HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

3D Scanning with NeRF for Movie Makers

Bora Kunter Sahin

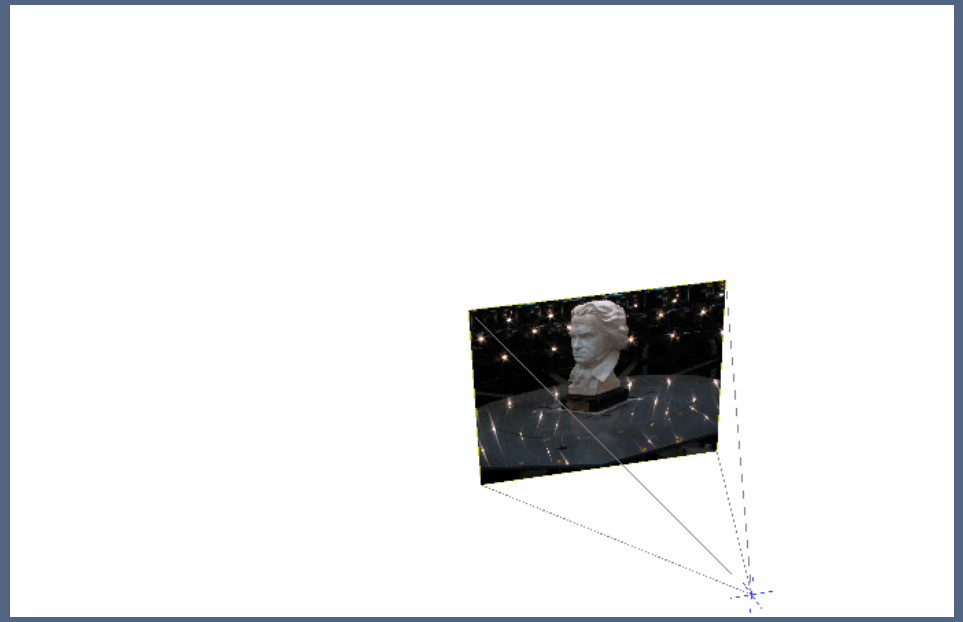
History

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

- Phases of development
  - Multi-View Photogrammetry(or so-called structure-from-motion)
  - Hardware scanning with depth sensors
  - AI-based 3D scanning(e.g. NeRF)

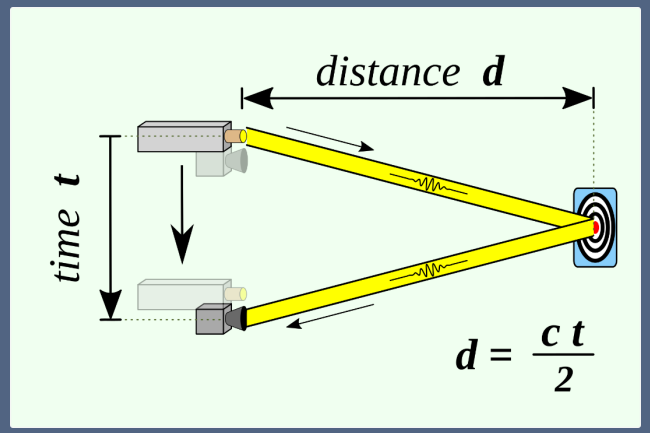
Structure-from-Motion

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

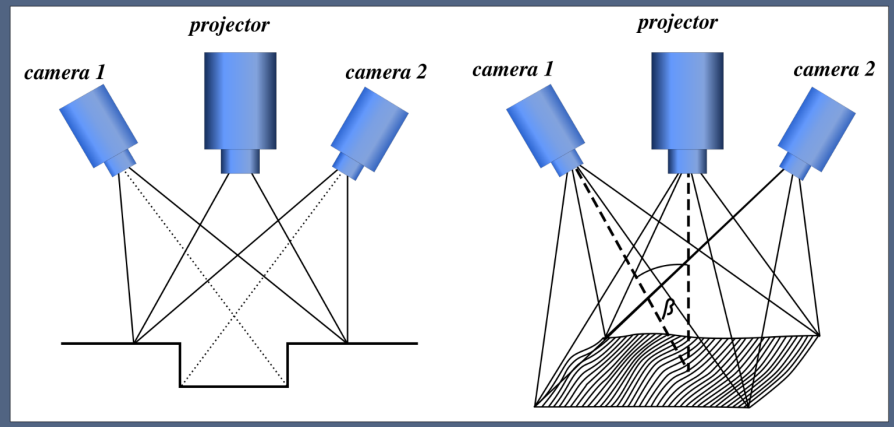


Source: https://cvg.cit.tum.de/research/image-based_3d_reconstruction/multiviewreconstruction

Depth Sensor Technology



Source: https://en.wikipedia.org/wiki/Lidar



Source: https://en.wikipedia.org/wiki/Structured-light_3D_scanner

AI-based Scanning

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

"If traditional 3D representations like polygonal meshes are akin to vector images, NeRFs are like bitmap images: they densely capture the way light radiates from an object or within a scene,"

-David Luebke, NVIDIA

Source: https://blogs.nvidia.com/blog/2022/03/25/instant-nerf-research-3d-ai/

Why NeRF?

