3D-Kernel Foveated Rendering for Light Fields

IEEE Transactions on Visualization and Computer Graphics (TVCG)

Xiaoxu Meng[1], Ruofei Du[2], Joseph F. JaJa[1] and Amitabh Varshney[1]

[1]University of Maryland, College Park *[2]Google LLC*

Content

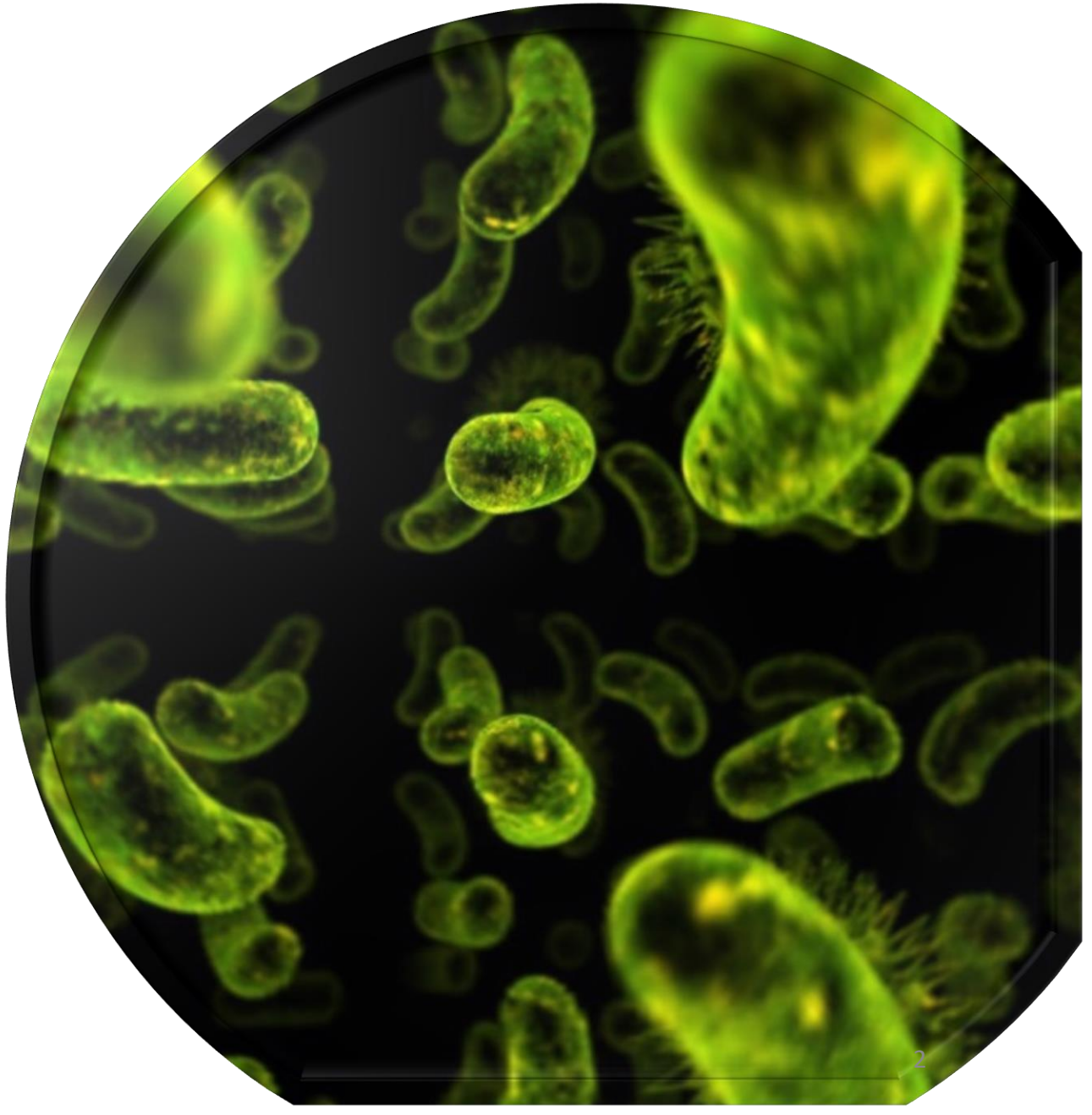
Motivation

Foveated Rendering

Our Approach

User Study

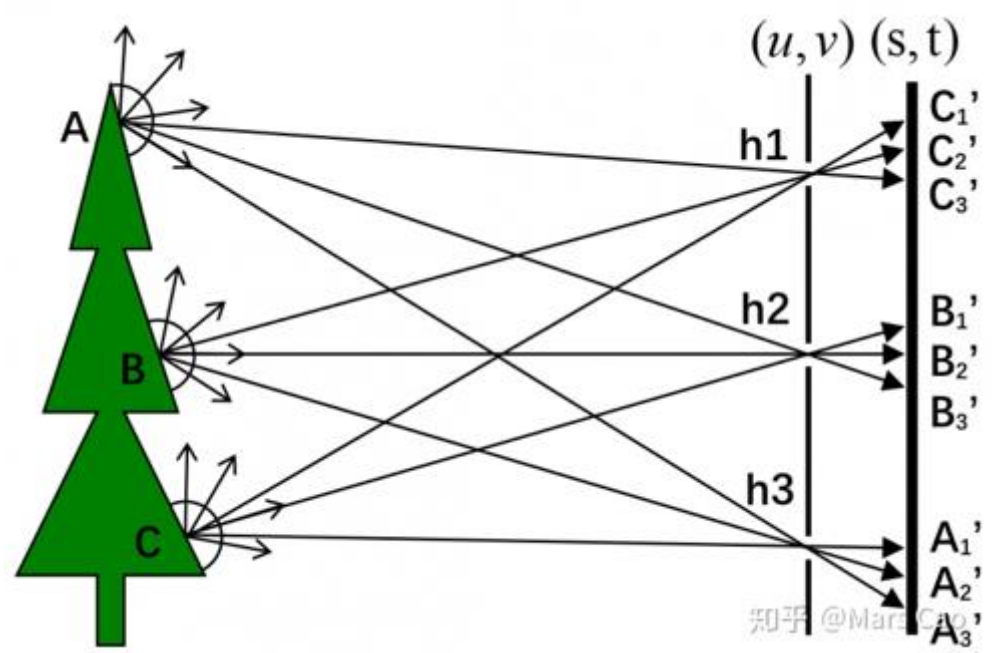
Rendering Acceleration



Light Field $L(u, v, s, t)$

uv - camera plane

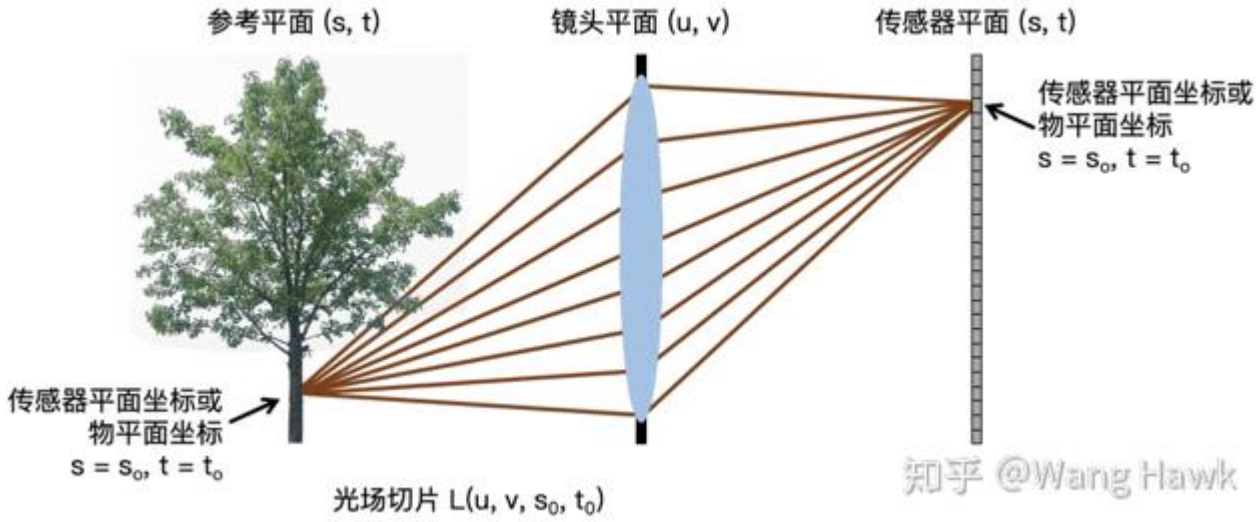
st - focal plane

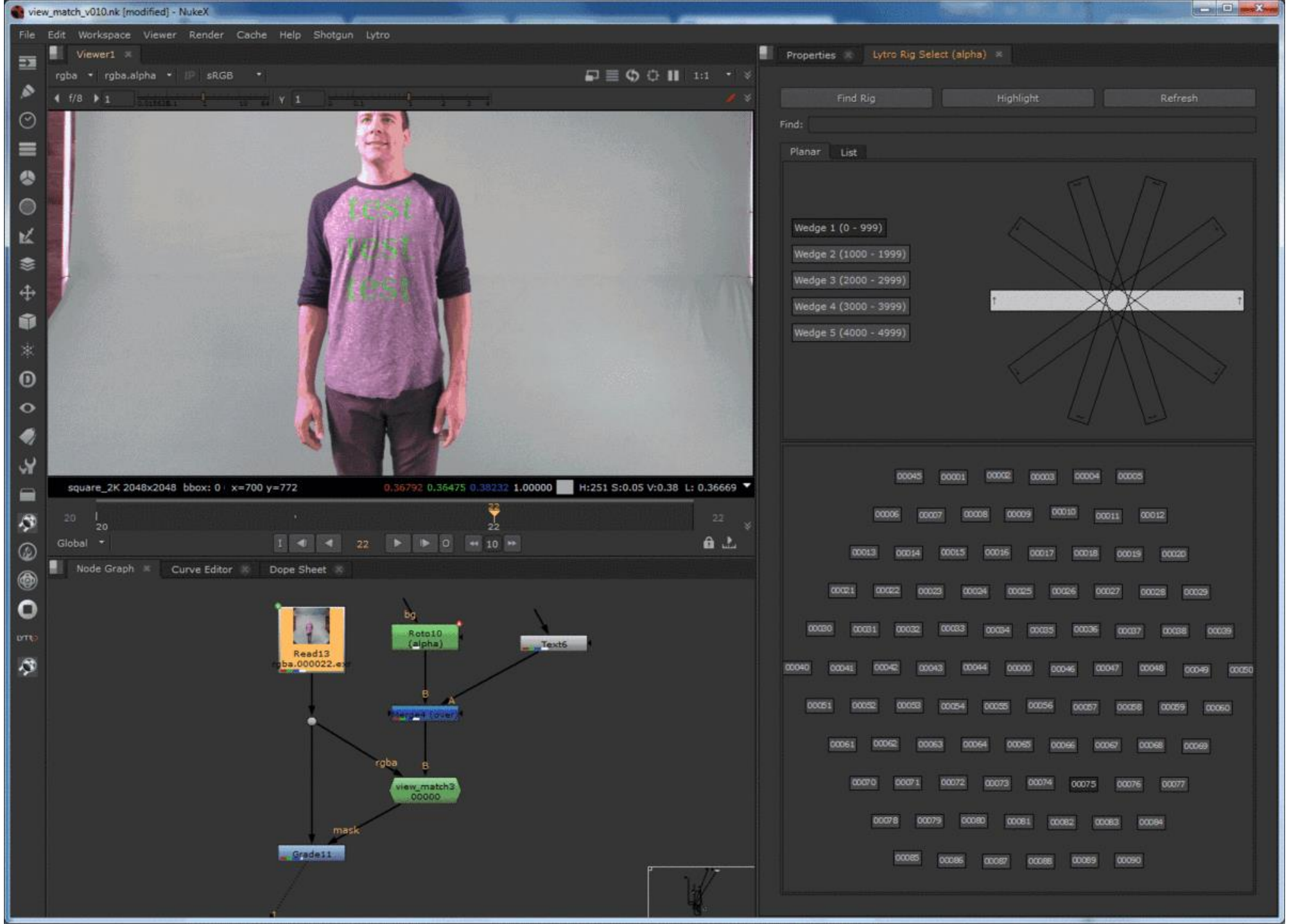


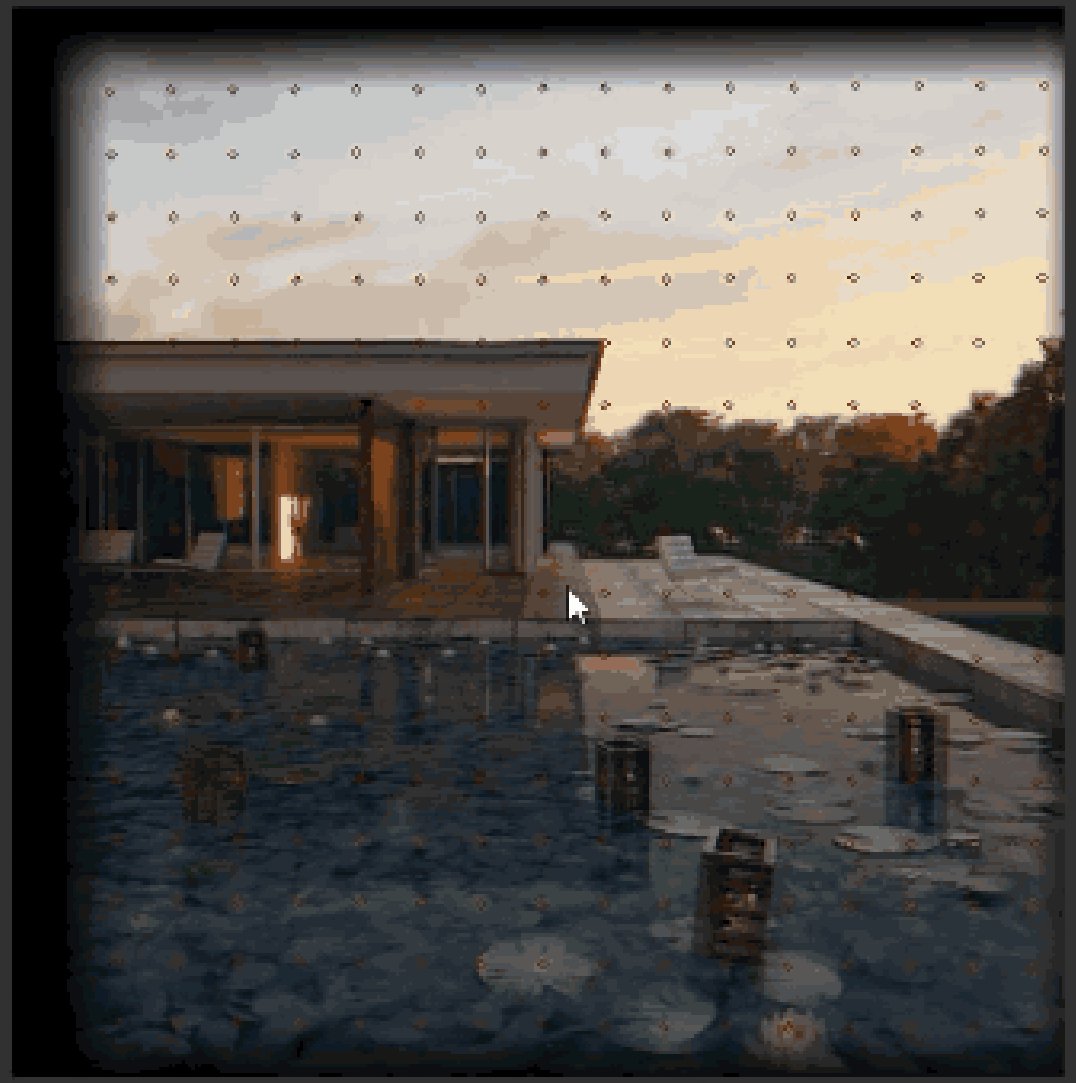
Light Field $L(u, v, s, t)$

uv - camera plane

st - focal plane




view_match_v010.nk [modified] - NukeX
File Edit Workspace Viewer Render Cache Help Shotgun Lytro
Viewer1
rgba rgba.alpha sRGB
test test test
square_2K 2048x2048 bbox: 0 x=700 y=772 0.36792 0.36475 0.38232 1.00000 H:251 S:0.05 V:0.38 L: 0.36669
Global
Node Graph
Curve Editor
Dope Sheet
Read13
bg
Roto10 (alpha)
Text6
B
A
rgba
B
view_match3 00000
mask
Grade11
Properties
Lytro Rig Select (alpha)
Find Rig
Highlight
Refresh
Find:
Planar
List
Wedge 1 (0 - 999)
Wedge 2 (1000 - 1999)
Wedge 3 (2000 - 2999)
Wedge 4 (3000 - 3999)
Wedge 5 (4000 - 4999)



Total: 57ms. Prepare: 0ms. Render: 54ms. Blit: 1ms.

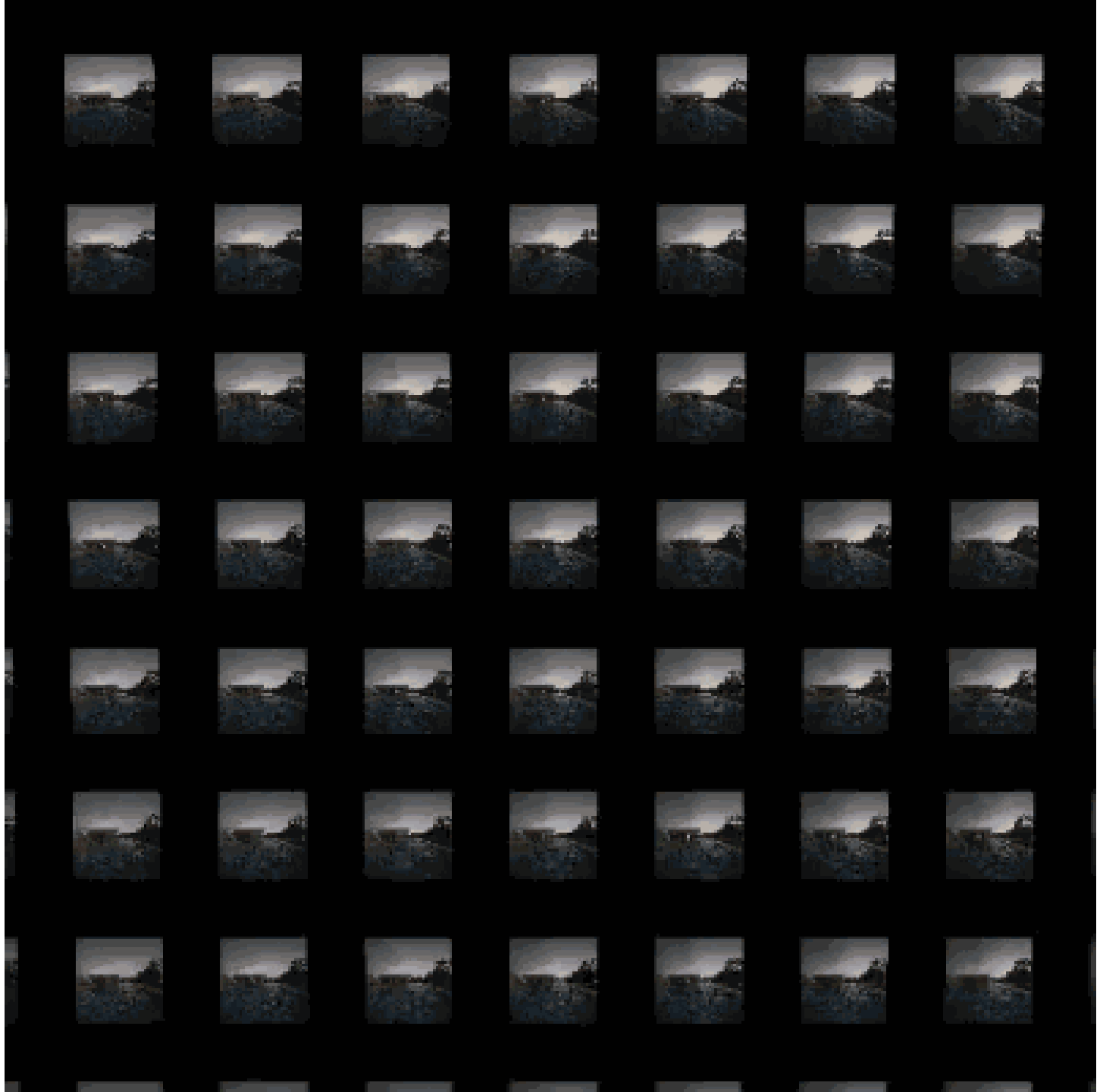
Light field info

Camera plane: 16 × 16
Capture interval: 10 cm
Image plane: 1024 × 1024
Capture density: 100 cameras/m²
Ray count: 268,44 Megarays
Fov: 90°

