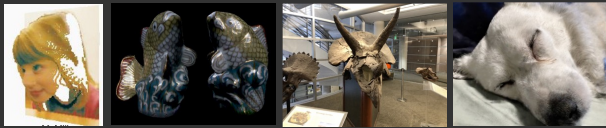


Image-Based Rendering

CSE 274, Lecture 1: Keynote and Logistics

Ravi Ramamoorthi
<http://www.cs.ucsd.edu/~ravir>



1

Instructor

Ravi Ramamoorthi <http://www.cs.ucsd.edu/~ravir>

- PhD Stanford, 2002 [with Pat Hanrahan, 2020 Turing Award]
- "Spherical Harmonic Lighting" widely used in games (e.g. Halo series), movies (e.g. Avatar), etc. (Adobe, ...)
- At Columbia 2002-2008, UC Berkeley 2009-2014
- "Monte Carlo denoising" inspired raytracing offline, real-time; consults at NVIDIA
- At UCSD since Jul 2014: Director, Center for Visual Computing
- "NeRF: Neural Radiance Fields for View Synthesis" widely cited IBR technique
- Awards for research: White House PECASE (2008), SIGGRAPH Significant New Researcher (2007), ACM Fellow (2018), Two Frontiers Science Awards (23,24)
- <https://www.youtube.com/watch?v=qpyCXqXGe7I>
- Computer Graphics online MOOC (CSE 167x) finalist for two edX Prizes.

2

Outline of Lecture

- Lecture: Image-Based Rendering: From View Synthesis to Neural Radiance Fields and Beyond (originally Eurographics Keynote Apr 24)
- Logistics of course

3

Virtual Experiences of Real-World Scenes



4

Input Images

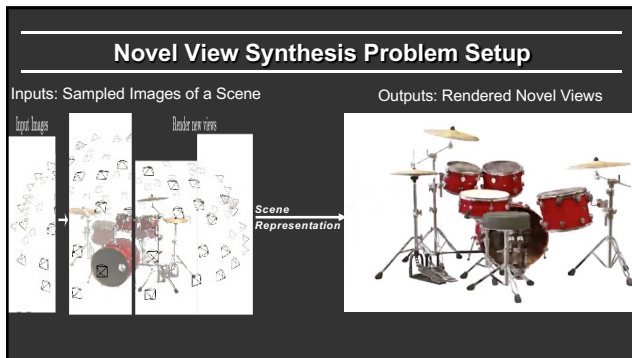


5

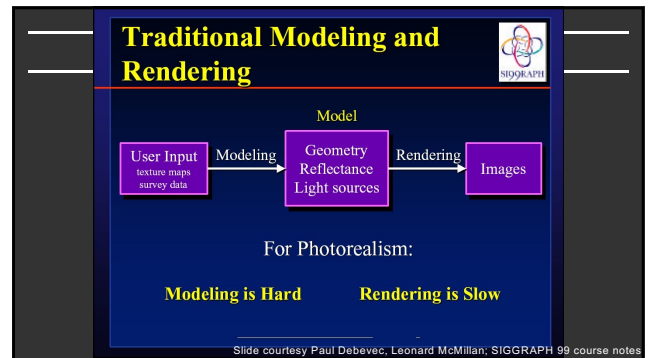
Output Virtual Experience



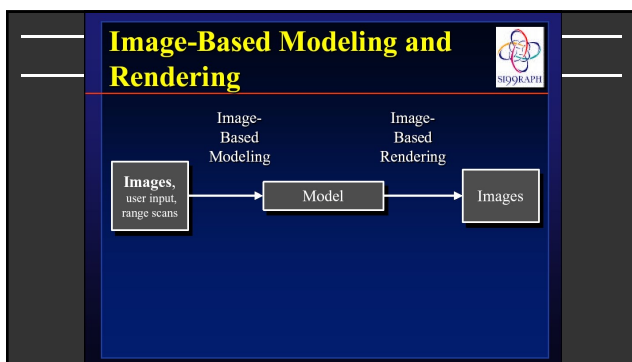
6



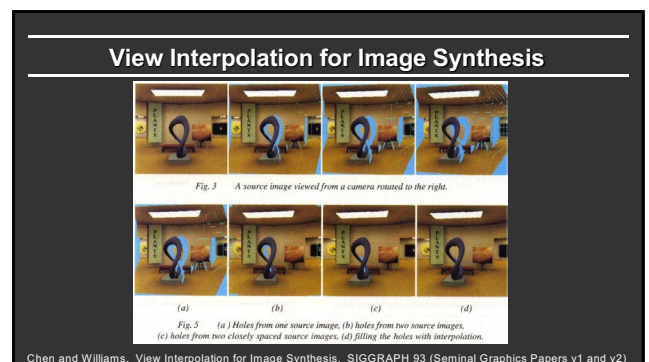
7



8



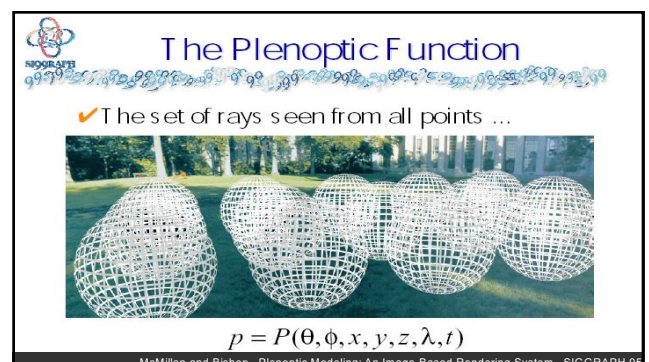
9



10





11



12

Light Fields


Gershun's and Moon's idea of a light field: 
Radiance as a function of a ray or line: $L(x, y, z, \theta, \phi)$

- In "free space" (no occluders) 5D reduces to 4D
 - Exterior of the convex hull of an object
 - Interior of an environment
- Images are 2D slices
 - Insert acquired imagery
 - Extract image from a given viewpoint 

From Levoy and Hanrahan, Light Field Rendering, SIGGRAPH 96



13

4D Light Field



14

Surface Light Fields and Reflectance Fields

Wood et al. Surface Light Fields for 3D Photography, SIGGRAPH 00
 Debevec et al. Acquiring the Reflectance Field of a Human Face, SIGGRAPH 00

15

Seminal Graphics Papers: Pushing the Boundaries

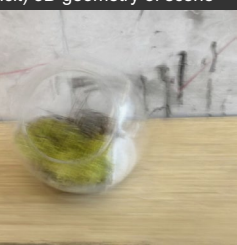
SIGGRAPH AWARD WINNERS (SNR, CGA, COONS)	
6. Imaging and Vision	Article 419
View Interpolation for Image Synthesis	Article 45:423-432
Plausibility Modeling: An Image-Based Rendering System	Article 46:433-440
Light Field Rendering	Article 47:441-452
The Lumigraph	Article 48:453-464
Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach	Article 49:465-476
Acquiring the Reflectance Field of a Human Face	Article 50:475-486
Surface Light Fields for 3D Photography	Article 51:487-496
Unstructured Lumigraph Rendering	Article 52:497-508
Fast Separation of Direct and Global Components of a Scene Using High-Frequency Illumination	Article 53:505-514
Photo Tourism: Exploring Photo Collections in 3D	Article 54:515-526

- Williams: Coons 03
- (an interesting story)
- Levoy CGA96; Hanrahan: CGA93, Coons03, Turing20
- Gortler: SNR 02; Szeliski: CGA 11; Cohen: CGA 98, Coons 19
- Debevec: SNR 01; (Malik: NAE 11)
- Debevec: SNR 01
- Salesin: CGA 00
- Gortler: SNR 02; Cohen: CGA 98, Coons 19
- Raskar: CGA 17; (Nayar: NAE 08)
- Snavely: SNR 14; Szeliski: CGA 11