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

- Relighting of the scene
- Fine details are captured
- Reflections can be handled correctly
- Transparency is handled correctly(e.g. glass surfaces)

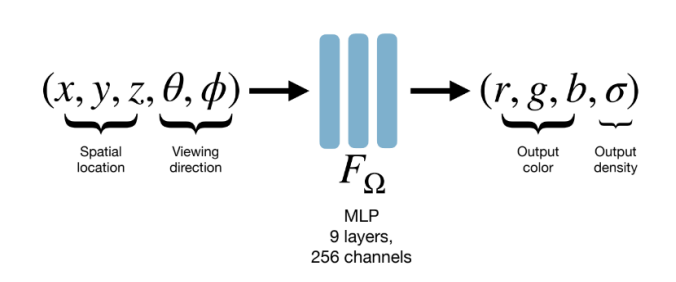
How does NeRF work?

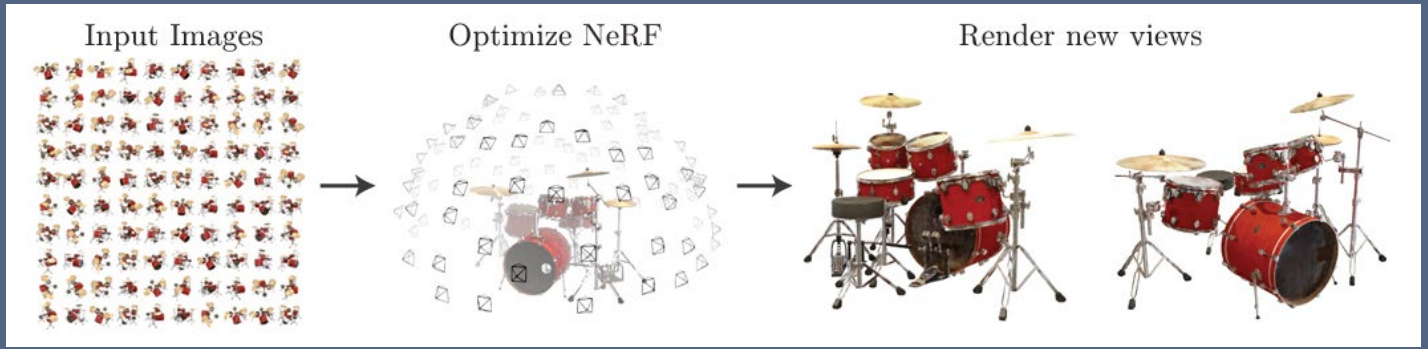
HOCHSCHULE FÜR FERNSEHEN UND FILM MÜNCHEN

1. For each pixel, march camera rays through scene to gather a set of samples at (**x, d**) locations.
2. Use (**x, d**) points and viewing directions at each sample as input to produce output (c,σ) values (essentially rgbσ).
3. Construct an image using classical volume rendering techniques.

Source: https://towardsdatascience.com/its-nerf-from-nothing-build-a-vanilla-nerf-with-pytorch-7846e4c45666

"NeRFs are a volumetric representation encoded into a neural network. They are not 3D meshes and they are not voxels. **For each point in space the NeRF represents a view dependent radiance.** More concretely each point has a density which describes how transparent or opaque a point in space is. They also have a view dependent color that changes depending on the angle the point is viewed."



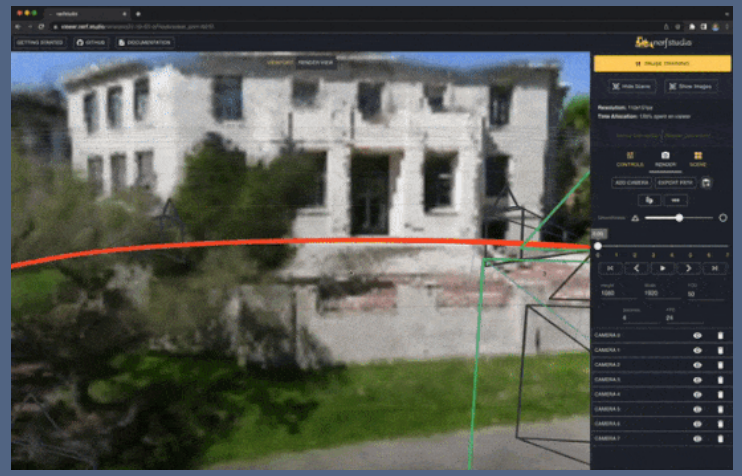


Source: https://arxiv.org/pdf/2003.08934.pdf

Tools

Instant Neural Graphics Primitives

nerfstudio



- “Nerfstudio provides a simple API that allows for a simplified end-to-end process of creating, training, and visualizing NeRFs”

