
Realistic Image Synthesis

- Perception-based Rendering
&
Advanced Displays-

Philipp Slusallek
Karol Myszkowski
Gurprit Singh
Corentin Salaun

Outline

- **Perceptually based adaptive sampling algorithm**
- **Eye tracking driven rendering**
- **Binocular 3D displays**
- **Autostereoscopic (Glass-free 3D) Displays**
 - Parallax Barriers
 - Integral Imaging
 - Multi-layer displays
 - Holographic displays
- **Head-Mounted Displays with accommodation cues**

A Perceptually Based Physical Error Metric for Realistic Image Synthesis

by Mahesh Ramasubramanian, Sumanta N. Pattanaik,
and Donald P. Greenberg
Siggraph 1999

Aims for perceptual accuracy

- Limitations of the human visual system...
perceptual accuracy < physical accuracy.
- Perceptual accuracy guides rendering, not physical accuracy.

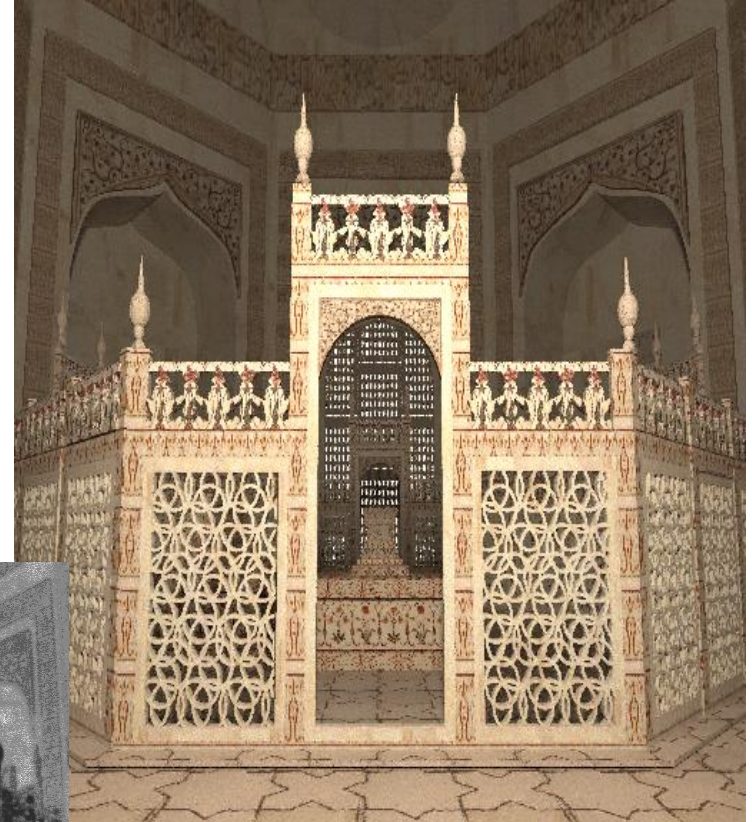
Perceptually Based Rendering

6% effort

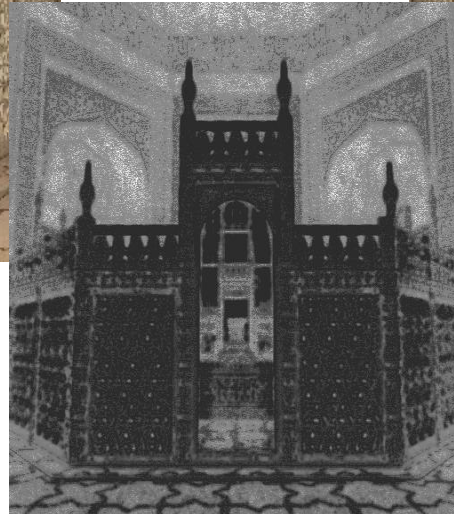


physically
accurate

effort
distribution
(darker
regions -
less effort)



perceptually
accurate



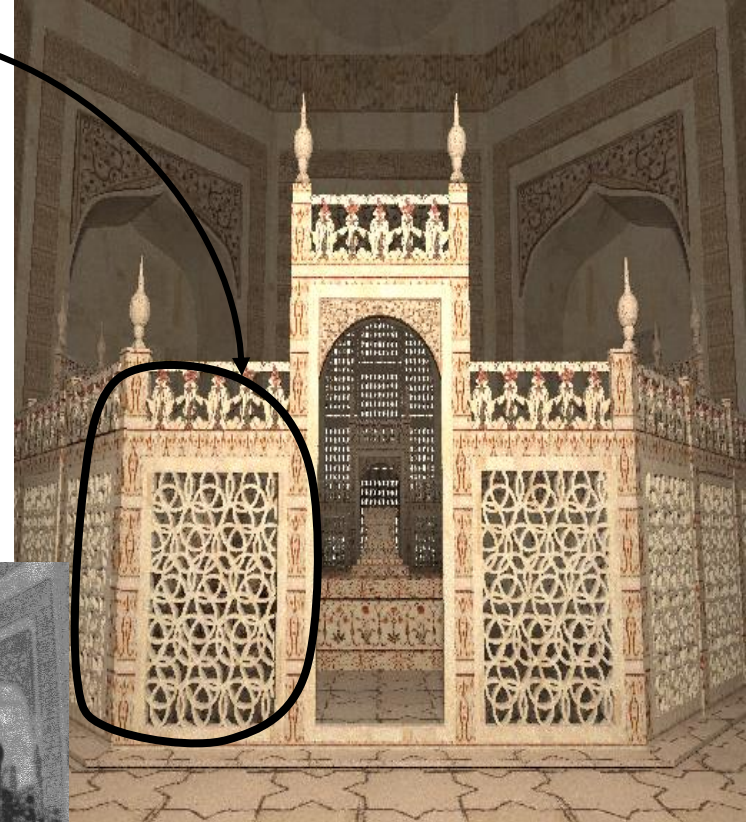
Perceptually Based Rendering

6% effort

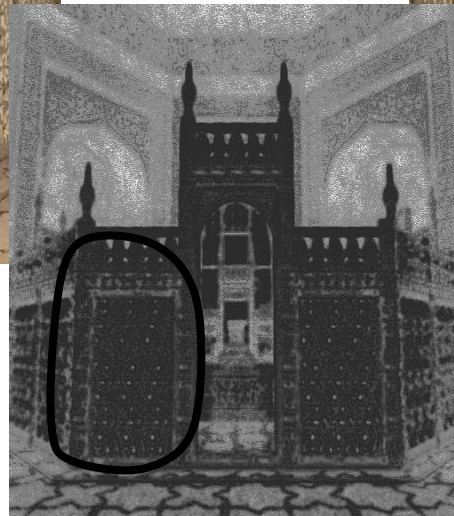


physically
accurate

effort
distribution
(darker
regions -
less effort)



perceptually
accurate



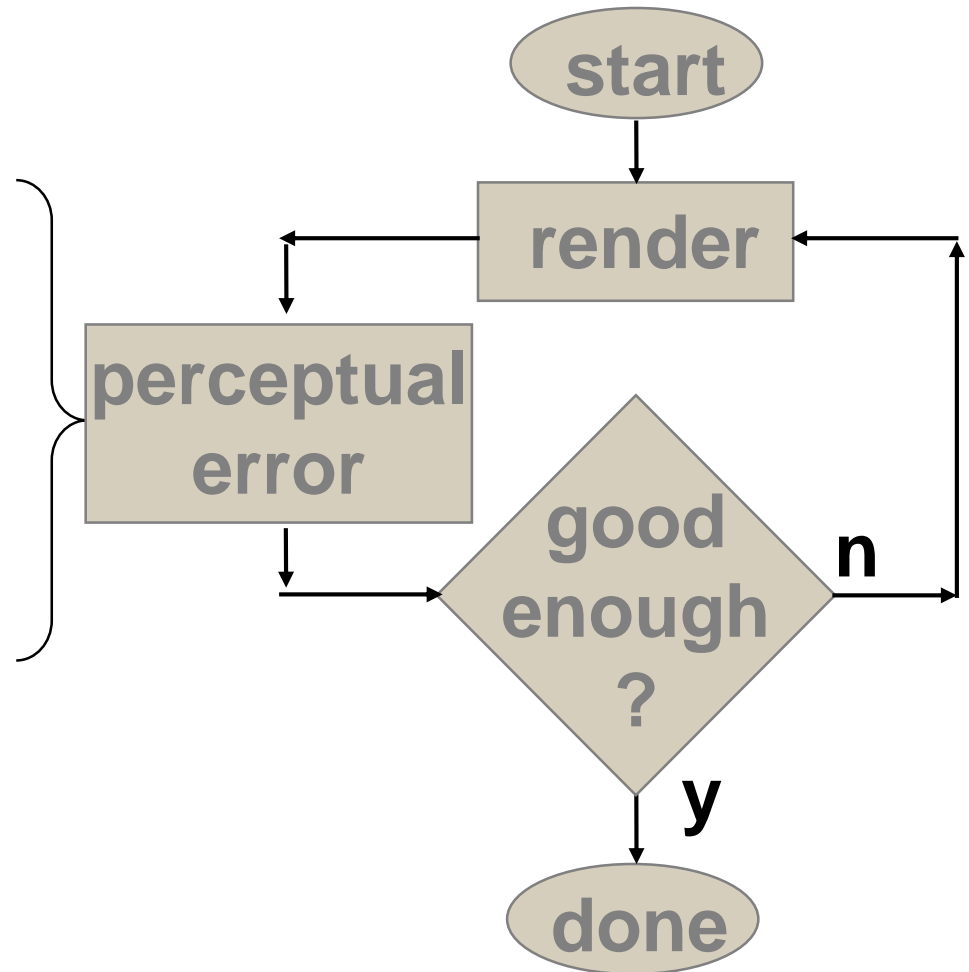
Perceptually Based Rendering

Traditional approach:

Pair of images to compare at each time step

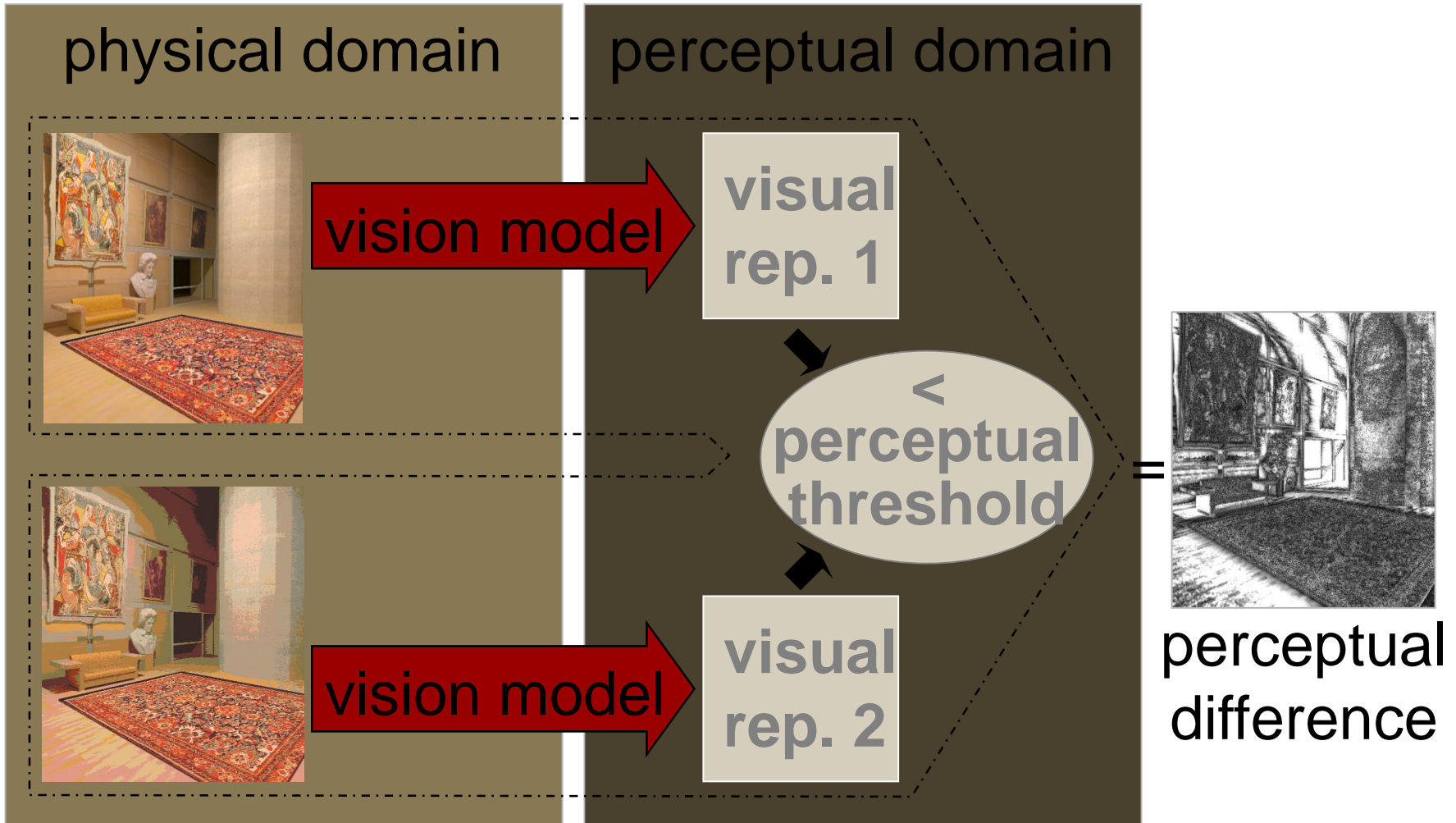
(a) intermediate images at consecutive time steps.

(b) upper and lower bound images at each time step.

