30 JULY – 3 AUGUST *Los Angeles*
SIGGRAPH2017

VIDEO FOR VIRTUAL REALITY

# LIGHT FIELD BASICS

JAMES TOMPKIN

BROWN
Computer Science

# WHAT IS A LIGHT FIELD?

‘Light field’ seems to have turned into a catch-all term for many advanced camera/display technologies.

# WHAT IS A LIGHT FIELD?

Has become a trade mark:

- Avegant Light Field Technology
- Light Field Lab
- LightField Studios


[Avegant / TechCrunch.com]

# WHAT IS A LIGHT FIELD?

- Volumetric video
  - But usually not a regular sampling like volumetric MRI data

- 4D video
  - But definitely not ‘4D cinema’

- 6DoF video
- Multi-camera / camera array
- Free viewpoint video
- Video-based rendering
- Lumigraph (+unstructured)
- …

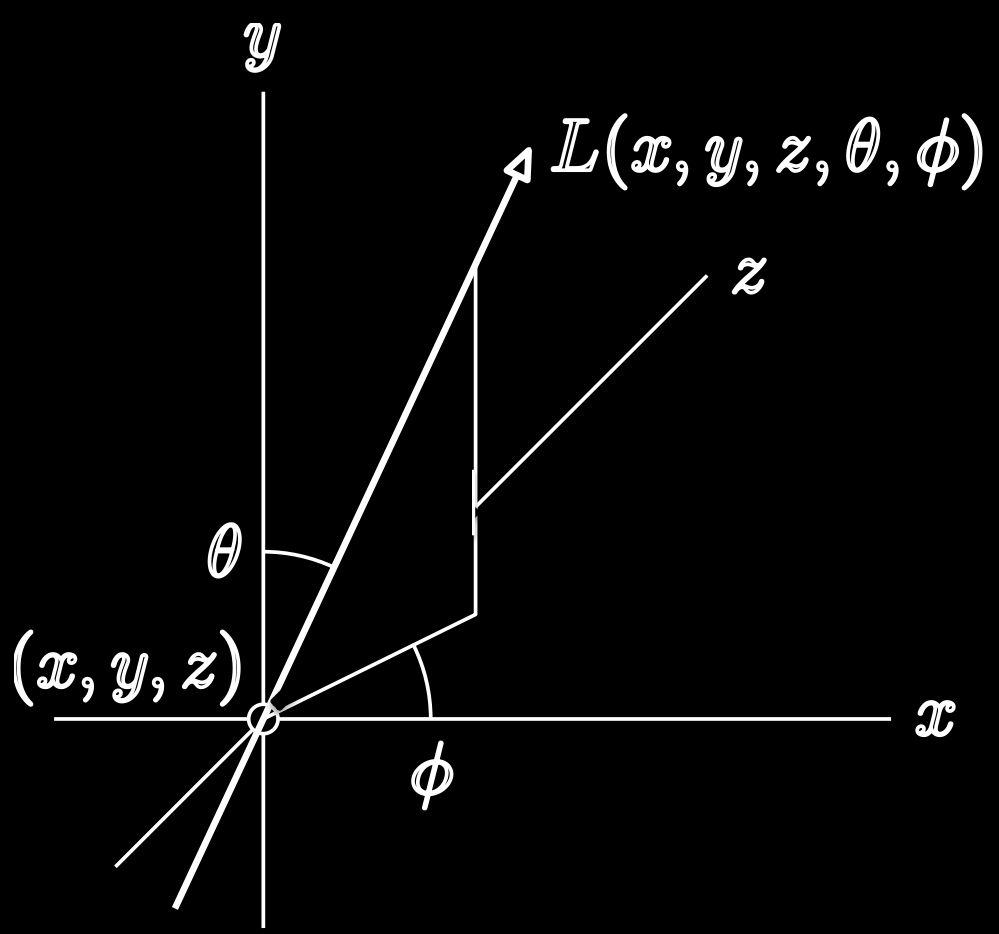
~~*“I’m in the movie!”*~~

# WHAT IS A LIGHT FIELD?

‘Light field’ seems to have turned into a catch-all term for many advanced camera/display technologies.

*How should we think about this?*

# PLENOPTIC FUNCTION



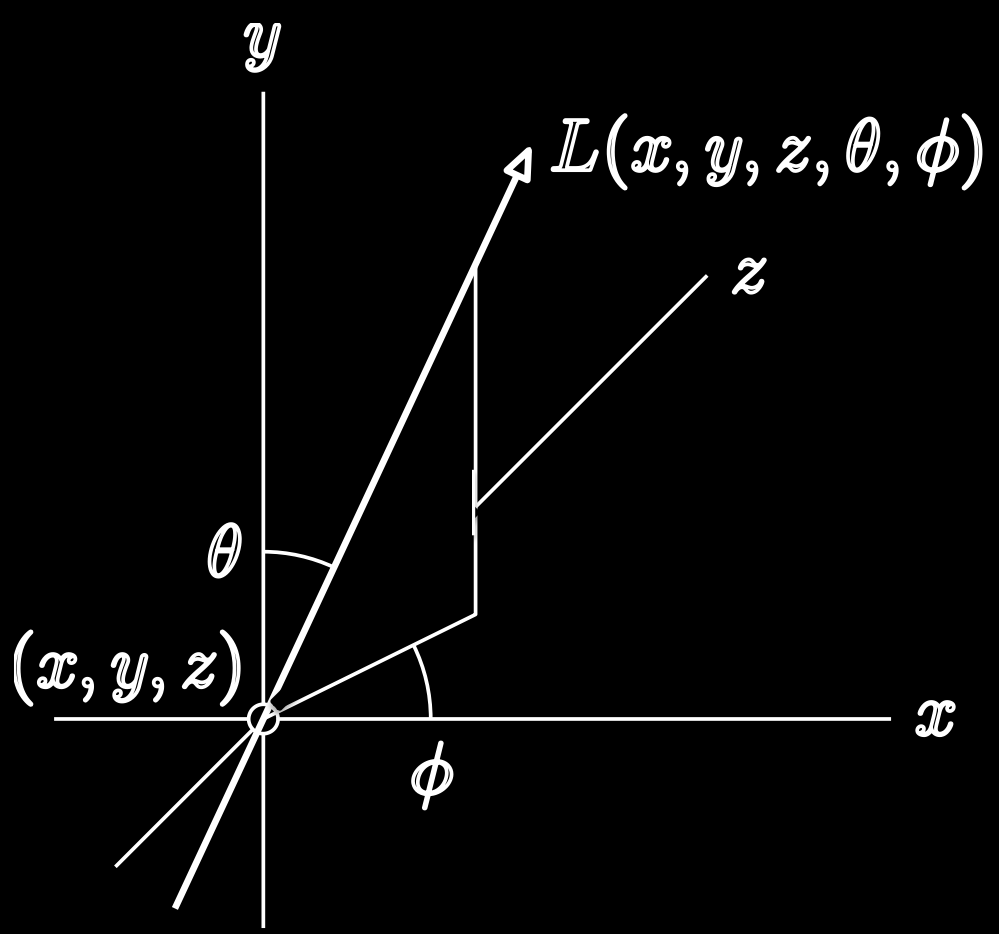

5+ dimensional function representing all light everywhere

- *x,y,z* -> position in space
- $\theta,\phi$ -> angle on sphere
- *L* -> amount of light (radiance)

[Wikipedia – ‘Light Field’ – Qef]

[Adelson and Bergen 1991, McMillan & Bishop 1995]

# PLENOPTIC FUNCTION




5+ dimensional function representing all light everywhere

- *x,y,z* -> position in space
- $\theta,\phi$ -> angle on sphere
- *L* -> amount of light (radiance)

- Video needs *t* -> time

[Wikipedia – 'Light Field' – Qef]

[Adelson and Bergen 1991, McMillan & Bishop 1995]

# WHAT IS A LIGHT FIELD?

‘Light field’ seems to have turned into a catch-all term for many advanced camera/display technologies.

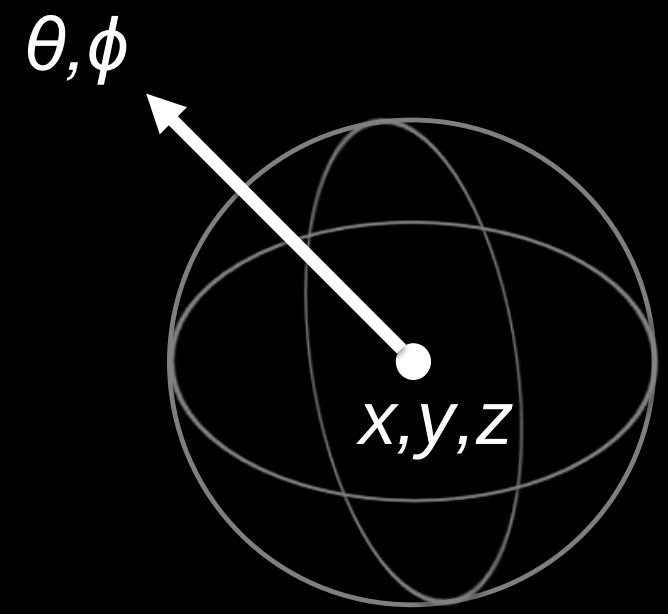
*How should we think about this?*

Different camera configurations provide different *samplings* of the plenoptic function.

