SAN ANTONIO

SIGGRAPH 2002

Visual Perception and LOD

Martin Reddy

Contents

- Why Care About Perceptual Issues?
- Overview of LOD Selection Criteria
- Example Perceptual LOD Systems
- Perceptual Theory
- Applying the Perceptual Theory
- Conclusions & Future Research

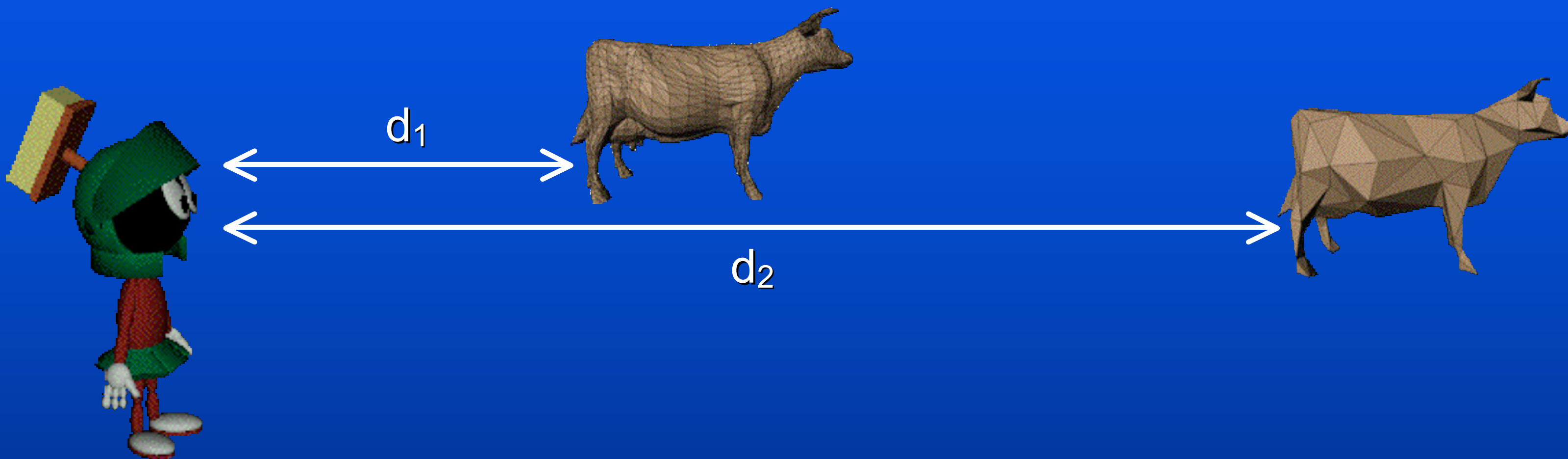
Introduction

- **What are we going to talk about?**
  - Guiding LOD selection with reference to models of visual perception
- **Why do we care about visual perception?**
  - Optimize computational resources
  - Minimize LOD popping effects
  - Develop a principled scheme for selecting LOD
- **Caveats!**
  - Visual perception is hard! Let’s do what we can.

Summary of Level of Detail

- **Primary LOD selection criteria**
  - Distance or Size
  - Velocity
  - Eccentricity
  - Depth of Field
- **Additional LOD constraints**
  - Fixed-frame rate schedulers (reactive or predictive)
  - Hysteresis (switching lag)
  - Priority schemes
  - Alpha-blended transitions (fading regions)
  - Geomorph transitions (morph geometry)

Distance LOD



Select resolution based upon the distance between an element and the viewpoint, i.e. coarser resolution for distant geometry.

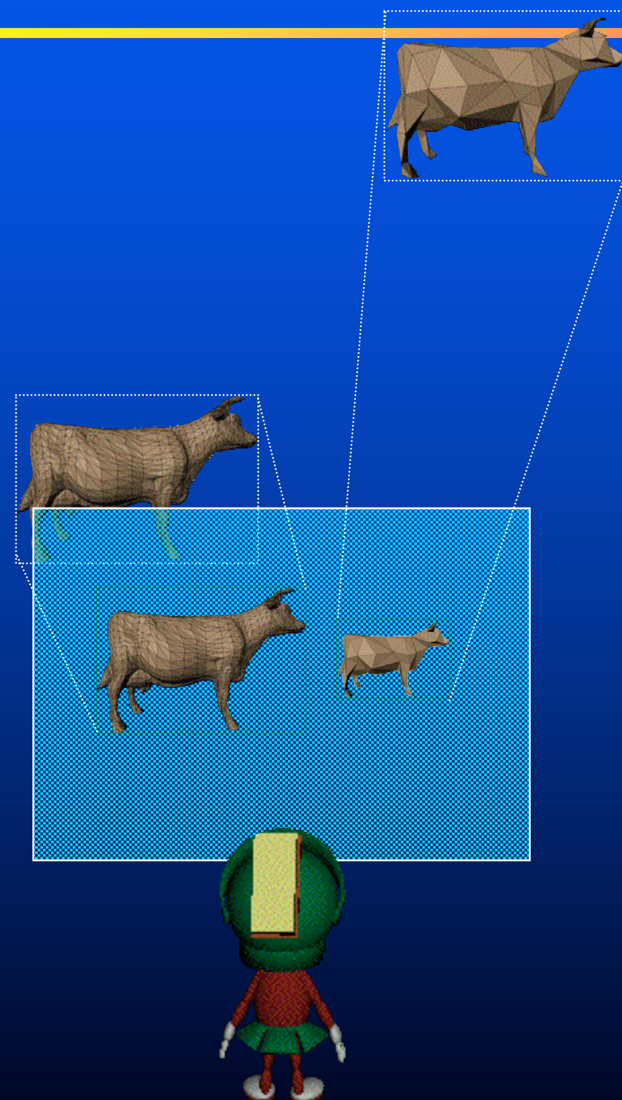
- Simple to calculate (3-D Euclidean distance)
- Scale dependent
- Resolution dependent
- Field of View dependent

Size LOD

Select resolution based upon the projected screen size (or area) of an element. Objects appear smaller as they move further away.

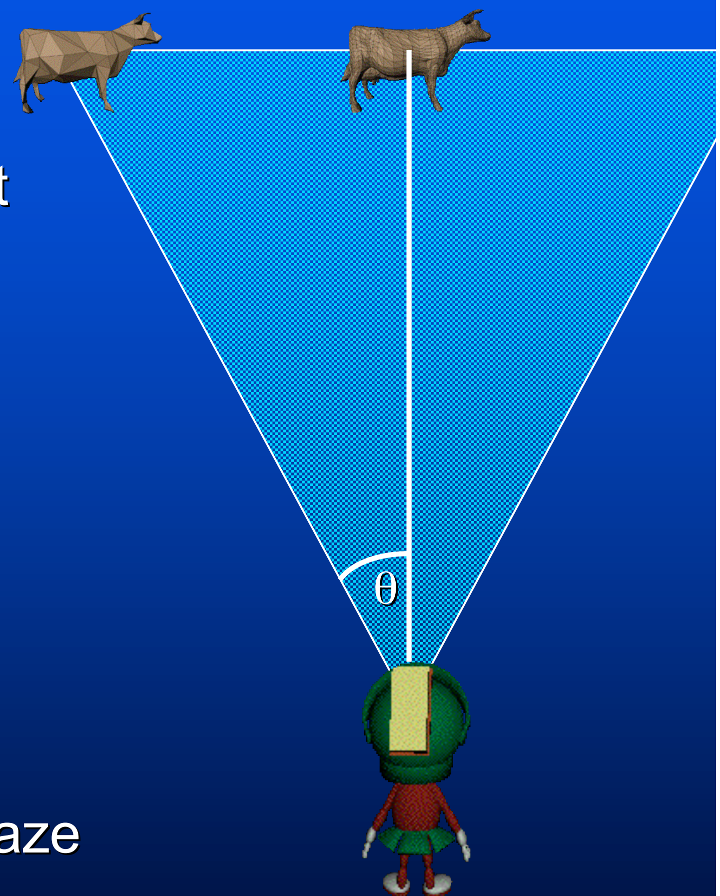
- Requires 3-D → 2-D projection
- Scale invariant
- Resolution invariant
- Field of View invariant

Bounding spheres or ellipsoids normally used instead of boxes as more efficient to calculate projected extent



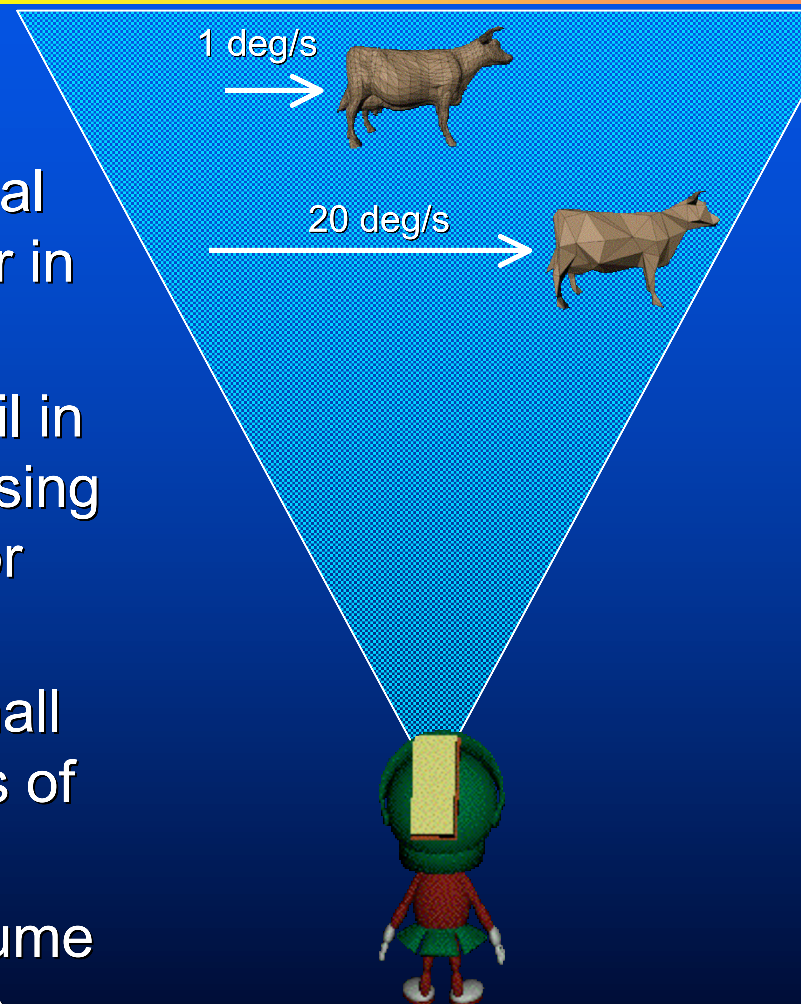
Eccentricity LOD

– Resolution is selected based upon the degree to which an element exists in the visual periphery, i.e. display elements that the user is looking at in high resolution.
– Humans can resolve less detail in their peripheral field due to:
  - more retinal photoreceptors (rods/cones) towards fovea
  - retinal and cortical cell receptive field sizes increases linearly with eccentricity
  - 80% of cortical cells devoted to central 10 degrees of vision
– Use eye tracking system to track user's gaze or assume user looking towards center of display

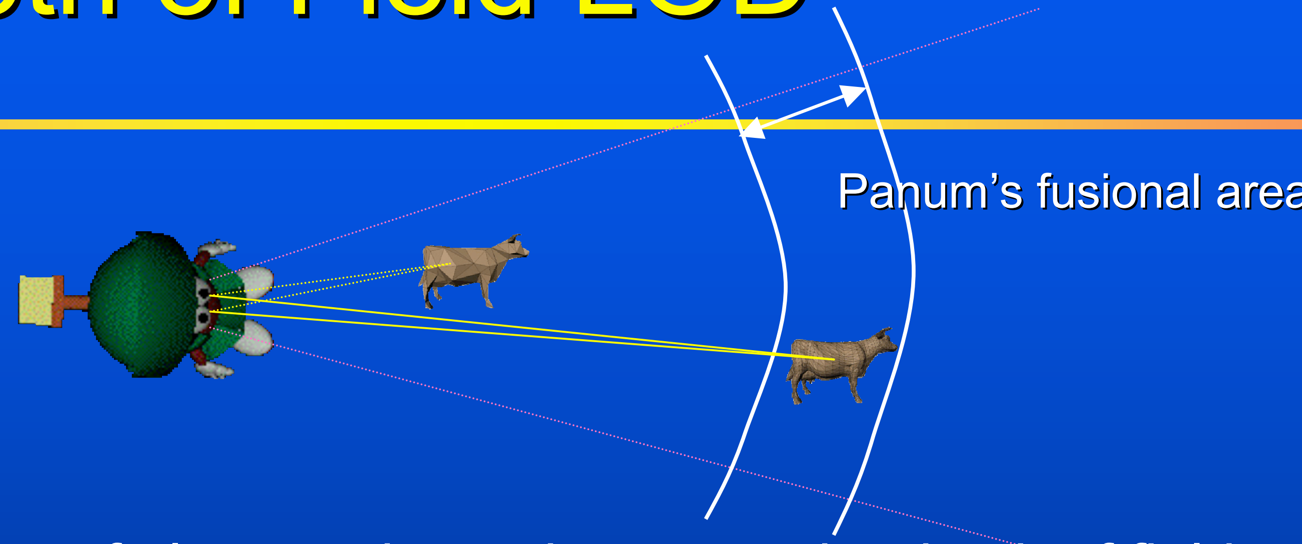


Velocity LOD

- Resolution based upon the angular velocity of an element across the visual field, i.e. faster moving objects appear in lower resolution
- Humans can resolve less spatial detail in objects moving across the retina, causing objects to blur as they move/ rotate, or the user's gaze moves
- It is believed visual information for small features are destroyed by the process of integrating stimulus energy over time
- Without eye tracking technology, assume angular velocity across display device



Depth of Field LOD



- Resolution of element dependent upon the depth of field focus of the user’s eyes, i.e. objects out with the fusional area appear in lower detail
- Under binocular vision, both eyes converge on object at certain distance in order to focus retinal image
- Objects in front or behind this fusional area are unfocused, suffering from double images
- Must track both eyes accurately to evaluate convergence distance

Funkhouser & Séquin

SIGGRAPH 1993 - Visualize Complex Virtual Envs.