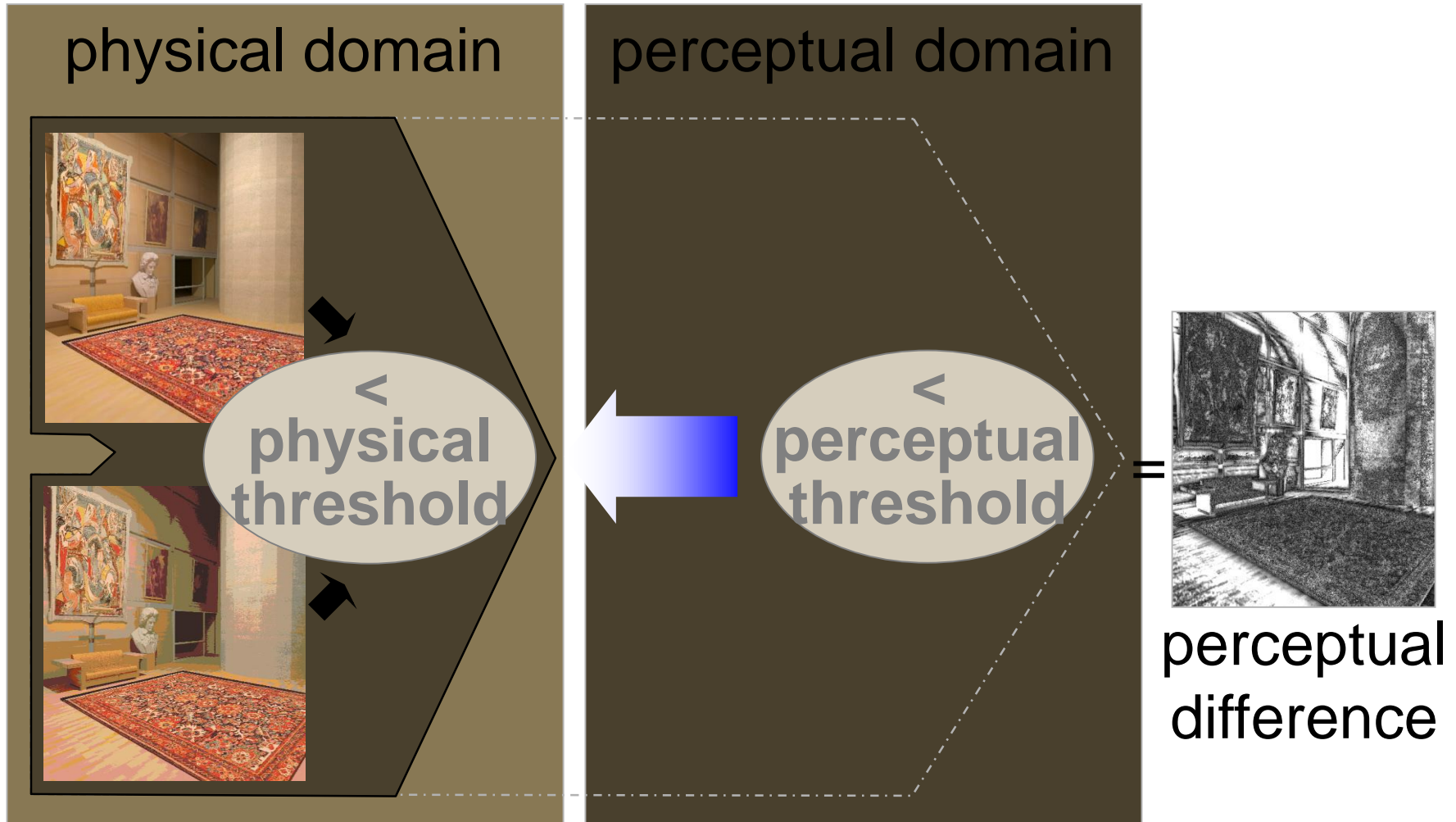


Perceptual Error Metric

Vision model - expensive



Perceptually Based Physical Error Metric



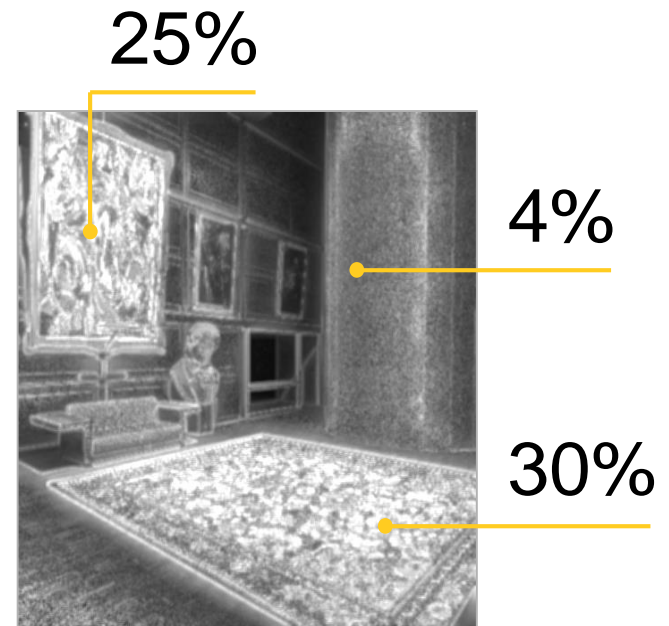
Physical Threshold Map

Predicted bounds of permissible luminance error



input image

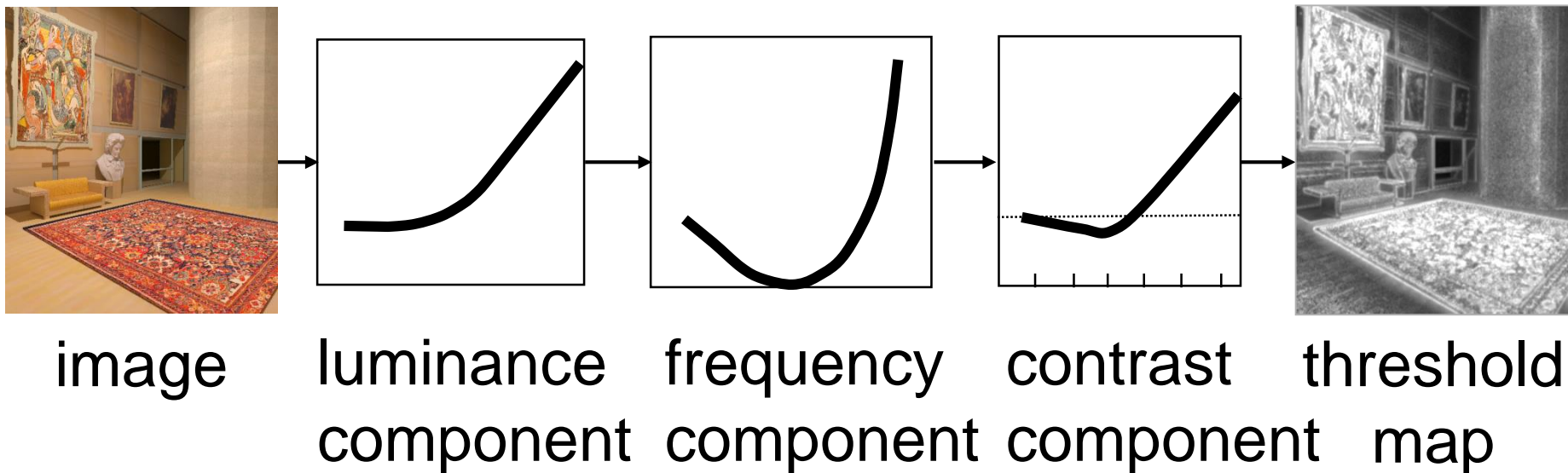
threshold
model



physical threshold
(brighter regions
- higher thresholds)

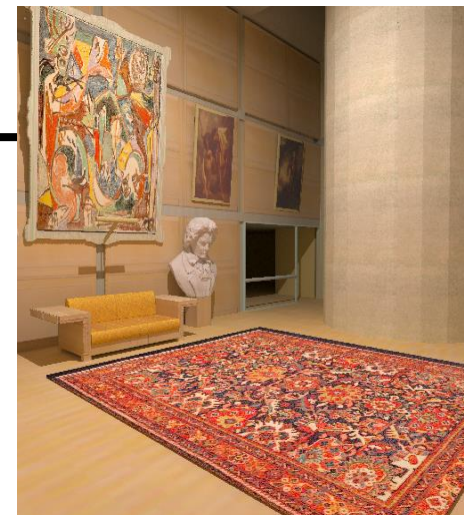
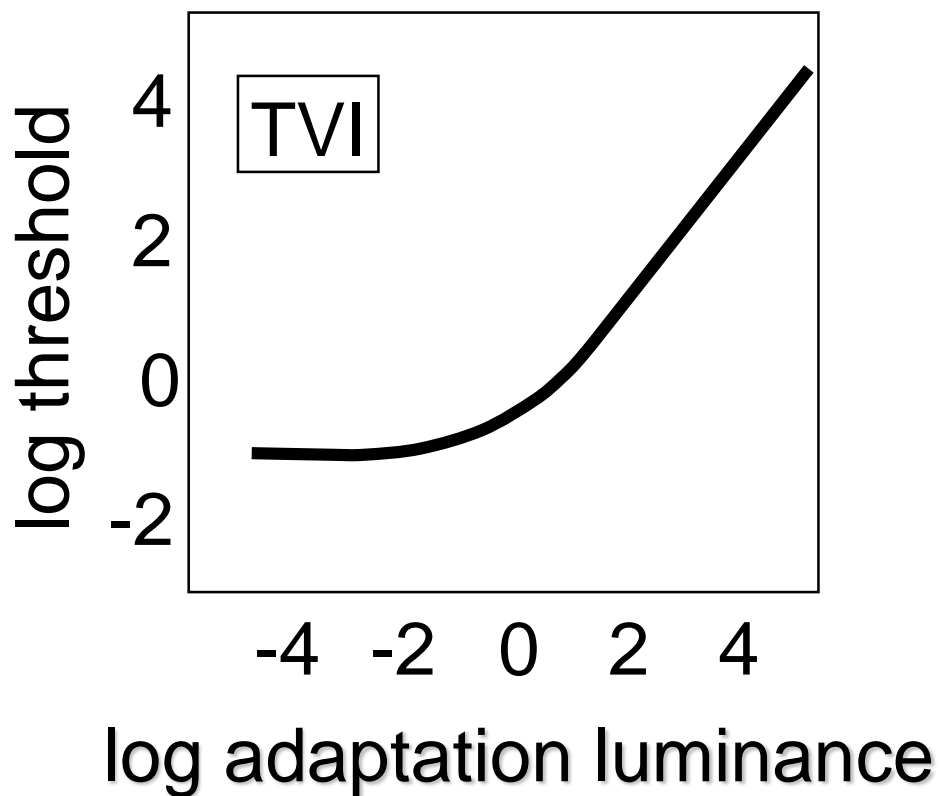
Threshold Model

Components



Threshold Model

1. Luminance component

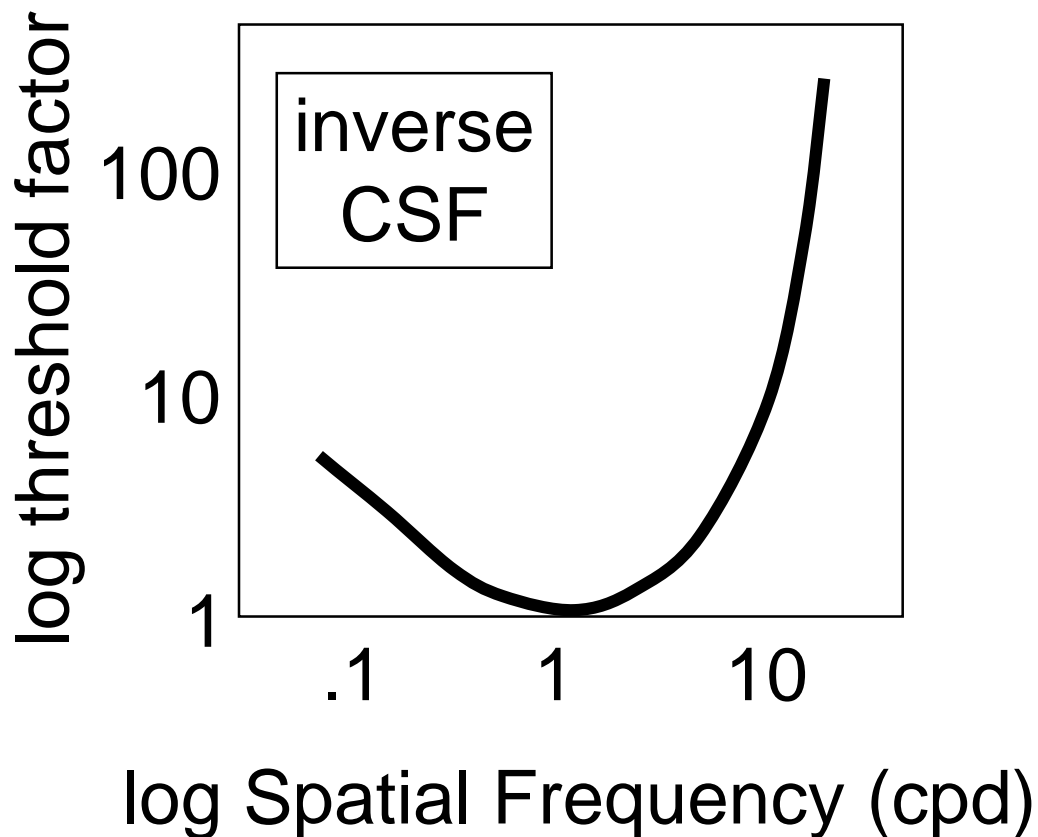


2%

threshold due to
luminance

Threshold Model

2. Frequency component



2%



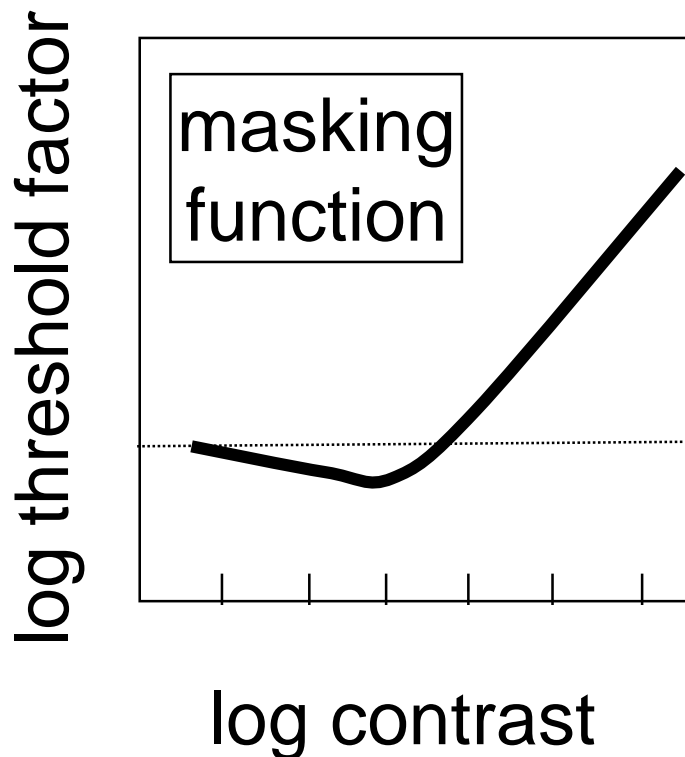
15%



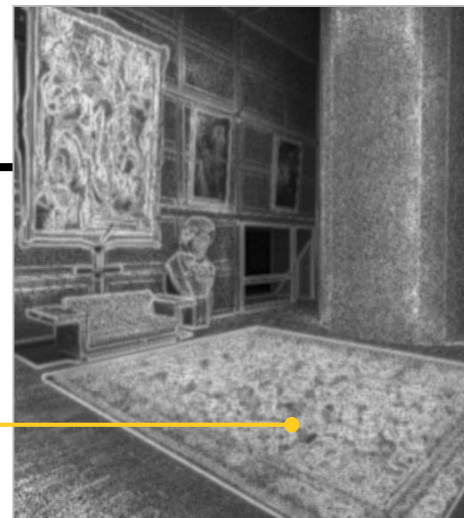
threshold due to
luminance + freq.

Threshold Model

3. Contrast component (visual masking)



15%



30%

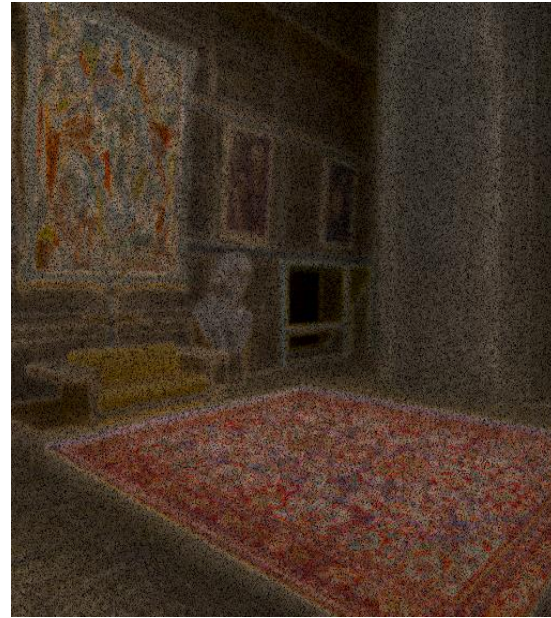


threshold due to
luminance + freq.
+ contrast

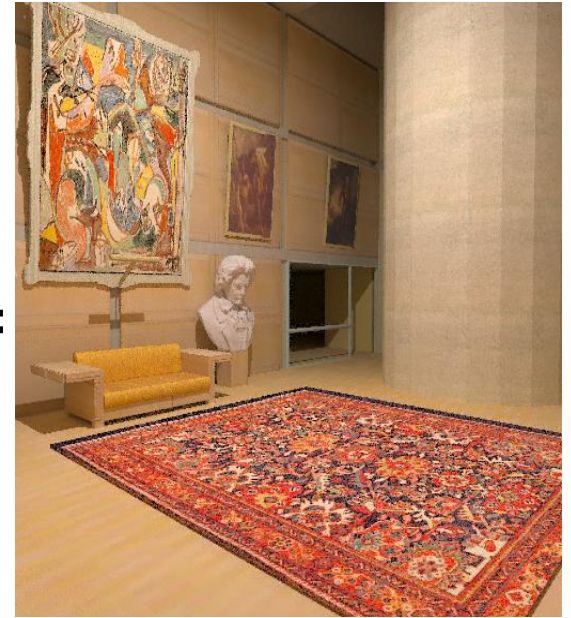
Validation



+



=

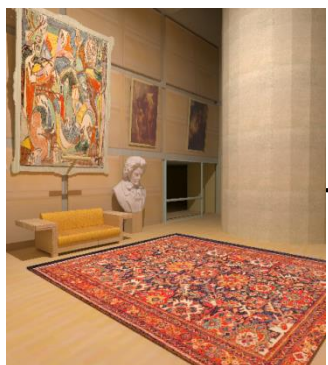


image

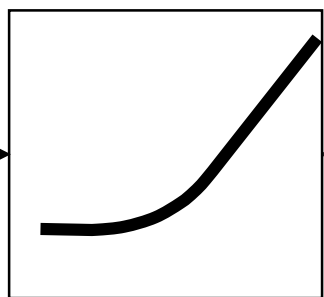
noise

image + noise

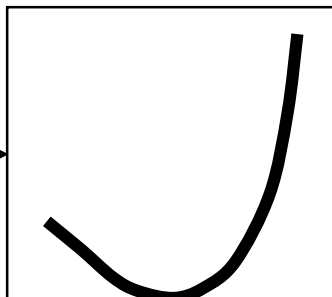
Threshold Model



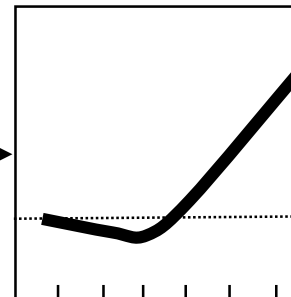
image



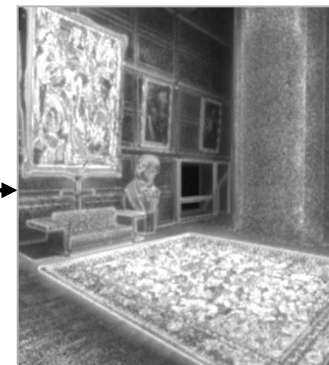
luminance
component



frequency
component

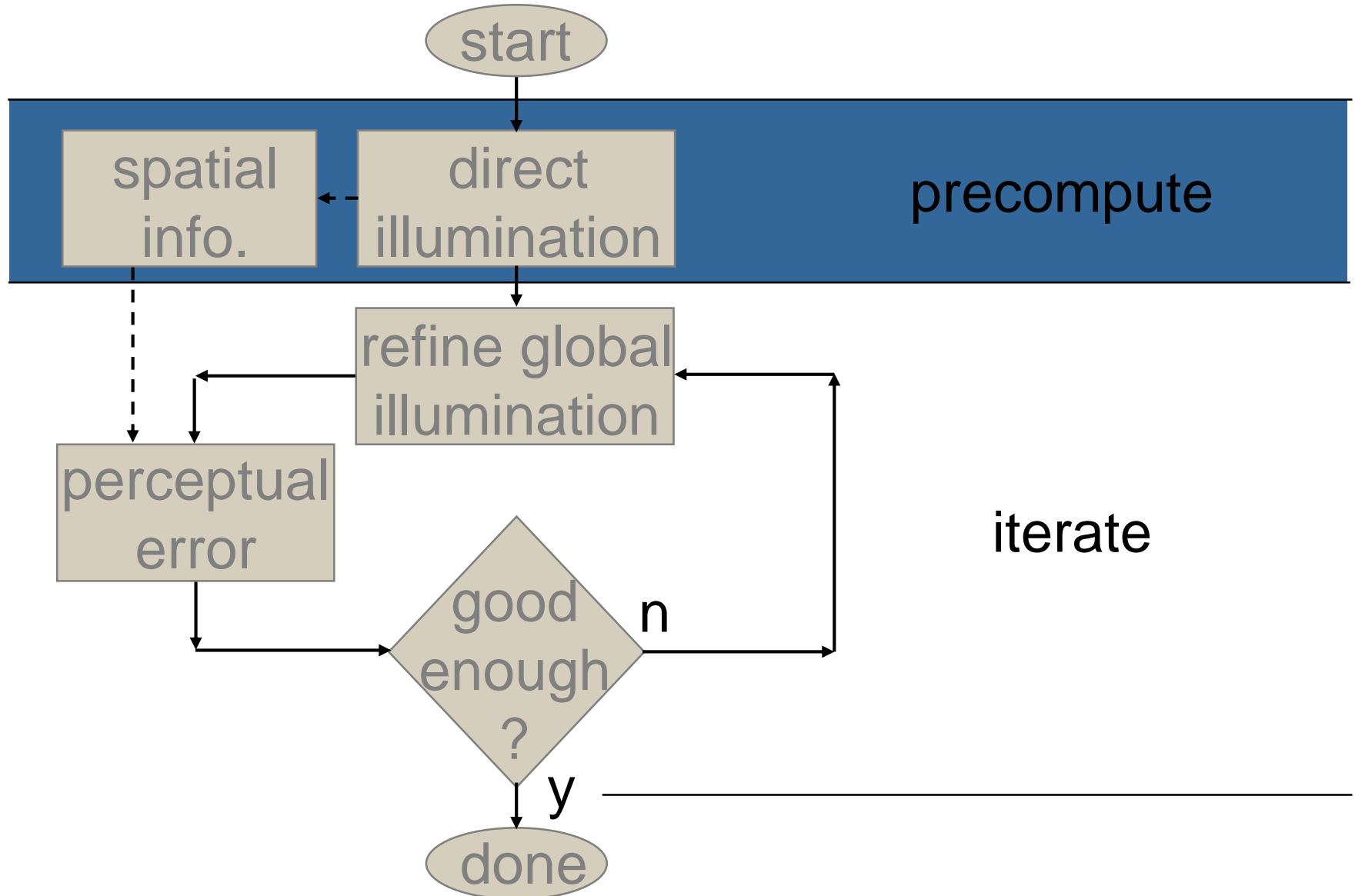


contrast
component



threshold
map

Adaptive Rendering Algorithm



Results

5% effort



reference
solution

effort
distribution
(darker
regions -
less effort)



adaptive
solution

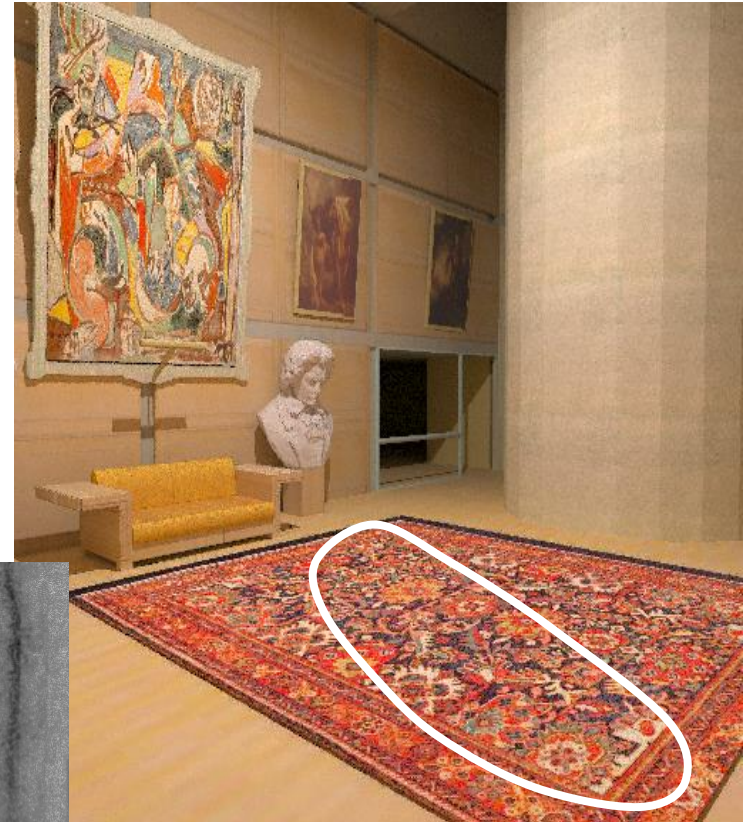
Results: Masking by Textures

5% effort



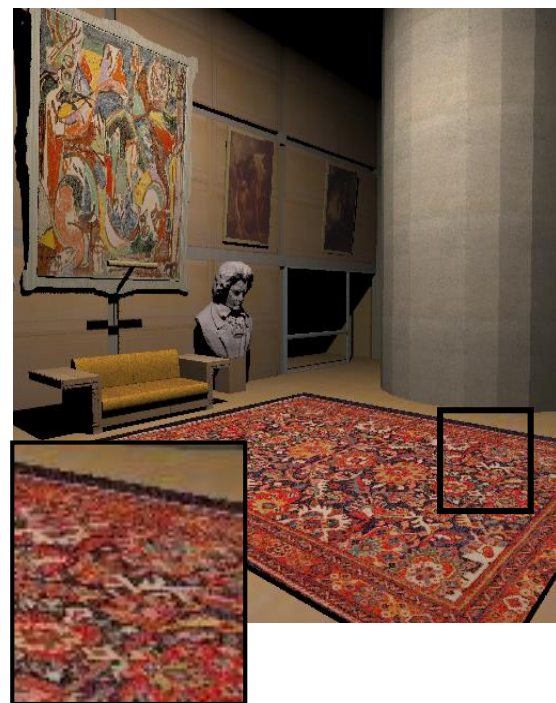
reference
solution

effort
distribution
(darker
regions -
less effort)



adaptive
solution

Results



direct
illumination

+



noisy

adaptive
indirect
illumination

=



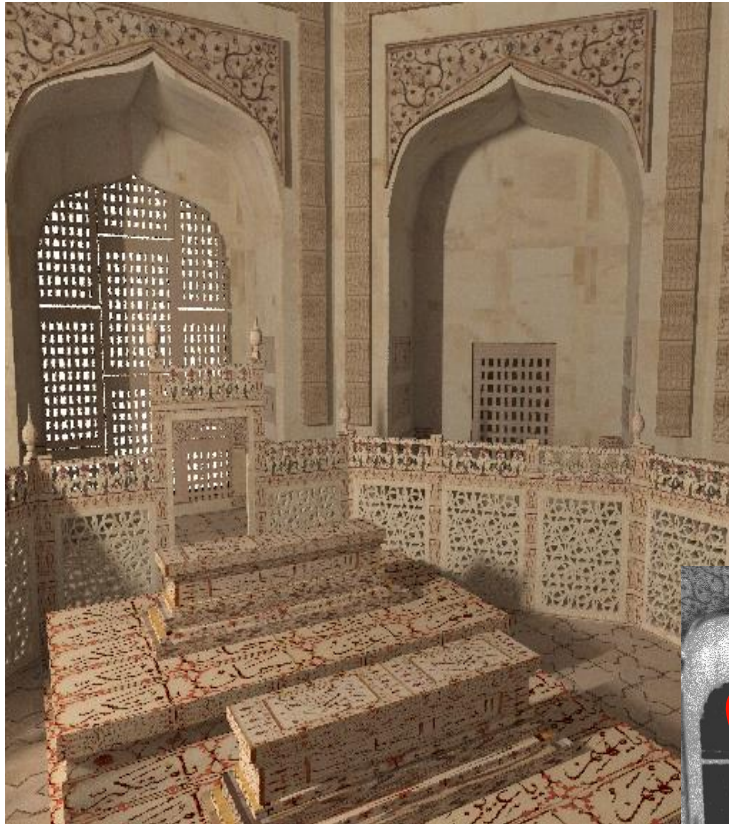
5% effort

masked

adaptive
global
illumination

Results: Masking by Geometry

5% effort

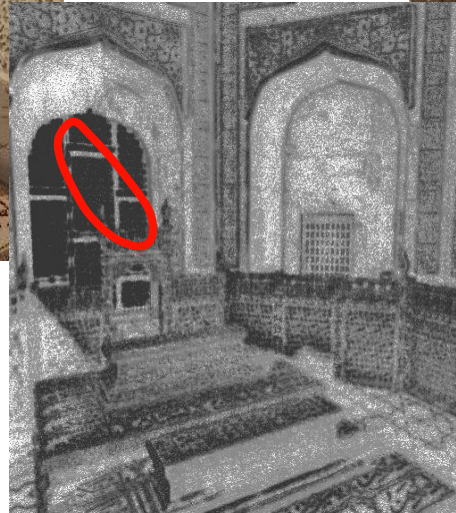


reference
solution

effort
distribution
(darker
regions -
less effort)

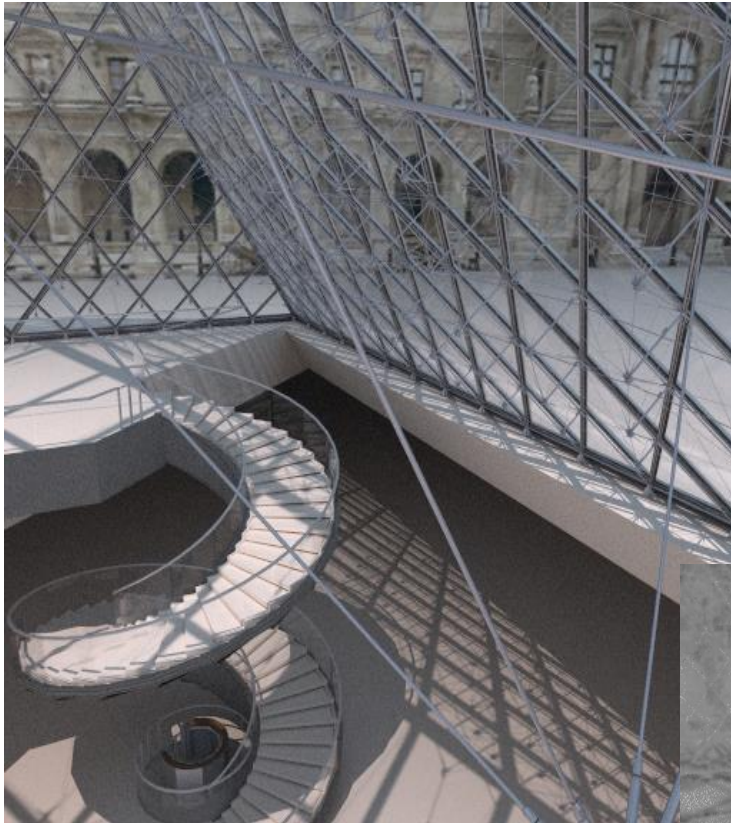


adaptive
solution



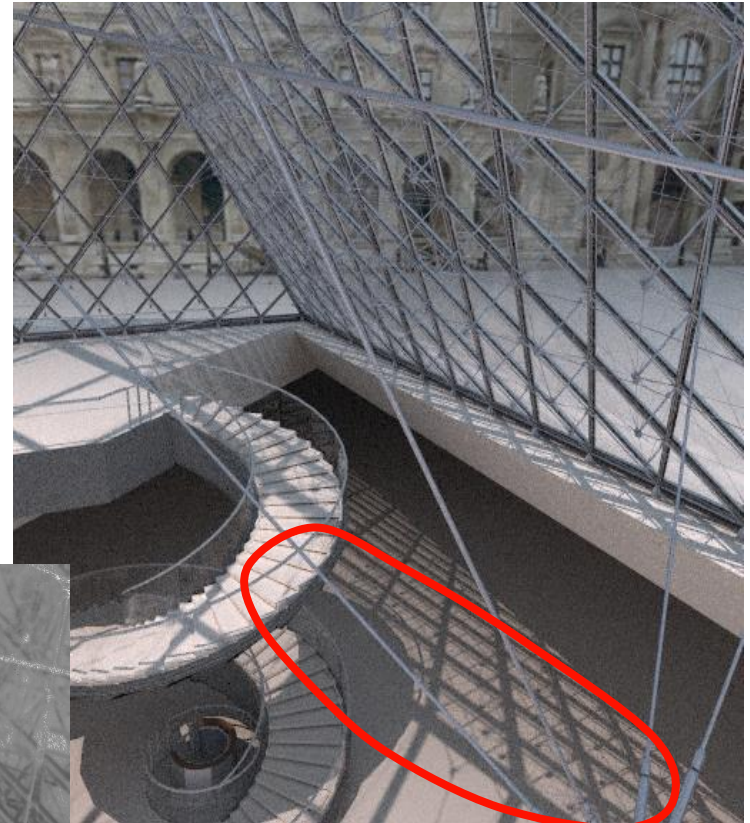
Results: Masking by Shadows

6% effort

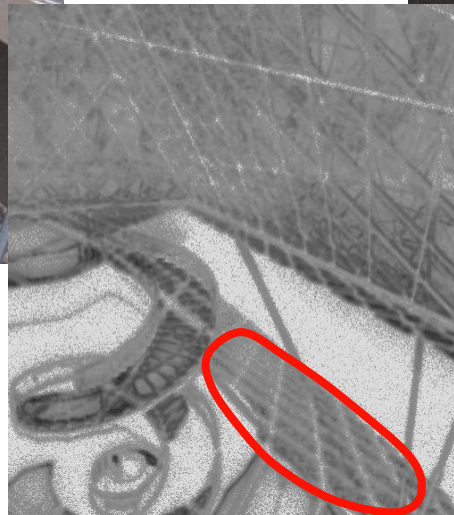


reference
solution

effort
distribution
(darker
regions -
less effort)



adaptive
solution



Eye Tracking - Motivation

1. Improving computational efficiency

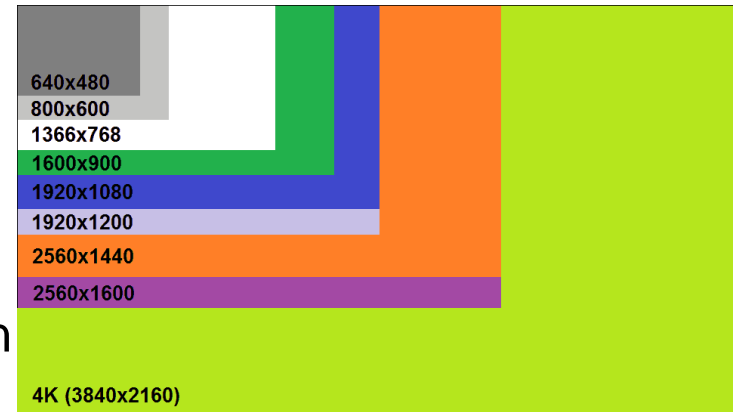
- There is a trend towards higher resolution displays
→ Higher computational requirement for 3D rendering
- Only a fraction of pixels is consciously attended and perceived in the full-resolution

2. Improving realism

- Eye is always focused on the screen plane; nevertheless, it is possible to simulate Depth-of-Field (DoF) effect by artificially blurring out-of-focus regions according to the gaze location

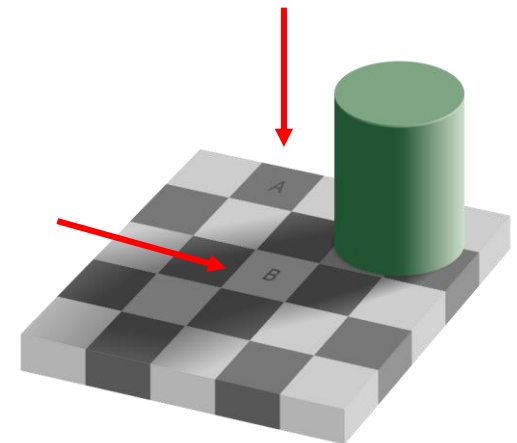
3. Improve perceived quality

- Human Visual System (HVS) has local adaptation property
- Perception of luminance, contrast and color are not absolute and highly dependent on both spatial and temporal neighborhood of the gaze location



Evolution of computer screen sizes.

<https://www.gamingscan.com/best-resolution-for-gaming/>

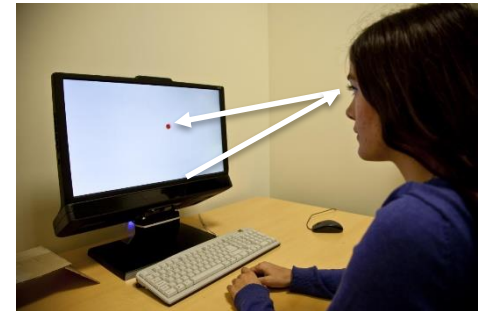
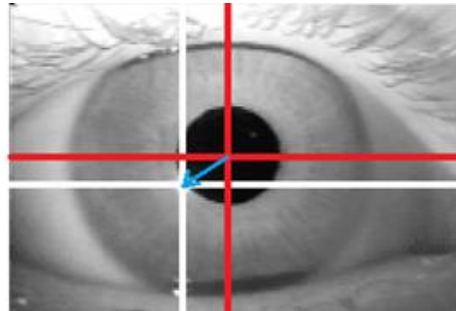
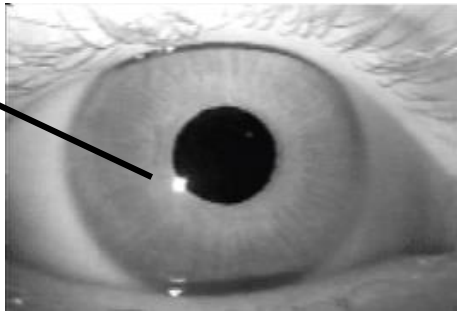


Checker shadow illusion:
A and B, as indicted by
the red arrows, the same
luminance [E. Adelson]

Eye Tracking

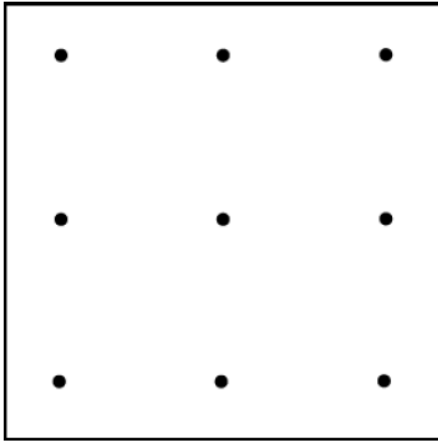
- **Basic Technology:**

Corneal Reflection (also known as “glint” or “1st Purkinje Reflection”)

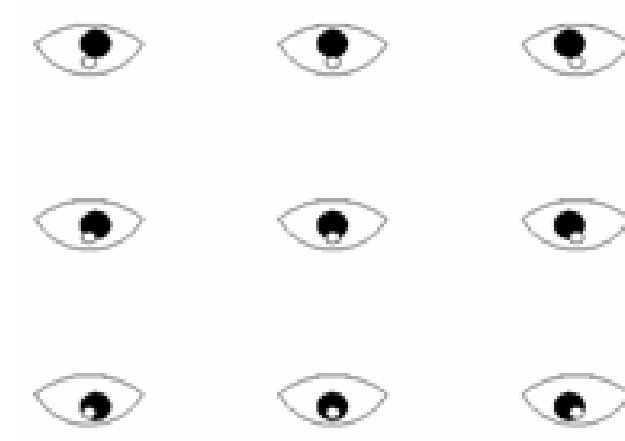


- **Eye trackers mostly operate using infrared imaging technology**
- **Once the pupil is detected the vector between the center of the pupil and the corneal reflection of the infrared light source is translated into the gaze location on screen coordinates**
- **Requires calibration at the beginning**

Eye Tracking



Sample 9-point calibration grid



Relative positions of the pupil and the corneal reflection

- **Individual calibration is necessary for each observer**
- **Relative location of the corneal reflection and the pupil is different among the population due to**
 - Difference in eye ball radius and shape
 - Eye-glasses

Images adapted from <http://wiki.cogain.org>

Eye Tracking



Chin-rest (EyeLink 1000/2000)



Glasses (SMI Eye Tracking Glasses)



Head-mounted displays (Oculus Rift)

- **Some of the other types of setups are used only for specific applications since they may be highly intrusive (e.g. chin-rest eye trackers) and not comfortable for the end-users in practice**
- **Head-mounted displays (HMD) offer 3D stereo and augmented reality capabilities in addition to eye tracking**

Types of Eye Motion

Type	Duration (ms)	Amplitude (1° = 60')	Velocity
Fixation	200-300	-	-
<i>Microsaccade</i>	10-30	10-40'	15-50°/s
<i>Tremor</i>	-	<1'	20'/sec
<i>Drift</i>	200-1000	1-60'	6-25'/s
Saccade	30-80	4-20°	30-500°/s
Glissade	10-40	0.5-2°	20-140°/s
Smooth Pursuit	variable	variable	10-30°/s

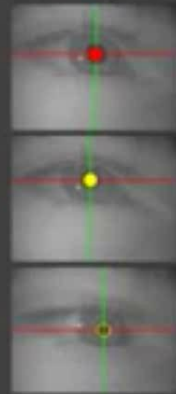
- While the mechanisms are not exactly known, it is thought that the brain performs visual suppression and compensation during **saccades** and smooth pursuits against motion blur on the retina.

