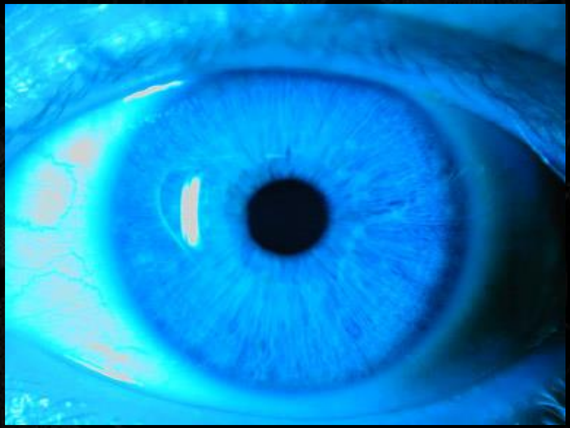
Detail to attention: Exploiting Visual Tasks for Selective Rendering



Kirsten Cater [1],

Alan Chalmers [1] and Greg Ward [2]

[1] University of Bristol, UK

[2] Anyhere Software, USA

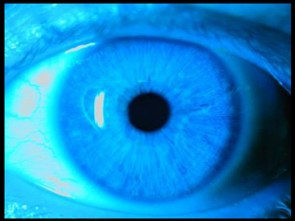
Contents of Presentation

- Introduction – Realistic Computer Graphics
- Flaws in the Human Visual System
- Magic Trick!
- Previous Work
- Counting Teapots Experiment
- Perceptual Rendering Framework
- Example Animation
- Conclusions

Realistic Computer Graphics

- Major challenge achieve realism at interactive rates
- How do we keep computational costs realistic?

The Human Visual System

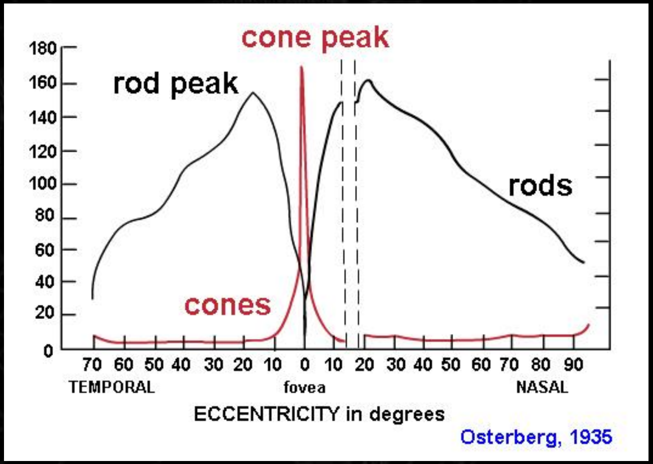


Good but not perfect!

Flaws in the human visual system:

- Change Blindness
- Inattentional Blindness

Avoid wasting computational time



Magic trick to Demonstrate Inattentional Blindness

Please choose one of the six cards below.

Focus on that card you have chosen.

Magic trick (2)

Can you still remember your card?

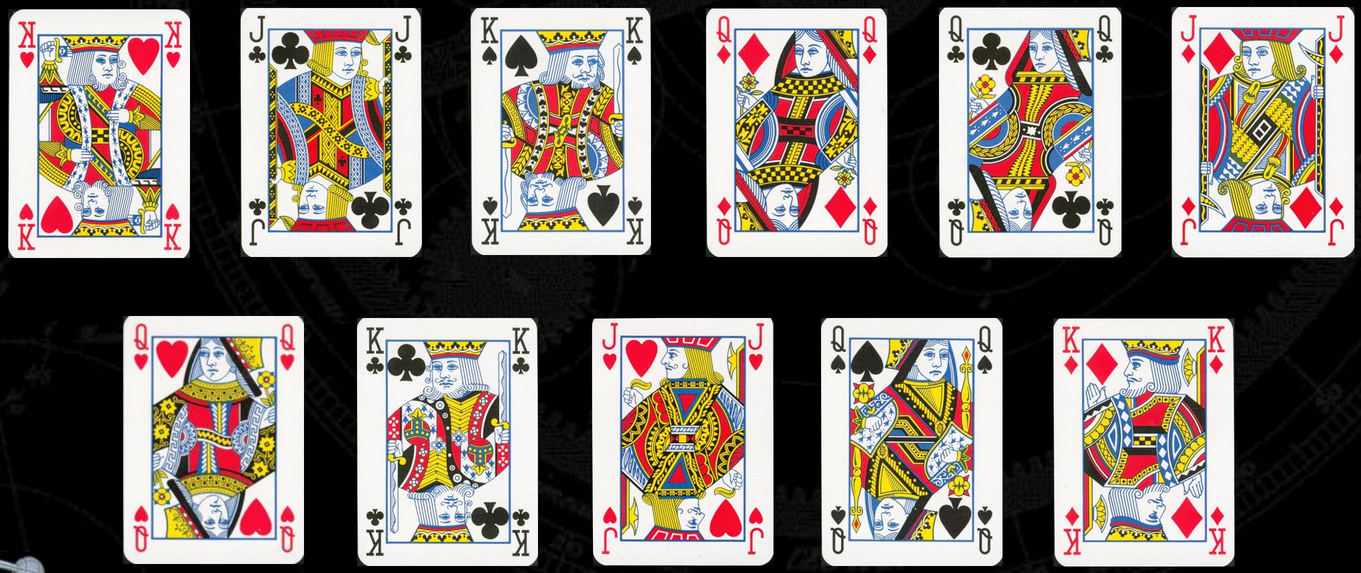
Magic trick (3)

Here are the remaining five cards, is your card there?

Did I guess right? Or is it an illusion?

