural SDFs for Real-Time View Synthesis

Yariv et al.

SIGGRAPH 2023

3D Gaussian Splatting for Real-Time Radiance Field Rendering

Kerbl et al.

SIGGRAPH 2023 (Best Paper Award)

2D Gaussian Splatting for Geometrically Accurate Radiance Fields

Huang et al.

SIGGRAPH 2024

Fast Training

Instant Neural Graphics Primitives with a Multiresolution Hash Encoding

Müller et al.

ACM TOG 2022

TensoRF: Tensorial Radiance Fields

Chen and Xu et al.

ECCV 2022

+ Factor Fields: A Unified Framework for Neural Fields and Beyond

Chen et al.

SIGGRAPH 2023

Antialiasing

Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields

Barron et al.

ICCV 2021 (Oral, Best Paper Honorable Mention)

+ Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields

Barron et al.

CVPR 2022 (Oral Presentation)

+ Zip-NeRF: Anti-Aliased Grid-Based Neural Radiance Fields

Barron et al.

ICCV 2023 (Oral Presentation, Best Paper Finalist)

Mip-NeRF v.s. Mip-NeRF 360 v.s. Zip-NeRF:

Common: Address the aliasing artifacts of NeRF.

Mip-NeRF: Mitigates aliasing artifacts at different resolutions by replacing point sampling with Gaussian sampling.

Mip-NeRF 360: Extends Mip-NeRF to unbounded scenes using a non-linear scene parameterization to allocate appropriate capacity for foreground and background.

Zip-NeRF: Addresses z-aliasing artifacts from Mip-NeRF 360's resampling and adapts to an efficient grid representation using multisampling within a conical frustum.

Mip-Splatting: Alias-free 3D Gaussian Splatting

Yu et al.

CVPR 2024 (Best Student Paper Finalist)

Note: For a paper bundle, you only need to present one of the papers in the bundle according to their preference, but you are encouraged to discuss the connections between the papers in the bundle.



Large (Unbounded) Scenes

MERF: Memory-Efficient Radiance Fields for Real-time View Synthesis in Unbounded Scenes

Reiser *et al.*

SIGGRAPH 2023

+ SMERF: Streamable Memory Efficient Radiance Fields for Real-Time Large-Scene Exploration

Duckworth and Hedman *et al.*

SIGGRAPH 2024 (Best Paper Honorable Mention)

MERF v.s. SMERF:

Common: Use compact representation to achieve high-quality real-time volumetric rendering.

MERF: Proposed a combination of a low-resolution 3D grid and a set of higher-resolution 2D planes.

SMERF: Supports real-time rendering on mobile devices; dedicates each viewpoint a MERF for large scenes.

Grid-guided Neural Radiance Fields for Large Urban Scenes

Xu *et al.*

CVPR 2023

Generalization

pixelNeRF Neural Radiance Fields from One or Few Images

Yu *et al.*

CVPR 2021

PixelSplat: 3D Gaussian Splats from Image Pairs for Scalable Generalizable 3D Reconstruction

Charatan *et al.*

CVPR 2024 (Oral)

(infers a 3D Gaussian scene from two input views in a single forward pass.)

LRM: Large Reconstruction Model for Single Image to 3D

Hong *et al.*

ICLR 2024 (Oral)

3D Generative Model

[Per-scene optimization: diffusion distillation]

DreamFusion: Text-to-3d using 2D diffusion

Poole et al.

ICLR 2023

+ **ProlificDreamer: High-Fidelity and Diverse Text-to-3D Generation with Variational Score Distillation**

Wang et al.

NeurIPS 2023 (Spotlight)

[Single-view image → Multi-view image → 3D reconstruction]

Cat3D: Create Anything in 3D with Multi-View Diffusion Models

Gao et al.

arXiv 2024

InstantMesh: Efficient 3D Mesh Generation from a Single Image with Sparse-view Large Reconstruction Models

Xu et al.

arXiv 2024

+ **LGM: Large Multi-View Gaussian Model for High-Resolution 3D Content Creation**

Tang et al.

ECCV 2024 (Oral)

+ **One-2-3-45++: Fast Single Image to 3D Objects with Consistent Multi-View Generation and 3D Diffusion**

Liu et al.

CVPR 2024

[Pose-free 3D Generation]

PF-LRM: Pose-Free Large Reconstruction Model for Joint Pose and Shape Prediction

Wang et al.

arXiv 2024

+ **SpaRP: Fast 3D Object Reconstruction and Pose Estimation from Sparse Views**

Xu et al.

ECCV 2024

PF-LRM v.s. SpaRP:

Common: 3D reconstruction from sparse unknown-posed images.

PF-LRM: Explicit matching through pointcloud + differentiable PnP solver.

SpaRP: Distill stable diffusion model to predict NOCS images for camera pose estimation.

[Native 3D Generation]

Splatter Image: Ultra-Fast Single-View 3D Reconstruction

Szymanowicz et al.

CVPR 2024

[Multi-view ImageNet]

EG3D: Efficient Geometry-aware 3D Generative Adversarial Networks

Chan et al.

CVPR 2022

3D generation on ImageNet

Skorokhodov et al.

ICLR 2023 (Oral)

Dynamic Scenes & Human

Shape of Motion: 4D Reconstruction from a Single Video

Wang et al.
arXiv 2024

+ MoSca: Dynamic Gaussian Fusion from Casual Videos via 4D Motion Scaffolds

Li et al.
arXiv 2024

K-Planes: Explicit Radiance Fields in Space, Time, and Appearance

Fridovich-Keil et al.
CVPR 2023

4K4D: Real-Time 4D View Synthesis at 4K Resolution

Xu et al.
CVPR 2024

Pose Estimation

COLMAP-Free 3D Gaussian Splatting

Fu et al.
CVPR 2024

Local-to-Global FlowCam: Training Generalizable 3D Radiance Fields without Camera Poses via Pixel-Aligned Scene Flow

Smith et al.
NeurIPS 2023

Lighting

TensoIR: Tensorial Inverse Rendering

Jin et al.
CVPR 2023

Relightable 3D Gaussian: Real-time Point Cloud Relighting with BRDF Decomposition and Ray Tracing

Zhang et al.
ECCV 2024

Physics Simulation

PhysGaussian: Physics-Integrated 3D Gaussians for Generative Dynamics

Xie et al.
CVPR 2024 (Highlight)

PhysAvatar: Learning the Physics of Dressed 3D Avatars from Visual Observations

Zheng et al.
ECCV 2024

Editing & Multi-modality

Instruct-NeRF2NeRF: Editing 3D Scenes with Instructions

Haque et al.
ICCV 2023 (Oral)

PlatoNeRF: 3D Reconstruction in Plato's Cave via Single-View Two-Bounce Lidar

Klinghoffer et al.
CVPR 2024 (Oral, Best Paper Award Finalist)

Robotics

LERF: Language Embedded Radiance Fields

Kerr *et al.*

ICCV 2023 (Oral)

+ LERF-TOGO: Language Embedded Radiance Fields for Zero-Shot Task-Oriented Grasping

Rashid *et al.*

CORL 2023 (Best Paper Finalist)

LERF v.s. LERF-TOGO:

Common: Embed language embeddings into 3D scene representation.

LERF: Enables pixel-aligned zero-shot queries on the distilled 3D CLIP embedding.

LERF-TOGO: Extends LERF to task-oriented grasping by adding DINO feature grouping.

Unifying 3D Representation and Control of Diverse Robots with a Single Camera

Li *et al.*

arXiv 2024

Surface Reconstruction

NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction

Wang *et al.*

NeurIPS 2021

+ NeuS2: Fast Learning of Neural Implicit Surfaces for Multi-view Reconstruction

Wang *et al.*

ICCV 2023

Gaussian Opacity Fields: Efficient and Compact Surface Reconstruction in Unbounded Scenes

Yu *et al.*

arXiv 2024

Differentiable Mesh Extraction

NeurCross: A Self-Supervised Neural Approach for Representing Cross Fields in Quad Mesh Generation

