Perception in Graphics, Visualization, Virtual Environments and Animation

Ann McNamara[*]
Department of Visualization
Texas A&M University

Katerina Mania[†]
Department of Electronic & Computer Engineering
Technical University of Crete

Diego Gutierrez[‡]
Department of Computer Science
Universidad de Zaragoza

September 16, 2011

[*] ann@viz.tamu.edu
[†] mania@ced.tuc.gr
[‡] diegog@unizar.es

Abstract

See http://web.me.com/annmcnamara/PerceptionCourse2011/Home.html for updates

This course presents timely, relevant examples on how researchers have leveraged perceptual information for optimization of rendering algorithms, to better guide design and presentation in display media, and for improved visualization of complex or large data sets. Each presentation will provide references and short overviews of cutting-edge current research pertaining to that area. We will ensure that the most up-to-date research examples are presented by sourcing information from recent perception and graphics conferences and journals such as ACM Transactions on Perception, paying particular attention work presented at the Symposium on Applied Perception in Graphics and Visualization.

About the Lecturers

Ann McNamara

Department of Visualization
Texas A&M University
3137 TAMU
College Station, TX 77843-3137

+1-979-845-4715
ann@viz.tamu.edu
http://www.viz.tamu.edu/people/ann

Ann McNamara received her undergraduate and graduate degrees from the University of Bristol, UK. Anns research focuses on the advancement of computer graphics and scientific visualization through novel approaches for optimizing an individuals experience when creating, viewing and interacting with virtual spaces. She investigates new ways to exploit knowledge of human visual perception to produce high quality computer graphics and animations more efficiently. She joined the faculty of the newly formed Department of Visualization at Texas A&M University in 2008, where she is currently an assistant professor. Ann serves on several IPCs including APGV.

Katerina Mania

Department of Electronic and Computer Engineering
Technical University of Crete
University Campus Kounoupidiana Chania Crete Greece

+30 28210 37222
k.mania@ced.tuc.gr
http://www.music.tuc.gr/kmania

Katerina Mania completed a B.Sc. in Mathematics, University of Crete, Greece, an M.Sc./Ph.D in Computer Science, University of Bristol, UK, funded by HP Labs. She worked at HP Labs as a researcher before serving on Faculty in the Department of Informatics, University of Sussex. Katerina spent her sabbatical at NASA Ames Research Centre (Advanced Displays and Spatial Perception Laboratory) in 2003. She is currently an Assistant Professor with tenure at the Technical University of Crete, Greece. Katerina was the program co-chair for APGV 2010 and Associate Editor of ACM Transactions on Applied Perception and Presence Teleoperators and Virtual Environments.

Diego Gutierrez

Universidad de Zaragoza
CPS (Edificio Ada Byron)
Maria de Luna, 1
50018 Zaragoza
SPAIN

+34-976-762-354
diegog@unizar.es
http://giga.cps.unizar.es/ diegog/

Diego Gutierrez received his PhD in Computer Science (Universidad de Zaragoza, Spain) on the topic of Curved photon mapping: global illumination in participating inhomogeneous media and also received the PhD Extraordinary Award granted by his university. Diego is a tenured Associate Professor (Professor Titular) in the Departamento de Informtica e Ingeniera de Sistemas of the Universidad de Zaragoza, and Affiliated Staff of the Computer Graphics Group of the University of Bristol. His areas of interest include computational photography, photorealistic rendering (actually, rendering in general), perception and the human visual system and High Dynamic Range imaging.

Course Overview

5 minutes: Welcome and Introductions

Ann McNamara
Welcome, overview of course & motivation for attending. Speaker Introductions

30 minutes: Overview of the Human Visual System

Ann McNamara
A look at Visual Attention, Visual Memory, and its Role in Visualization.

30 minutes: Perceptually Motivated Rendering

Ann McNamara
Overview of how knowledge from perceptual research feeds into optimized rendering algorithms.

45 minutes: Simulation and Virtual Environments

Katerina Mania
Perceptually-based Optimizations & Fidelity Metrics for Simulation Technology.

30 minutes: HDR, Illumination and Image Processing

Diego Gutierrez
Discussion of HDR for imagery and video, and also a closer look at the role perception plays in Image Processing.

45 minutes: Realistic Characters, Faces, Animation

Diego Gutierrez
The role of perception in the design and realization of realistic characters, facial animation and animation in general.

30 minutes: New Trends in Perception and Graphics Research

Katerina Mania
A summary of cutting edge perceptual research selected from APGV 2011.

10 minutes: Conclusion, Questions & Answers

All
Wrap up, review, questions and discussion.

Contents

1 Introduction

1.1 Motivation

The advent of affordable display technology, and seamless integration of real-world scenes with computer graphics fuels our continuing ability to create and display stunning realistic imagery. With the arrival of new technology, algorithms and display methods comes the realization that gains can be made by tailoring output to the intended audience; humans. Human beings have an amazingly complex perceptual systems, which have the ability to quickly capture and process vast amounts of complex data. With all its capability however, the Human Visual System (HVS) has some surprising nuances and limitations that can be exploited to the benefit of numerous graphics applications. This short course will provide insight into those aspects of the HVS and other perceptual systems that can serve as both a guide and yard-stick to further the development and evaluation of computer graphics imagery and presentations. The literature on perception provides a rich source of knowledge that can be applied to the realm of computer graphics for immediate and direct benefit, generating images that not only exhibit higher quality, but use less time and resources to process. In addition, knowledge of the HVS serves as a guide on how best to present the images to fulfill the application at hand.

1.2 Course Overview

We will present timely, relevant examples on how researchers have leveraged perceptual information for optimization of rendering algorithms, to better guide design and presentation in (3D stereoscopic) display media, and for improved visualization of complex or large data sets. Each section will provide references and short overviews of cutting-edge current research pertaining to that area. We will ensure that the most up-to-date research examples are presented by sourcing information from recent perception and graphics conferences and journals such as ACM Transactions on Perception, paying particular attention work presented at the 2010 and 2011 Symposium on Applied Perception in Graphics and Visualization.

1.3 Focus Areas

We will focus on key areas in which perceptual knowledge has been successfully interleaved with computer graphics.

1.3.1 Perceptually Motivated Visualization

Discussion of recent research pertaining to psychophysics and application to scientific and information visualization. A closer look at visual attention and visual memory will provide the framework for steering perceptually informed visualizations.

1.3.2 Exploitation of the limitations of the HVS to reduce rendering times

while improving resulting image quality. This includes real-time and non-real time graphics, image quality metrics and high dynamic range imagery.

1.3.3 Exploration of incorporating perceptual and cognitive aspects to Virtual Environments (VEs).

Such principles could be applied to selective real-time rendering algorithms, positive transfer of training as well as to optimizations for latency degradations and predictive tracking.

1.3.4 Insights into High Dynamic Range Imagery, Illumination and Image Processing.

Discussion of the now ubiquitous HDR for imagery and video, and also a closer look at the role perception plays in Image Processing.

1.3.5 Realistic Characters, Faces and Animation.

In this section we will explore the role of perception in the design and realization of characters, facial animation and animation in general.

1.3.6 New Trends in Perception and Graphics Research

A look to the future - current trends and possible emerging trends.

1.4 Summary

In summary, this course represents a whirlwind tour of insights into how the eye and brain capture and process visual information through our perceptual systems, and how we can use those insights to further advance many areas in computer graphics.

Ann McNamara

2 Overview of the Human Visual System

2.1 Introduction

NOTE: This section has been kindly reproduced from "Perceptually Motivated Graphics", Siggraph 2010 Course Notes, authored by Christopher Healey.

A fundamental goal of visualization is to produce images of data that support visual analysis, exploration and discovery, and identifying novel insights. An important consideration during visualization design is the role of human visual perception [79, 111, 115, 126]. How we "see" details in an image can directly impact a user's efficiency and effectiveness. This article surveys research on attention and visual perception, with a specific focus on results that have direct relevance to visualization and visual analytics. We discuss theories of low-level visual perception, then show how these findings form a foundation for more recent work on visual memory and visual attention.

2.2 Visual Attention and Preattentive Processing

For many years vision researchers have been investigating how the human visual system analyzes images. An important initial result was the discovery of a limited set of visual properties that are detected very rapidly by low-level and fast-acting visual processes. These properties were initially called preattentive, since their detection seemed to precede focused attention. We now know that attention plays a critical role in what we see, even at this early stage of vision. The term preattentive continues to be used, however, since it conveys an intuitive notion of the speed and ease with which these properties are identified.

Typically, tasks that can be performed on large multi-element displays in less than 200–250 milliseconds (msec) are considered preattentive. Eye movements take at least 200 msec to initiate, and random locations of the elements in the display ensure that attention cannot be prefocused on any particular location, yet viewers report that these tasks can be completed with very little effort. This suggests that certain information in the display is "seen" in parallel by low-level visual processes.

A simple example of a preattentive task is the detection of a red circle in a group of blue circles (Fig. 1). The target object has a visual property red that the blue distractor objects do not. A viewer can tell at a glance whether the target is present or absent. Here the visual system identifies the target through a difference in hue, specifically, a red target in a sea of blue distractors. Hue is not the only visual feature that is preattentive. For example, viewers can just as easily find a red circle in a background of red squares. Here, the visual system identifies the target through a difference in curvature (or form).

A unique visual property in the target—a red hue or a curved form—allows it to "pop out" of a display. A conjunction target made up of a combination of non-unique features normally cannot be detected preattentively. For example,

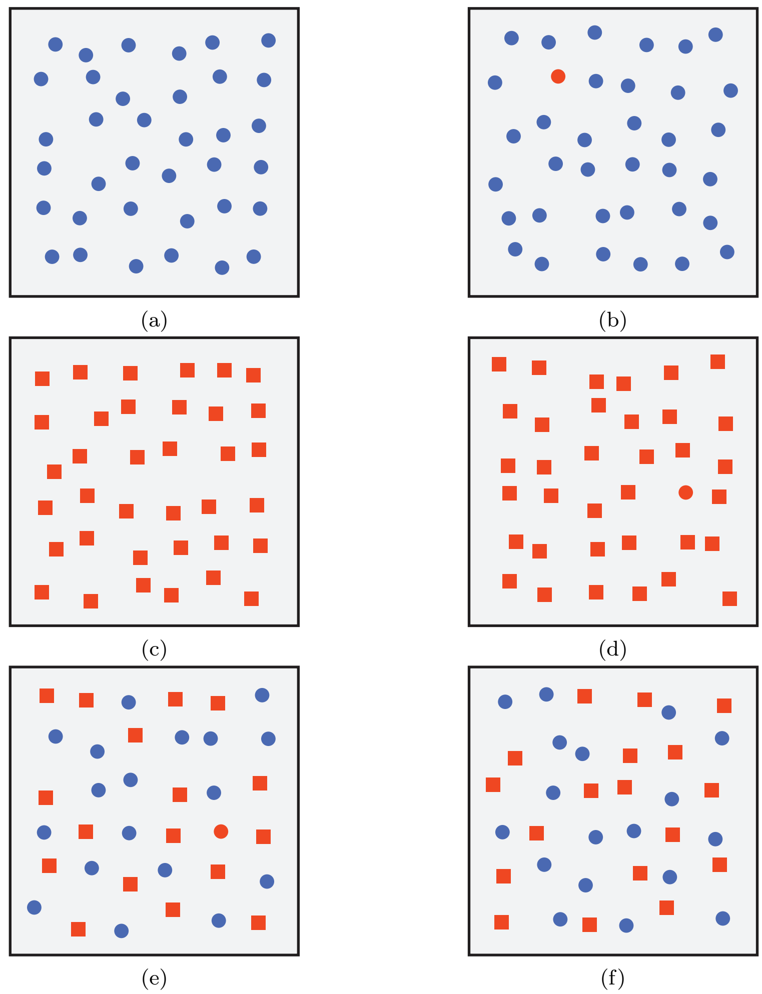


Figure 1: Target detection: (a) hue target red circle absent; (b) target present; (c) shape target red circle absent; (d) target present; (e) conjunction target red circle present; (f) target absent

consider combining the two backgrounds and searching for a red circle in a sea of blue circles and red squares. The red circle target is made up of two features: red and circular. One of these features is present in each of the distractor objects—red squares and blue circles. The visual system has no unique visual property to search for when trying to locate the target. A search for red items always returns true because there are red squares in each display. Similarly, a search for circular items always sees blue circles. Numerous studies have shown that a conjunction target cannot be detected preattentively. Viewers must perform a time-consuming serial search through the display to confirm its presence or absence.

If low-level visual processes can be harnessed during visualization, it can draw attention to areas of potential interest in a display. This cannot be accomplished in an ad-hoc fashion, however. The visual features assigned to different data attributes—the data-feature mapping—must take advantage of the strengths of our visual system, must be well-suited to the analysis needs of the viewer, and must not produce visual interference effects (e.g., conjunction search) that could mask information.

2.3 Theories of Preattentive Processing

A number of theories have been proposed to explain how preattentive processing occurs within the visual system: feature integration, textons, guided search, and boolean maps. We provide an overview of these theories, then discuss briefly feature hierarchies, which describes situations where the visual system favors certain visual features over others, and ensemble coding, which shows that viewers can generate summaries of the distribution of visual features in a scene, even when they are unable to locate individual elements based those same features.

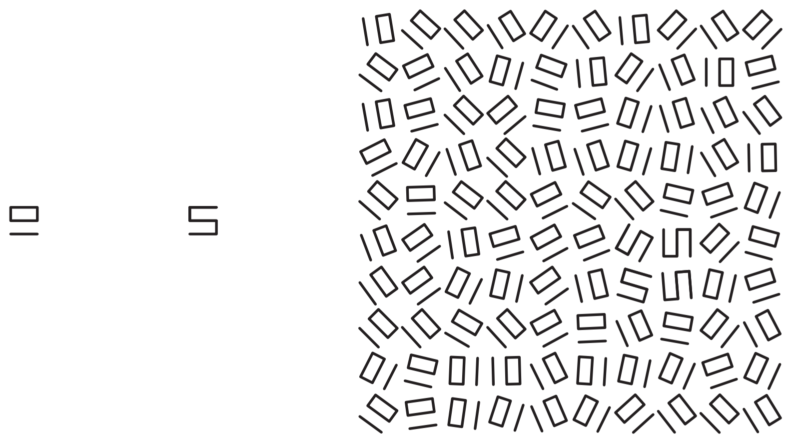
2.3.1 Feature Integration

Anne Treisman was one of the original researchers to document the area of preattentive processing [118, 116, 117]. In order to explain the phenomena, Treisman proposed a model low-level human vision made up of a set of feature maps and a master map of locations. Each feature map registers activity for a specific visual feature. Treisman suggested a manageable number of feature maps, including one for each of the opponent colors, as well as separate maps for orientation, shape, and texture. When the visual system first sees an image, all the features are encoded in parallel into their respective maps. A viewer can access a particular map to check for activity, and perhaps to determine the amount of activity. The individual feature maps give no information about location, spatial arrangement, or relationships to activity in other maps, however.

2.3.2 Textons

Bela Julész was also instrumental in expanding our understanding of what we "see" in an image. Jullész initially focused on statistical analysis of texture patterns [55, 56, 57, 58, 59]. His goal was to determine whether variations in a particular order statistic were detected by the low-level visual system, for example contrast—a first-order statistic—orientation and regularity—a second-order statistic—and curvature—a third-order statistic. Based on these findings, Julész suggested that the early visual system detects a group of features called textons, which fall into three general categories:

1. Elongated blobs—line segments, rectangles, or ellipses—with specific properties of hue, orientation, width, and so on.



(a) (b) (c)

Figure 2: Textons: (a,b) two textons A and B that appear different in isolation, but have the same size, number of terminators, and join points; (c) a target group of B-textons is difficult to detect in a background of A-textons when random rotation is applied

2. Terminators—ends of line segments.
3. Crossings of line segments

Julész believed that only a difference in textons or in their density could be detected preattentively (Fig. 2). No positional information about neighboring textons is available without focused attention. Like Treisman, Julész suggested that preattentive processing occurs in parallel and focused attention occurs in serial.

2.3.3 Guided Search

More recently, Jeremy Wolfe has proposed a theory that he calls "guided search." He hypothesized that an activation map based on both bottom-up and top-down information is constructed during visual search. Attention is drawn to peaks in the activation map that represent areas in the image with the largest combination of bottom-up and top-down influence [132, 133, 131].

As with Treisman, Wolfe believes early vision divides an image into individual feature maps. In his theory, there is one map for each feature type—a color map, an orientation map, and so on. Within each map a feature is filtered into multiple categories. Bottom-up activation follows feature categorization. It measures how different an element is from its neighbors. Top-down activation is a user-driven attempt to find items with a specific property or set of properties. The activation map is a combination of bottom-up and top-down activity. Hills in the activation map mark regions that generate relatively large amount of bottom-up or top-down influence, but without providing information about

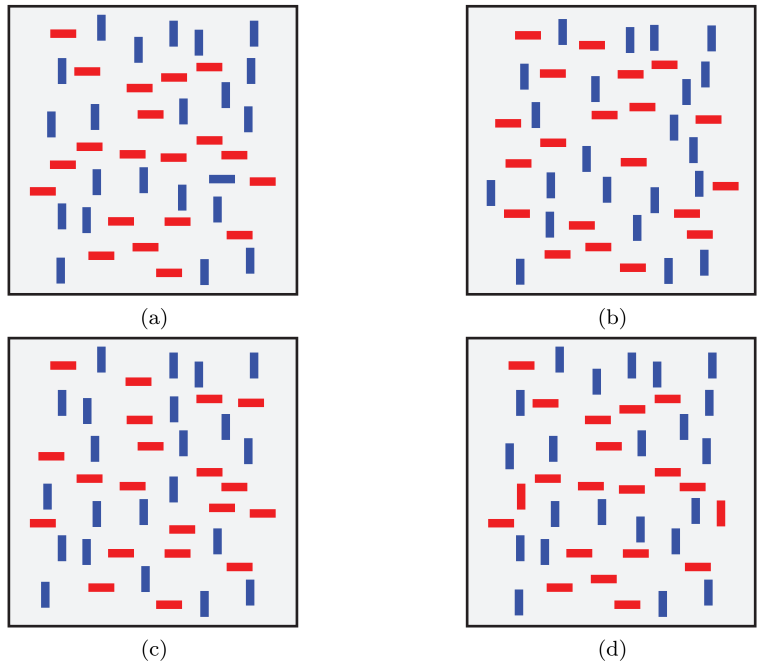


Figure 3: Conjunction search with boolean maps: (a–b) blue horizontal target, select "blue" objects, then search within for a horizontal target, present in (a), absent in (b); (c–d) red vertical target, select "red" objects, then search within for a vertical target, absent in (c), present in (d)

the source of a hill. A subject's attention is drawn from hill to hill in order of decreasing activation.

2.3.4 Boolean Maps

A more recent model of low-level vision has been presented by Huang et al. [52, 53]. This theory carefully divides visual search into two parts: selection and access. Selection involves choosing a set of objects from a scene. Access determines what properties of the selected objects a viewer can apprehend. Although both operations are implicitly present in previous theories, they are often described as a whole and not as separate steps.

Huang et al. suggest that the visual system can divide a scene into exactly two parts: selected elements and excluded elements. This is the "boolean map" that underlies their theory. The visual system can then access certain properties of the selected elements in the map. Once a boolean map is created, two properties are available to a viewer: the label for any feature in the map, and the spatial location of the selected elements (Fig. 3). Boolean maps can be created in two ways. First, a viewer can specify a single value of an individual feature to select all objects that contain that feature. Second, union or intersection can be applied to two existing maps. In either case, only the result is retained, since evidence suggests that a viewer can only hold and access one boolean map at a time. Viewers can chain these operations together to search for targets in a fairly complex scene.

2.3.5 Ensemble Coding

Existing characterizations of preattentive vision have focused on how low level-visual processes can be used to guide attention to specific location or object in a larger scene. An equally important characteristic of low-level visual processes is their ability to generate a quick summary of how simple visual features are distributed across the field of view. The ability of humans to register a rapid and in-parallel summary of a scene in terms of its simple features was first reported by Ariely [4]. He demonstrated that observers could extract the average size of a large number of dots from only a single glimpse at a display. Yet, when observers were tested on the same displays and asked to indicate whether a single dot of a given size was present, they were unable to do so. This suggests that there is a preattentive mechanism that records summary statistics of visual features without retaining information about the constituent elements that generated the summary.

This ability to rapidly identify scene-based averages may offer important advantages in certain visualization environments. For example, given a stream of real-time data, ensemble coding would allow viewers to observe the stream at a high frame rate, yet still identify individual frames with interesting distributions of visual features (i.e. attribute values). Ensemble coding would also be critical for any situation where viewers want to estimate the amount of a particular data attribute in a display. These capabilities were hinted at in a paper by Healey et al., but without the benefit of ensemble coding as a possible explanation.

2.3.6 Feature Hierarchies

One promising strategy for multidimensional visualization is to assign different visual features to different data attributes. This allows multiple data values to be shown simultaneously in a single image. A key requirement of this method is a data-feature mapping that does not produce visual interference. One example of interference is a conjunction target. Another example is the presence of feature hierarchies that appears to exist in the visual system. For certain tasks one visual feature may be “more salient” than another. Researches in psychophysics and visualization have demonstrated a hue-shape hierarchy: the visual system favors color over shape [15, 16, 17, 48, 49]. Background variations in hue interfere with a viewer‘s ability to identify the presence of individual shapes and the spatial patterns they form. If hue is held constant across the display, these same shape patterns are immediately visible. The interference is asymmetric: random variations in shape have no effect on a viewer‘s ability to see color patterns. Similar luminance-hue and hue-texture hierarchies have also been identified.

2.4 Visual Memory

Preattentive processing asks in part: "What visual properties draw our eyes, and therefore our focus of attention to a particular object in a scene?" An equally interesting question is: "What do we remember about an object or a scene when we stop attending to it and look at something else?" Many viewers assume that as we look around us we are constructing a high-resolution, fully detailed description of what we see. Researchers in psychophysics have known for some time that this is not true. In fact, in many cases our memory for detail between glances at a scene is very limited. Evidence suggests that a viewer's current state of mind can play a critical role in determining what is seen and what is not.

We present three theories that demonstrate and attempt to explain this phenomena: change blindness, inattentional blindness, and attentional blink. Understanding what we remember as we focus on different parts of a visualization is critical to designing visualizations that encourage locating and retaining the information that is most important to the viewer.

2.4.1 Change Blindness

New research in psychophysics has shown that an interruption in what is being seen—a blink, an eye saccade, or a blank screen—renders us "blind" to significant changes that occur in the scene during the interruption [104, 69, 105, 110]. This change blindness phenomena can be illustrated using a task similar to one shown in comic strips for many years. A viewer is shown two pairs of images. A number of significant differences exists between the images. Many viewers have a difficult time seeing any difference and often have to be coached to look carefully to find it. Once they discover it, they realize that the difference was not a subtle one. Change blindness is not a failure to see because of limited visual acuity; rather, it is a failure based on inappropriate attentional guidance. Some parts of the eye and the brain are clearly responding differently to the two pictures. Yet, this does not become part of our visual experience until attention is focused directly on the objects that vary.

The presence of change blindness has important implications for visualization. The images we produce are normally novel for our viewers, so prior expectations cannot be used to guide their analyses. Instead, we strive to direct the eye, and therefore the mind, to areas of interest or importance within a visualization. This ability forms the first step towards enabling a viewer to abstract details that will persist over subsequent images.

2.4.2 Inattentional Blindness

A related phenomena called inattentional blindness suggests that viewers fail to perceive objects or activities that occur outside of the focus of attention [69]. This phenomena is illustrated through an experiment conducted by Neisser [90, 109]. His experiment superimposed video streams of two basketball games. Players wore white shirts in one stream and black shirts in the other. Subjects

(a)

(b)

Figure 4: Change blindness, a major difference exists between the two images

attended to one team—either white or black—and ignored the other. Whenever the subject's team made a pass, they were told to press a key. After about 30 seconds of video, a third stream was superimposed showing a woman walking through the scene with an open umbrella. The stream was visible for about 4 seconds, after which another 25 seconds of basketball video was shown. Following the trial, only six of twenty-eight naive observers reported seeing the woman. When subjects only watched the screen and did not count passes, 100% noticed the woman.

Additional issues with relevance to visualization are also being investigated. Most et al. are studying the relationship between inattentional blindness and attentional capture, the ability of an object to draw the focus of attention without a viewer's active participation. Researchers are also studying how perceptual load affects inattentional blindness. Finally, results suggest meaningful objects (e.g., a person's name or a happy face icon) may be easier to notice.

2.4.3 Attentional Blink

In each of the previous methods for studying visual attention, the primary emphasis is on how human attention is limited in its ability to represent the details of a scene (change blindness) and in its ability to represent multiple objects at the same time (inattentional blindness). But attention is also severely limited in its ability to process information that arrives in quick succession, even when that information is presented at a single location in space. The attentional blink paradigm is currently the most widely used method to study the availability of attention across time. Its name—"blink"—derives from the finding that when two targets are presented in rapid succession, the second of the two targets cannot be detected or identified when it appears within approximately 100–500 msec following the first target [14, 101]. This suggests that that attention operates over time like a window or gate, opening in response to finding a visual item that matches its current criterion or template and then closing shortly thereafter to consolidate that item as a distinct object or event from others. The attentional blink is an index of the "dwell-time" needed to consolidate a rapidly presented visual item into visual short term memory.

2.5 Conclusions

This presentation surveys past and current theories of low-level visual perception and visual attention. Initial work in preattentive processing identified basic visual features that can implicitly or explicitly capture a viewer's focus of attention. More recent work has extended this to study limited visual memory for change—change blindness and attentional blink—and being "blind" to objects that are outside the focus of attention—inattentional blindness. Each of these phenomena have significant consequences for visualization. We strive to produce images that are salient and memorable, and that guide attention to locations of importance within the data. Understanding what the visual seems sees and does not see is critical to designing effective visual displays.

Ann McNamara

3 Perceptually Motivated Rendering

3.1 Visual Perception in Realistic Image Synthesis

Realism is often a primary goal in computer graphics imagery. We strive to create images that are perceptually indistinguishable from an actual scene. Rendering systems can now closely approximate the physical distribution of light in an environment. However, physical accuracy does not guarantee that the displayed images will have an authentic visual appearance. In recent years the emphasis in realistic image synthesis has begun to shift from the simulation of light in an environment to images that look as real as the physical environment they portray. In other words the computer image should be not only physically correct but also perceptually equivalent to the scene it represents. This implies aspects of the Human Visual System (HVS) must be considered if realism is required. Visual perception is employed in many different guises in graphics to achieve authenticity [92, 7]. Certain aspects of the HVS must be considered to identify the perceptual effects that a realistic rendering system must achieve in order to effectively reproduce a similar visual response to a real scene. This section outlines the main characteristics of the HVS and the manner in which knowledge about visual perception is increasingly appearing in state-of-the-art realistic image synthesis. Perception driven rendering algorithms are described, which focus on embedding models of the HVS directly into global illumination computations in order to improve their efficiency.

3.1.1 Visual Perception

Perception is the process by which humans, and other organisms, interpret and organize sensation in order to understand their surrounding environment. Sensation refers to the immediate, relatively unprocessed result of stimulation of sensory receptors. Perception, on the other hand, is used to describe the ultimate experience and interpretation of the world and usually involves further processing of sensory input. Sensory organs translate physical energy from the environment into electrical impulses processed by the brain. In the case of vision light, in the form of electromagnetic radiation, activates receptor cells in the eye triggering signals to the brain. These signals are not understood as pure energy, rather, perception allows them to be interpreted as objects, events, people and situations.

3.1.2 The Human Visual System

Vision is a complex process that requires numerous components of the human eye and brain to work together. Vision is defined as the ability to see the features of objects we look at, such as color, shape, size, details, depth, and contrast. Vision begins with light rays bouncing off the surface of objects. These reflected light rays enter the eye and are transformed into electrical signals. Millions of signals per second leave the eye via the optic nerve and travel to the visual area of the brain. Brain cells then decode the signals providing us with sight. The response of the human eye to light is a complex, still not well understood process. It is difficult to quantify due to the high level of interaction between the visual system and complex brain functions. A sketch of the anatomical components of the human eye is shown in Figure 5 The main structures are the iris, lens, pupil, cornea, retina, vitreous humor, optic disk and optic nerve.

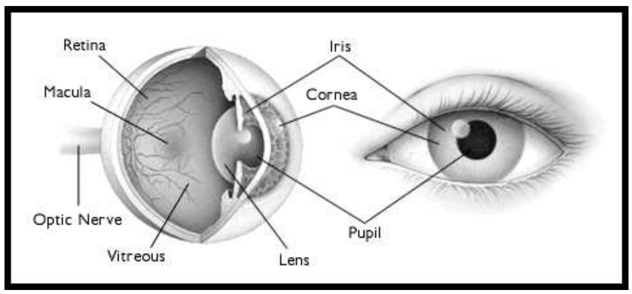


Figure 5: Cross section of the human eye

The path of light through the visual system begins at the pupil, is focused by the lens, then passes onto the retina, which covers the back surface of the eye. The retina is a mesh of photoreceptors, which receive light and pass the stimulus on to the brain. The internal structure of the human eye, a sphere, typically 12mm in radius, is enclosed by a protective membrane, the sclera. At the front of the sclera lies the cornea, a protruding opening, and an optical system comprising the lens and ciliary muscles which change the shape of the lens providing variable focus. Light enters the eye though the lens and proceeds through the vitreous humor, a transparent substance, to the rear wall of the eye, the retina. The retina has photoreceptors coupled to nerve cells, which intercept incoming photons and output neural signals. These signals are transmitted to the brain through the optic nerve, connected to the retina at the optic disk or papilla, more commonly known as the blind spot. The retina is composed of two major classes of receptor cells known as rods and cones. The rods are extremely sensitive to light and provide achromatic vision at low (scotopic) levels of illumination. The cones are less sensitive than the rods but provide color vision at high (photopic) levels of illumination. A schematic drawing of rod and cone cells is shown in Figure 5. Cones are nerve cells that are sensitive to light, detail, and color. Millions of cone cells are packed into the macula, aiding it in providing the visual detail needed to scan the letters on an eye chart, see a street sign, or read the words in a newspaper. Rods are designed for night vision. They also provide peripheral vision, but they do not see as acutely as cones. Rods are insensitive to color. When a person passes from a brightly lit place to one that is dimly illuminated, such as entering a movie theatre during the day, the interior seems very dark. After some minutes this impression passes and vision becomes more distinct. In this period of adaptation to the dark, the eye becomes almost entirely dependent on the rods for vision, which operate best at very low light levels. Since the rods do not distinguish color, vision in dim light is almost colorless.

Cones provide both luminance and color vision in daylight. They contain three different pigments, which respond either to blue, red, or green wavelengths of light. A person who is missing one or more of the pigments is said to be color-blind and has difficulty distinguishing between certain colors, such as red from green. These photoreceptor cells are connected to each other and the ganglion cells which transmit signals to and from the optic nerve. Connections are achieved via two layers, the first and second synaptic layers. The interconnections between the rods and cones are mainly horizontal links, indicating a preferential processing of signals in the horizontal plane.

Normal daytime vision, where the cones predominate visual processing, is termed photopic, whereas low light levels where the rods are principally responsible for perception is termed scotopic vision. When both rods and cones are equally involved then vision is termed mesopic. Visual acuity is the ability of the Human Visual System (HVS) to resolve detail in an image. The human eye is less sensitive to gradual and sudden changes in brightness in the image plane but has higher sensitivity to intermediate changes. Acuity decreases with increase in distance. Visual acuity can be measured using a Snellen Chart, a standardized chart of symbols and letters. Visual field indicates the ability of each eye to perceive objects to the side of the central area of vision. A normal field of vision is 180 degrees .

3.1.3 Contrast

Contrast is defined as:

$$lmax = \frac{lmax - lmin}{lmax + lmin} \tag{1}$$

where lmax and lmin are the maximum and minimum luminance. Human brightness sensitivity is logarithmic, so it follows that for the same perception, higher brightness requires higher contrast. Apparent brightness is dependent on background brightness. This phenomenon, termed simultaneous contrast, is illustrated in Figure 6. Despite the fact that all centre squares are the same brightness, they are perceived as different due to the different background brightness.

Depth Perception is the ability to see the world in three dimensions and to perceive distance. Images projected onto the retina are two-dimensional, and from these flat images vivid three dimensional worlds are constructed. Binocular Disparity and monocular cues provide information for depth perception. Binocular disparity is the difference between the images projected onto the left and right eye. The brain integrates these two images into a single three dimensional image to allow depth and distance perception. Monocular cues are cues to depth that are effective when viewed with only one eye, including interposition,

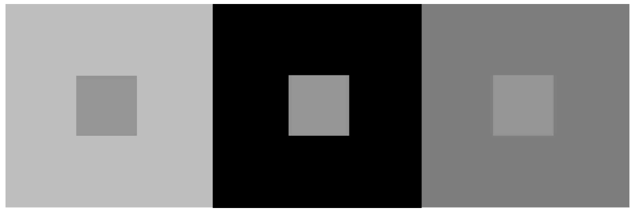


Figure 6: Simultaneous Contrast. Despite the fact that all centre squares are the same brightness, they are perceived as different due to the different background brightness.

atmospheric perspective, texture gradient, linear perspective, size cues, height cues and motion parallax.

3.1.4 Constancy

Perceptual Constancy is a phenomenon which enables the same perception of an object despite changes in the actual pattern of light falling on the retina. Psychologists have identified a number of perceptual constancies including lightness constancy, color constancy, size constancy and shape constancy. Lightness Constancy: The term lightness constancy describes the ability of the visual system to perceive surface lightness correctly despite changes in the level of illumination.
Color Constancy: Closely related to lightness constancy, this is the ability of the HVS to perceive the correct color of an object despite changes in illumination.
Shape Constancy: Objects are perceived as having the same shape regardless of changes in their orientation. -example with cube, from front and side
Size Constancy: This is the tendency to perceive objects as staying the same size despite changes in viewing distance.

3.1.5 Human Visual Perception

A number of psychophysical experimental studies have demonstrated many features of how the HVS works. However, problems arise when trying to generalize these results for use in computer graphics. This is because, often, experiments are conducted under limited laboratory conditions and are typically designed to explore a single dimension of the HVS. As described earlier, the HVS comprises complex mechanisms, which rather than working independently, often features work together, and therefore it makes sense to examine the HVS as a whole. Instead of reusing information from previous psychophysical experiments, new experiments are needed. Some examples will support this. Figure 2.6: When a black and white patterned top shown on the left is rotated at 5-10 revolutions per second, colored rings can be seen. The light intensity distribution of the rotating pattern as a function of time is shown on the right. Spatiotemporal interactions between antagonistic, spectrally opponent color mechanisms account for this phenomenon. A Benhams disk is a flat disc, half of which is black and the other half has three sets of lines like the grooves on a record but more spaced out, Figure 7. When the disk is spun a human observer sees red, yellow and green rings, despite the fact that there are no colors in the pattern. The curves on the right of the pattern begin to explain what happens. Each curve plots the temporal light intensity distribution at the different radii from the centre, created when the top is spun. These changing light patterns produce spatiotemporal interaction in the HVS that unbalance antagonistic, spectrally-opponent mechanisms to create the appearance of colored rings. This illusion demonstrates that, although it may be convenient to model the HVS in terms of unidimensional responses to motion, pattern and color, human percepts are in fact the product of complex multidimensional response.

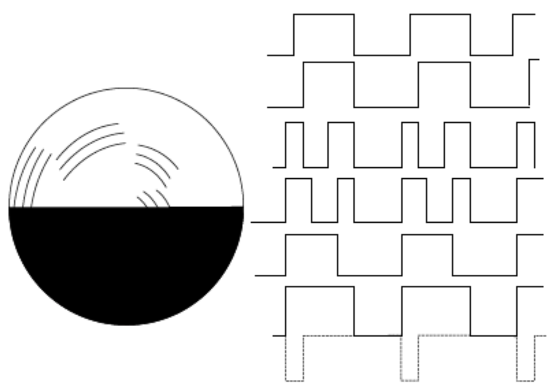


Figure 7: When a black and white patterned top shown on the left is rotated at 5-10 revolutions per second, colored rings can be seen. The light intensity distribution of the rotating pattern as a function of time is shown on the right. Spatiotemporal interactions between antagonistic, spectrally opponent color mechanisms account for this phenomenon.

A second example, Figure 8, shows the panels in checkerboard block on the left and a flat pattern on the right, which have the same reflectance, but differences in their three-dimensional organization means they are perceived differently. The two panels marked with Xs have the same reflectance, but on the block they appear to have different reflectance under different levels of illumination. Conversely, the two panels marked with Os have different reflectance values but on the block appear to be the same color due to the different illumination conditions. This demonstrates the complexity of interactions between apparent reflectance, apparent illumination and apparent shape that can dramatically affect human perception.

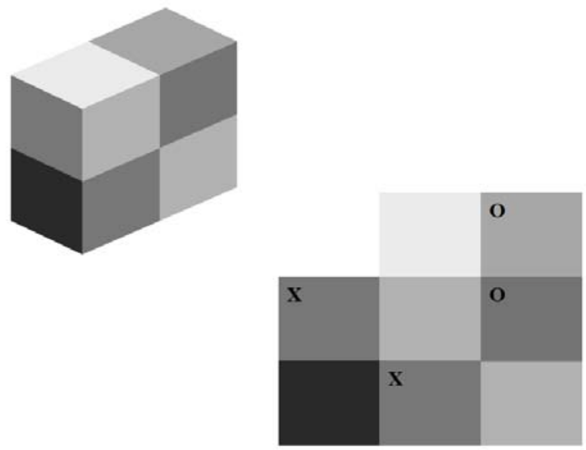


Figure 8: Interaction between apparent reflection, apparent illumination and apparent three-dimensional shape. Corresponding panels in the two patterns have the same physical reflectance. Differences in the perceived spatial organization of the patterns produces differing interpretations in terms of lightness (apparent reflectance) and brightness (apparent illumination).

3.1.6 Lightness Perception

Gilchrist [43, 44, 18] justified the systematic study of lightness error as an understanding of the HVS. He found that there are always errors when judging lightness, and these errors are not random, but systematic. The pattern of these systematic errors therefore provide a signature of the visual system. He defines a lightness error as any difference between the actual reflectance of a target surface and the reflectance of the matching chip selected from a Munsell chart. The task defined for the psychophysical experiments described later in this thesis involves asking human observers to match the reflectance of real world objects to a Munsell chart, which gives a measure of errors in lightness matching. The observer is then asked to match the reflectance of simulated objects (in a computer generated rendition of the real world) to the same Munsell chart. This gives a measure of lightness errors with respect to the computer image. There are limitations on the HVS, so there will be errors (systematic errors) in both cases. For the rendered image to be deemed a faithful representation, both sets of lightness errors should be close to each other.

Gilchrist (1977) [45] showed that the perception of the degree of lightness of a surface patch (i.e. whether it is white, gray or black) is greatly affected by the perceived distance and orientation of the surface in question, as well as the perceived illumination falling on the surface -where the latter was experimentally manipulated through a variety of cues such as occlusion, or perspective. Perception of the lightness of patches varying in reflectance may thus be a suitable candidate for the choice of visual task. It is simple to perform, and it is known that lightness constancy depends on the successful perception of lighting and the 3D structure of a scene, for example Figure 9 When viewed in isolation, the patches on the top left hand corner appear to be of different luminance. However, when examined in the context of the entire scene, it can be seen that the patches have been cut from the edge of the stairwell, and are perceived as an edge where the entire stairwell has the same luminance. Lightness has been applied when developing Tone Mapping techniques for High Dynamic Range Imagery [61, 62].

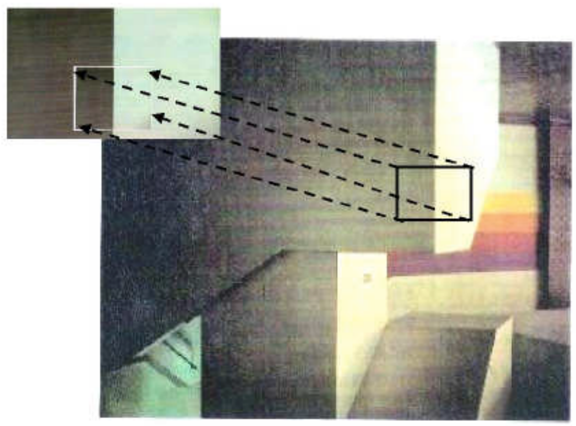


Figure 9: Importance of depth perception for lightness constancy.

3.2 Perceptually driven rendering

Recent years have seen an increase in the application of visual perception to computer graphics. As mentioned earlier, in certain applications it is important that computer images should not only be physically correct but also perceptually equivalent to the scene it is intended to represent. Realism implies computational expense, and research is beginning to emerge to investigate how knowledge of the human visual system can be used to cut corners and minimize rendering times by guiding algorithms to compute only what is necessary to satisfy the observer. Perception based image quality metrics, which can be used to evaluate, validate and compare imagery have been presented [99, 97, 80, 76, 86].

Even for realistic image synthesis there may be little point spending time or resources to compute detail in an image that would not be detected by a human observer. By eliminating any computation spent on calculating image features which lie below the threshold of visibility, rendering times can be shortened leading to more efficient processing. Because the chief objective of physically based rendering is realism, incorporating models of HVS behavior into rendering algorithms can improve performance, as well as improving the quality of the imagery produced. So by taking advantage of the limitations of the human eye, just enough detail to satisfy the observer can be computed without sacrificing image quality. Several attempts have been made to develop image synthesis algorithms that detect threshold visual differences and direct the algorithm to work on those parts of an image that are in most need of refinement.

Raytracing produces an image by computing samples of radiance, one for each pixel in the image plane. Producing an anti-aliased image is difficult unless very high sampling densities are used. Mitchell [85] realized that deciding where to do extra sampling can be guided by knowledge of how the eye perceives noise as a function of contrast and color. Studies have shown that the eye is most sensitive to noise in intermediate frequencies [124]. While frequencies of up to 60 cycles per degree (cpd) can be visible, the maximum response to noise is at approximately 4.5 cpd, so sampling in regions with frequency above this threshold can be minimized, without affecting the visual quality of the image. Mitchell begins by sampling the entire image at low frequency then uses an adaptive sample strategy on the image according to the frequency content. This results in a non uniform sampling of the image, which enables aliasing noise to be channelled into high frequencies where artifacts are less conspicuous. However, non-uniform sampling alone doesnt eliminate aliasing, just changes its characteristics to make it less noticeable. Mitchell applies two levels of sampling. To decide whether the high sampling density should be invoked the variance of samples could be used [89], but this is a poor measure of visual perception of local variation. Instead Mitchell chooses to use contrast to model the non-linear response of the eye to rapid variations in light intensity:

As each sample consists of three separate intensities for red, green and blue, three separate contrasts can be computed for each of them. These three contrasts are tested against separate thresholds, 0.4, 0.3 and 0.6 for red, green and blue respectively, and super-sampling is done if any one exceeds the threshold. The contrast metric is then used to determine when the high sampling density should be invoked. This test is most sensitive to green in accordance with the human eyes response to noise as a function of color. Multi stage filters are then used to reconstruct the non-uniform samples into a digital image. Although this idea has the beginnings of a perceptual approach, it is at most a crude approximation to the HVS. Only two levels of sampling are used and it doesnt account for visual masking [1].

The HVS exhibits different spatial acuities in response to different colors. Evidence exists that color spatial acuity is less than monochrome spatial acuity. Exploiting this poor color spatial acuity of the HVS, Meyer and Liu [84] developed an adaptive image synthesis algorithm which uses an opponents processing model of color vision [61] comprising chromatic and achromatic color channels. Using a Painter and Sloan [122] adaptive subdivision, a k-D [2] tree representation 3 of the image is generated. Areas of the image containing high frequency

[1] The presence of high spatial frequency in an image can mask the presence of other high frequency information

[2] A KD Tree is a data structure that is used in computer science during orthogonal range searching

information are stored at the lower levels of the tree. They then modified a screen subdivision raytracer to limit the depth to which the k-D tree must be descended to compute the chromatic color channels. The limit is determined by psychophysical results describing the color spatial frequency. They achieved a modest saving in computational effort and showed, using a psychophysical experiment, that decreasing the number of rays used to produce the chromatic channels had less of an effect on image quality than reducing the number of rays used to create the achromatic channels. This was the first work to attempt to minimize the computation of color calculations, as opposed to just decreasing costly object intersection calculations.

Bolin and Meyer [9] took a frequency based approach to raytracing, which uses a simple vision model, making it possible for them to control how rays are cast in a scene. Their algorithm accounts for the contrast sensitivity, spatial frequency and masking properties of the HVS. The contrast sensitivity response of the eye is non-linear. So, when deciding where rays should be cast, the algorithm deems a luminance difference at low intensity to be of greater importance than the same luminance difference at high intensity. The spatial response of the HVS is known to be less for patterns of pure color than for patterns that include luminance differences. This means that it is possible to cast fewer rays into regions with color spatial variations than are cast in regions with spatial frequency variations in luminance. Finally, it is known that the presence of high spatial frequency can mask the presence of other high frequency information (masking). When used in conjunction with a Monte Carlo raytracer, more rays are spawned when low frequency terms are being determined than when high frequency terms are being found. Using this strategy, the artifacts that are most visible in the scene can be eliminated from the image first, then noise can be channelled into areas of the image where artifacts are less conspicuous. This technique is an improvement on Mitchells method because the vision model employed accounts for contrast sensitivity, spatial frequency and masking.

Despite the simplicity of the vision models used in these approaches, the results are promising, especially as they demonstrate the feasibility of embedding HVS models into the rendering systems to produce more economical systems without forfeiting image quality. Fueled by the notion that more sophisticated models of the HVS would yield even greater speedup, several researchers began to introduce more complex models of the HVS into their global illumination computations.

Myszkowski [88] applied a more sophisticated vision model to steer computation of a Monte Carlo based raytracer. Aiming to take maximum advantage of the limitations of the HVS, his model included threshold sensitivity, spatial frequency sensitivity and contrast masking. A perceptual error metric is built into the rendering engine allowing adaptive allocation of computation effort into areas where errors remain above perceivable thresholds and allowing computation to be halted in all other areas (i.e. those areas where errors are below the perceivable threshold and thus not visible to a human observer). This perceptual error metric takes the form of Dalys [25] Visible Difference Predictor (VDP).

Bolin and Meyer [10] devised a similar scheme, also using a sophisticated vision model, in an attempt to make use of all HVS limitations. They integrated a simplified version of the Sarnoff Visible Discrimination Model (VDM) into an image synthesis algorithm to detect threshold visible differences and, based on those differences direct subsequent computational effort to regions of the image in most need of refinement. The VDM takes two images, specified in CIE XYZ color space, as input. Output of the model is a Just Noticeable Difference (JND) map. One JND corresponds to a 75% probability that an observer viewing the two images would detect a difference [81]. They use the upper and lower bound images from the computation results at intermediate stages and used the predictor to get an error estimate for that stage. Drettakis et al introduced a perceptual rendering pipeline which takes into account visual masking due to contrast and spatial frequency[29]. Scenes are split into layers to account for inter-object masking. Using a perceptually driven level of detail algorithm the layers are then used to choose an appropriate level of detail for each object based on predicted contrast and and spatial masking. A subsequent user study showed that their algorithmic choices corresponded well with perceived differences in the images. Masking has also been used in geometric modeling, Lavoué et. al. [65] introduced the notion of *roughness* for a 3D mesh. Roughness gives a measure of geometric noise on the surface, based on this noise masking can be invoked to hide geometric distortions.

Applying a complex vision model at each consecutive time step of image generation requires repeated evaluation of the embedded vision model. The VDP can be expensive to process due to the multi-scale spatial processing involved in some of its components. This means that in some cases the cost of recomputing the vision model may cancel the savings gained by employing the perceptual error metric to speed up the rendering algorithm. To combat this, Ramasubramanian [100] introduced a metric that handles luminance-dependent processing and spatially-dependent processing independently, allowing the expensive spatially-dependent component to be precomputed. Ramasubramanian developed a physical error metric that predicts the perceptual threshold for detecting artifacts in the image. This metric is then used to predict the sensitivity of the HVS to noise in the indirect lighting component. This enables a reduction in the number of samples needed in areas of an image with high frequency texture patterns, geometric details, and direct lighting variations, giving a significant speedup in computation.

Using validated image models that predict image fidelity, programmers can work towards achieving greater efficiencies in the knowledge that resulting images will still be faithful visual representations. Also in situations where time or resources are limited and fidelity must be traded off against performance, perceptually based error metrics could be used to provide insights into where computation could be economized with least visual impact.

In addition to Tone Mapping Operators (TMOs) being useful for rendering calculated luminance to the screen [120, 63, 26, 103], they are also useful for giving a measure of the perceptible difference between two luminances at a given level of adaptation. This function can then be used to guide algorithms, such as discontinuity meshing, where there is a need to determine whether some process would be noticeable or not to the end user.

Gibson and Hubbold [42] have used features of the threshold sensitivity displayed by the HVS to accelerate the computation of radiosity solutions. A perceptually based measure controls the generation of view independent radiosity solutions. This is achieved with an a priori estimate of real-world adaptation luminance, and uses a TMO to transform luminance values to display colors and is then used as a numerical measure of their perceived difference. The model stops patch refinement once the difference between successive levels of elements becomes perceptually unnoticeable. The perceived importance of any potential shadow falling across a surface can be determined, this can be used to control the number of rays cast during visibility computations. Finally, they use perceptual knowledge to optimize the element mesh for faster interactive display and save memory during computations. This technique was used on the adaptive element refinement, shadow detection, and mesh optimization portions of the radiosity algorithm.

Discontinuity meshing is an established technique used to model shadows in radiosity meshes. It is computationally expensive, but produces meshes which are far more accurate and which also contain fewer elements. Hedley et al. [50] used a perceptually informed error metric to optimize adaptive mesh subdivision for radiosity solutions, the goal being to develop scalable discontinuity meshing methods by considering visual perception. Meshes were minimized by discarding discontinuities which had a negligible perceptible effect on a mesh. They demonstrated that a perception-based approach results in a greater reduction in mesh complexity, without introducing more visual artifacts than a purely radiometrically-based approach.

Farrugia and Peroche [36] used a perceptual metric for discontinuity refinement to develop a progressive radiance evaluation based on the work of Guo et. al. [46]. Guo used an iterative process to construct an irregular subdivision of the image in blocks which refer to smooth regions, or discontinuous regions to build a Directional Coherence Map (DCM). As the algorithm proceeds the current DCM dictates where new samples are taken. A contrast based perceptual heuristic based on contrast over samples corresponding to the corners of each block is used. Farrugia and Peroch extend this by applying a visual differences predictor based on Pattanaik et al.s Multiscale Model of Adaptation and Spatial Vision to classify their subdivision cells [94]. To speed up computation they apply the metric over each cell pair using a statistical approach based on Albin et. al. [2].

Recognizing that the illumination on a surface can be split into separable components, which can be individually computed, Stokes et. al [113] introduced a new approach which applied a perceptual metric on each component. After a suite of psychophysical experiments to probe various global illumination scenarios they determined limited contribution of light path interaction and fitted a mathematical model used to guide rendering based on the metric predicted relative importance of each component as a function of visible surface materials.

Ramanarayanan et. al. noticed that when viewing an aggregate, observers attend less to individual objects and focus more on overall properties such as numerosity, variety, and arrangement. They also noted that rendering and modeling costs increase with aggregate complexity, exactly when observers are attending less to individual objects. They presented new aggregate perception metrics to simplify scenes by substituting geometrically simpler aggregates for more complex ones without changing appearance [98].

Ramanarayanan [99] worked toward developing a perceptual metrics based on higher order aspects of visual coding and introduced the term "Visual equivalence". Images are visually equivalent if they convey the same impressions of scene appearance, even if they are visibly different. They conducted a series of psychophysical experiments to investigate how object geometry, material, and illumination interact influence appearance. In their paper they characterized conditions under which two classes of transformations on illumination maps (blurring and warping) yield images that are visually equivalent to reference solutions, and from this developed a metric to predict visual equivalence.

3.3 Conclusion

Using validated image models that predict image fidelity, programmers can work toward achieving greater efficiencies in the knowledge that resulting images will still be faithful visual representations. Also in situations where time or resources are limited and fidelity must be traded off against performance, perceptually based error metrics could be used to provide insights into where computation could be economized with least visual impact.

Some of the applications of visual perception in computer graphics were explored. For many applications computer imagery should not only be physically correct but also perceptually equivalent to the scene it represents. Knowledge of he HVS can be employed to greatly benefit the synthesis of realistic images at various stages of production. Global illumination computations are costly in terms of computation. There is a great deal of potential to improve the efficiency of such algorithms by focusing computation on the features of a scene which are more conspicuous to the human observer. Those features that are below perceptual visibility thresholds have no impact on the final solution, and therefore can be omitted from the computation, increasing efficiency without causing any perceivable difference to the final image. Perceptual metrics involving advanced HVS models can be used to determine the visible differences between a pair of images. These metrics can then be used to compare and evaluate image quality. They can also be used within the rendering framework to steer computation into regions of an image which are in most need of refinement, and to halt computation when differences in successive iterations of the solution become imperceptible.

Katerina Mania

4 Perceptually Motivated Simulation and Virtual Environments

4.1 Introduction

Computer graphics algorithms have for long dealt with simulation of physics: simulation of the geometry of a real-world space, simulation of the light propagation in a real environment and simulation of motor actions with appropriate tracking. Perception principles have subsequently been incorporated into rendering algorithms [82], in order to save rendering computation, mainly following the generic idea of "do not render what we cannot see" [77, 54, 68]. However, with Virtual Environment (VE) simulator technologies aiming at simulating real-world task situations, the research community is challenged to produce a much more complex system which is perceptually optimized. We do not necessarily require accurate simulation of physics to induce reality. Much less detail is often adequate [127, 37, 72, 73].

4.2 Perceptually-based Selective Rendering

Perception principles have been incorporated into rendering algorithms in order to optimize rendering computation and produce photorealistic images from a human rather than a machine point of view. In order to economize on rendering computation, selective rendering guides high level of detail to specific regions of a synthetic scene and lower quality to the remaining scene, without compromising the level of information transmitted. Scene regions that have been rendered in low and high quality can be combined to form one complete scene. Such decisions are guided by predictive attention modeling, gaze or task-based information.

In order to economize on rendering computation, previous research dealing with interactive synthetic scenes has been focused on either rendering in high quality the 2-3 degrees foveal region of vision and with less detail the periphery of vision based on gaze information [70], or rendering in high quality the foveal area based on a-priory knowledge of the viewers task focus [19, 114]. Gaze-dependent rendering encounters difficulties of maintaining display updates free of visual artifacts after a fast (4ms) eye saccade. Such processes are quite computationally demanding, however, if the speed gaze-to-rendering issue is resolved, task performance results are indistinguishable to a fully fledged, high resolution real-time environment. It has also been proposed to assign selective high quality rendering in the visual angle of the fovea (2o) centered on the users task focus [19, 114]. This approach, however, cannot be applied when there is no overt task to be conducted. Moreover, there is no acceptable model of comparing or predicting task-relevant saccades. Following a different approach, Haber et al. [47] suggested rendering the informative areas of a scene in varying quality based on saliency models. Such models aim to predict the visual features that involuntarily attract visual attention such as object edges, sudden color changes or movements. It was proposed that the most noticeable areas as derived from saliency modeling should be rendered in higher quality. Bottom-up visual attention models are not shown to predict attention regions successfully [77]. Correlation between actual human and computationally-derived scan-paths was found to be much lower than predicted when carrying out a real-world task such as making a cup of tea [35]. Moreover, we have no generally accepted model of comparing scan paths.

A comprehensive approach should be task and gaze-independent, simulating cognitive processes rather than predicting attention employing bottom-up processes such as saliency models. A recent selective rendering approach exploits existing research on memory schemata which could ultimately guide selective rendering based on spatial cognition processes. Schemata are knowledge structures based on the notion that an individuals prior experience will influence how he or she perceives, comprehends and remembers new information. When participants are exposed to a large amount of information in a scene, cognitive psychologists have suggested that schemata are used to guide the search for information in memory [13]. A general premise derived from this research is that information which is not related to the schema being used in retrieval will be harder to recall than information which is schema related. In terms of real world scenes, schemata represent the general meaning of a scene such as office, theatre etc. Schemata influence memory of the objects in a given context according to their association with the schema in place. When being exposed to a synthetic environment, similar information should be transmitted between the simulated scene and the real- world scene, both depicting a specific schema. This would, in due course, indicate which objects or areas in a synthetic scene could be rendered in lower quality without affecting information uptake but at the same time reducing computational complexity [87].

Flannery and Walles [39] investigated how schema theories apply to real versus virtual memories. Participants were instructed to explore either a virtual or a similar real environment for 20 seconds, without prior knowledge that their memory of the space would be subsequently assessed. Participants then completed a recognition task. Recognition scores revealed that participants had better recognition for consistent objects, but were more confident for the recognition of the inconsistent objects.

Previous work [71, 72], included a preliminary investigation of the effect of object type (consistent vs. inconsistent) and shadows (flat-shaded scene vs. radiosity scene) on object memory recognition in a VE. The computer graphics simulation was displayed on a Head Mounted Display (HMD) utilizing stereo imagery and head tracking. Thirty-six participants across three conditions of varied rendering quality of the same space were exposed to the computer graphics environment and completed a memory recognition task. The high-quality and mid-quality conditions included a pre-computed radiosity simulation of an academics office (with 80% and 40% radiosity iterations computed respectively). The low-quality condition consisted of a flat-shaded version of the same office.

Results revealed that schema consistent elements of the scene were more likely to be recognized than inconsistent information. Overall, higher confidence ratings were assigned to consistent rather than inconsistent items. Total object recognition was better for the scene including rough shadows (mid-quality condition) compared to the flat-shaded scene. The presence of accurate shadow information, though, did not affect recognition of consistent or inconsistent objects, therefore lower quality of rendering was adequate for better memory recognition of consistent objects. This study was limited to the investigation of subtle shadow variations. Another experimental study employed a more extreme set of rendering types: wireframe with added color, and full radiosity [87] (Figure 3). The proportion of inconsistent/consistent objects was varied, and object recognition tests ensured that all objects were easily recognized in all conditions. The results showed a significant interaction between rendering type, object type, and consistency ratio. This suggests that inconsistent objects are only preferentially remembered if the scene looks normal or if there are many such objects in an abnormal scene such as in the wireframe condition. It was also shown that memory performance is better for the inconsistent objects in the radiosity rendering condition compared to the wireframe condition. We conclude that memory for objects can be used to assess the degree to which the context of a VE appears close to expectations.

Despite contradictory results in literature as detailed above, it seems that perceptual information can be complemented by involuntary knowledge based on past experience. Experimental studies in synthetic scenes have revealed that consistent objects which are expected to be found in a scene can be rendered in lower quality without affecting information uptake taking advantage of such expectations, whereas inconsistent items which are salient would require a high level of rendering detail in order for them to be perceptually acknowledged [71]. Therefore, by exploiting schema theory, it is possible to reduce computational complexity, producing scenes from a cognitive point of view without affecting information uptake and resulting in an entirely novel and interdisciplinary approach which is gaze, task and saliency-model independent. A novel selective rendering system has been presented that exploits schema theory by identifying the perceptual importance of scene regions [135]. Objects that have been rendered in low and high quality are incorporated in a scene based on schema expectations. The rendered quality of these objects will change in real time, dependent on user navigation and interaction [87].

High level visual cognition is that which takes place late in the HVS, that is parietal and temporal cortex and into the frontal lobes when decisions based on visual information need to be made. Low level visual cognition is that which takes place in the occipital lobe, early on in the visual processing stream, e.g. the visual signal is received in the retinae, and initially passed through the Lateral Geniculate Nucleus to the occipital lobe at the back. Thus, these higher decision processes are unaffected by changes in the experiments, whilst normal (visual) cognition occurs as long as the scene is realistic. Taken together, the results of previous studies investigating the effect of schemas on object recognition suggest that high-level visual cognition is generally unaffected by ubiquitous

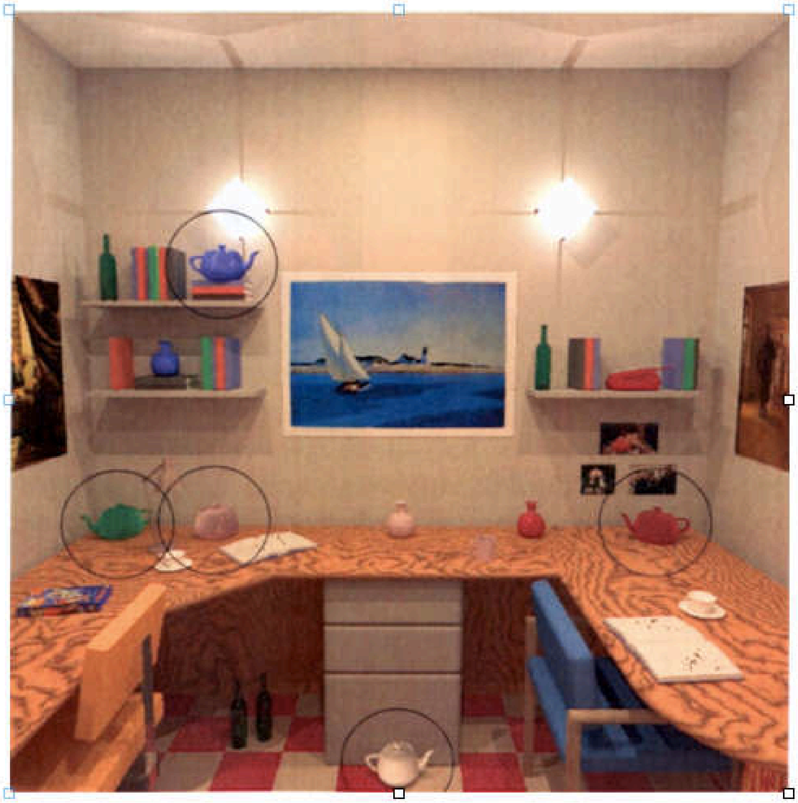


Figure 10: Task-based rendering [19].

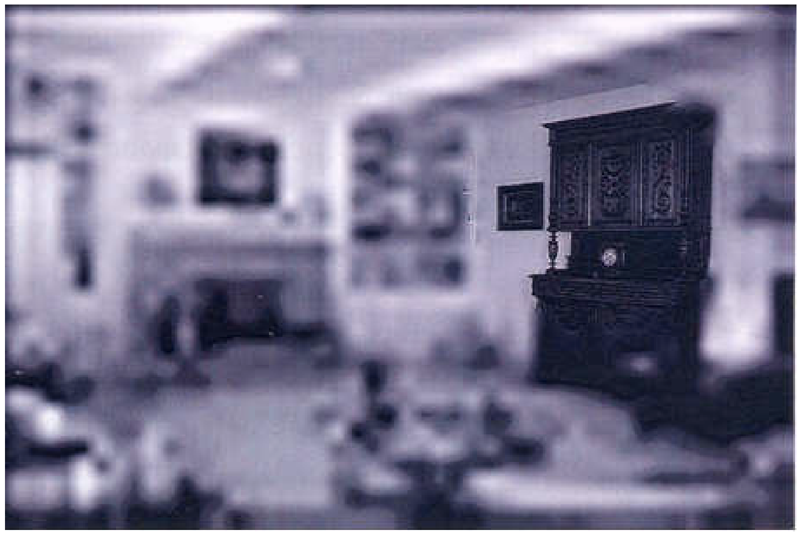


Figure 11: Gaze-based rendering [70].

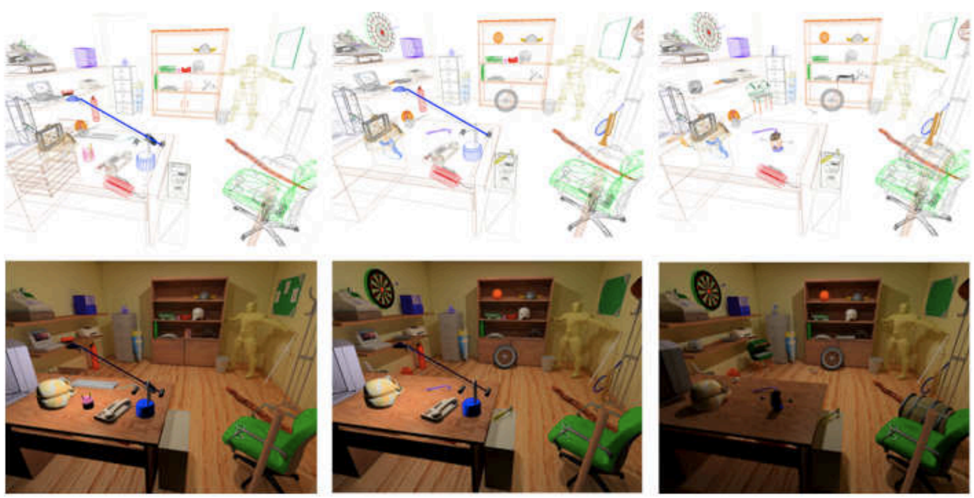


Figure 12: Schema memory experimental studies [87].

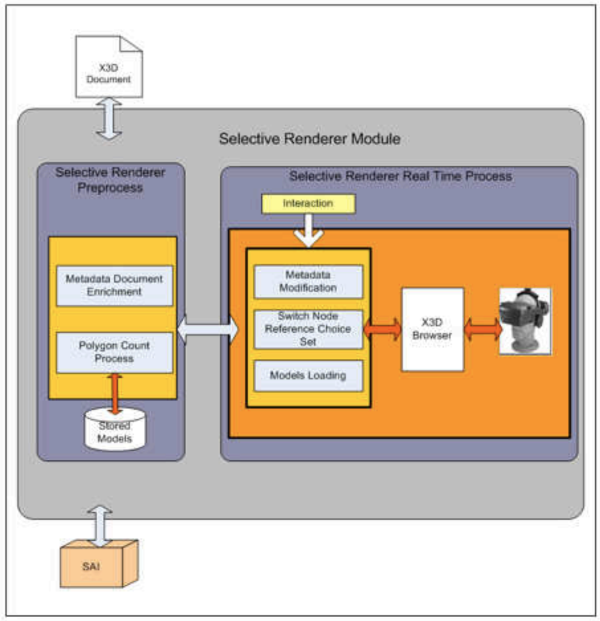


Figure 13: Selective Renderer Architecture [135].

graphics manipulations such as polygon count and depth of shadow rendering; normal cognition operates as long as the scenes look acceptably realistic. On the other hand, when the overall realism of the scene is greatly reduced, such as in the wireframe condition, then visual cognition becomes abnormal. Specifically, effects that distinguish schema-consistent from schema-inconsistent objects change because the whole scene now looks incongruent, and we have shown that this effect is not due to a failure of basic recognition. Thus, a recipe for anyone wishing to use such displays in studies of visual cognition is to construct environments which look acceptably realistic in terms of polygon count but need not be of very high quality. This relates to the kinds of high-level visual cognitive effects we have studied here, such as object congruence. Lower-level effects, such as recognition, can be dissociated from these high-level effects (but make their presence felt when the scene is further degraded, e.g. when color is removed from the wireframe scenes). Thus, high-level processes need somewhat greater realism than low-level ones.

4.3 Behavioral Fidelity of Simulations based on Space Memory

The entertainment world appears to consider highly realistic visual quality one of the keys to success, with cinematic quality graphics claimed for the next generation of gaming consoles. On the other hand, when interactive immersive

Virtual Environments (VEs) are implemented for training rather than entertainment purposes, visual quality might not be as significant. If the training is to be effective, the skills acquired must transfer into the real world at appropriate levels of performance. A VE with maximum visual and interaction fidelity would result in a transfer of information equivalent to real world training since the environments would be indistinguishable [73]. Visual fidelity refers to the degree to which visual features in the VE conform to visual features in the equivalent real environment [123, 106]. Functional realism refers to the communication of similar information in the real and virtual world rather than the aesthetics or physics, in the sense that users are able make the same judgments and perform the same tasks as in the real world [37, 91].

It is tempting to replicate the real world as accurately as possible in order to provide equivalent experiences [67]. Whilst arguably ideal, it is not yet computationally feasible for this to occur. Trade-offs between visual/interaction fidelity and computational complexity should be applied to a simulation system without detracting from its training effectiveness [72, 125]. There is, therefore, a call for efficient techniques assessing the fidelity of a VE and determine its relationship with performance in order to economize on rendering computation without compromising the level of information transmitted (functional realism) [37].

The utility of Virtual Environment (VE) technologies for training systems such as flight simulators is predicated upon the accuracy of the spatial representation formed in the VE. Spatial memory tasks, therefore, are often incorporated in benchmarking processes when assessing the fidelity of a VE simulation for training. Spatial awareness is significant for human performance efficiency of such tasks as it is dependent on spatial knowledge of an environment [64, 28, 130]. A central research issue, therefore, is how an interactive synthetic scene is cognitively encoded and how recognition and memory of such worlds transfer to real world conditions [72, 1, 38]. Previous research has examined the variables that communicate transfer of spatial knowledge acquired in a simulation environment, in the real-world and discuss the form and development of spatial awareness in VE training compared to either real-world training or training with maps, photographs and blueprints [8, 6]. The suitability of VE systems as effective training mediums was examined and was concluded to be as effective as map or blueprint training. Configurational knowledge acquisition based on estimation of absolute distances and directions between known points could yield training effects similar to training with photographs and real world training [8]. Furthermore, estimation of travel distance from optic flow is subject to scaling when compared to static intervals in the environment, irrespective of additional depth cues [40]. Past research often aims to identify the minimum system characteristics relevant to rendering computations and interaction interfaces that would yield the maximum performance on a task or the greatest sense of presence. For example, search objects rendered in global or ambient illumination have been shown to take significantly longer to identify than those rendered through a local illumination model [134]. What if the visual fidelity of a system should be assessed across a range of applications and tasks? Could we interrogate the human cognitive systems that are activated when training within VE scenes of varied visual or interaction fidelity in order to identify whether such responses are transferable to the real-world task situation simulated? Which simulation characteristics should we optimize in order to match the capabilities of the VE system to the requirements of these cognitive systems?

Because of the wide-range of VE applications and differences in participants across their backgrounds, abilities and method of processing information, an understanding of how spatial knowledge is acquired within a VE, complementing spatial memory performance per se, is significant. Common strategies may be revealed across a range of applications and tasks. Recent research focuses upon the effect of rendering quality (flat-shaded vs radiosity) on object-location recognition memory and its associated awareness states while spatial knowledge is transferred from a synthetic training environment into a real-world situation. The main premise of this work is that accuracy of performance per se is an imperfect reflection of the cognitive activity that underlies performance on memory tasks [74]. The framework to be presented has been drawn from traditional memory research adjusted to form an experimental procedure [119, 11, 27].

Accurate recognition memory can be supported by: a specific recollection of a mental image or prior experience (remembering); reliance on a general sense of knowing with little or no recollection of the source of this sense (knowing); guesses. Gardiner and Richardson-Klavehn [41] explained the remembering as personal experiences of the past that are recreated mentally. Meanwhile knowing refers to other experiences of the past but without the sense of reliving it mentally. The work of Tulving [119] first suggested that remembering and knowing were measurable constructs. Through a series of experiments, Tulving [119] reported that participants find it easy to distinguish between experiences of remembering and knowing when self-reporting their experiences. The sense of knowing has since been further divided into two related concepts. The correct answer may be just known without the associated recollection of contextual detail associated with remembering or the answer may feel more familiar than a uninformed guess, but cannot be considered as being known (familiar).

According to this theoretical framework derived from memory psychology, measures of the accuracy of memory can be complemented by self-report of states of awareness such as remember, know, familiar and guess during recognition [24]. Previous studies have investigated the relationship between recognition memory and simulation environments of varied visual and interaction fidelity. Such work by the authors of this paper revealed varied distribution of awareness states whilst overall accuracy remained the same across experimental conditions suggesting that measurement of awareness states acts as a useful additional measure to supplement the information provided by accuracy [75, 72].

A different study employing the same methodology aims to interrogate the mental processes associated with obtaining spatial knowledge during exposure to a simulated scene while transferring such knowledge in the real-world scene simulated [74]. An object-memory task was performed in the simulated real-world environment immediately after VE training and a retention test was conducted one week after the VE exposure. The virtual scene was rendered with one of two levels of visual fidelity (flat shaded vs. radiosity rendering) and displayed on a stereo Head Mounted Display (HMD). The experimental scene consisted of a room depicting an academics office. Central to this work is identifying whether high fidelity or low fidelity scenes are associated with stronger visually induced recollections represented by self-report of the remember awareness state. A secondary, exploratory goal is to investigate the effect of schemas on memory recognition post VE exposure. Memory recognition studies in synthetic scenes have demonstrated that low interaction fidelity interfaces such as the mouse compared to head tracking as well as low visual fidelity scenes provoked a higher proportion of visually-induced recollections associated with the remember awareness state, while there was no effect of condition upon memory recognition performance [75, 72].

Broadly, desirable influences on recognition memory and the associated cognitive states may be ultimately identified and generalized to aid specific applications. It could be true, for instance, that for flight simulation applications it is crucial for trainees to refer to mental images associated with instruments as opposed to recollections that are confident but not accompanied by mental imagery when training is transferred into a real-world flight situation. The study presented, therefore, explores the effect of training in immersive environments of varied visual fidelity on the distribution of memory awareness states measured in a real-world task [74]. The fact that it has been shown that interfaces of low interaction or visual fidelity induce a higher number of recollections based on mental imagery when compared with systems of high visual or interaction fidelity, may relate to attentional resources directed to systems that vary strongly from the real-world. The results demonstrated that participants who trained in the low fidelity simulation reported a larger proportion of correct remember responses while conducting the memory recognition task in the real-world situation compared to participants trained in the high fidelity simulation. These results were consistent with previous findings that associated a larger proportion of correct remember responses with low visual and interaction fidelity simulations [75, 72]. The results observed consistently in previous studies was also observed in this study despite the fact that participants physically performed the task in the real-world room after training in its simulated counterpart consisting of an ecologically plausible training scenario.

Recent developments in psychological research have shown that distinctive information or experiences generate more awareness states associated with remembering. For example, participants who are shown typical and distinctive faces are more likely to recognize the distinctive faces in a later memory test with an accompanying experience of remembering [12]. Similar results have also been found using other stimuli such as forenames [11] . In the current context, a low fidelity rendered simulation could be considered as being more distinctive than a high fidelity rendered simulation because of its variation from real. Given that these are immersive environments, distinctiveness in this instance would be judged relative to reality. The less real the environment is, the more distinctive it can be considered. It would be expected that a more distinctive immersive environment, e.g. a low fidelity one would result in more remember responses than a less distinctive immersive environment, e.g. a high fidelity one. It is worth noting that distinctiveness in this sense may not only refer to visual distinctiveness but to motor responses to the environments [72]. The important variable therefore appears to be differentiation relative to multiple aspects of reality, e.g. visual appearance of, and, motor responses within. Here, higher confidence scores associated with the flat-shaded condition compared to confidence of recollections after training in the radiosity condition further support this suggestion.

Whilst the relationship between distinctiveness and memory may prove useful in explaining these effects it is important to consider what cognitive processes may underlie such a relationship. Previous psychological research has indicated that remember responses require more attentional processing in the first instance than those based on familiarity [93, 12]. A tentative claim would therefore be: immersive environments that are distinctive recruit more attentional resources. This additional attentional processing may bring about a change in participants subjective experiences of remembering when they later recall the environment. This change would therefore lead to an increase in the experience of remembering. Interestingly, this effect was not observed during the retest that revealed similar proportions of awareness states distributed across the viewing conditions. It is likely that the fidelity of the training environment only affects awareness states when transfer of training is tested immediately. As time goes by, the enhanced attentional resources associated with low fidelity environments do not influence the long-term memories associated with the training simulation.

Moreover, it is found here that more correct know responses are reported after training in the high fidelity rendered simulation than in the low fidelity rendered simulation. This would suggest a shift from remember responses to know responses. Memories that are accompanied with a feeling of remembering for participants in the low fidelity simulation are only accompanied with a feeling of knowing in the high fidelity simulation. In line with suggestions made above, this could be explained on the basis of reduced attentional processing of these items in the high fidelity simulation.

4.4 Investigating Perceptual Sensitivity to Head Tracking Latency

Virtual Environment (VE) latency is the time lag between a users action in a Virtual Environment and the systems response to this action. This lag typically takes the form of a transport delay and arises from the sum of times associated with measurement processes of the various input devices, computation of the VE contents and interaction dynamics, graphics rendering, and finite data transmission intervals between these various components. The VE and human factors literature has established that these delays have a significant impact on user performance [33], [32] and user impressions of simulation fidelity [31, 33, 60, 78, 1]. Latency negatively affects user performance in 3D object placement tasks [66, 128].

Excessive latency has long been known to hinder operator adaptation to other display distortions such as static displacement offset [51]. Latency also degrades manual performance, forcing users to slow down to preserve manipulative stability, ultimately driving them to adopt a move and wait strategy [107] , [108]. Operator compensation for a delay usually requires the ability to predict the future state of a tracked element.

Interest has more recently been directed toward the subjective impact of system latency relevant to virtual reality simulations. Latency as well as update rate have been considered as factors affecting the operators sense of presence in the environment [129, 121]. In a recent study, lower latencies were associated with a higher self-reported sense of presence and a statistically higher change in heart rate for users while in a stress-inducing (fear of heights), photorealistic environment involving walking around a narrow pit [83].

Since the combination of sensing, computation, rendering, and transmission delay is unavoidable in most VE, tele-operation, and augmented reality applications, interest naturally is directed to how detectable differing levels of latency might be. Both the quantification of perceptual sensitivity to latency and description of the mechanism by which VE latency is perceived will be essential to guide system design in the development of countermeasures such as predictive compensation [5, 60].

Previous research has also focused on the precision, stability, efficiency and complexity of operation interaction and performance with latency-plagued systems [78]. Additionally, the first measures of human operators discrimination of the consequences of latency during head or hand tracked movements have been provided [34, 31] Related investigations have explored the hypothesis that observers do not explicitly detect time delay, but rather detect the consequences of latencyi.e., they use the artifact motion of the VE scene (away from its normally expected spatially stable location) caused by system time lags [1]. Relevant perceptual thresholds (i.e., Just Noticeable Difference or JND) were identified to average 8-17ms, depending on viewing condition. This psychometric quantity appeared to be invariant across different pedestals (33, 100 and 200ms, standard stimuli). The apparent invariance of the detection function in [34, 31, 1] demonstrated that the classic Webers Law of psychophysics (that JND is linearly proportional to the magnitude of the standard stimulus) did not hold for latency. In other words, observers of long latency VEs will be as sensitive to changes in latency as those who use prompter, more advanced systems. It can also be inferred that the same sensitivity would also apply for comparisons against zero latency pedestal.

Regan et al.[102] found 70.7% latency thresholds averaging 15ms for a specialized non-immersing CRT display. By making assumptions of Gaussian psychometric functions and zero response bias for two-interval forced-choice judgments with balanced presentation order, the 70.7% threshold from [102] can be equated with a JND of 18.6ms.

Allison et al. 2001 observed on the other hand that with large virtual objects occupying the full Head Mounted Display (HMD) Field-of-View (FOV), 50%

thresholds for perceived image instability (oscillopsia) were found to be 180-320ms depending on head motion velocity [3]. This threshold indicates the latency level at which observers were equally likely as not to say the image was unstable and represents their average response bias or preference. Such response biases may be attributable to, among other things, the amount of observer training before the data was collected and the type of judgment task required. In the case of Allison et. al. [3] participants performed single interval judgmentsi.e., they did not compare each presentation against a standard stimulus but relied on their own internal notion of when an image was no longer stable. Data from citeellis99a, [31], [1] shows their participants response bias ranged between 40 and 70ms for a two-interval judgment of whether the stimulus was the same as or different than the pedestal standard. In contrast, the participants in [102] were forced to choose which of the two stimulus intervals was actually the one with added latency, which though not reported, leads to a presumption of zero bias.

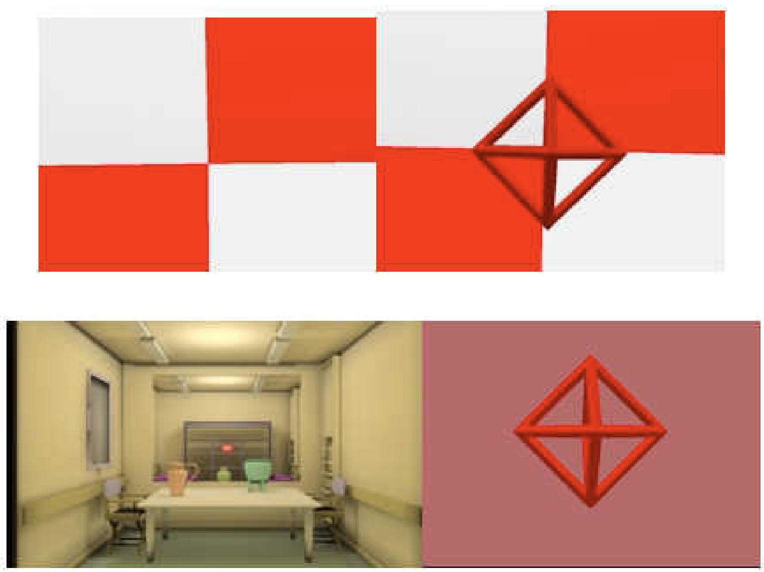


Figure 14: Experimental conditions [30], [73].

The much higher threshold reported by Allison et al. [3] might also be attributable to the fact that their participants viewed a textured virtual background (the inside surface of red and white faceted sphere) that completely enveloped their head and thus always occupied their entire FOV. Surrounding observers with such a geometrically structured environment contributes to the phenomenon of visual capture. The term visual capture implies that when concurrent multi-sensory spatial information is available, the observer will weight the visual channel more heavily in constructing a percept. It has been demonstrated that, even with very simple VE graphics in an HMD, visually discrepant information will bias proprioceptive and vestibular feedback of static head pitch angle [91]. Since awareness of VE image instability relies on visual, vestibular, and proprioceptive information, the full structured background viewed in Allison et al.s study [3] may have diminished their observers sensitivity to latency induced oscillopsia. Furthermore, without the inclusion of nearer objects in their environment, participant head movement does not trigger motion of scene contents relative to the background and thus does not provide cues through internal image shear.

One aim of ongoing research on latency perception at NASA Ames Research Centre has been to quantify the latency that a VE system can exhibit without being perceptible to the user. In our prior studies, we employed very sparse environments containing only a single simple object such as a faceted sphere [31], [34] or a hollow-framed octahedron [1]against an empty black background. Synthetic environments with differing levels of graphical complexity with the goal of extending the generality of our results for participant sensitivity latency in VEs have been also employed by Mania et al.[73].

The focus of the [73] study is in describing observer sensitivity to latency differences during head movements in an immersive VE representing a real-world space (room, building, etc.)sensitivity that has not been measured in previous research. On the one hand, because there could be an inherent association with how the real world is perceived, we might expect observers to be more sensitive to the visual consequences of latency when viewing a scene representing what could be a real-world space rather than a sparse, simplified scene with only one or two artificial objects. On the other hand, an enveloping structured scene could promote visual capture, thereby degrading observers sensitivity to VE latency.

During an earlier study more fully reported in [30], a simple white-red checker sphere surrounding the observer, such as that used in [3] and/or a hollow-frame octahedron in front of the observer, as in [1] served as the VEs visual content. Participants were asked to compare two sequential stimulus presentations while moving their head in a constant pattern and report whether the stimuli differed in the visible consequences of the experimentally manipulated VE latency. The study presented here employed the same experimental methodology, but instead, the visual scene was a pre-computed radiosity rendering of two interconnected rooms that include real-world objects. Here, we also statistically compare sensitivity results derived from Ellis et al. [30] and the study presented in this paper. Both studies also explore whether relative motion shear between more than one artificial object in the VE could be a mechanism contributing to observer perception of head tracking latency.

In summary, results from these studies conducted at NASA Ames Research Centre suggest that virtual environment system designers should expect observers who are not burdened with any other performance tasks to generally be able to notice differences in latency as low as 15ms, regardless of the relative location of objects in the scene, the meaningfulness of the scene context in relation to the real world, or possibly even the degree of photorealism in their rendering. These results will also serve as performance guidelines to aid in the design of predictive compensation algorithms.

Diego Gutierrez

5 HDR, Illumination and Image Processing

High Dynamic Range: Case study on Reverse Tone Mapping

Some slides from Masia et al. [SIGGRAPH Asia 2009] and
Banterle et al. [SIGGRAPH Asia 2011 (courses)]

Background – HDR

The dynamic range of a scene refers to the difference between the highest and lowest luminance values in it.

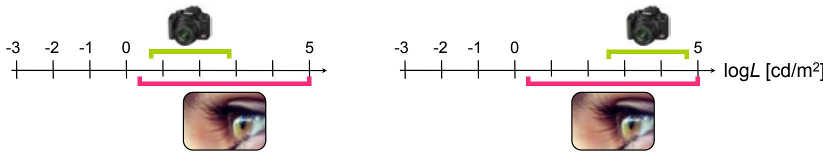
Starlight 10^{-3} cd/m^2

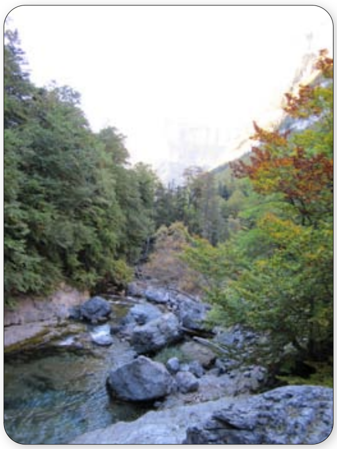
10^{-1} Moonlight

10^{2} Indoor lighting

Direct sunlight 10^{5} cd/m^2

Background – HDR





Traditional LDR images:

- 8 bits/channel
- 2 log-units of luminance

➡ **They fail to faithfully recreate the real world**

Background – HDR

HDR images can be obtained by combining standard photographs with different amounts of exposure [Debevec and Malik 1997]

[Images from Cambridge in Colour]

Background – Tone mapping

HDR

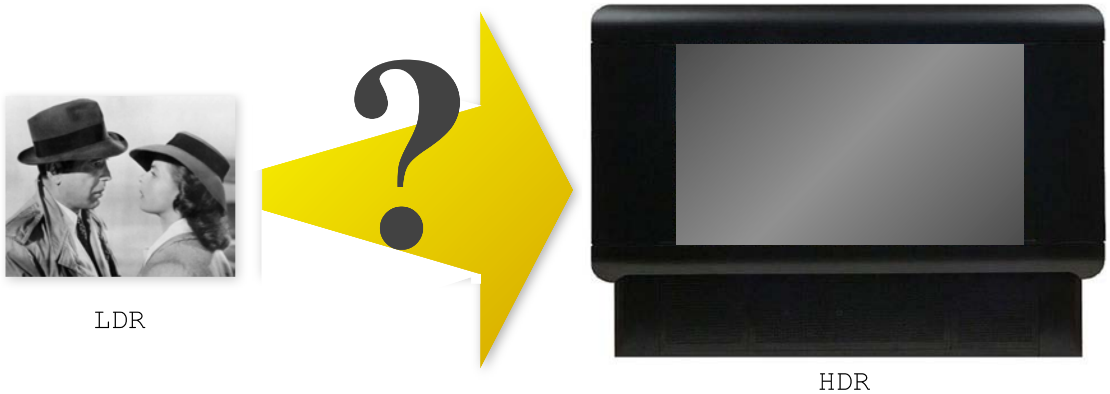
LDR

Dozens of algorithms in the last 10 years (and more to come)

Background – HDR

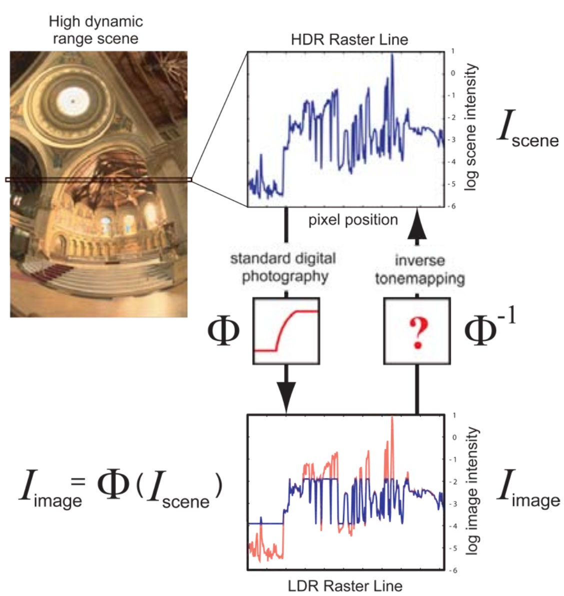
Dynamic range of displays is increasing more and more… so what do we do with all the LDR material?

Background – Reverse tone mapping



Just a few algorithms in the last years

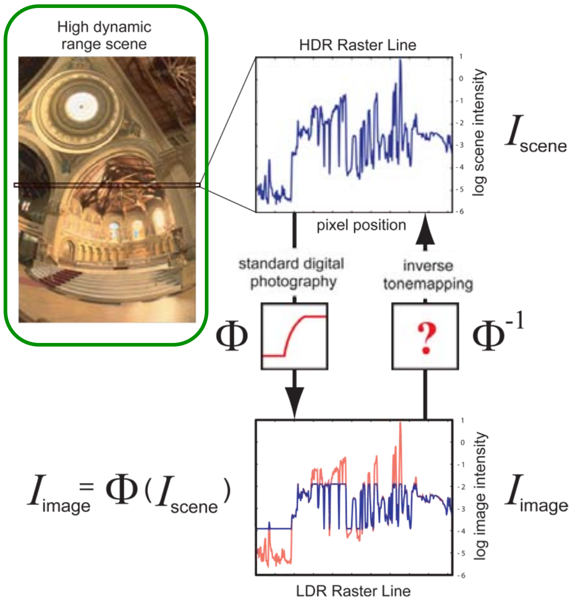
Background – Reverse tone mapping



Nonlinearities,
sensor noise,
lens flare,
blooming,
sensor saturation…

Ill-posed problem!

Background – Reverse tone mapping

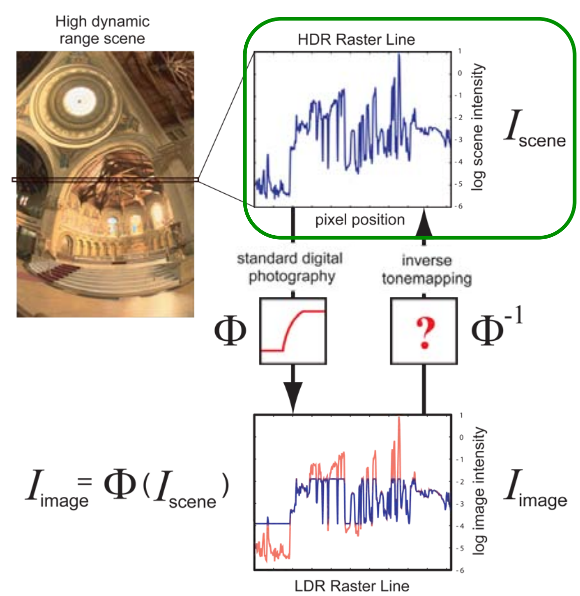


Nonlinearities,
sensor noise,
lens flare,
blooming,
sensor saturation…

Ill-posed problem!

Conclusions

Background – Reverse tone mapping

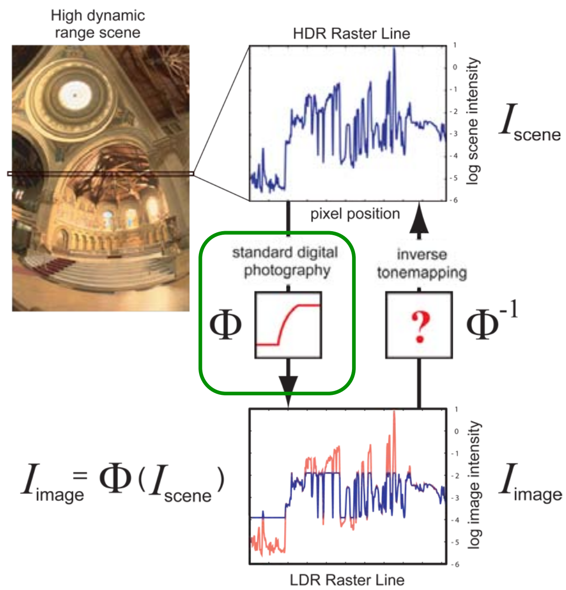


Nonlinearities,
sensor noise,
lens flare,
blooming,
sensor saturation…

Ill-posed problem!

SIGGRAPH ASIA 2009

Background – Reverse tone mapping

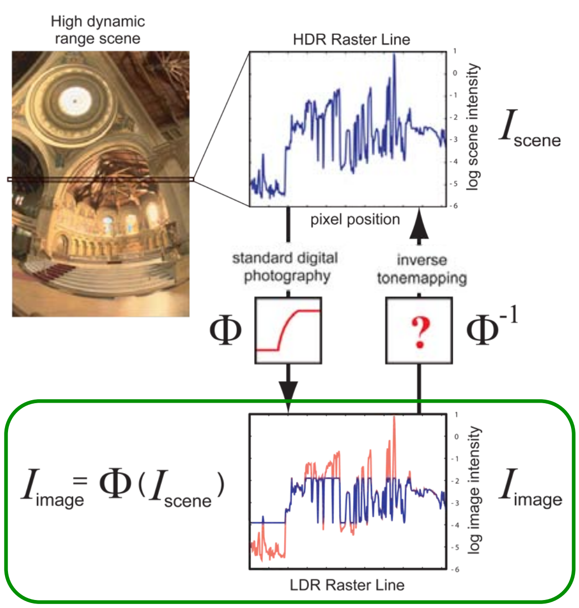


Nonlinearities, sensor noise, lens flare, blooming, **sensor saturation**…

Ill-posed problem!

Conclusions

Background – Reverse tone mapping

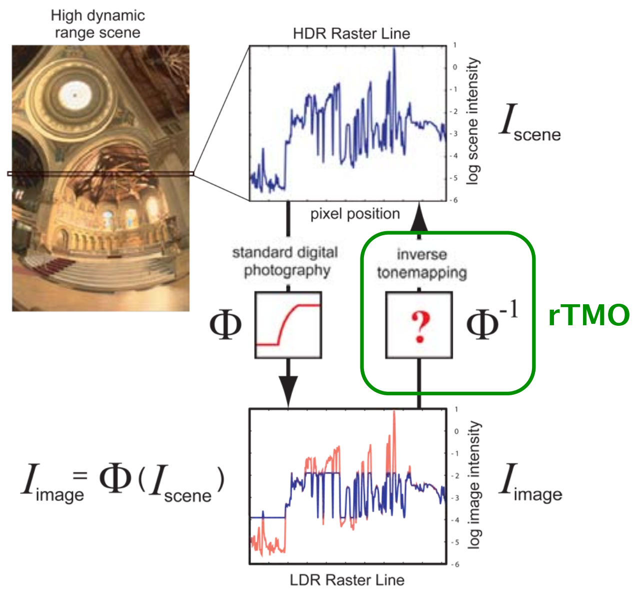


Nonlinearities, sensor noise, lens flare, blooming, **sensor saturation**…

Ill-posed problem!

SIGGRAPH ASIA 2009

Background – Reverse tone mapping



Nonlinearities,
sensor noise,
lens flare,
blooming,
sensor saturation…

Ill-posed problem!

Previous approaches (rTMOs)

[Daly and Feng 2003; 2004]

[Banterle et al. 2006; 2007; 2008]

[Meylan et al. 2006; 2007]

[Wang et al. 2007]

[Rempel et al. 2007]

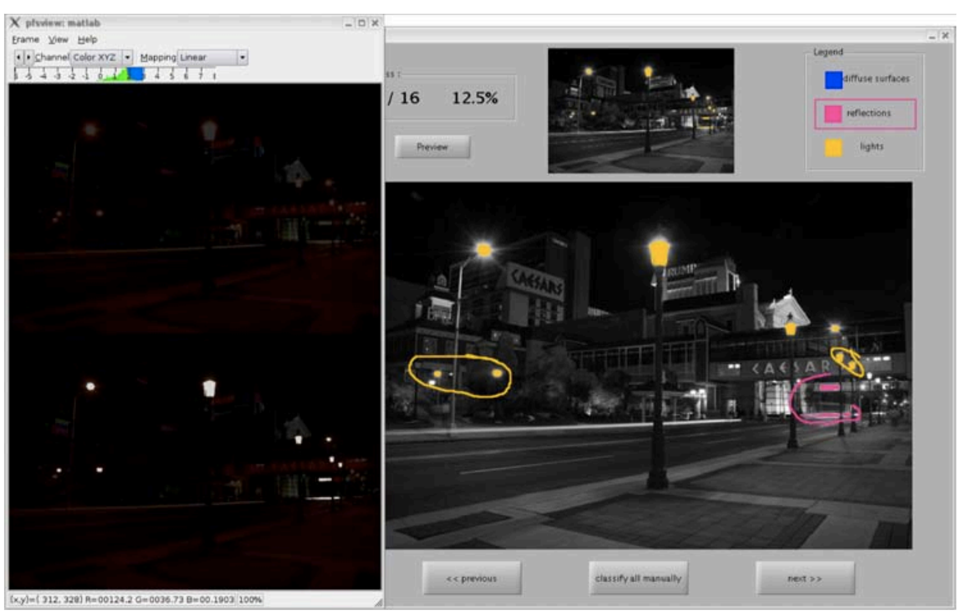
[Didyk et al. 2008]

[Kovaleski and Oliveira 2009]

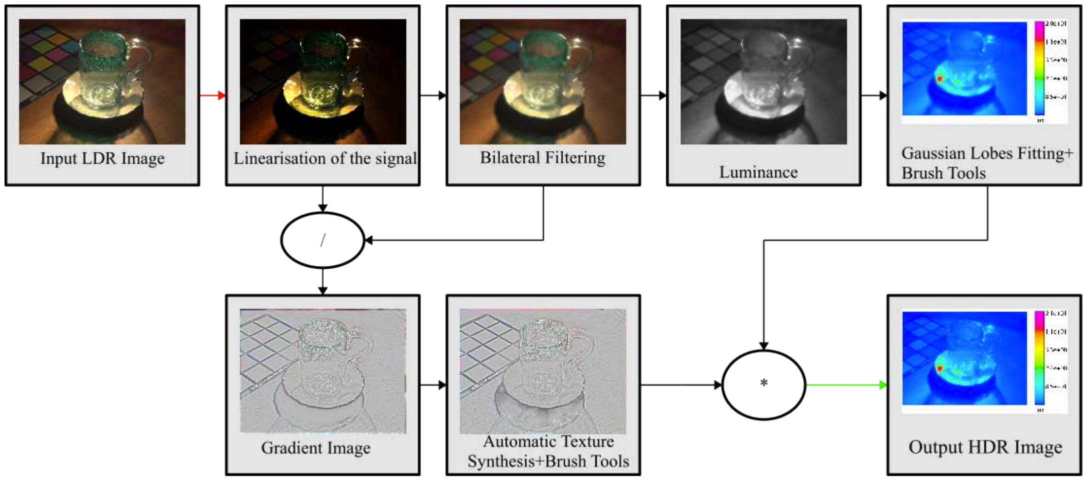
…

- Expansion Methods:
  - Global Methods
  - Expand Map Methods
  - Classification Methods
  - User Based Methods
- Evaluation:
  - Psychophysical Experiments
  - Computational Metrics
- Conclusions

Classification Methods:



User Based Methods: Hallucination



Previous approaches

[Banterle et al. 2006; 2007]

Source
LDR image

Inverse of
Reinhard's TMO

+

Expand map

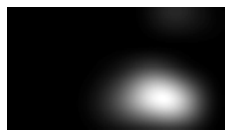
Previous approaches

[Rempel et al. 2007] (LDR2HDR)

Source LDR image

Inverse gamma and noise filtering

+



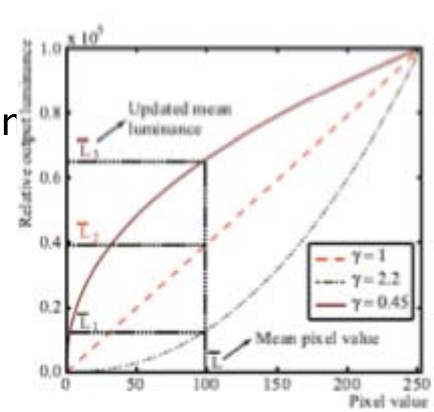
Smooth brightness enhancement

+

Edge stopping function

Previous approaches – [Akyüz et al. 2007]

- Keep it simple – why not a simple linear scaling?
- Performed a psychophysical evaluation
- **“A simple linear scale can provide an HDR experience”** based on psychophysically experir



3 key observations

Conclusions

Our approach

3 key observations

1 Simple global operators may be good enough

SIGGRAPHASIA2009

Our approach

3 key observations

1 Simple global operators may be good enough

2 Perception as a key issue to gain insight into the problem

Our approach

3 key observations

1 Simple global operators may be good enough

2 Perception as a key issue to gain insight into the problem

3 No previous addressing of exposure

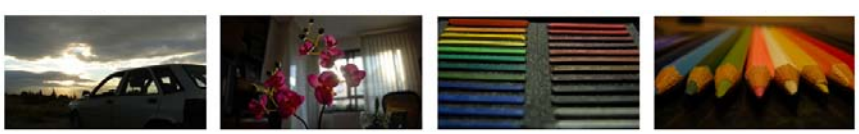
Experiment One – Objective

Study how existing rTMOs work across exposure

Experiment One – Stimuli

Study how existing rTMOs work across exposure

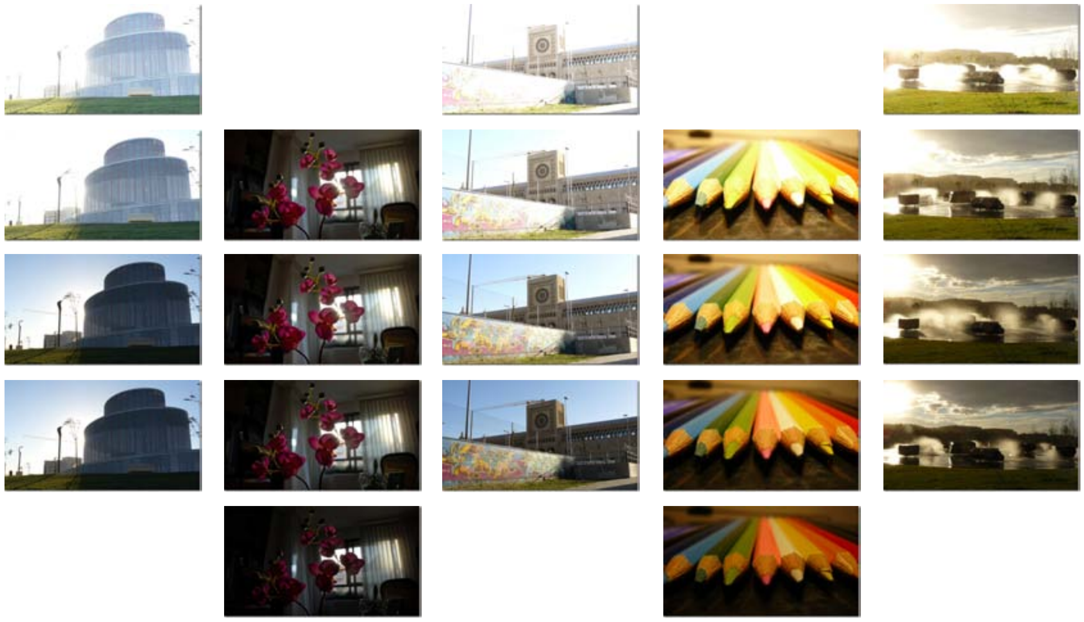
Bright series



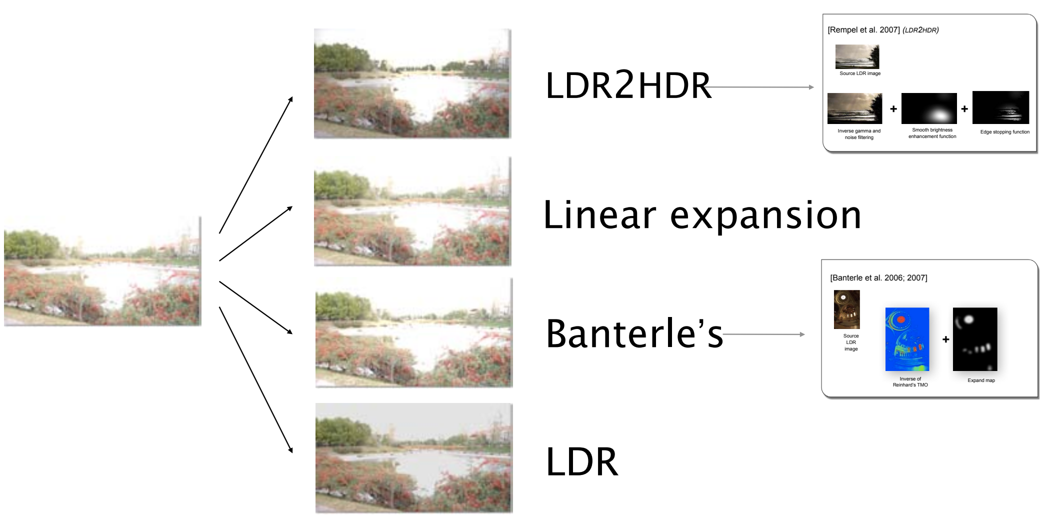
Dark series

[Martin et al. 2008]

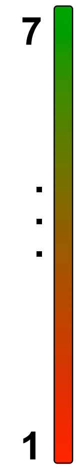
Experiment One – Stimuli



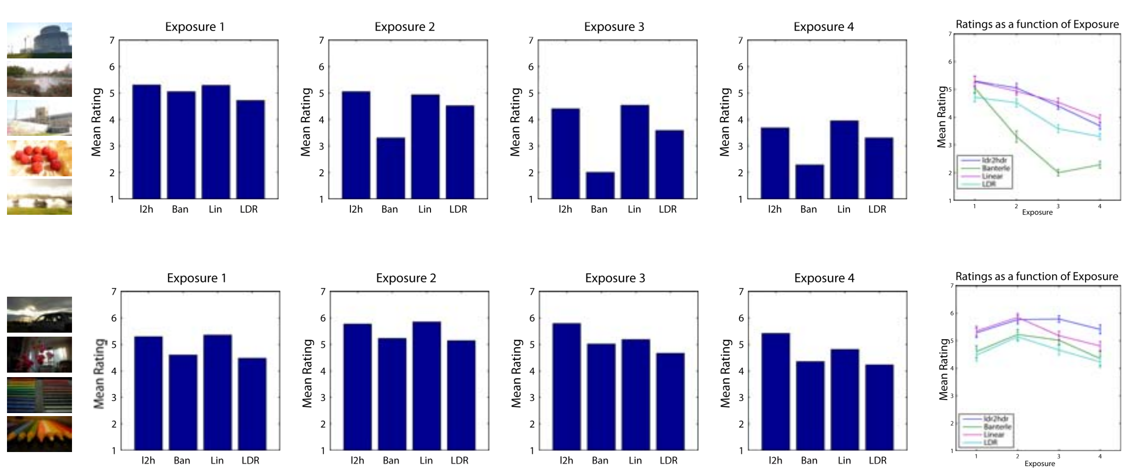
Experiment One – Stimuli



Experiment One – Procedure



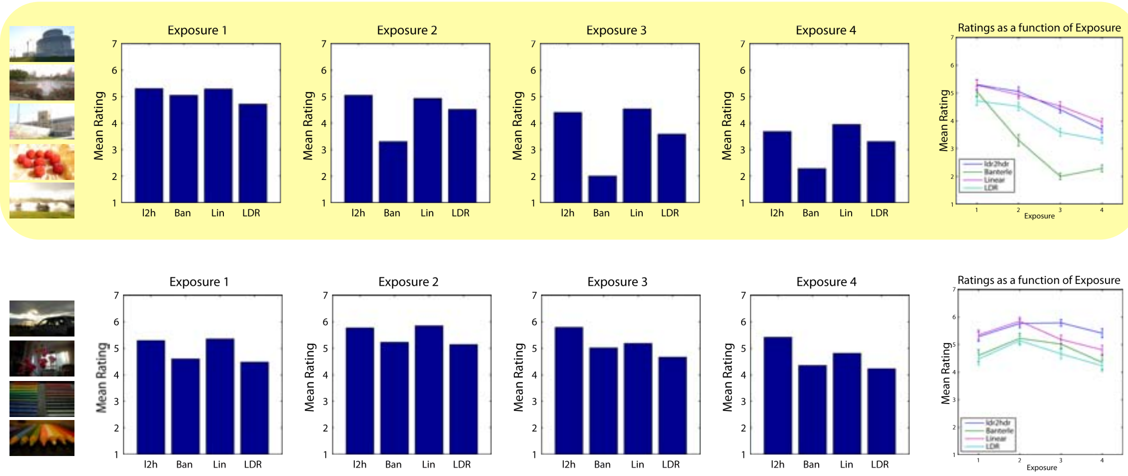
Experiment One – Results



Mean ratings

Conclusions

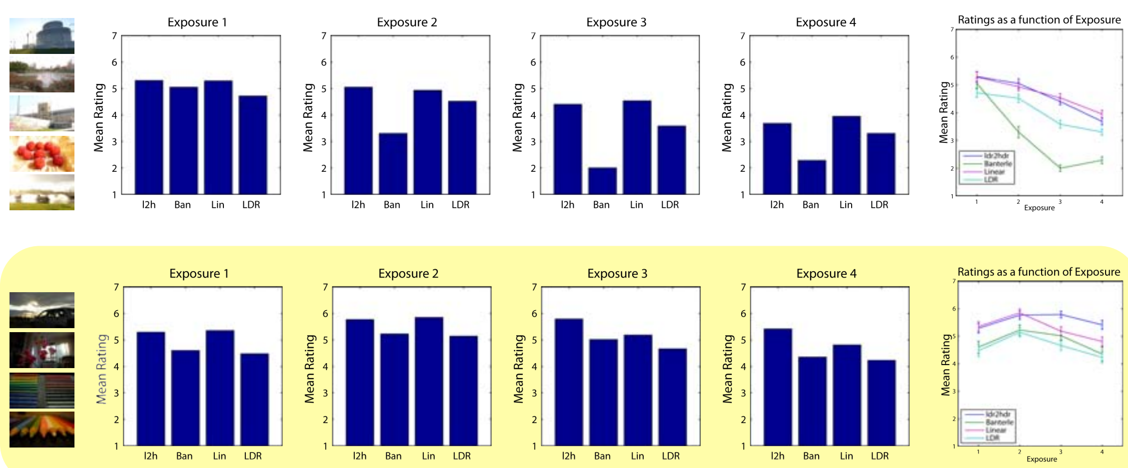
Experiment One – Results



Mean ratings

Conclusions

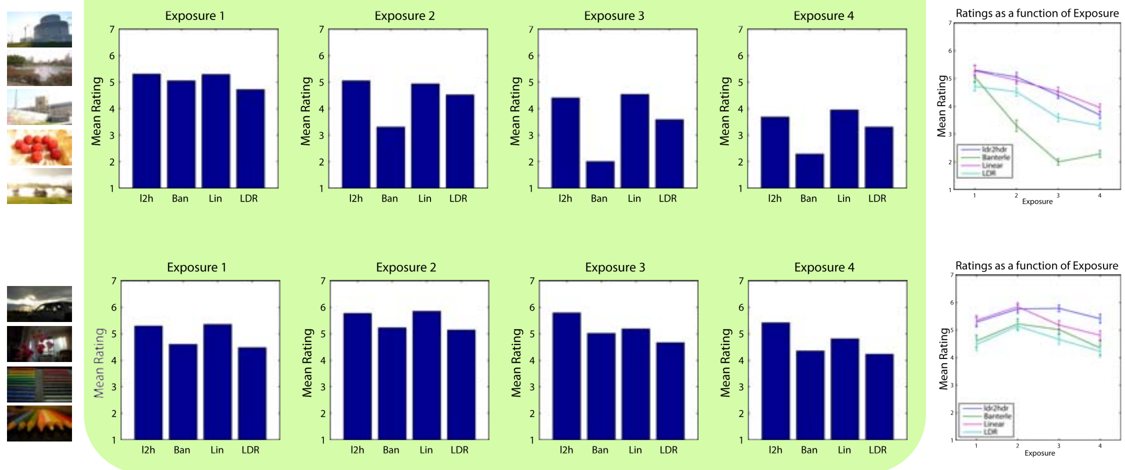
Experiment One – Results



Mean ratings

Experiment One – Results

Exposure

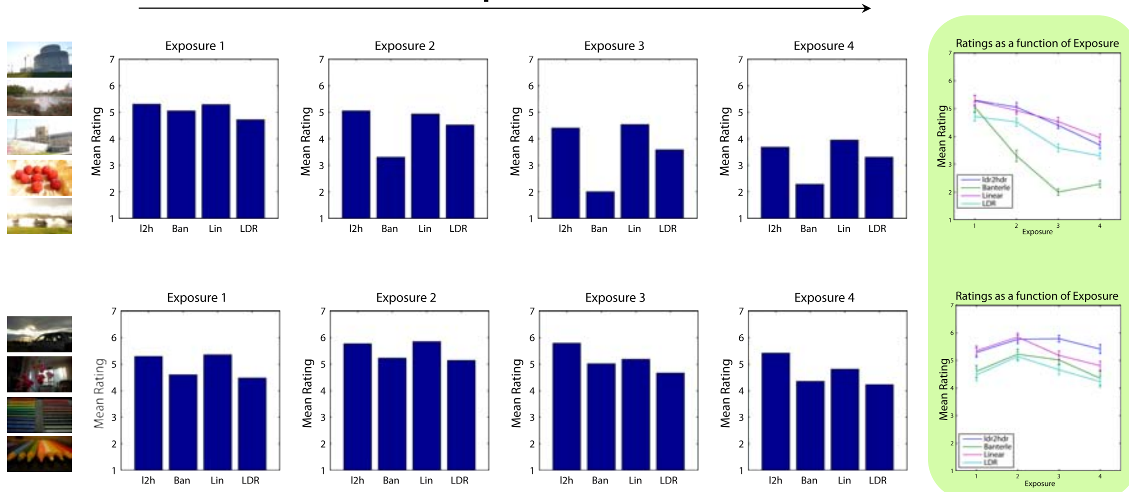


Mean ratings

Conclusions

Experiment One – Results

Exposure

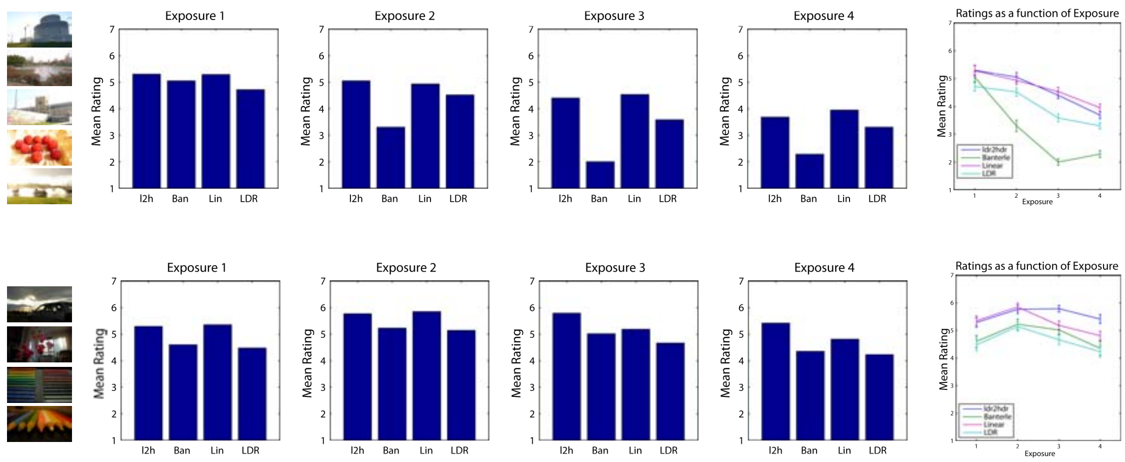


Mean ratings

SIGGRAPH ASIA 2009

Experiment One – Results

Exposure

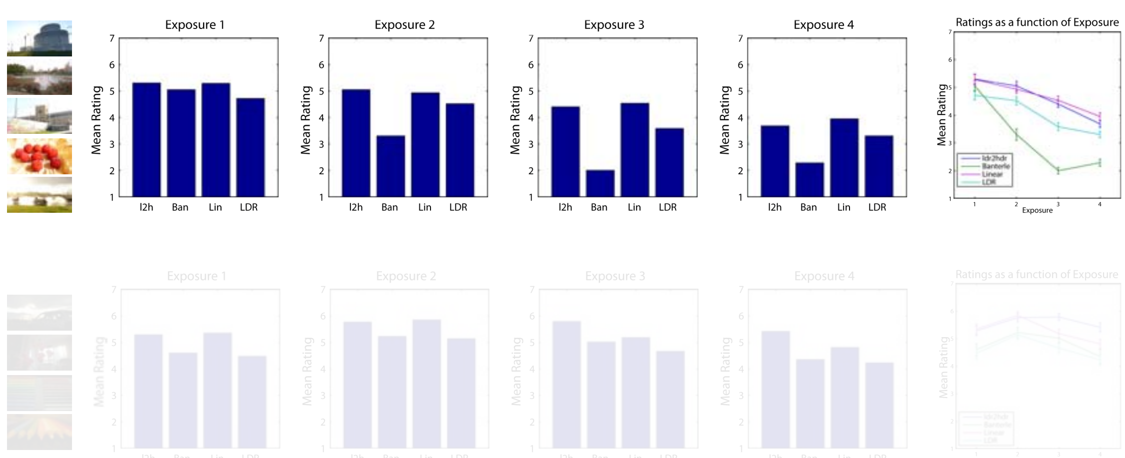


Mean ratings

Conclusions

Experiment One – Results

Exposure



Mean ratings

SIGGRAPH ASIA 2009

Experiment One – Results

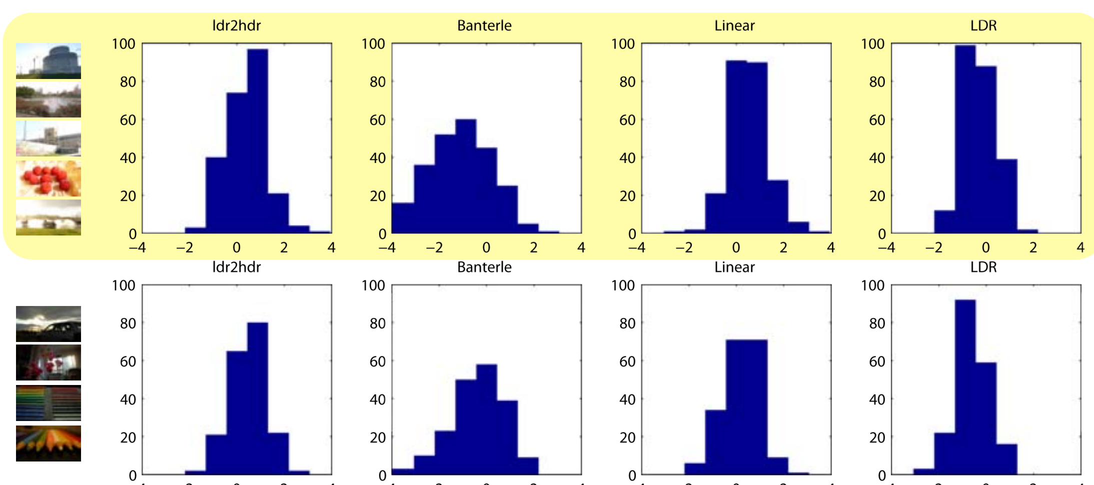
- Further analysis → for each rTMO for each image we compute the outlier index:

$$\text{Outlier index } (\text{rTMO}_i) = (\text{rTMO}_i \text{ Rating}) - (\text{Median rating across rTMOs})$$

- And we plot the corresponding histograms…

Conclusions

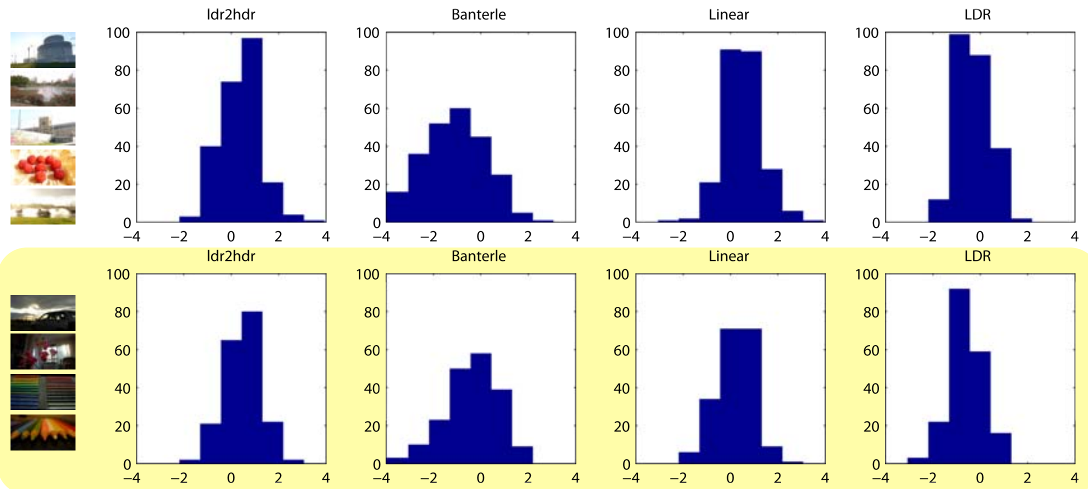
Experiment One – Results



Distribution of outlier indices

SIGGRAPHASIA2009

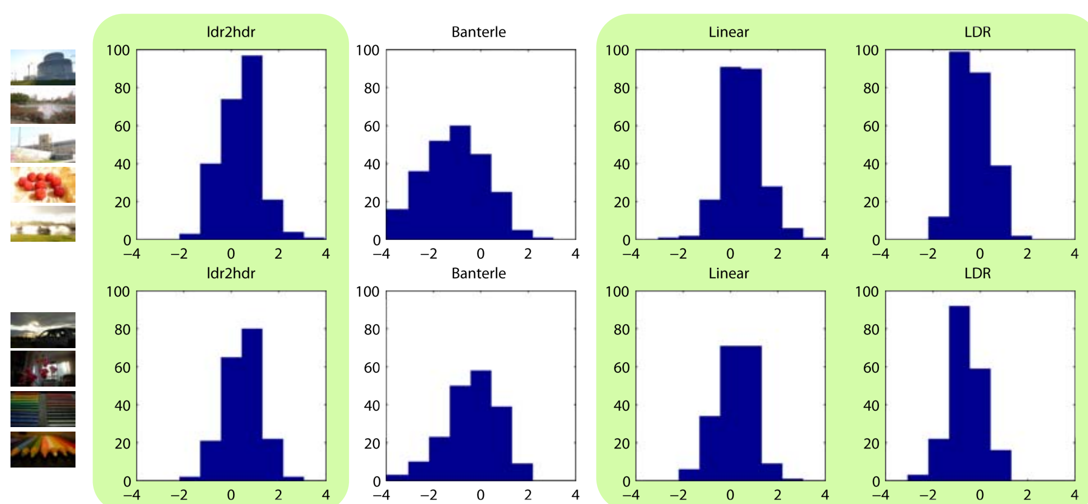
Experiment One – Results



Distribution of outlier indices

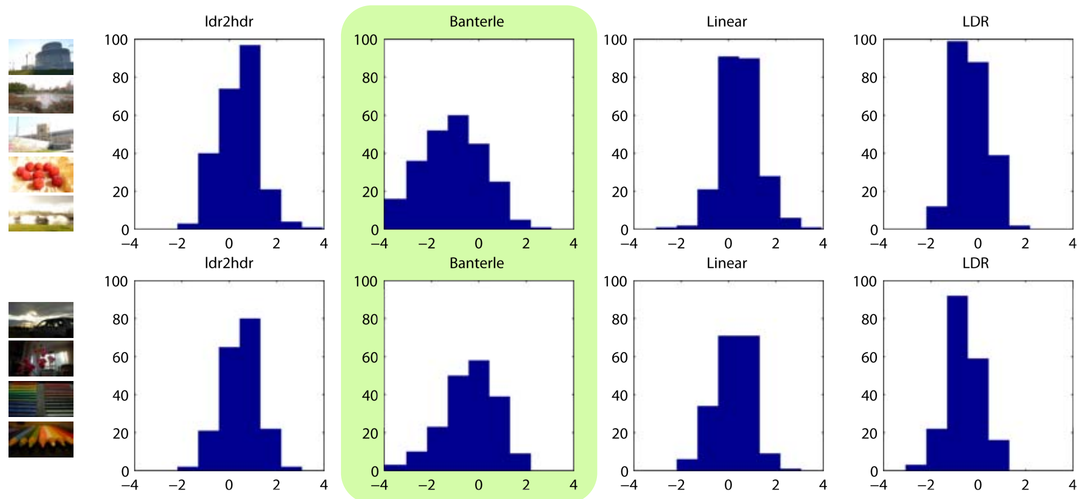
Conclusions

Experiment One – Results



Distribution of outlier indices

Experiment One – Results

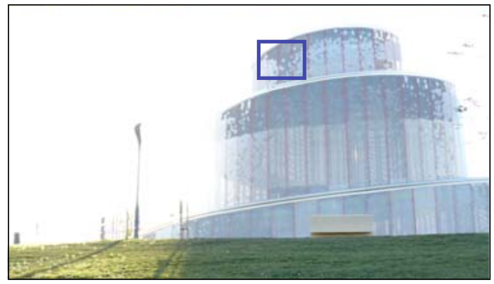


Distribution of outlier indices

Experiment Two

Experiment Two

Artifacts introduced by rTMOs also visible in LDR renditions of the images



Experiment Two

HDR

vs.

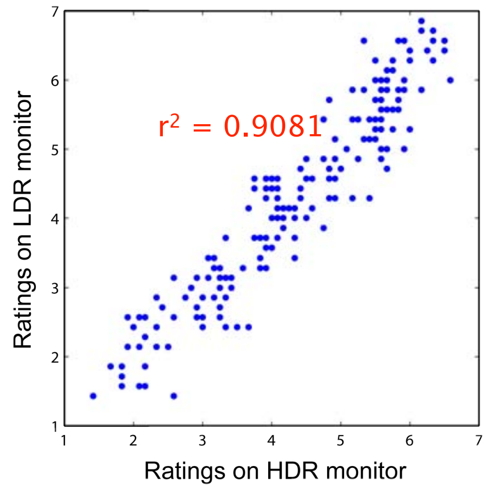
LDR

Experiment Two

HDR

vs.

LDR

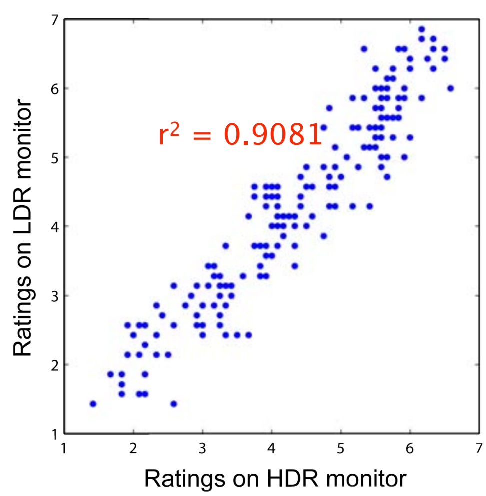


Experiment Two

HDR

vs.

LDR



γ Expansion for Over-exposed Content

- rTMO for bright images with large saturated areas
- Simple global rTMOs for imperfect input images

Conclusions

γ Expansion for Over-exposed Content

- rTMO for bright images with large saturated areas
- Simple global rTMOs for imperfect input images

- Darker HDR depictions are usually preferred for bright input images **[Meylan et al. 2006]**
- Contrast enhancement usually improves image quality **[Rempel et al. 2007; Banterle et al. 2009]**

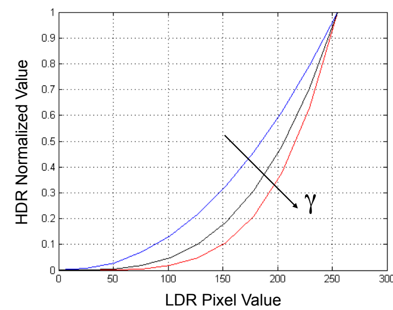
SIGGRAPHASIA2009

γ Expansion for Over-exposed Content

- rTMO for bright images with large saturated areas
- Simple global rTMOs for imperfect input images

γ expansion

✓ Darkens overall appearance
✓ Increases contrast
» Noise?



γ Expansion for Over-exposed Content

Ok, but… How do we choose γ for each image?

γ Expansion for Over-exposed Content

Ok, but… How do we choose γ for each image? With the **key** of the image

Conclusions

γ Expansion for Over-exposed Content

Ok, but… How do we choose γ for each image? With the **key** of the image

1. Key *[Akyüz and Reinhard 2006]* $k = \dfrac{\log L_{avg} - \log L_m}{\log L_M - \log L_m}$

$$\log L_{avg} = (\textstyle\sum_{x,y} \log(L(x,y) + \delta))/n$$

2. Pilot study adjusting γ manually to get best depiction

3. Best fit (see also [Masia et al. 2011])

SIGGRAPHASIA2009

γ Expansion for Over-exposed Content

Subjective evaluation – Experiment One (again)

γ Expansion for Over-exposed Content

Linear

LDR2HDR

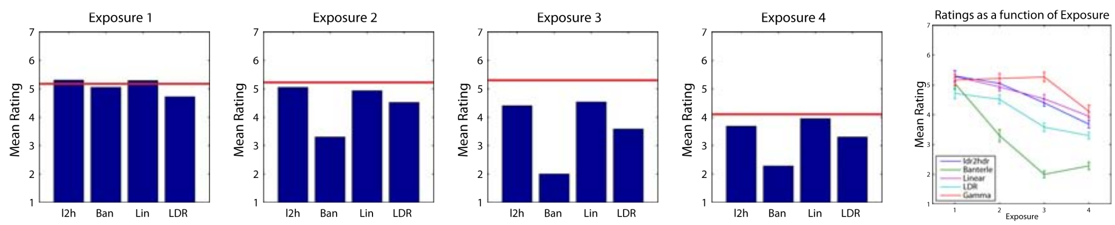
γ expansion

Banterle's

γ Expansion for Over-exposed Content

Subjective evaluation – Experiment One (again)

Exposure



Mean ratings – BRIGHT SERIES

γ Expansion for Over-exposed Content

Objective evaluation

γ Expansion for Over-exposed Content

Objective evaluation ➡ Dynamic range independent quality assessment **[Aydin et al. 2008]**

Conclusions

γ Expansion for Over-exposed Content

Objective evaluation ➡ Dynamic range independent quality assessment **[Aydin et al. 2008]**

■ Contrast reversal

SIGGRAPH ASIA 2009

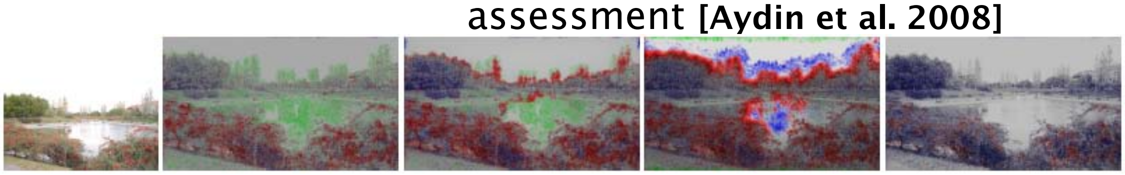
γ Expansion for Over-exposed Content

Objective evaluation → Dynamic range independent quality assessment **[Aydin et al. 2008]**

- Contrast reversal
- Loss of visible contrast

γ Expansion for Over-exposed Content

Objective evaluation → Dynamic range independent quality assessment **[Aydin et al. 2008]**



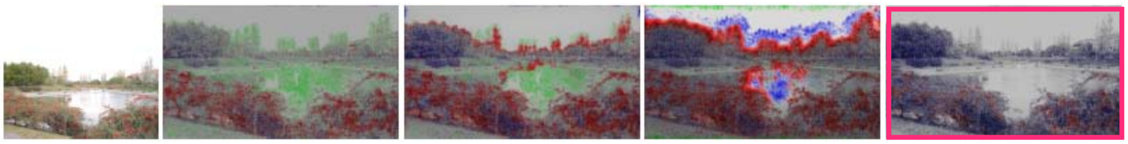
Source LDR image | Linear expansion | LDR2HDR | Banterle's operator | γ expansion

Source LDR image | Linear expansion | LDR2HDR | Banterle's operator | γ expansion

- Contrast reversal
- Loss of visible contrast
- Amplification of invisible contrast

γ Expansion for Over-exposed Content

Objective evaluation



Source LDR image | Linear expansion | LDR2HDR | Banterle's operator | **γ expansion**

Source LDR image | Linear expansion | LDR2HDR | Banterle's operator | **γ expansion**

- Contrast reversal
- Loss of visible contrast
- Amplification of invisible contrast

Recap & Conclusions

- Previous works on rTMO design have assumed correctly exposed input images
- Experiment One → for bright overexposed images the consensual approach of detecting and boosting saturated areas could be improved
- A γ expansion was proposed for these cases instead
  - Subjective evaluation
  - Objective evaluation
- Experiment Two → spatial artifacts are more important than inaccuracy in reproduced intensity levels

The future…

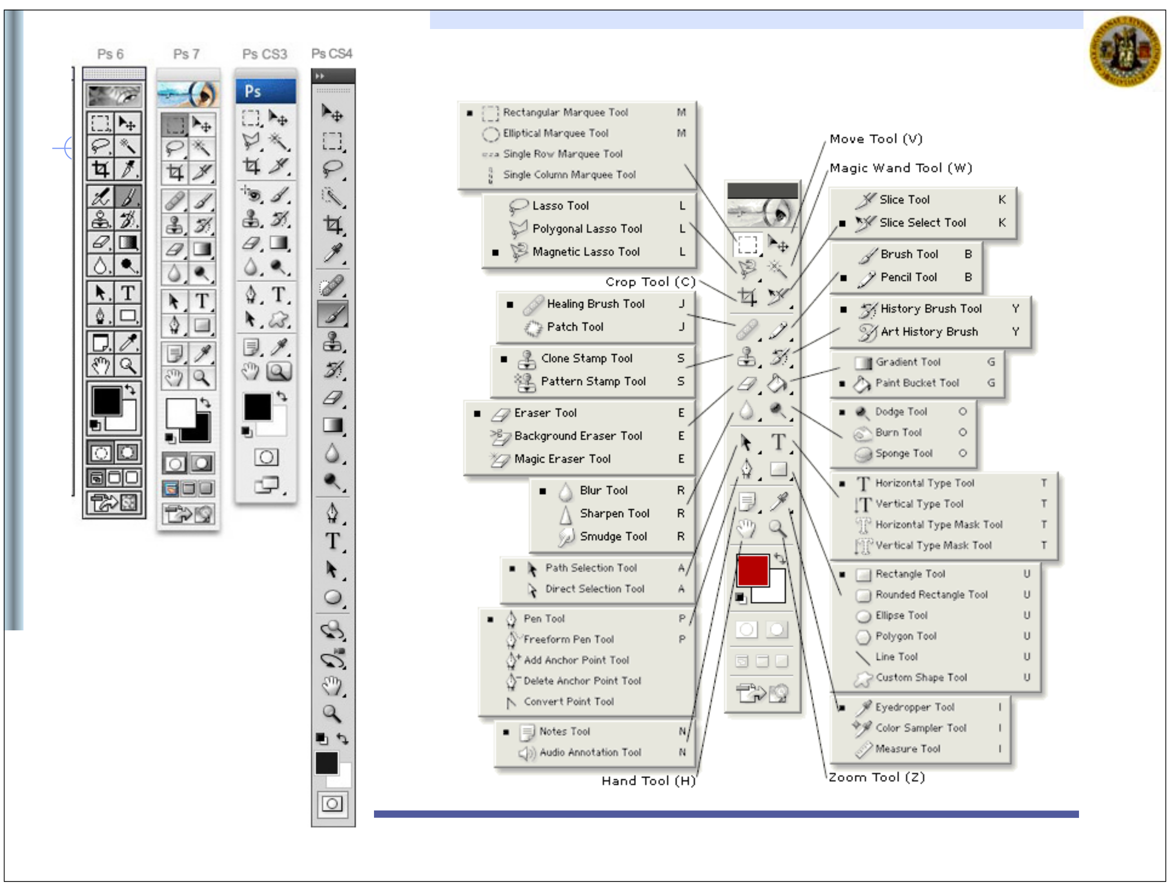
• This study aims to be valuable for further development of HDR display technology and for the design of future rTM algorithms

• Understanding the problem in its early stages is crucial for the development of new strategies → further tests are desirable

Diego Gutiérrez
Universidad de Zaragoza

Perception-based image editing

The first photograph: Joseph Nicéphore Niépce,1826. Source:http:// palimpsest.stanford.edu/byorg/abbey/an/an26/an26-3/ an26-307.html.



Ps 6
Ps 7
Ps CS3
Ps CS4
Rectangular Marquee Tool M
Elliptical Marquee Tool M
Single Row Marquee Tool
Single Column Marquee Tool
Lasso Tool L
Polygonal Lasso Tool L
Magnetic Lasso Tool L
Crop Tool (C)
Healing Brush Tool J
Patch Tool J
Clone Stamp Tool S
Pattern Stamp Tool S
Eraser Tool E
Background Eraser Tool E
Magic Eraser Tool E
Blur Tool R
Sharpen Tool R
Smudge Tool R
Path Selection Tool A
Direct Selection Tool A
Pen Tool P
Freeform Pen Tool P
Add Anchor Point Tool
Delete Anchor Point Tool
Convert Point Tool
Notes Tool N
Audio Annotation Tool N
Hand Tool (H)
Move Tool (V)
Magic Wand Tool (W)
Slice Tool K
Slice Select Tool K
Brush Tool B
Pencil Tool B
History Brush Tool Y
Art History Brush Y
Gradient Tool G
Paint Bucket Tool G
Dodge Tool O
Burn Tool O
Sponge Tool O
Horizontal Type Tool T
Vertical Type Tool T
Horizontal Type Mask Tool T
Vertical Type Mask Tool T
Rectangle Tool U
Rounded Rectangle Tool U
Ellipse Tool U
Polygon Tool U
Line Tool U
Custom Shape Tool U
Eyedropper Tool I
Color Sampler Tool I
Measure Tool I
Zoom Tool (Z)

ORIGINAL SHOT 1

ORIGINAL SHOT 2

FINAL, DIGITALLY COMBINED SHOT

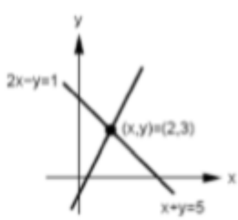
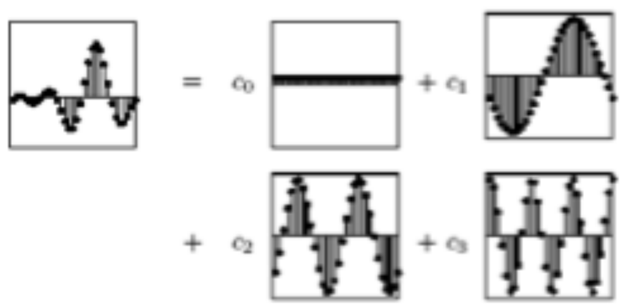
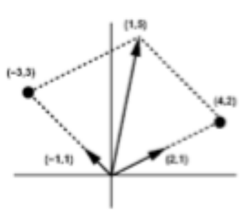


Figure 0.1 "Row" solution





$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{pmatrix}$$

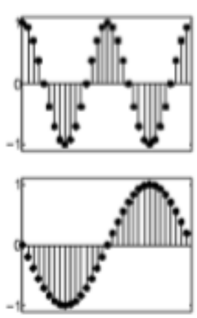
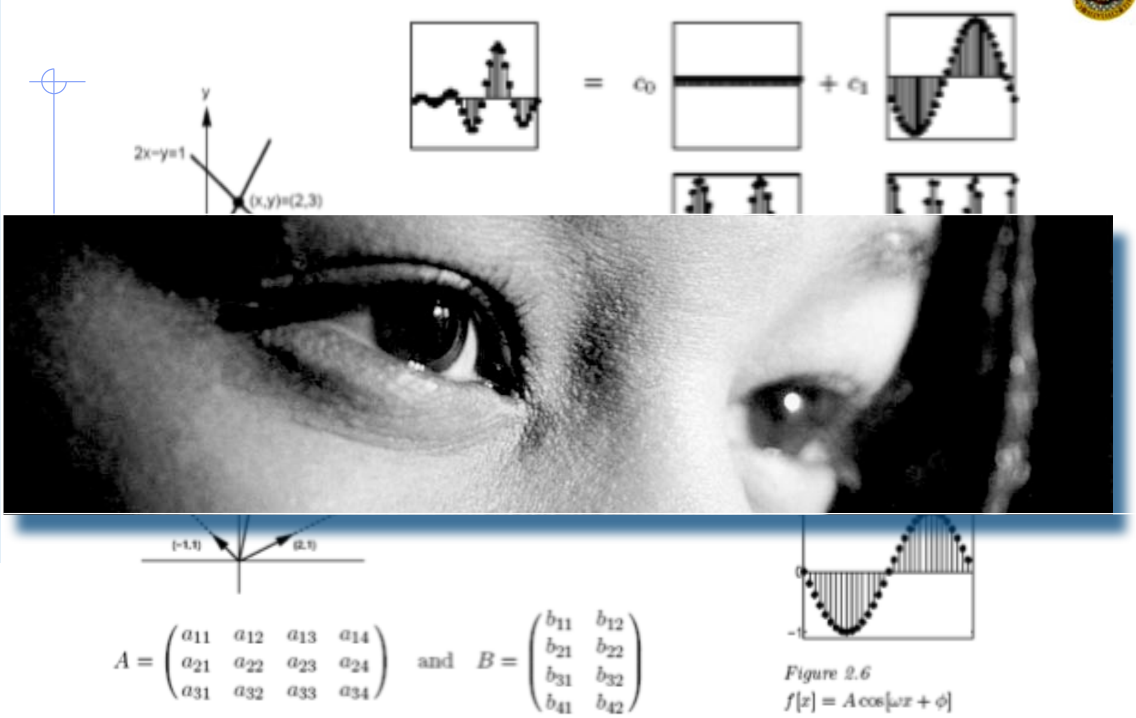
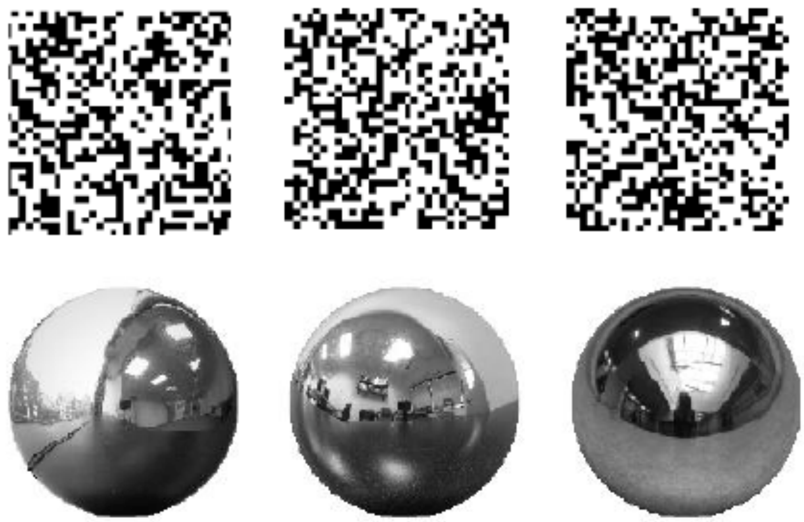


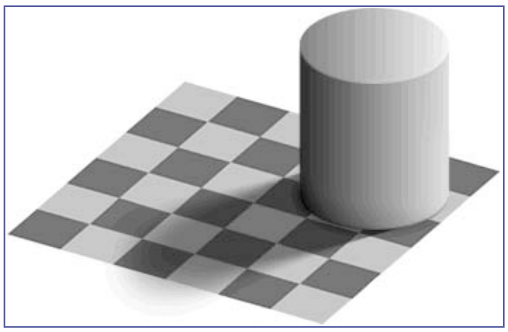
Figure 2.6
$f[x] = A\cos[\omega x + \phi]$

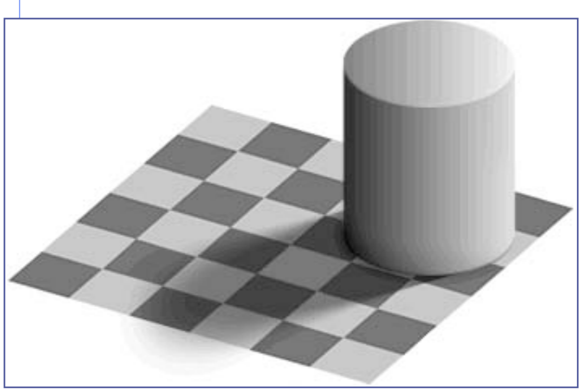


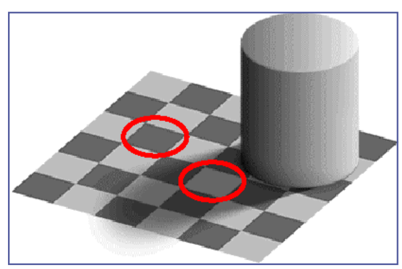
- [Adelson 2008]

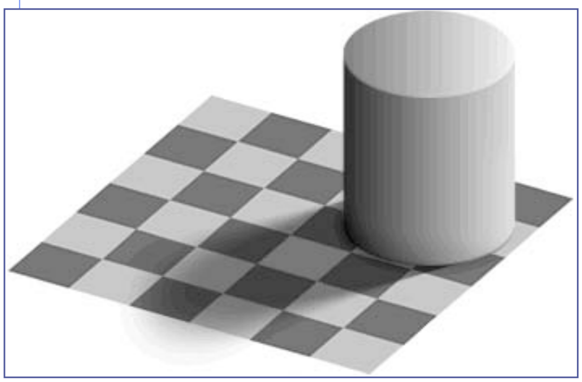
- [Adelson 2008]

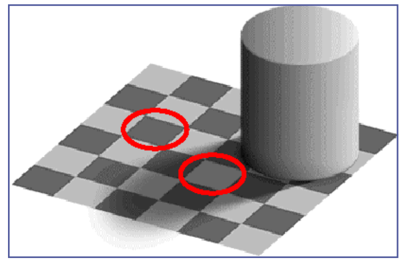


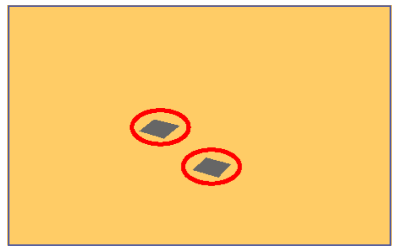




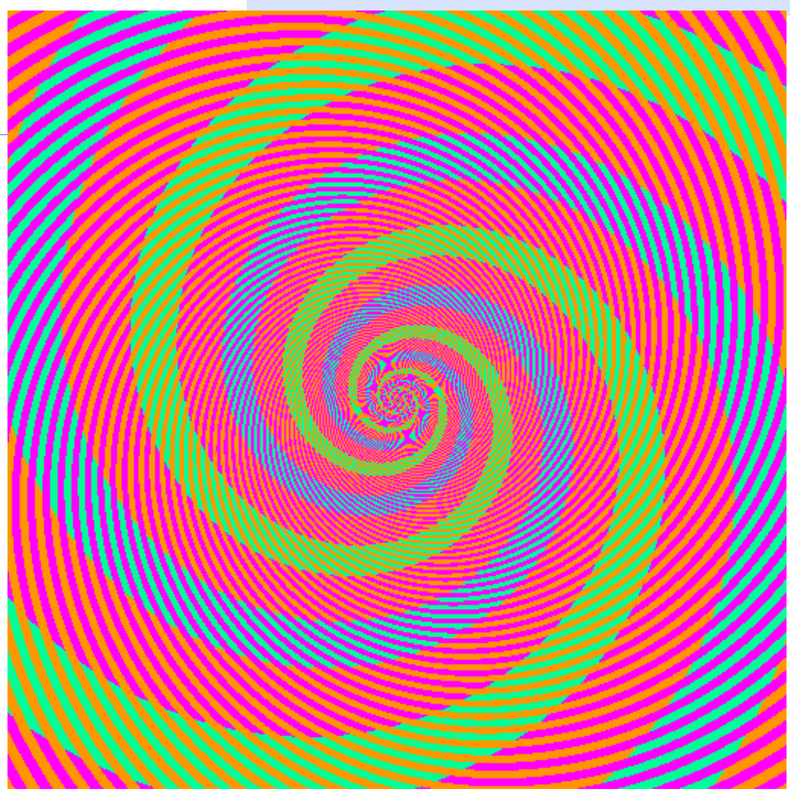




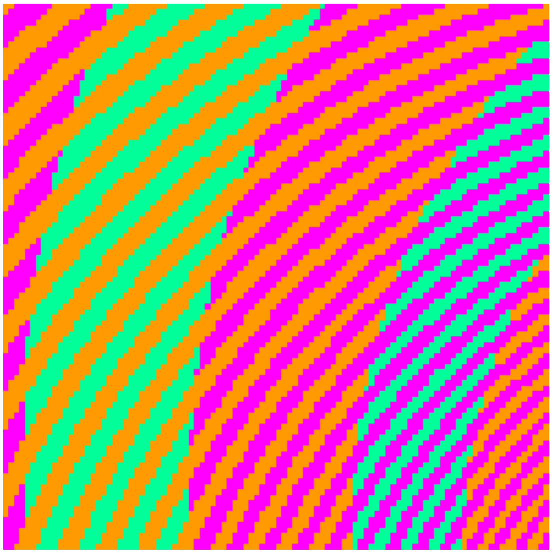


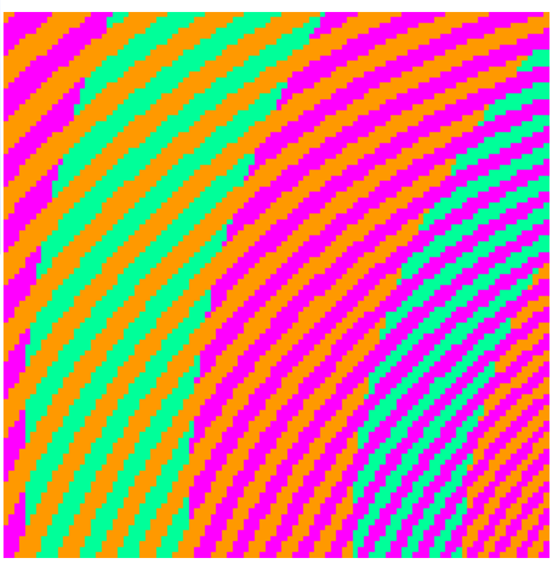


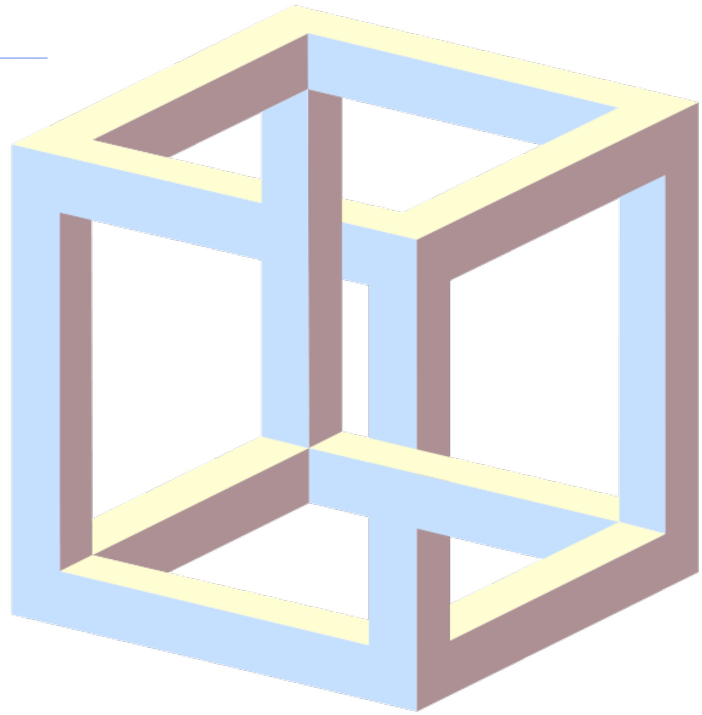
◆ [Mitra et al. 2009]

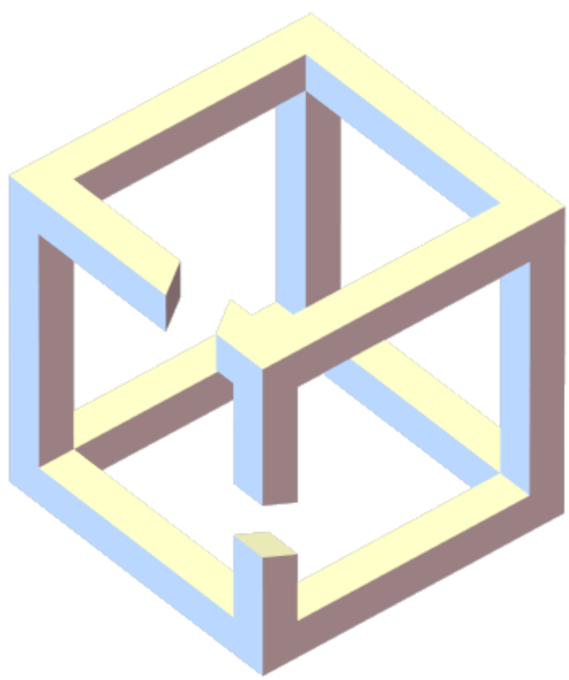


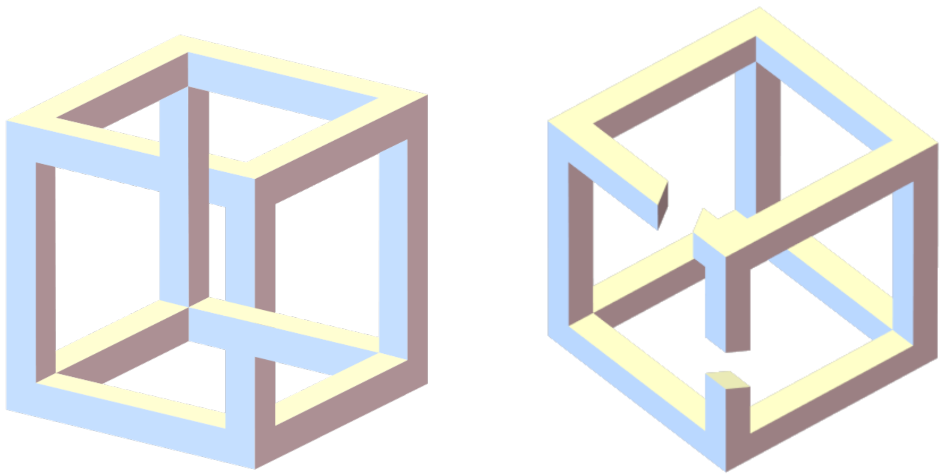
http://blogs.discovermagazine.com/badastronomy/2009/06/24/the-blue-and-the-green/







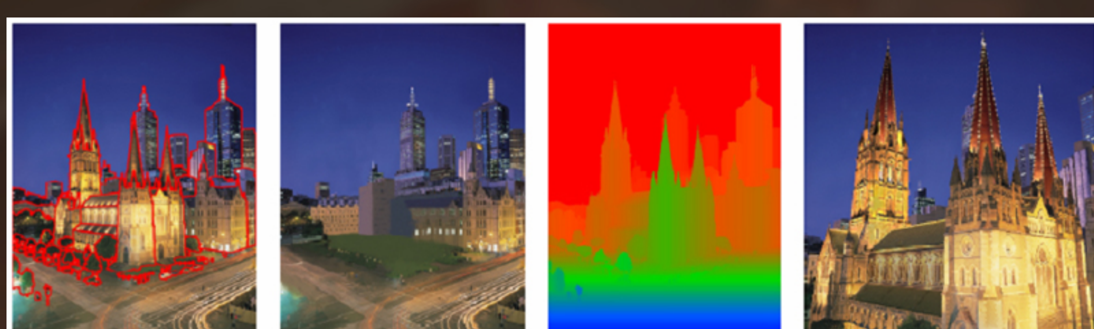




http://www.youtube.com/watch?v=PRRWGTridAY

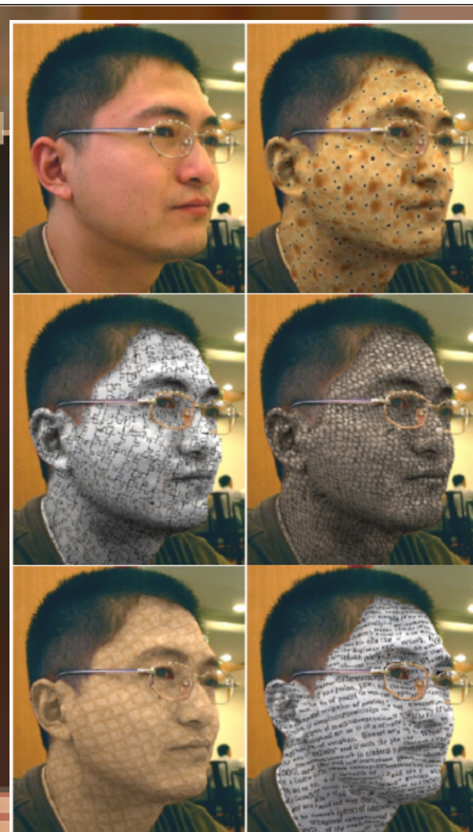
Overview

- [Oh et al. 2001]



Overview

- [Fang and Hart 2004]
- [Zelinka et al. 2005]
- [Fang and Hart 2006]



Overview

- [Khan et al. 2006]

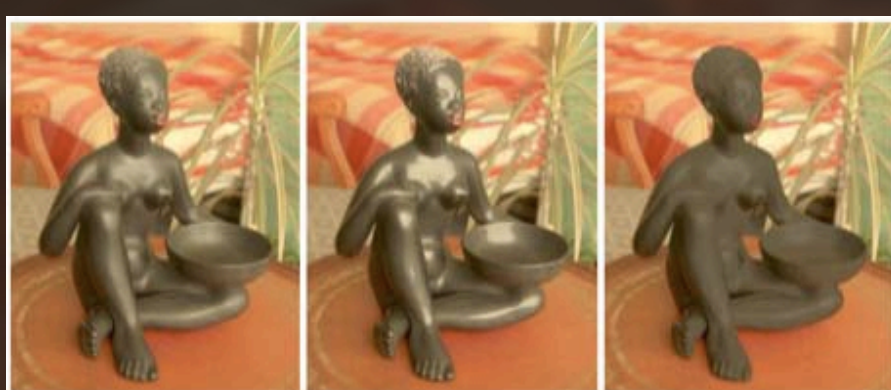


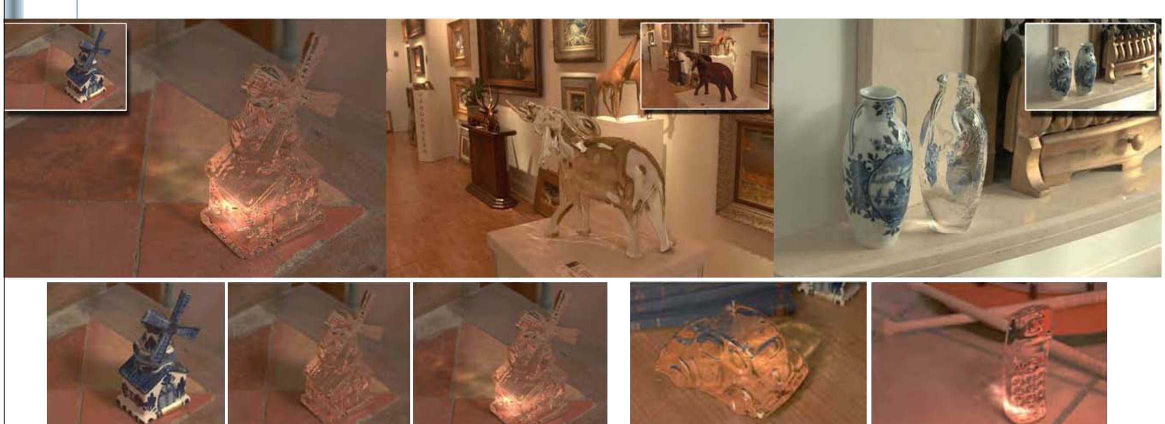
Image-Based Material Editing

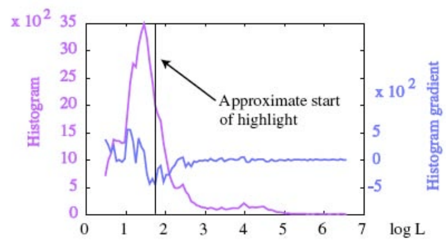
Erum Arif Khan[1] Erik Reinhard[2,1] Roland W. Fleming[3] Heinrich H. Bülthoff[3]
[1] University of Central Florida, [2] University of Bristol, [3] Max Planck Institute for Biological Cybernetics

Depicting procedural caustics in single images

D. Gutiérrez, J. López, F. Serón, M. Sánchez E. Reinhard
Universidad de Zaragoza University of Bristol

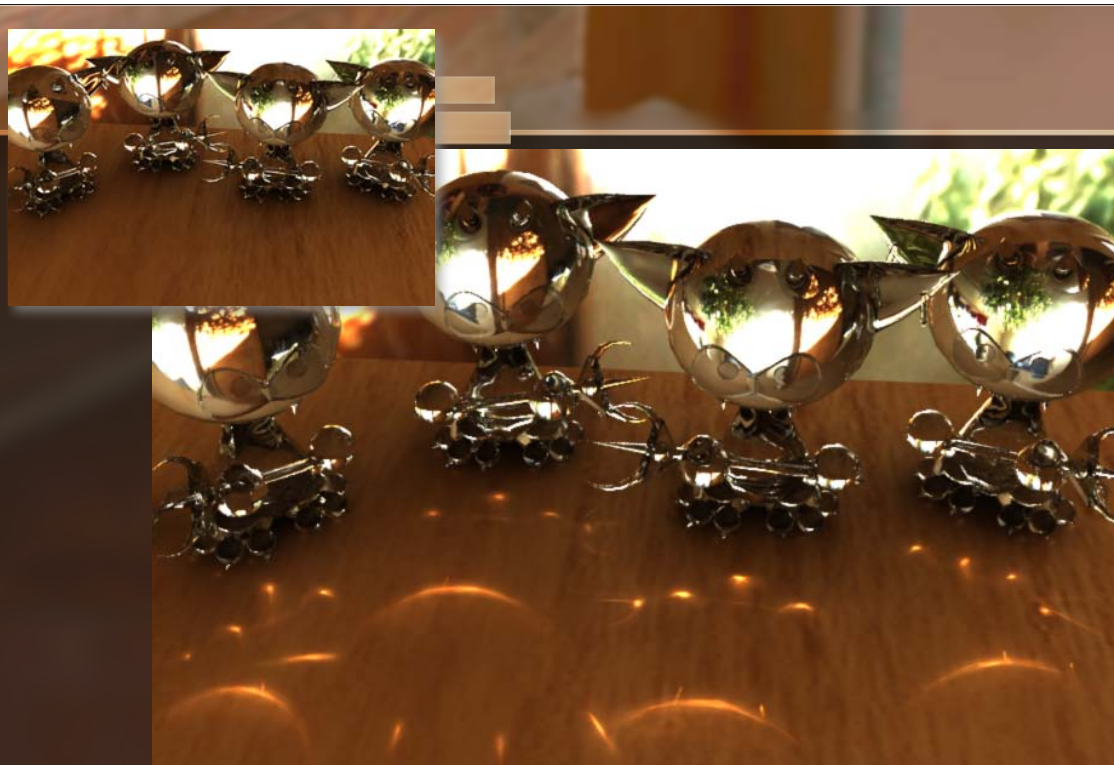
ACM Transactions on Graphics (SIGGRAPH Asia) 2008





Overview

- Why is this difficult?
- It´s difficult enough in 3D: L S* D E

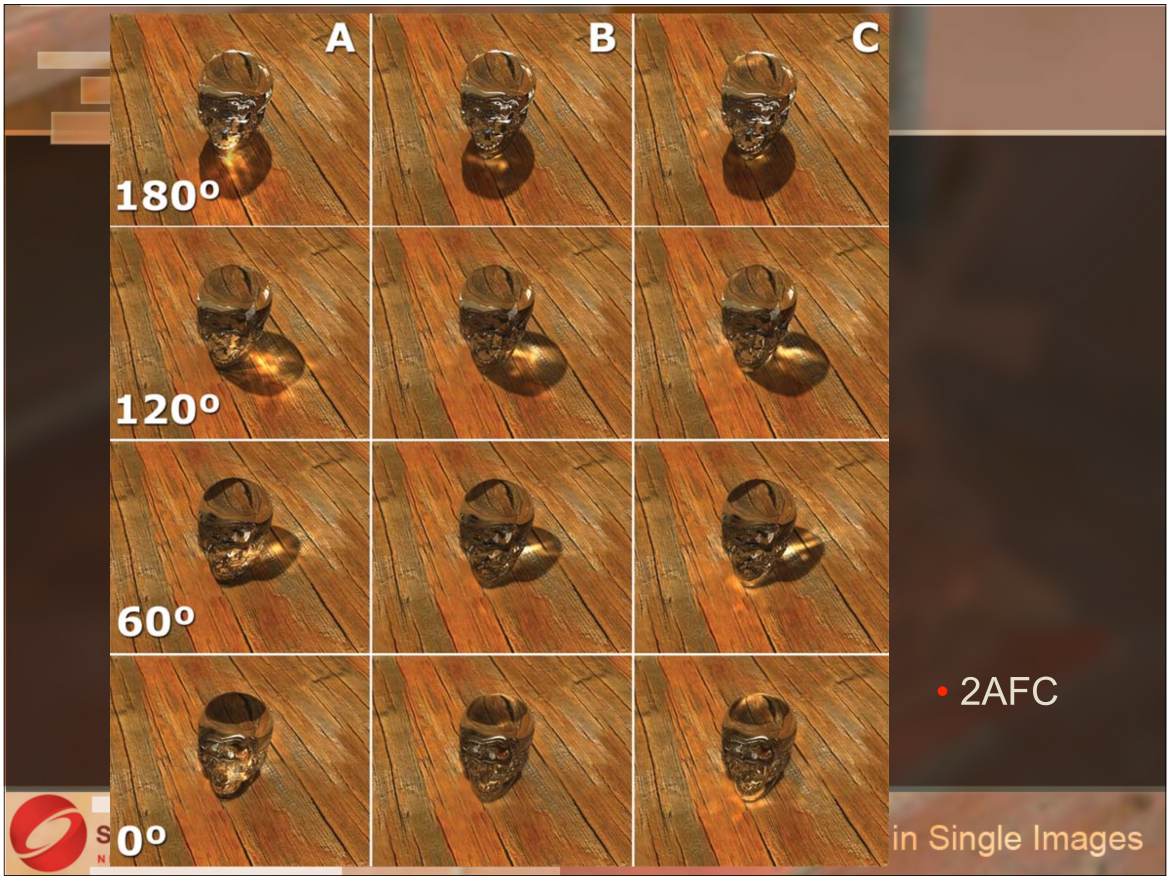


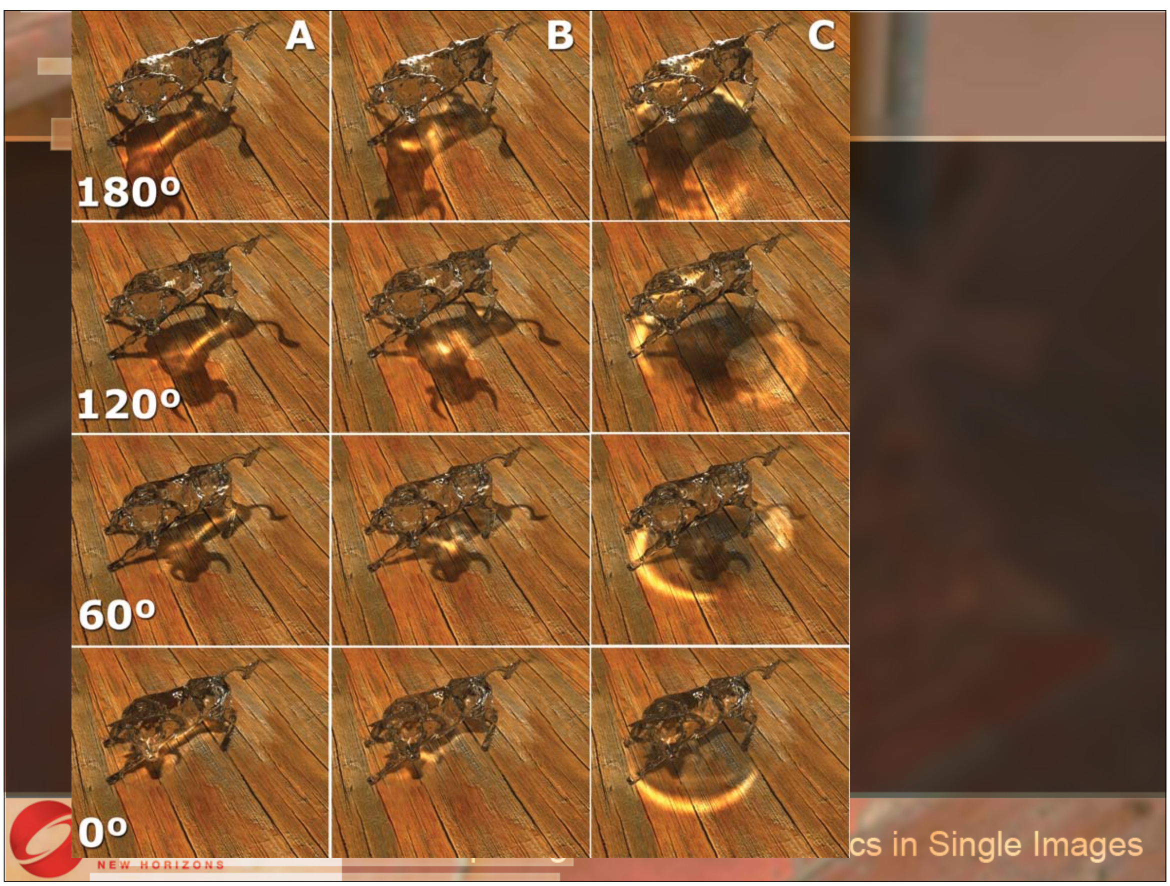
Results

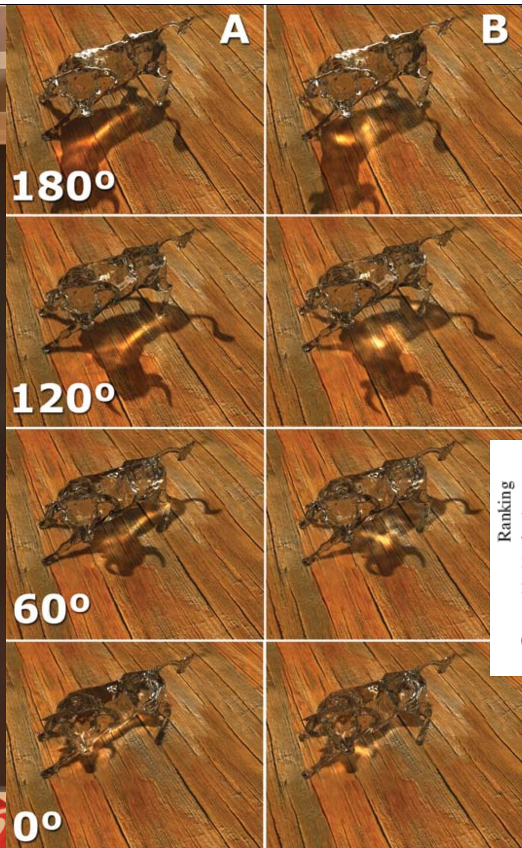
Psychophysics: Experiment 1

- As visually plausible as a full 3D photon mapping simulation?
- How would a simpler approach work?

SIGGRAPHASIA2008
NEW HORIZONS
Depicting Procedural Caustics in Single Images


A
B
C
180º
120º
60º
0º
• 2AFC
in Single Images

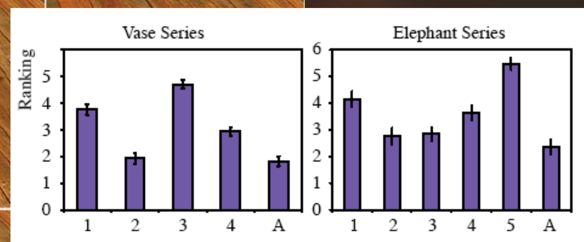

A
B
C
180º
120º
60º
0º
NEW HORIZONS
cs in Single Images



Scene	ξ	u	Angle	ξ	u
Skull	0.790	-0.068	0	0.903	0.040
Vertebrae	0.903	-0.047	60	0.903	0.044
Dolphin	0.966	0.240	120	0.909	0.021
Bull	0.972	0.249	180	0.914	0.020

Skull:	K PM R
Vertebrae:	K PM R
Dolphin:	K PM R
Bull:	K PM R

K = Kovesi Phase Symmetry
PM = Photon Mapping
R = Revolution Method



cs in Single Images

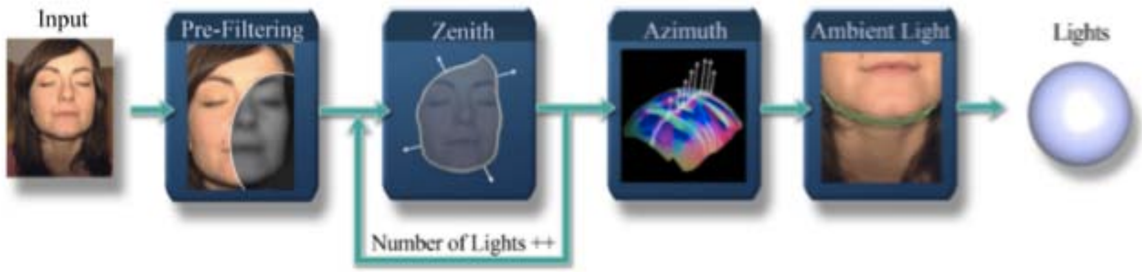
NEW HORIZONS

Compositing through light detection and relighting

J. López, D. Gutiérrez E. Reinhard S. Hadap
U. Zaragoza U. Bristol Adobe Inc.

Computers & Graphics (to appear, 2010)

Adobe



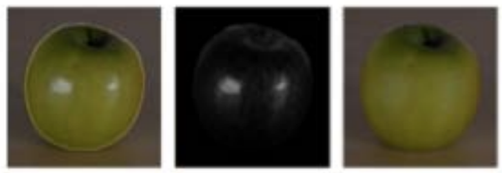
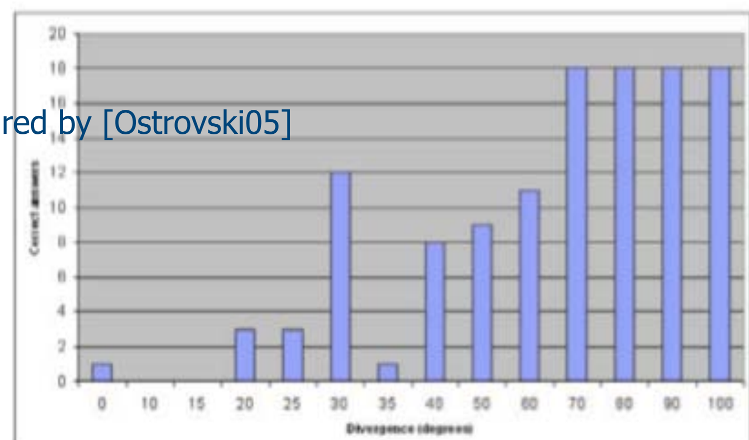
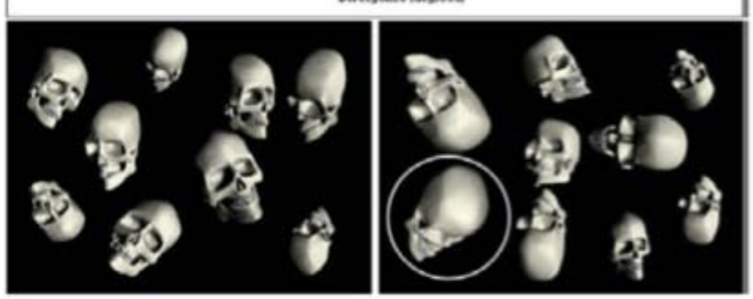
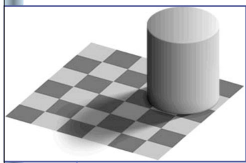
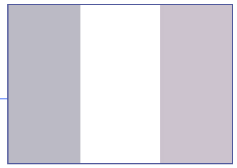
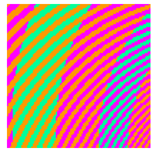
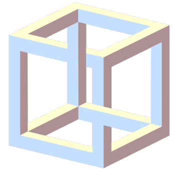


Figure 4: *Highlight removal. Left: Input image with the filtered area outlined in white. Middle: Image with the transformed color space. Right: Results.*

- Inspired by [Ostrovski05]





Adobe

Painting with light

J. López, D. Gutiérrez S. Hadap K. Anjyo

U. Zaragoza Adobe Inc. OLM Digital

Submitted to NPAR 2010

Diego Gutierrez

6 Realistic Characters, Faces and Animation

Characters & faces

Diego Gutierrez
Universidad de Zaragoza

Characters & faces

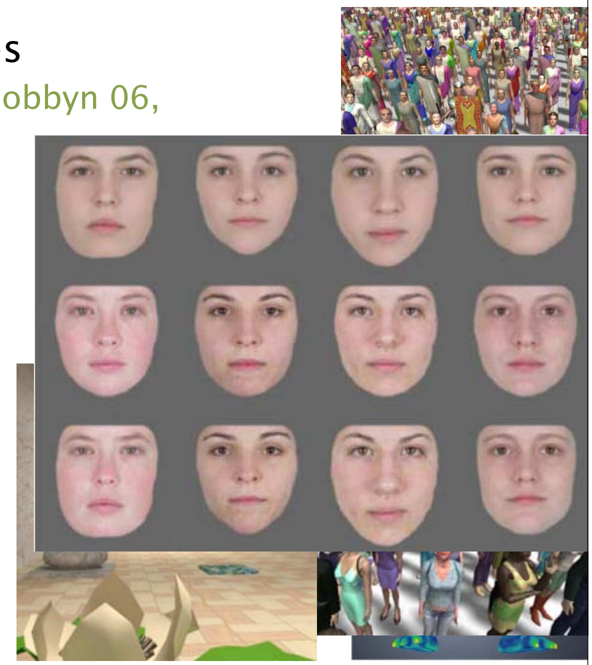
Diego Gutierrez
Universidad de Zaragoza

Many thanks to all the slides contributors:
R. McDonnell, C. Ennis, H. W. Jensen, J. Jimenez

Download the latest version from:
http://web.me.com/annmcnamara/PerceptionCourse2011/Home.ht

Background

- Crowd Variety Techniques
  - Tecchia 02, Gosselin 05, Dobbyn 06,
  - Maim 07, Thalmann 07
- Visual Attention
  - Yarbus 67, Itti 98, Yee 01
- Perception of Faces
  - Farah 92, Hochstein 04,
  - Nothdurft 93, Sinha 06



Clone Attack: Perception of Crowd Variety

R. McDonnell et al. 2008

Eye-catching Crowds: Saliency Based Selective Variation

R. McDonnell et al. 2009

Seeing is believing: Body motion dominates in multisensory conversations

C. Ennis et al. 2010

Results

- Appearance clones much easier to detect than motion clones

Body-part Saliency

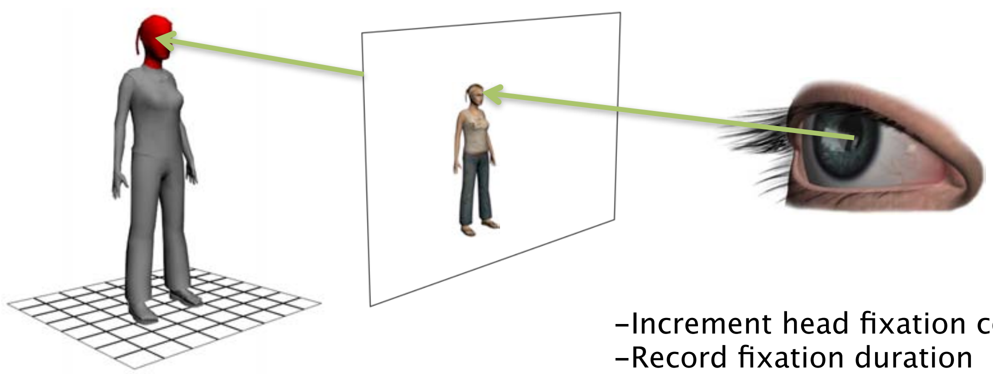
- What are the most salient parts of virtual characters?

Faces?

Hair style?

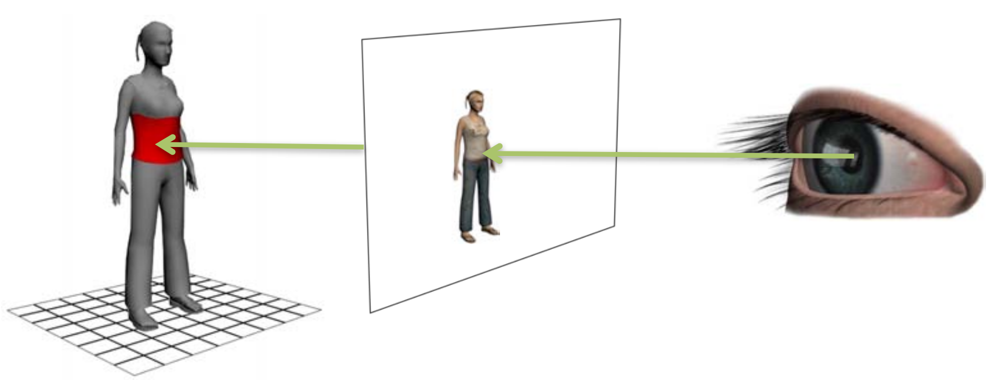
Clothing?

Setup

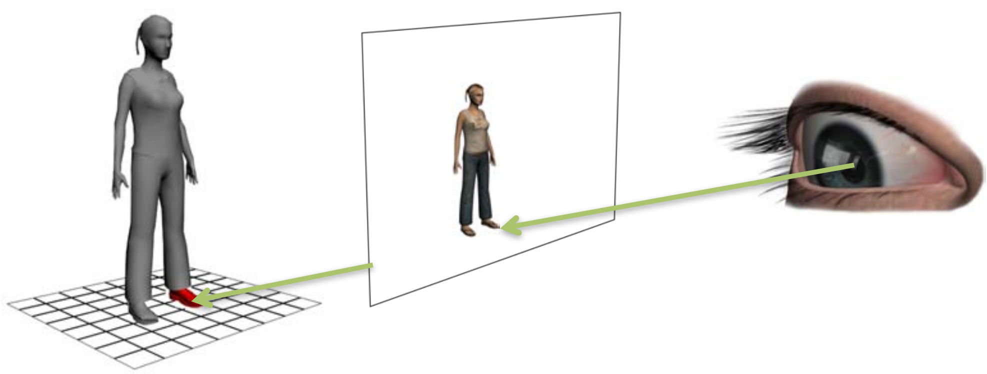


-Increment head fixation count
-Record fixation duration
-Add to running count for head

Setup



Setup



Full Variation

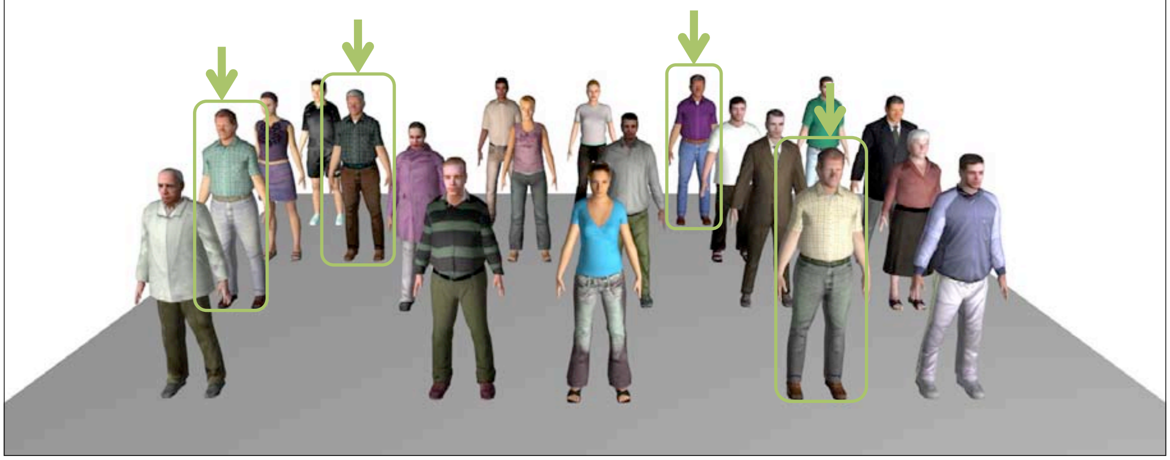
Selective Variation

Inverse-Selective Variation

No Variation

Trial Example

Trial Example



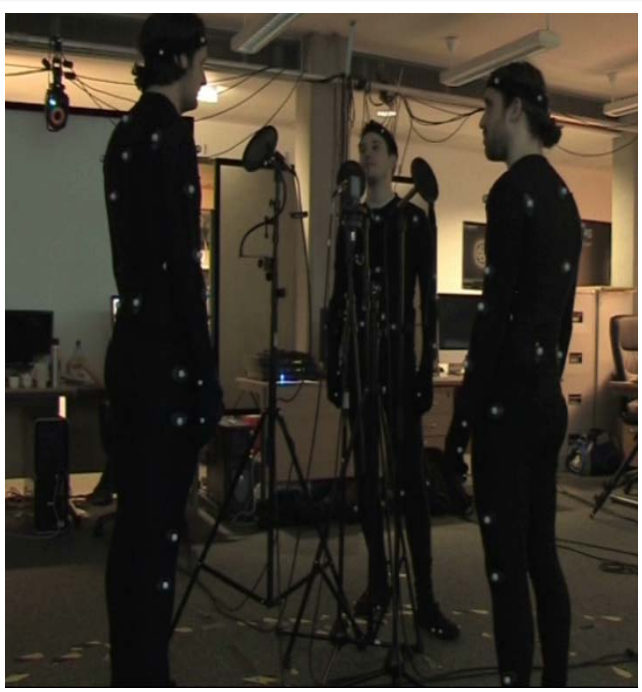
Motivation

- Perception of conversing groups
  - Cut scenes in video games
  - Real time conversational groups
  - Multisensory perception

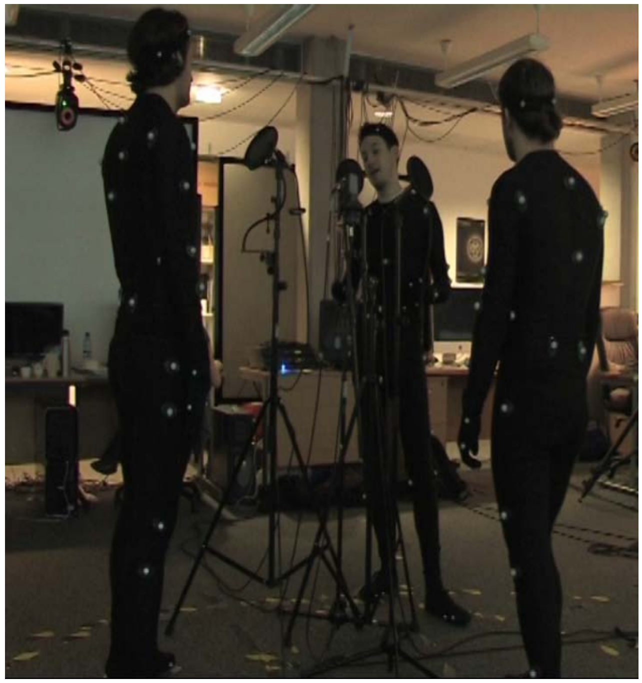
Stimuli Creation

- 2 actor sets:
  - 3 Male
  - 3 Female
- Motion capture:
  - 13 cameras
  - 52 markers per actor
- Audio recording:
  - Omni-directional microphone
  - 3 Condenser microphones

Debates



Dominant Conversations



Virtual Characters

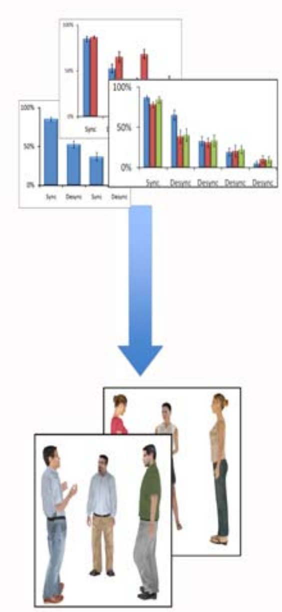
Experiment Design

- Does the presence of **audio** affect sensitivity to desynchronization?
  - 19 participants (11M, 9F)
- Will a richer **localized audio** signal help mask desynchronization better than mono audio?
  - 13 participants (8M, 5F)
- If there is **no audio** present, will we find similar results?
  - 12 participants (10M, 2F)

Guidelines

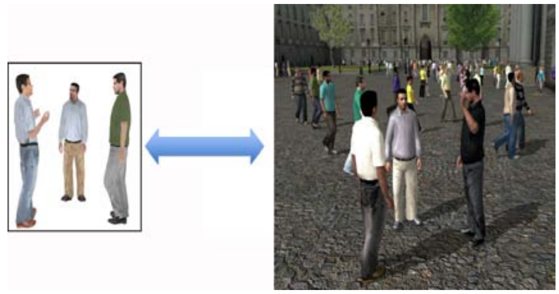
- Audio does not mask desynchronization in general
- Match one character to audio
- Ensure talker / listener coordination
  - Dominant easiest to mix and match

Results hold for males and females



Future Work

- Scaling to large crowds and outdoors
- Detailed animation for close-ups
- Group dynamics



Faces

•

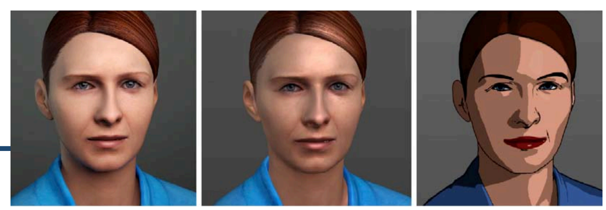
Images from [Miranda et al. 2011]

Face Reality: Investigating the Uncanny Valley for virtual faces

Rachel McDonnell*
Trinity College Dublin

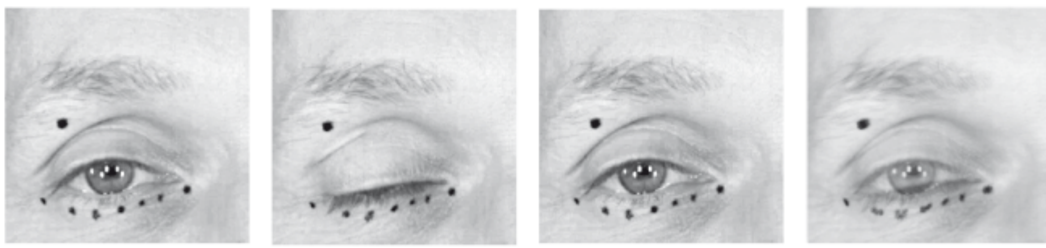
Martin Breidt†
Max Planck Institute for Biological Cybernetics

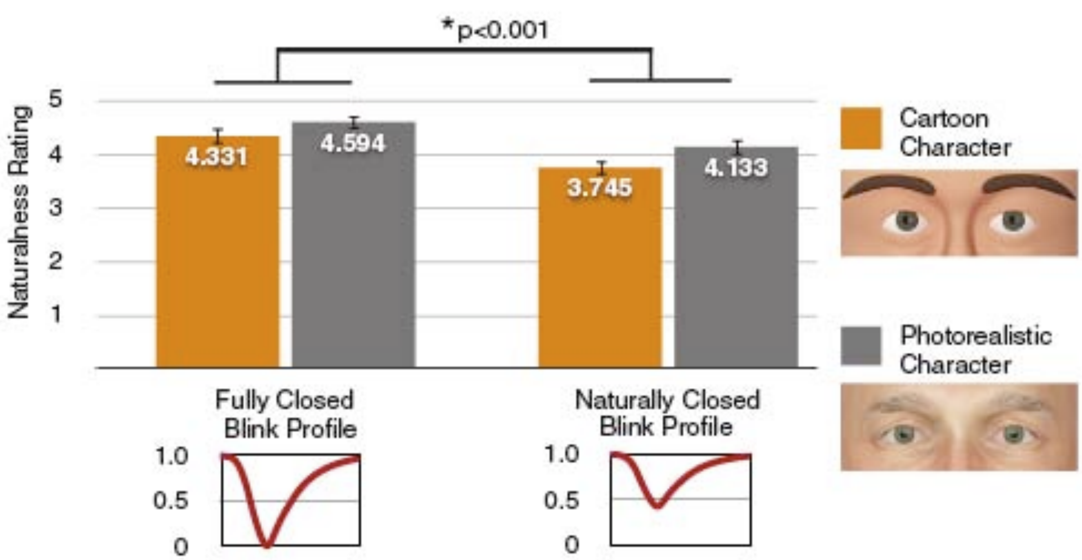
Figure 1: *Our virtual model rendered in different visual styles: (left) High Quality, (middle) Game Quality, (right) NPR.*

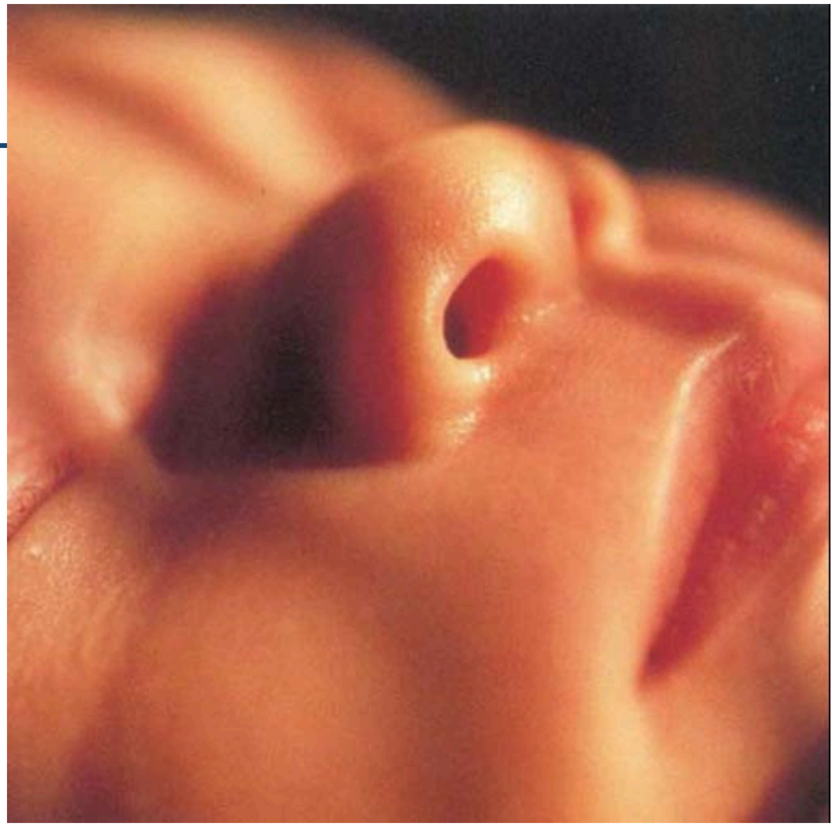


- Lying or telling the truth?
- High quality (HQ) was caught lying more often than NPR
- Subtle cues easier to detect in HQ?

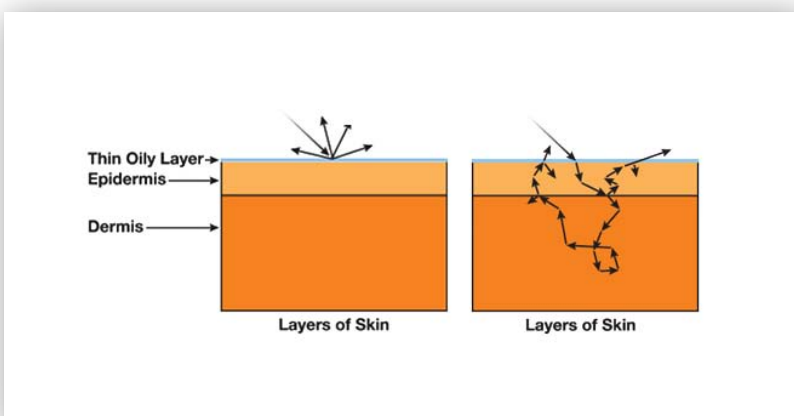
Modeling and animating eye blinks
Trutoiu et al. 2011







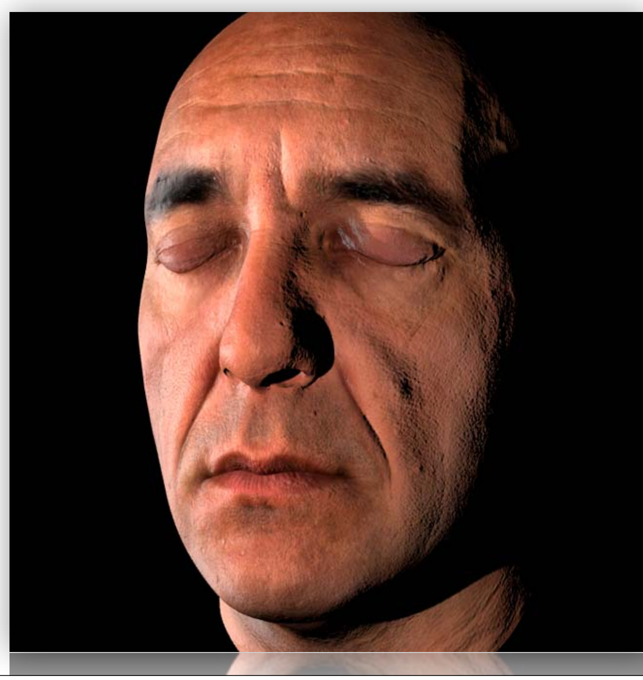
Subsurface scattering



Subsurface scattering

Slide by Henrik Wann Jensen

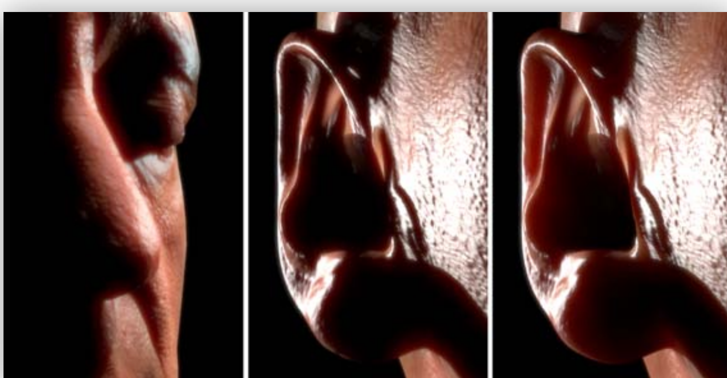
Subsurface scattering

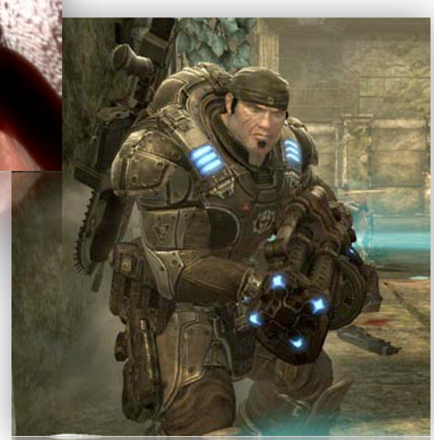


Screen-Space Perceptual Rendering of Human Skin
Jorge Jimenez, Veronica Sundstedt and Diego Gutierrez
ACM Transactions on Applied Perception, Vol. 6(4), Article 23

- 2AFC (no reference)
- Task description
  - Choose the image which you think look most realistic (i.e. most like real human skin)
  - Unlimited viewing time (10 sec recommended)
  - Person has the same age (45 years old)

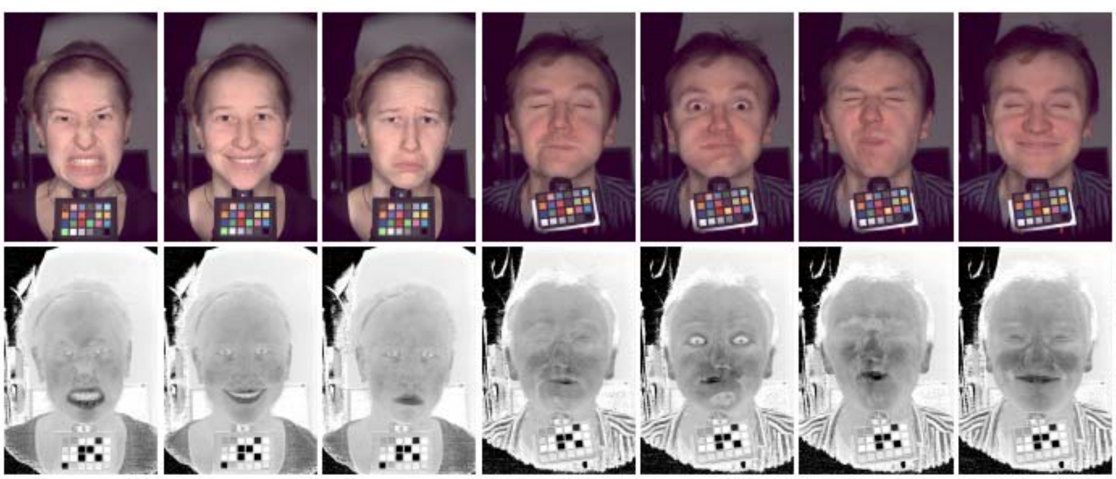
Faces



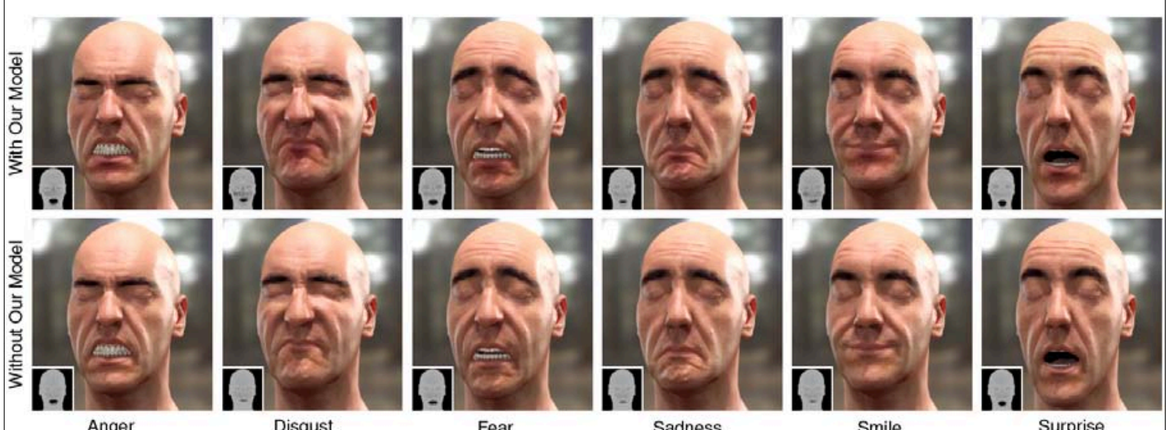


A Practical Appearance Model for Dynamic Facial Color
Jorge Jimenez, Timothy Scully, Nuno Barbosa, Craig Donner, Xenxo Alvarez, Teresa Vieira, Paul Matts, Veronica Orvalho, Diego Gutierrez and Tim Weyrich
ACM Transactions on Graphics, Vol. 29(5) (SIGGRAPH Asia 2010)

Faces



Faces



With Our Model
Without Our Model
Anger
Disgust
Fear
Sadness
Smile
Surprise

APGV

7 Trends from the 8th annual ACM/Eurographics Symposium on Applied Perception in Graphics and Visualization 2011

7.1 APGV Proceedings and Resources

In 2001 a small group of researchers gathered in Snowbird for a “campfire” on Graphics and Perception, there was a certain air of excitement that things were getting started. The growing interest in this area, and the realization that perception is playing an increasingly important role in graphics and visualization lead to the establishment of a symposium dedicated to perceptual research in graphics and visualization, called APGV - Applied Perception in Graphics and Visualization. The goal for this symposium is to have it serve as an inclusive forum where researchers working at the intersection of perception, graphics and visualization can come together to share ideas and results. APGV hopes to provide a great opportunity for people not only to acquire new knowledge, but also to seek new partnerships and collaborations. Now in it’s 7th year papers of the years have represented active interdisciplinary efforts. A wide range of topics have treated . including color, shape, motion, distance judgments, virtual reality, and haptics, as well as application areas such as product design and medicine. Please see http://www.apgv.org and this years proceedings for further information, [22, 20, 21, 23, 96, 95, 112].

References

[1] B. D. Adelstein, T. G. Lee, and S. R. Ellis. Head tracking latency in virtual environments: Psychophysics and a model. *In Proc. of the 47th Annual Human Factors and Ergonomics Society meeting*, pages 2083–2087, 2003.

[2] Stéphane Albin, Gilles Rougeron, Bernard Péroche, and Alain Trémeau. Quality image metrics for synthetic images based on perceptual color differences. *IEEE Transactions on Image Processing*, 11(9):961–971, 2002.

[3] Robert S. Allison, Laurence R. Harris, Michael Jenkin, Urszula Jasiobedzka, and James E. Zacher. Tolerance of temporal delay in virtual environments. In *VR '01: Proceedings of the Virtual Reality 2001 Conference (VR'01)*, page 247, Washington, DC, USA, 2001. IEEE Computer Society.

[4] D. Ariely. Seeing sets: Representation by statistical properties. *Psychological Science*, 12(2):157–162, 2001.

[5] Ronald Azuma and Gary Bishop. Improving static and dynamic registration in an optical see-through hmd. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 197–204, New York, NY, USA, 1994. ACM.

[6] J. H. Bailey and B. G. Witmer. Learning and transfer of spatial knowledge in a virtual environment. *Proc. of the Human Factors and Ergonomics Society 38th Annual Meeting*, pages 1158–1162, 1994.

[7] Dirk Bartz, Douglas Cunningham, Jan Fischer, and Christian Wallraven. Star state of the art report the role of perception for computer graphics, 2008.

[8] James P. Bliss, Philip D. Tidwell, and Michael A. Guest. The effectiveness of virtual reality for administering spatial navigation training to firefighters. *Presence*, 6(1):73–86, 1997.

[9] Mark R. Bolin and Gary W. Meyer. A perceptually based adaptive sampling algorithm. *Computer Graphics*, 32(Annual Conference Series):299–309, August 1998.

[10] M.R. Bolin and G.W. Meyer. A frequency based ray tracer. In *ACM SIGGRAPH '95 Conference Proceedings*, pages 409–418, 1995.

[11] K. R. Brandt, J.M. Gardiner, and C. N. Macrae. The distinctiveness effect in forenames: The role of subjective experiences and recognition memory. *British Journal of Psychology*, pages 269–280, 2006.

[12] K. R. Brandt, C.N. Macrae, A.M. Schloerscheidt, and A. B. Milne. Do i know you? target typicality and recollective experience. *Memory*, 11(1):89–100, 2003.

[13] W. Brewer and J. Treyens. Role of schemata in memory for places. *Cognitive Psychology*, 13:207–230, 1981.

[14] D. E. Broadbent and M. H. P. Broadbent. From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception and Psychophysics*, 42(4):105–113, 1987.

[15] T. C. Callaghan. Dimensional interaction of hue and brightness in preattentive field segregation. *Perception & Psychophysics*, 36(1):25–34, 1984.

[16] T. C. Callaghan. Interference and domination in texture segregation: Hue, geometric form, and line orientation. *Perception & Psychophysics*, 46(4):299–311, 1989.

[17] T. C. Callaghan. Interference and dominance in texture segregation. In D. Brogan, editor, *Visual Search*, pages 81–87. Taylor & Francis, New York, New York, 1990.

[18] J. Cataliotti and A. Gilchrist. Local and global processes in lightness perception. In *Perception and Psychophysics*, volume 57(2), pages 125–135. Perception, 1995.

[19] K. Cater, A. Chalmers, and G. Ward. Detail to attention: exploiting visual tasks for selective rendering. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering*, pages 270–280, Aire-la-Ville, Switzerland, Switzerland, 2003. Eurographics Association.

[20] Heinrich Conference Chair-Bülthoff and Tom Conference Chair-Troscianko. Apgv '05: Proceedings of the 2nd symposium on applied perception in graphics and visualization. *A Coro na, Spain*, 2005.

[21] Roland W. Conference Chair-Fleming and Sunghee Conference Chair-Kim. Apgv '06: Proceedings of the 3rd symposium on applied perception in graphics and visualization. *Boston, Massachusetts*, 2006.

[22] Victoria Conference Chair-Interrante, Ann Conference Chair-McNamara, Heinrich Program Chair-Bulthoff, and Holly Program Chair-Rushmeier. Apgv '04: Proceedings of the 1st symposium on applied perception in graphics and visualization. *Los Angeles, California*, 2004.

[23] Christian Conference Chair-Wallraven and Veronica Conference Chair-Sundstedt. Apgv '07: Proceedings of the 4th symposium on applied perception in graphics and visualization. *Tubingen, Germany*, 2007.

[24] et AL. Conway, M. A. Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology : General*, 126(4):393 413, 1997.

[25] Scott Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. In *Digital images and human vision*, pages 179–206, Cambridge, MA, USA, 1993. MIT Press.

[26] Kate Devlin, Alan Chalmers, Alexander Wilkie, and Werner Purgathofer. Star: Tone reproduction and physically based spectral rendering. In Dieter Fellner and Roberto Scopignio, editors, *State of the Art Reports, Eurographics 2002*, pages 101–123, Vienna, September 2002. The Eurographics Association.

[27] S. A. Dewhurst, S. J. Holmes, K. R. Brandt, and G. M. Dean. Measuring the speed of the conscious components of recognition memory: Remembering is faster than knowing. *Consciousness and Cognition*, 15:147–162, 2006.

[28] Huong Q. Dinh, Neff Walker, Chang Song, Akira Kobayashi, and Larry F. Hodges. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *VR '99: Proceedings of the IEEE Virtual Reality*, page 222, Washington, DC, USA, 1999. IEEE Computer Society.

[29] George Drettakis, Nicolas Bonneel, Carsten Dachsbacher, Sylvain Lefebvre, Michael Schwarz, and Isabelle Viaud-Delmon. An interactive perceptual rendering pipeline using contrast and spatial masking. In *Rendering Techniques (Proceedings of the Eurographics Symposium on Rendering)*. Eurographics, June 2007.

[30] S. R. Ellis, K. Mania, B. D. Adelstein, and M. I. Hill. Generalizability of latency detection in a variety of virtual environments. *In Proc. of the 48th Annual Human Factors and Ergonomics Society meeting*, 2004.

[31] S. R. Ellis, M. J. Young, S. M. Ehrlich, and B. D. Adelstein. Discrimination of changes of rendering latency during voluntary hand movement. *In Proc. of the 43th Annual Human Factors and Ergonomics Society meeting*, pages 1182–1186, 1999.

[32] S.R. Ellis, A. Wolfram, and B.D. Adelstein. Large amplitude three-dimensional tracking in augmented environments: a human performance trade-off between system latency and update rate. *In Proc. of the 46th Annual Human Factors and Ergonomics Society meeting*, 2002.

[33] Stephen R. Ellis, F. Breant, Brian M. Menges, Richard H. Jacoby, and Bernard D. Adelstein. Operator interaction with visual objects: Effect of system latency. *HCI (2)*, pages 973–976, 1997.

[34] Stephen R. Ellis, Mark J. Young, Bernard D. Adelstein, and Sheryl M. Ehrlich. Discrimination of changes in latency during head movement. *HCI (2)*, pages 1129–1133, 1999.

[35] Land. M. F. Motion and vision: why animals move their eyes. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioural Physiology*, 185(4):341–352, 1999.

[36] