

Image-Based Rendering

CSE 274, Lecture 8: Deep Learning for IBR

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

1

Local Light Field Fusion

Practical View Synthesis with Prescriptive Sampling Guidelines

Ben Mildenhall<sup>1</sup>

Pratul Srinivasan<sup>1</sup>

Rodrigo Ortiz-Cayon<sup>2</sup>

Nima Kalantari<sup>3</sup>

Ravi Ramamoorthi<sup>4</sup>

Ren Ng<sup>1</sup>

Abhishek Kar<sup>2</sup>

<sup>1</sup>Berkeley

UNIVERSITY OF CALIFORNIA

<sup>2</sup>FUSION

<sup>3</sup>TEXAS A&M

UNIVERSITY

<sup>4</sup>UC San Diego

Paper and code available at: [fyusion.com/LLFF](http://fyusion.com/LLFF)

2

Virtual Experiences of Real-World Scenes

3

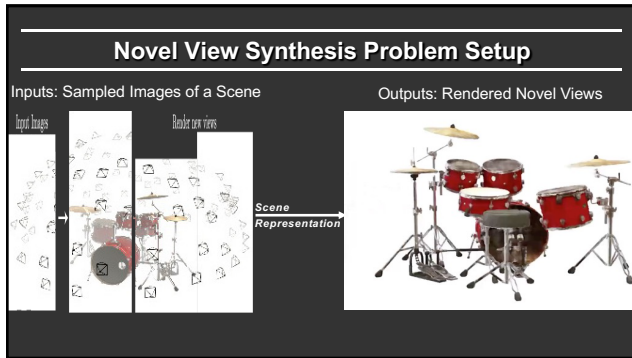
Input Images

4

Output Virtual Experience

5

6



7

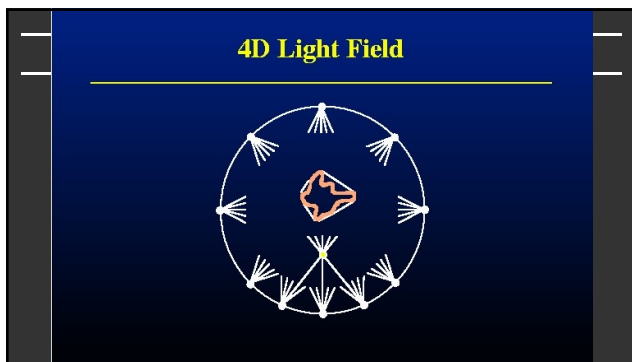
### Light Fields

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

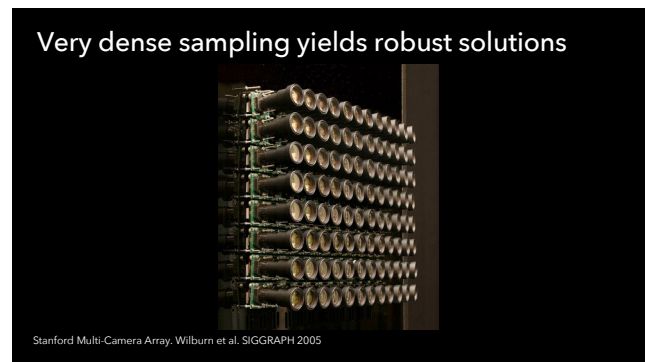
- In "free space" (no occluders) 5D reduces to 4D
  - Exterior of the convex hull of an object
  - Interior of an environment
- Images are 2D slices
  - Insert acquired imagery
  - Extract image from a given viewpoint

From Levoy and Hanrahan, Light Field Rendering, SIGGRAPH 96

8



9



10



11

How do we achieve robust high-quality performance with sparser casual input view sampling?

12

### Promote each sampled view to a local light field



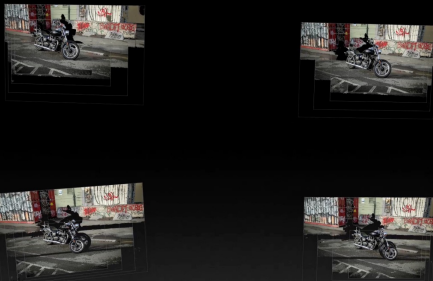
13

### Promote each sampled view to a local light field



14

### Blend nearby local light fields to render novel views



15

### Multiplane Image (MPI) as local light field representation

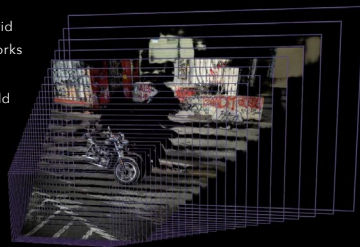


Lacroute and Levoy, SIGGRAPH 1994, Golland and Szeliski, IJCV 1999, Zhou et al. SIGGRAPH 2018

16

### MPI local light field representation well suited for deep learning pipeline

1. Samples stored on regular grid
  - Can use convolutional networks
2. 3D representation of light field
  - Consistent across views
3. Rendering is differentiable
  - Supervise by held-out views



Zhou et al. SIGGRAPH 2018

17

### Promote each sampled view to a local light field



18

Use neighboring views as inputs to predict MPI for each view



19

Use neighboring views as inputs to predict MPI for each view



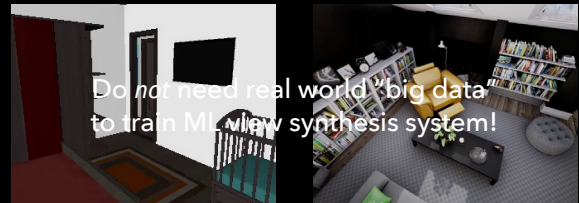
20

Use neighboring views as inputs to predict MPI for each view



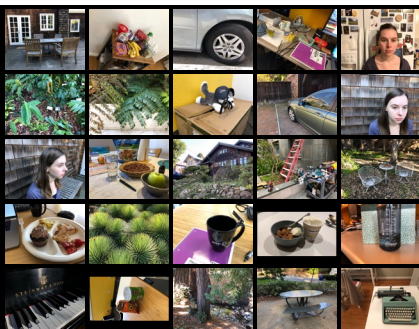
21

99% of our training data is synthetically rendered



Our main synthetic dataset takes under a day to generate  
We first train on synthetic data and then fine-tune on a small real dataset (25 scenes)

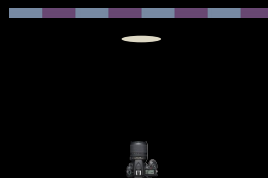
22



Real finetuning data

23

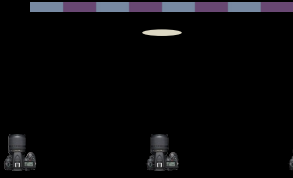
How densely do we need to sample input views to reconstruct the light field?



24

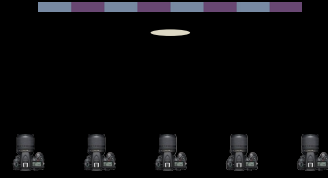


How densely do we need to sample input views to reconstruct the light field?



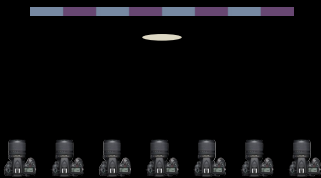
25

How densely do we need to sample input views to reconstruct the light field?



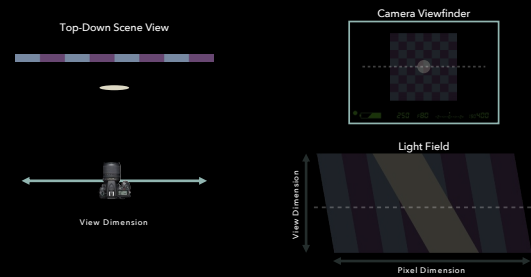
26

How densely do we need to sample input views to reconstruct the light field?



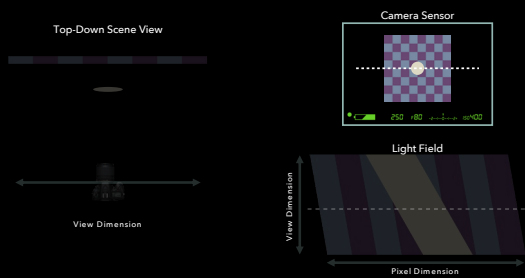
27

Visualizing a light field: each scene point lies on line based on depth



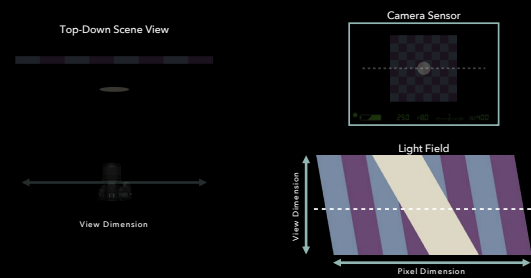
28

Visualizing a light field: each scene point lies on line based on depth

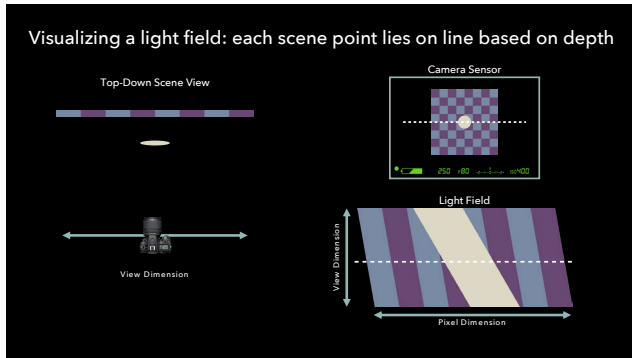


29

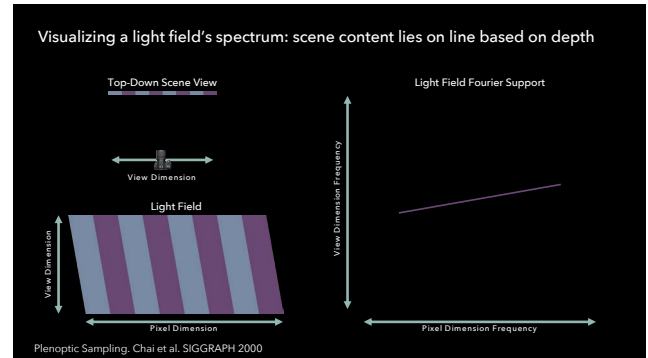
Visualizing a light field: each scene point lies on line based on depth