Magic trick – The Explanation

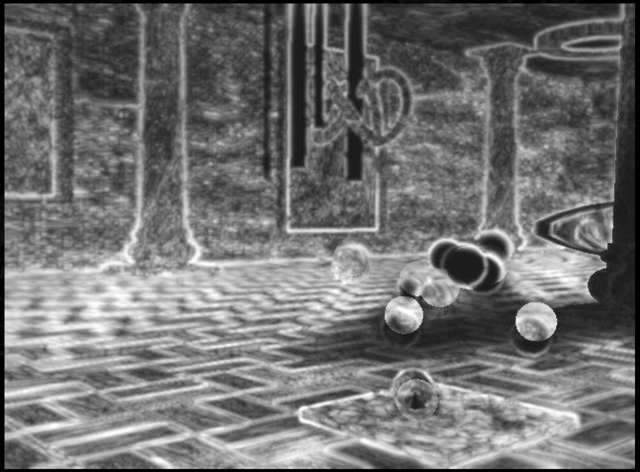
- You just experienced ***Inattentional Blindness***
- ***None*** of the original six cards was displayed!



Previous Work

Bottom-up Processing – Stimulus Driven

- Perceptual Metrics – Daly, Myszkowski et al., etc.
- Peripheral Vision Models – McConkie, Loschky et al., Watson et al., etc.
- Saliency Models – Yee et al., Itti & Koch, etc.

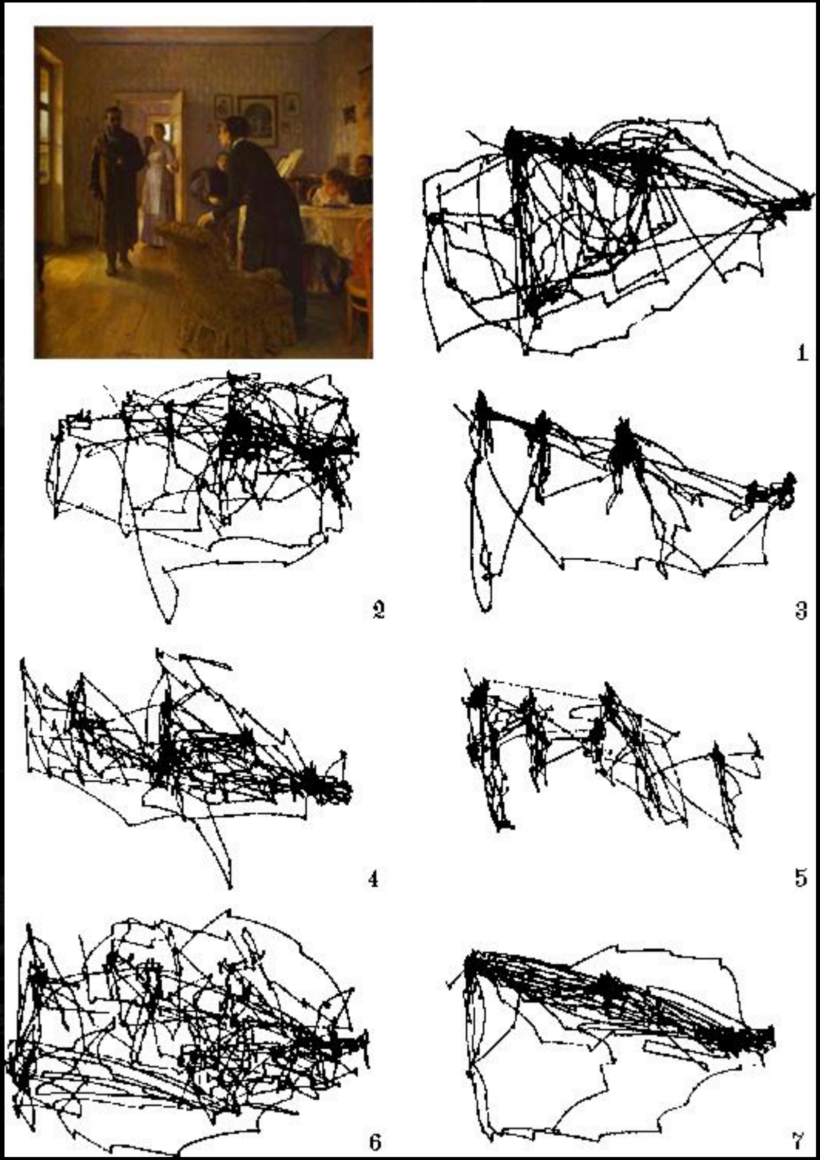


Top-down processing – Task Driven

Yarbus (1967) recorded observers' fixations and saccades whilst answering specific questions given to them about the scene they were viewing.

Saccadic patterns produced were dependent on the question that had been asked.

Theory – if we don't perceive parts of the scene what's the point rendering it to such a high level of fidelity?!



Counting Teapots Experiment

- Inattentional Blindness – more than just peripheral vision methods – don’t notice an object’s quality if it’s not related to task at hand, even if observers fixate on it.
- Top-down processing i.e. task driven unlike saliency models which are stimulus driven, bottom-up processing.

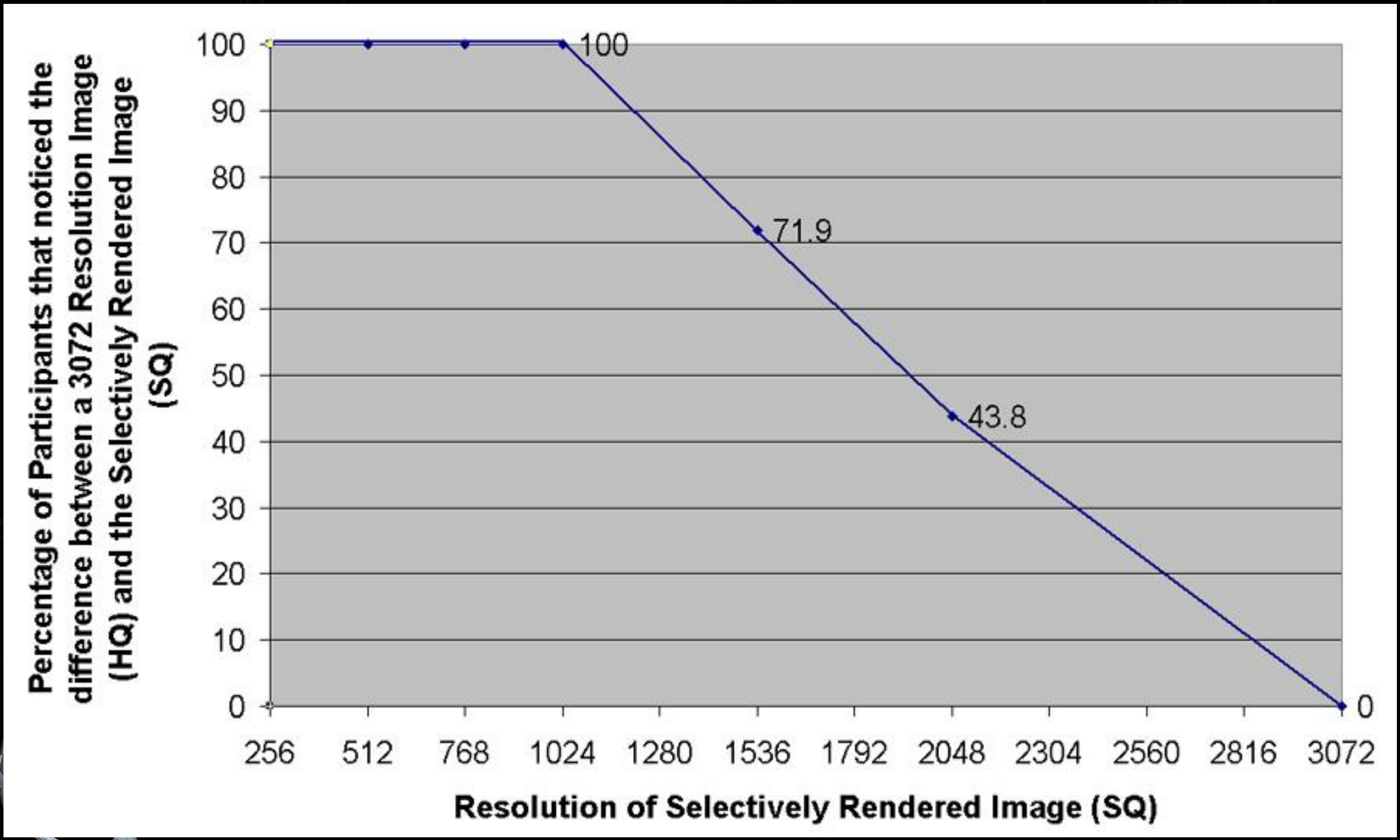
Experiment: Get participants to count teapots for two images then ask questions on any observed differences between the images. Also jogged participants’ memory getting them to choose which image they saw from a choice of two.