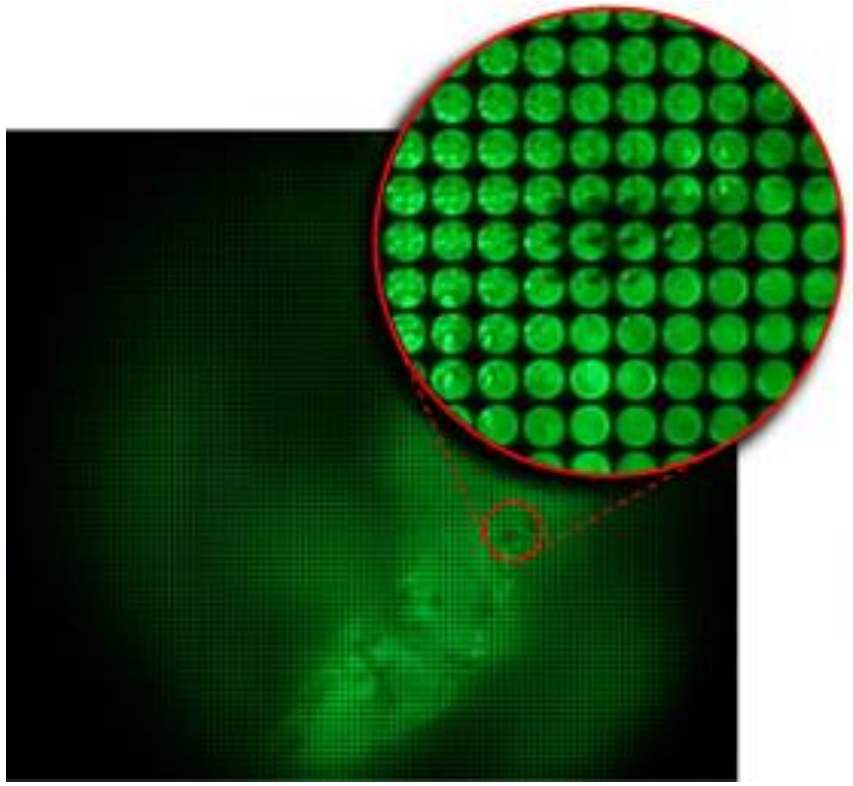
Viewpoint camera

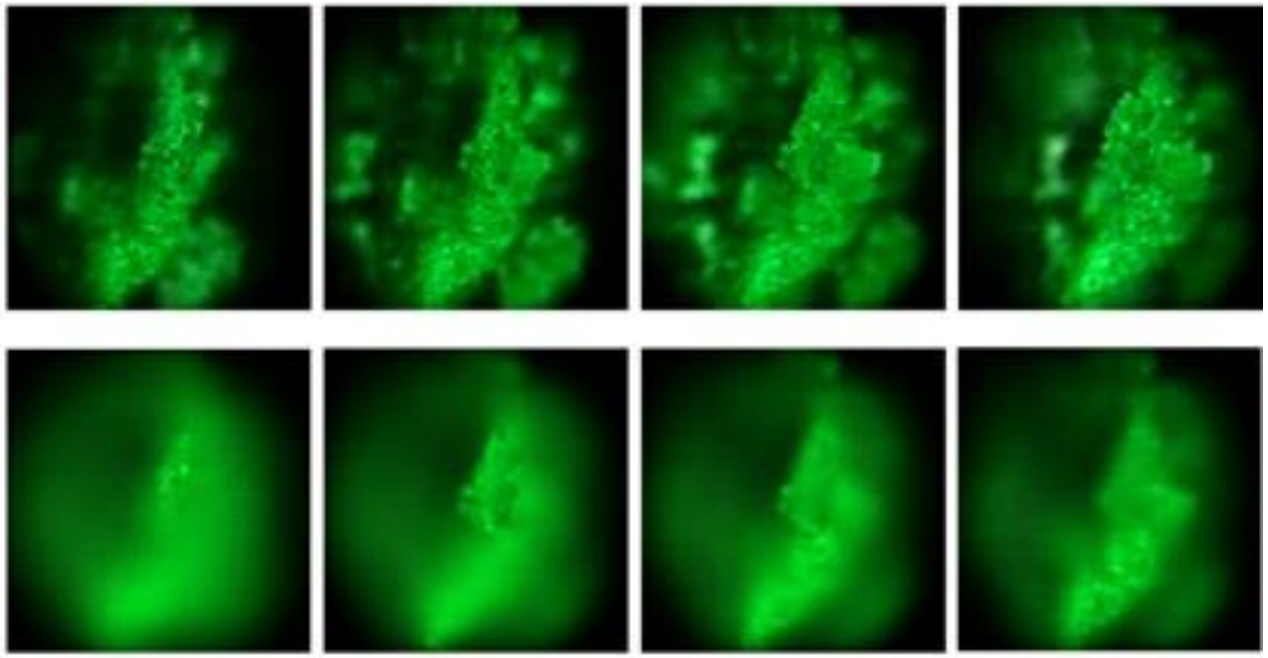
Position (m): X: 0,296 Y: -0,748 Z: -1,936
Rotation: X: 0,000 Y: -0,138 Z: 0,000
Image size (px): 512 × 512
FOV (°): 47,000
Focusing plane (m): 15,000
Aperture (cm): 0,200 Bilinear

Geometry




Light field “autostereoscopic”



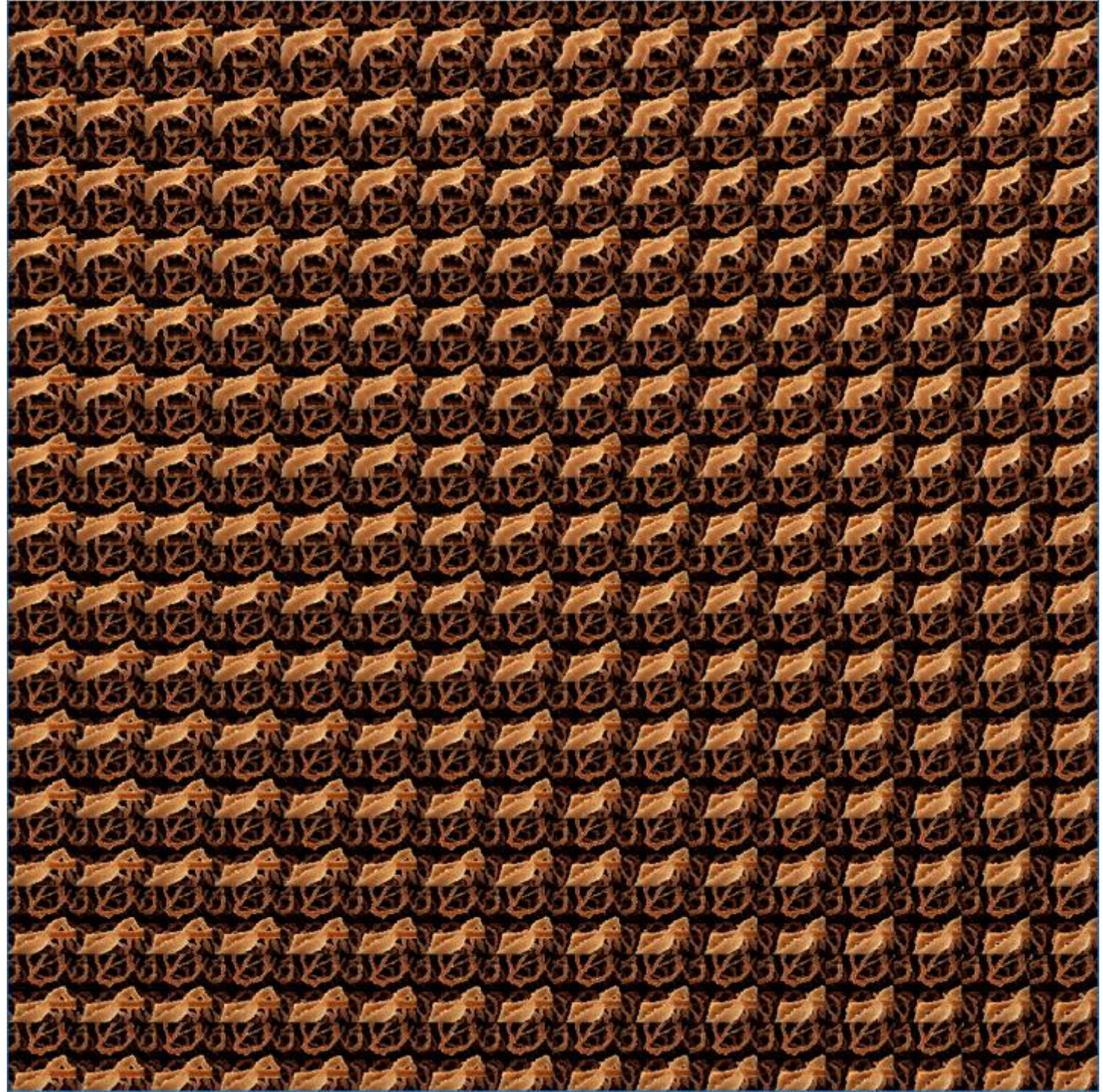


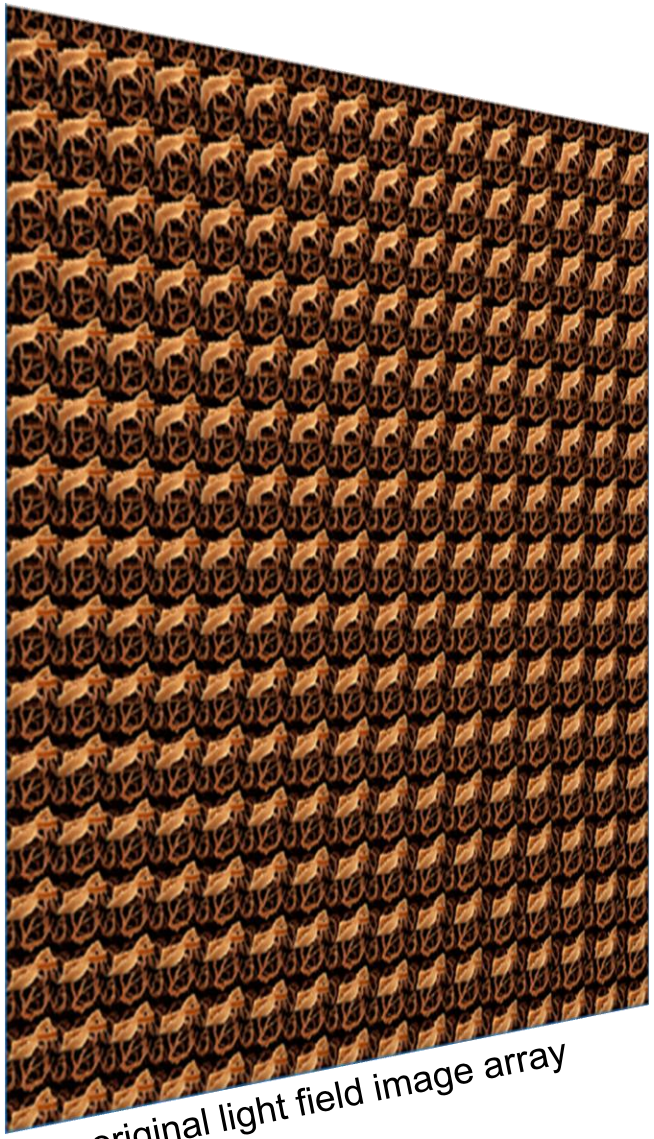
a light field captured by photographing a speck of fluorescent crayon wax through a microscope objective and microlens array

"Since microscopes are inherently orthographic devices,

perspective views represent a new way to look at microscopic specimens." [1]

[1] Marc Levoy, Ren Ng, Andrew Adams, Matthew Footer, and Mark Horowitz. 2006. Light field microscopy. *ACM Trans. Graph.* 25, 3 (July 2006), 924-934. DOI: https://doi.org/10.1145/1141911.1141976

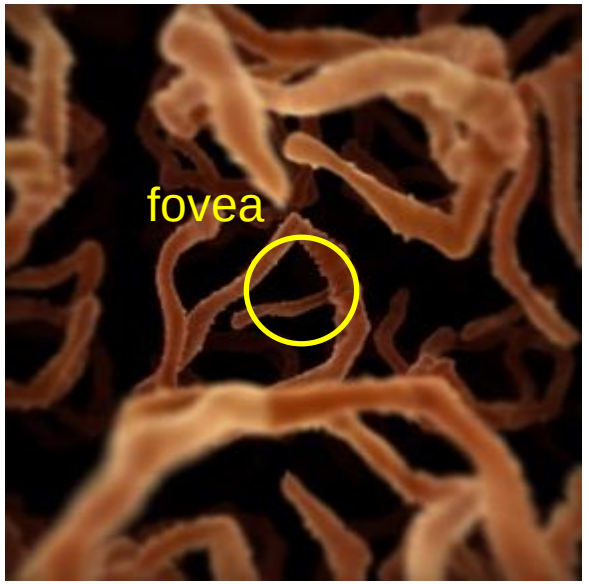




original light field image array

foveated rendering

Screen



Content

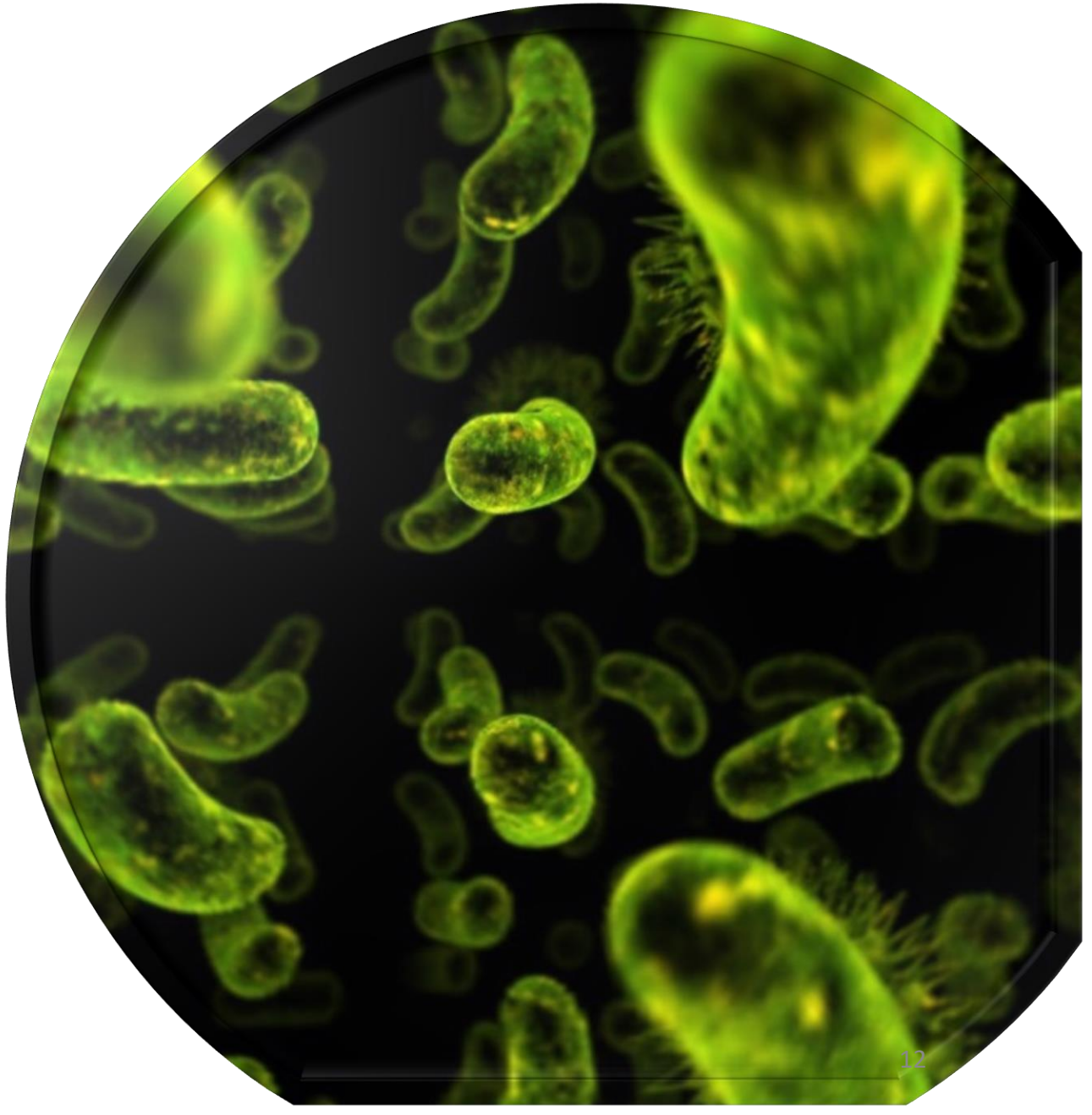
Motivation

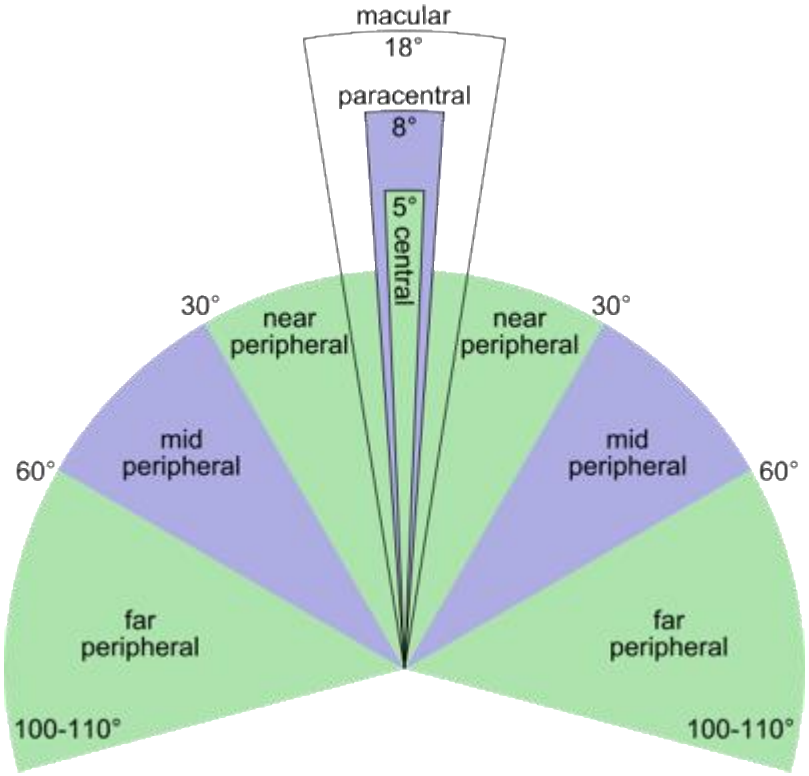
Foveated Rendering

Our Approach

User Study

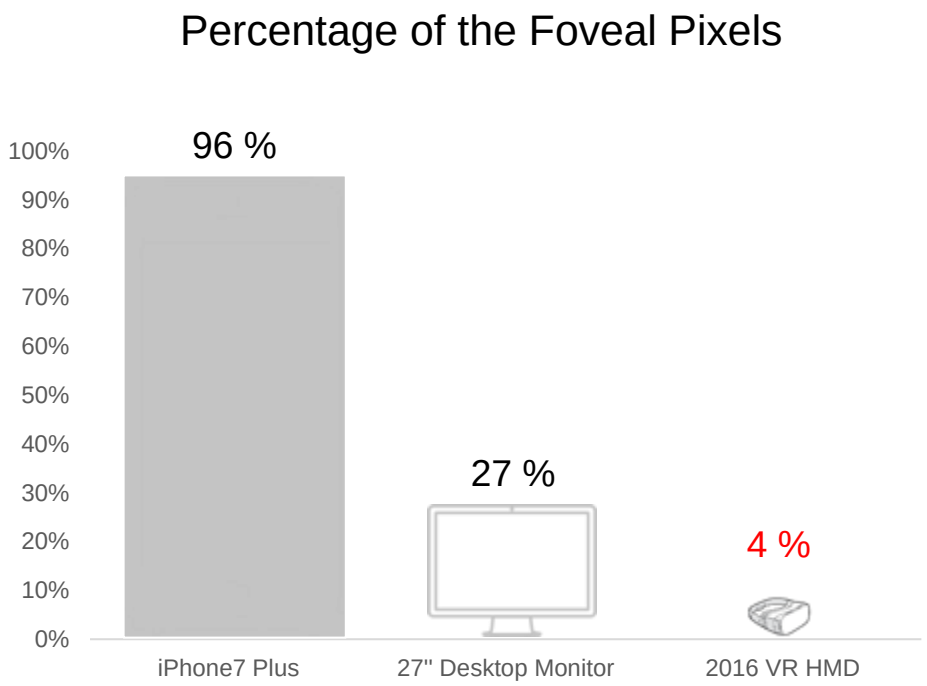
Rendering Acceleration





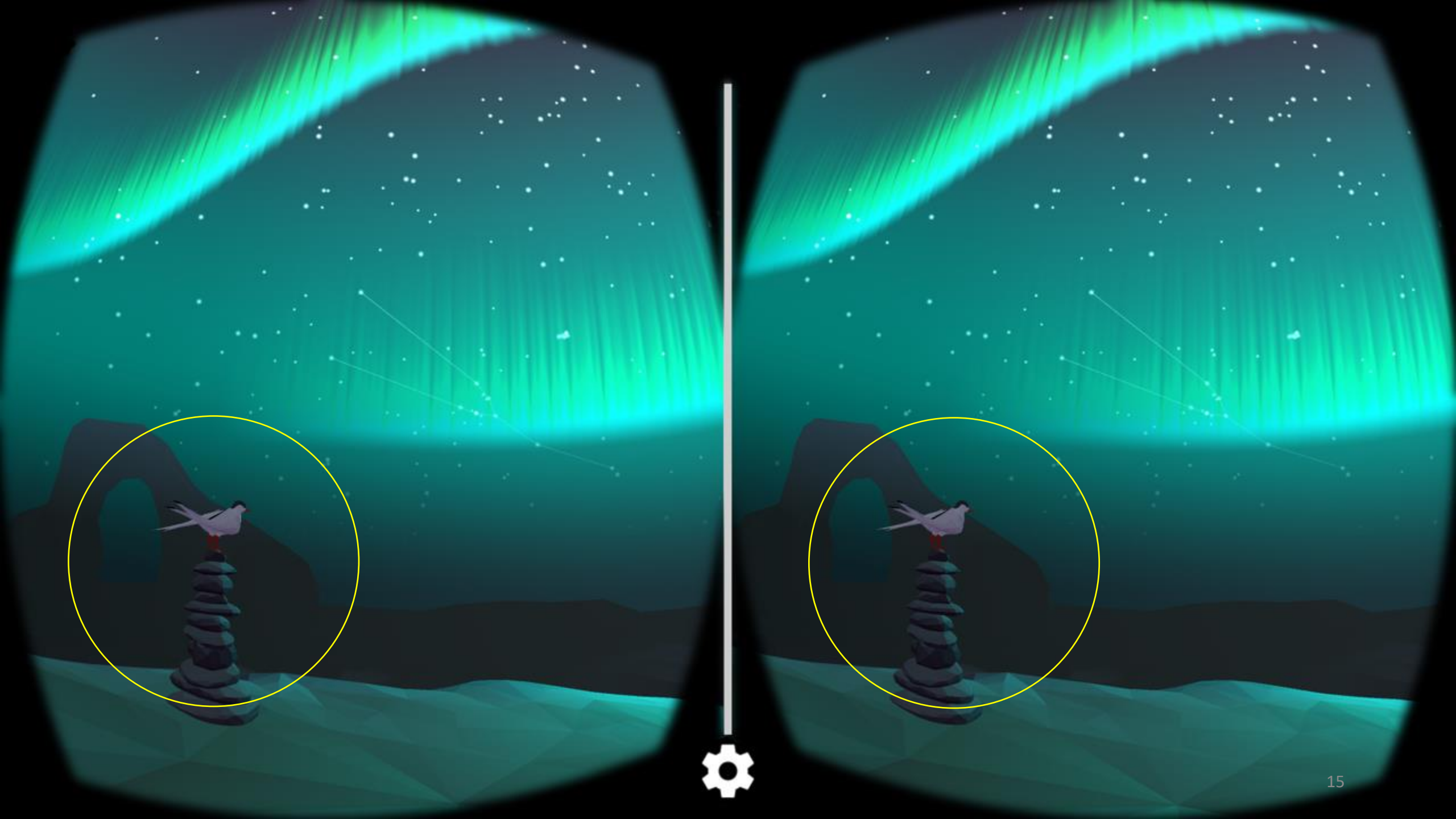
***foveal region**:* the human eye detects significant high-fidelity detail
***peripheral region**:* the human eye detects little high-fidelity detail