[Chai et al., 2000, Lin et al. 2004]

# ALL TECHNIQUES ARE SAMPLINGS

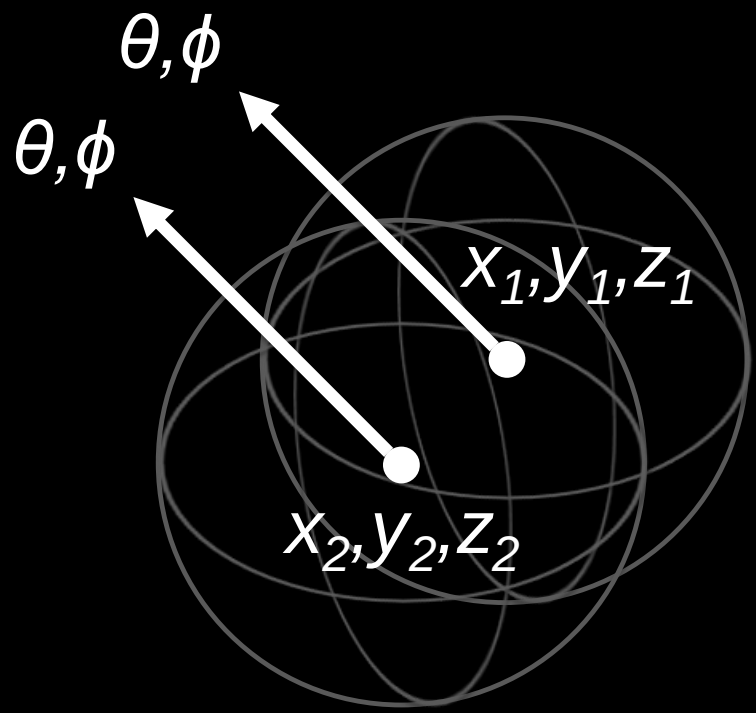
- Idealized 360° video
  - Single *x,y,z* sample
  - ‘Complete’ sampling of $\theta,\phi$



# ALL TECHNIQUES ARE SAMPLINGS

- Idealized stereo 3D 360° video
  - Two *x,y,z* samples
  - ‘Complete’ sampling of $\theta,\phi$

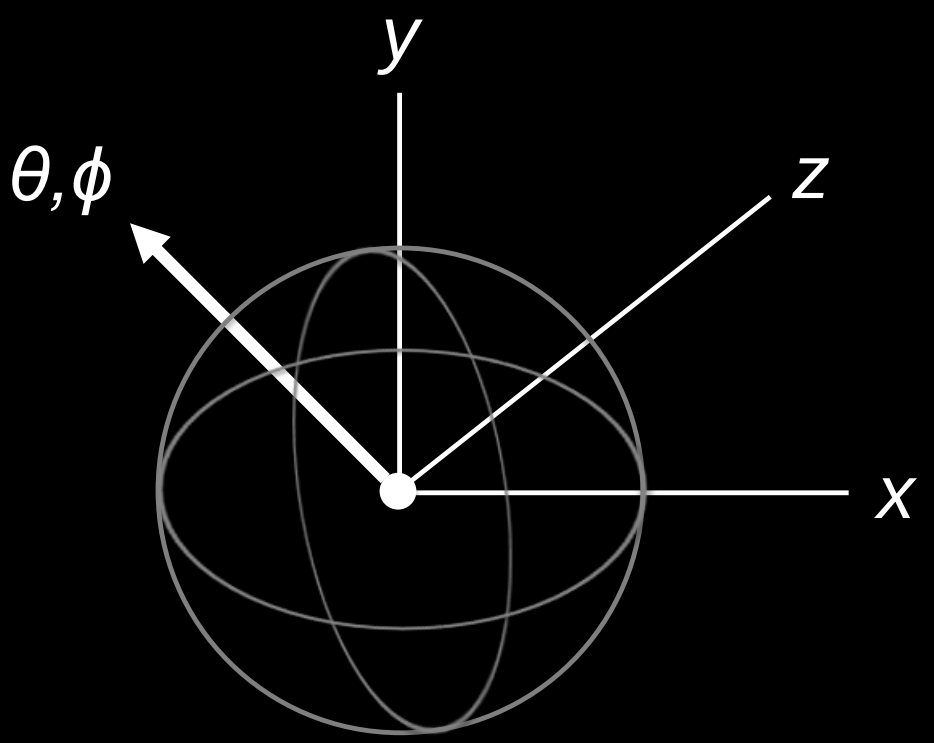
- *What happens when you tilt your head?*



# ALL TECHNIQUES ARE SAMPLINGS

- A light field generally implies
  - ‘Many’ *x,y,z* samples
  - Some sampling of *θ,ϕ*

- *Many options for how to sample!*



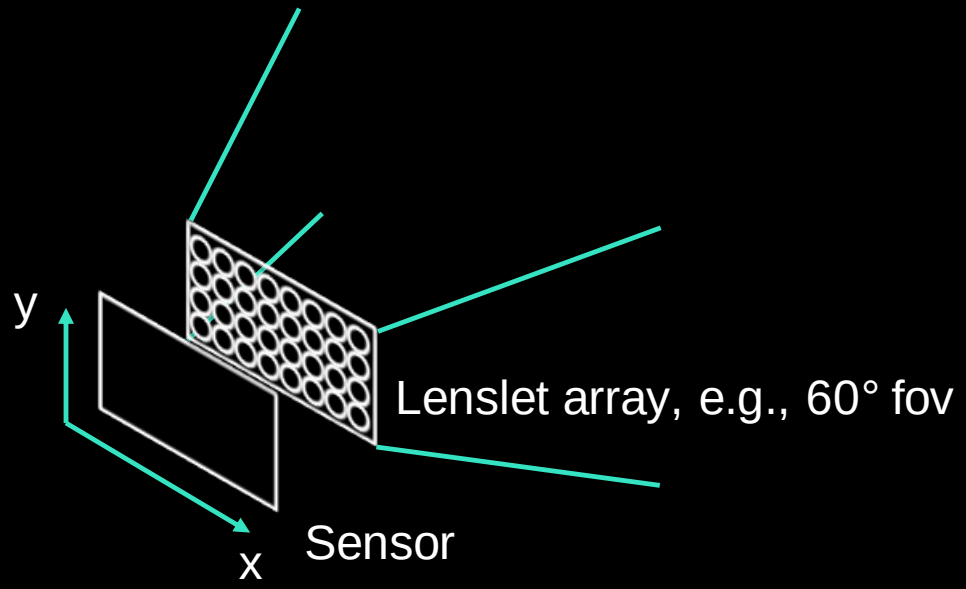
# THE RIGHT WAY TO THINK ABOUT LIGHT FIELDS

- A space of possible samplings!
  - Leads to a space of possible camera setups.
  - Each is some trade-off in the sampling space.

- Different configurations for different applications.
  - Each fails to sample some rays…
  - …but we have a suite of algorithms to (try to) accommodate.

# EXAMPLE: LYTRO CINEMA

[Image © Lytro]



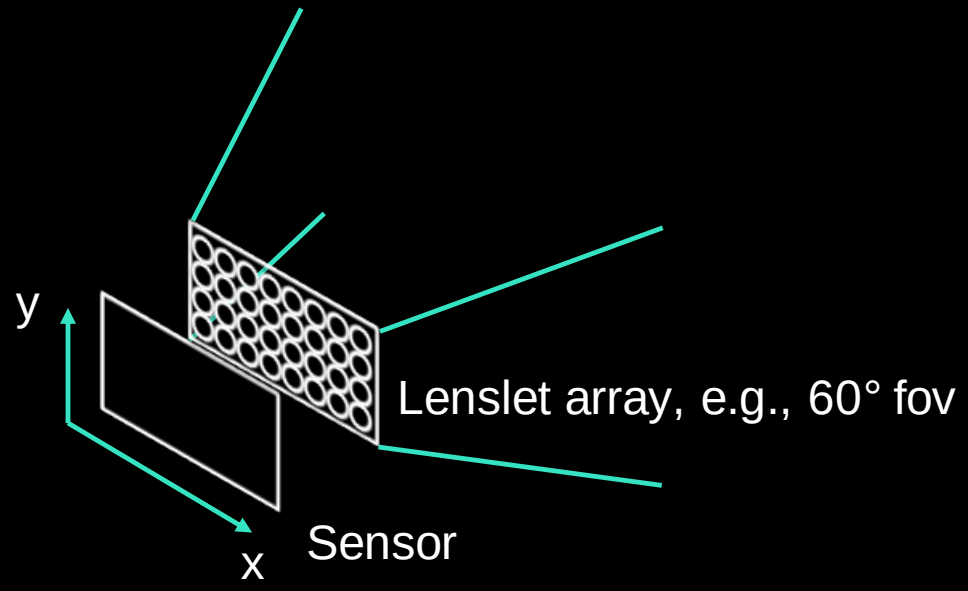

Lenslet-based light field imager
755 MPixel sensor

[Levoy and Hanrahan 1996, Gortler et al. 1996, Ng et al. 2005]

# EXAMPLE: LYTRO CINEMA

[Image © Lytro]

Lenslet-based light field imager
755 MPixel sensor



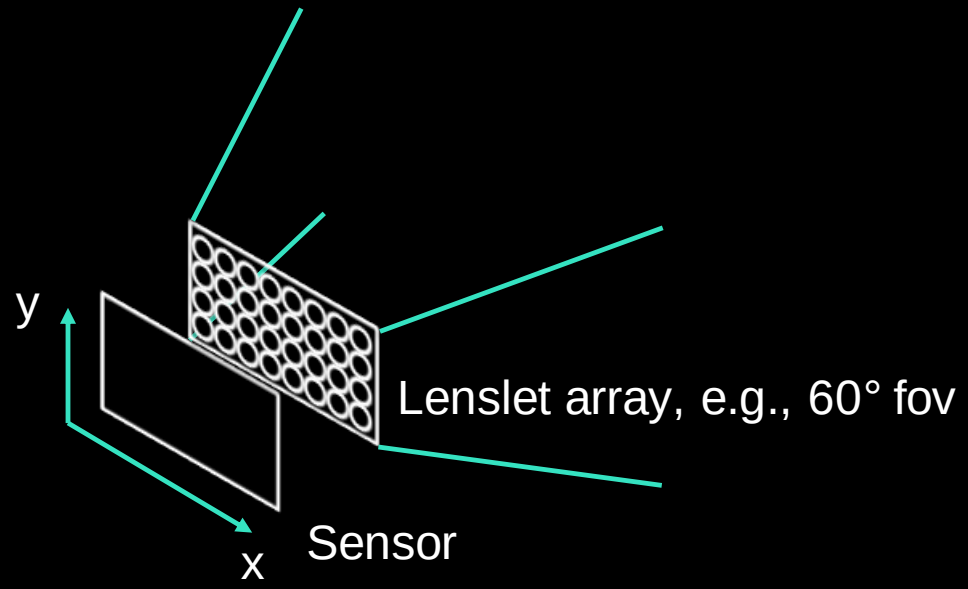

- Small baseline – dense sampling
  - Little parallax/head motion
- Small $\theta,\phi$ extent
  - Only sees part of scene