Counting Teapots Experiment

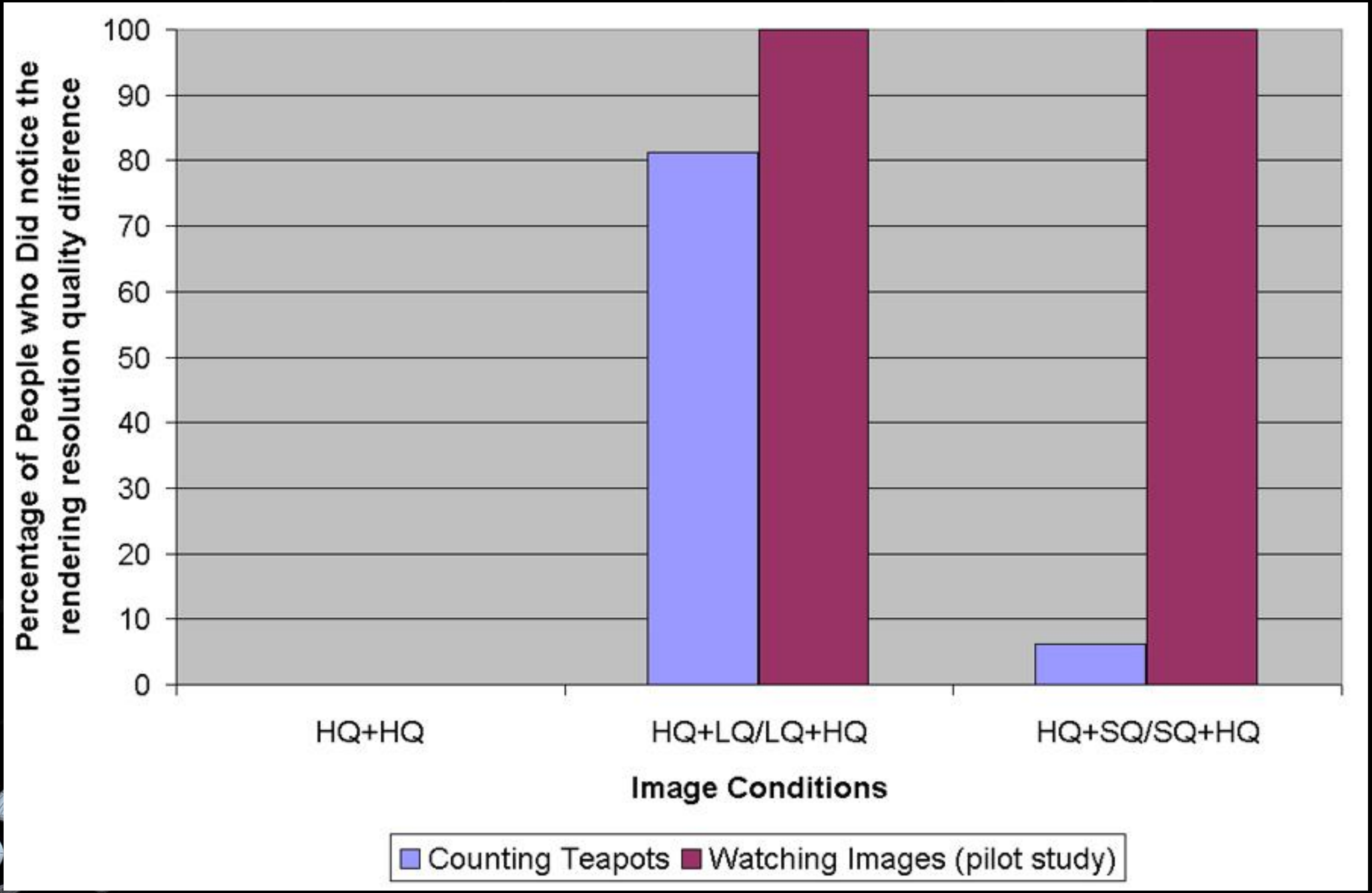
3072 x 3072

1024 x 1024

Counting Teapots Experiment –
Pre-Exp to find the Detectable Resolution

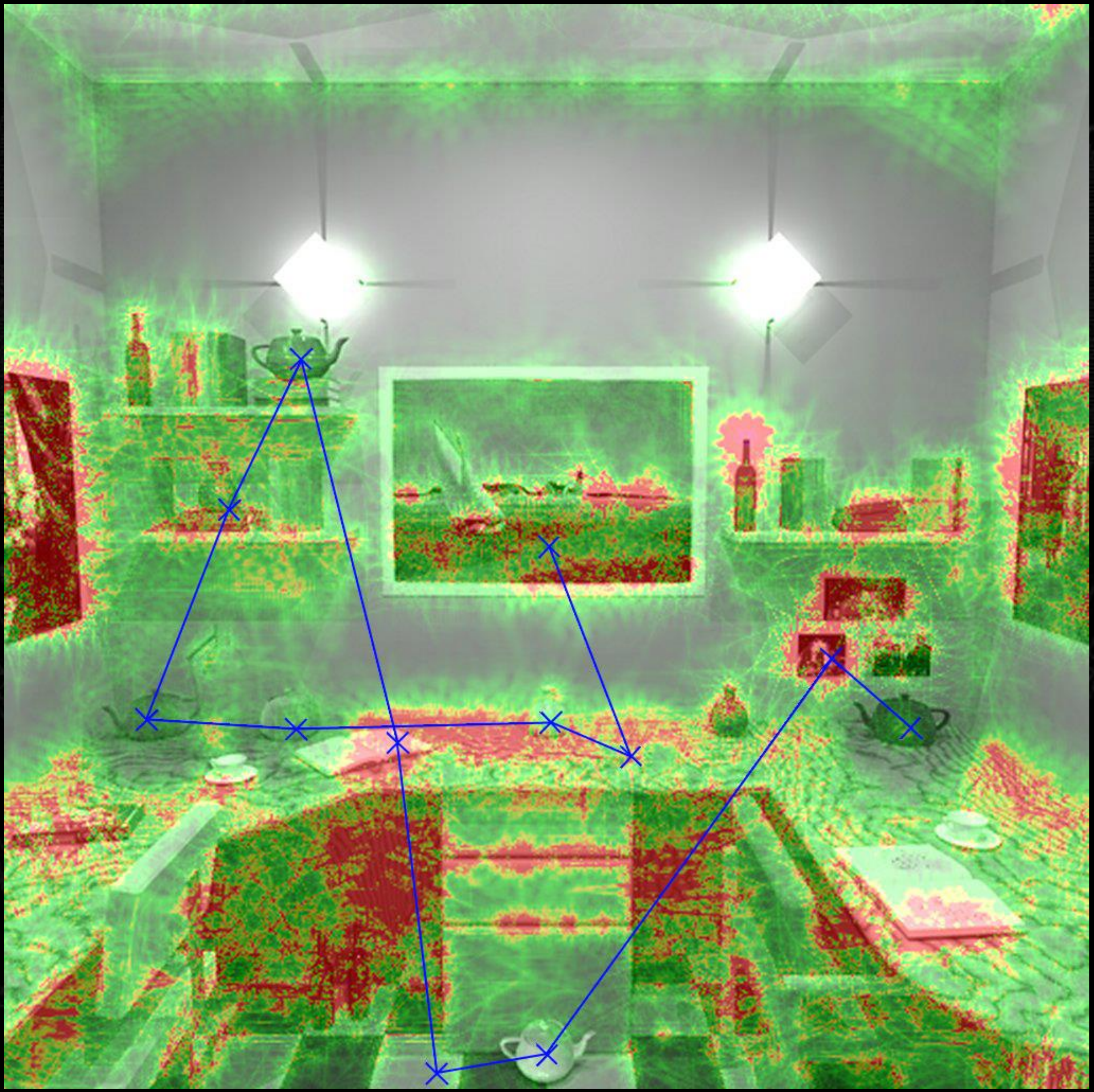


Counting Teapots Experiment



Eye Tracking Verification

Eye Tracking Verification with VDP



Counting Teapots Conclusion

Failure to distinguish difference in rendering quality between the teapots (selectively rendered to high quality) and the other low quality objects is NOT due to purely peripheral vision effects.