- Achieve (predictive) fixed frame rate by

  Maximise $\sum_{s} \text{Benefit}(\,\textit{Object}, \textit{Lod}, \textit{Algorithm}\,)$

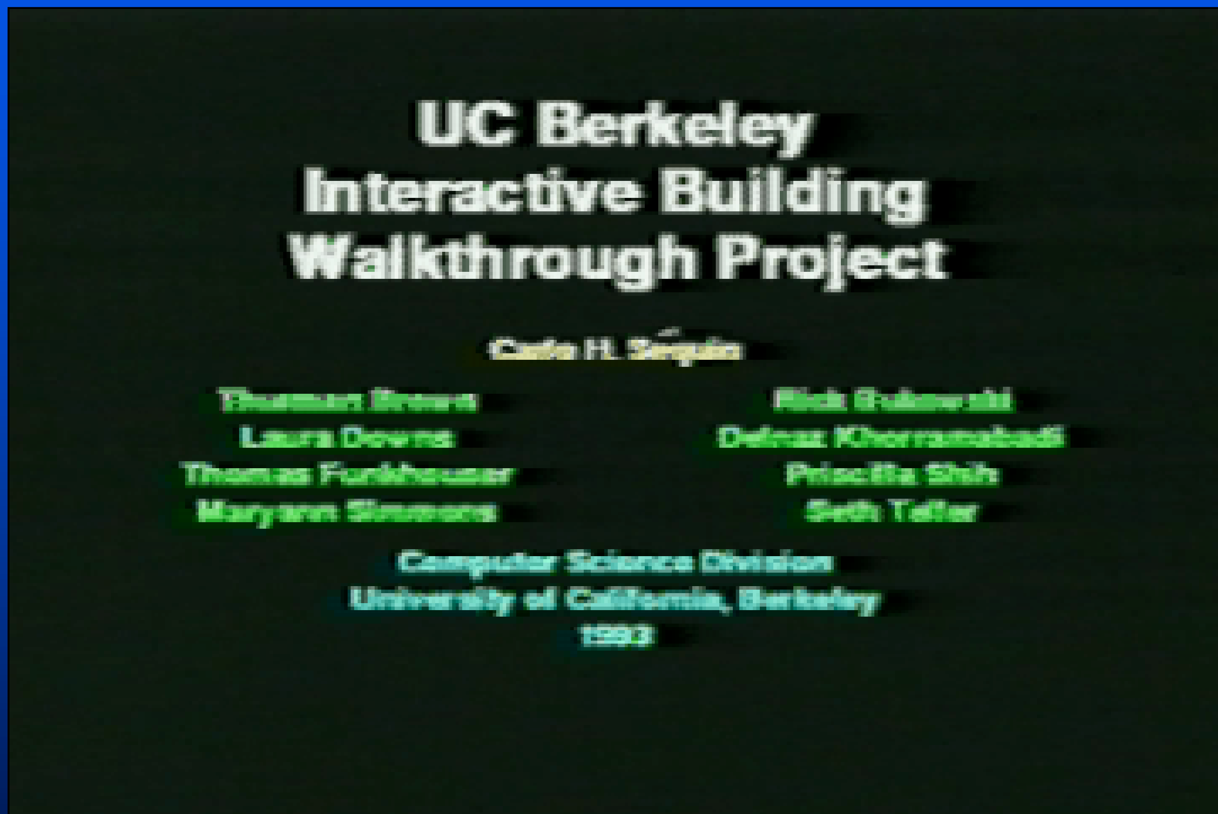
  Subject to $\sum_{s} \text{Cost}(\,\textit{Object}, \textit{Lod}, \textit{Algorithm}\,) \leq \text{TargetFrameRate}$

- Benefit = contribution to model perception:
  - Size: larger objects contribute more to image
  - Accuracy: no of verts/polys, shading model, etc.
  - Priority: account for inherent importance
  - Eccentricity: based on distance from center of display
  - Velocity: ratio of apparent speed to average polygon size
  - Hysteresis: use state from previous frame

- No head/eye tracking.

  No results on perceptual criteria.

Funkhouser & Séquin



"Adaptive Display Algorithm"

Funkhouser and Séquin

(University of California, Berkeley)

Hitchner & McGreevy

SPIE (1993) - NASA VPE Testbed

- Achieve (reactive) fixed frame rate by:

  detail = geometry primitives per unit area

  interest = importance to the user

  – Eccentricity factor: $Fn_E = \gamma_{static}$ / eccentricity

  – Velocity factor: $Fn_V = \gamma_{dynamic}$ / velocity

  – Distance factor: $Fn_D = \beta$ / velocity

  (where γ_{static}, $\gamma_{dynamic}$, and β are arbitrary scaling factors)

- Used a head-mounted display with 6 degree-of-freedom head tracker.

Ohshima et al.

VRAIS (1996) - Gaze-Directed Rendering

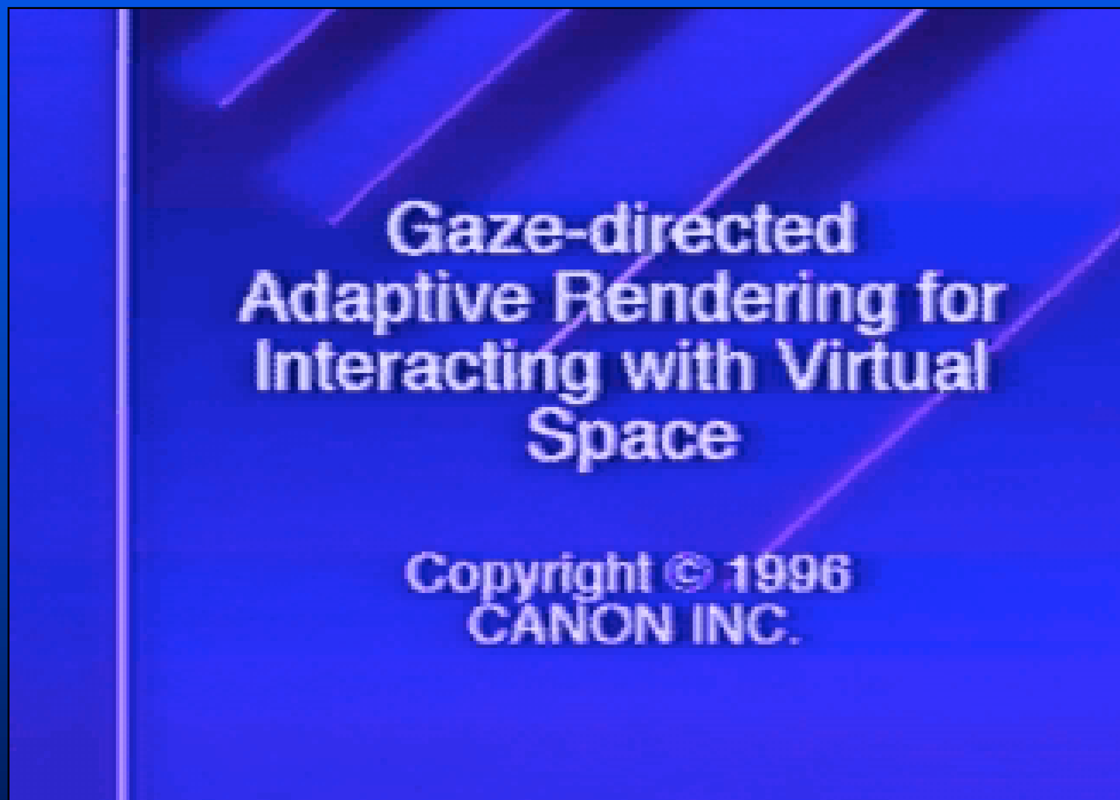
- Used 3 criteria to evaluate visual acuity:
  - eccentricity: $g(\Delta\phi) = 1 - \Delta\phi/c_2$ $(0 \le \Delta\phi \le c_2$; else 0)
  - velocity: $f(\theta) = \exp(-(\theta-\alpha)/c_1)$ $(\alpha < \theta$; else 1)
  - depth of field: $h(\Delta\varphi) = \exp(-(\Delta\varphi-b)/c_3)$ $(b < \Delta\varphi$; else 1)

    arbitrary scaling factors set as follows:

    c_1 = 6.2 deg, c_2 = 180 deg, c_3 = 0.62 deg, b = 0 deg.

    α = visual angle occupied by object
- Saccadic suppression:
  - skip rendering when gaze velocity > 180 deg/s
- Used head tracker as a substitute for eye tracker. Used 60 deg projection screen.

Example: Ohshima et al.



"Gaze-directed Adaptive Rendering"

Ohshima, Yamamoto, and Tamura

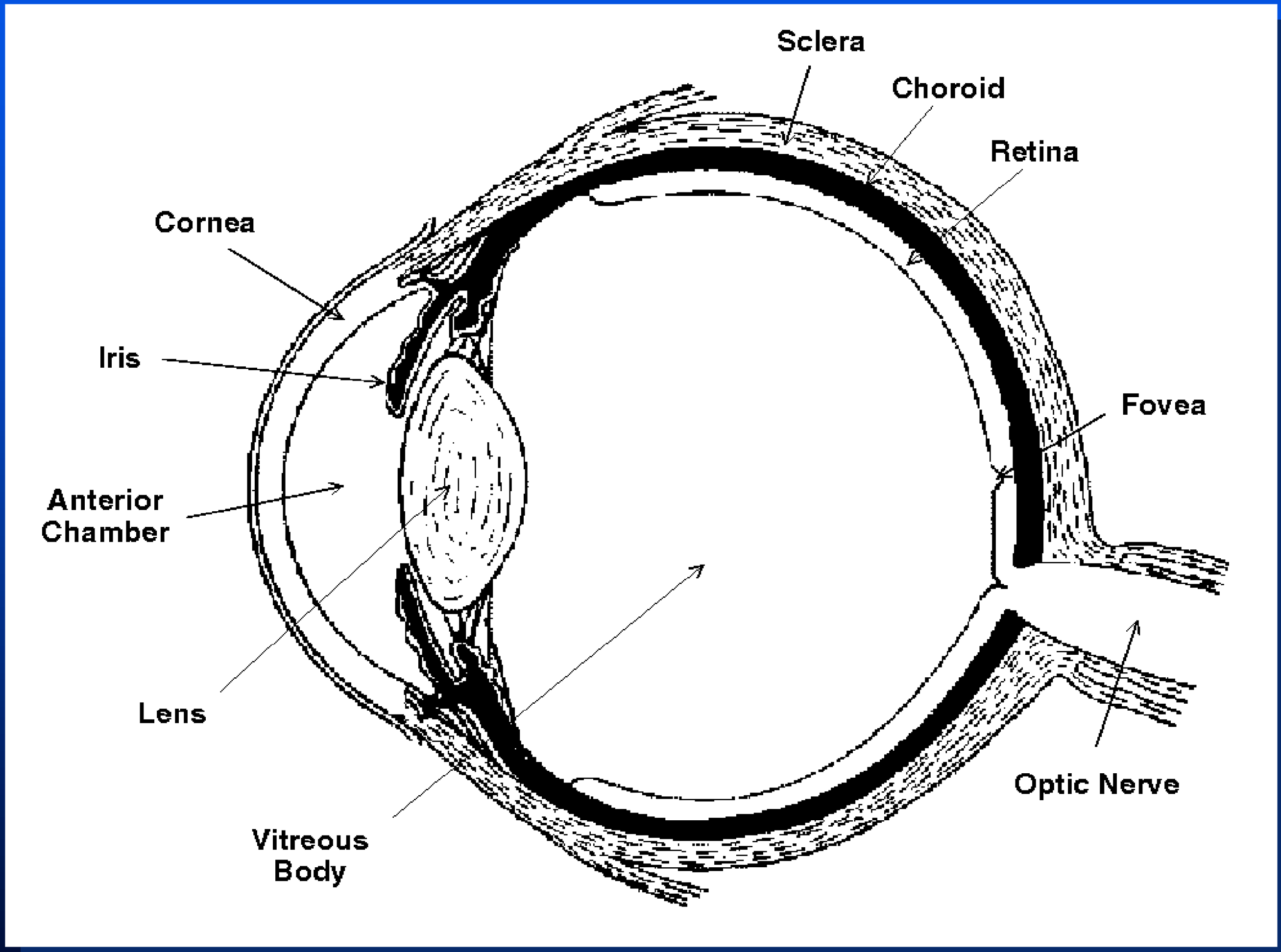
(Canon Inc.)