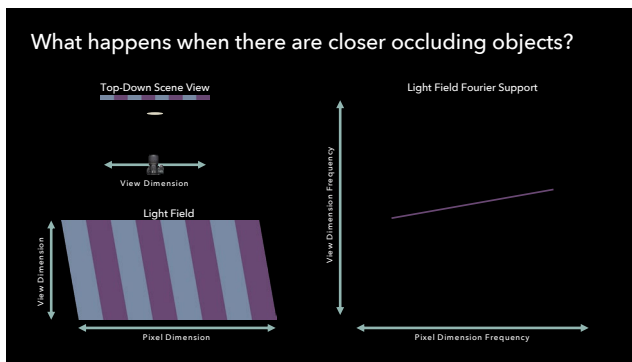
30



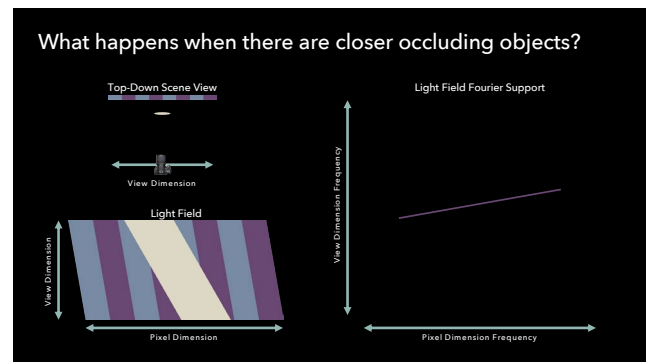
31



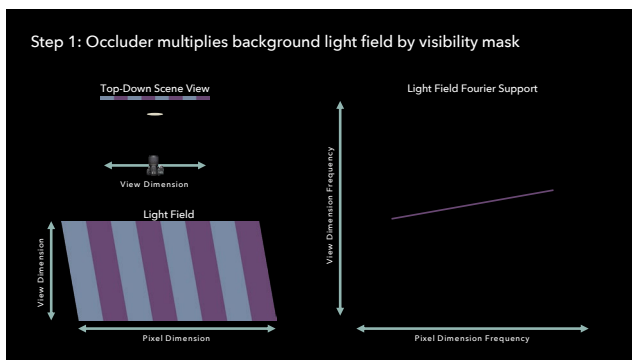
32



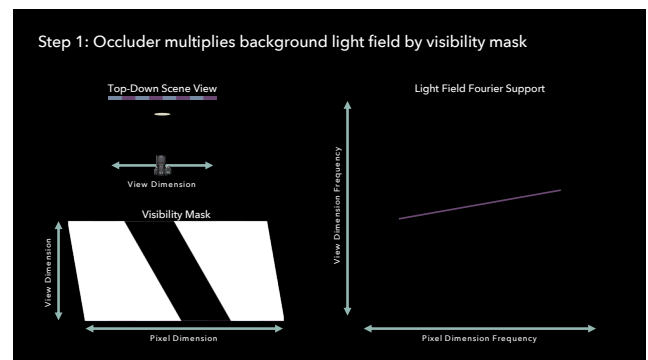
33



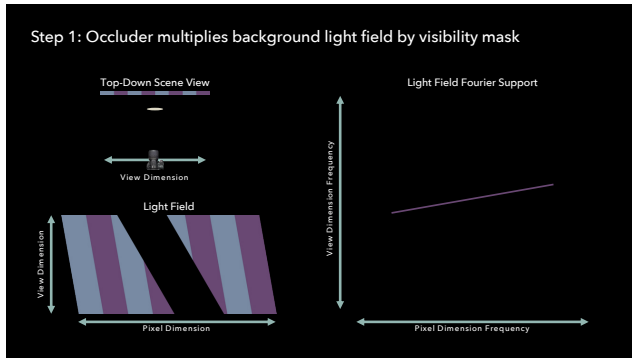
34



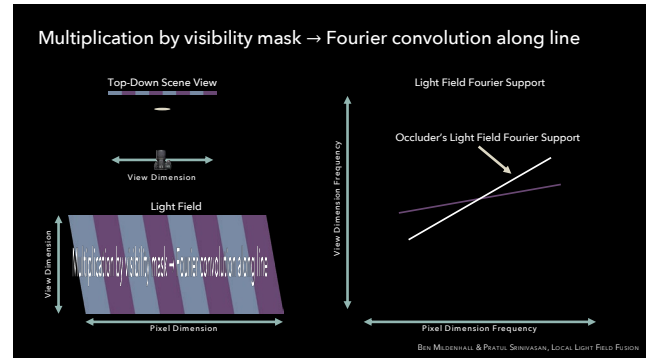
35



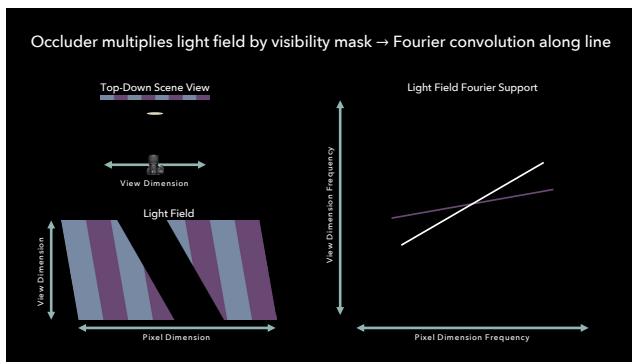
36



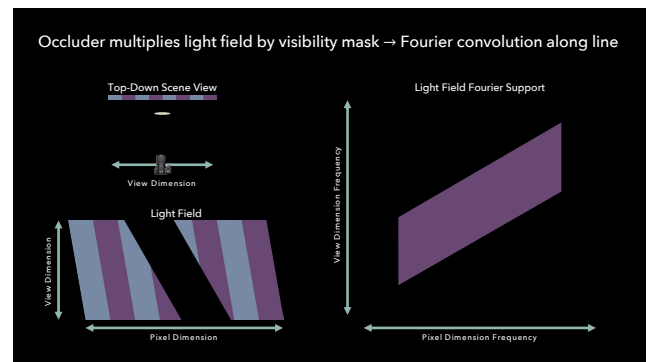
37



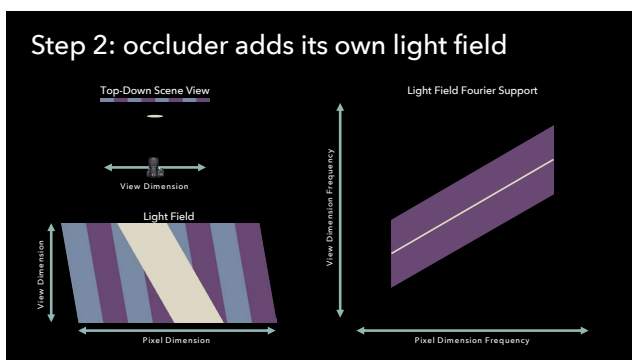
38



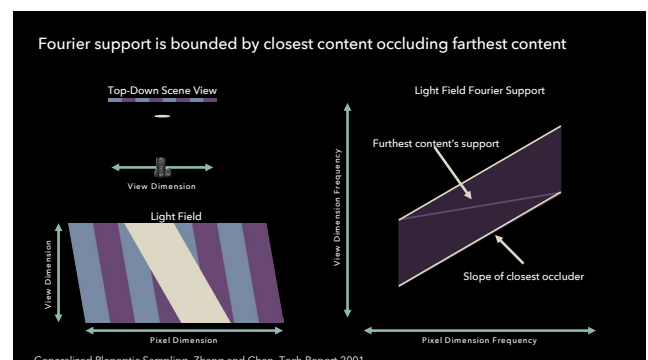
39