[Levoy and Hanrahan 1996, Gortler et al. 1996, Ng et al. 2005]

# EXAMPLE: LYTRO CINEMA

[Image © Lytro]

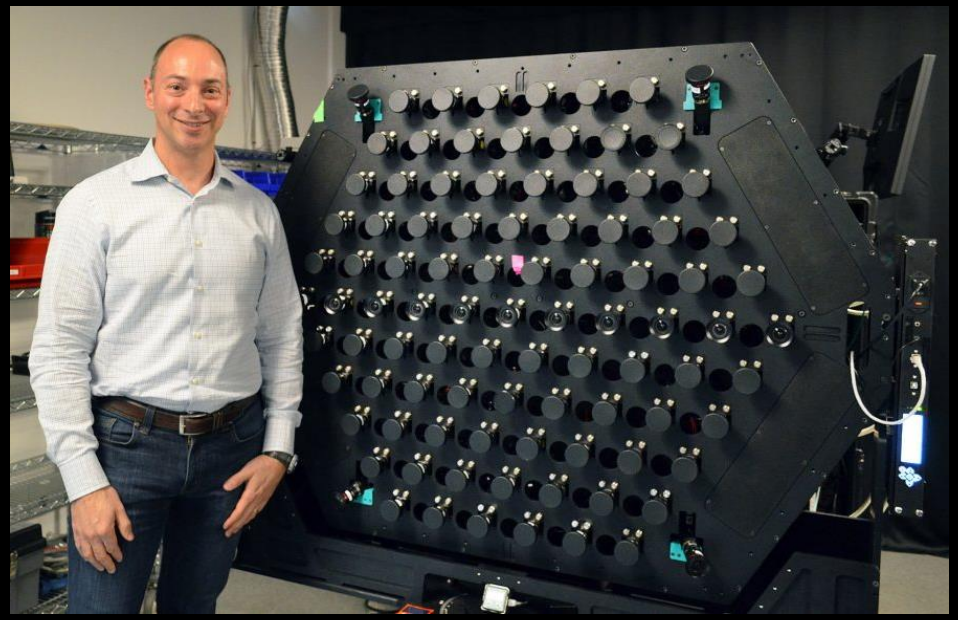
Lenslet-based light field imager
755 MPixel sensor




- Still a light field camera…
….but not intended for VR.
- Helps post-production for synthetic shutter and aperture, scene segmentation.

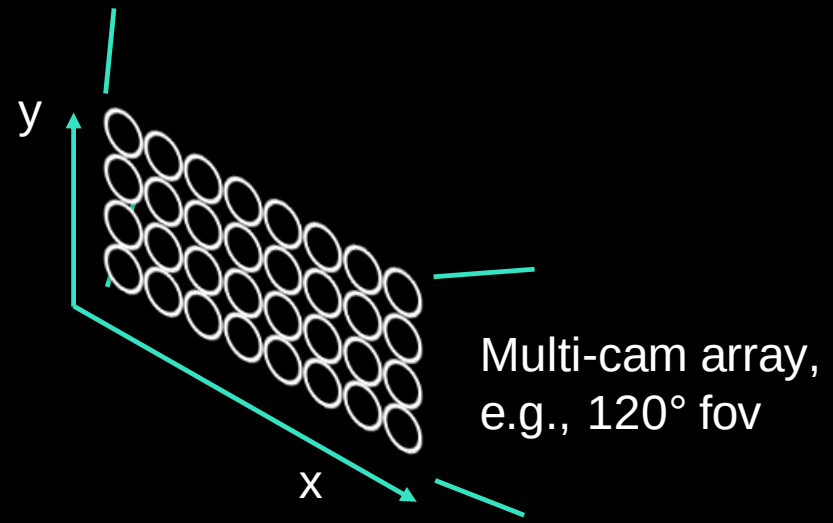
[Levoy and Hanrahan 1996, Gortler et al. 1996, Ng et al. 2005]

# EXAMPLE: LYTRO IMMERGE

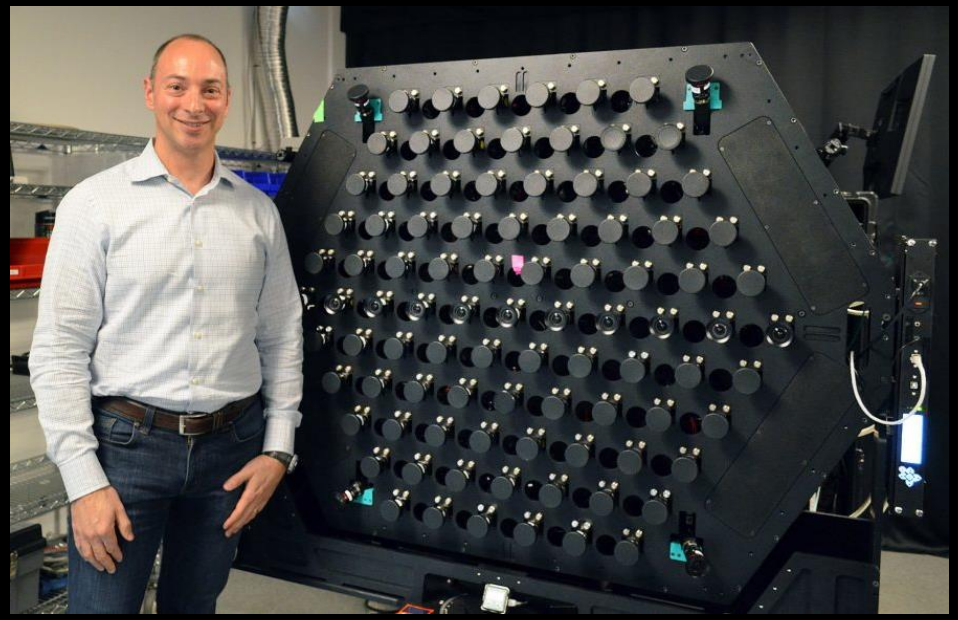


[Image © Lytro]

Multi-camera light field imager
95 cameras

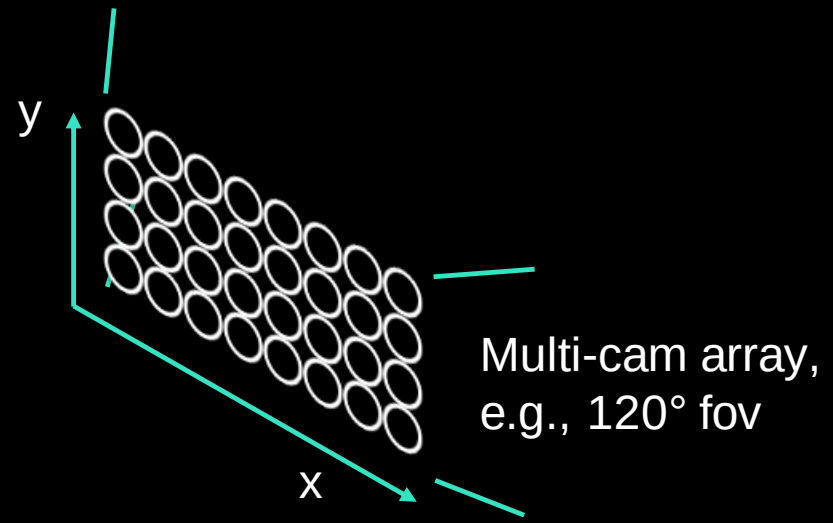


# EXAMPLE: LYTRO IMMERGE



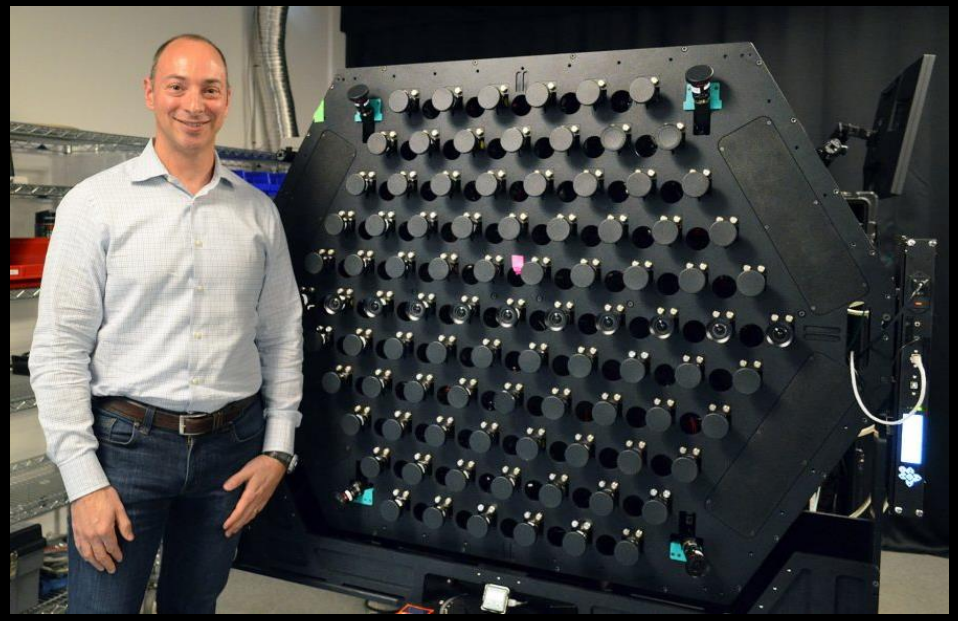
[Image © Lytro]

Multi-camera light field imager
95 cameras



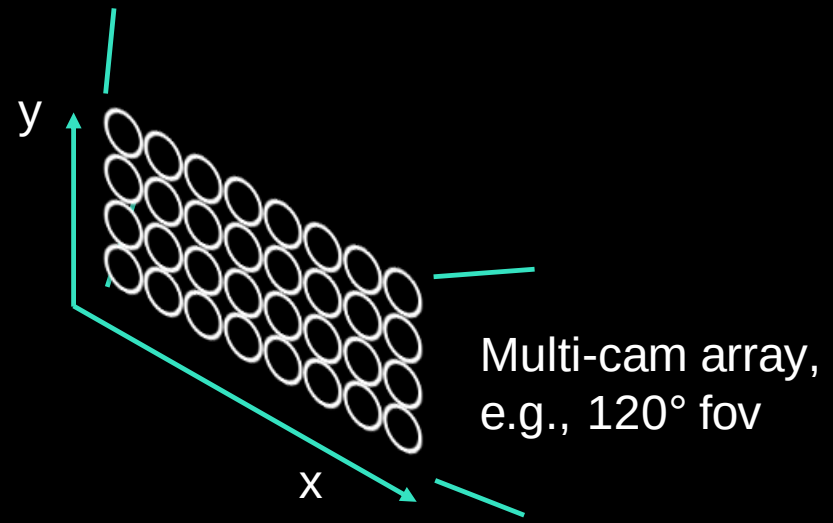

- Larger baseline
  - Sufficient parallax/head motion for seated(+) VR
- Still not 360° $\theta,\phi$ extent
  - Wider FOV lenses

# EXAMPLE: LYTRO IMMERGE



[Image © Lytro]

Multi-camera light field imager
95 cameras

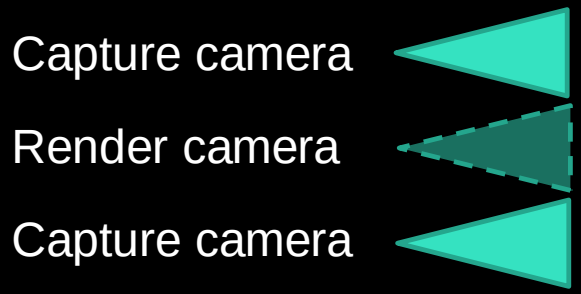



- Better for VR
  - Larger ‘headbox’
  - But need to interpolate farther between views.

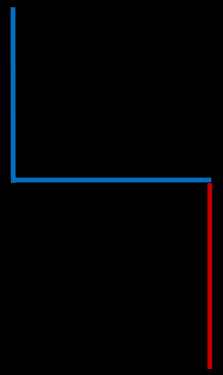
# VIEW INTERPOLATION

- What happens when my VR head goes ‘in between’ the cameras?
  - Must resample captured rays to render a novel view.

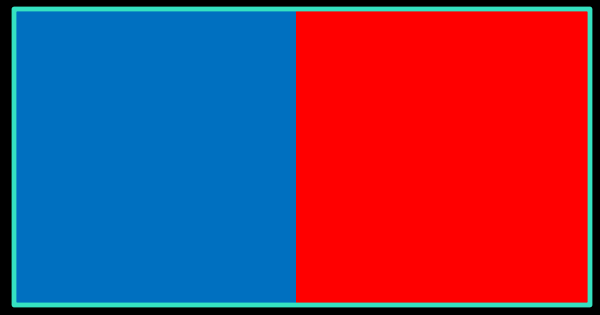
Dense camera sampling, e.g., Lytro Cinema




World



Interpolated frame

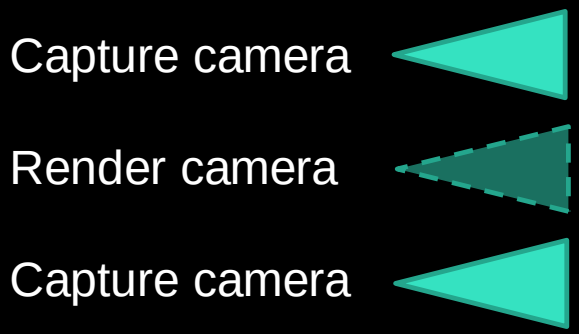


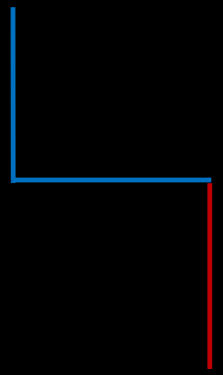
Looks ok if sampling is dense / object is far away.

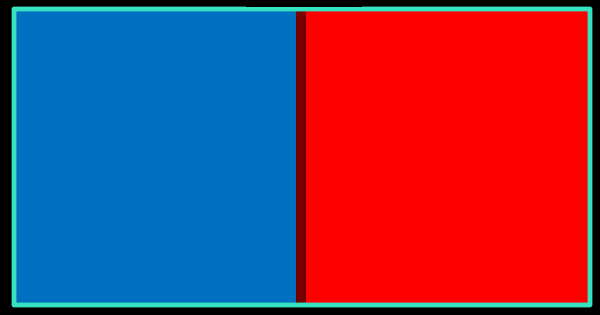
[Chai et al., 2000, Lin et al. 2004]

# VIEW INTERPOLATION

- What happens when my VR head goes ‘in between’ the cameras?
  - Must resample captured rays to render a novel view.



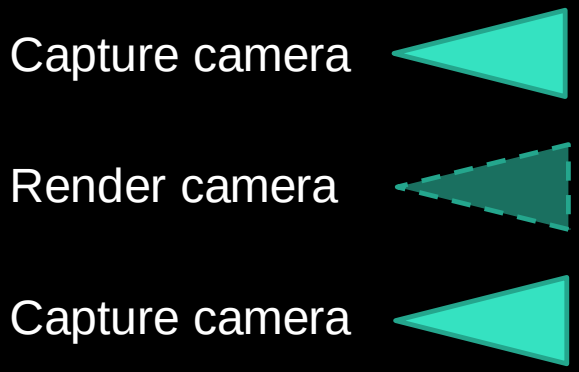



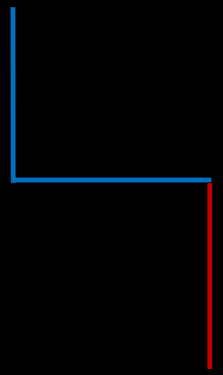

More and more ghosting.

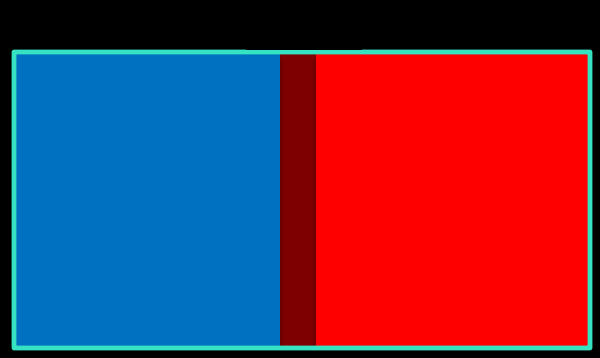
[Chai et al., 2000, Lin et al. 2004]

# VIEW INTERPOLATION

- What happens when my VR head goes ‘in between’ the cameras?
  - Must resample captured rays to render a novel view.



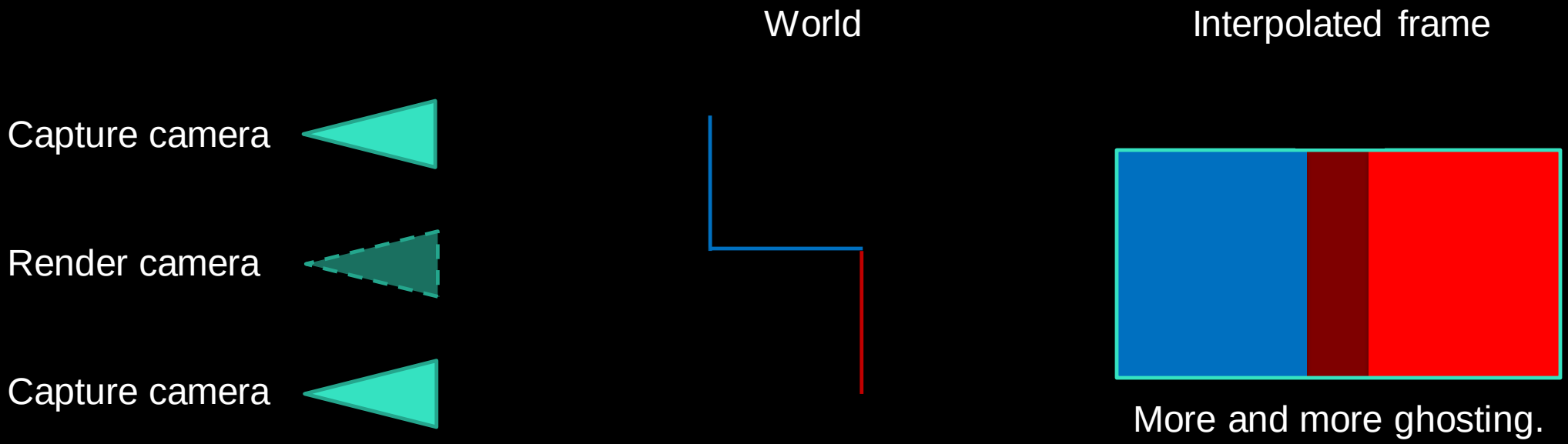




[Chai et al., 2000, Lin et al. 2004]

# VIEW INTERPOLATION

- What happens when my VR head goes ‘in between’ the cameras?
  - Must resample captured rays to render a novel view.

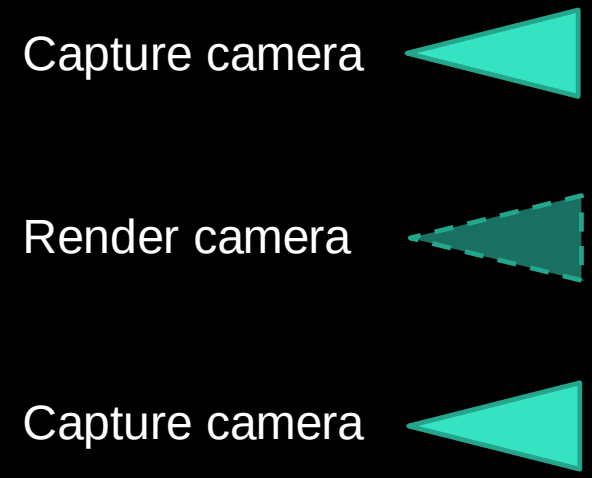



[Chai et al., 2000, Lin et al. 2004]

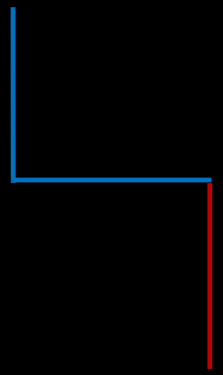
# VIEW INTERPOLATION

- What happens when my VR head goes ‘in between’ the cameras?
  - Must resample captured rays to render a novel view.

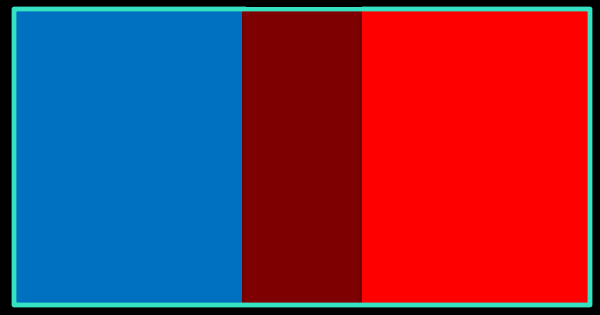
Sparser camera sampling, e.g., Lytro Immerge




World



Interpolation is untenable.

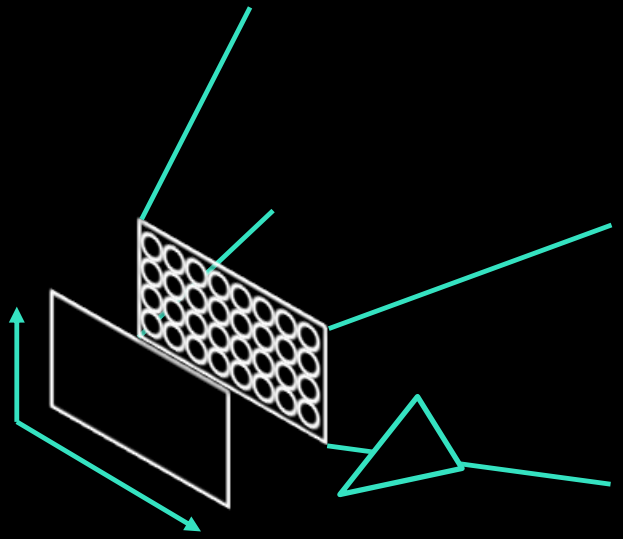


Must do something else.

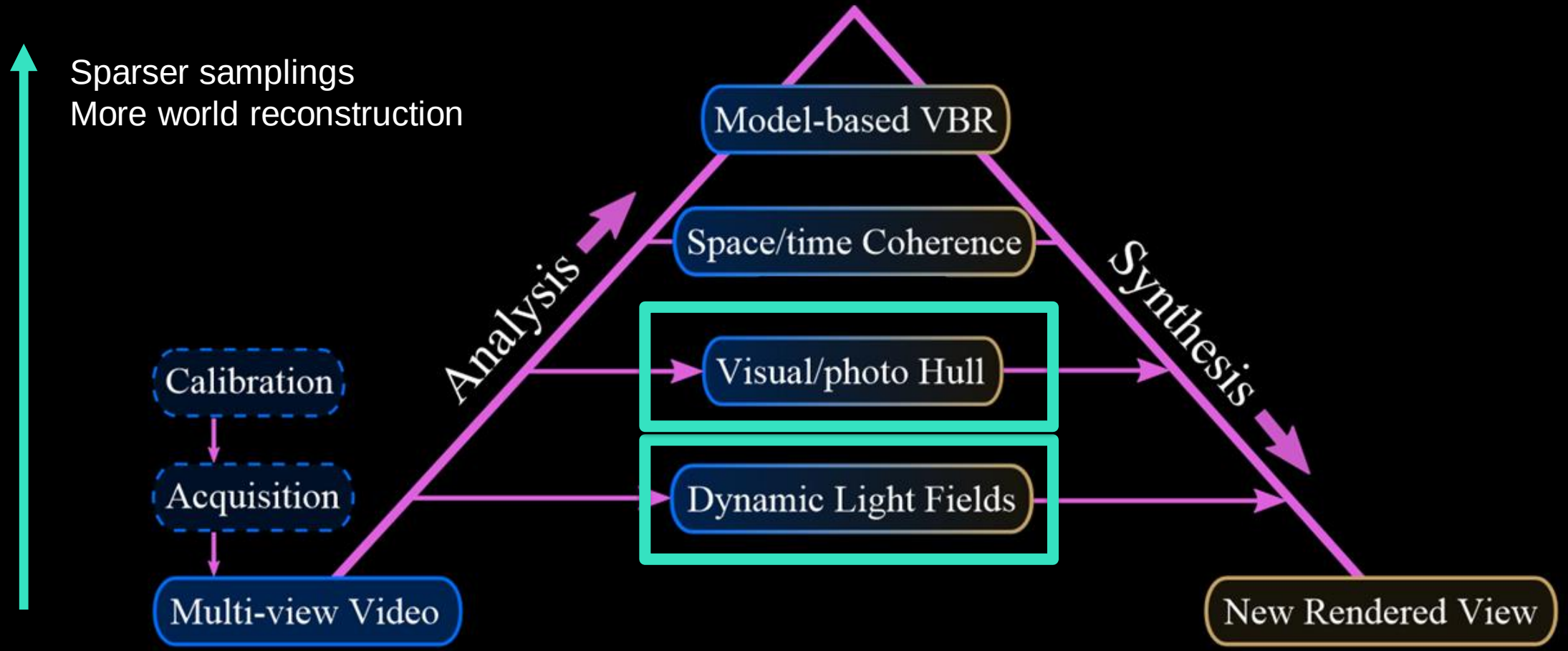
[Chai et al., 2000, Lin et al. 2004]

# VIEW INTERPOLATION

- What happens when my VR head goes ‘beyond’ my baseline?
  - Reveal space not captured by any cameras.
  - Can ‘fake it’ or ‘hallucinate’ – *inpainting only goes so far.*
  - Basically impossible with sampling-based approach.



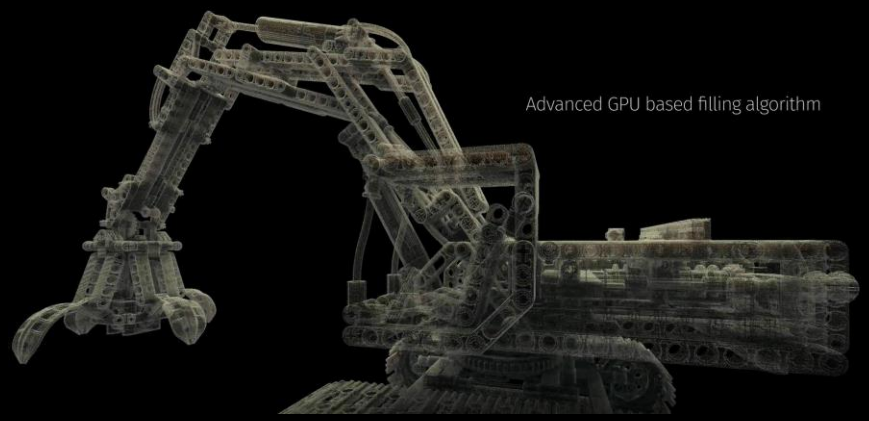
# PYRAMID OF ANALYSIS / SYNTHESIS




[Magnor et al., Video-based Rendering, SIGGRAPH 2005 Course]

# COMMON REPRESENTATIONS

## Point Clouds




[Point-based Graphics, Gross and Pfister 2007]
e.g., Nurulize ‘Atom View’

## Dynamic Mesh + View-dependent Texture

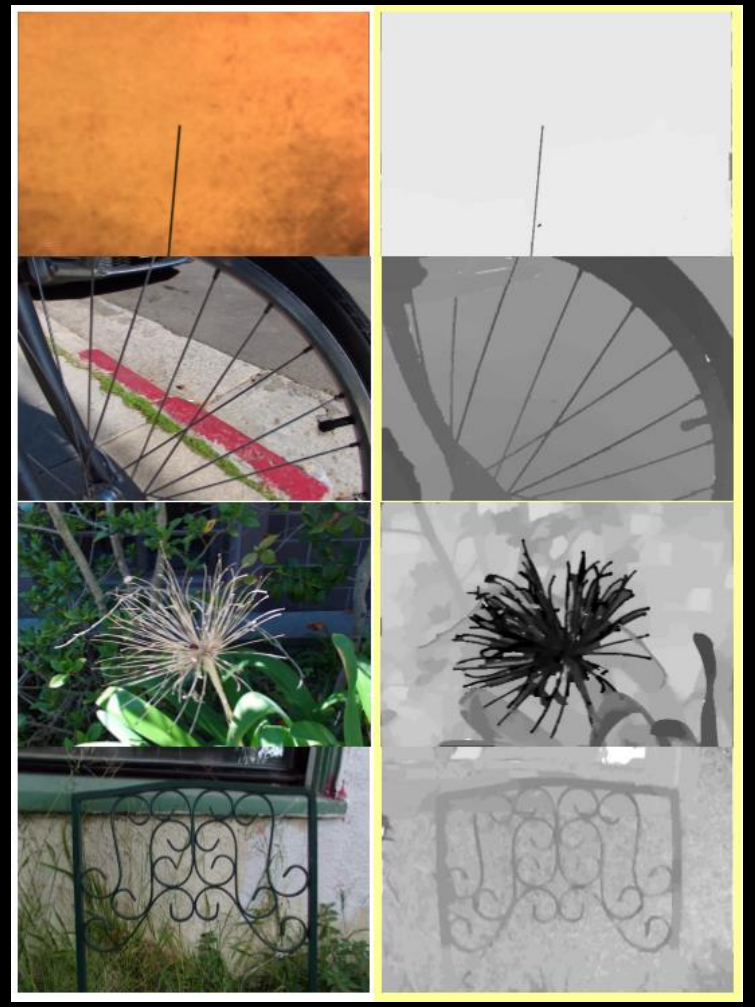
[Collet et al., High-quality Streamable Free-viewpoint Video, ACM ToG 2015]

# DEPTH MAP EXTRACTION

Point clouds come from computing world depth.
Multi-view stereo.

Many techniques:

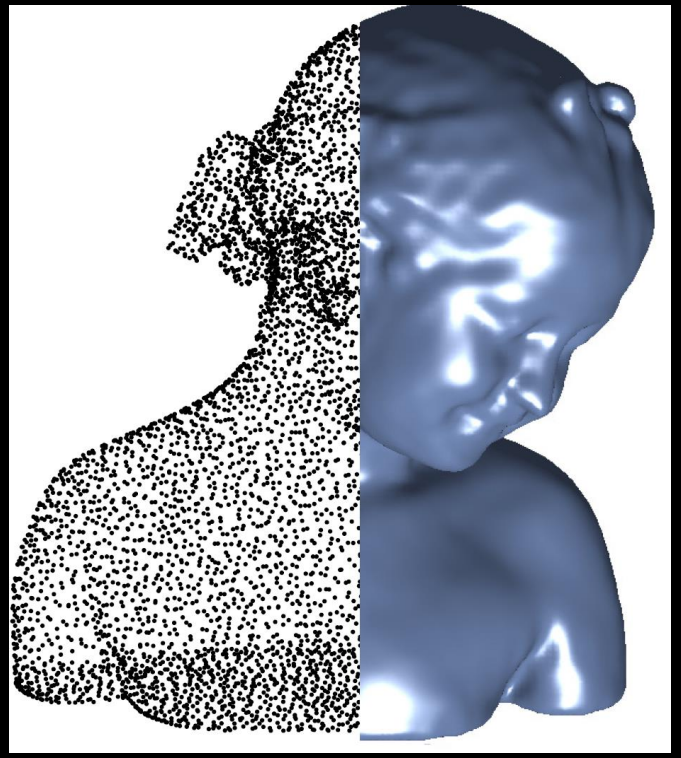
- Yucer et al., SIGGRAPH 2016
- Wang et al., TPAMI 2016
- Jeon et al., CVPR 2015
- Kim et al., SIGGRAPH 2013, 3DV 2016



LF input Wang et al.

# PHOTOGRAMMETRY / GEOMETRY RECONSTRUCTION

- Merge depth maps / fit surface to point cloud.



[© Mathworks]

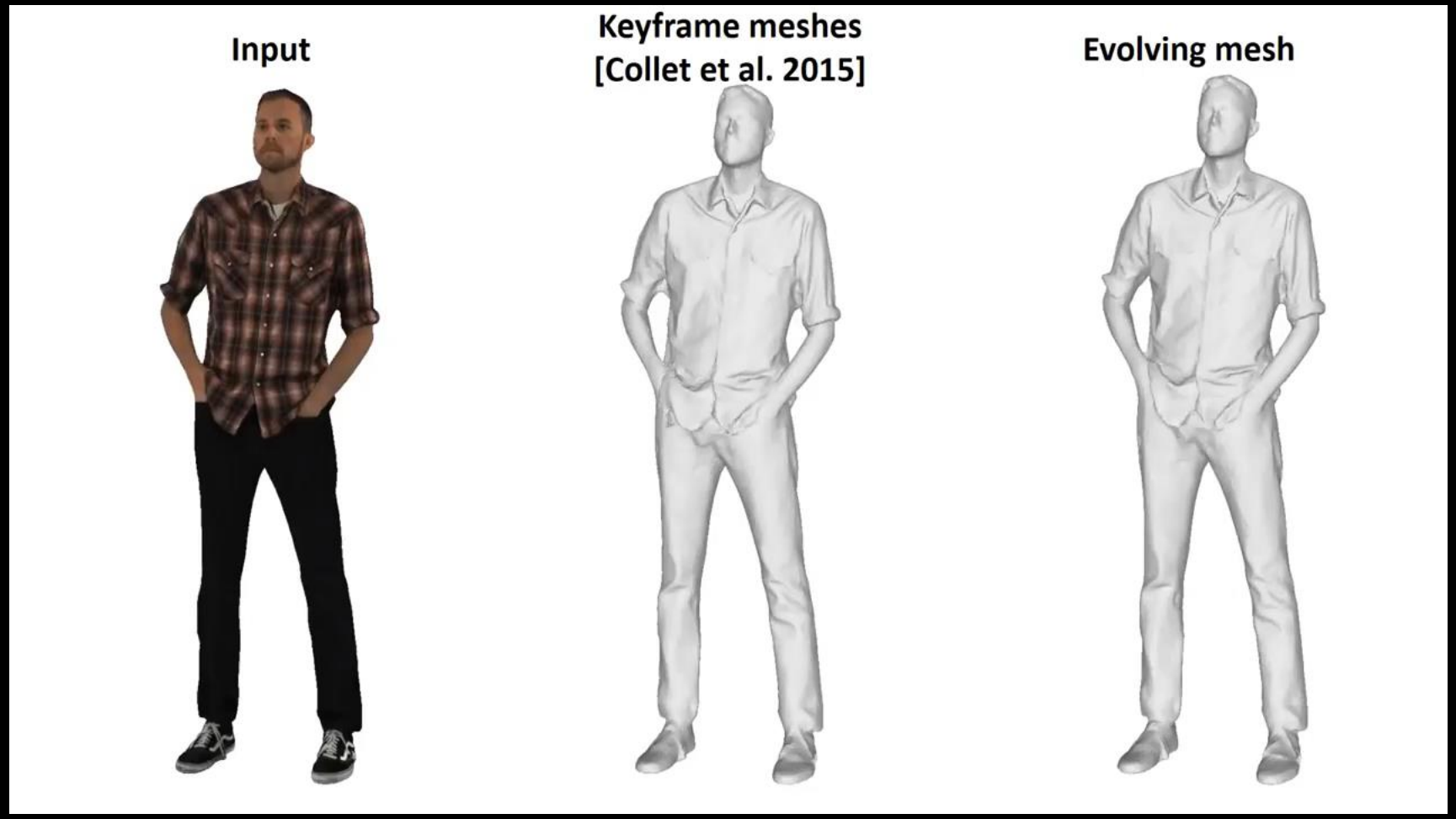
One reconstruction per frame.

Errors: Geometry edge flickering, phantom volumes…

Tools: RealityCapture, ImageModeler, PhotoScan.

[Kazhdan et al., Poisson Surface Reconstruction, SGP 2006]

# TIME-EVOLVING MESHES




Prada et al.,
**Spatiotemporal atlas parameterization for evolving meshes.**
SIGGRAPH 2017