The observers fixate on the low quality objects, but because these objects are not relevant to the task at hand they fail to notice the reduction in rendering quality!

Perceptual Rendering Framework

Use results/theory from experiment to design a "*Just in time*" animation system.

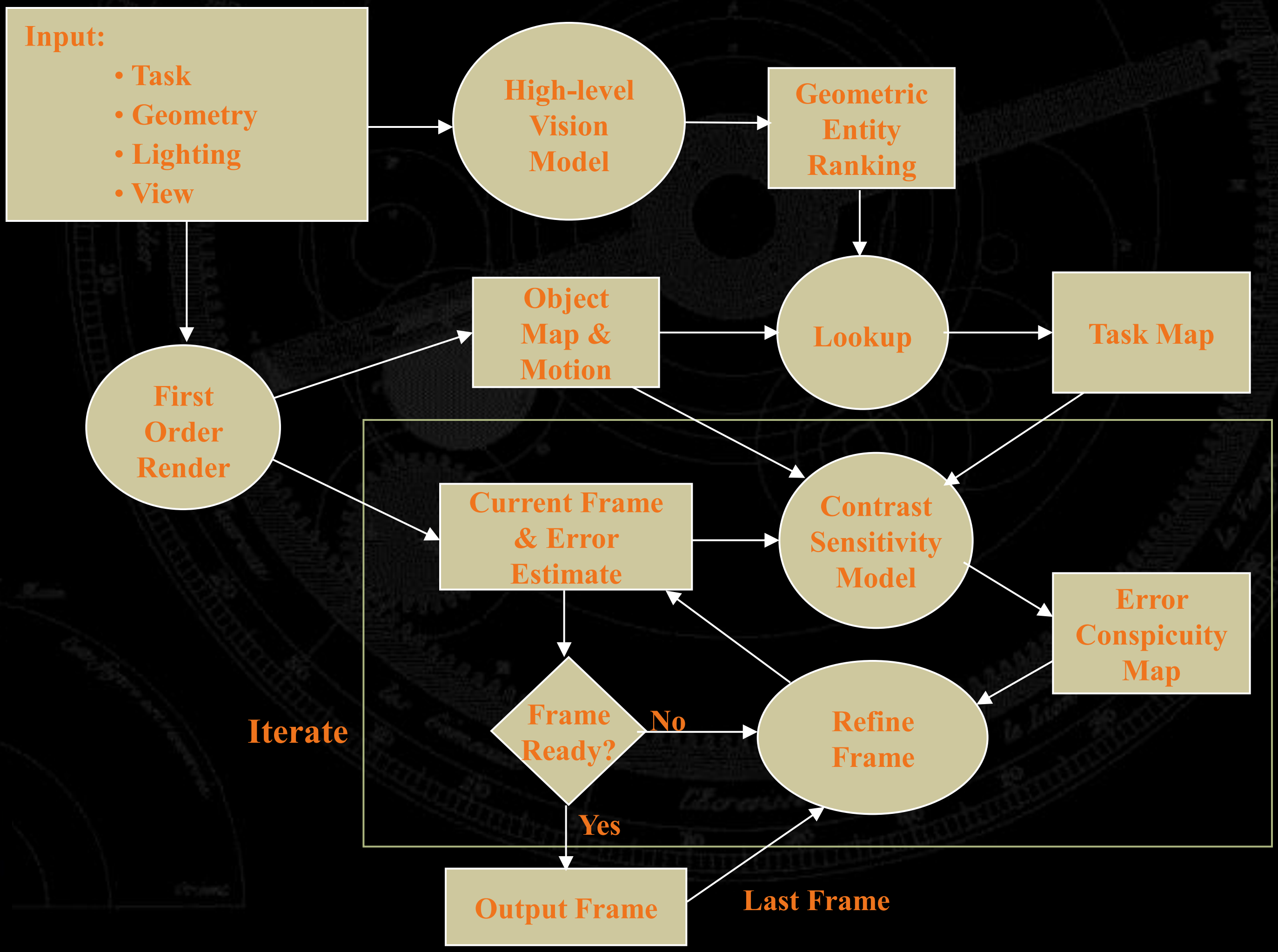
Exploits inattentional blindness and IBR.