40

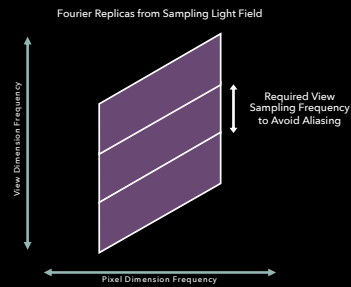


41



42

Nyquist rate view sampling means packing parallelograms as closely as possible



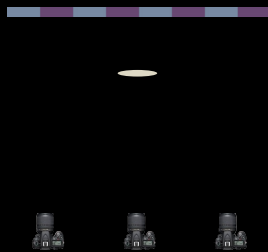
43

Nyquist rate means 1 pixel max disparity between adjacent views

Capturing a scene with closest content 0.5 m away for viewing on current VR headsets requires an image every millimeter!

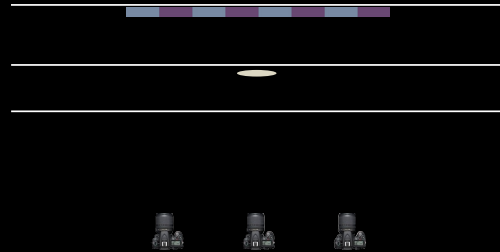
44

What if we have *ideal* MPI layers for each view?