Example: Lindstrom & Turk

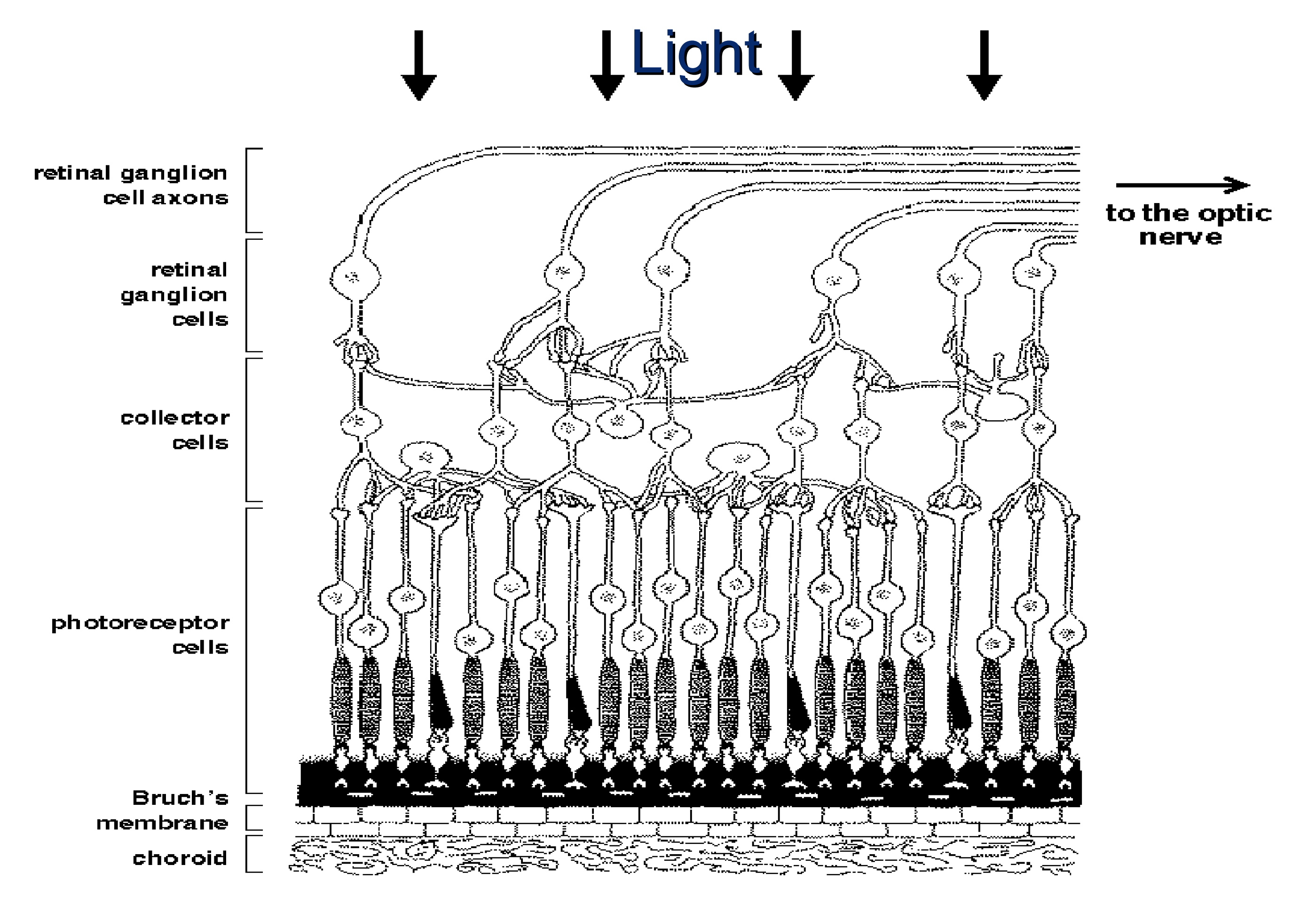
ACM Trans. (2000) - Image-Driven Simplification

- Render each LOD off-screen and analyze images to decide which parts of the model to simplify
  - Guides simplification based upon the visual effect of the reduction rather than some geometric metric
  - Uses a sphere of cameras to capture multiple viewpoints
  - Deals with surface properties and textures
  - Uses RMS error of luminances to compute image distances (fast but not perceptually based)
  - Not real-time (several secs to mins or hours)

The Eye


Sclera
Choroid
Retina
Cornea
Iris
Fovea
Anterior Chamber
Lens
Optic Nerve
Vitreous Body

The Retina


Light
retinal ganglion cell axons
retinal ganglion cells
collector cells
photoreceptor cells
Bruch's membrane
choroid
to the optic nerve

The Limits of Vision

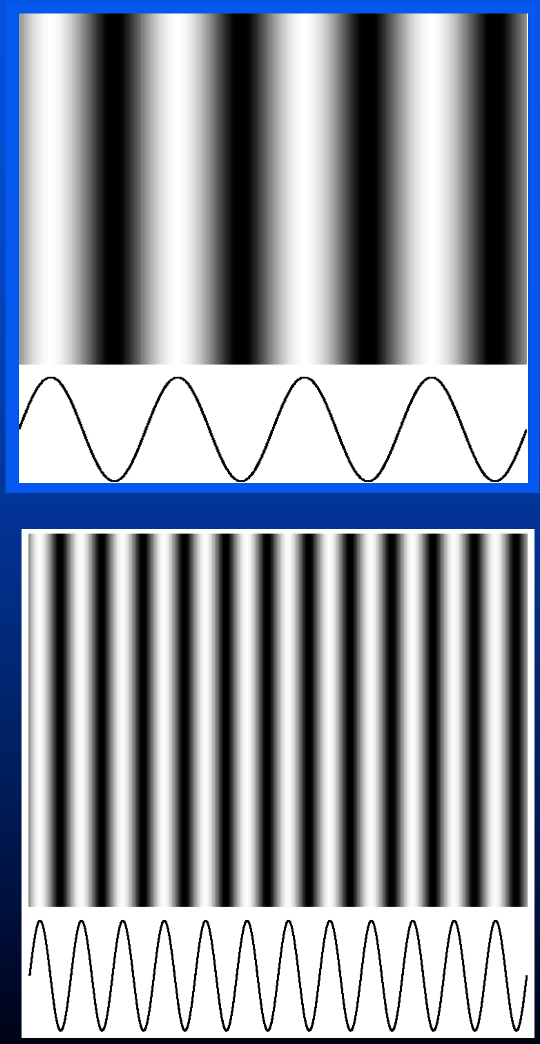
- **Visual acuity**
  - Retina can resolve detail of around 0.5 min of arc
  - 130 million photoreceptors / 1 million ganglion cells
- **Peripheral Vision**
  - Highest sensitivity to spatial detail at fovea (the central 4 to 5 degrees of vision)
  - 35-fold reduction from fovea → periphery

- **Motion Sensitivity**
  - Eye less sensitive to detail moving across retina
  - Fast moving objects become “blurred”

Measuring Limits of Vision

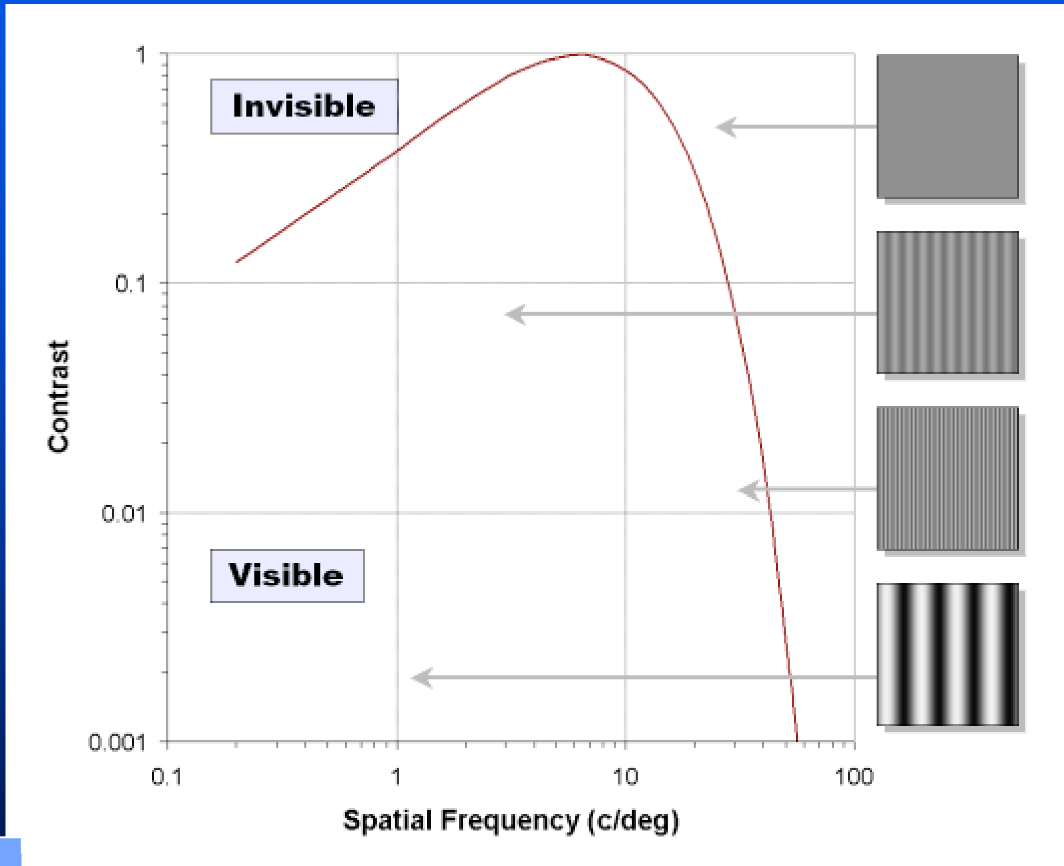
“Contrast Grating” used to analyze contrast sensitivity . Can vary:

- Spatial frequency (bar spacing) - cycles per deg (c/deg)
- Contrast (amplitude)
- Orientation
- Velocity
- Eccentricity



Modeling Limits of Vision

- Results of Contrast Grating tests can be modeled with a Contrast Sensitivity Function
- CSF defines the bandwidth of vision



$A(\alpha) = 2.6\,(0.0192 + 0.133\,\alpha\,) \times e^{-(0.144\,\alpha)^{1.1}}$

where,

α = spatial frequency (c/deg)

Velocity CSF

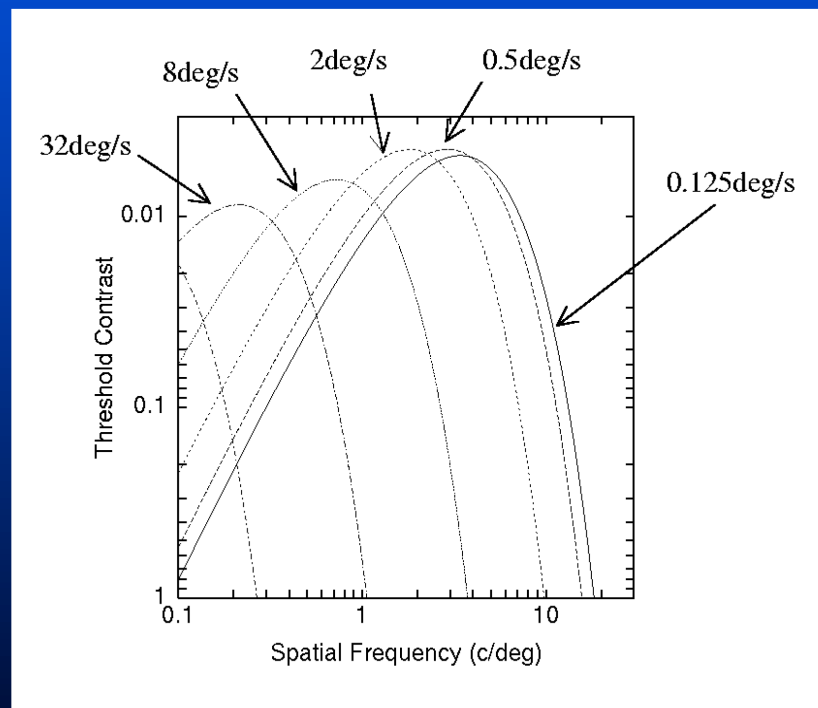
- Kelly (1979) developed an equation to predict the CSF for various stimulus velocities

$$G(\alpha,v) = [250.1+299.3\ |\log_{10}(v/3)|^{3}] \times v\alpha^{2}10^{-5.5\alpha(v+2)/45.9}$$

where,

α = spatial frequency (c/deg)

v = velocity (deg/s)