16

View Synthesis till mid 2010s: Classical Geometry

- Reconstruct (implicit) 3D geometry of scene

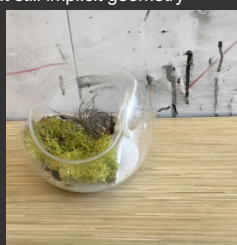


Penner and Zhang 2017

17

View Synthesis in mid 2010s: Deep Learning

- Deep learning but still implicit geometry




Flynn et al. 2016, Kalantari et al. 2016

18

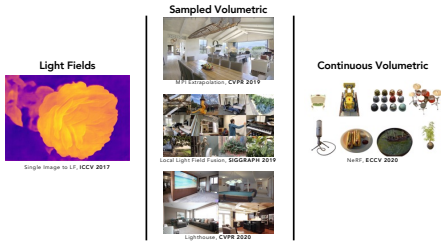
View Synthesis end 2010s: Deep Learning + MPIs

- Multi-Plane Images (Szeliski Golland 99, Zhou 18): Sampled Volume Representation



Mildenhall, Srinivasan et al. 2019

19




Slide courtesy Pratul Srinivasan

20

PREDICTING SCENE DEPTHS AND LIGHT FIELD FROM SINGLE IMAGE

- Input: single image




Synthesizing a 4D RGBD Light Field from a Single Image. Srinivasan et al. ICCV 2017

21

PREDICTING SCENE DEPTHS AND LIGHT FIELD FROM SINGLE IMAGE

- Output: dense multiview depths




Synthesizing a 4D RGBD Light Field from a Single Image. Srinivasan et al. ICCV 2017

22

PREDICTING SCENE DEPTHS AND LIGHT FIELD FROM SINGLE IMAGE

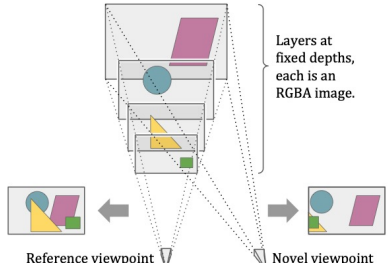
- Output: local light field



Synthesizing a 4D RGBD Light Field from a Single Image. Srinivasan et al. ICCV 2017

23

Representing scenes with frustum-sampled RGBA layers (a.k.a. Multi-plane Image or MPI)