Reference: Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. OUP Oxford.

Eye Tracking in Action

Bayesian Identification of Fixations, Saccades, and Smooth Pursuits

An example of I-BDT classification



Fixation = Solid Red Circle

Saccade = Solid Yellow Circle

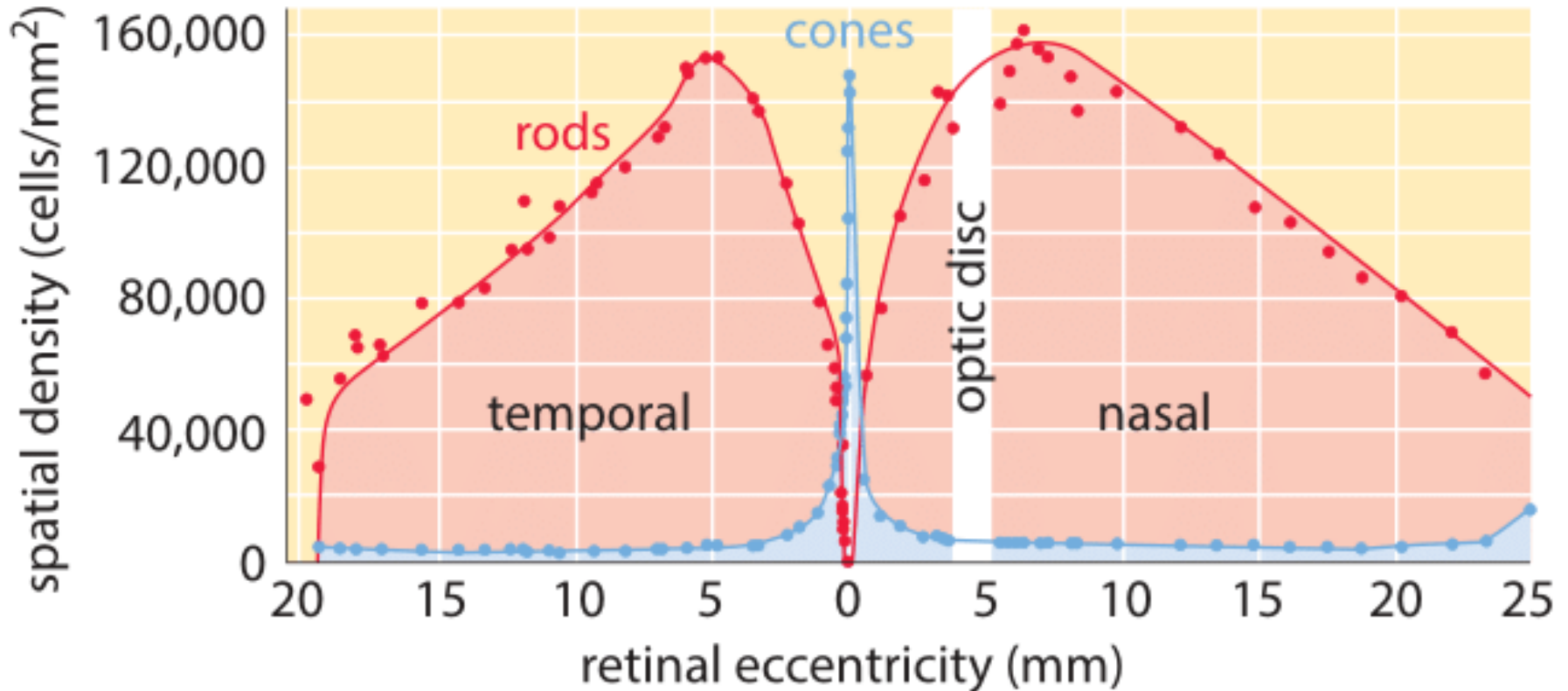
Smooth Pursuit = Hollow Yellow Circle

Original framerate: 30 Hz
Playback framerate: 10 Hz

Adapted from T. Santini, W. Fuhl, T. Kübler, and E. Kasneci. Bayesian Identification of Fixations, Saccades, and Smooth Pursuits ACM Symposium on Eye Tracking Research & Applications, ETRA 2016.

Visual Acuity

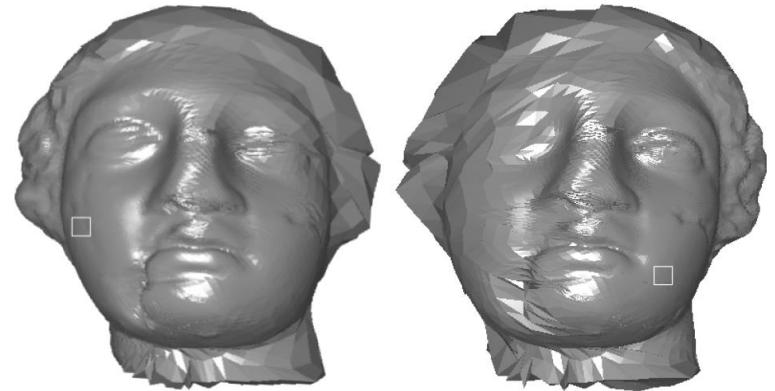
- **Distribution of photoreceptor cells in the retina**



Adapted from R. W. Rodieck, *The First Steps of Seeing*, Sinauer Associates, 1998.

Level-of-Detail Rendering

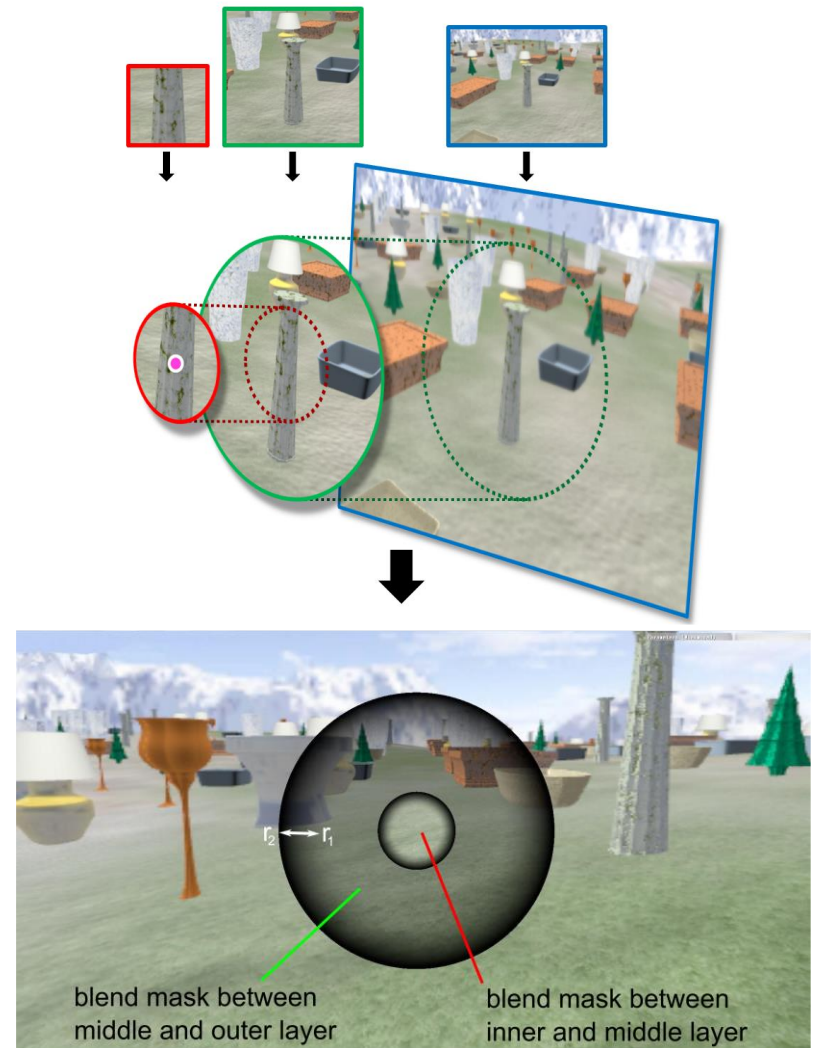
- **The model resolution may be degraded according to the visual angle and the acuity of HVS at the given angle**
 - Mesh structure of the model is partitioned into tiles using Voronoi diagram
 - Tiles are mapped to planar polygons
 - Remeshing into multiresolution form



Adapted from Murphy, Hunter, and Andrew T. Duchowski. "Gaze-contingent level of detail rendering." EuroGraphics 2001 (2001).

Foveated 3D Graphics

- Screen-based (in contrast to model-based methods)
- Human eye has full acuity in around 5° foveal region
- The efficiency of image generation can be improved by maintaining high image resolution only around the gaze location
- Using 60Hz monitor and Tobii X50 eye tracker with 50Hz sampling frequency and 35ms latency caused artifacts for the observer
- Results using 120Hz monitor and Tobii TX300 with 300Hz sampling frequency and 10ms latency were tolerable



Foveated 3D Graphics

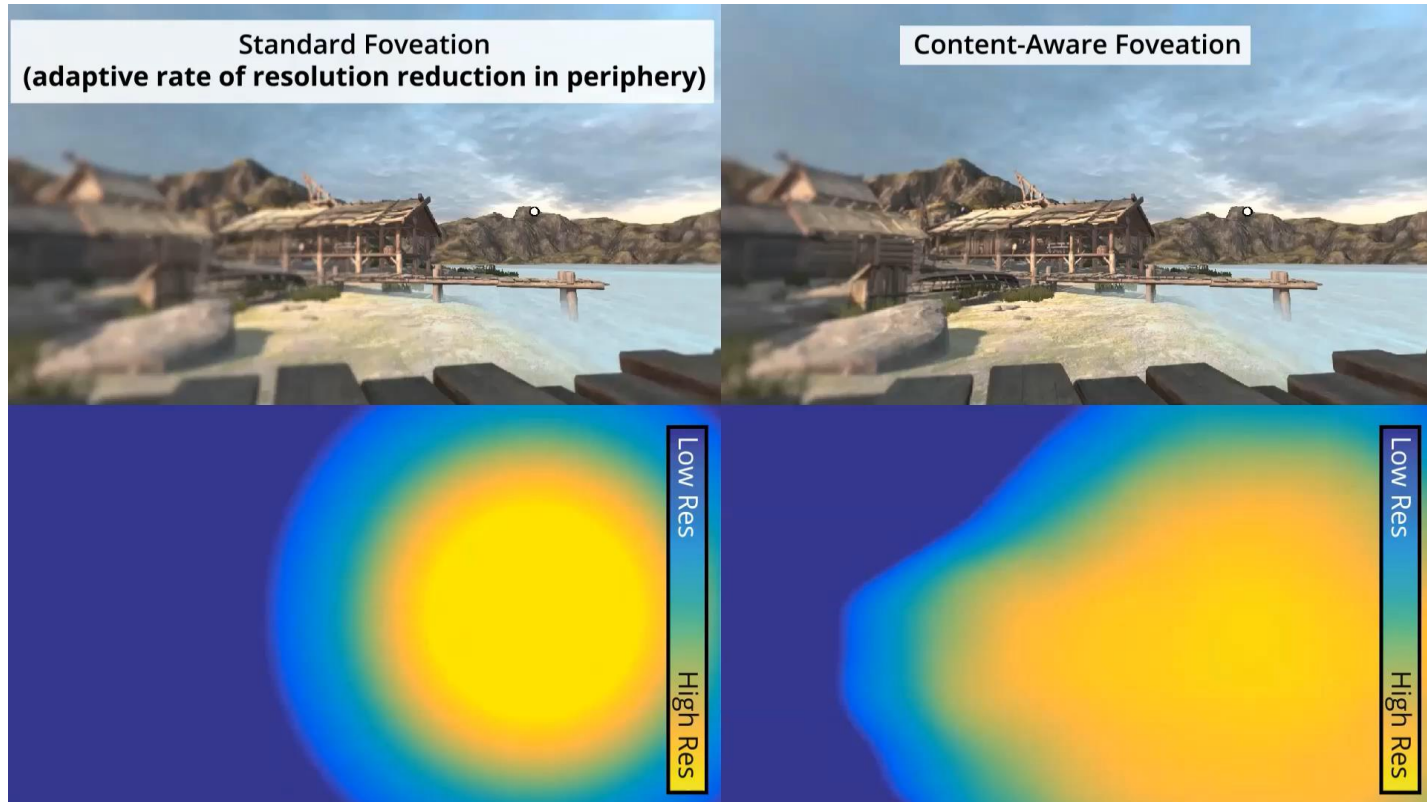


Video adapted from <http://research.microsoft.com>

Luminance-Contrast-Aware Foveated Rendering



Luminance-Contrast-Aware Foveated Rendering



Effect of Depth-of-Field

- Improves the rendering realism and enhances the depth perception



(a) Image focused on objects at shallow depth (flower)



(b) Image focused on objects at large depth (Main Quad)



(c) Image with everything in focus



Images adapted from Gupta, Kushagr, and Suleman Kazi, "Gaze Contingent Depth of Field Display", 2016. Video adapted from Mantiuk, Radoslaw, Bartosz Bazyluk, and Rafal K. Mantiuk. "Gaze-driven Object Tracking for Real Time Rendering." Computer Graphics Forum. Vol. 32. No. 2pt2. Blackwell Publishing Ltd, 2013.

Depth-of-Field Rendering

- **Circle of Confusion :**

$$CoC = a \cdot \left| \frac{f}{d_0 - f} \right| \cdot \left| 1 - \frac{d_0}{d_p} \right|$$

a - diameter of the lens aperture

f - focal length of the lens

d_0 - distance between the focal plane and lens

d_p - distance from an object to the lens

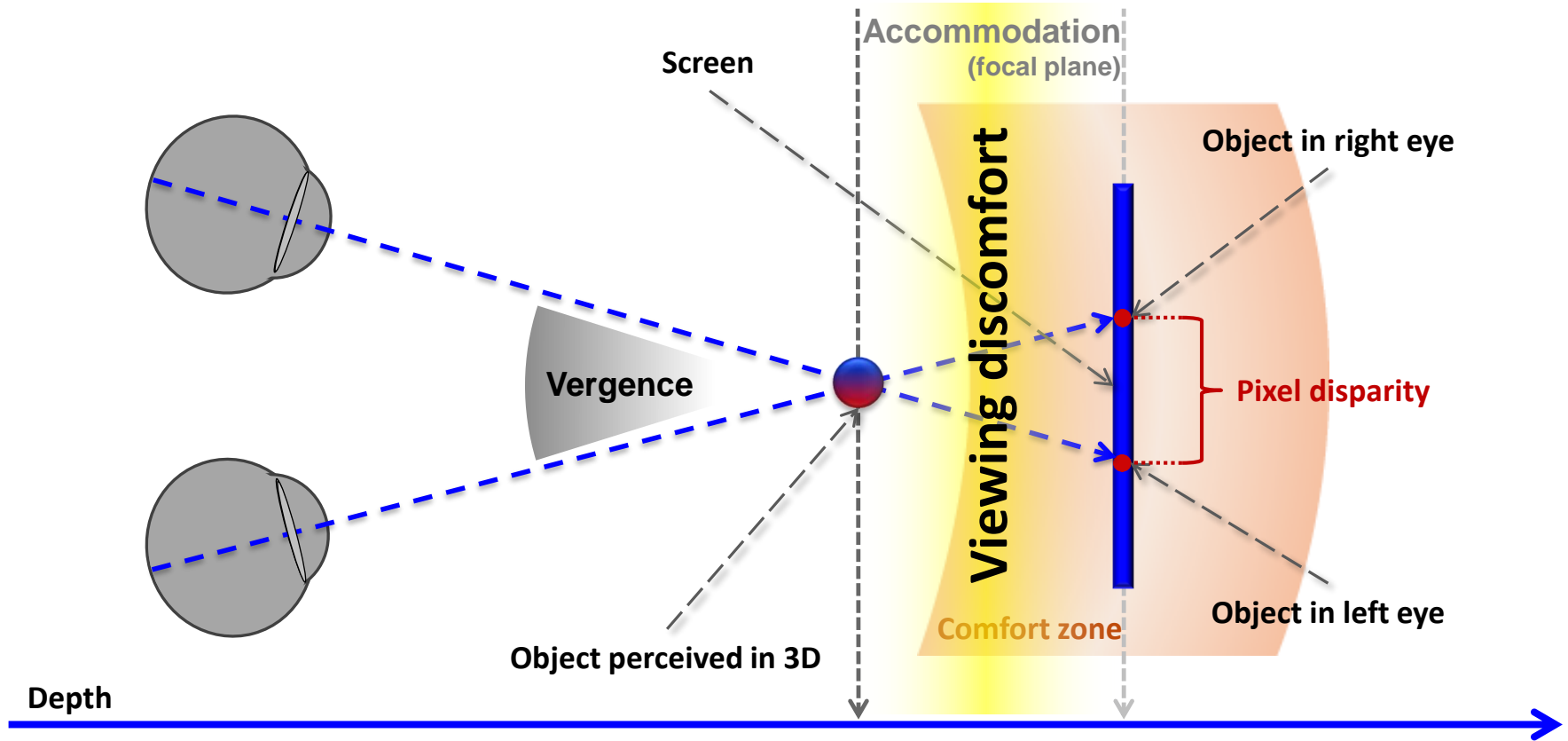
- d_p is obtained from reverse mapping of the z-buffer