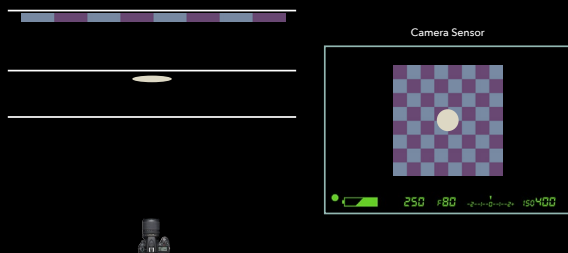
45

What if we have *ideal* MPI layers for each view?



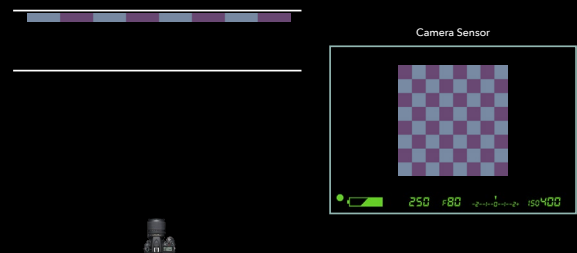
46

What if we have *ideal* MPI layers for each view?



47

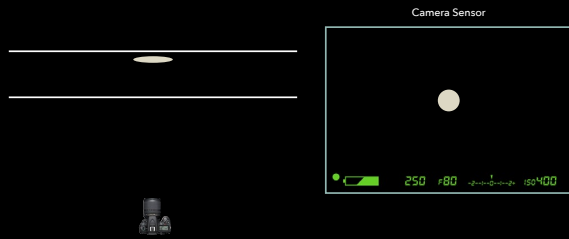
What if we have *ideal* MPI layers for each view?



48

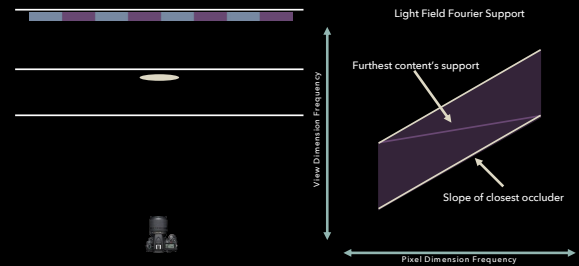


What if we have *ideal* MPI layers for each view?



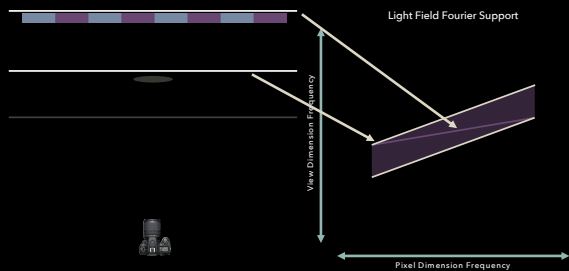
49

$D$  layers  $\rightarrow$  light field from each layer has  $D \times$  smaller Fourier support



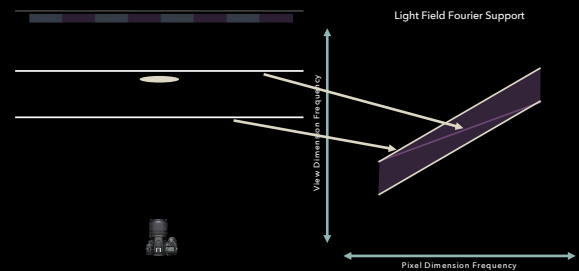
50

$D$  layers  $\rightarrow$  light field from each layer has  $D \times$  smaller Fourier support