Dong *et al.*

arXiv 2024

Flexible Isosurface Extraction for Gradient-Based Mesh Optimization

Shen *et al.*

SIGGRAPH 2023

Before the seminar

- Read the papers of the week.
- Submit at least two questions for discussion before the seminar to a Google form ([https://docs.google.com/forms/d/e/1FAIpQLSfSxryv JO9Ffbd7iKCIqnczqPWJUqv3OGFI6K-2sAKOJmBYQ/viewform](https://docs.google.com/forms/d/e/1FAIpQLSfSxryvJO9Ffbd7iKCIqnczqPWJUqv3OGFI6K-2sAKOJmBYQ/viewform)). This is important – your contribution will be marked. The deadline for submitting questions is one hour before each class session (so Monday 2:30 PM and Wednesday 2:30 PM).

During the seminar (Starting from Sept 9, two rounds)

- Overview (10 minutes)
 - The instructor or TAs give a brief introduction on the topic.
- 2x Presentations (each 25 minutes, 25 % of grade):
 - Two pre-assigned participants present the paper of their choice.
 - 5 minutes on motivation, background and related work.
 - 20 minutes of presentation of the paper.
- Discussion and Feedback (30 minutes, 25% of grade across weeks):
 - One participant is assigned at random at the beginning of the seminar to lead the discussion. Everyone leads the discussion at least once in the seminar series.
 - The discussion leader receives a digest of the submitted questions just before the seminar.
 - The discussion leader raises questions appropriately throughout the discussion, covers future work aspects, and finally provides a summary of the strengths and weaknesses of the techniques and of the discipline.
 - The students provide feedback to the presenting student on their presentation with respect to what has worked well, and what could be improved and how.

Grading Criteria

Form (30%) <i>To time? Verbal speed & clarity? Body posture? Engagement with audience?</i>	Moderation (30%) <i>Integrates questions well? Pushes forward discussion? Good summary? Strengths and weaknesses of paper?</i>	Practice Lecture (30%) <i>Listen attentively to the lecture? Engage in small coding exercises?</i>
Content (50%) <i>Structure/storyline? Main points? Paper connections? Valid conclusions?</i>	Questions (70%) <i>One question per paper (two questions per class) should be submitted at least one hour before the class during which the paper will be presented. (However, students are permitted to submit questions late (but before the discussion), up to two occurrences, without facing any penalties)</i>	Discussion on your favorite papers (35%) <i>Each person has 5 minutes to present their favorite paper. Is your presentation clear? Explain clearly why you chose it and what you like about it?</i>
Answers (20%) <i>Good answers to questions? Knowledgeable?</i>		Brainstorming (35%) <i>Actively participate in the discussion? Contribute your own ideas or opinions?</i>

2x Presentations
(50% of grade)

Discussion
(25% of grade)

Other Activity Participation
(25% of grade)

TODOs After this Class

1. **Paper Selection and Registration:** [Important! Deadline: Sept 3]

Please select and register for the two papers you would like to present using the following Excel link:

<https://docs.google.com/spreadsheets/d/1FJueXqWnKWoYOGZTNiwp2qRmSP0u1H6ayEYE5j3lb0/edit?gid=0#gid=0>

2. **Presentation Preparation:**

- Ensure you are fully prepared **one class before your scheduled class for presentation.**
- Upload your slides to the Google folder (<https://drive.google.com/drive/folders/1NO-JdWIRtKiLGZOMQxCOUs0AjdtrypY>) **at least one hour before the class prior to your assigned class for presentation.** This is important in case of an emergency requiring us to reschedule your talk.
- For example, if you're presenting on Monday, upload your slides by the previous Wednesday at 2:30 PM. If presenting on Wednesday, upload by Monday at 2:30 PM.



3. **Class Participation:**

- Before each class, please read the papers that will be discussed and submit two questions **at least one hour before the class** using the following link:

<https://docs.google.com/forms/d/e/1FAIpQLSfSxryvJO9Ffdb7iKCIqnczqPWJUqv3OGFI6K-2sAKOJmBYQ/viewform>

Acknowledgments

- Advances in Neural Rendering
- Neural Fields in Visual Computing and Beyond
- awesome-NeRF: a curated list of awesome neural radiance fields papers
- MPII Summer Semester 2023: Computer Vision and Machine Learning for Computer Graphics

Any Questions?