- **Addresses the artifacts due to the depth discontinuity near object boundaries by spreading the blur outside the object boundary**



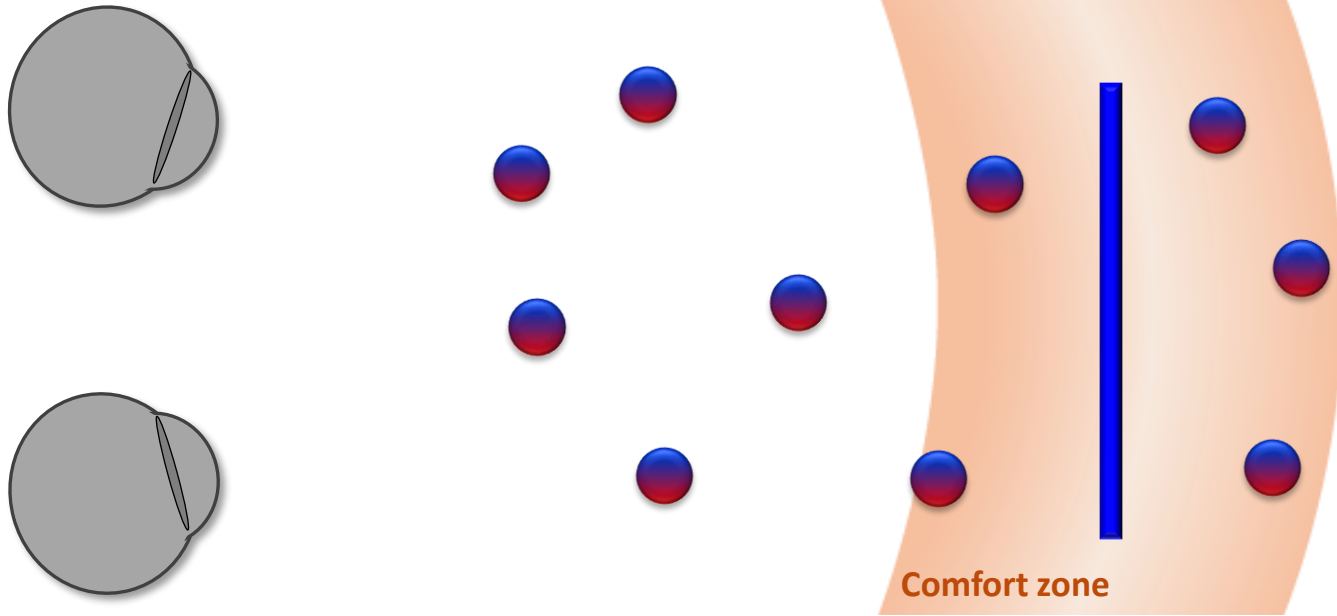
Vergence-accommodation Conflict

Stereo 3D: Binocular Disparity



Vergence-accommodation Conflict

Depth Manipulation



Scene manipulation
~~Viewing discomfort~~ Viewing comfort

Vergence-accommodation Conflict

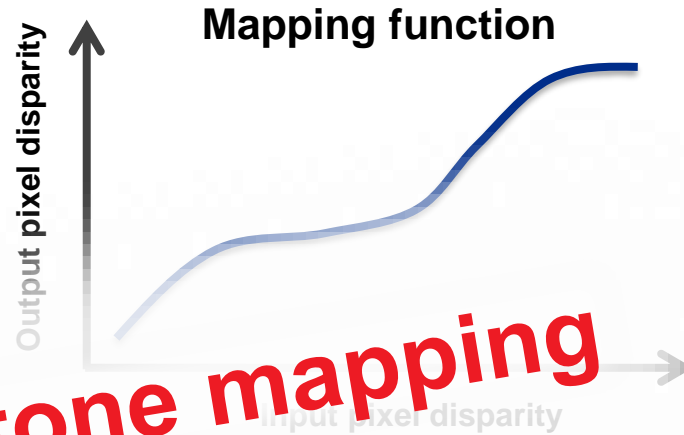
Depth Manipulation



Pixel disparity map



Modified pixel disparity



Similar to tone mapping

Function:

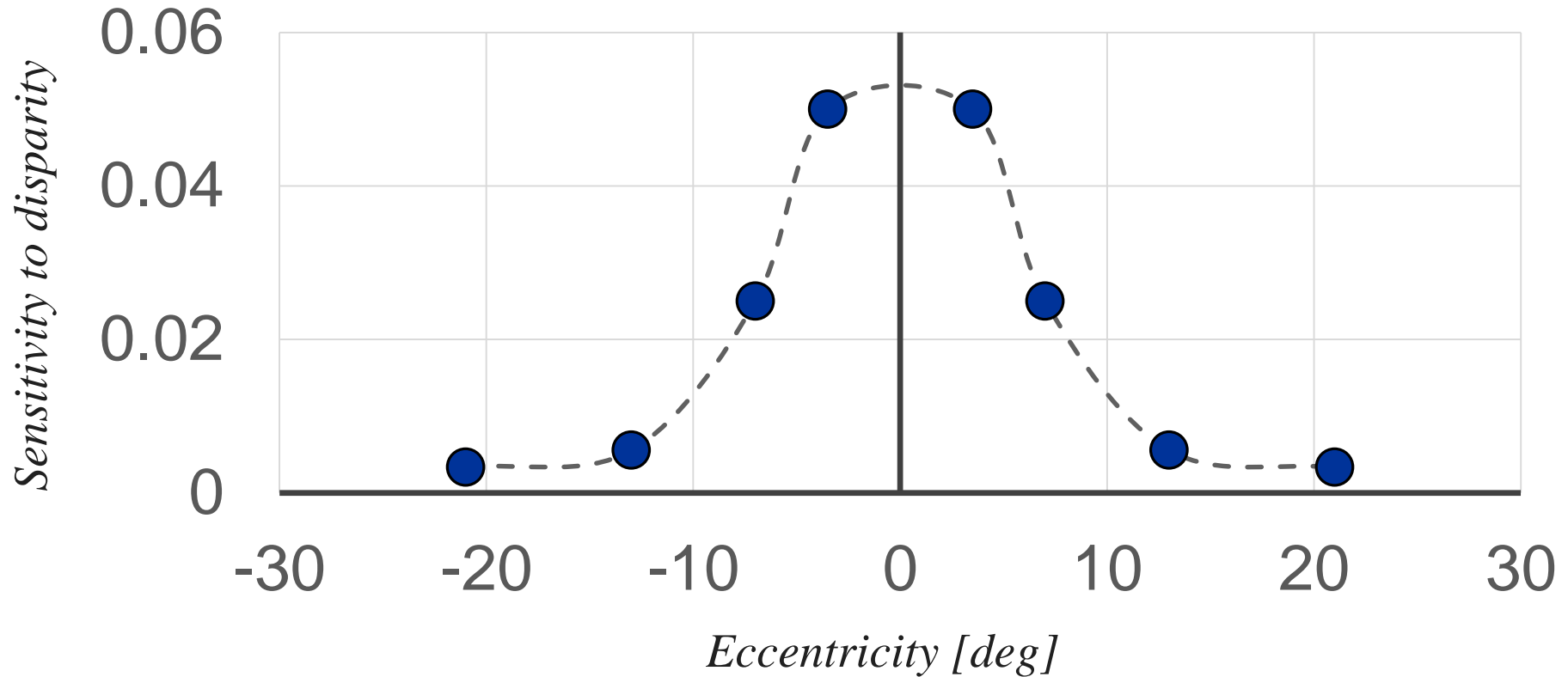
- Linear
- Logarithmic
- Content dependent

Other possibilities:

- Gradient domain
- Local operators

"Nonlinear Disparity Mapping for Stereoscopic 3D" [Lang et al. 2010]

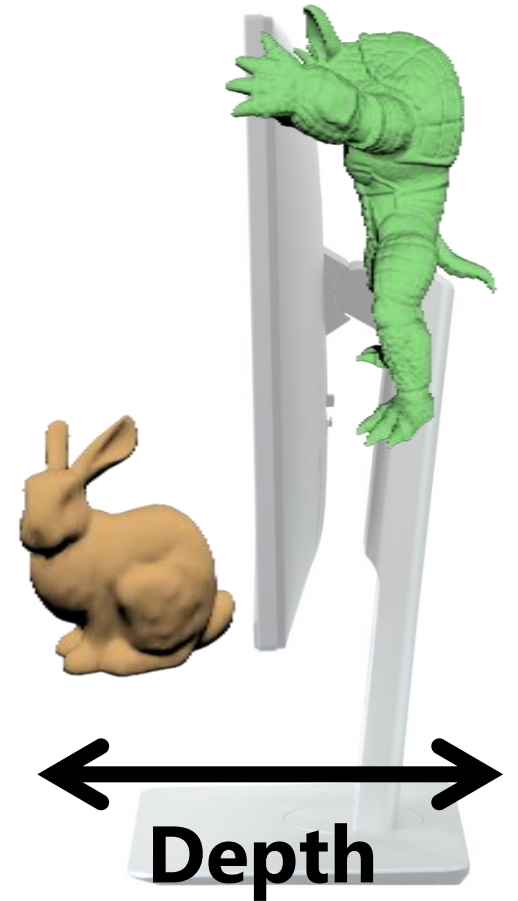
Disparity Perception (Stereo 3D)



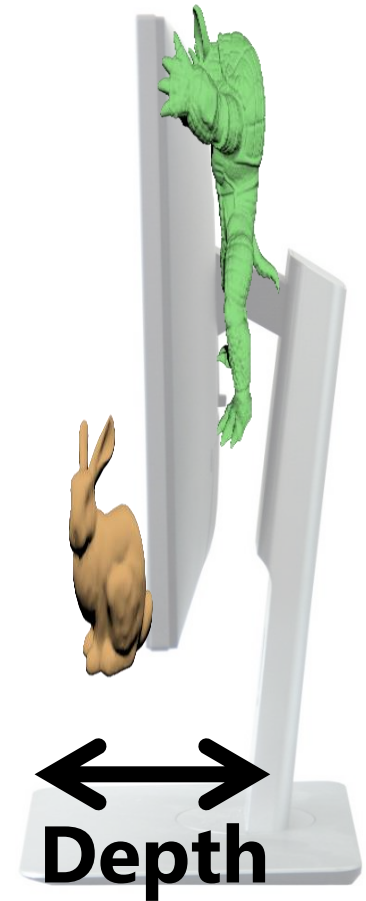
Replotted from Figure 3 of Simon J.D Prince, Brian J Rogers

Sensitivity to disparity corrugations in peripheral vision, Vision Research, Volume 38, Issue 17, September 1998

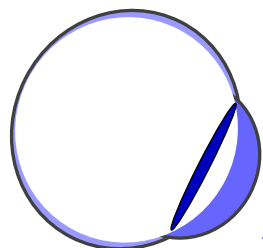
Vergence-accommodation Conflict



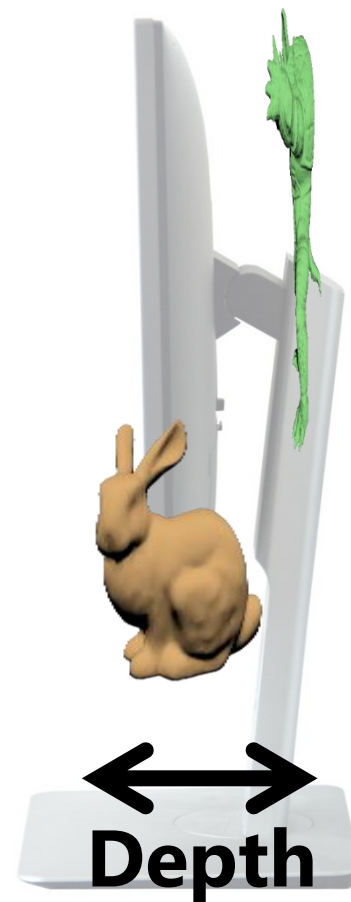
Vergence-accommodation Conflict



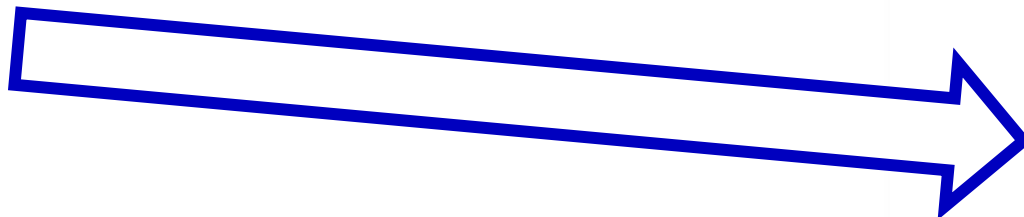
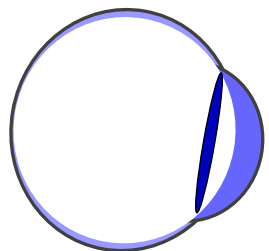
Vergence-accommodation Conflict



✓ More depth

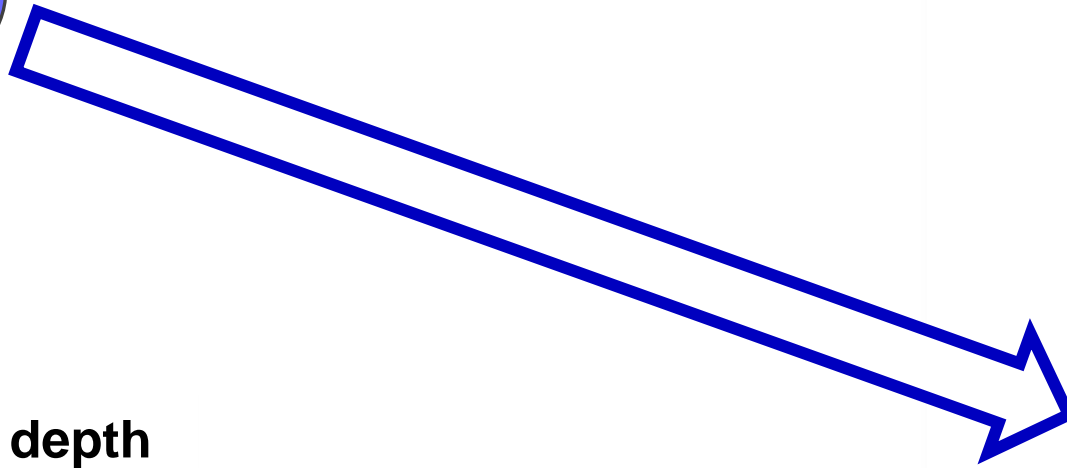
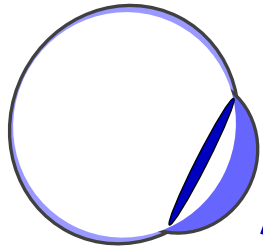


Vergence-accommodation Conflict

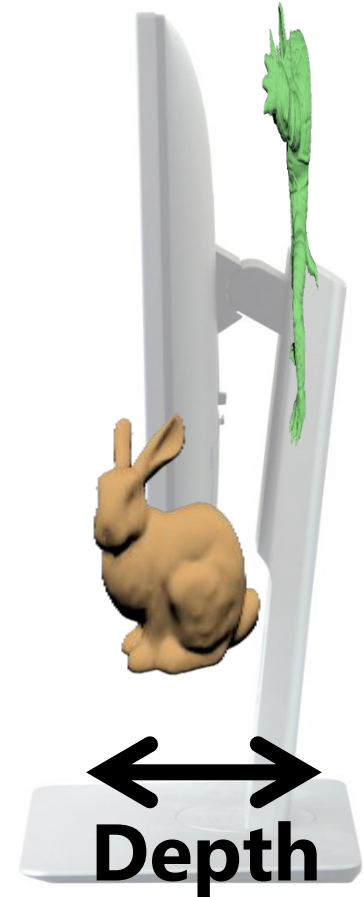


- ✓ More depth
- ✓ More comfort

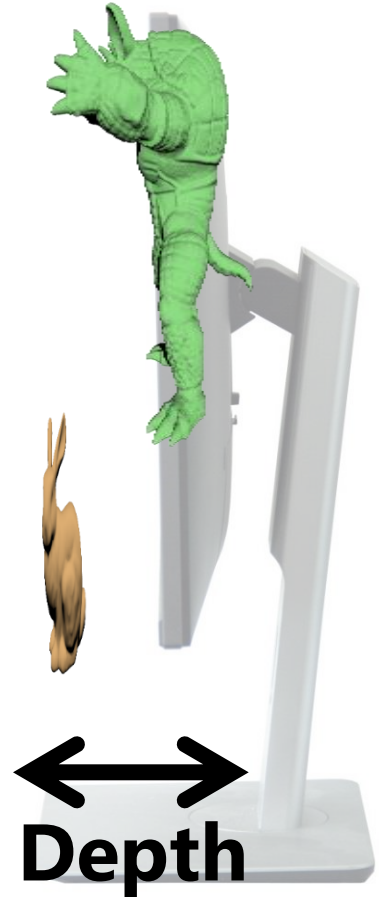
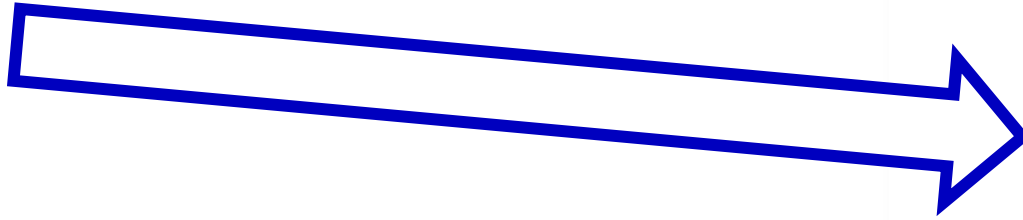
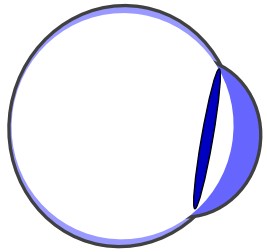
Vergence-accommodation Conflict