Eccentricity CSF

- Rovamo & Virsu (1979) developed equations to model the decline of sensitivity with eccentricity for the 4 principal half-meridians of the retina

Nasal: $Mn(E) = 1 / (1+0.33E+0.00007E^3)$

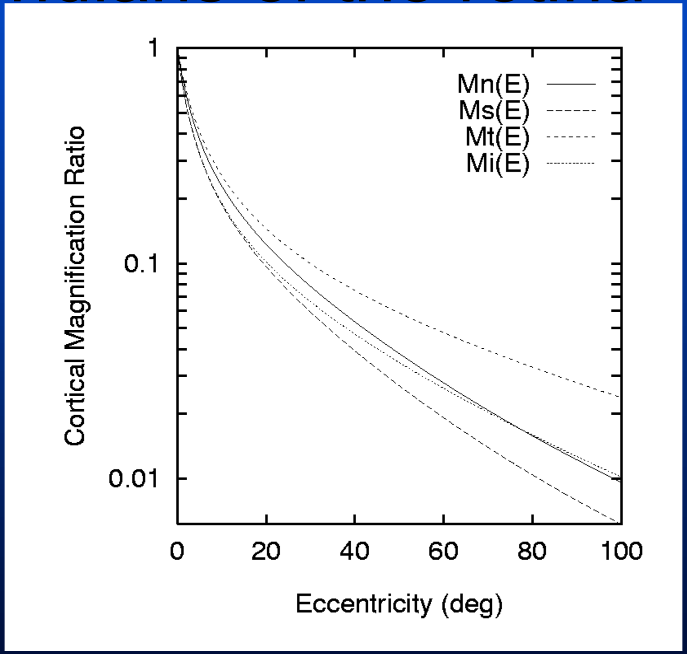
Superior: $Ms(E) = 1 / (1+0.42E+0.00012E^3)$

Temporal: $Mt(E) = 1 / (1+0.29E+0.000012E^3)$

Inferior: $Mi(E) = 1 / (1+0.42E+0.000055E^3)$

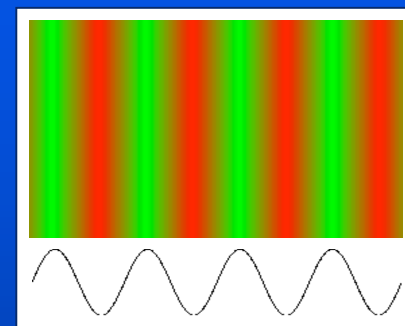
where,

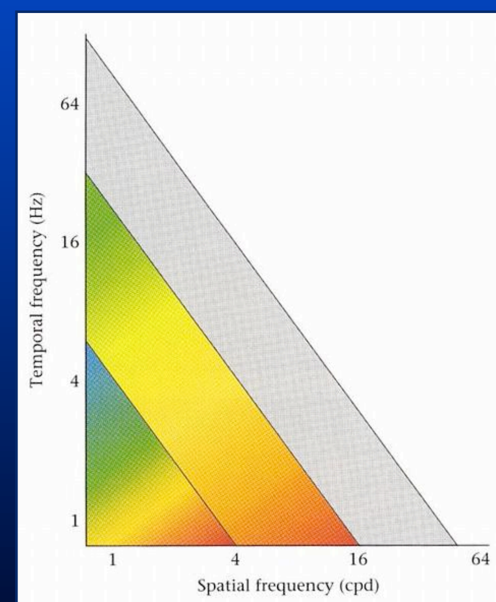
E = eccentricity (deg)



The CSF and 3D Graphics

- Color v Greyscale Gratings
  - Luminance channel more effective than chromatic channels for
    - Form detection
    - Motion
    - Stereoscopic depth
  - For example
    - Luminance upper resolution = 60 c/deg
    - Red/Green upper resolution = 12 c/deg
  - Therefore, use (simpler) achromatic CSF threshold models rather than chromatic ones





(Wandel, 1995)

Other CSF Factors

- Background illumination
  - Contrast sensitivity degrades in dim conditions
- Display Device Settings
  - Brightness, contrast, color, and gamma
- Viewer's level of light adaption
  - Photoreceptor range and pupil dilation controlled by a feedback loop
- Viewer's visual system efficiency
  - e.g., myopia causes light to converge in front of retina
- Viewer's age
  - Contrast sensitivity less developed in infants & declines with old age

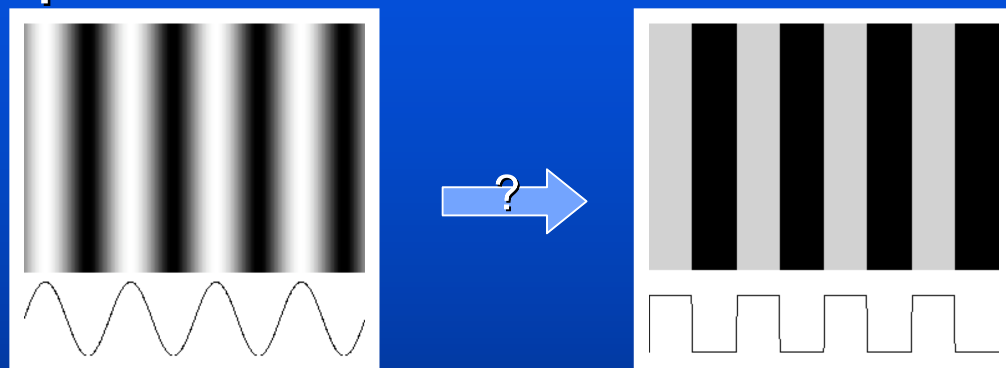
Other CSF Factors (cont.)

- Viewer's emotional state
  - Affects dilation of pupils: smaller pupil = less light = drop in visual acuity
- Auditory Stimuli?
  - Recent Nature paper shows visual perception affected by a adding an audible beep during task

Therefore, perceptual data are normally based upon a "Standard Observer", assuming ideal environmental and viewer conditions.

The CSF and 3D Graphics

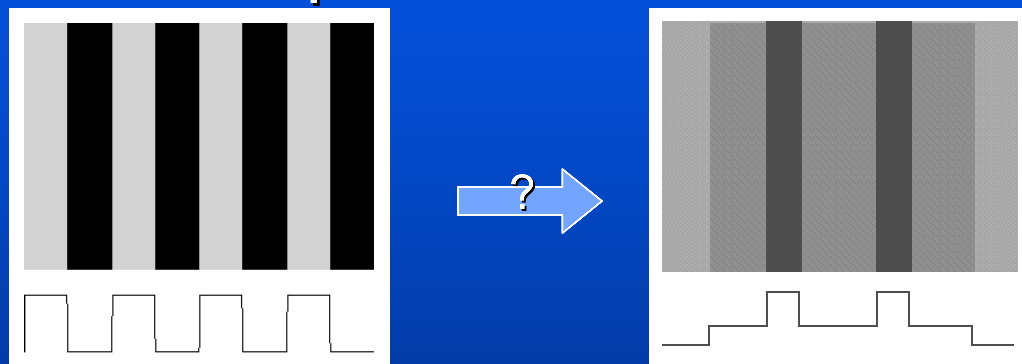
- Sine v Square Waveform



  – Above the peak frequency, the amplitude of square wave CSF is largely determined by the fundamental sine wave
  – The limit of vision is the same in each case

The CSF and 3D Graphics

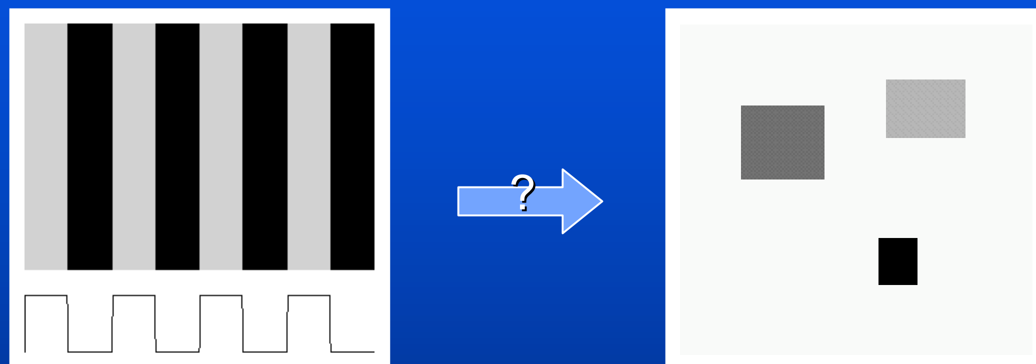
- Harmonic v Complex Waveform



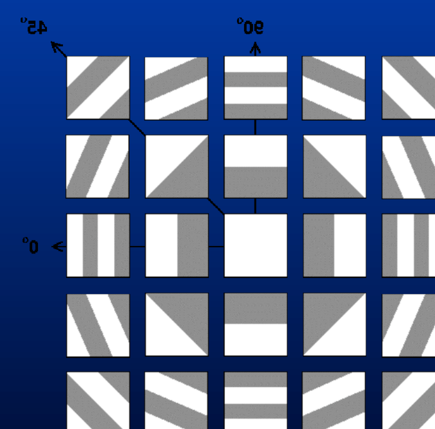
  – Visibility of complex grating is characterized by the independent contributions from each harmonic component
  – below-threshold high-frequency components can be removed without perceivable change

The CSF and 3D Graphics

- 2-D v 1-D Waveform



  - – Introduce an orientation parameter to describe 2-D features, e.g. (2 c/deg, 90 deg)



Visual Perception Software

Constable

"Salisbury Cathedral from the Bishops Gardens" (1826)

100 x 80 degrees FOV

No eccentricity blurring

No velocity blurring

Visual Perception Software

Constable

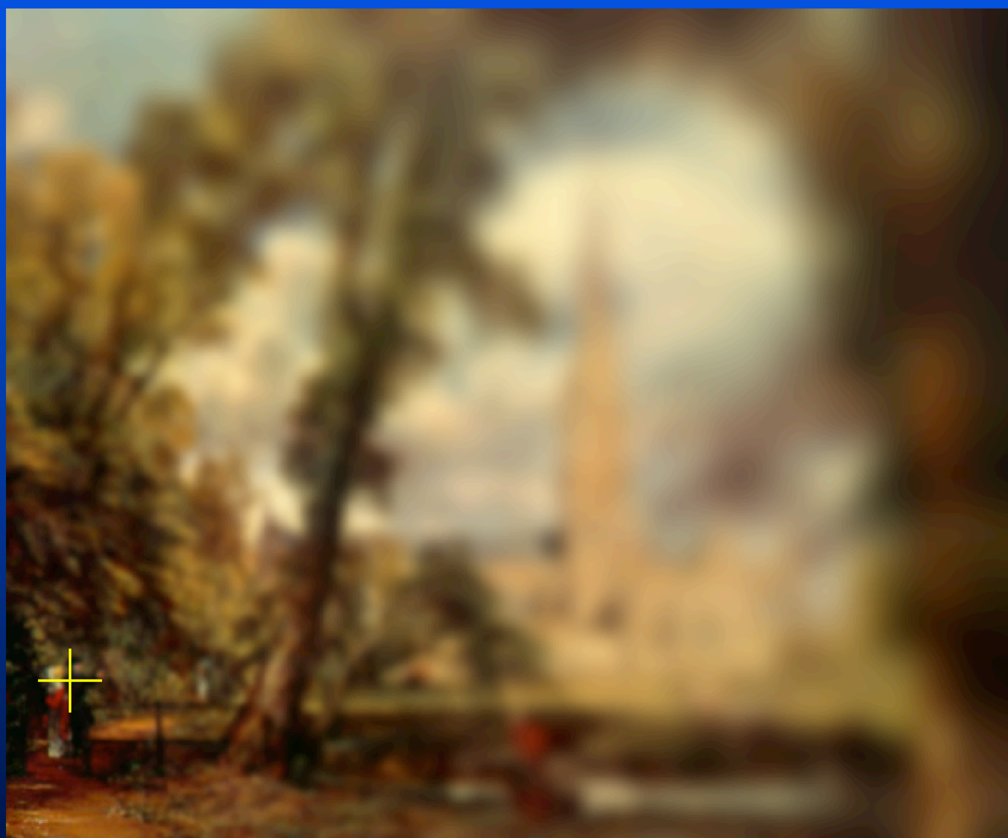
"Salisbury Cathedral from the Bishops Gardens" (1826)