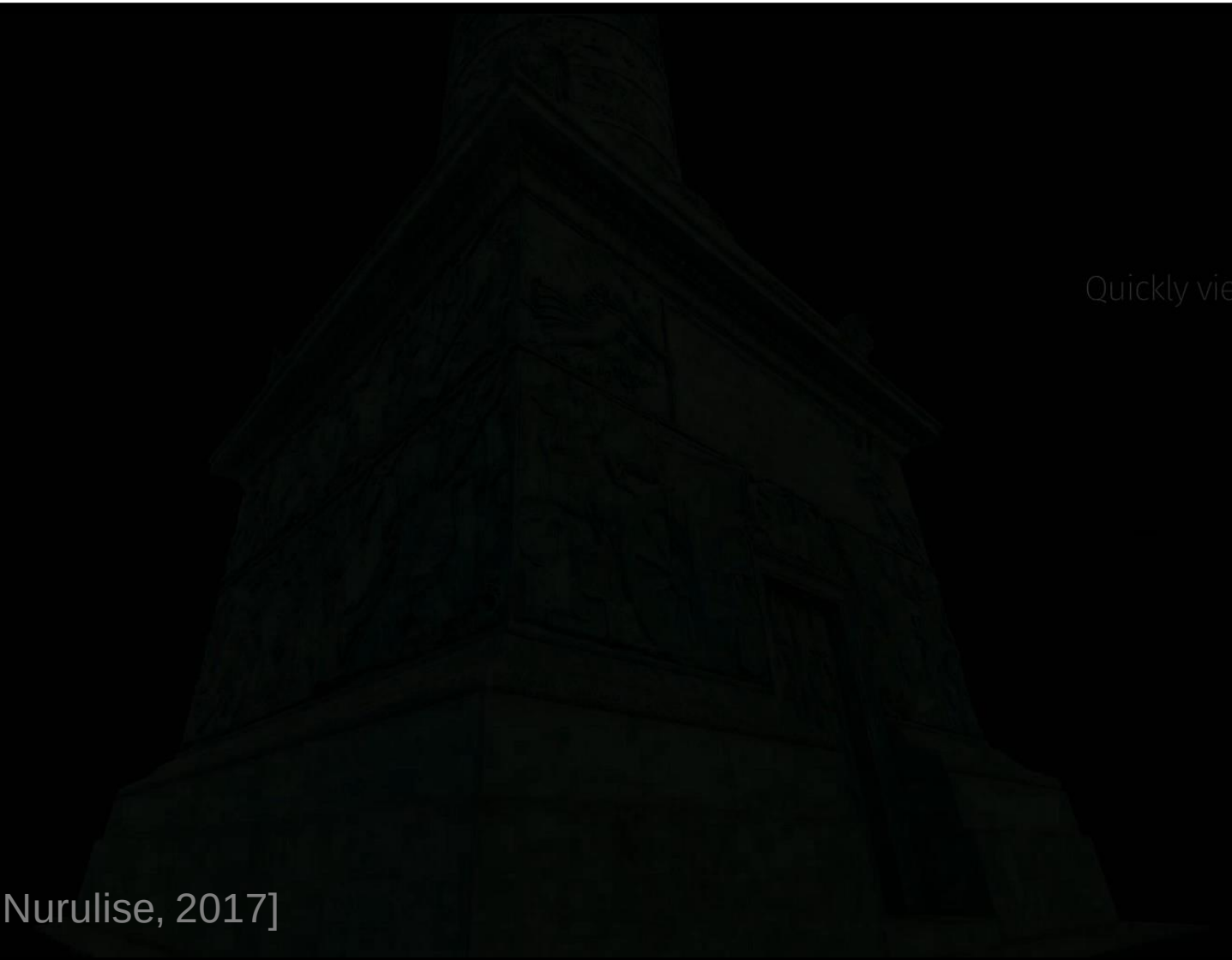
# OTHER TECHNOLOGIES:
# LIDAR, DEPTH CAMERAS, OTHER SCANNERS

- Even with dense sampling, sometimes we can’t reconstruct depth or geometry:
  - Complex materials
  - Dynamic events

- Use other modes to fill in:
  - Geometry from LIDAR to generate ‘clean backplate’ for dynamic depth.
  - Fit known geometric models to scene.


[FARO Focus laser scanner]

# DO WE EVEN NEED SURFACES?



Quickly view and use photogrammetry

[© Nurulise, 2017]

[Point-based Graphics, Gross and Pfister 2007]

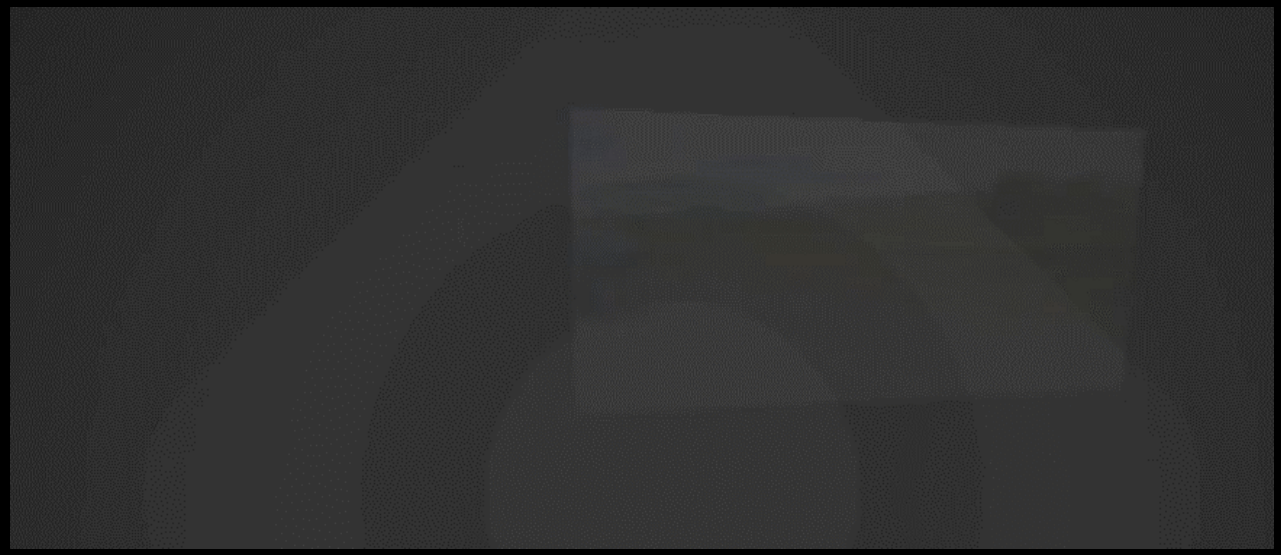
# RECORDING LIGHT FIELDS

- A lot of data!
  - 100x the cameras = 100x the data
  - VR needs high frame rate – 90+ fps
  - 100 camera 1080p 8bit RGB 90fps
    = 52 Gbytes per second RAW
  - Terabytes of data for even a movie short.

# PLAYING LIGHT FIELDS

- Real-time rendering in ‘game engine’ – 90 Hz
- High data bandwidth, complex geometry.
  - Big machine.

- Alternative: holographic video



[OTOY holographic videos - https://home.otoy.com/render/light-fields/]

# LIGHT FIELD VIDEO GOALS

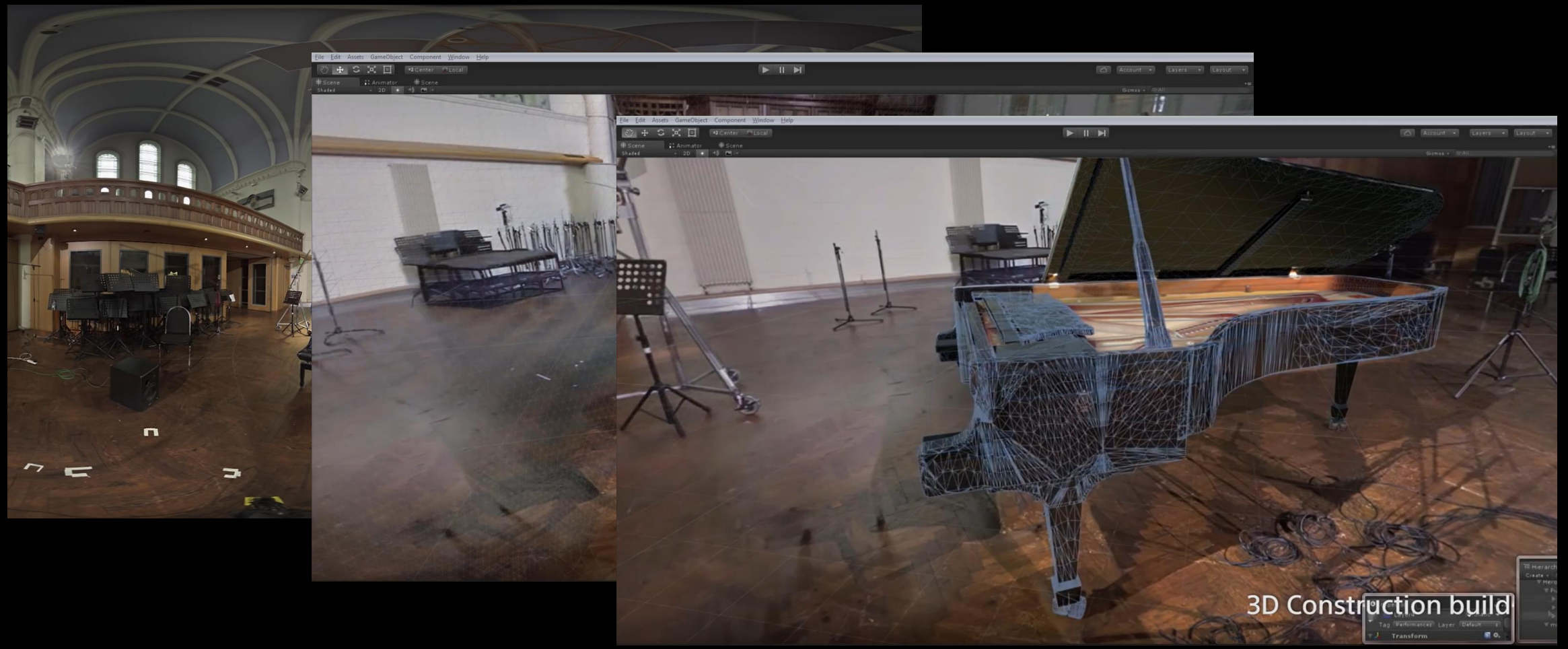
- High immersion
  - 6 degrees of movement freedom.
  - Correct stereo vision everywhere, not just horizontal.

- High realism
  - Looks closer to real life than a real-time rendering
  - But…artifacts often not ‘solid’, e.g., geometric edge flickering.
    - Starts to lose the look of ‘video’.

# LYTRO ‘MOON’


[Images © Lytro - http://blog.lytro.com/vr-announcing-moon/]

# SONY ‘JOSHUA BELL’




[Images © Sony - https://youtu.be/poI6E8O4jmQ]

# SONY ‘JOSHUA BELL’

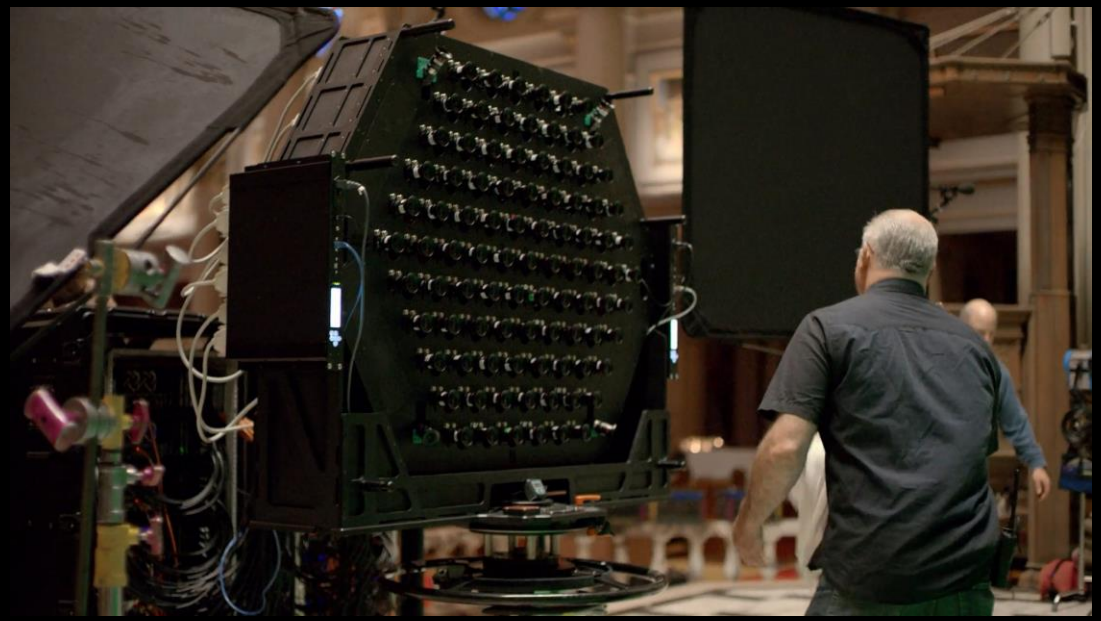

[Images © Sony - https://youtu.be/poI6E8O4jmQ]