- ✓ **More depth**
- ✓ **More comfort**
- ✓ **Seamless**



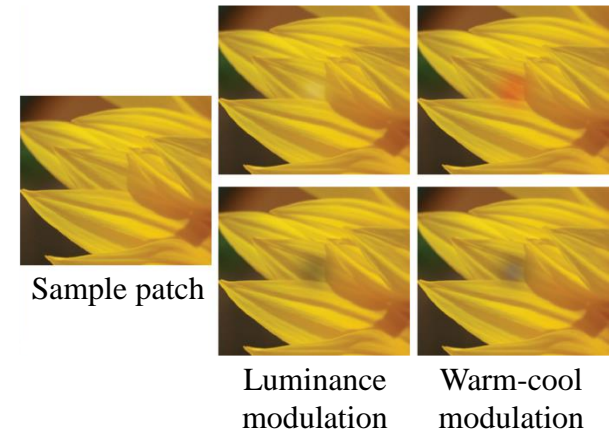
Vergence-accommodation Conflict



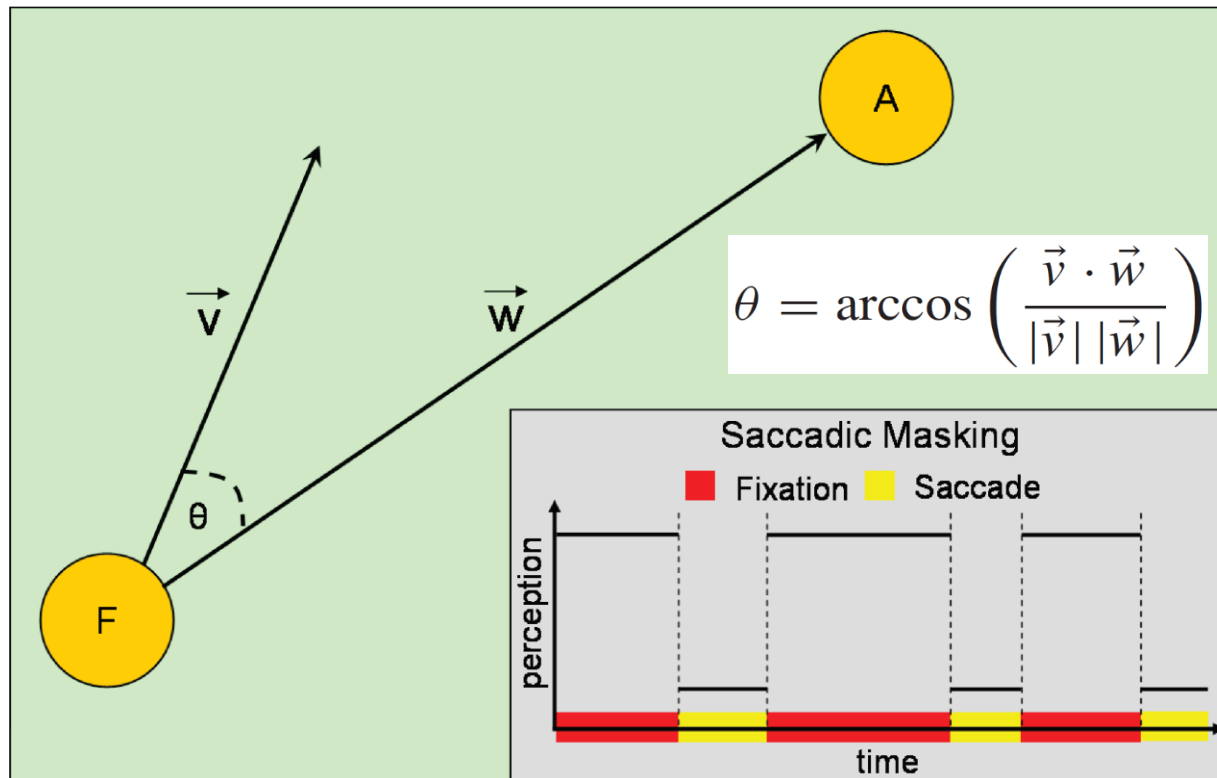
- ✓ More depth
- ✓ More comfort
- ✓ Seamless
- ✓ Low cost

Subtle Gaze Direction

- When viewing an image low-acuity peripheral vision detects areas of interest, then HVS directs gaze to those locations
- HVS is very sensitive to changes in luminance (Spillmann et al. 1990) and opponent color channels (Hurvich and Jameson 1957)
- Introduces subtle image modulation to control the gaze direction of the observer
- Luminance and warm-cool modulations are studied and both are found successful



Subtle Gaze Direction

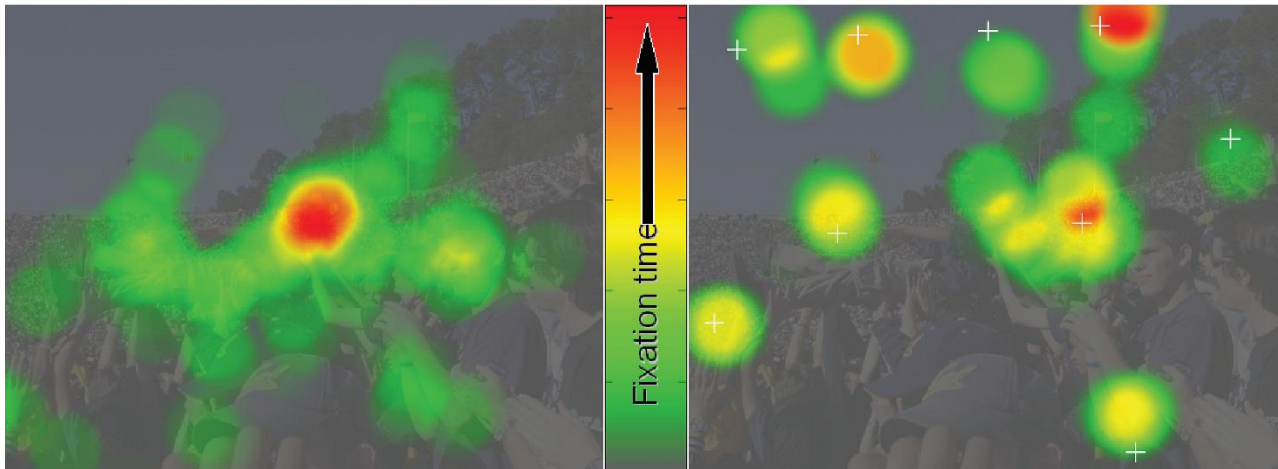


F: Fixation point, **A:** Predetermined Area of Interest

Goal: To direct the user attention to from **F** to **A**

Modulation is applied to A and θ is monitored real-time.
When $\theta \leq 10^\circ$, the modulation is terminated immediately.

Subtle Gaze Direction



Top: Input image, **Left:** No modulation, **Right:** Modulation at white crosses

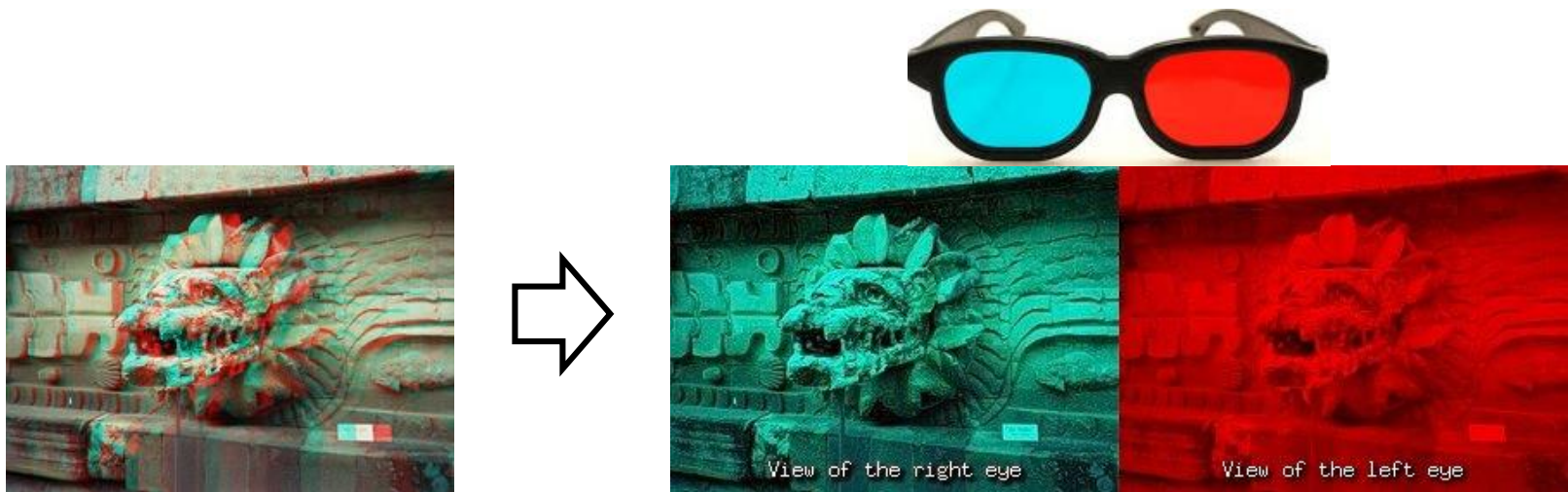
Images adapted from Bailey, R., McNamara, A., Sudarsanam, N., & Grimm, C. (2009). Subtle gaze direction. ACM Transactions on Graphics (TOG), 28(4), 100.

Binocular 3D Displays

- **Capable of providing sense of 3D by simulating binocular disparity**
 - Color Anaglyphs
 - Polarization
 - Shutter Glasses
 - Head-Mounted Displays
- **They mostly do not provide accommodation depth cue**

Color Anaglyphs

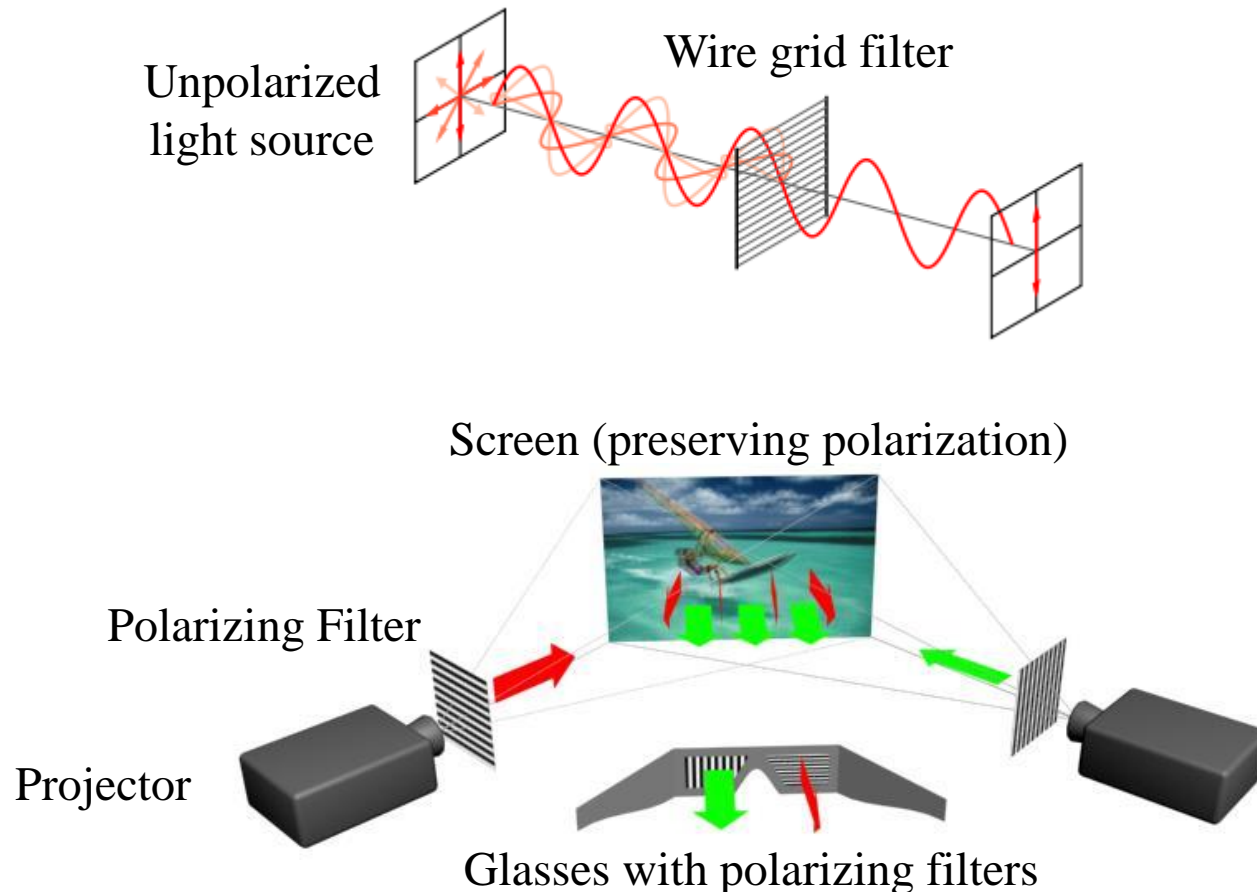
- **Left and right images are filtered using different colors (usually complementary):**
 - Red – Green, Red – Cyan, Green – Magenta
 - Amber – Blue (ColorCode 3D, patented [Sorensen et al. 2004])
- **Limited color perception (since each eye sees only a subset of whole colorspace)**



Images adapted from http://axon.physik.uni-bremen.de/research/stereo/color_anaglyph/

Polarization

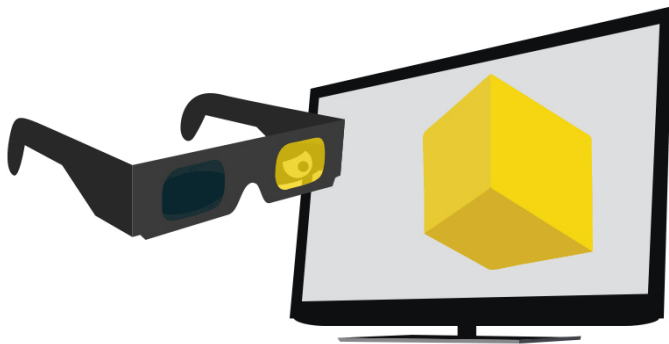
- Usually a wire grid filter converts the unpolarized light beam to a polarized one



Images adapted from https://cpinettes.u-cergy.fr/S6-Electromag_files/fig1.pdf

Shutter Glasses

- Exploits the “memory effect” of the Human Visual System [Coltheart 1980]
- Glasses have shutters which operate in synchronization with the display system
- Left and right eye images are shown in alternation
- Color neutral; however, temporal resolution is reduced



IR receiver for synchronization

Images adapted from https://en.wikipedia.org/wiki/Active_shutter_3D_system

Head-Mounted Displays

- **Separate displays for the left and right eye**
- **May provide current orientation of the head (and update the stimuli accordingly to provide a VR)**



Images adapted from <http://www.oculus.com>

Autostereoscopic Displays

- **Stereo displays which are viewable without special glasses or head-wear equipment**
- **Simulate an approximate lightfield with a finite number of views**
 - Parallax Barriers
 - Integral Imaging
 - Multi-layer Displays

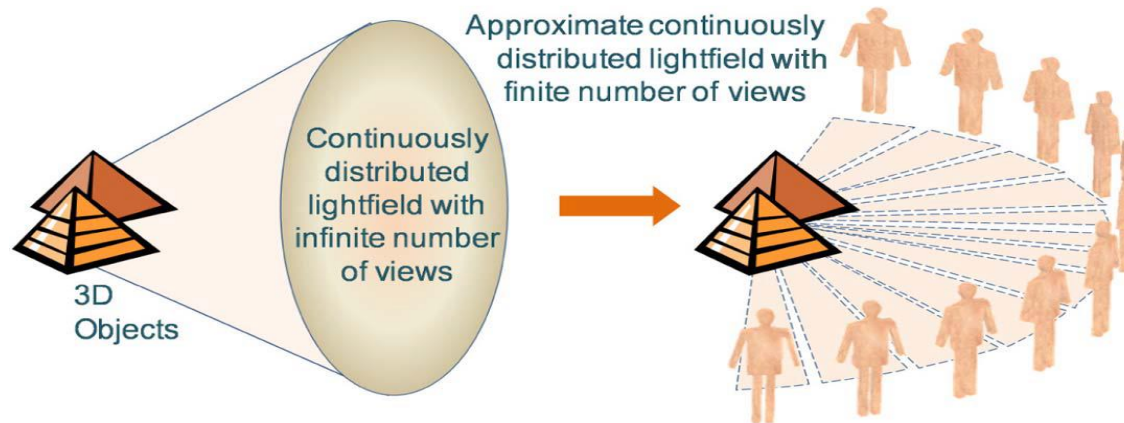
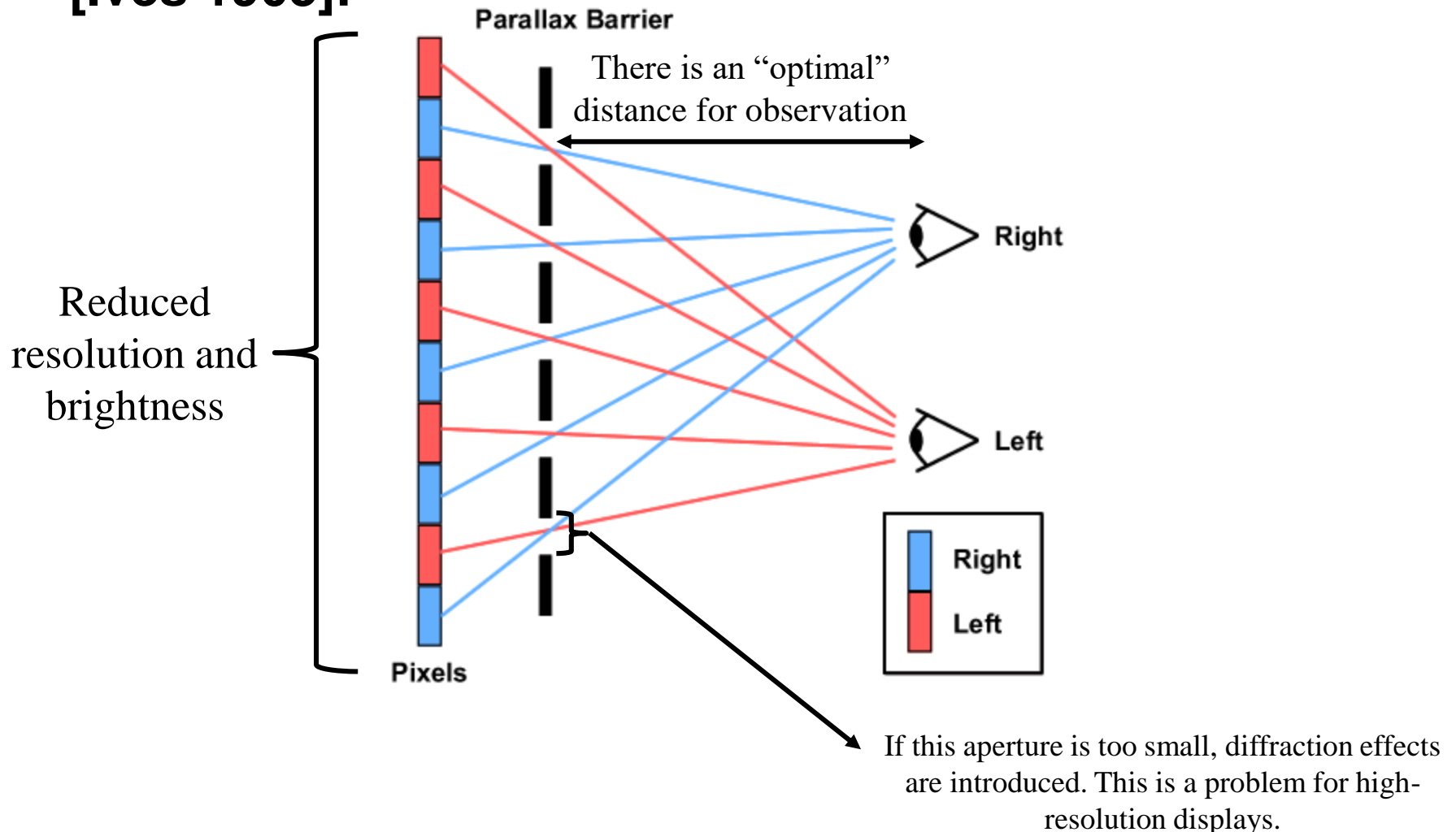


Image adapted from Geng, Jason. "Three-dimensional display technologies." *Advances in optics and photonics* 5.4 (2013): 456-535.