100 x 80 degrees FOV

Eccentricity blurring

No velocity blurring

Visual Perception Software



Constable

"Salisbury Cathedral from the Bishops Gardens" (1826)

100 x 80 degrees FOV

Eccentricity blurring

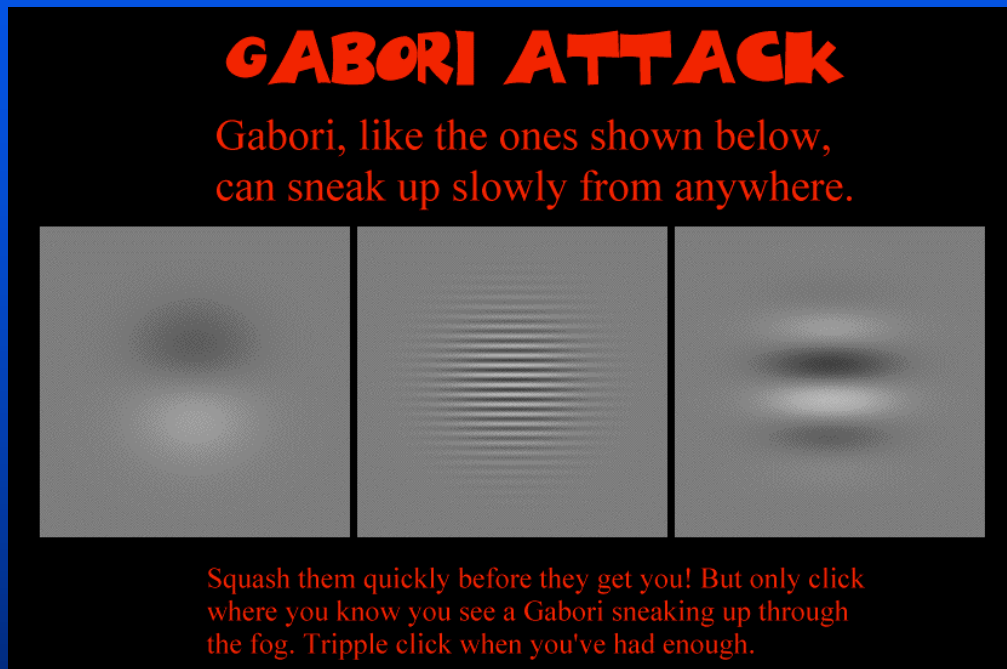
Velocity = 50 deg/s

Visual Perception Software

PERCEPT

- Visualize the limits of vision in an intuitive manner
- Calculate highest spatial frequency of each pixel given its eccentricity and a constant user-specified velocity. Then blur this pixel using an equivalently-sized Gaussian filter
- Based on Kelly, Rovamo & Virsu models
- Open Source, C code (GPL)
- **http://www.ai.sri.com/~reddy/percept/**

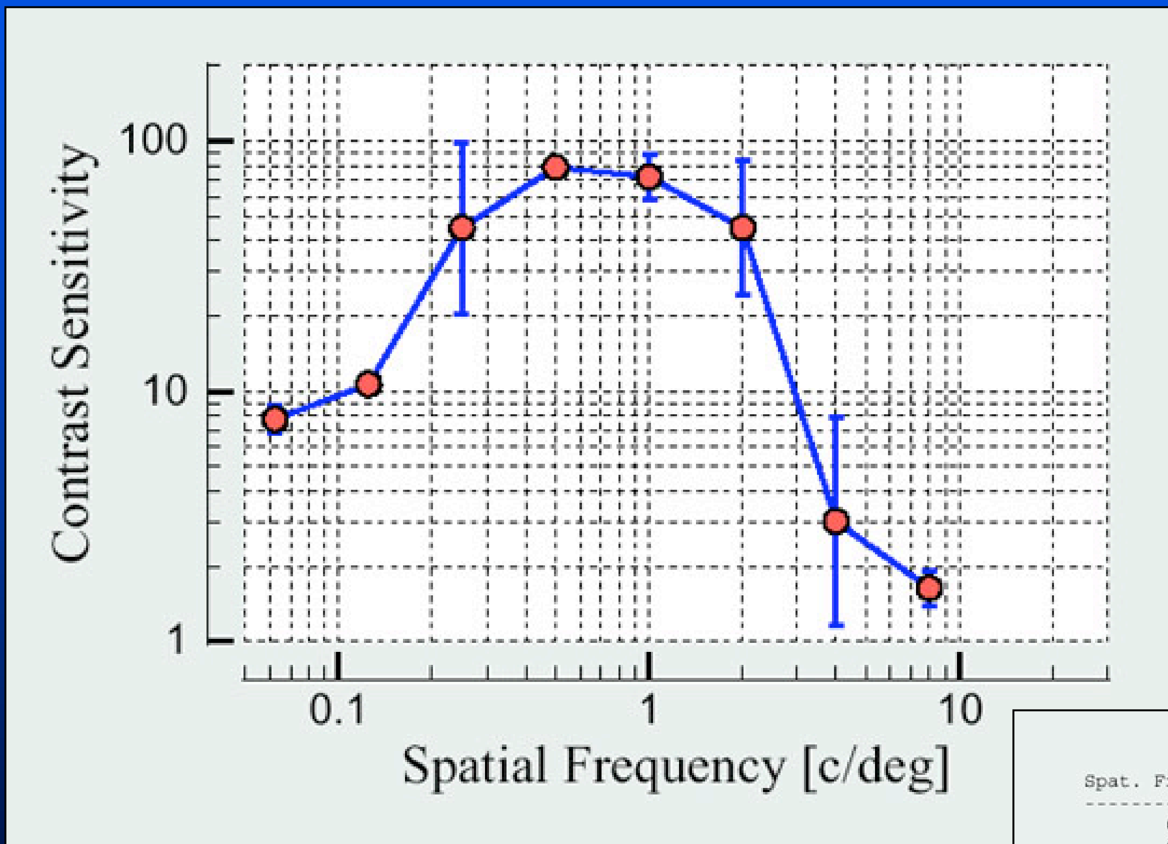
Visual Perception Software



A program for testing contrast sensitivity functions that is presented as a simple video game. From Berkeley VSOC.

http://vsoc.berkeley.edu/vsoc/

Visual Perception Software

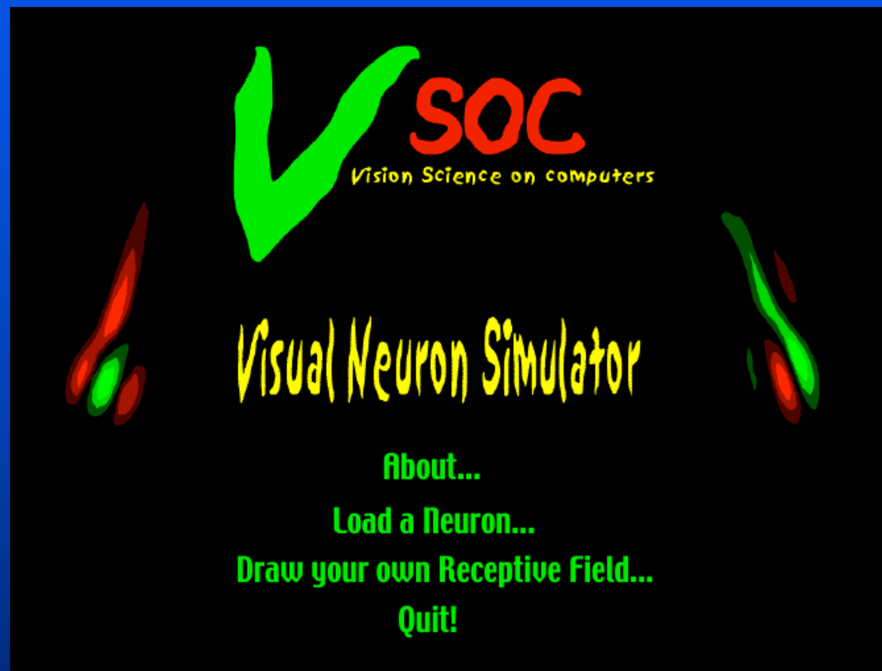


Example results from a Gabori Attack session

Spat. Freq.[c/deg]	Sensitivity	Std.Err.[dB]	N
0.063	7.74	1.081	2
0.125	10.69	0.721	2
0.250	45.03	6.847	2
0.500	78.63	0.011	2
1.000	72.00	1.802	2
2.000	45.03	5.406	2
4.000	3.03	8.289	2
8.000	1.64	1.442	2

Contrast Calib=1.390

Visual Perception Software



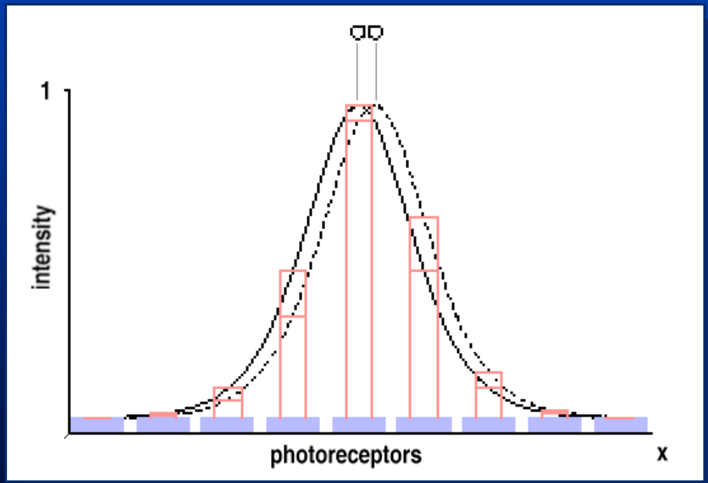
A real-time visual neuron simulator. Explore the receptive fields of artificial neurons. From Berkeley VSOC.

http://vsoc.berkeley.edu/vsoc/

What about hyperacuity?

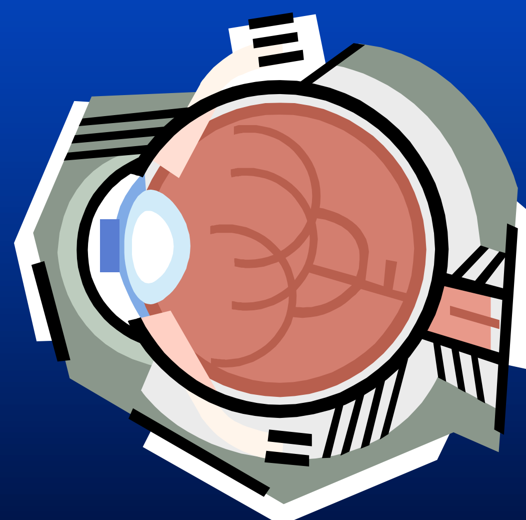
- Hyperacuity describes the paradox that certain stimuli can be perceived that are smaller than the size of a single photo-receptor cell
- Photoreceptors subtend 25-30 sec of arc (= 60 c/deg)
- But it is possible to discriminate the non co-linearity of two thick abutting lines to a resolution of 2-5 sec of arc (Vernier Acuity)

Due to differences in mean distribution of light sampled over a number of photoreceptors
Degrades markedly with ecc.
We are concerned with detection not discrimination though



What about saccades?

- A saccade is a rapid reflex movement of the eye to fixate a target onto the fovea (from French *saccader*, “to jerk”)
- We do not appear to perceive detail during a saccade
- Saccades can occur at velocities of up to 800 deg/s
- Duration can be many milliseconds:
  - duration (ms) = 20 + angularDist * 2
  - e.g. 10 deg saccade lasts ~40 ms

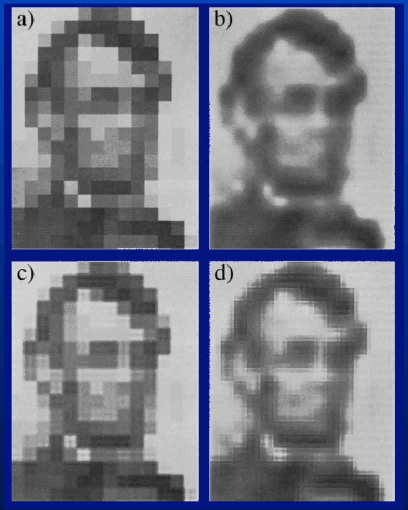


Ohshima et al.’s (1996) system suspended rendering at gaze velocities > 180 deg/s

What about visual masking?

The presence of one visual pattern can affect the visibility of another pattern, e.g. a large adjacent stimulus (in time or space) can cause the threshold of a smaller stimulus to be increased - the smaller stimulus needs to be more intense for it to be visible

- Also, the detection threshold of a stimulus varies inversely as a function of its distance from an edge (larger errors can be tolerated around an edge)
- Ferwerda et al. (1997) developed a visual masking model for computer graphics

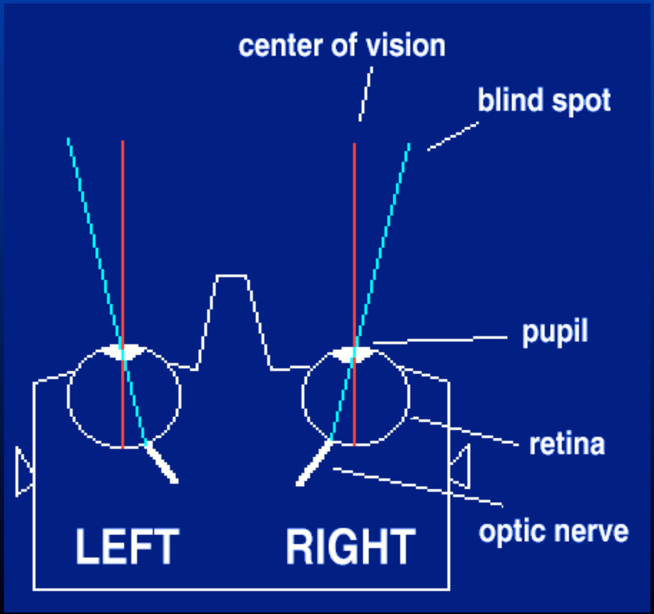


Harmon & Julesz (1973)

What about the blind spot?

– The blind spot is caused by the area of the retina where all axons of the retinal ganglion cells meet to form the optic nerve. There are no photoreceptors in this region.
– Can we reduce detail if an object falls onto the blind spot?
– Size of blind spot = 5 - 7 deg
– Located at ~17 deg eccentricity

We have 2 eyes! At least one eye will always detect the object, so don't bother with the blind spot



Applying Perceptual Metrics

- **Monitor object statistics**

  Calculate the projected size (deg), velocity (deg/s), and eccentricity (deg) of any part of an object. (Relative to the display, the user’s head, or the user’s gaze.)

- **Measure perceived detail in imagery**

  Describe the perceived spatial detail of any part of an object in terms of its spatial frequencies (c/deg)

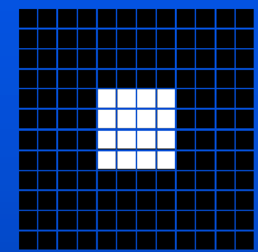
- **Model user’s visual acuity**

  Use mathematical model to estimate the contrast sensitivity of the user under various conditions (velocity, eccentricity, etc.)

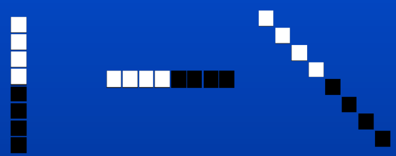
Perceptual LOD Systems

- Reddy (1997, 2001)
  - Discrete and view-dependent LOD Systems
  - Calculate spatial frequency profiles for each LOD off-line using a sphere of cameras
  - Ignore contrast (assume worst-case scenario)
- Luebke (2000, 2002)
  - View-dependent LOD System
  - Calculate spatial frequency induced by folding a node in real-time
  - Incorporate contrast and silhouette conditions
  - Use to control appearance preservation

Calculating c/deg - Reddy

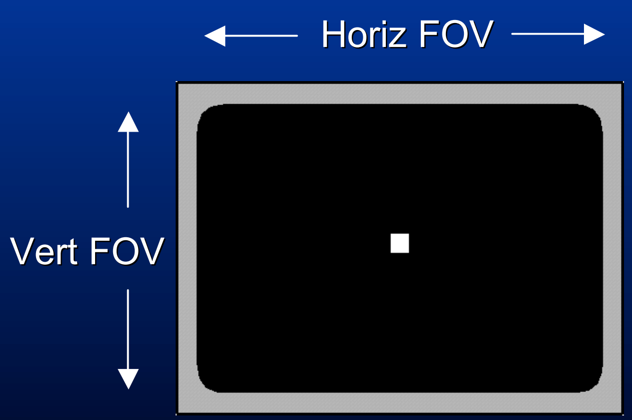


Isolate visual feature



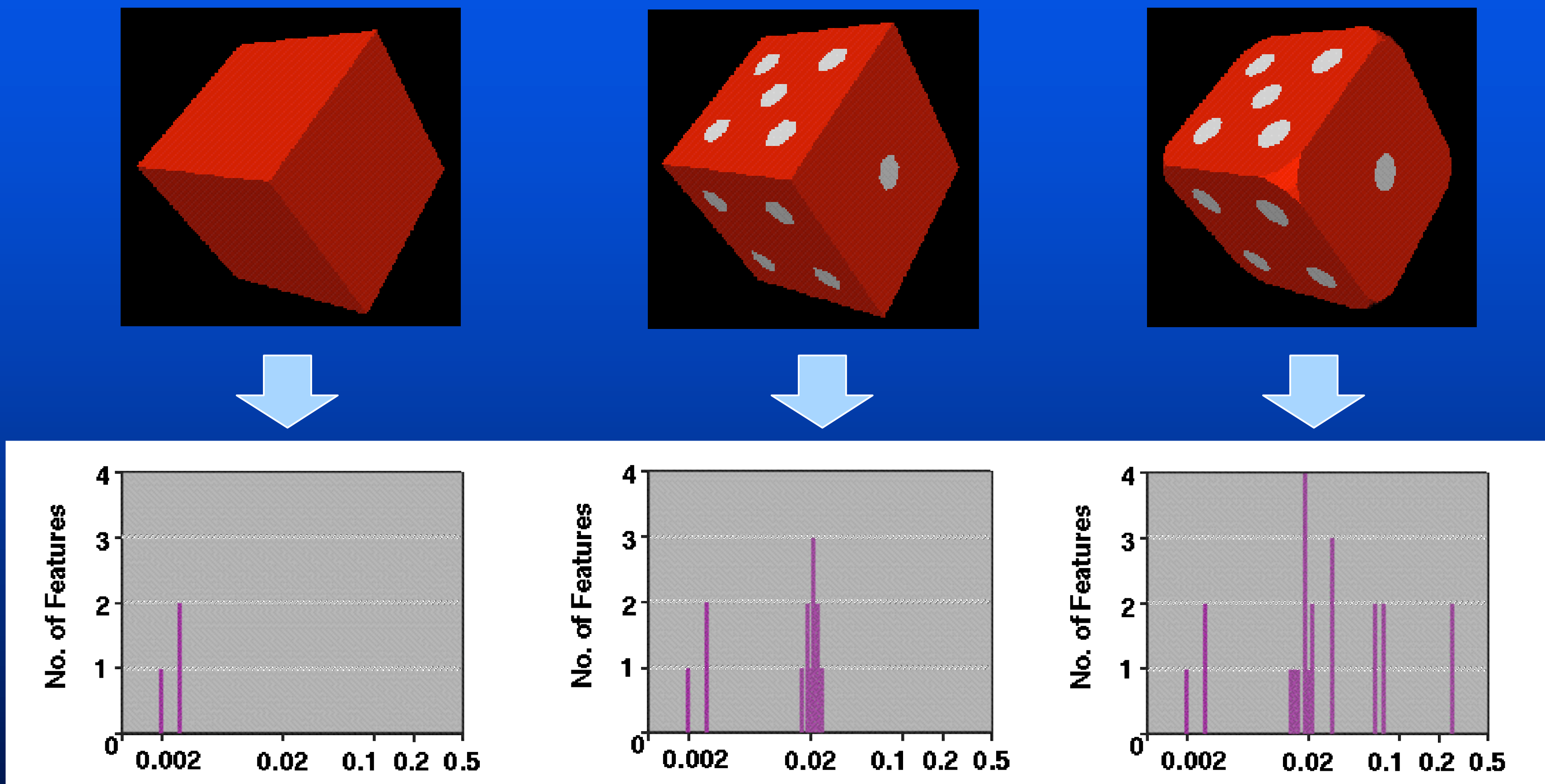
Extract relative fundamental spatial frequencies (c/pixel)

- 1 feature = 1/2 contrast cycle, so 4 pixels = 1/8 c/pixel



Apply Field of View scaling (c/deg)

Calculating c/deg - Reddy



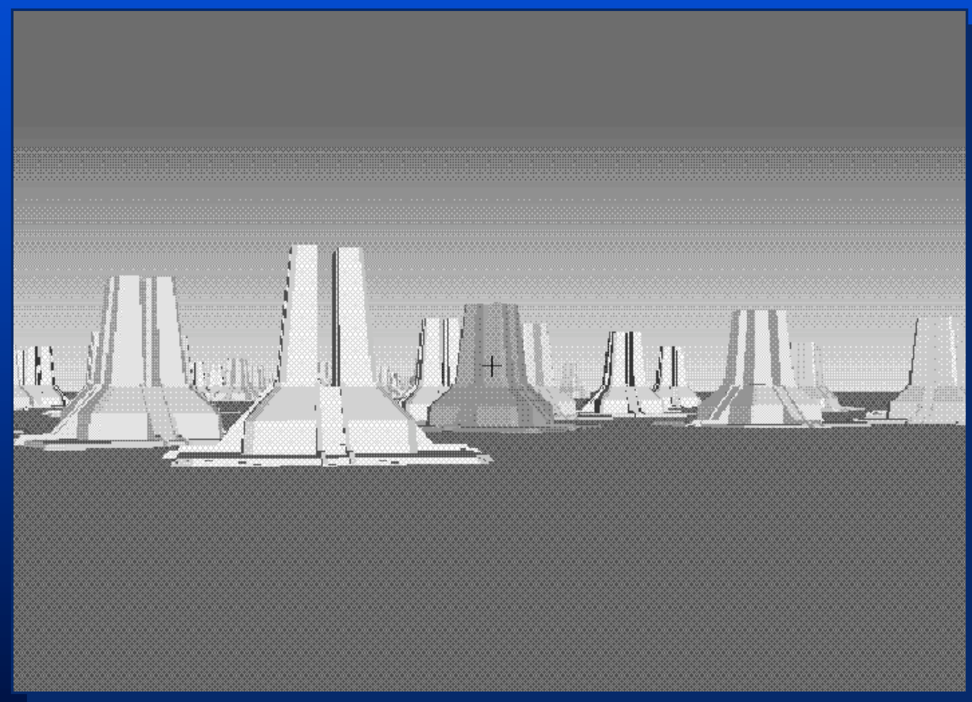
x-axis = relative spatial frequency (c/pixel).

Scale by field of view for c/deg

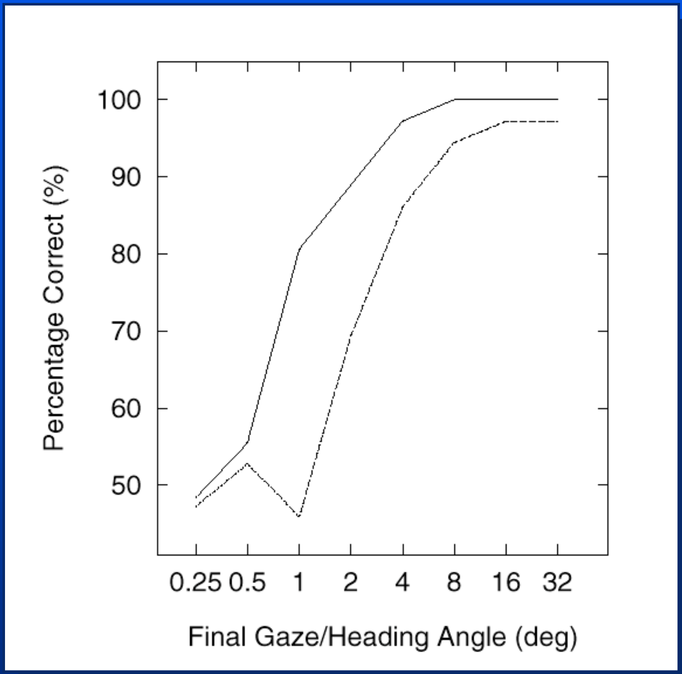
Sample Results - Reddy

- Passive psychophysical navigation task used to evaluate benefit of perceptual criteria