Stereo Magnification. Zhou et al. 2018

24

Casual Capture for Light Field Synthesis



25

Promote each sampled view to local light field



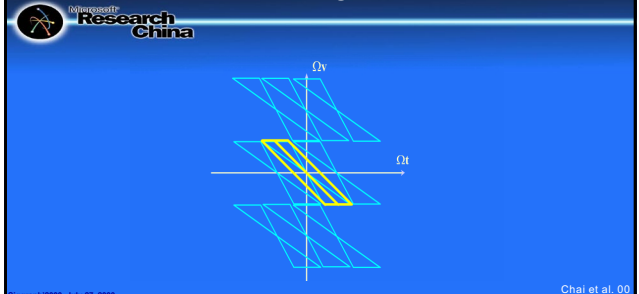
26

Blend nearby local light fields to render novel views



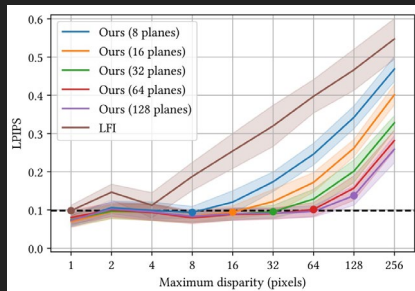
27

Light Field Reconstruction



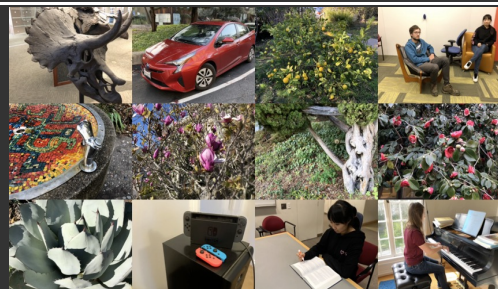
28

ML PIPELINE GETS US SURPRISINGLY CLOSE TO THEORETICAL BOUND

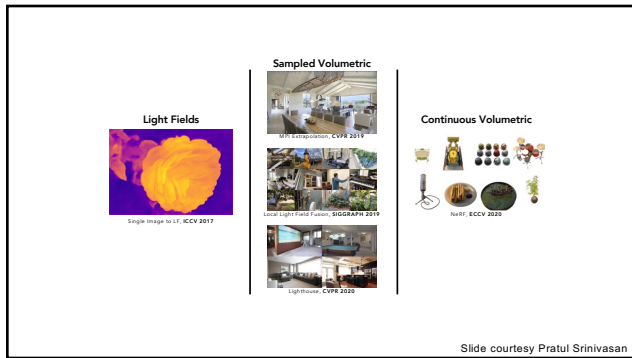


29

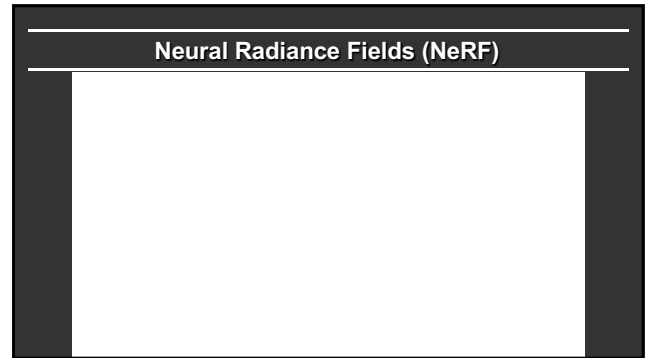
Local Light Field Fusion



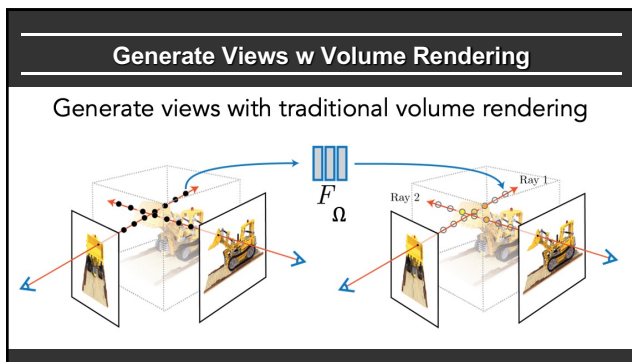
30



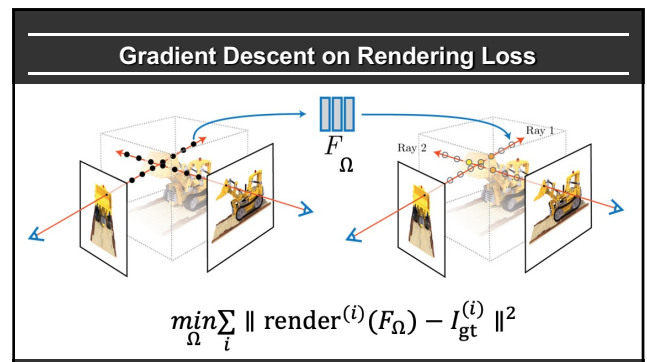
31



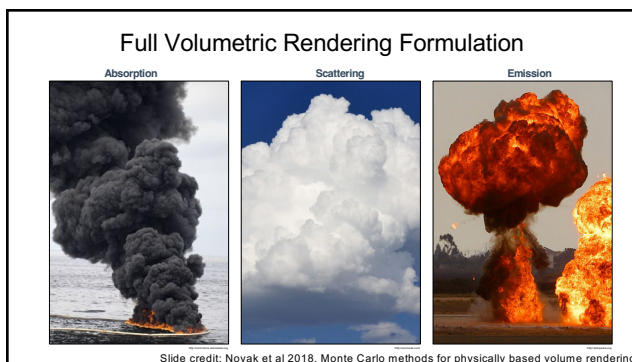
32



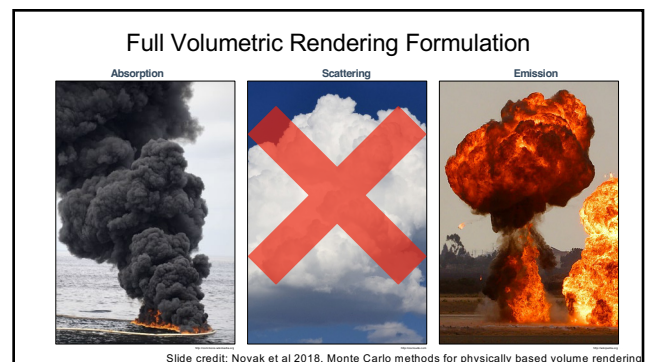
33



34



35



36

Summary: volume rendering integral estimate

Rendering model for ray $\mathbf{r}(t) = \mathbf{o} + t\mathbf{d}$:

$$\mathbf{c} \approx \sum_{i=1}^m T_i \alpha_i \mathbf{c}_i$$

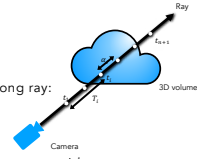
weights colors

How much light is blocked earlier along ray:

$$T_i = \prod_{j=1}^{i-1} (1 - \alpha_j)$$

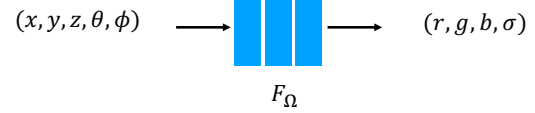
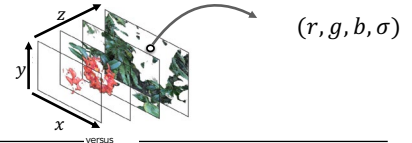
How much light is contributed by ray segment i :

$$\alpha_i = 1 - \exp(-\sigma_i \delta_i)$$



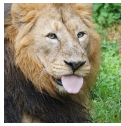
37

Use neural network to replace large N-d array



38

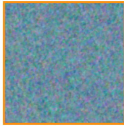
Preserve High-Frequency Features



Ground truth image



Neural network output without high frequency mapping



Neural network output with high frequency mapping

Fourier Features Let Networks Learn High Frequency Functions in Low Dimensional Domains. Tancik et al. Neurips 20

39

More detail than implicit scene MLP

SRN [Sitzmann et al. 2019]

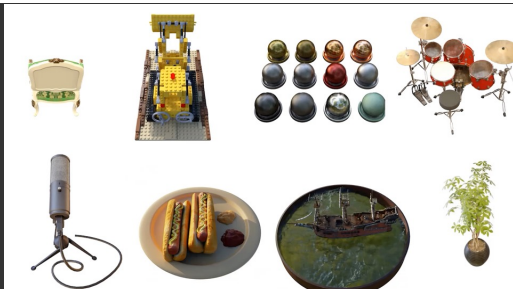