Parallax Barriers

- Occlusion-based working principle and key features [Ives 1903]:



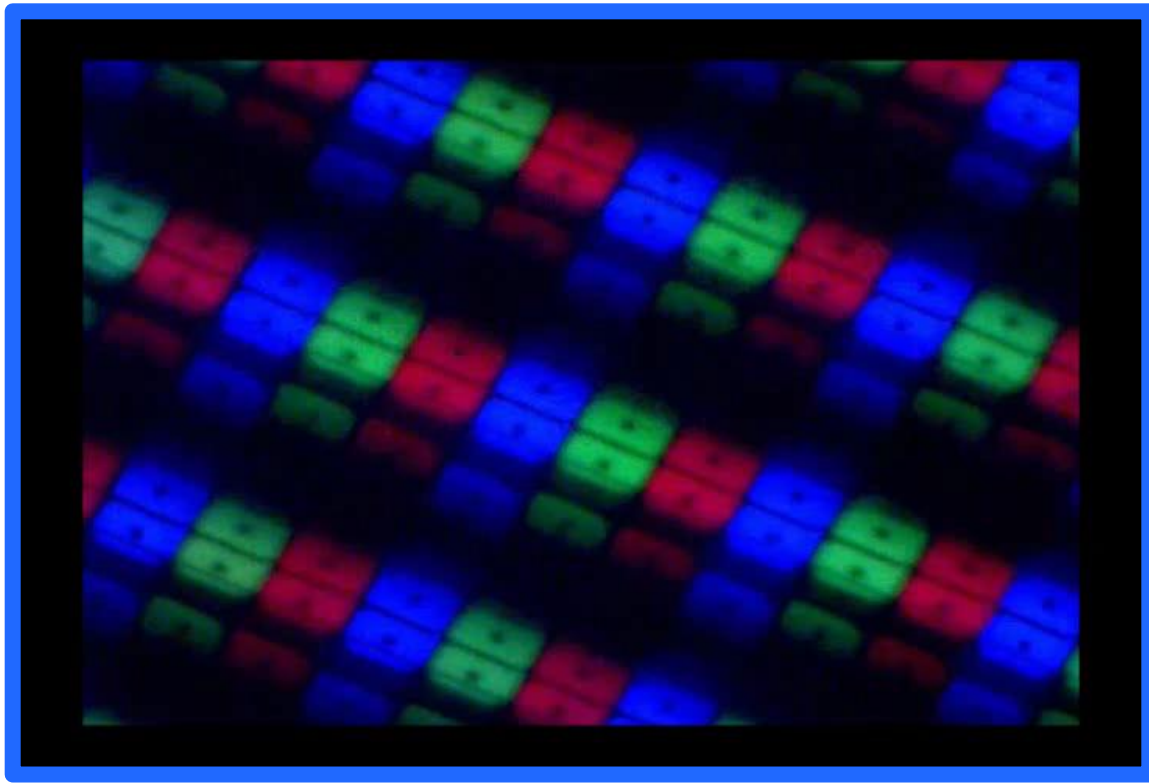
Parallax Barriers



Video adapted from: <http://www.youtube.com/watch?v=sxF9PGRiabw> “Glasses-Free 3D Gaming for \$5 (Parallax Barrier)”

Parallax Barriers

- It is possible to switch between 2D and 3D modes
- Parallax barrier of Nintendo 3DS turning on/off under microscope:

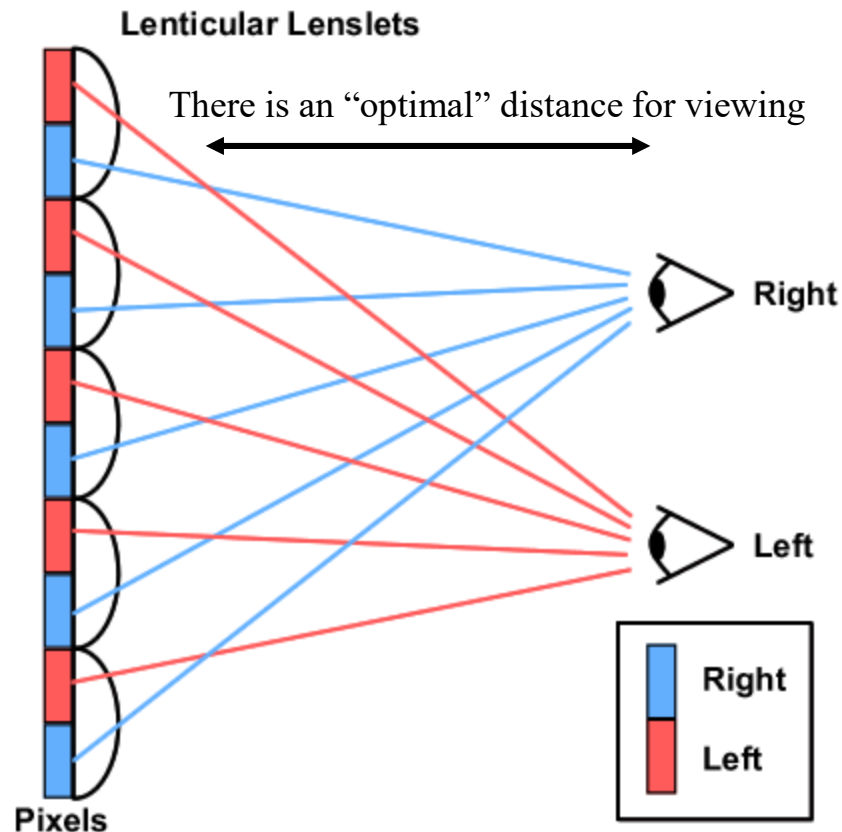


Video adapted from: <https://www.youtube.com/watch?v=D-LzRT7Bvc0>

Integral Imaging

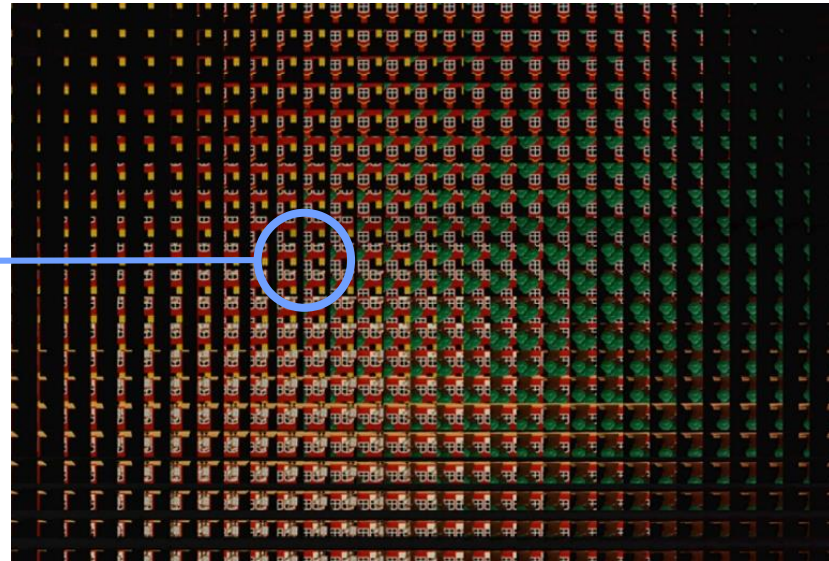
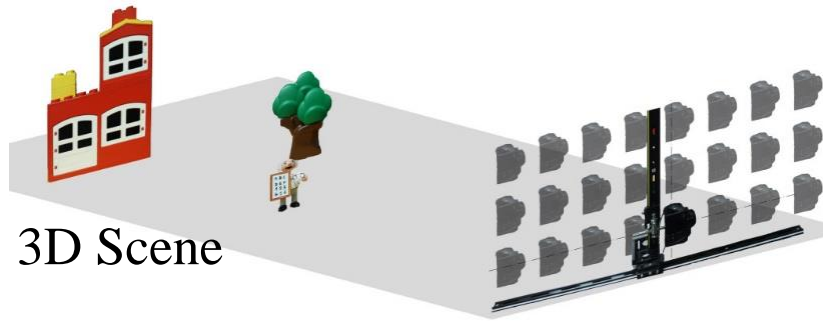
- **Refraction-based working principle [Lippmann 1908]:**

Reduction in resolution and brightness is still a problem.



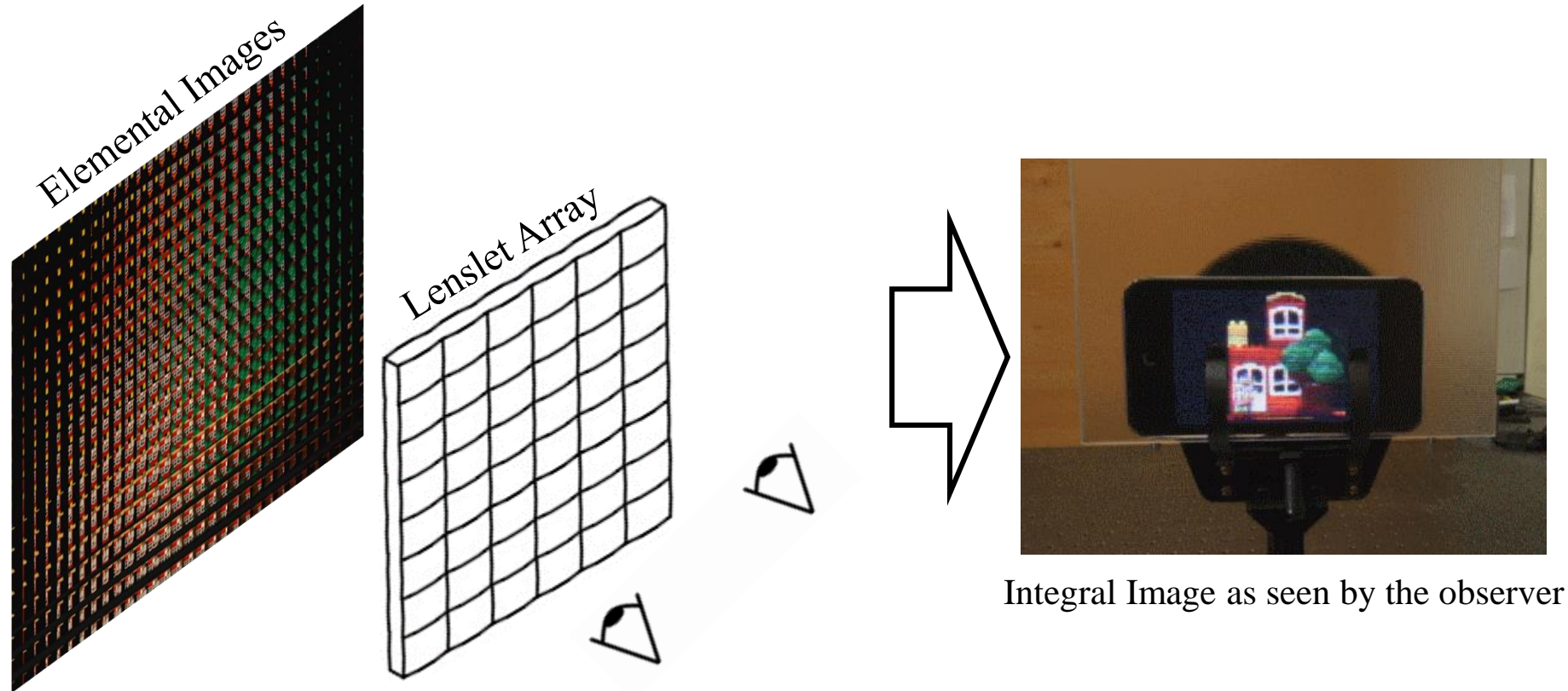
It is possible to reproduce parallax, perspective shift and accommodation depth cues.

Integral Imaging



Images adapted from Martinez-Corral, Manuel, et al. "3D integral imaging monitors with fully programmable display parameters."

Integral Imaging

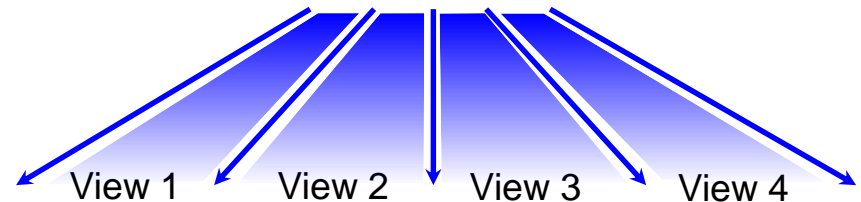


Images adapted from Martinez-Corral, Manuel, et al. "3D integral imaging monitors with fully programmable display parameters."

Multi-view Autostereoscopic Display

- **Smooth transitions**

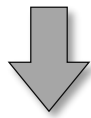
Multi-view autostereoscopic display



„Antialiasing for automultiscopic 3D displays” [Zwicker et al. 2006]

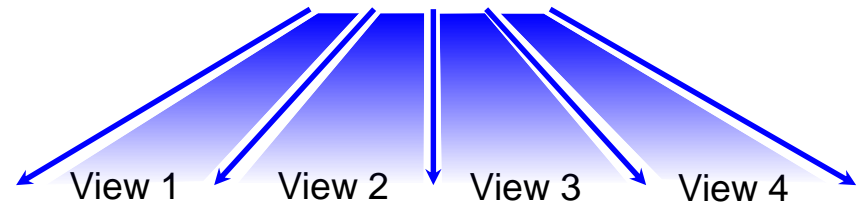
Multi-view Autostereoscopic Display

- **Smooth transitions**
- **Blur increases with depth**



Weaker depth percept

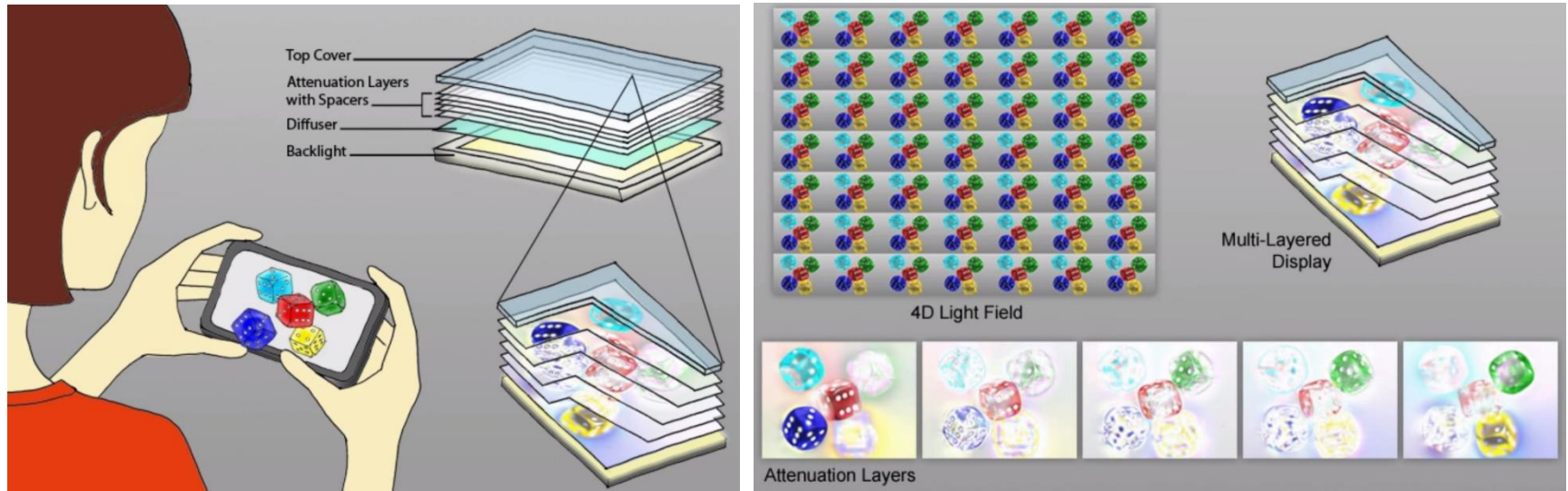
Multi-view autostereoscopic display



„Antialiasing for automultiscopic 3D displays” [Zwicker et al. 2006]

Multi-layer Displays

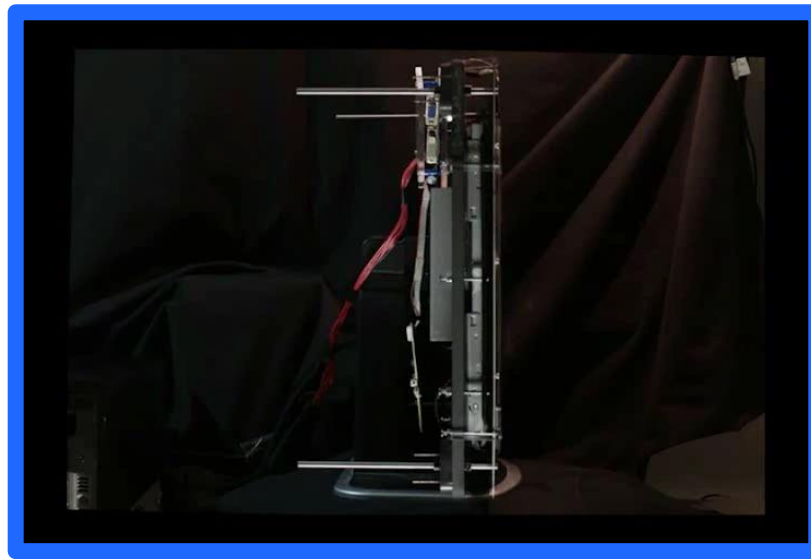
- Improved resolution over parallax barriers and lenslet arrays
- Provides a solution to accommodation-vergence conflict



Images adapted from Wetzstein, Gordon, et al. "Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays." ACM Transactions on Graphics (ToG). Vol. 30. No. 4. ACM, 2011.