- Virtual reality is a challenging workload
- Most VR pixels are peripheral

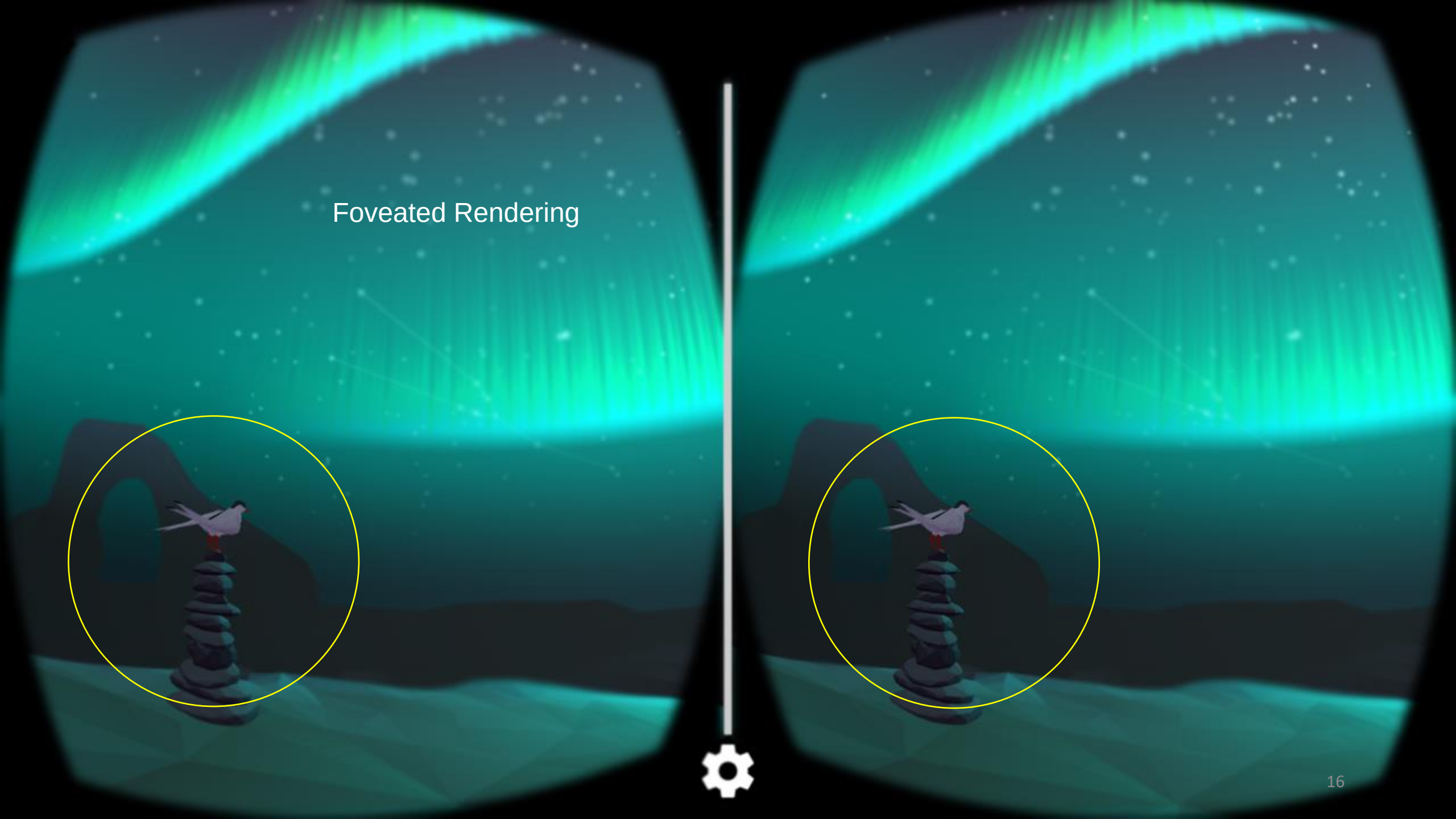


- Virtual reality is a challenging workload
- Most VR pixels are peripheral

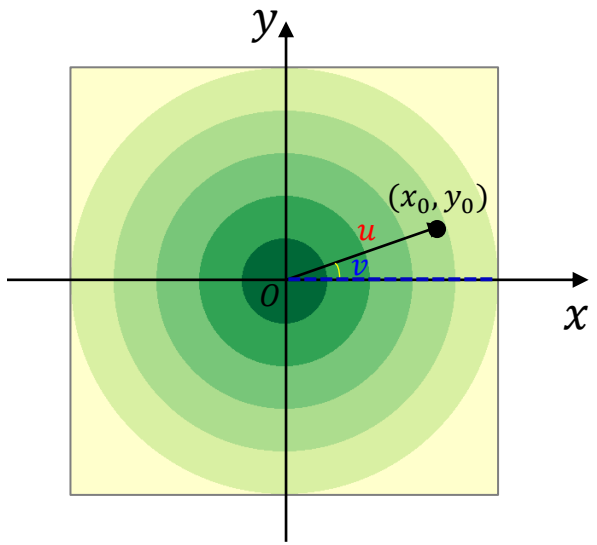
* Data from Siggraph 2017, by Anjul Patney, August 2017



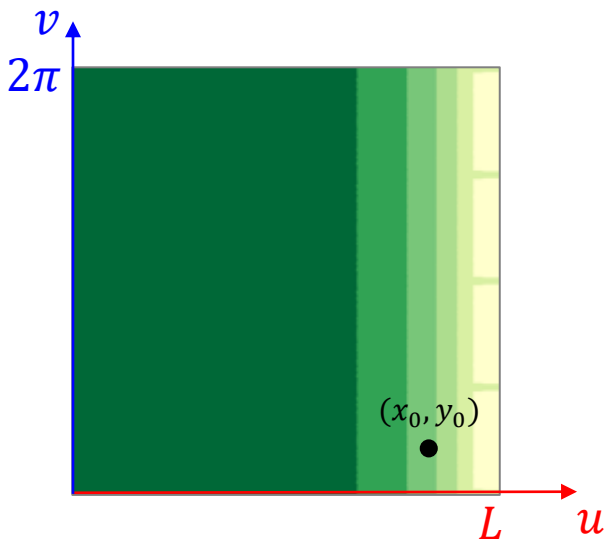
Foveated Rendering



Log-polar mapping [Araujo and Dias 1996]



Cartesian coordinates
(x, y)



Log-polar coordinates
(u, v)

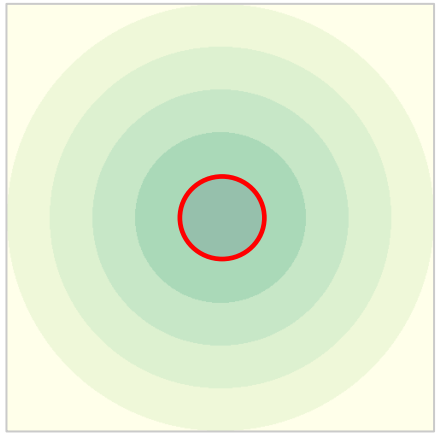
Log-polar Mapping

$$u = \frac{\log\sqrt{x^2 + y^2}}{L} \cdot w$$

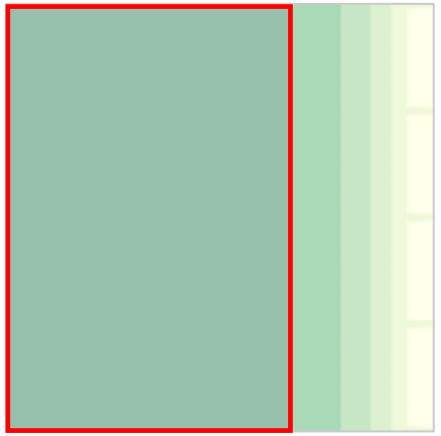
$$v = \frac{(\arctan\frac{y}{x} + \mathbf{1}\,[y - 0] \cdot 2\pi)}{2\pi} \cdot h$$

- $W: screen\ width \quad H: screen\ height \quad w: buffer\ width \quad h: buffer\ height$
- $\mathbf{1}\,[y < 0] = \begin{cases} 1 & y < 0 \\ 0 & y > 0 \end{cases}$
- $L = \log\sqrt{W^2 + H^2}$

Log-polar mapping [Araujo and Dias 1996]



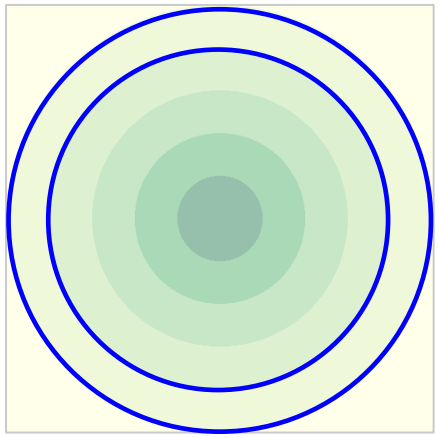
Cartesian coordinates
(x, y)



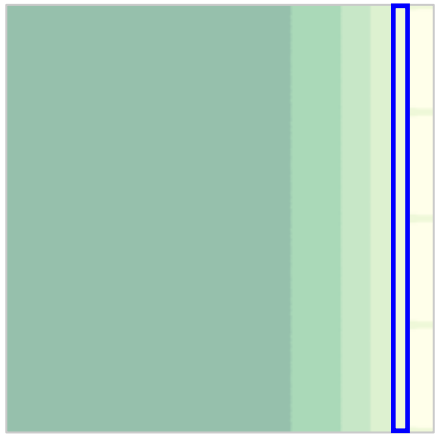
Log-polar coordinates
(u, v)

The closer the point is to the center, the higher pixel density it preserves.

Log-polar mapping [Araujo and Dias 1996]



Cartesian coordinates
(x, y)



Log-polar coordinates
(u, v)

Points in the peripheral regions are naturally compressed in the log polar domain.

Log-polar mapping [Araujo and Dias 1996]

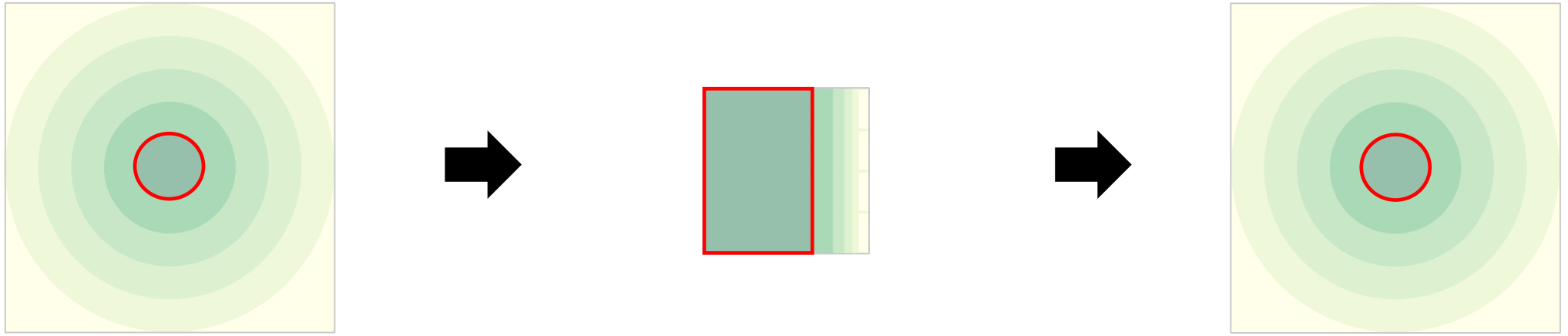
Cartesian coordinates
(x, y)

Log-polar coordinates
(u, v)

Cartesian coordinates
(x, y)

We can recover the image by performing the inverse transformation

Log-polar mapping [Araujo and Dias 1996]



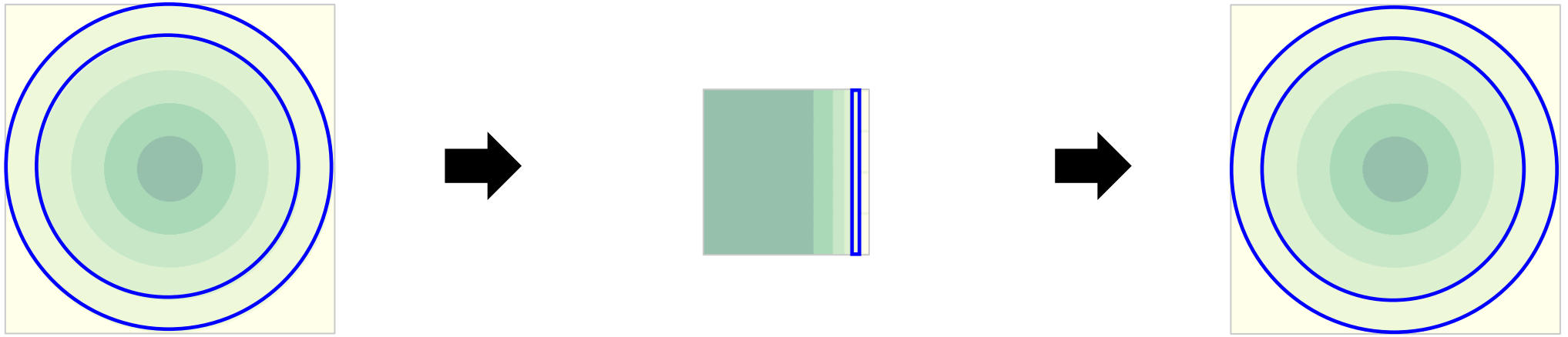
Cartesian coordinates
(x, y)

Log-polar coordinates
(u, v)

Cartesian coordinates
(x, y)

With smaller log-polar buffer, the foveal preserves the original pixel density while the peripheral regions are compressed.

Log-polar mapping [Araujo and Dias 1996]



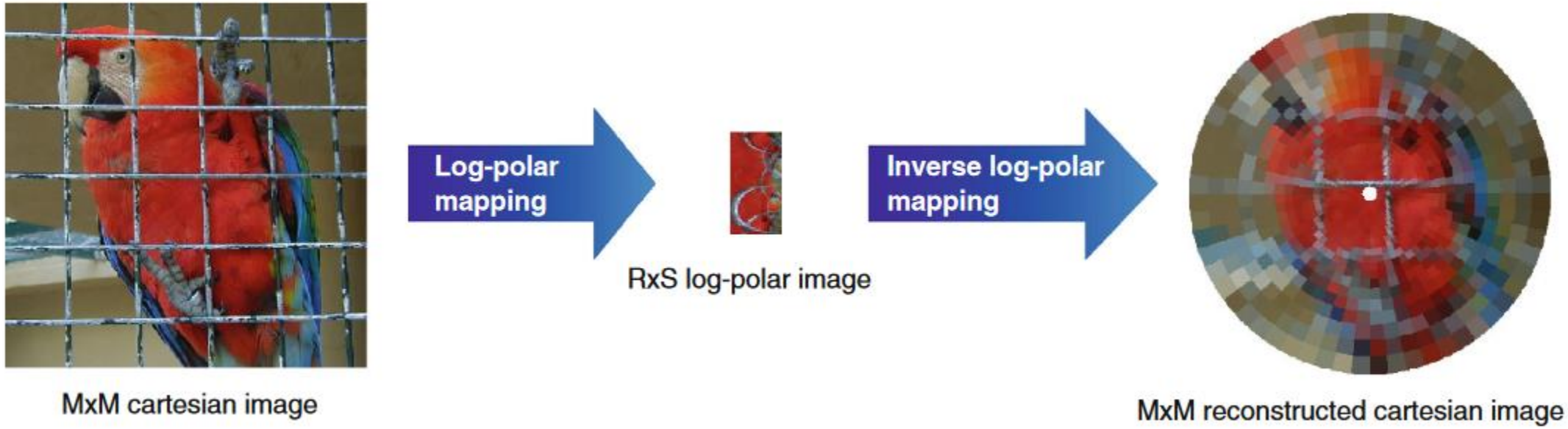
Cartesian coordinates
(x, y)

Log-polar coordinates
(u, v)

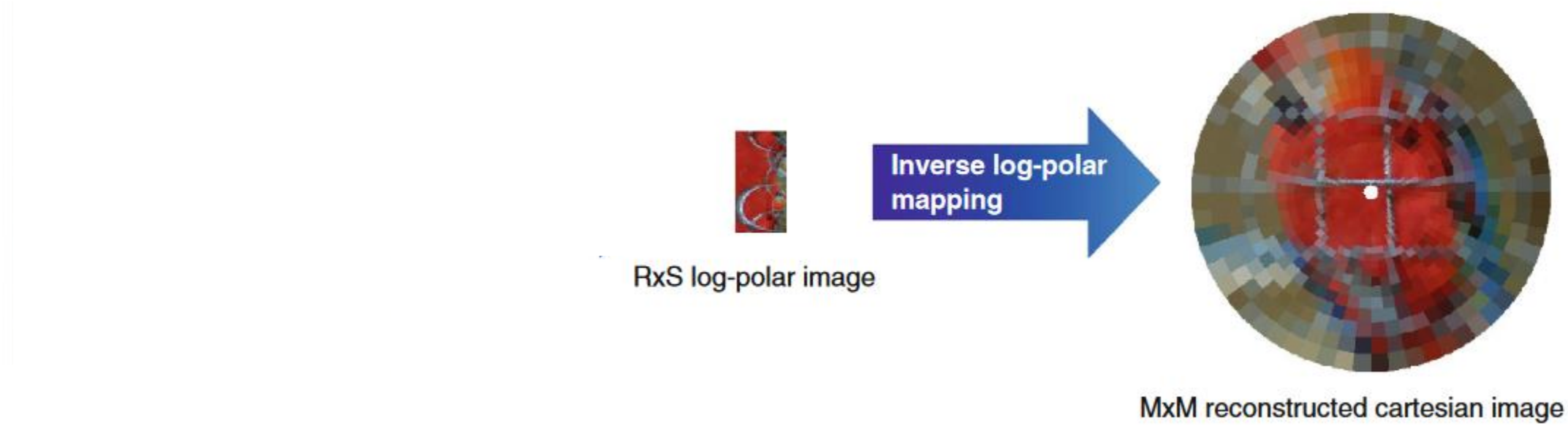
Cartesian coordinates
(x, y)

With smaller log-polar buffer, the foveal preserves the original pixel density while the peripheral regions are compressed.

Log-polar Mapping for 2D Image [Antonelli et al. 2015]

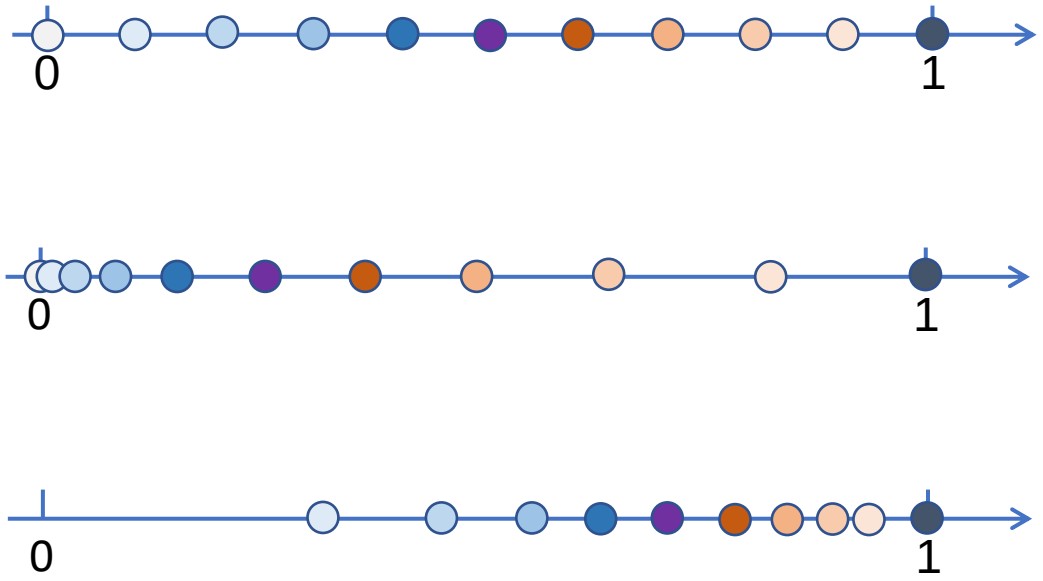


Log-polar Mapping for 2D Image



What if we directly render in the log-polar domain to reduce the rendering cost?
How to improve the quality in the peripheral regions?