Supported Methods

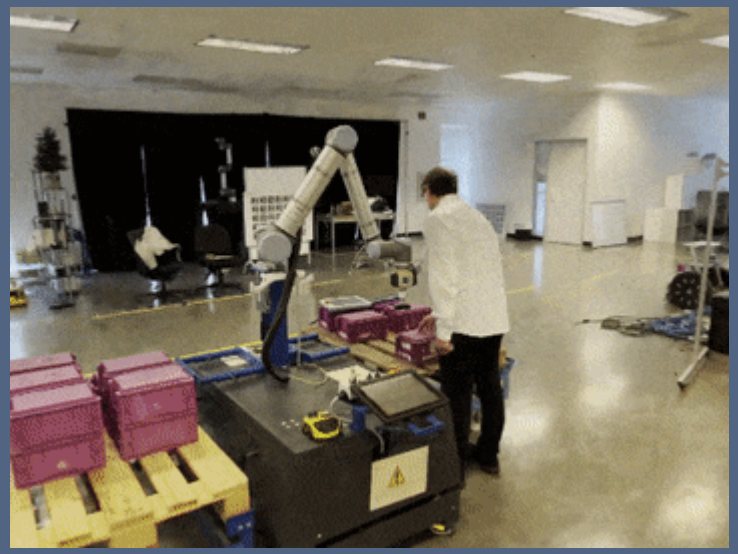
- **Nerfacto**: our de facto NeRF method combines modules focused on quality with modules focused on faster rendering. Nerfstudio easily lets us experiment with the best of both worlds!
- NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis
- Instant NGP: Instant Neural Graphics Primitives with a Multiresolution Hash Encoding
- Mipnerf: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields
- NerfW: Neural Radiance Fields for Unconstrained Photo Collections
- Semantic NeRF: In-Place Scene Labelling and Understanding with Implicit Scene Representation

Installation

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

- Easier using a Docker image
- https://docs.nerf.studio/en/latest/quickstart/installation.html

Instant NeRF



- “In all tasks, our encoding and its efficient implementation provide clear benefits: rapid training, high quality, and simplicity”

Comparison of NeRF Methods

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

- NeRF: default, slow training and inference times, not recommended(„For most tasks, using the original NeRF model is likely not a good choice and hence we provide implementations of various other NeRF related models…”)
- InstantNeRF: reduced computational complexity, faster training times
- mipNeRF: is better at representing fine details, more eligible for smaller objects to capture details
- NeRFW: better for complex environments, corrections to local illumination properties
- Nerfacto: new method of nerfstudio combining the methods above

How to get good results?

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

If any of the following assumptions are broken, the reconstructions may fail completely or contain artifacts such as excess geometry.

- Camera poses are known
- Scene is static, objects do not move
- The scene appearance is constant (ie. exposure doesn't change)
- Dense input capture (Each point in the scene should be visible in multiple images)

Source: https://docs.nerf.studio/en/latest/nerfology/methods/nerf.html

Sample Result

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN

- Trained only for ***10 mins***
- Video: only 20 seconds

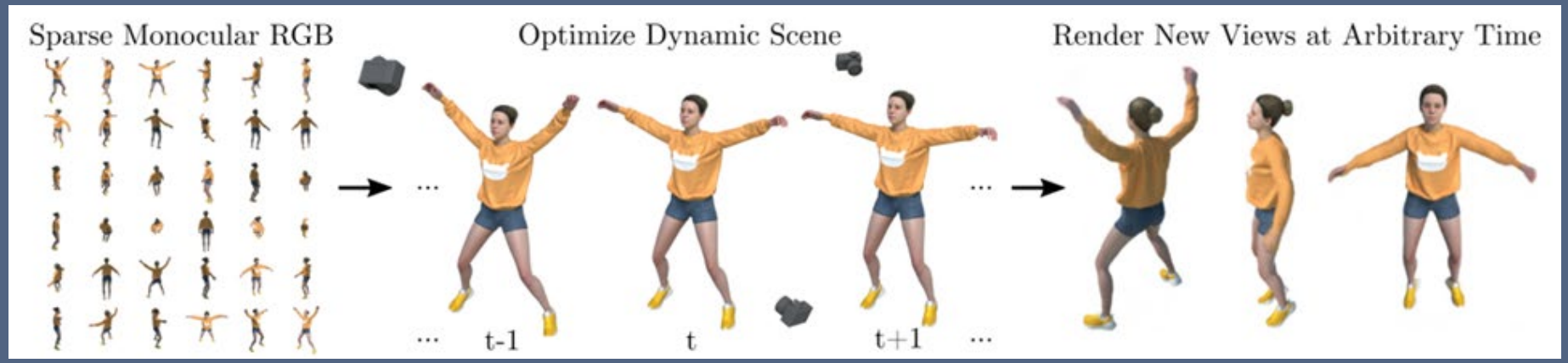
- Up until now, NeRF can be used for static scenes only...

NeRF for Dynamic Scenes

HOCHSCHULE FÜR FERNSEHEN UND FILM MÜNCHEN

- We need to handle the time aspect correctly

D-NeRF


Sparse Monocular RGB
Optimize Dynamic Scene
Render New Views at Arbitrary Time
t-1
t
t+1

TiNeuVox

HOCHSCHULE
FÜR FERNSEHEN UND
FILM MÜNCHEN