# SONY ‘JOSHUA BELL’

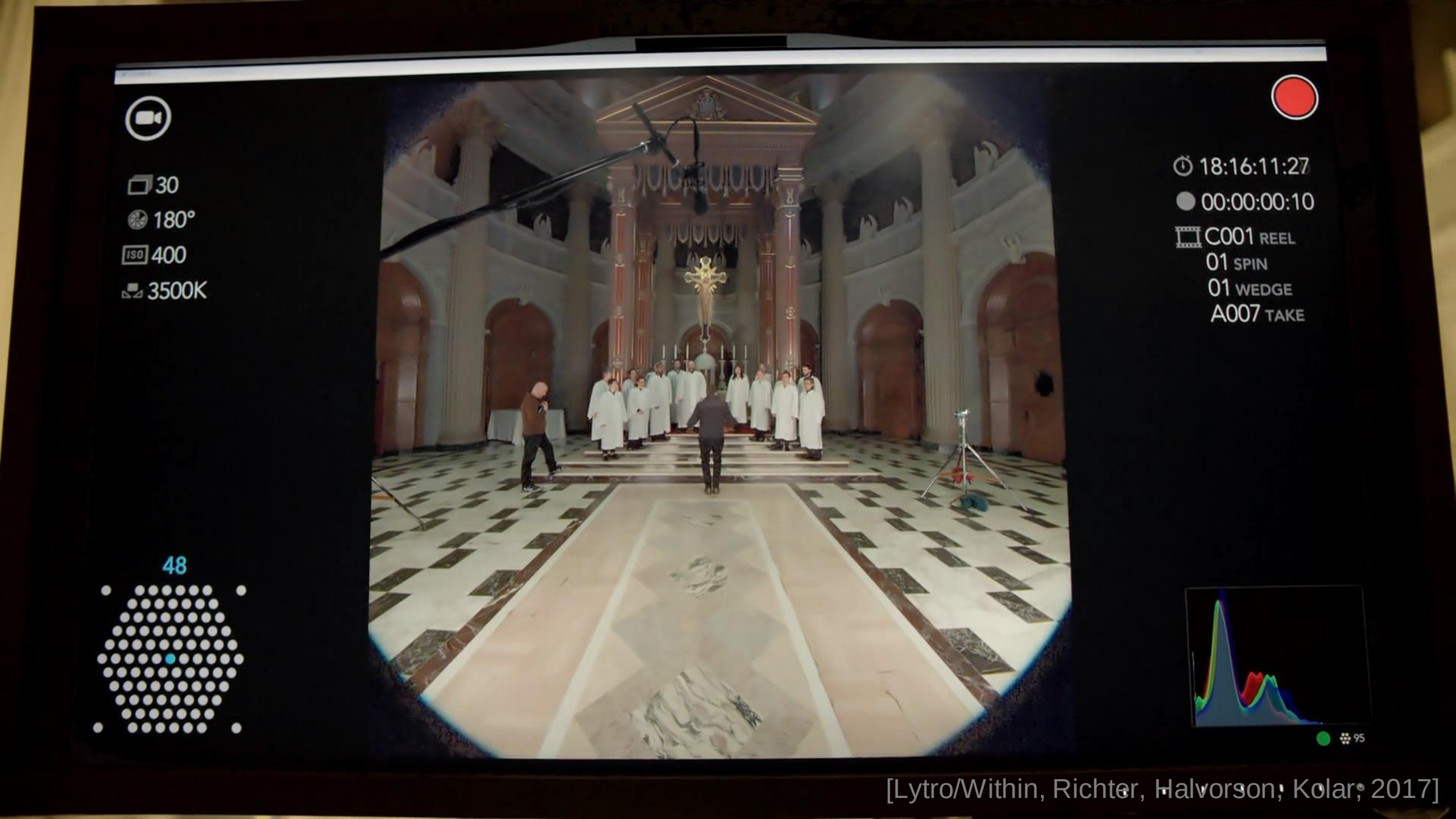

[Images © Sony - https://youtu.be/poI6E8O4jmQ]

# VR VILLAGE @ SIGGRAPH 2017


Hallelujah: Creating a Breakthrough VR Experience with Lytro Immerge



[Lytro/Within, Richter, Halvorson, Kolar, 2017]




[Lytro/Within, Richter, Halvorson, Kolar, 2017]

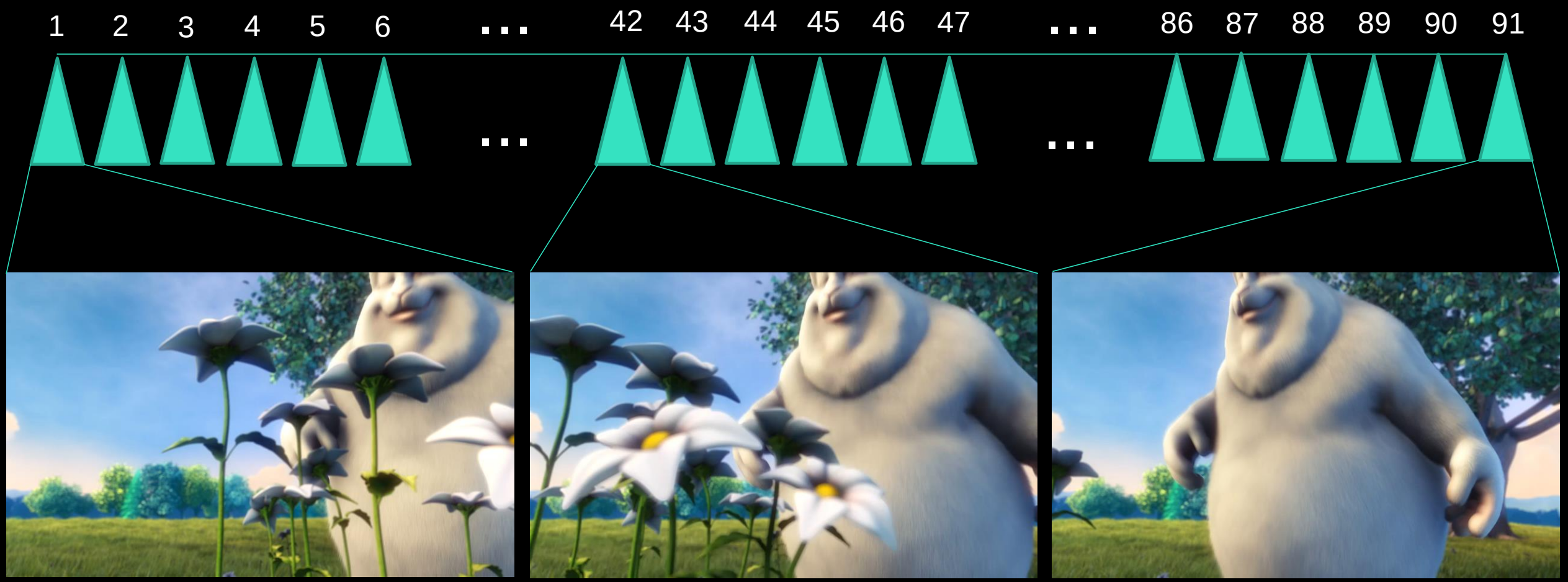

[Lytro/Within, Richter, Halvorson, Kolar, 2017]


[Lytro/Within, Richter, Halvorson, Kolar, 2017]

# EDITING LIGHT FIELDS

- Many operations we take for granted on 1-camera video are not well defined on light fields!
- Any editing operation must be efficient.
- Example: filtering

# EXAMPLE UNPROCESSED LIGHT FIELD VIDEO




[Big Buck Bunny / Kovács et al. / Holografika]

# LIGHT FIELD VIDEO – CYCLE THROUGH CAMERA VIEWS

[Bonneel, Tompkin, Wang et al., Eurographics 2017,
Consistent Video Filtering for Camera Arrays]

[Big Buck Bunny / Kovács et al. / Holografika]

# LIGHT FIELD VIDEO FILTERING
# INTRINSIC DECOMPOSITION (SHADING)

*Inconsistency across space and across time!*

[Bonneel, Tompkin, Wang et al., Eurographics 2017,
Consistent Video Filtering for Camera Arrays]

[Big Buck Bunny / Kovács et al. / Holografika]

# LIGHT FIELD VIDEO FILTERING
# INTRINSIC DECOMPOSITION (SHADING)



Input

Processed

[Bonneel, Tompkin, Wang et al., Eurographics 2017, Consistent Video Filtering for Camera Arrays]

[Big Buck Bunny / Kovács et al. / Holografika]

# LIGHT FIELD VIDEO FILTERING
# SPATIO-TEMPORAL CONSISTENCY

Our result

[Bonneel, Tompkin, Wang et al., Eurographics 2017,
Consistent Video Filtering for Camera Arrays]

[Big Buck Bunny / Kovács et al. / Holografika]

# 6DoF Production

Jordan
Halsey

VR Playhouse