Change the Distribution of Pixels (1D)



$K(x) = x^2$

$K(x) = x^{\frac{1}{2}}$

Kernel Log-polar Mapping

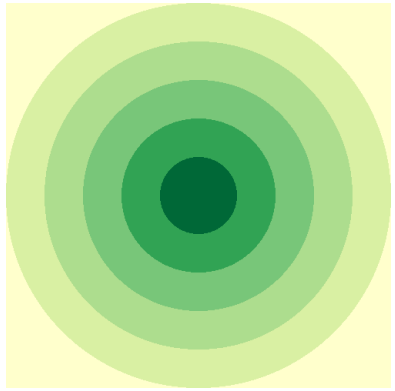


Image in
Cartesian Coordinates

$K(x) = x$

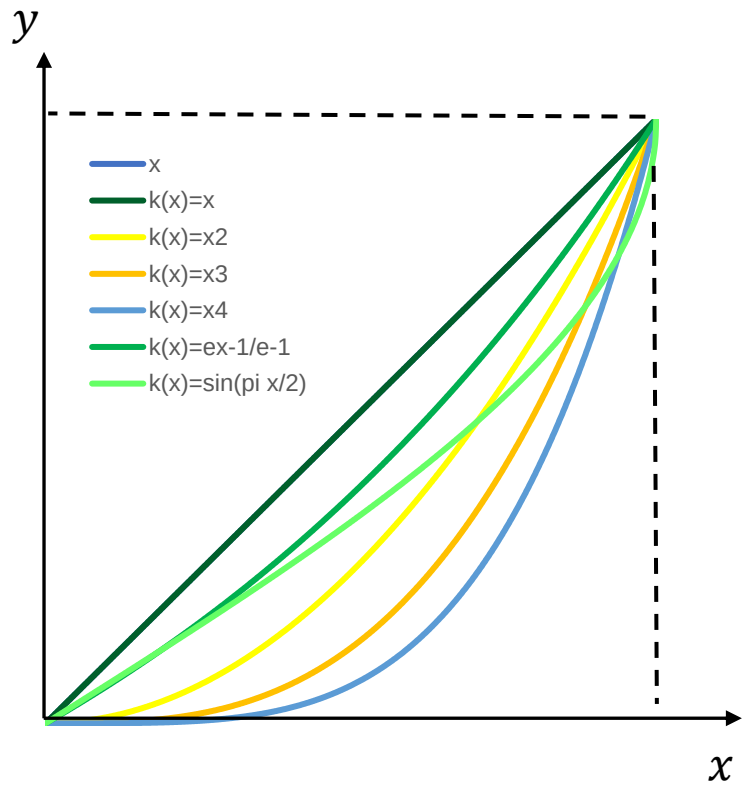
$K(x) = x^2$

$K(x) = x^3$

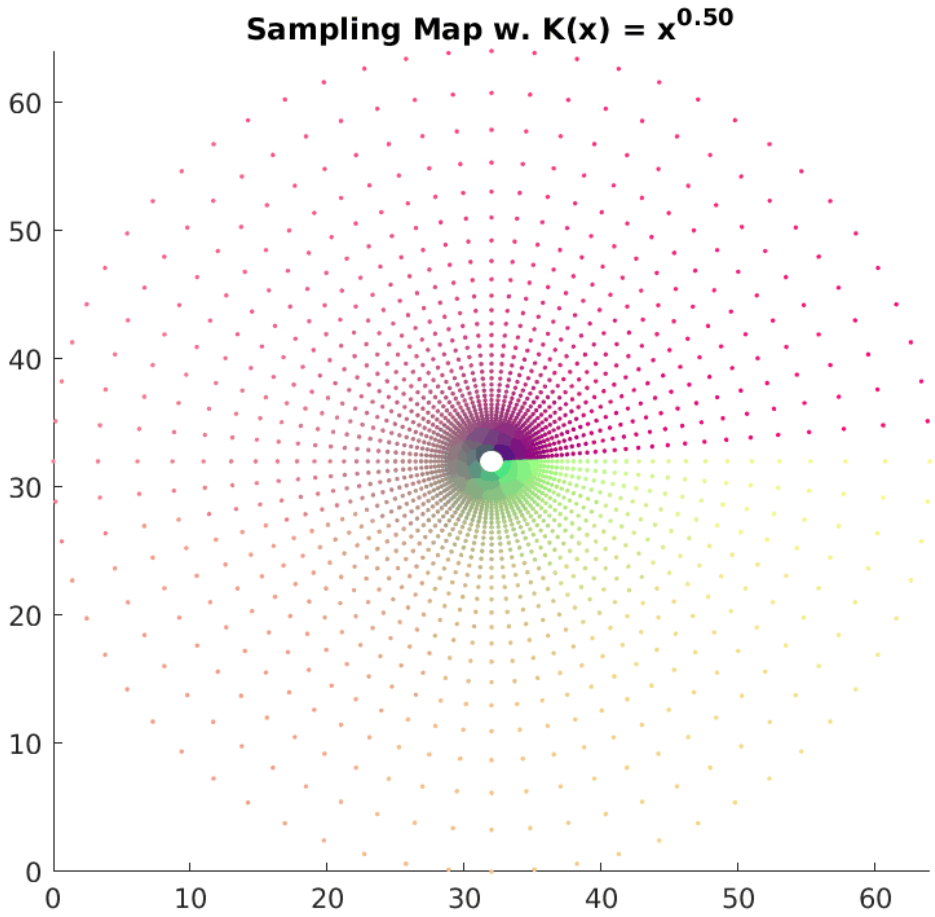
$K(x) = x^4$

$K(x) = \frac{e^x - 1}{e - 1}$

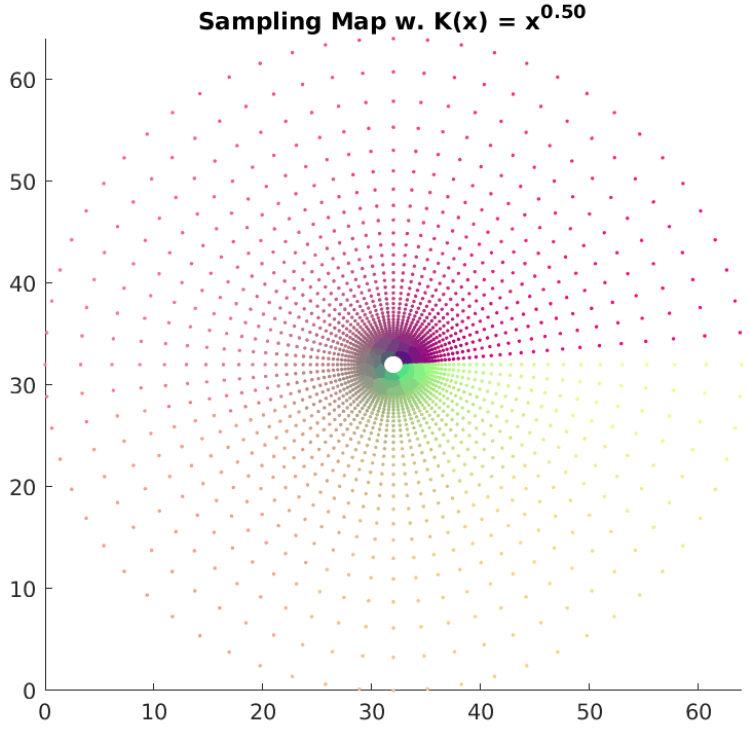
$K(x) = \sin(\frac{\pi}{2}x)$

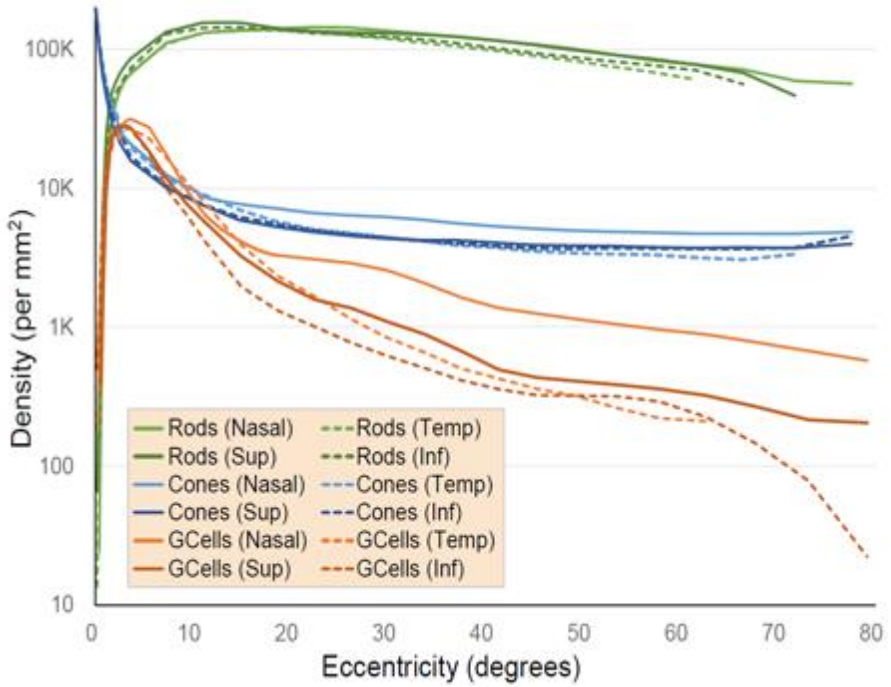


Change the Distribution of Pixels (2D)

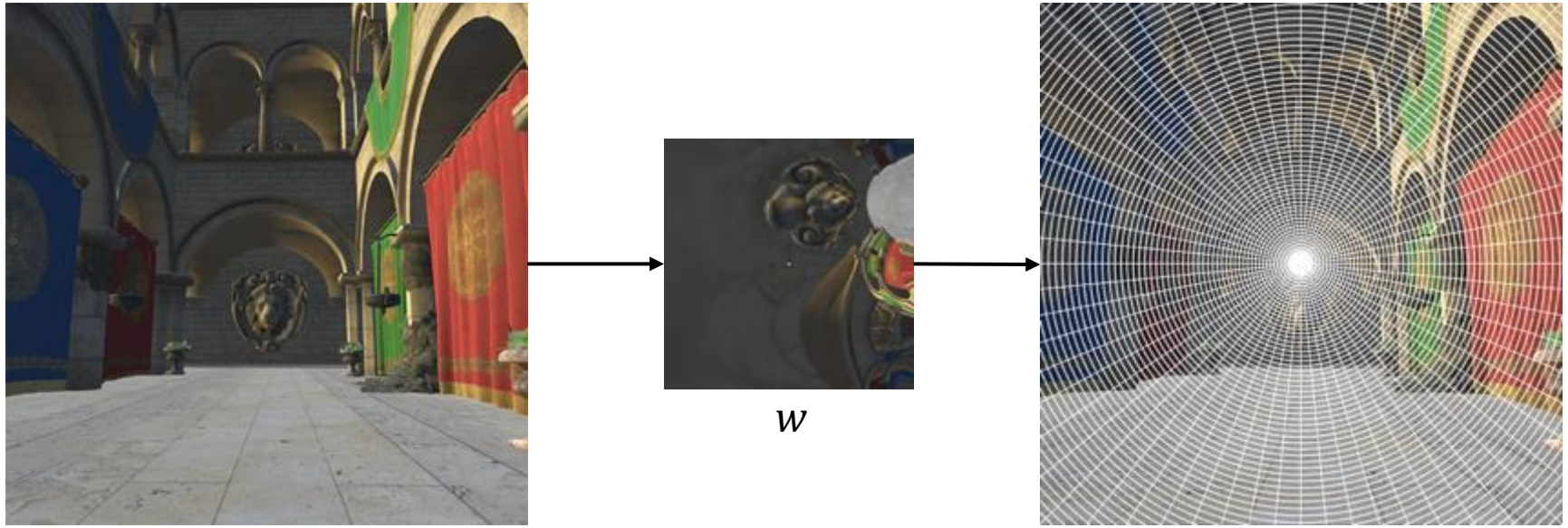


Kernel Foveated Rendering





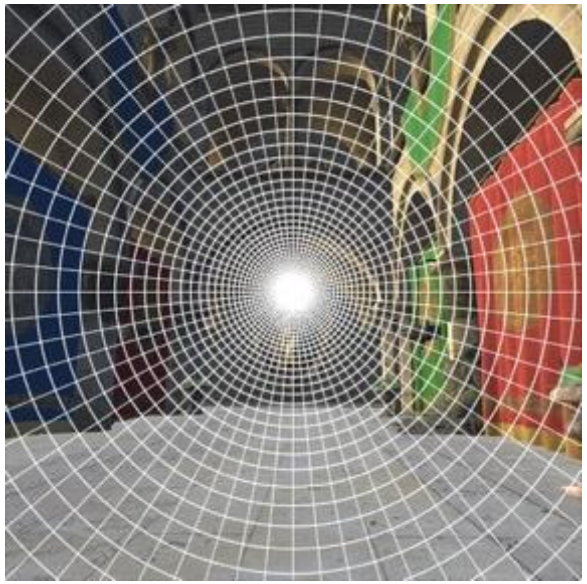
$$Distribution\ of\ pixels \xrightarrow{mimic} Distribution\ of\ photoreceptors\ in\ the\ human\ retina$$