Generalizes to other rendering techniques

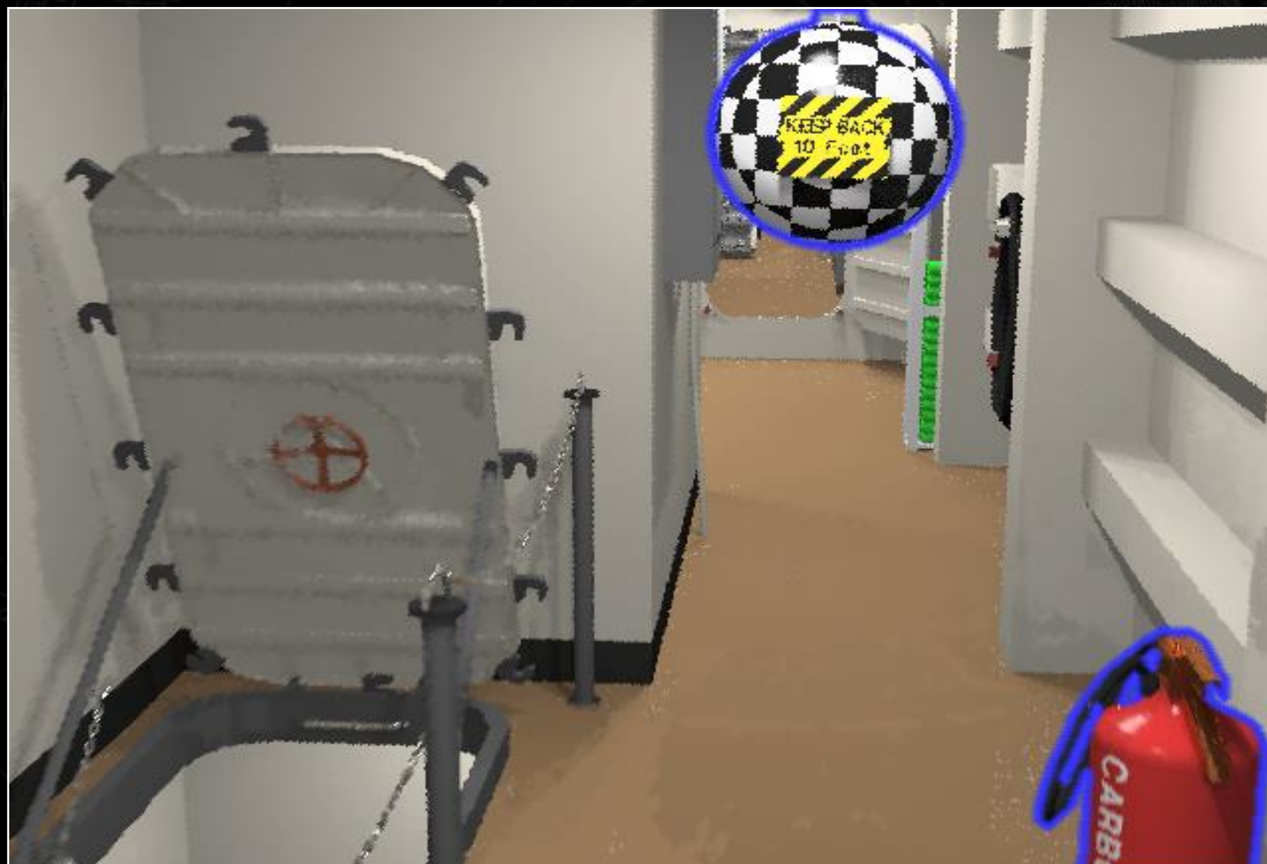
- Demonstration system uses *Radiance*
- Potential for real-time applications

Error visibility tied to attention and motion.

Rendering Framework



Example Frame w/ Task Objects



Example Animation

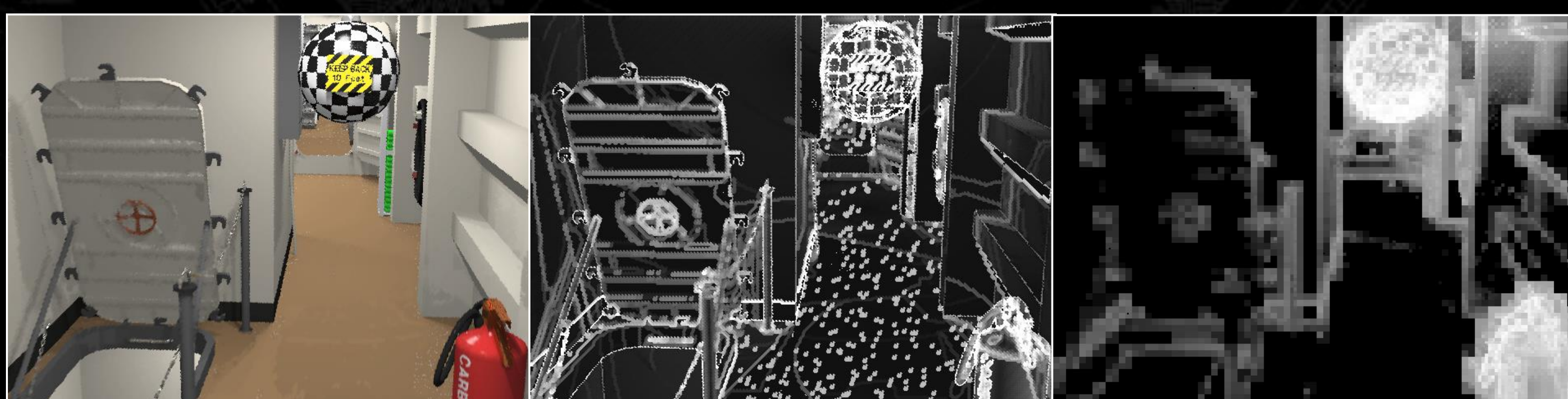
Main

Bonus

DVD

The following animation was rendered at two minutes per frame on a 2000 model G3 laptop computer.

Many artifacts are intentionally visible, but less so if you are performing the task.



Error Map Estimation

- Stochastic errors may be estimated from neighborhood samples.
- Systematic error bounds may be estimated from knowledge of algorithm behavior.
- Estimate accuracy is not critical for good performance.

Initial Error Estimate

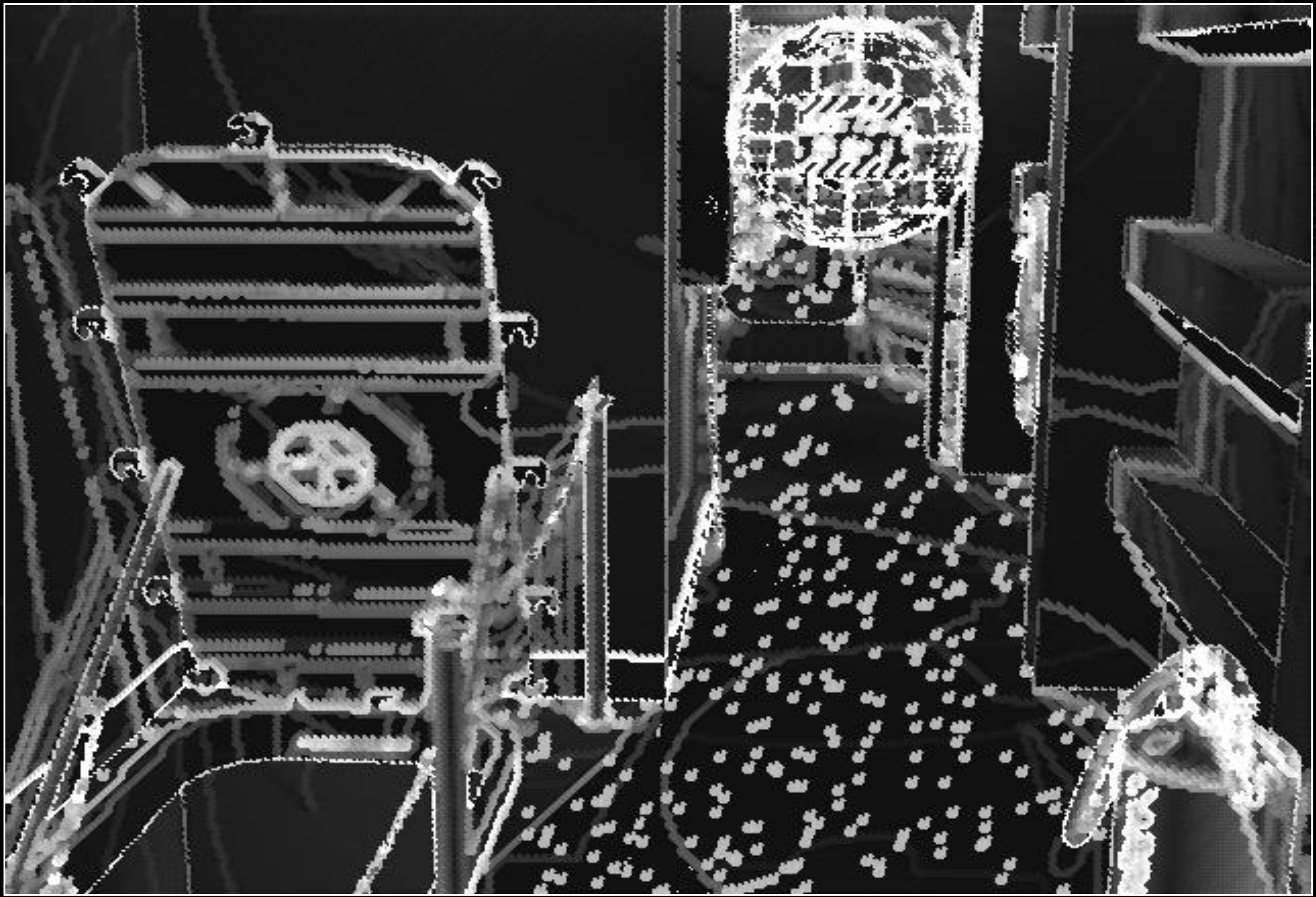


Image-based Refinement Pass

Since we know exact motion, IBR works very well in this framework.

Select image values from previous frame

- Criteria include coherence, accuracy, agreement

Replace current sample and degrade error

- Error degradation results in sample retirement

Contrast Sensitivity Model

Additional samples are directed based on Daly's CSF model:

$$CSF(\rho, v_R) = k \cdot c_0 \cdot c_2 \cdot v_R \cdot (c_1 2\pi\rho)^2 \exp\left(-\frac{c_1 4\pi\rho}{\rho_{max}}\right)$$

where:

ρ is spatial frequency

v_R is retinal velocity

$$k = 6.1 + 7.3\left|\log(c_2 v_R / 3)\right|^3$$

$$\rho_{max} = 45.9/(c_2 v_R + 2)$$

$$c_0 = 1.14, \quad c_1 = 0.67, \quad c_2 = 1.7 \quad \text{for CRT at 100 cd/m2}$$

Error Conspicuity Model

Retinal velocity depends on task-level saliency:

$$v_R = \left| v_I - \min(v_I \cdot S / S_{\max} + v_{\min},\ v_{\max}) \right|$$

where:

v_I = local pixel velocity (from motion map)
S = task-level saliency for this region
S_{max} = max. saliency in this frame, but not less than 1/0.82
v_{min} = 0.15°/sec (eye drift velocity)
v_{max} = 80°/sec (movement-tracking limit)

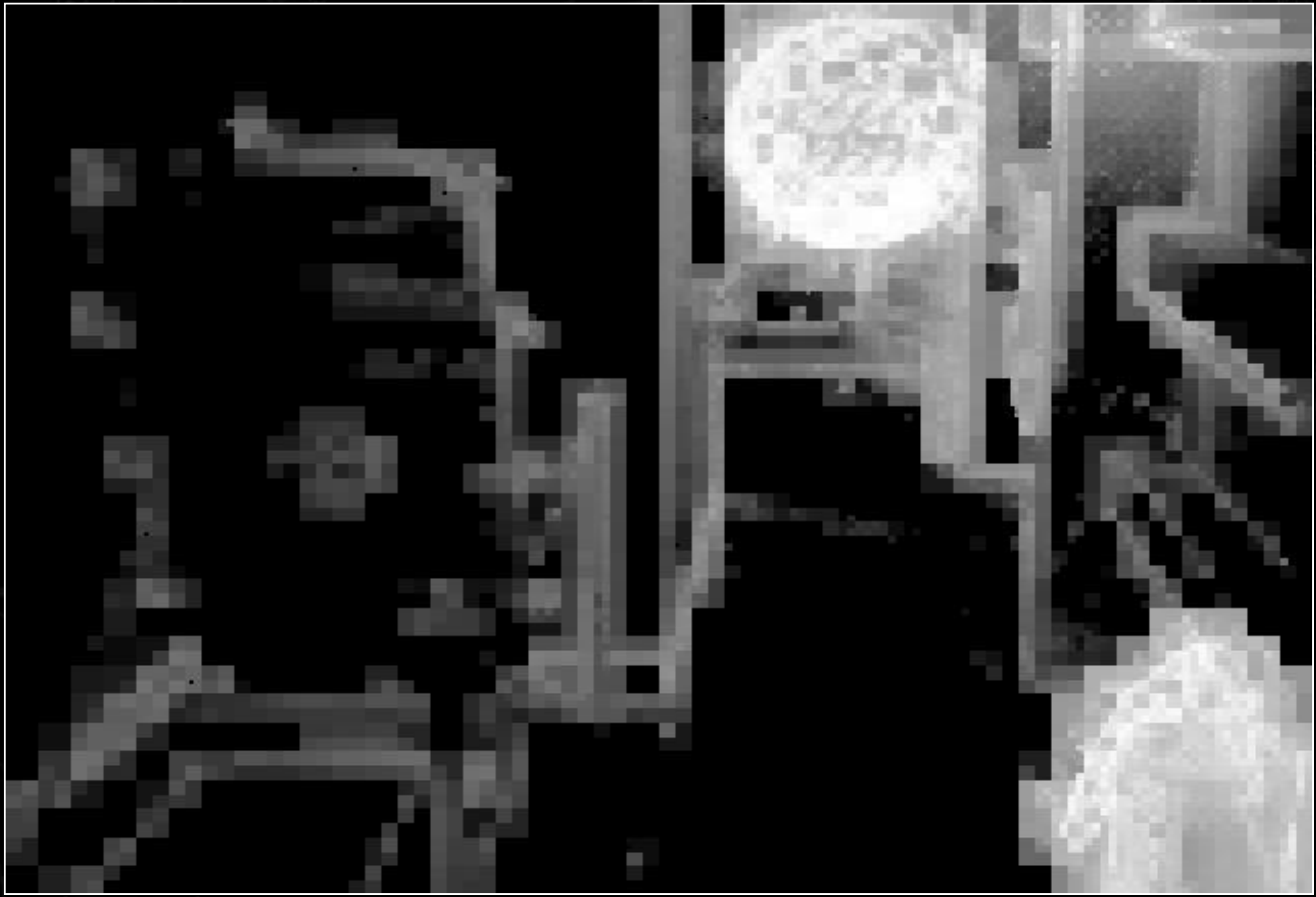
$$EC = S \cdot \max(E \cdot CSF / ND - 1,\ 0)$$

Error Conspicuity

where:

E = relative error estimate for this pixel
ND = noticeable difference threshold

Error Conspicuity Map



Implementation Example

Compared to a standard rendering that finished in the same time, our framework produced better quality on task objects.

Rendering the same high quality over the entire frame would take about 7 times longer using the standard method.

Framework rendering

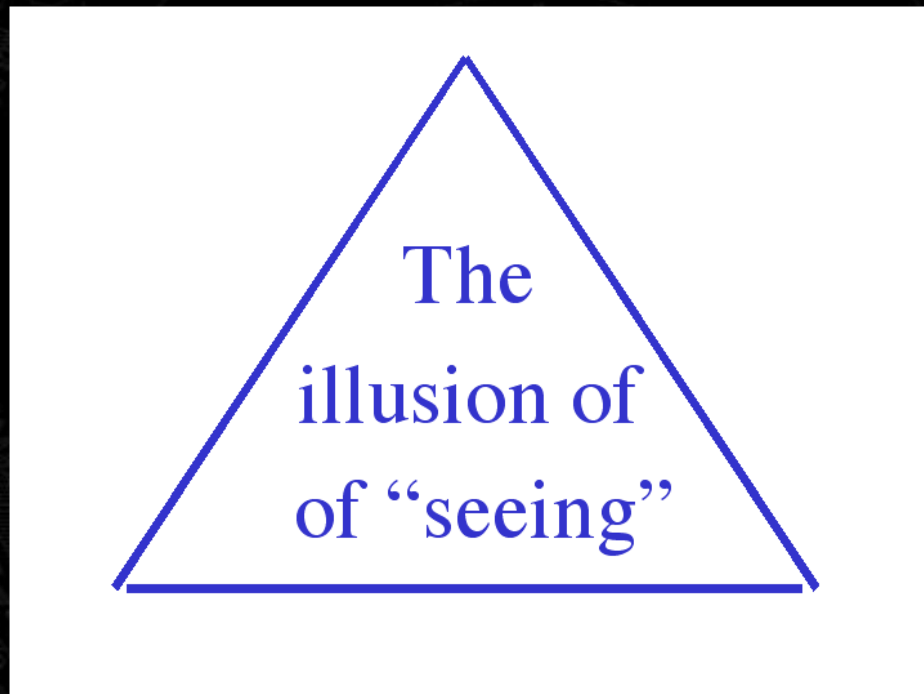
Standard rendering

Overall Conclusion

By taking advantage of flaws in the human visual system we can dramatically save on computational time by selectively rendering the scene where the user is not attending.

Thus for VR applications where the task is known a priori the computational savings, by exploiting these flaws, can be dramatic.

The End! Thank you ☺



cater@cs.bris.ac.uk, alan@cs.bris.ac.uk,
gward@lmi.net