NeRF



Nearest Input

40

Results



41

View-Dependent w Directional Dependence




42

Detailed Scene Geometry, Occlusion


44

Detailed Scene Geometry, Occlusion

A photograph of a museum exhibit featuring a large, detailed model of a mammoth skull with prominent tusks, displayed on a pedestal. The scene includes other exhibits, glass railings, and a modern interior with large windows and stairs in the background.

45


NeRFs in the Metaverse

A photograph of a man with curly grey hair and glasses, wearing a light blue button-down shirt. He is shown in profile, looking towards the right. The background is a blurred office environment with modern lighting and equipment.

46

2022: NVIDIA Instant NGP — Real Time NeRFs

TIME Magazine Named NVIDIA Instant NeRF a Best Invention of 2022



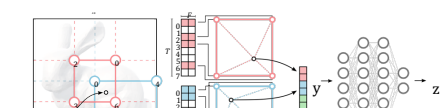
The original neural radiance fields paper from 2020 took many hours to train a model rendered at 0.03fps

With Instant NGP, NeRFs train in seconds, and render at real time rates (30 fps or faster)

Thomas Muller, Alex Evans, Christopher Schied, and Alex Keller 22

47


Multiresolution Hash Encoding



The diagram illustrates the Multiresolution Hash Encoding process. It starts with an input image x on the left. This image is processed by a series of operations represented by a grid of nodes and arrows. The process involves a sequence of operations, including a red path and a blue path, leading to a central point. The output of this process is a vector ξ , which is then used to generate a set of "Remaining inputs" (represented by a vertical stack of colored squares). These inputs are fed into a neural network structure (represented by a series of circles) to produce the final output z .

48

Reducing compute-per-sample: learned hash grids (Instant NGP)



Elapsed training time: 0 seconds

Müller, et al 2022, Instant Neural Graphics Primitives with a Multiresolution Hash Encoding; EG New Researcher 2024

49

3D Gaussian Splats for Radiance Fields



Kerbl, Kopanas, Leimkuhler, Drettakis 23; EG Distinguished Career Award 24

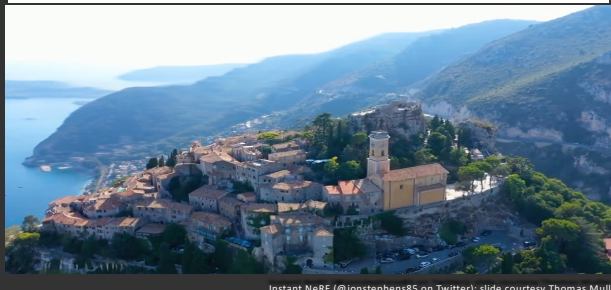
50

NeRFs for Digital Twins



51

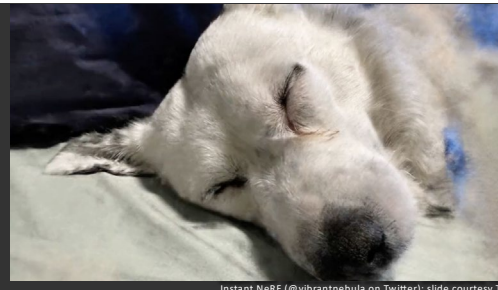
NeRFs for Landscapes



Instant NeRF (@jonsthenens85 on Twitter); slide courtesy Thomas Muller

52

NeRFs for Furry Volumes

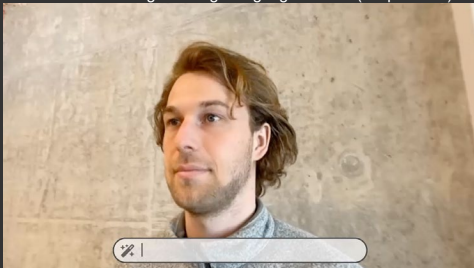


Instant NeRF (@vibrantnebula on Twitter); slide courtesy Thomas Muller

53

NeRFs + Language Models

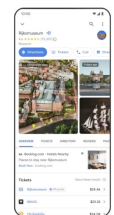
- Text-based editing with large language models (Haque et al.)



54

NeRFs in Production (Google, Luma)

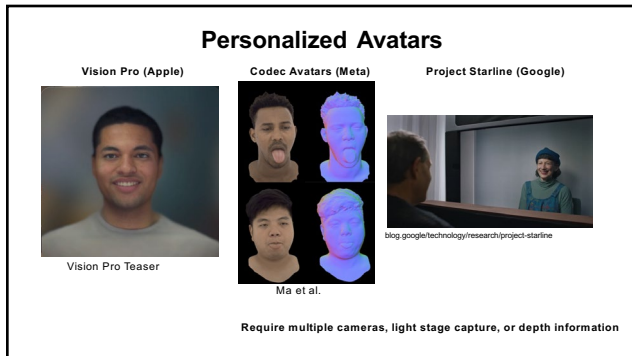
Google Street View



Luma AI iPhone app



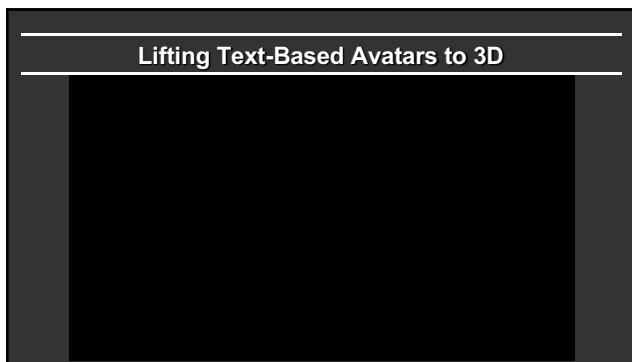
55



56



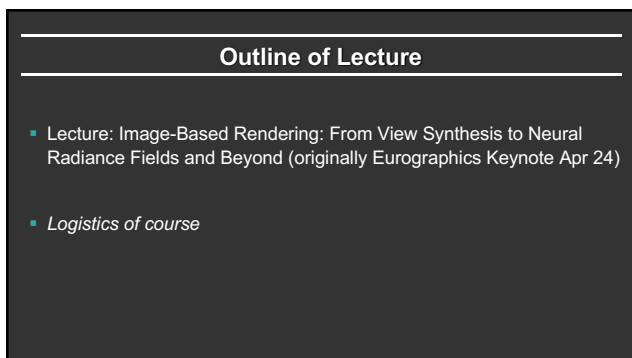
57



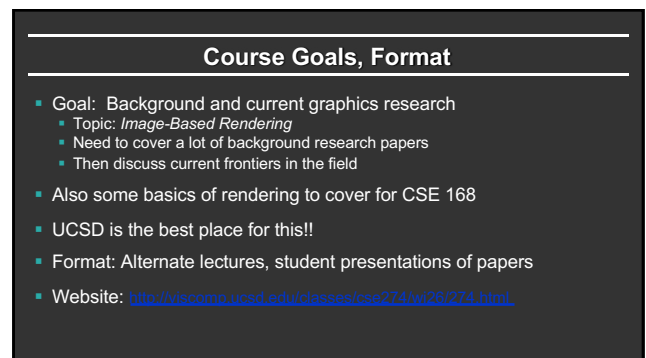
58



59



60



61

Course Logistics

- No textbooks. Required readings are papers online (and some handouts)
 - Handouts at <http://viscomp.ucsd.edu/classes/cse274w/26/readings>.
- Office hours: after class or email. My contact info is on my webpage: <http://www.cs.ucsd.edu/~ravir>.
- Zoom: <https://ucsd.zoom.us/j/ravir11> (if we need to move to remote instruction temporarily or permanently)
- TA: Nithin Raghavan n2raghavan@ucsd.edu. Office hours: see website or email for another time.
- Should count for PhD, MS, BS electives in graphics and vision, see me if there is a problem or you need a certification

62

Requirements

- Pass-Fail (2 units)
 - Show up to class regularly
 - Present 1 or 2 paper(s) if needed
 - Prefer you do this rather than just sit in
- Grades (4 units)
 - Attend class, participate in discussions
 - Present 1 or 2 papers (please do this well)
 - If class is large, groups of 2 can present 1 paper
 - Project (key part of grade)

63

Project

- Wide flexibility if related to course. Can be done groups of 2
 - Default: Implement (part of) one of papers and produce an impressive demo for image-based rendering or view synthesis
 - See/e-mail me re ideas
 - Best projects will go beyond simple implementation (try something new, some extensions)
- ML support only through basic DSMLP or AWS Educate through UCSD
- We may have some old GPUs but largely on own (or do non-ML project)
- Alternative (less desirable): Summary of 3+ papers in an area
 - Best projects will explore links/framework not discussed by authors, and suggest future research directions

64

Prerequisites

- Strong interest in graphics, vision, want to learn about image-based rendering
- Computer graphics experience (167 or equivalent)
 - Experience with rendering not required; first few weeks will cover basics. But also doesn't hurt, consider UCSD online CSE 168. And computer vision.
- Course will move quickly
 - Covering recent and current active research
 - Some material quite technical
 - Considerable background material is covered
 - Assume some basic knowledge
 - Many topics. Needn't fully follow each one, but doing so is most rewarding.

65

Assignment this week

- E-mail me (ravir@cs.ucsd.edu) [only if not done already]
 - Name, e-mail, status (Senior, PhD etc.)
 - Will you be taking course grades or P/F
 - Background in graphics/any special comments
 - *Optional for all: Papers you'd like to present FCFS (only those that say "presented by students")*

66

Questions?

67