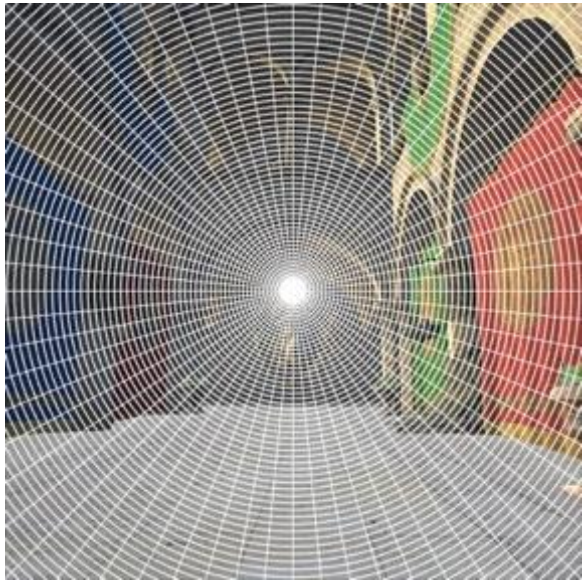
Kernel log-polar Mapping

- *Define buffer parameter σ*

$$\sigma = \frac{W}{w}$$



Result of log-polar
$(K(x) = x)$



Result of kernel log-polar
$(K(x) = x^4)$

Kernel log-polar Mapping

- *Define buffer parameter σ*

$$\sigma = \frac{W}{w}$$

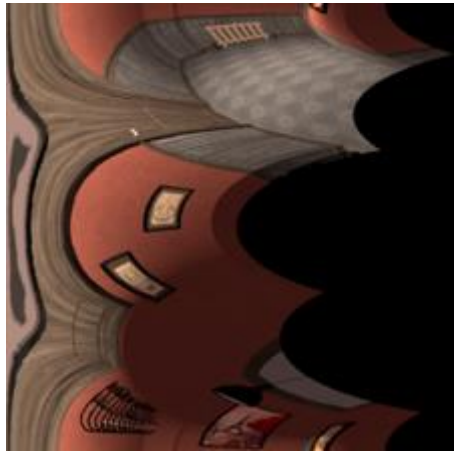
- *Define kernel function parameter α*

$$K(x) = x^{\alpha}$$

Buffer parameter
σ

Original Frame

Buffer



Screen

Sample Map

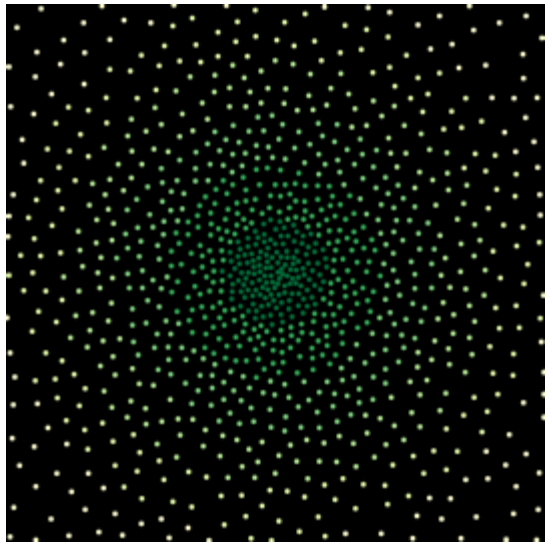
$$\sigma = 1.2$$

Original Frame

Buffer

Screen

Sample Map



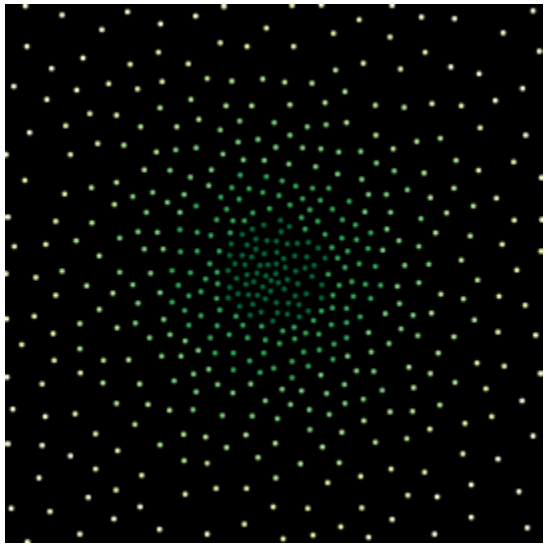
$$\sigma = 1.7$$

Original Frame

Buffer

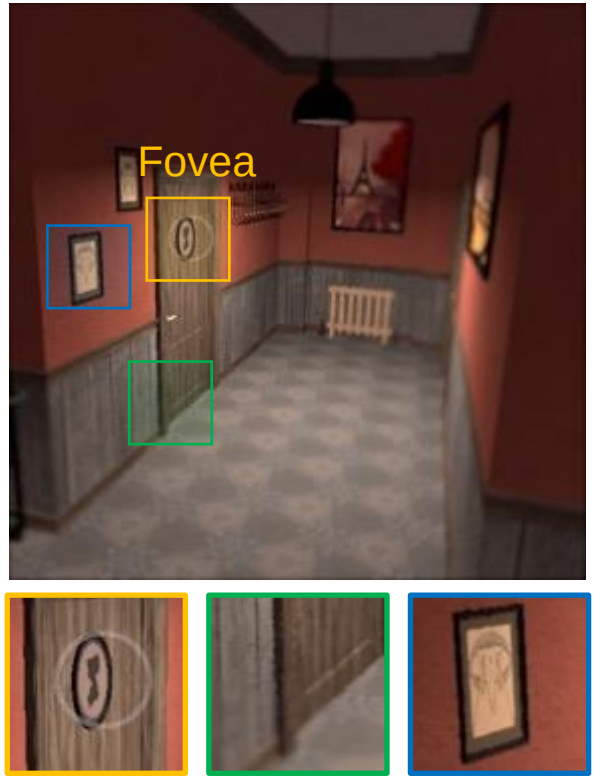
Screen

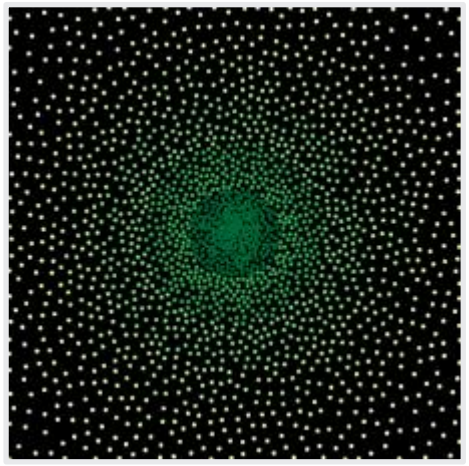
Sample Map



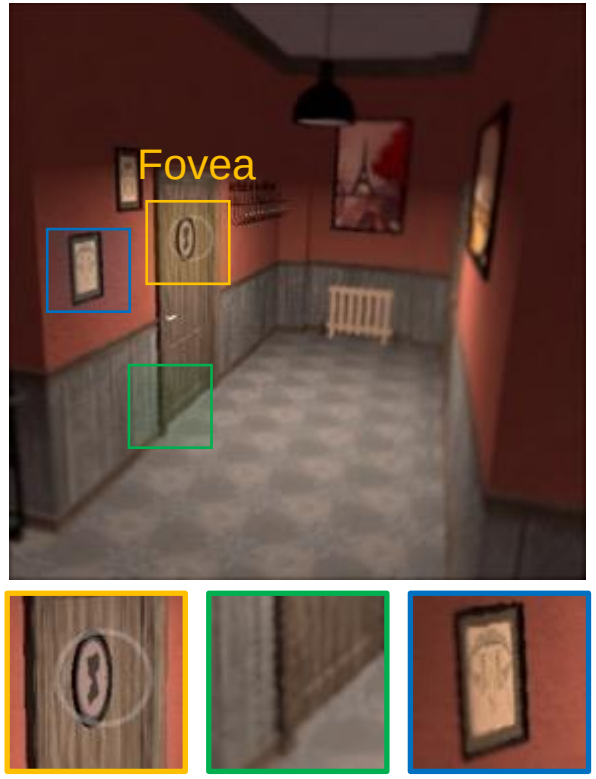
$$\sigma = 2.4$$

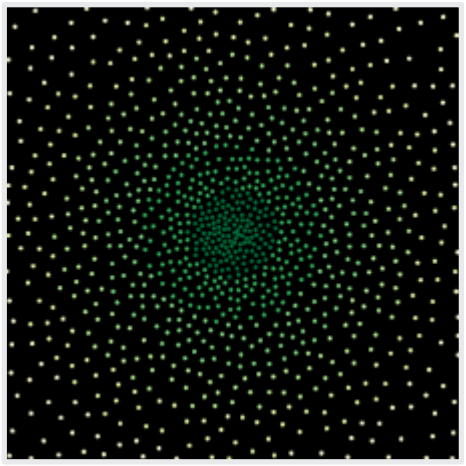
$\sigma = 1.2$



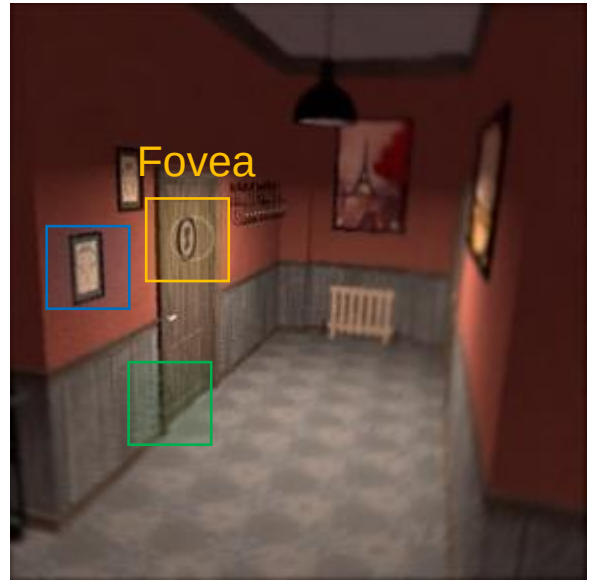


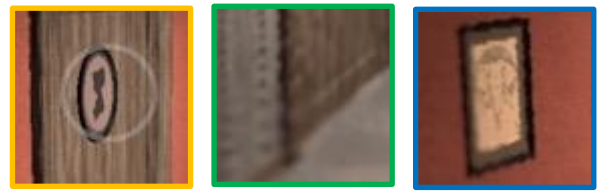
$\sigma = 1.7$

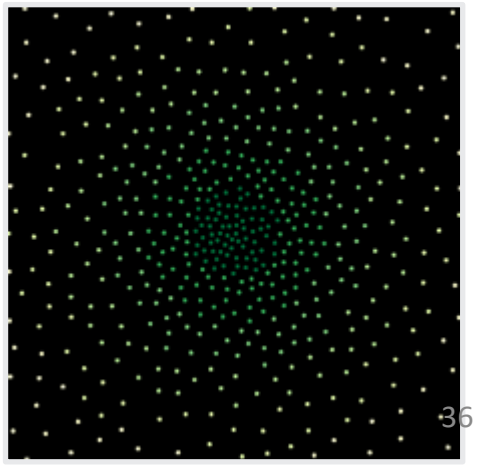




$\sigma = 2.4$







kernel function parameter

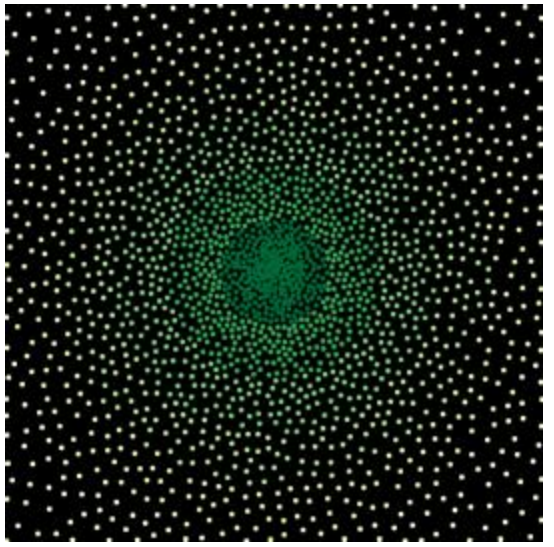
α

Original Frame

Buffer

Screen

Sample Map



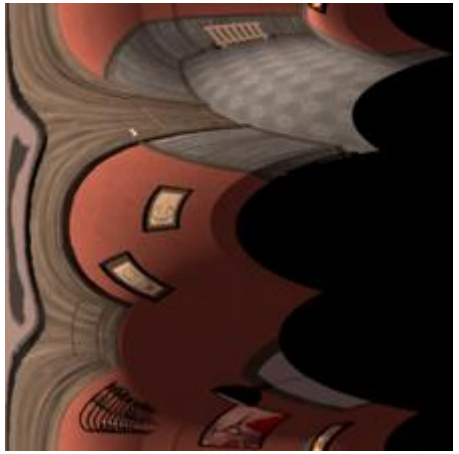
$$\alpha = 1$$

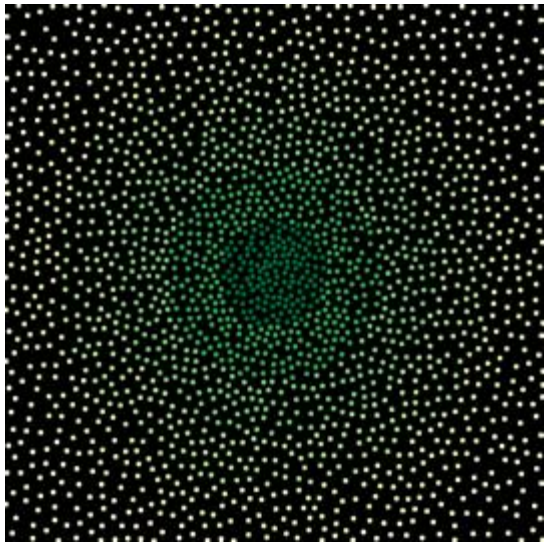
Original Frame

Buffer

Screen

Sample Map





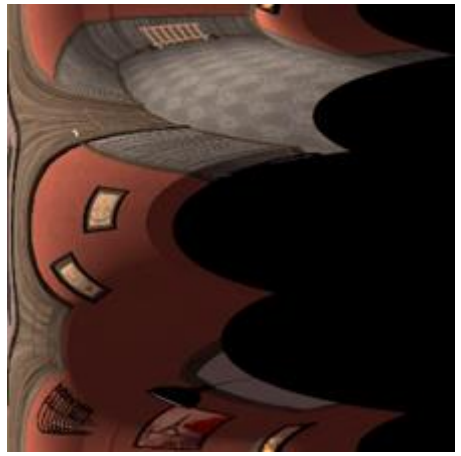
$$\alpha = 4$$

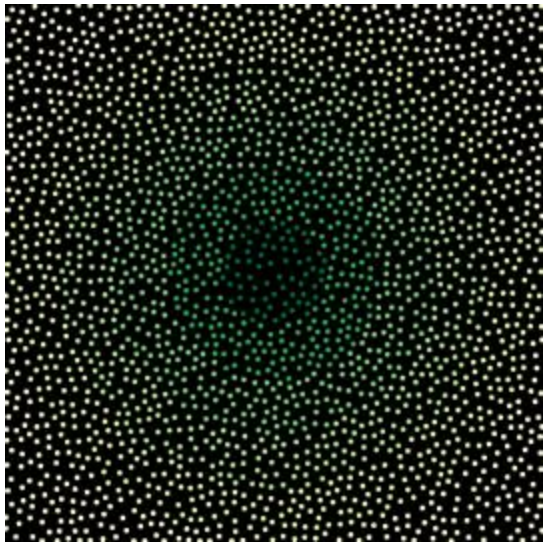
Original Frame

Buffer

Screen

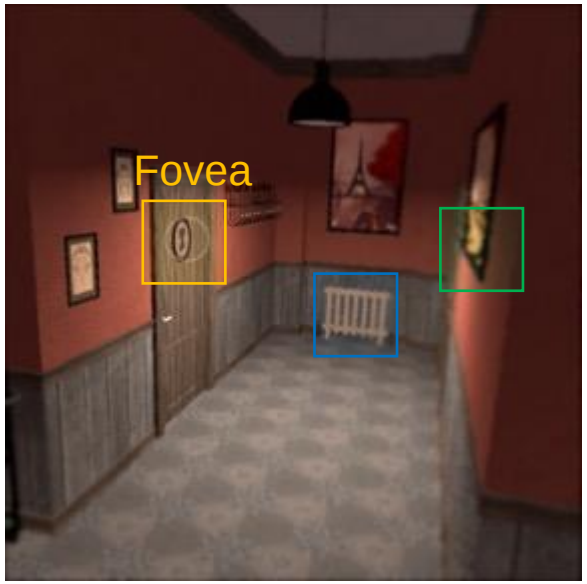
Sample Map



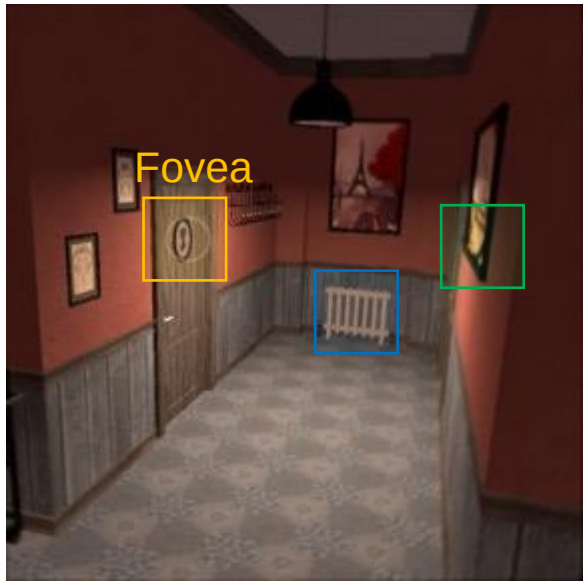


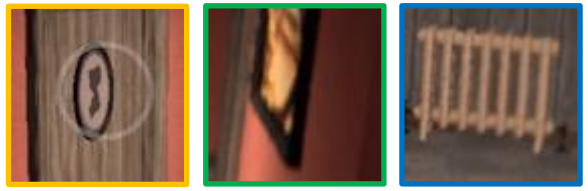
$$\alpha = 6$$

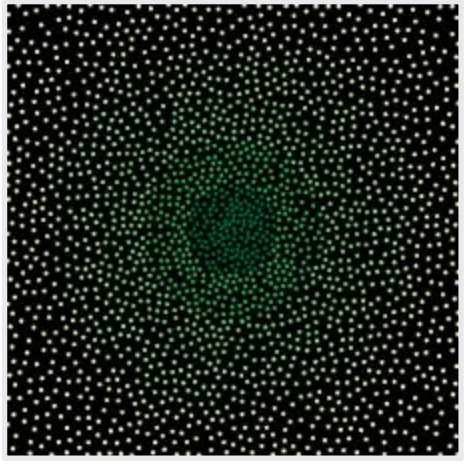
$\alpha = 1$



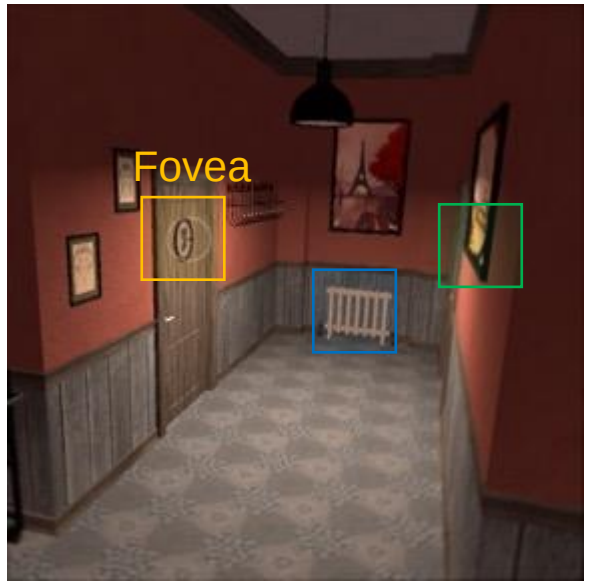
$\alpha = 4$

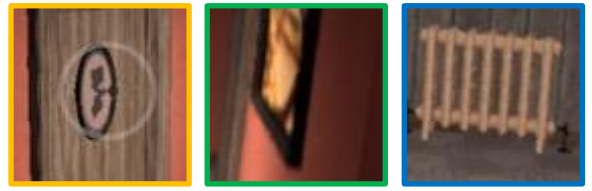


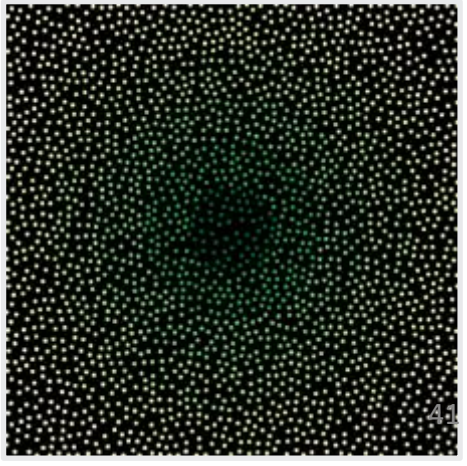




$\alpha = 6$



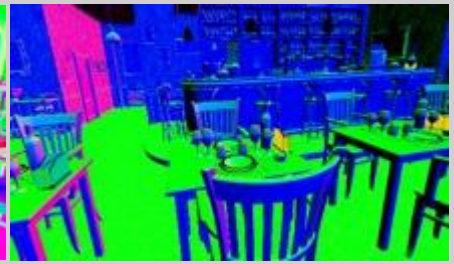




G-buffer

World position

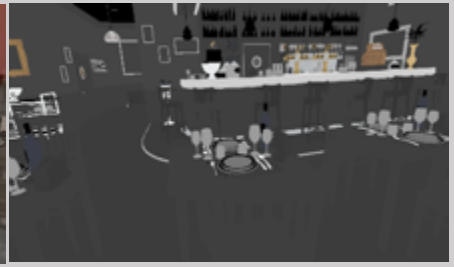
Bit tangent



Normal

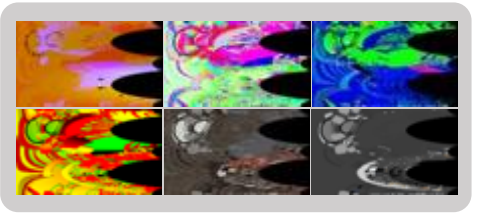
Texture coordinates

Albedo map



Roughness, ambient, and refraction maps

Kernel log-polar transformation



LP-buffer

($\sigma = 3.0$)

Shading & internal anti-aliasing

Inverse kernel log-polar transformation & post anti-aliasing

Screen

Sampling Rate Comparison between Different Rendering Approaches

# Samples	Original Light Field Rendering	KFR
Pass 1	-	$\frac{n^2}{\sigma_0^2}$
Pass 2	-	≈ 0
Total	n^2	$\frac{n^2}{\sigma_0^2}$