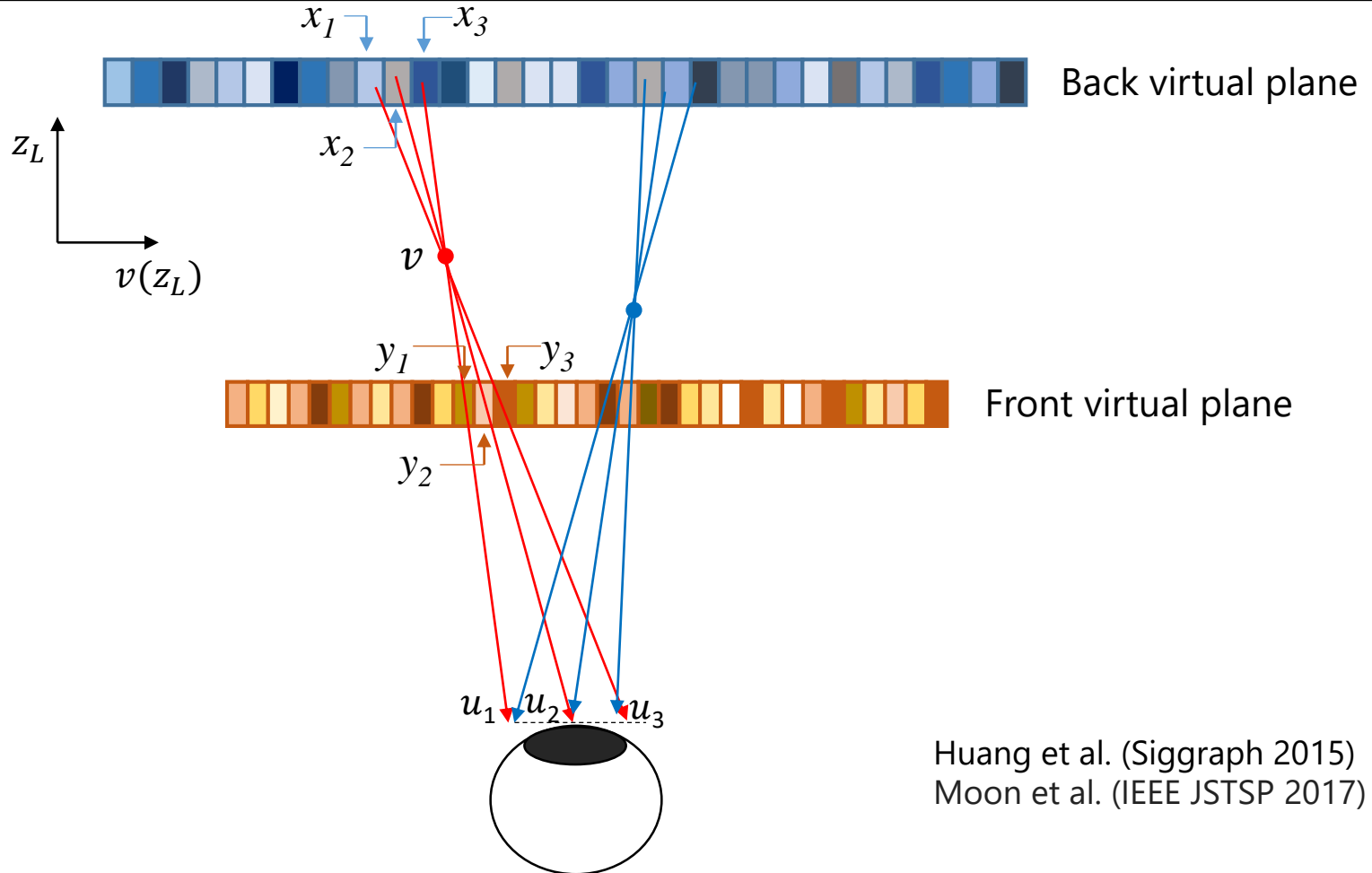
Tensor Displays

- **Lightfield emitted by a multi-layer display is represented by a tensor where rays span a 2D plane in 3D tensor space**
- **Target lightfield is decomposed into Rank-1 tensors using Nonnegative Tensor Factorization**
- **Rank-1 tensors are shown in quick succession with a high refresh rate, which are perceptually averaged over time by the Human Visual System**



Video adapted from Wetzstein, Gordon, et al. "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting." (2012).

Rendering images in Tensor Displays



Target Light-fields: $L(v, u_1) = L(v, u_2) = L(v, u_3) = R$

Optimization equation : $L(v, u_1) = x_3 \times y_1$

$L(v, u_2) = x_2 \times y_2$

$L(v, u_3) = x_1 \times y_3$

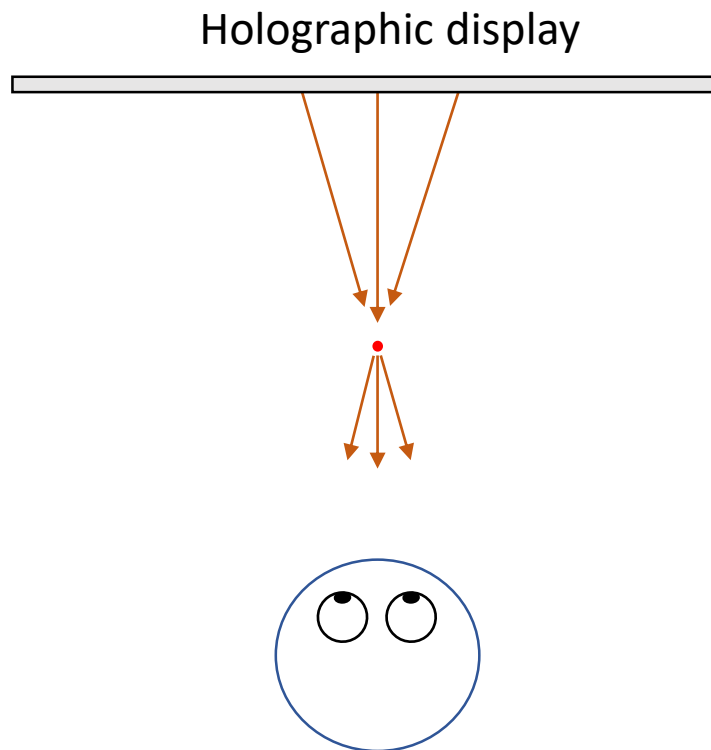
1
5
3

Lightfield Displays



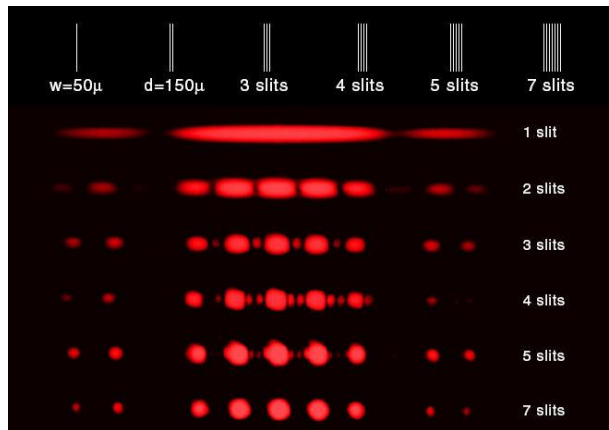
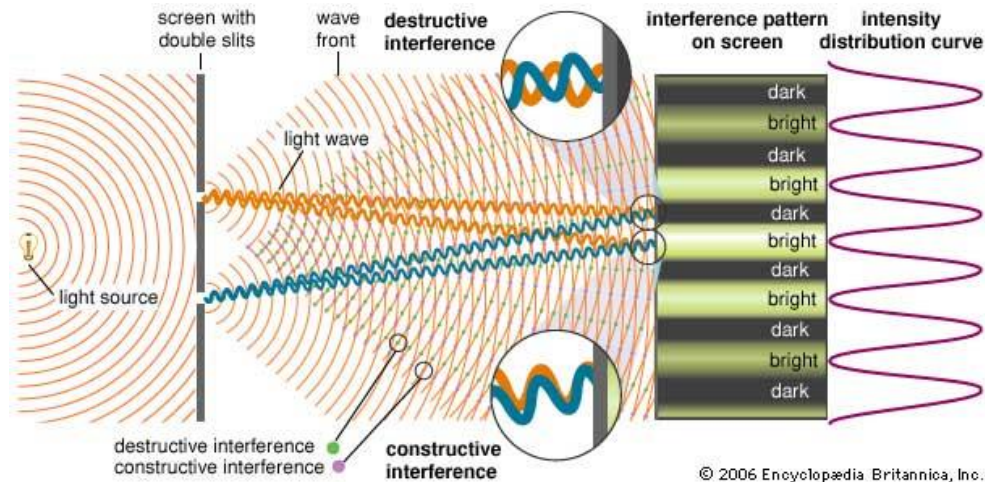
Holographic display

Holographic display : generating 3D images **in the air** without any scatterer



What is the meaning of “focusing the light”?

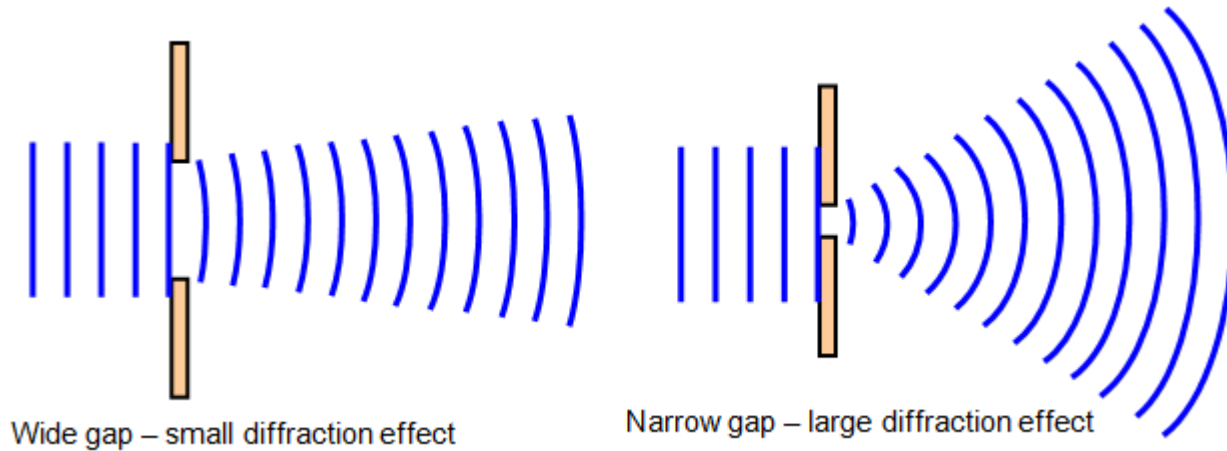
Focusing == interference



<http://labman.phys.utk.edu/phys136>

Focusing = constructive interference of multiple pixels
(but it requires coherent light sources such as laser)

Smaller pixel size == Large diffraction angle

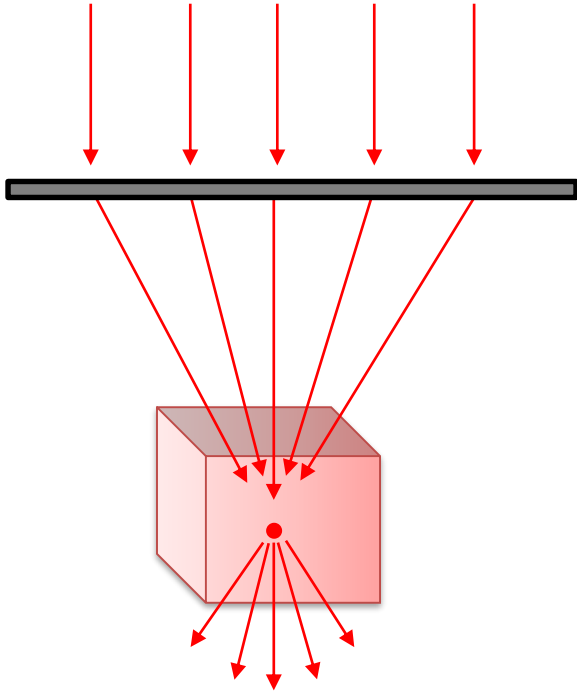


	Pixel size	Viewing angle
LCD monitor	200 μm	0.1°
LCoS Spatial light modulator	16 μm	2°
Ideal pixel size	1 μm	30°

http://www.schoolphysics.co.uk/age14-16/Wave%20properties/text/Diffraction_/index.html

Ultimate 3D display: Holographic display

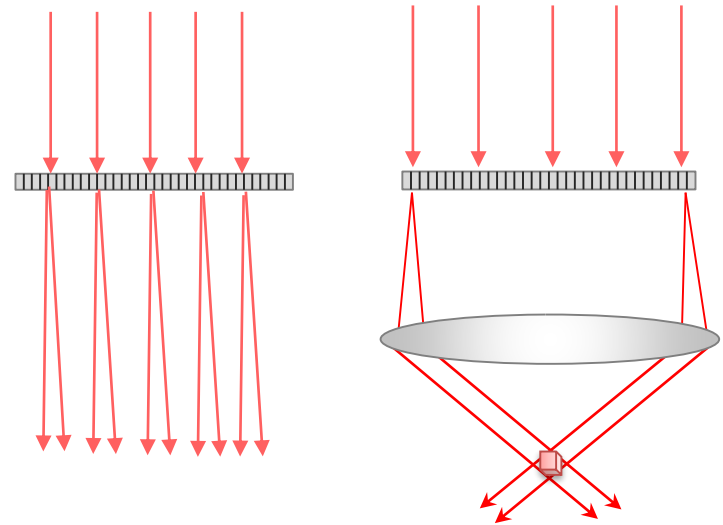
Ideal holographic monitor



Pixel size : $1 \mu m$
Screen size : 30 cm x 30 cm
Resolution : 300000 x 300000

Viewing angle : 30°
Image size : 30 cm x 30 cm

Current holographic monitor



Pixel size : $16 \mu m$
Screen size : 1 cm x 1 cm
Resolution : 1024 x 768

Viewing angle : 2°
Image size : 1 cm x 1 cm

Displays Comparison

	Pictorial Cues	Disparity	Motion Parallax	Glasses-free	Accommodation
2D Display	✓	✗	✗	✓	✗
Stereoscopic Display	✓	✓	✗	✗	✗
Head-mounted Display	✓	✓	✓	✗	✗
Autostereoscopic Display	✓	✓	✓	✓	✗
Light field Display	✓	✓	✓	✓	✓
Holographic Display	✓	✓	✓	✓	✓

References

- Geng, Jason. "Three-dimensional display technologies." *Advances in optics and photonics* 5.4 (2013): 456-535.
- Ives, Frederic E. "Parallax stereogram and process of making same." U.S. Patent No. 725,567. 14 Apr. 1903.
- Lippmann, Gabriel. "Epreuves reversibles donnant la sensation du relief." *J. Phys. Theor. Appl.* 7.1 (1908): 821-825.
- Wilburn, Bennett, et al. "High performance imaging using large camera arrays." *ACM Transactions on Graphics (TOG)*. Vol. 24. No. 3. ACM, 2005.
- Ng, Ren, et al. "Light field photography with a hand-held plenoptic camera." *Computer Science Technical Report CSTR 2.11* (2005): 1-11.
- Coltheart, M. "The persistences of vision." *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 290.1038 (1980): 57-69.
- Akeley, Kurt, et al. "A stereo display prototype with multiple focal distances." *ACM transactions on graphics (TOG)*. Vol. 23. No. 3. ACM, 2004.
- Sorensen, Svend Erik Borre, Per Skafte Hansen, and Nils Lykke Sorensen. "Method for recording and viewing stereoscopic images in color using multichrome filters." U.S. Patent No. 6,687,003. 3 Feb. 2004.
- Love, Gordon D., et al. "High-speed switchable lens enables the development of a volumetric stereoscopic display." *Optics express* 17.18 (2009): 15716-15725.
- Wetzstein, Gordon, et al. "Layered 3D: tomographic image synthesis for attenuation-based light field and high dynamic range displays." *ACM Transactions on Graphics (ToG)*. Vol. 30. No. 4. ACM, 2011.
- Wetzstein, Gordon, et al. "Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting." (2012).
- Narain, Rahul, et al. "Optimal presentation of imagery with focus cues on multi-plane displays." *ACM Transactions on Graphics (TOG)* 34.4 (2015): 59.
- Maimone, Andrew, et al. "Holographic near-eye displays for virtual and augmented reality". *ACM Transactions on Graphics* 36, 4 (2017),
- Xiao et al, "DeepFocus : Learned Image Synthesis for Accommodation-Supporting Displays". *ACM Transactions on Graphics (TOG)* (2018)
- Lanman, D. and Luebke, "Near-eye light field displays". *ACM Transactions on Graphics* 32, 6, 1–10. (2013)
- Cholewiak et al, "ChromaBlur: Rendering Chromatic Eye Aberration Improves Accommodation and Realism in HMDs". *ACM Transactions on Graphics*, (2017)
- Yu et al, "A Perception-driven Hybrid Decomposition for Multi-layer Accommodative Displays" *IEEE Transactions on Visualization and Computer Graphics* (2019)