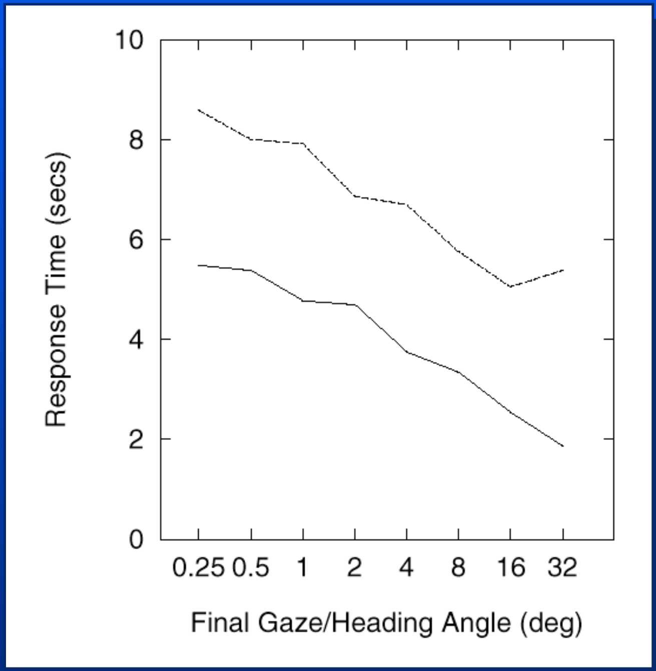
4 LODs for each object
LOD varied by size, velocity, & eccentricity
User focused on crosshair in center
Desktop configuration with bite bar to constrain head movements



Sample Results - Reddy



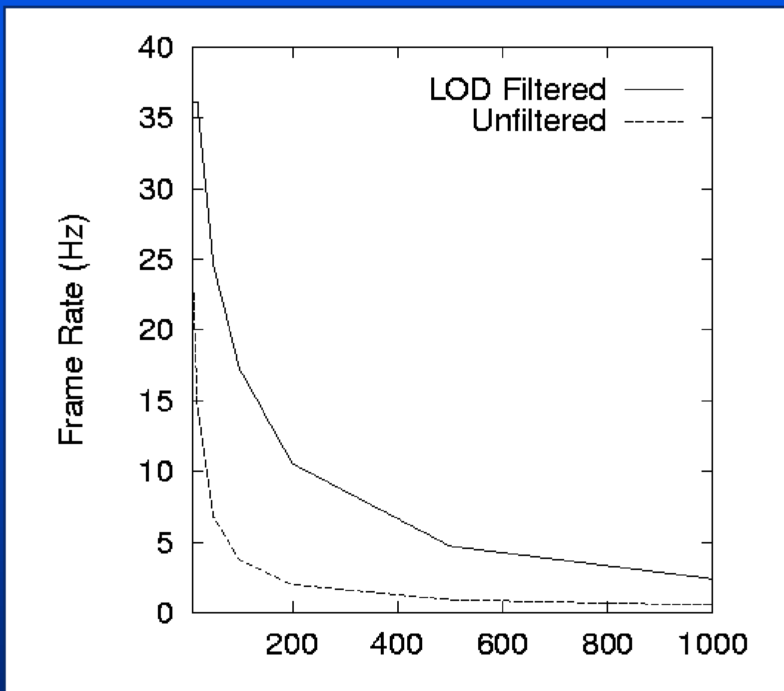
User task proficiency increased 2.8 times (at the 95% performance threshold)



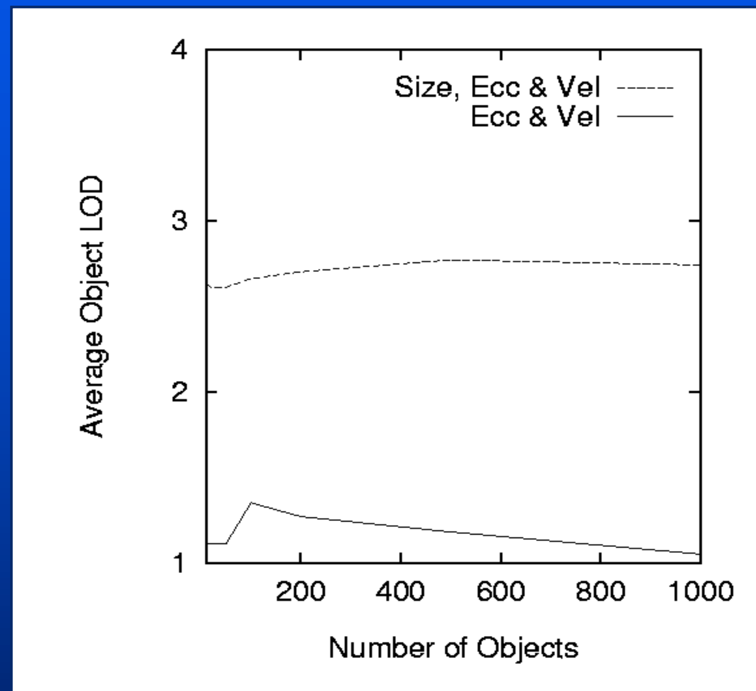
Average response time improved 1.67 times

Solid line = perceptual LOD optimizations. Broken line = no LOD.

Sample Results - Reddy

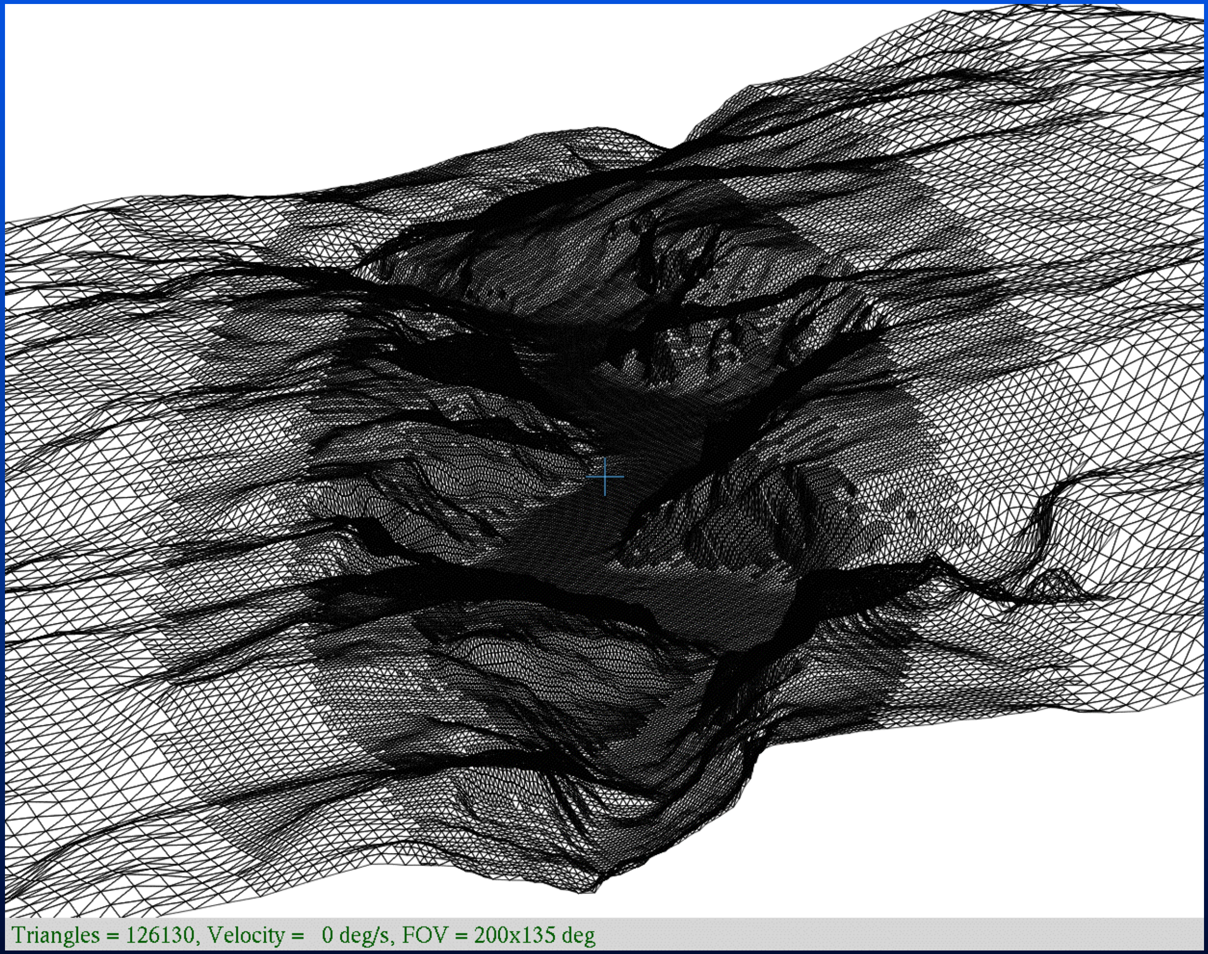


4-5 times improvement in frame rate when using LOD



Size LOD accounts for 90-95% of improvement

View-Dependent Perceptual LOD



Original Model

1,116,720 tris

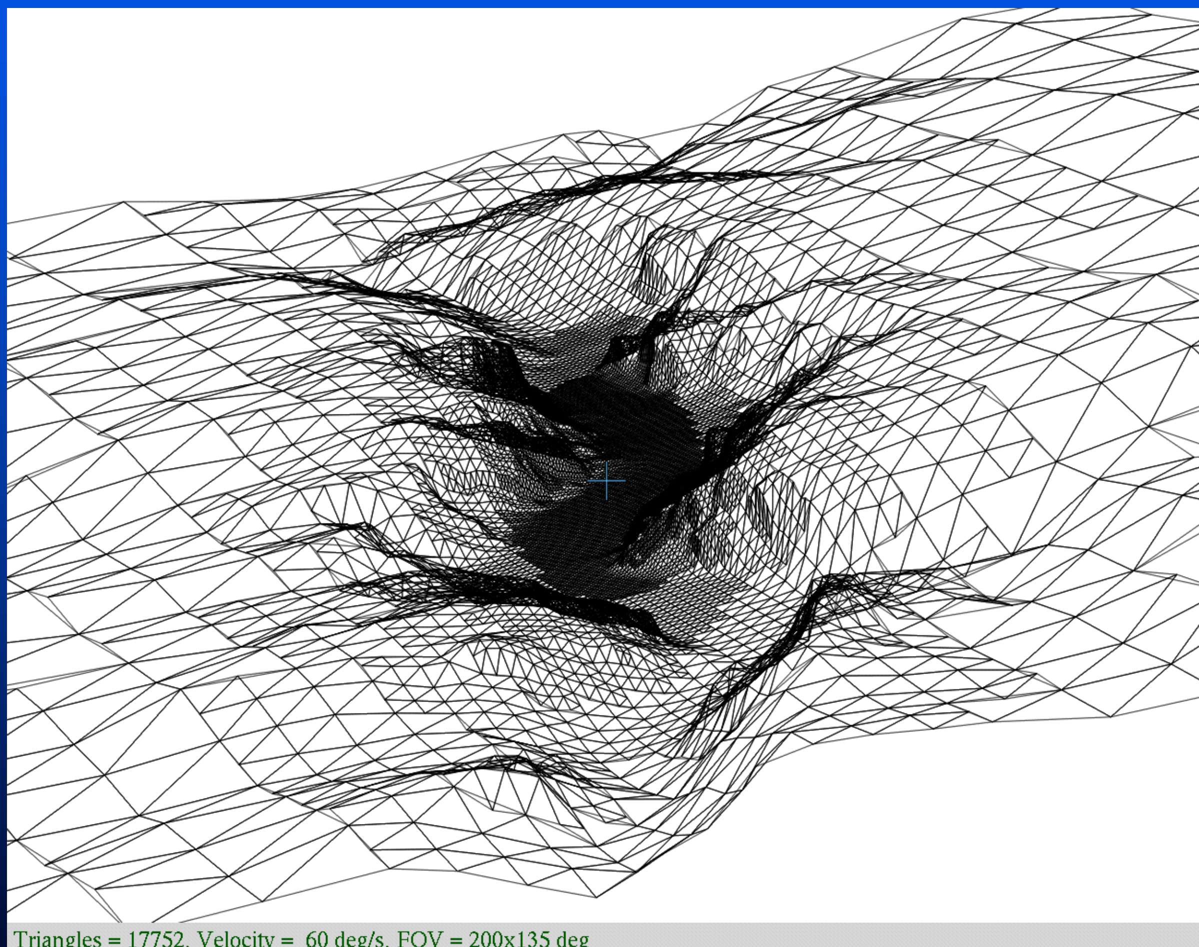
Eccentricity Opt

126,130 tris

Field of View

200x135 deg

View-Dependent Perceptual LOD



Original Model

1,116,720 tris

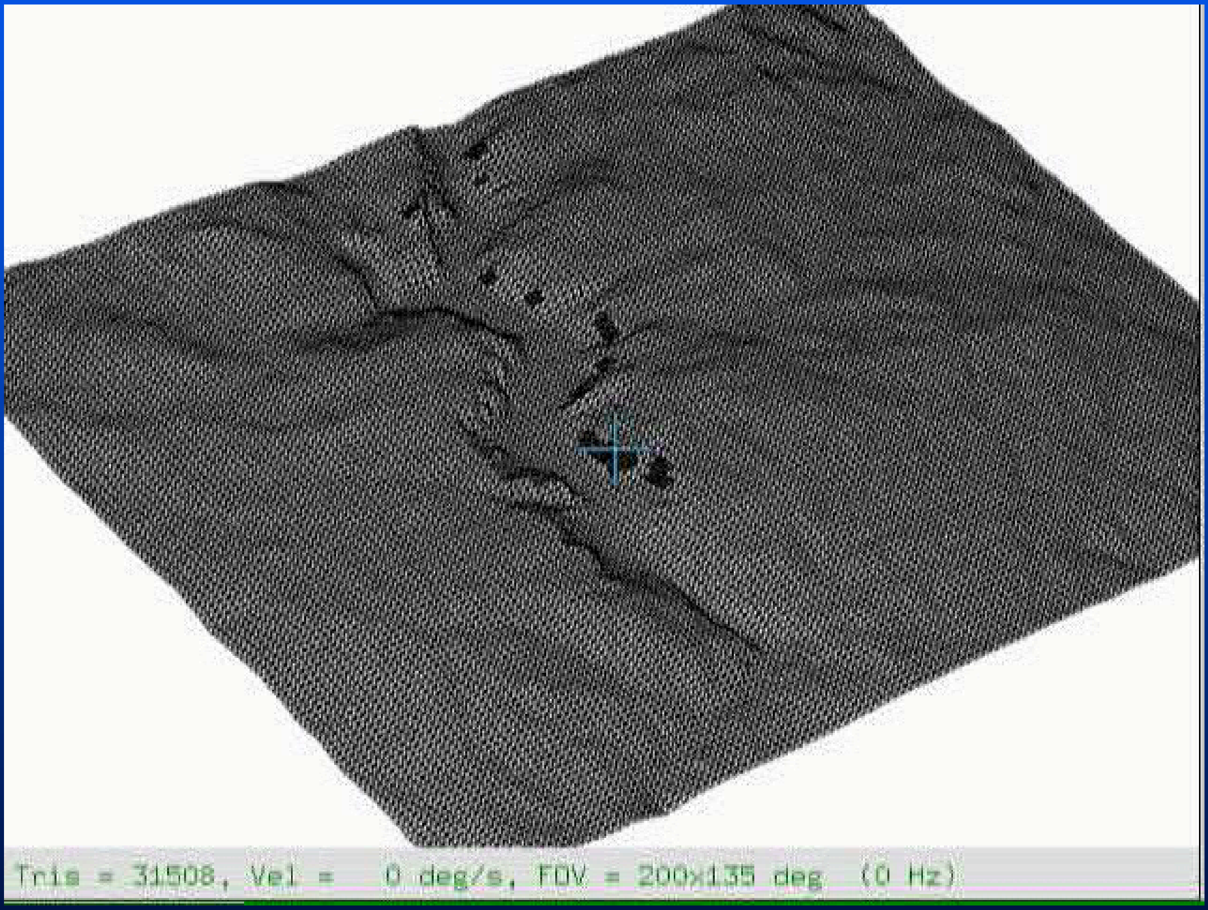
Ecc + 60 deg/s

17,752 tris

Field of View

200x135 deg

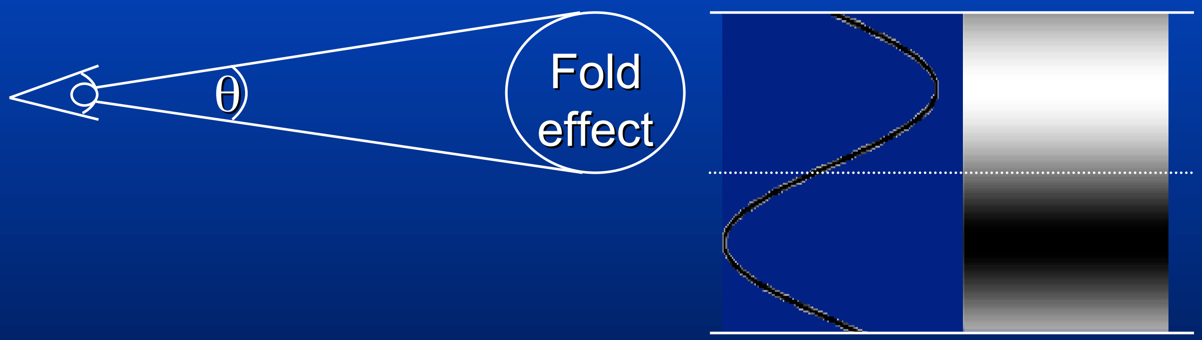
View-Dependent Perceptual LOD



Calculating c/deg - Luebke

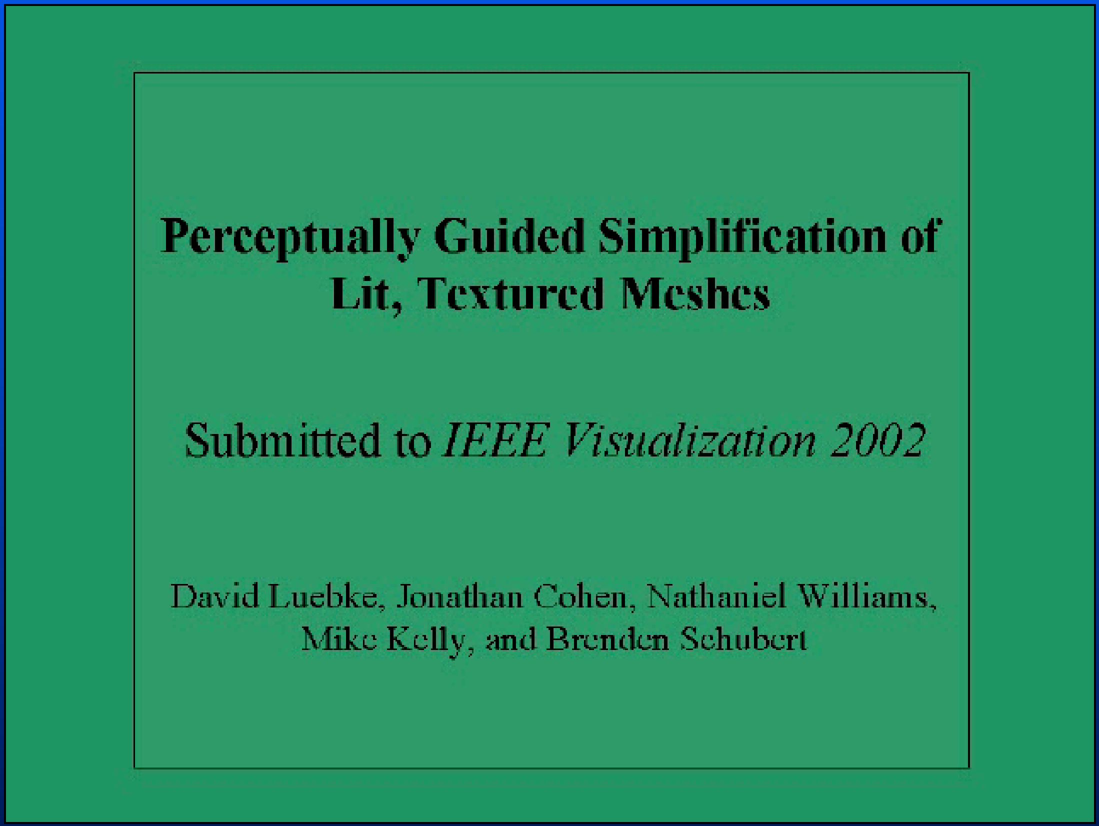
– Use bounding spheres to model extent of change caused by the folding a vertex in a hierarchical vertex tree

– Compute spatial frequency of change based upon angular projection of sphere (θº fold extent gives frequency of one cycle per 2 θº)



Contrast calculated by comparing intensities of all vertices in original and simplified surface (if silhouette edge, compare against brightest and darkest intensities in the scene).

Movie - Luebke



"Perceptually Guided Simplification of Lit, Textured Meshes"

Luebke et al.

Enabling Technologies

- Head tracking
  - Commonly used in virtual reality systems
  - Fast head movement = high angular velocity
- Eye Tracking
  - Required for true perceptual LOD optimizations
  - Generally cumbersome and suffers from problems of lag, drift, resolution, etc.
  - But perhaps head tracking is enough:
    - Resting gaze orientation ~= head orientation
    - Most saccades occur with 15 deg of gaze point

Conclusions

- Often, Perceptual LOD more applicable to immersive systems:
  - head/eye tracking better than center of display
  - extra benefit when user moves head/gaze
  - one display per person (multiple viewers possible though)
  - though... perceptual models can help appearance preservation too
- Results on Perceptual Criteria
  - Velocity and eccentricity optimizations should be used in conjunction for maximum benefit
  - Distance / Size LOD offers the most advantage (e.g. around 90-95% in a non-immersive, *discrete* LOD system).
  - View-dependent LOD system best for max. resolution reduction
  - Supporting Velocity LOD can give big wins in a dynamic environ.
- Less reduction than a non-perceptual system, but can do perceptually linear fixed frame rate also.

Further Research Areas

- Need to think about temporal effects of switching between two different representations (flicker frequency). The peripheral field is highly sensitive to flicker. Does this even matter?
- Need better perceptual metrics to assess the spatial frequency and contrast of a computer-generated image or of a polygonal model.

Need more results on the benefit of using various perceptual criteria under different viewing conditions and display devices (e.g. immersive systems).

Web Resources

Slides for this presentation
http://www.lodbook.com/

Perceptually Modulated LOD Thesis
http://martinreddy.net/thesis/

David Luebke's VDS Library
http://vdslib.virginia.edu/

Mike Krus' LOD Resources
http://www.multimania.com/krus/CG/LODS/

Lee Bull (Pip)'s LOD World
http://www.cs.ucl.ac.uk/staff/P.Bull/lod/lod.html

Multiresolution Modeling
http://www.cs.cmu.edu/~garland/multires/

LOD Bibliography

- Funkhouser, T. A. and Séquin, C. H. (1993). “Adaptive Display Algorithm for Interactive Frame Rates During Visualization of Complex Virtual Environments”. SIGGRAPH ‘93 Proceedings, 27: 247-254.
- Hitchner, L. E. and McGreevy, M. W. (1993). “Methods for User-Based Reduction of Model Complexity for Virtual Planetary Exploration”. Proceedings of the SPIE, 1913: 622-636.
- Ohshima, T., Yamamoto, H. and Tamura, H. (1996). “Gaze-Directed Adaptive Rendering for Interacting with Virtual Space”. Proceedings of VRAIS ‘96, pp. 103-110.
- Koller, D., Lindstrom, P., Ribarsky, W., Hodges, L., Faust, N. and Turner, G. (1995). “Virtual GIS: A Real-Time 3D Geographics Information System”. Proceedings of Visualization ‘95, pp 94-100.
- Astheimer, P. and Pöche, M-L. (1994). “Level-of-Detail Generation and its Application to Virtual Reality”. Proceedings of VRST ‘94, pp. 299-309.
- Watson, B. A., Walker, N., Hodges, L. and Reddy, M. (1997). “An Evaluation of Level of Detail Degradation in Head-Mounted Display Peripheries”. Presence: Teleoperators and Virtual Enviornments, 6(6).
- Lindstrom, P. and Turk, G. (2000). “Image-Driven Simplification”. ACM Transactions on Graphics, 19(3): 204-241.
- Reddy, M. (1997). “Perceptually Modulated Level of Detail for Virtual Environments”. PhD Thesis CST-134-97, University of Edinburgh, UK.

Perception Bibliography

- Kelly, D. H. (1979). “Motion and Vision II: Stabilized Spatio-Temporal Threshold Surface”. Journal of the Optical Society of America, 69(10): 1340-1349
- Rovamo, J. and Virsu, V. (1979). “An Estimation and Application of the Human Cortical Magnification Factor”, Experimental Brain Research, 37: 495-510.
- Nakayama, K. (1990). “Properties of Early Motion Processing: Implications for the Sensing of Egomotion”. In The Perception and Control of Self Motion, Lawrence Erlbaum, Hillsdale, NJ, pp. 69-80.
- Campbell, F. W. and Robson, J. G. (1968). “Application of Fourier Analysis to the Visibility of Gratings”. Journal of Physiology, 197: 551-566.
- Burr, D. C. and Ross, J. (1982). “Contrast Sensitivity at High Velocities”. Vision Research, 22: 479-484.
- Koenderink, J. J., Bouman, M. A., de Mesquita, A. E. B. and Slappendel, S. (1978). “Perimetry of Contrast Detection Thresholds of Moving Spatial Sine Wave Patterns”. Journal of the Optical Society of America, 68(6): 845-854.
- Lamming, D. (1991). “On the Limits of Visual Detection”. In Vision and Visual Dysfunction: Limits of Vision, vol 5. MacMillan Press Ltd., chapter 3, pp 15-22.
- Tyler, C. W. (1985). “Analysis of Visual Modulation Sensitivity. II. Peripheral Retina and the Role of Photoreceptor Dimensions”. Journal of the Optical Society of America, A2(3): 393-398.
- Wandel, B. A. (1995). “Foundations of Vision”. Sinauer Associates Inc., Sunderland, MA.