- Display resolution is $n \times n$

Content

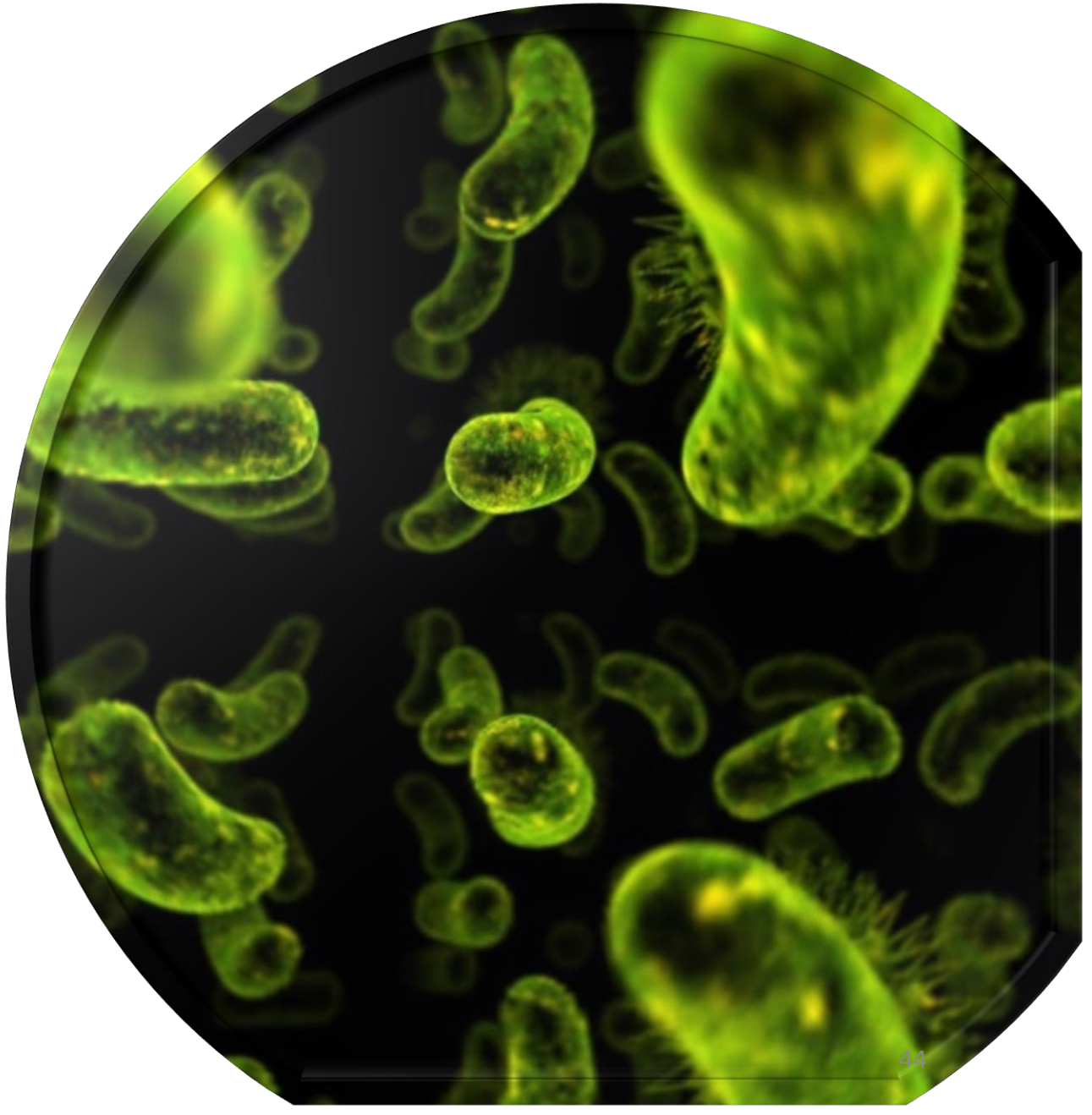
Motivation

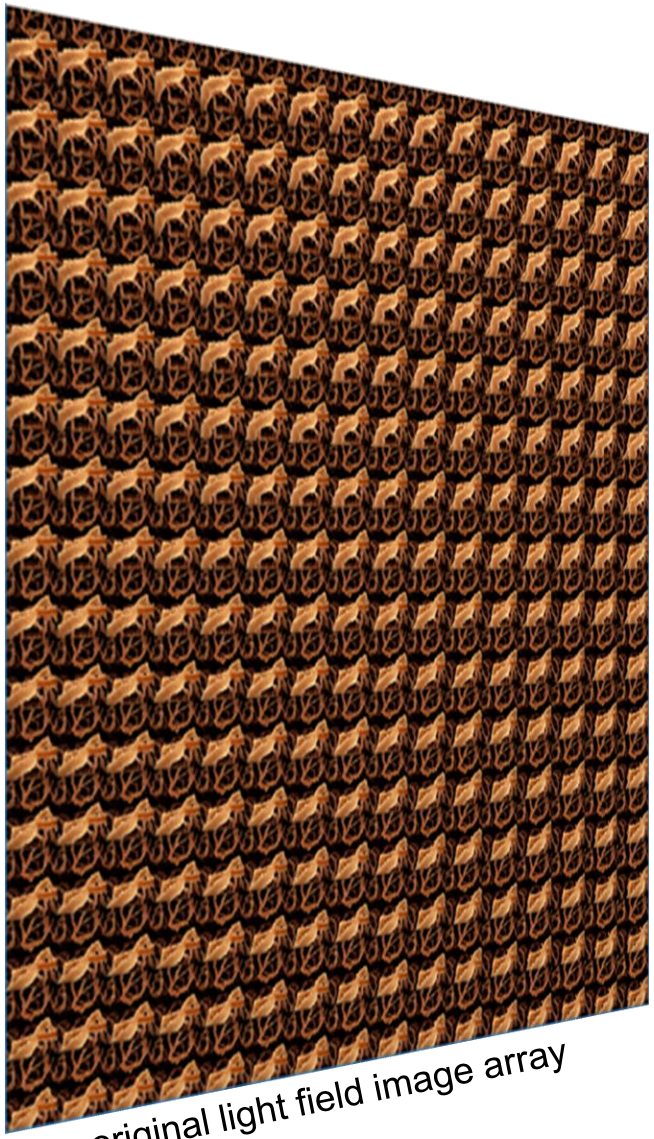
Foveated Rendering

Our Approach

User Study

Rendering Acceleration

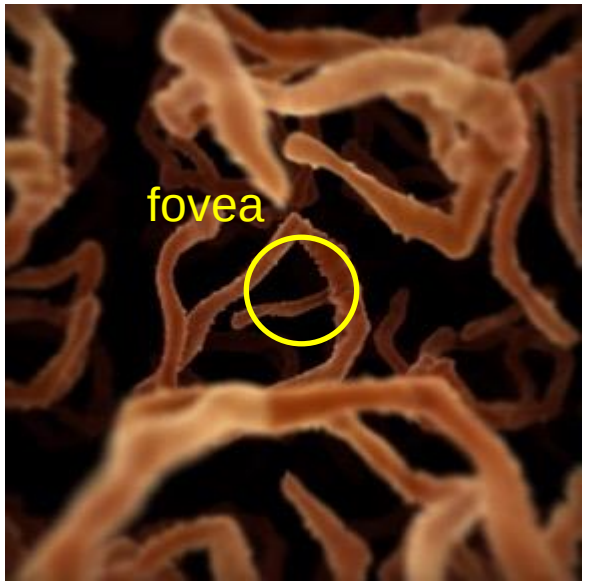


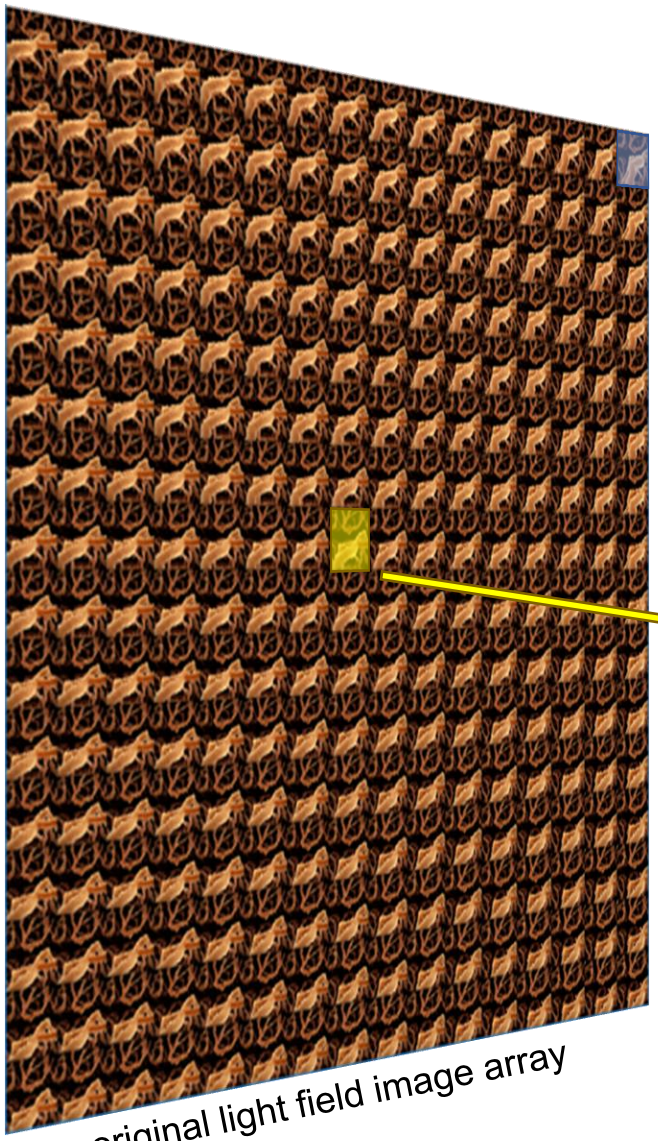


original light field image array

Better
Than
foveated rendering

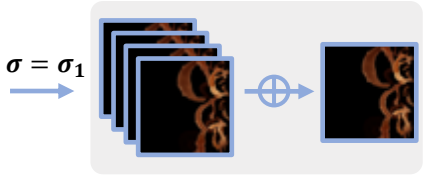
Screen



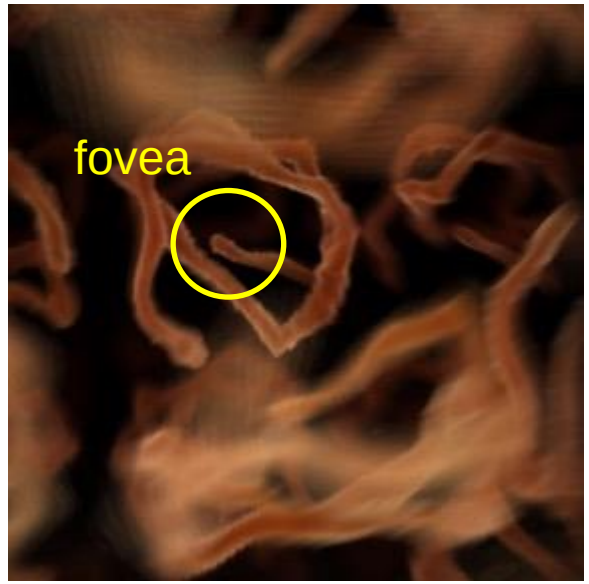


original light field image array

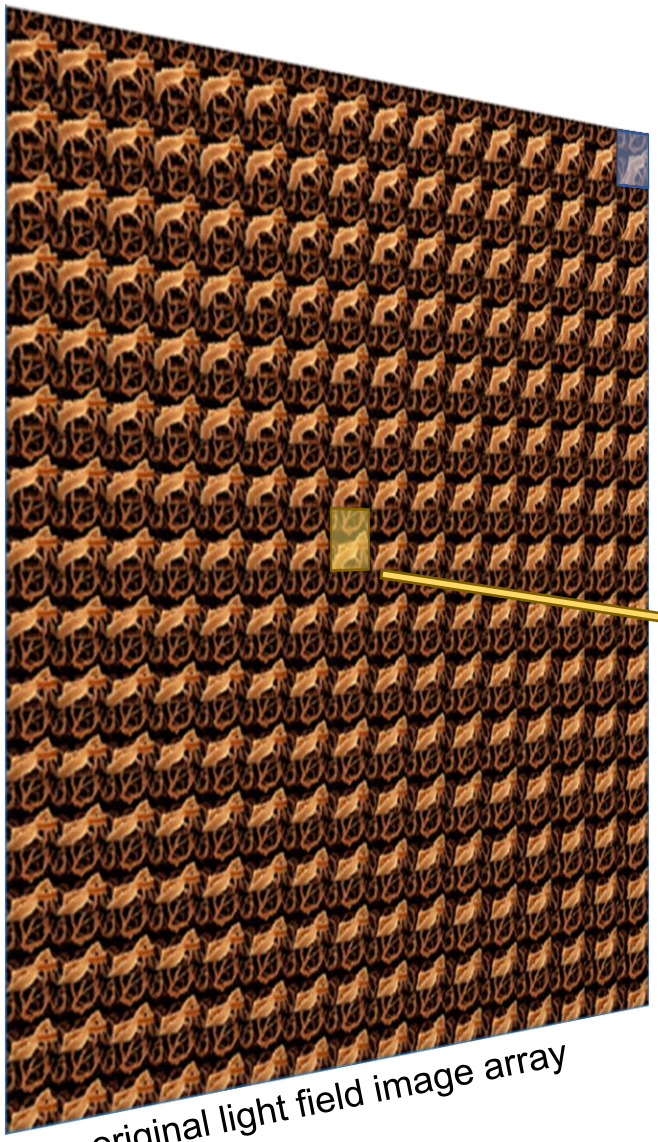
LP - buffers



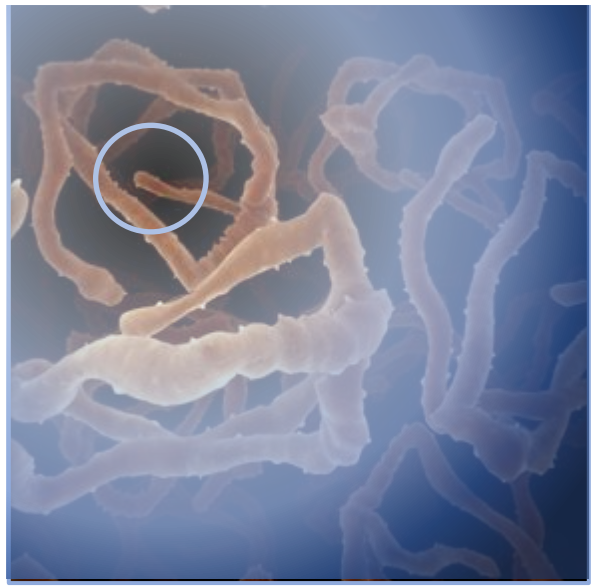
Screen

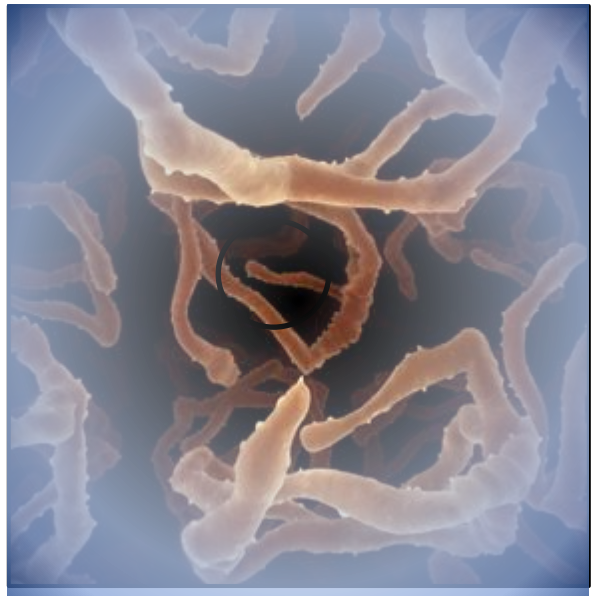


recovered visualization of the light field microscopy

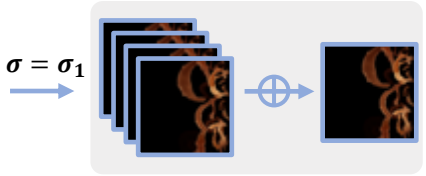


original light field image array

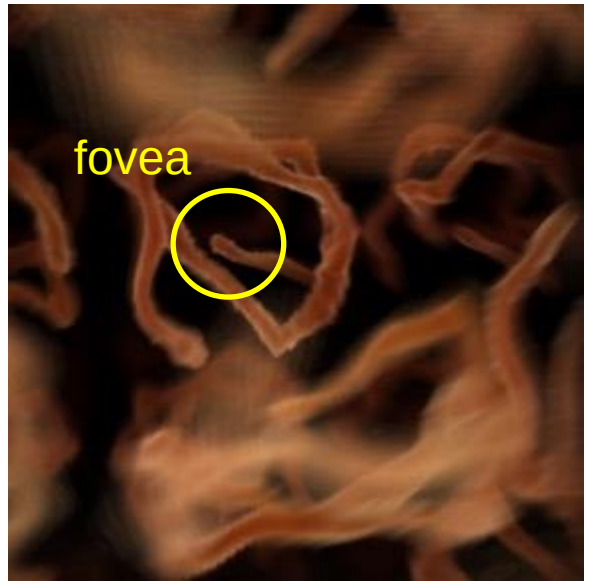




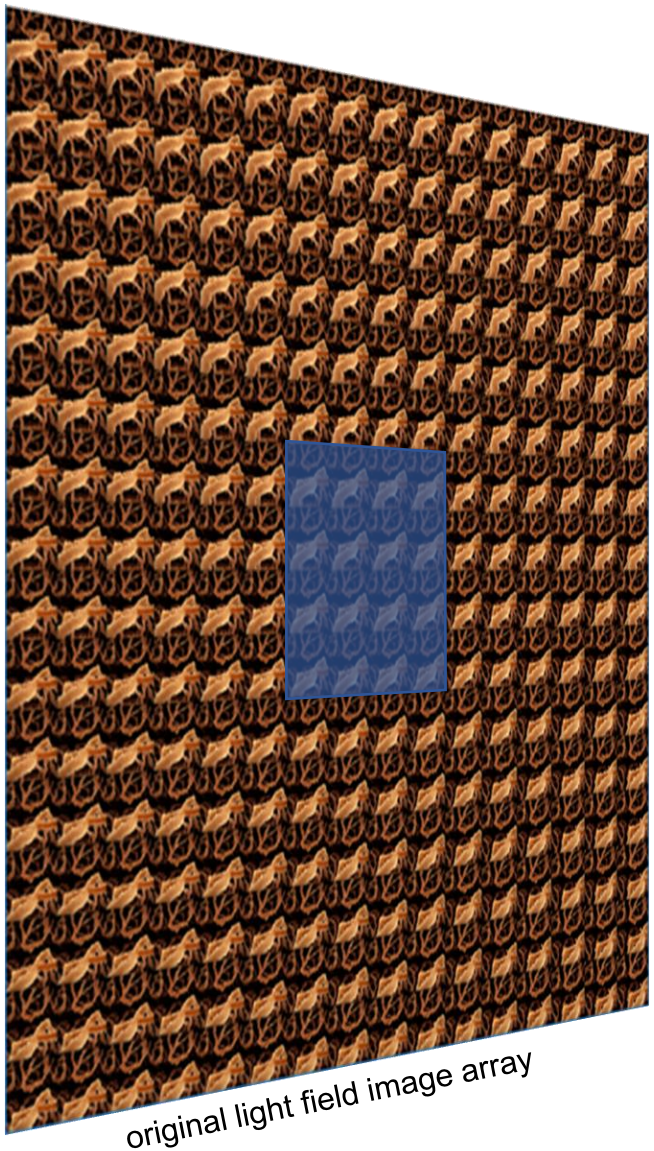
LP - buffers



Screen



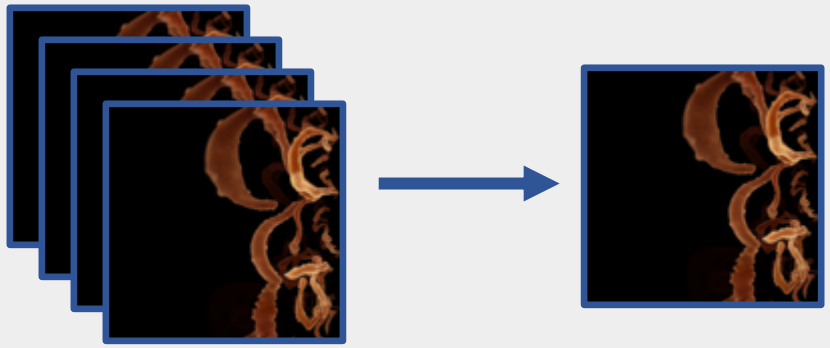
recovered visualization of the light field microscopy



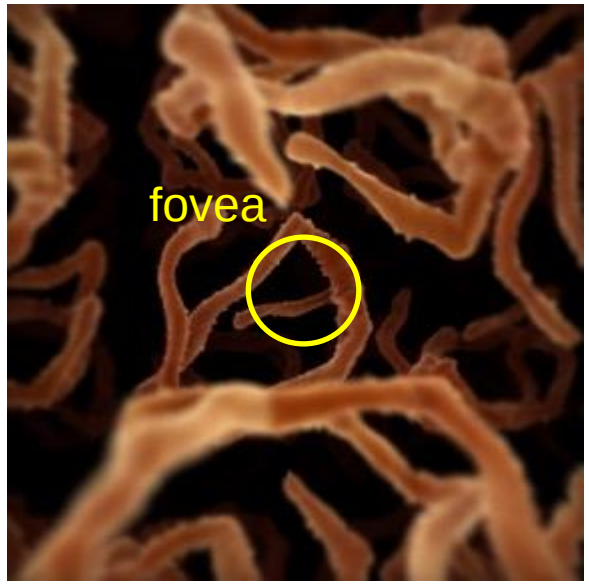
original light field image array

$\sigma = \sigma_0$

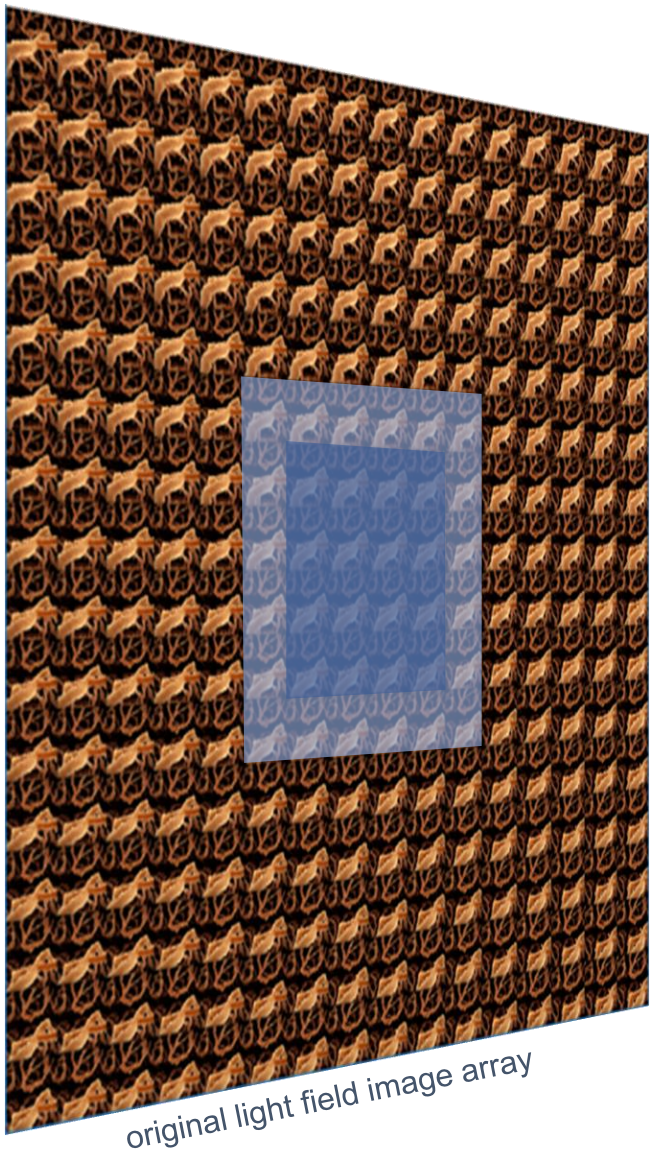
3D LP - buffers



Screen



recovered visualization of
the light field microscopy



original light field image array

$\sigma = \sigma_0$

$\sigma = \sigma_1$

3D LP - buffers

Screen

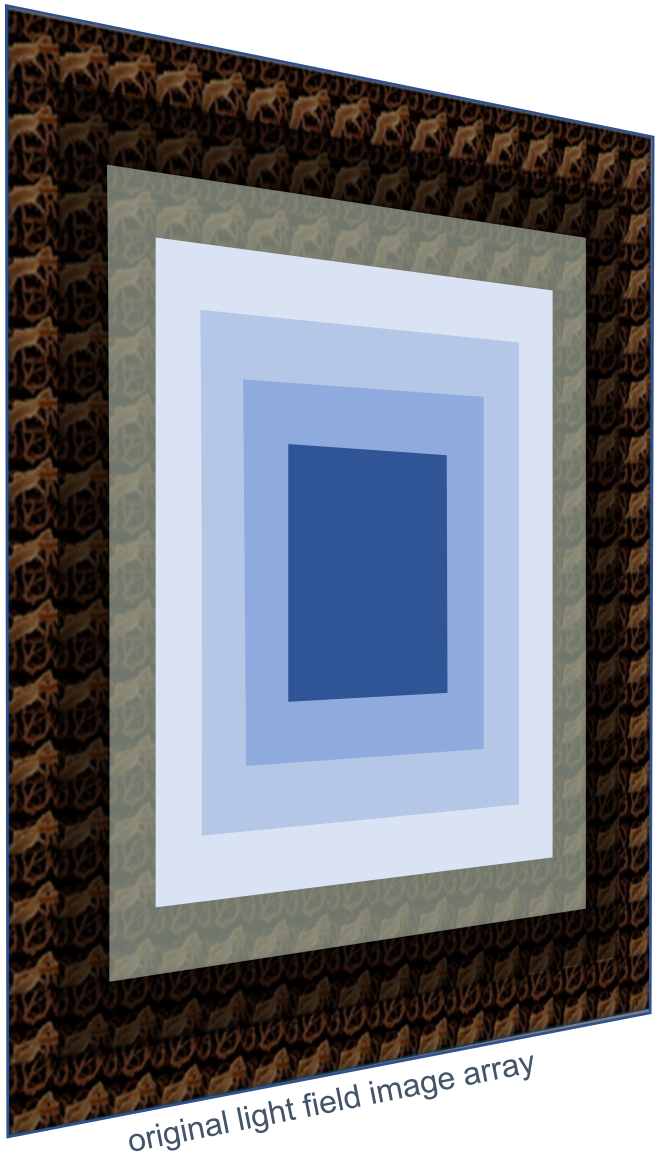
recovered visualization of
the light field microscopy

3D kernel log-polar transformation with different σ for different frames

sum up the frames with the same σ into one texture

inverse log-polar transformation and post-anti-aliasing

3D LP - buffers

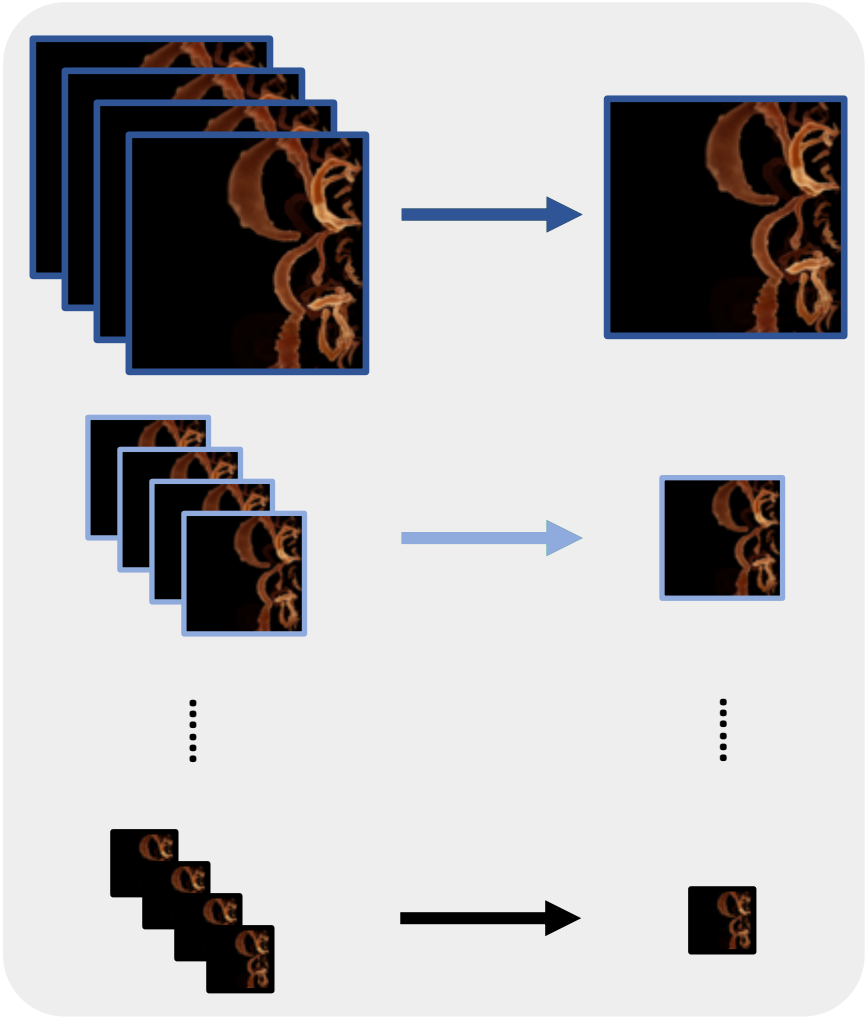


original light field image array

$\sigma = \sigma_0$

$\sigma = \sigma_1$

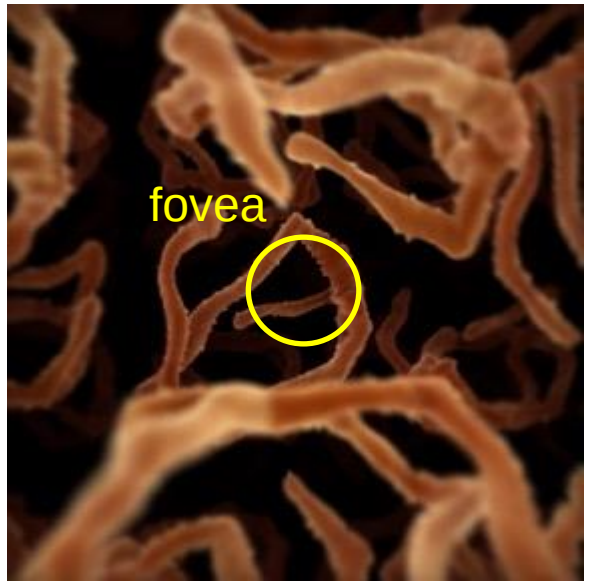
$\sigma = \sigma_1$



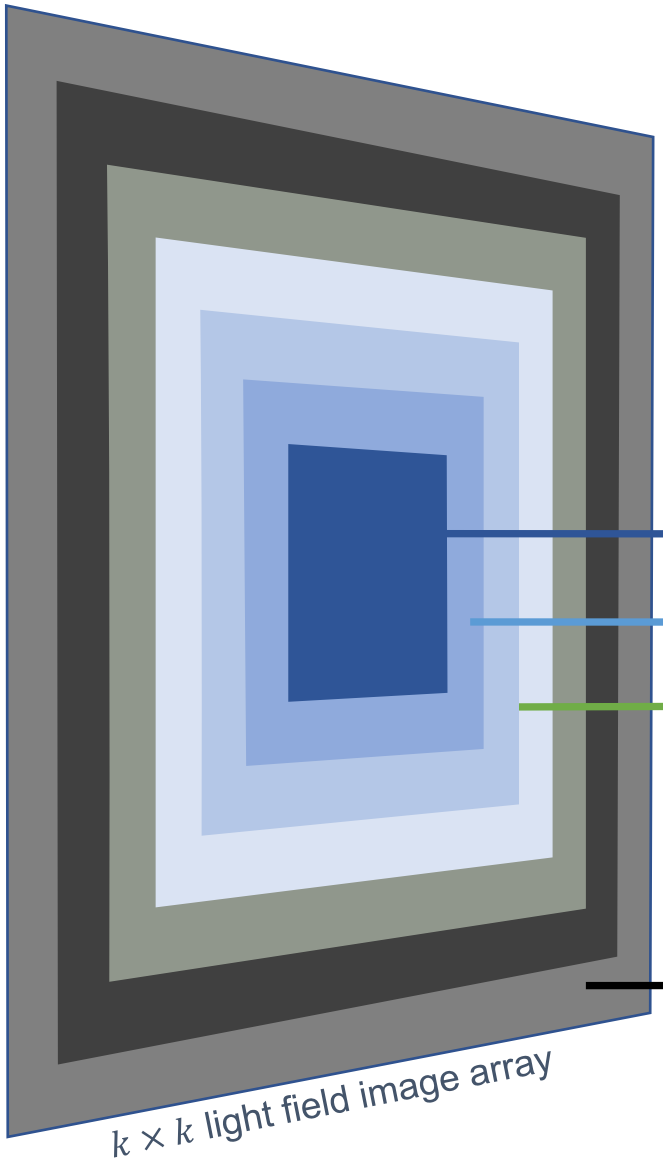
transformed image sub-arrays in kernel log-polar coordinates

transformed average image in kernel log-polar coordinates

Screen



recovered visualization of the light field microscopy with speedup of 7.3x compared with original light field microscopy



foveal region	$d < r_0$	σ_0	k_0 frames
peripheral region 1	$r_0 \leq d < r_1$	σ_1	k_1 frames
peripheral region 2	$r_1 \leq d < r_2$	σ_2	k_2 frames
⋮	⋮	⋮	⋮
peripheral region N	$r_{N-1} \leq d < r_N$	σ_N	k_N frames

$$k \times k = k_0 + k_1 + k_2 + k_N$$

Sampling Rate Comparison between Different Rendering Approaches

# Samples	Original Light Field Rendering	KFR	3D-KFR
Pass 1	-	$\frac{k^2}{\sigma_0^2} \times n^2$	$(\frac{k_0}{\sigma_0^2} + \frac{k_1}{\sigma_1^2} + \cdots + \frac{k_N}{\sigma_N^2}) \times n^2$
Pass 2	-	n^2	$(1 + N) \times n^2$
Total	$k^2 \times n^2$	$\left(\frac{k^2}{\sigma_0^2} + 1\right) \times n^2 \approx \frac{k^2}{\sigma_0^2} \times n^2$	$(\frac{k_0}{\sigma_0^2} + \frac{k_1}{\sigma_1^2} + \cdots + \frac{k_N}{\sigma_N^2} + 1 + N) \times n^2$