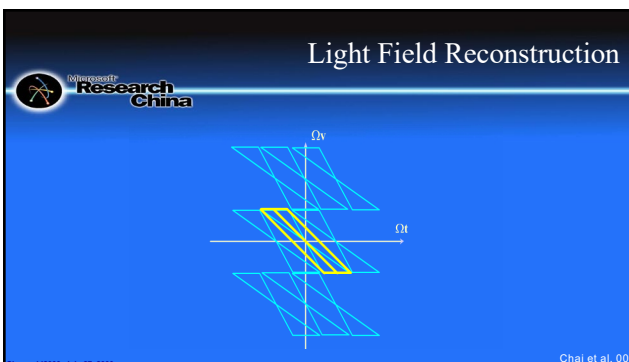
51

$D$  layers  $\rightarrow$  light field from each layer has  $D \times$  smaller Fourier support



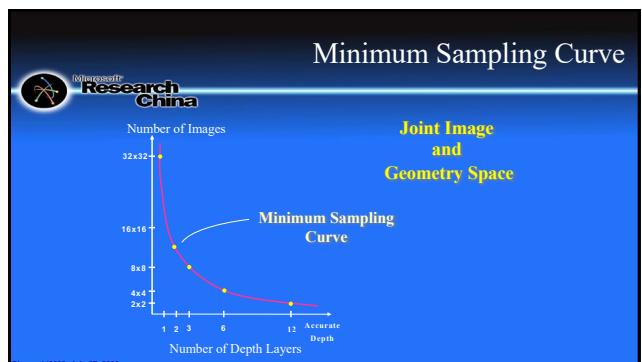
52

Light Field Reconstruction

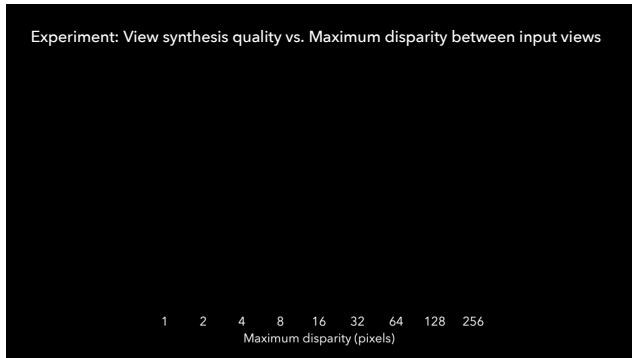


53

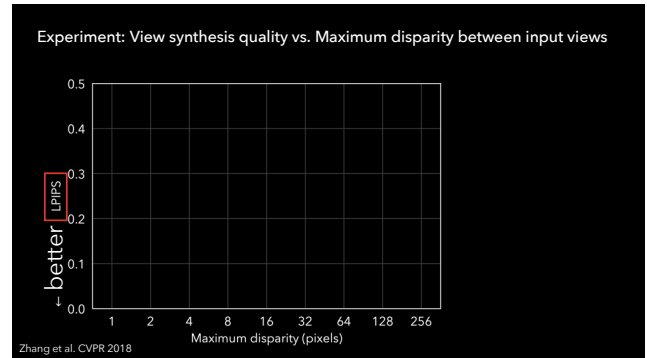
Minimum Sampling Curve



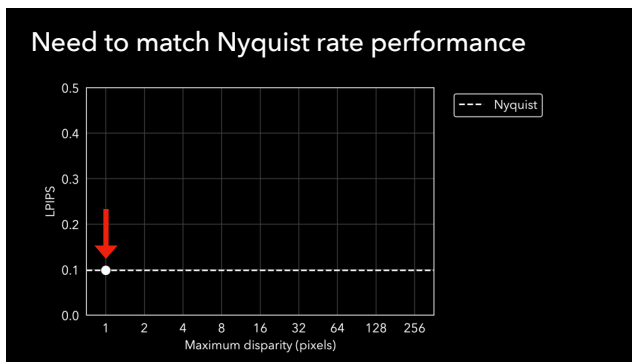
54