- For a $k \times k$ light field with image resolution of $n \times n$
- $k \times k = k_0 + k_1 + \cdots + k_N$
- Display resolution is $n \times n$

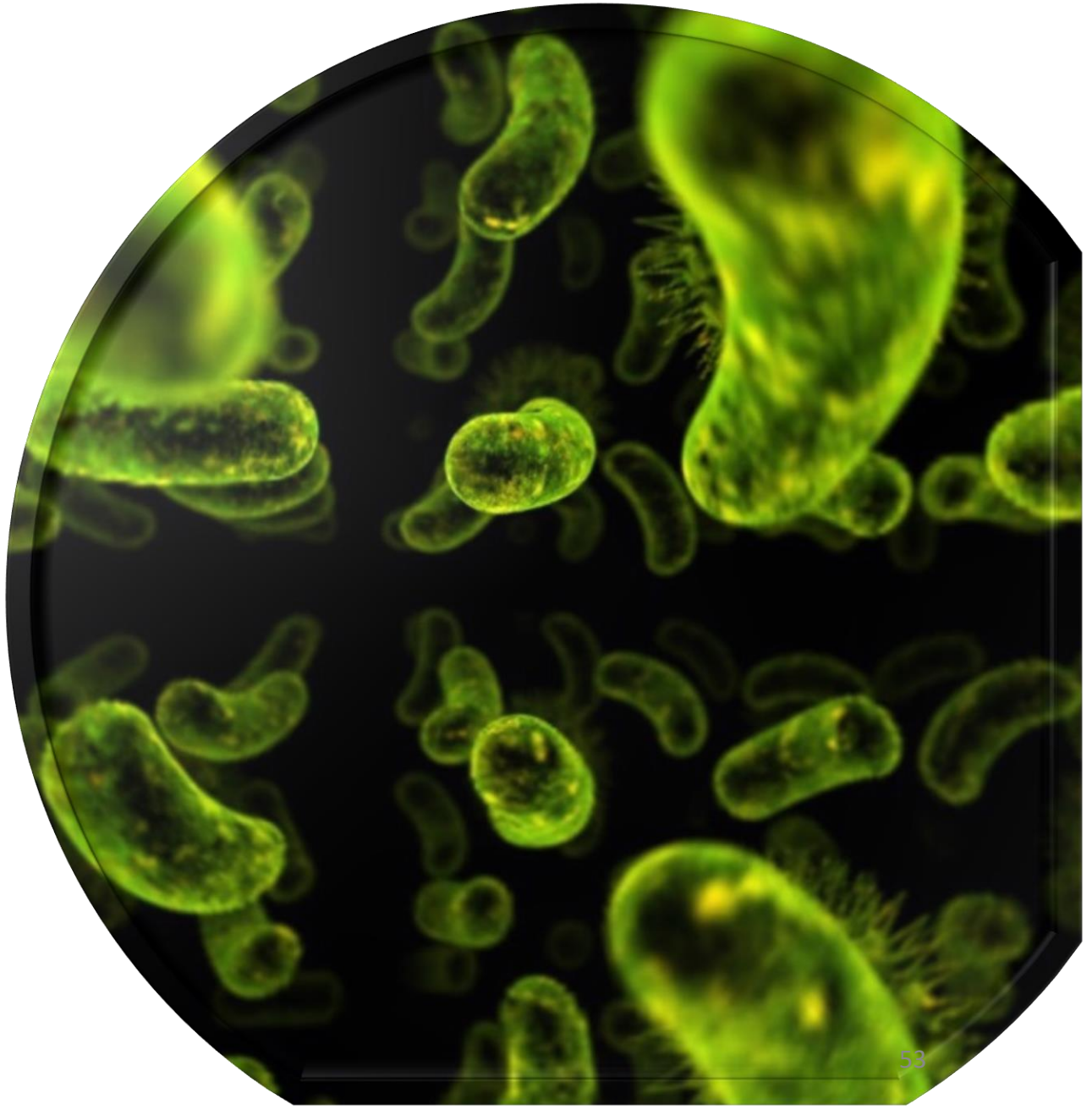
Content

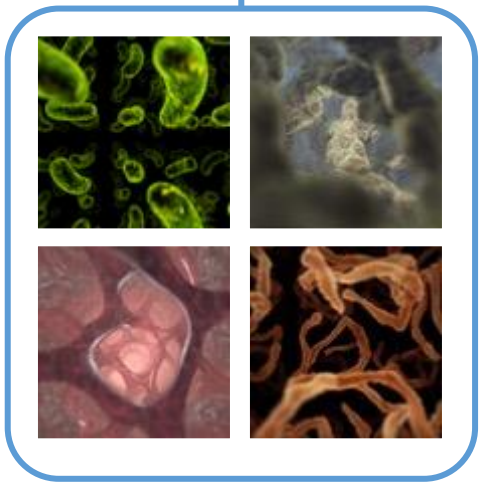
Motivation

Our Approach

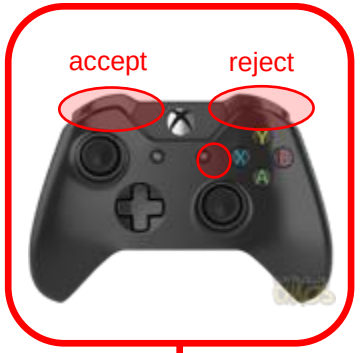
User Study

Rendering Acceleration





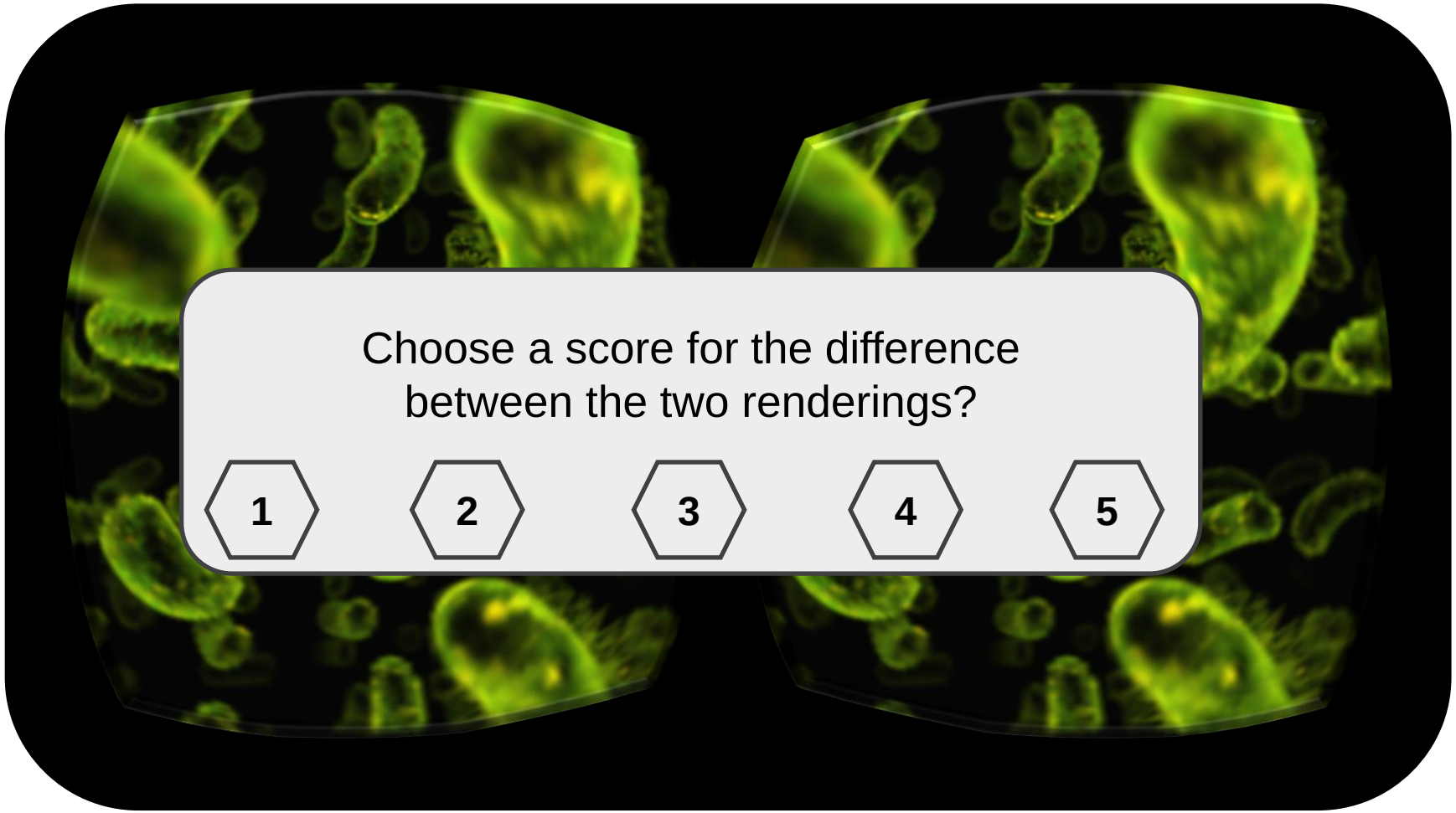
Resolution: 2560×1440
Field of view: up to 100 degrees



$\sigma \in [1.2, 3.6]$ step size: 0.2

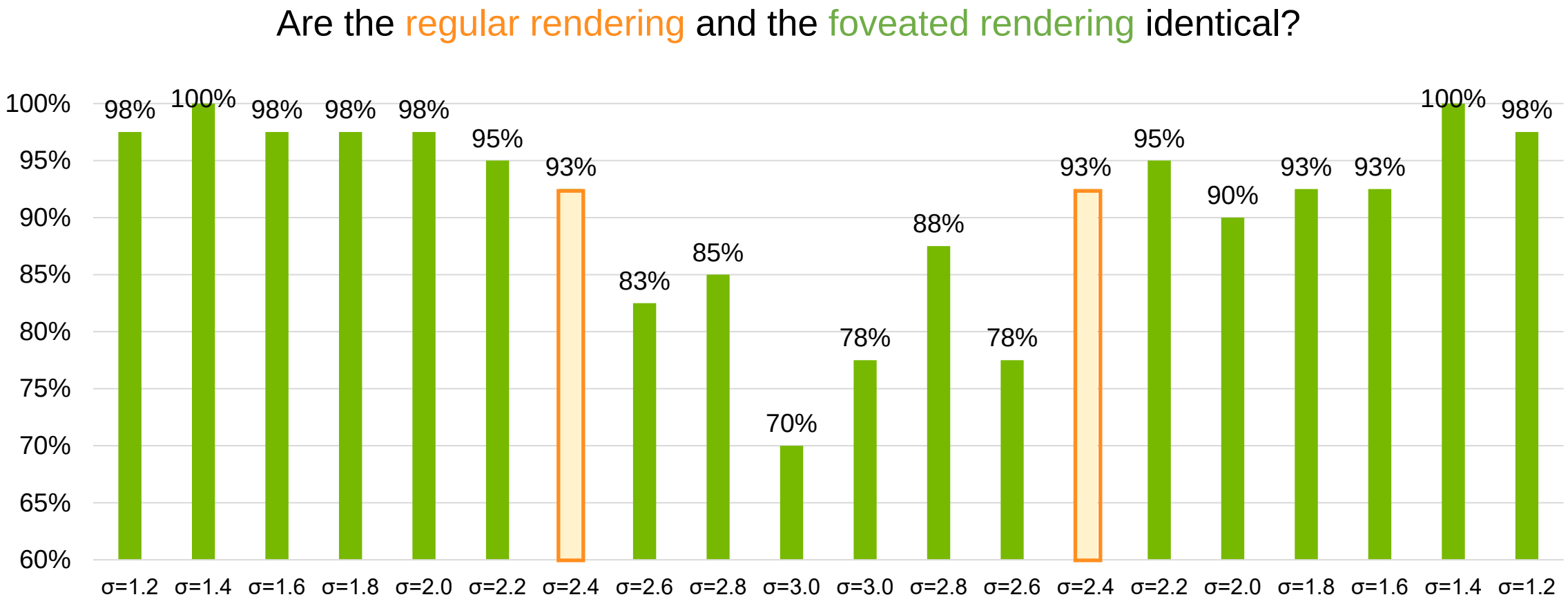
Pair Test

1. Ground Truth
2. Foveated Rendering
3. Evaluation



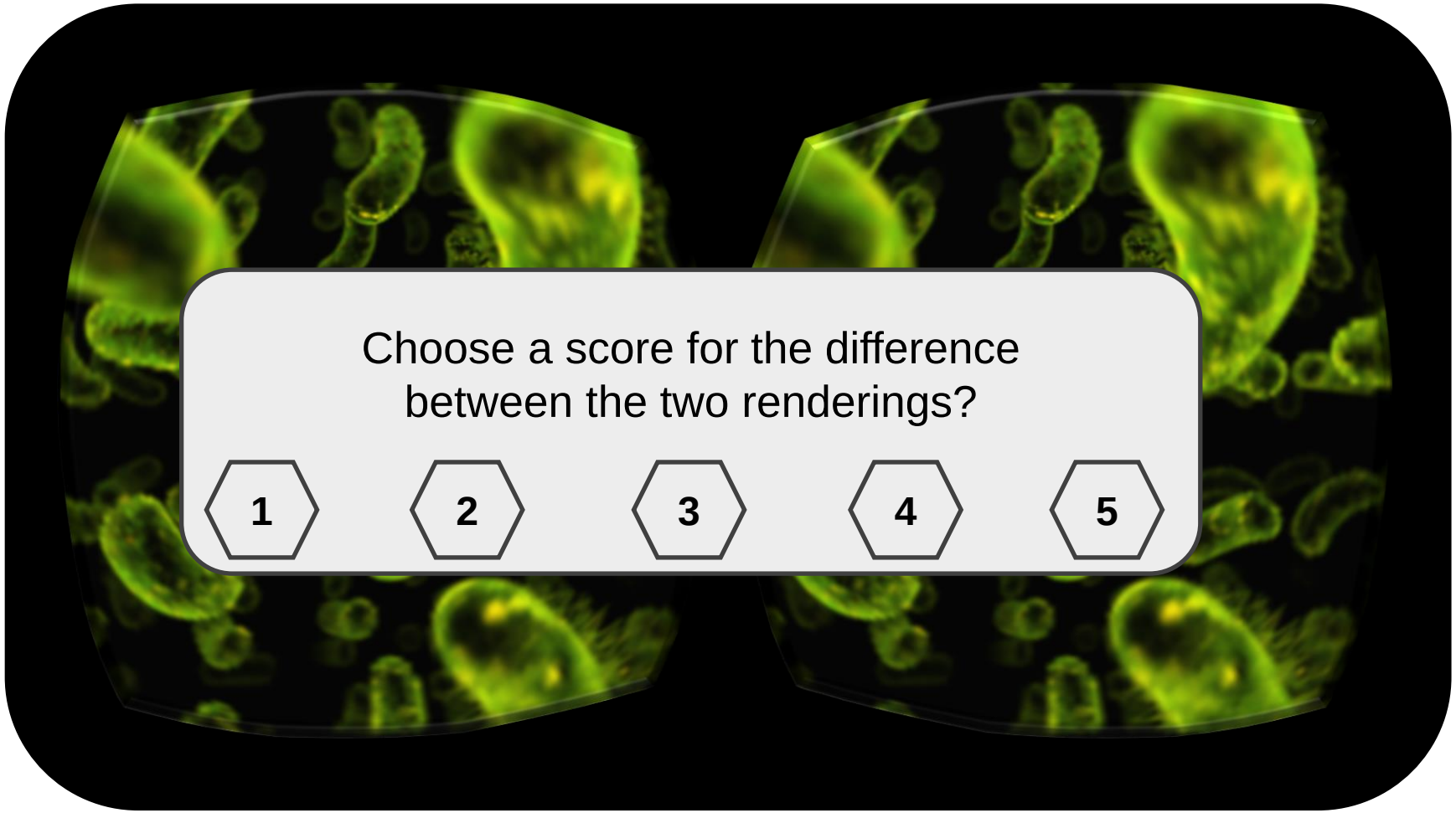
σ	1.2	1.4	…	2.8	**3.0**	2.8	…	1.4	1.2

Result – Pair Test



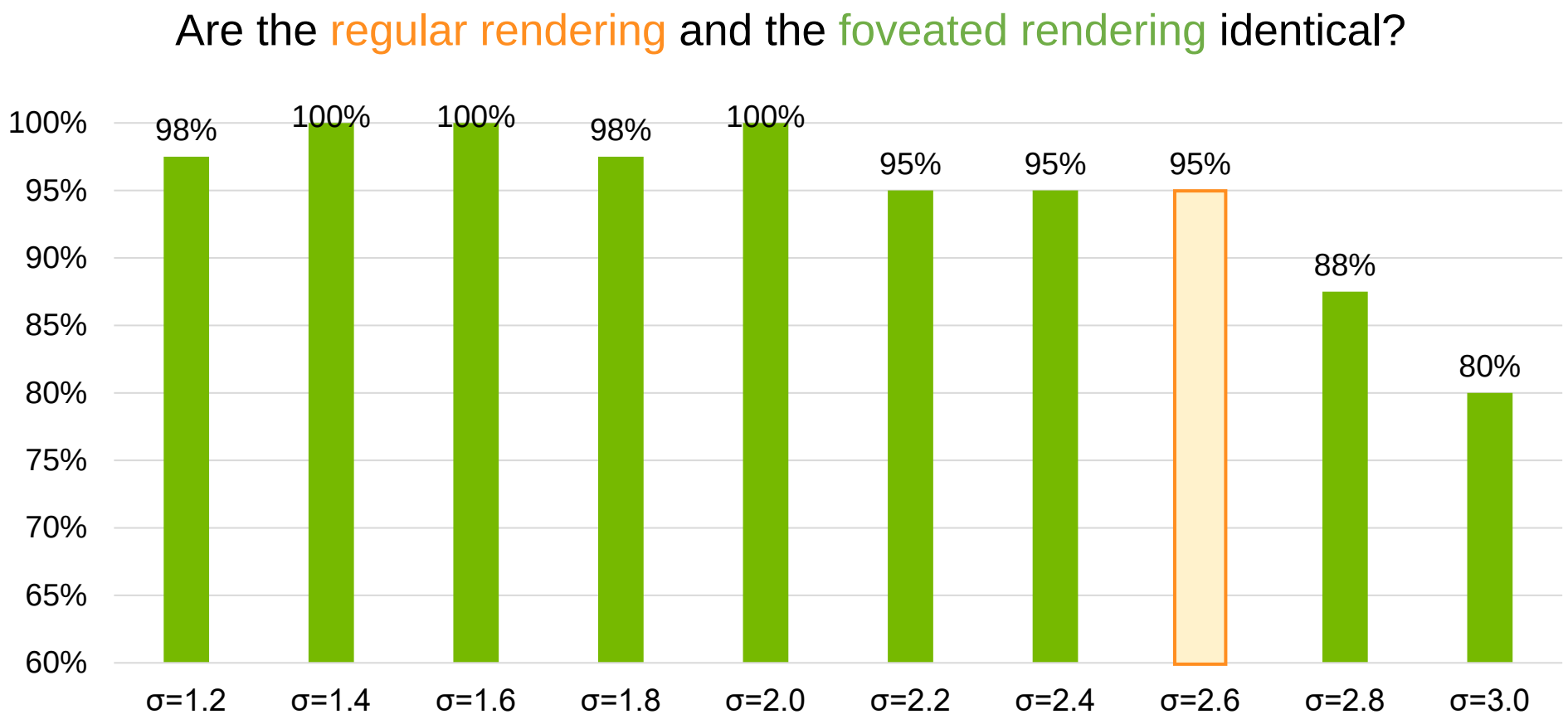
Random Test

1. Ground Truth
2. Foveated Rendering
3. Evaluation



Shuffled σ	2.6	1.4	…	2.0	**3.0**	1.8	…	2.2	1.6

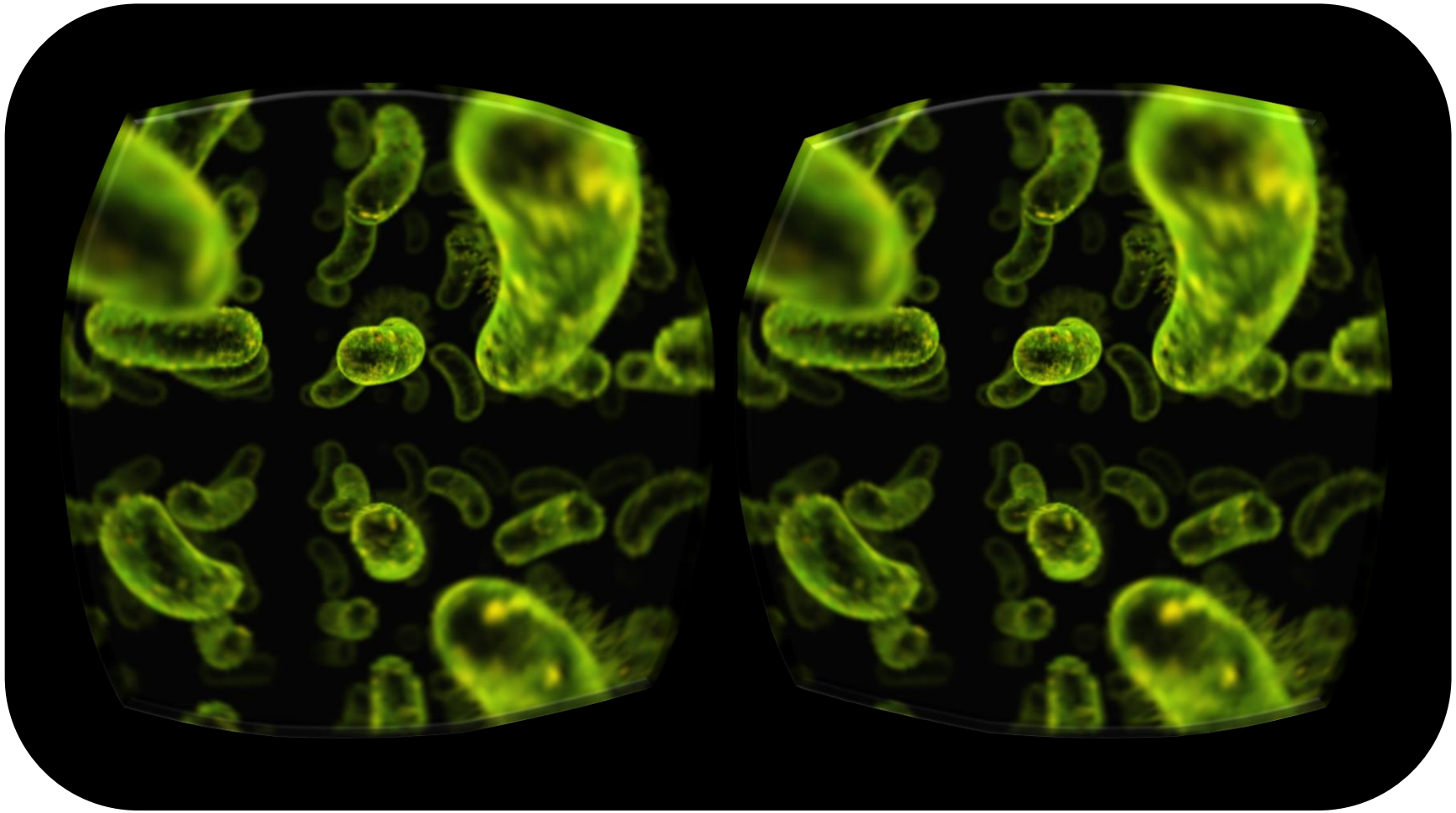
Result – Random Test



Slider Test

1 Ground Truth

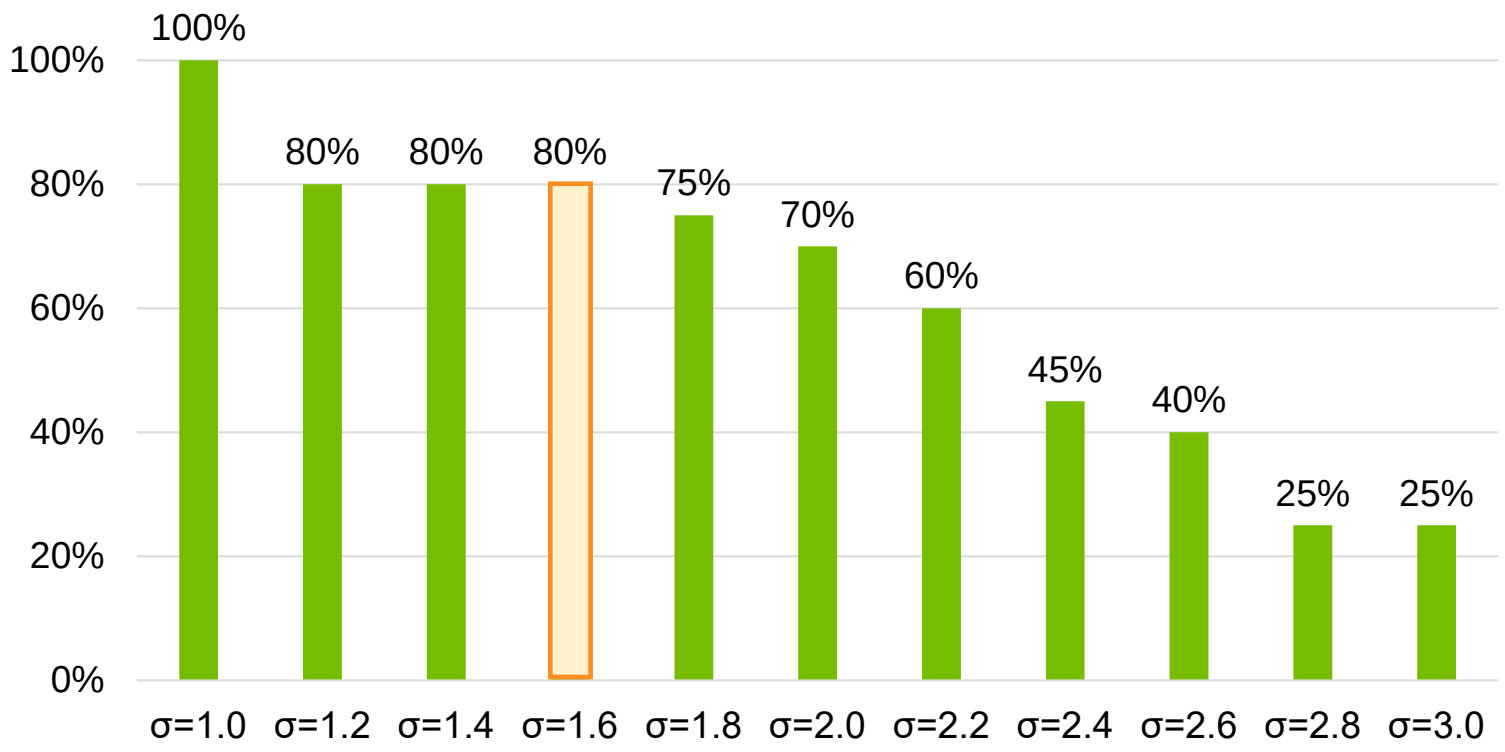
2 Foveated Rendering



1.2 — σ — 3.0

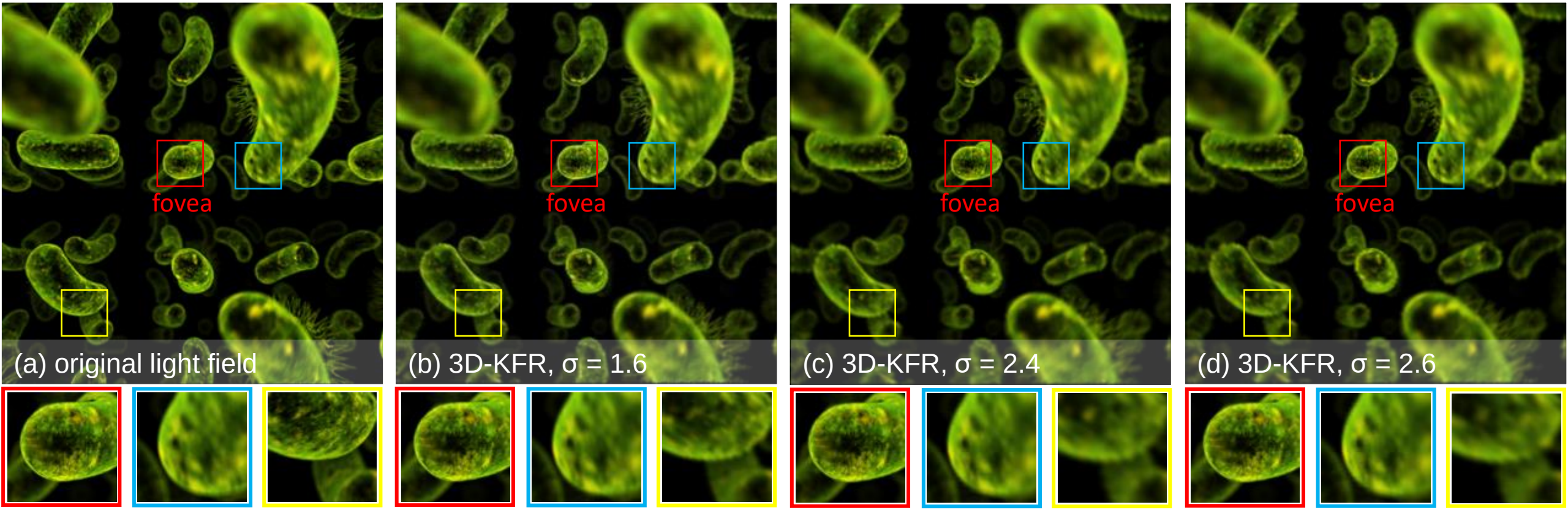
Result – Slider Test

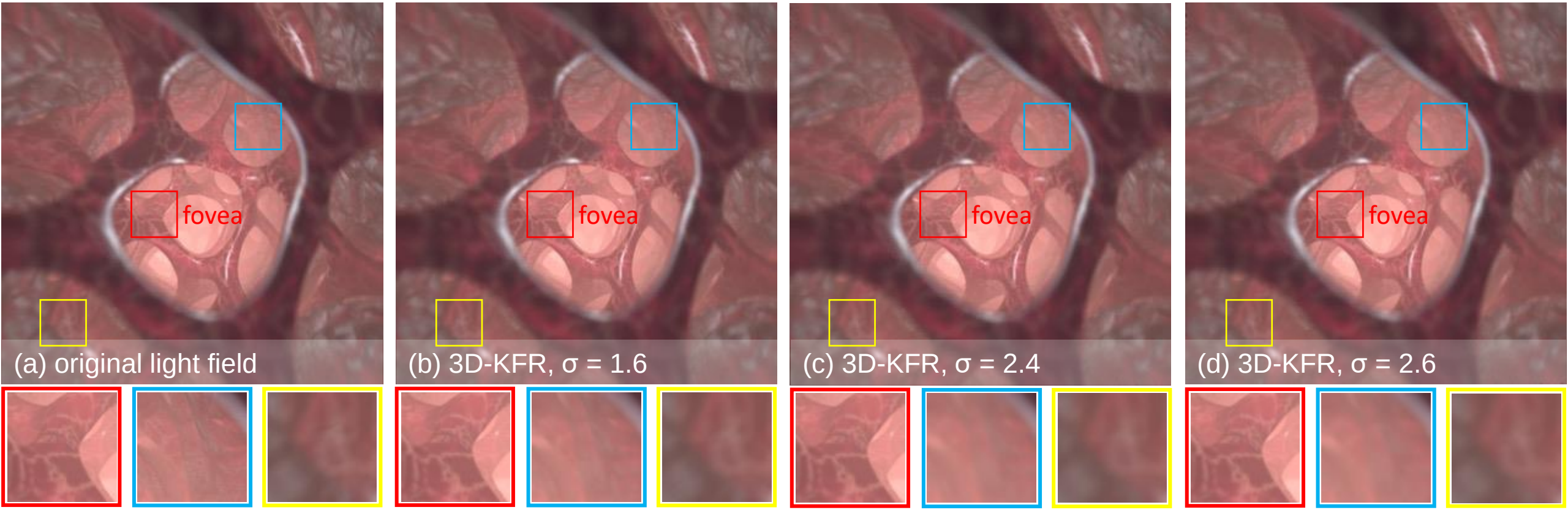
Are the regular rendering and the foveated rendering identical?

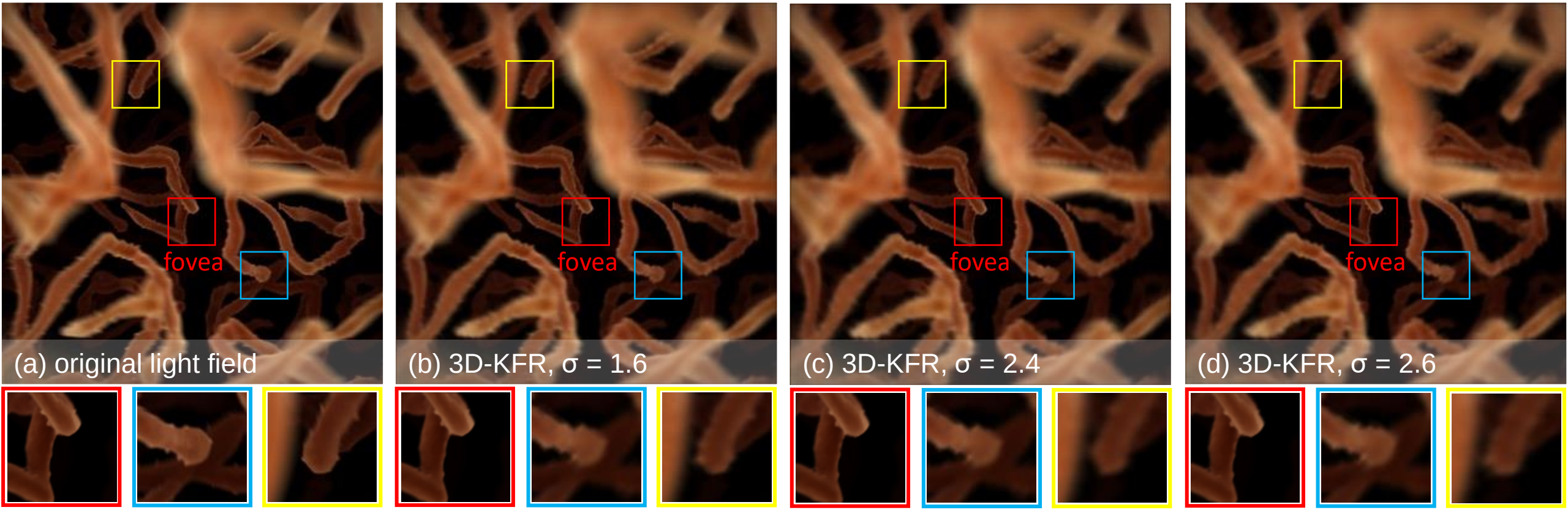


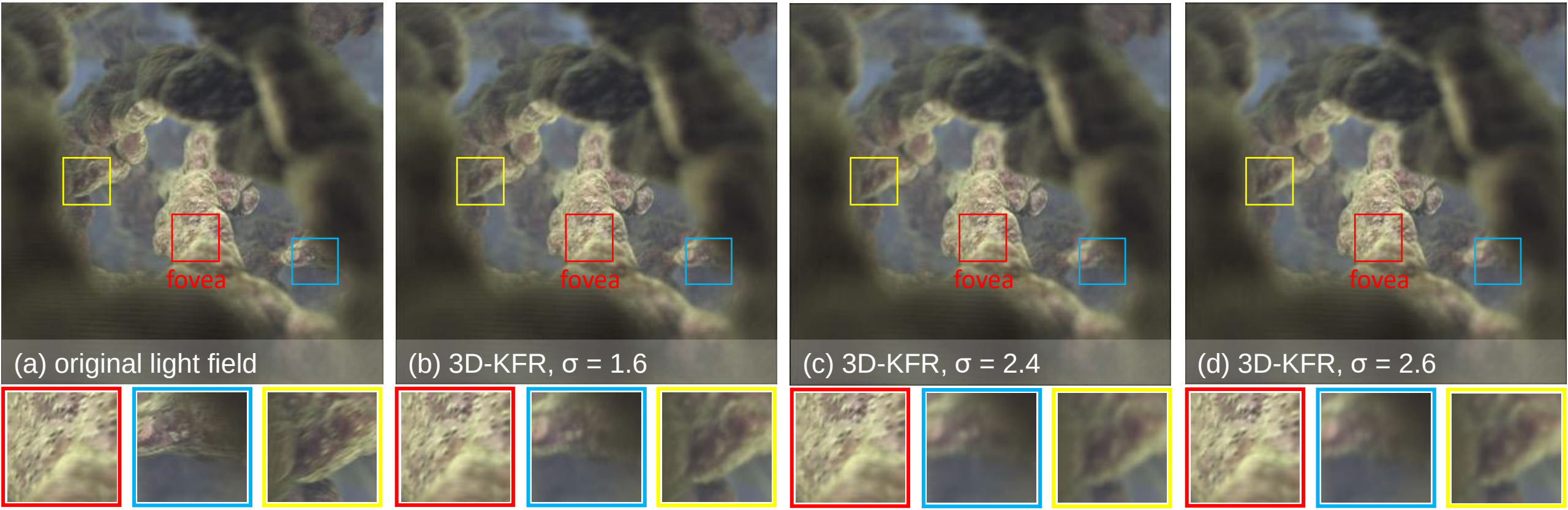
Rendering Acceleration

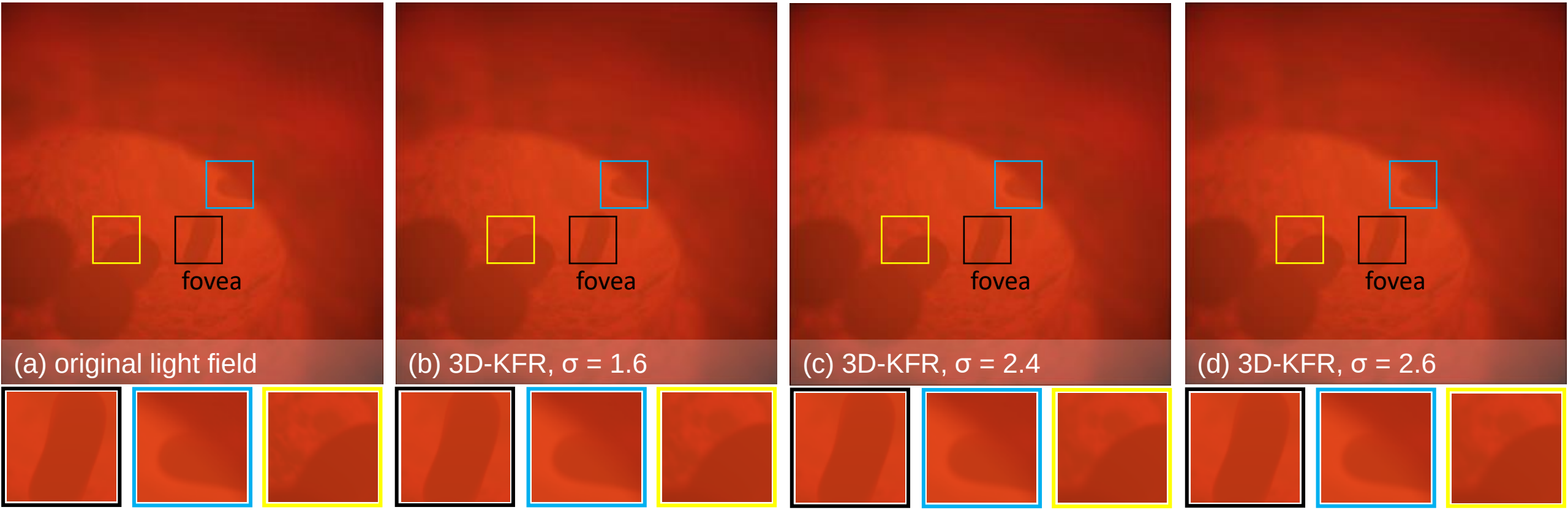
Resolution	Rendering Time of Ground Truth	$\sigma = 1.6$		$\sigma = 2.4$		$\sigma = 2.6$	
		Rendering Time of 3D KFR	Speedup	Rendering Time of 3D KFR	Speedup	Rendering Time of 3D KFR	Speedup
$20 \times 20 \times 1024 \times 1024$	$66.83\ ms$	$19.27\ ms$	$\mathbf{3.47}\times$	$10.22\ ms$	$\mathbf{6.54}\times$	$9.39\ ms$	$\mathbf{7.11}\times$
$21 \times 21 \times 1024 \times 1024$	$74.17\ ms$	$22.39\ ms$	$\mathbf{3.31}\times$	$11.90\ ms$	$\mathbf{6.24}\times$	$10.39\ ms$	$\mathbf{7.14}\times$
$22 \times 22 \times 1024 \times 1024$	$92.33\ ms$	$28.26\ ms$	$\mathbf{3.27}\times$	$14.65\ ms$	$\mathbf{6.30}\times$	$12.64\ ms$	$\mathbf{7.30}\times$
$23 \times 23 \times 1024 \times 1024$	$100.26\ ms$	$30.64\ ms$	$\mathbf{3.27}\times$	$16.30\ ms$	$\mathbf{6.15}\times$	$13.95\ ms$	$\mathbf{7.18}\times$
$24 \times 24 \times 1024 \times 1024$	$122.29\ ms$	$35.92\ ms$	$\mathbf{3.40}\times$	$19.09\ ms$	$\mathbf{6.41}\times$	$16.79\ ms$	$\mathbf{7.28}\times$
$25 \times 25 \times 1024 \times 1024$	$138.93\ ms$	$41.42\ ms$	$\mathbf{3.35}\times$	$21.96\ ms$	$\mathbf{6.33}\times$	$19.09\ ms$	$\mathbf{7.28}\times$

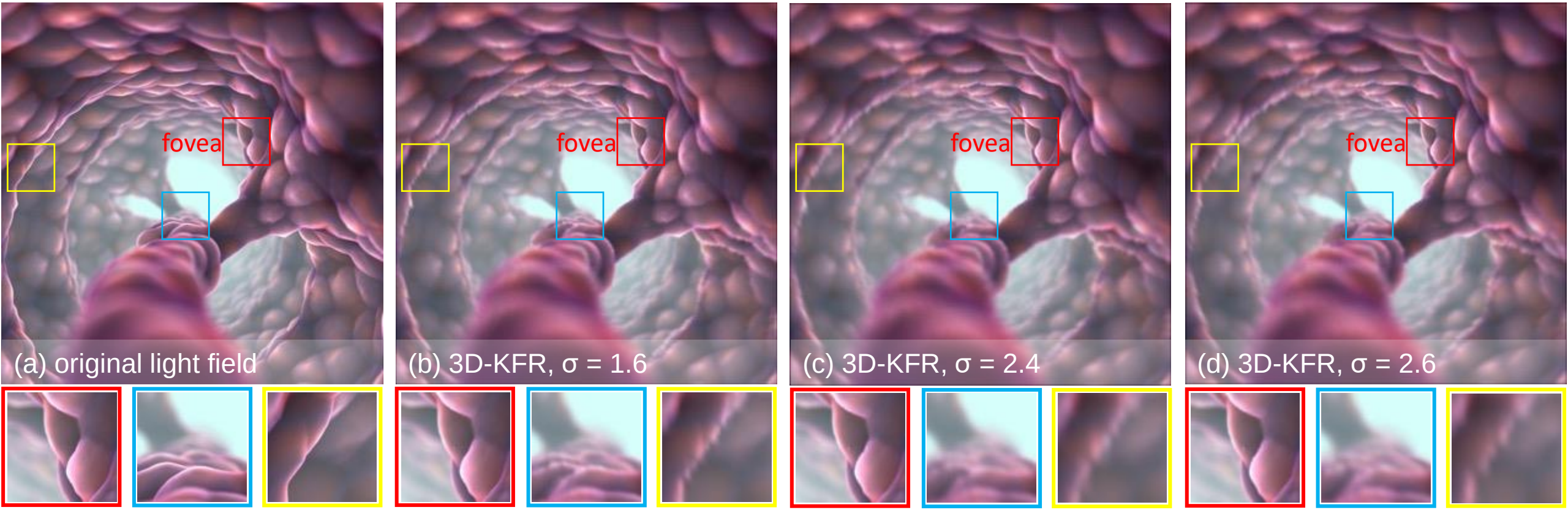

fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6


fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6


fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6


fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6


fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6


fovea
fovea
fovea
fovea
(a) original light field
(b) 3D-KFR, σ = 1.6
(c) 3D-KFR, σ = 2.4
(d) 3D-KFR, σ = 2.6

Summary

- 3D-Kernel log-polar transformation for light fields
  - 3.3X - 7.3X speedup
- User study
  - Determine parameters to maximize perceptual realism and minimize computation
- Light field dataset
  - Open Source: https://foveation.umiacs.umd.edu/3D_KFR