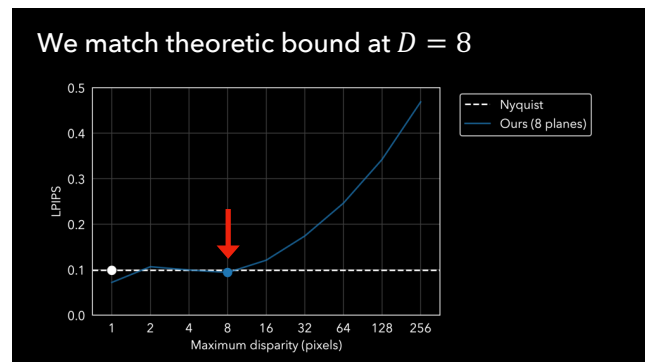
55



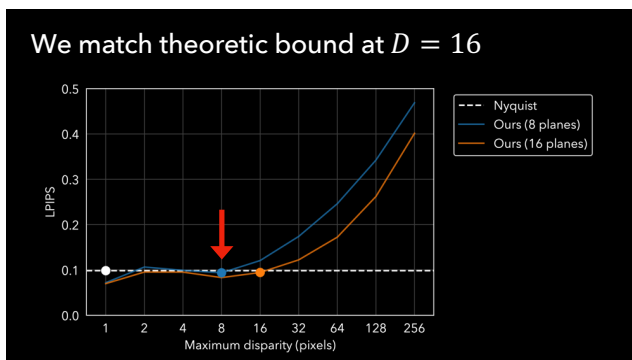
56



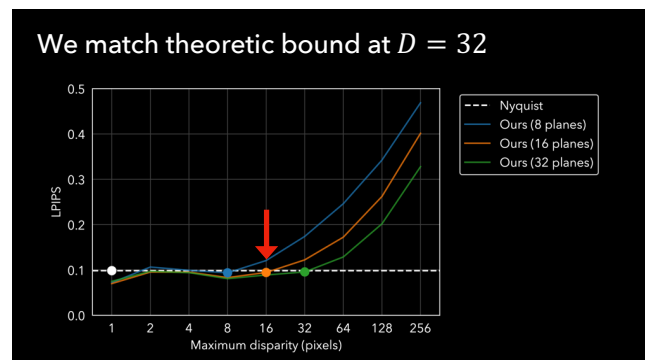
57



58

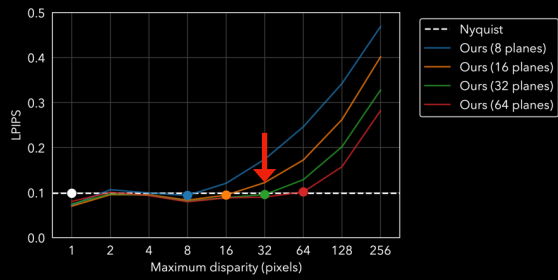


59



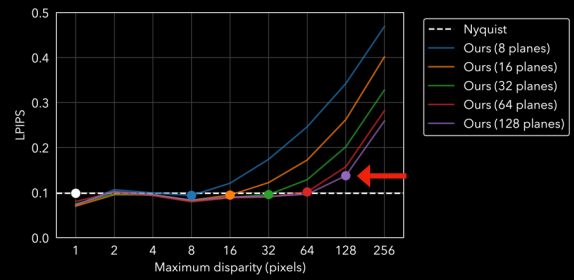
60

We match theoretic bound at  $D = 64$



61

Bound no longer holds at  $D = 128$



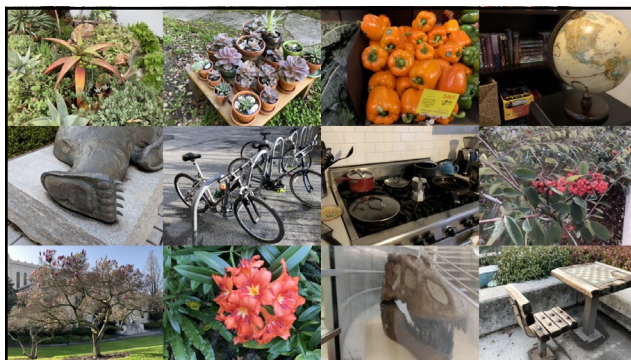
62



63



64



65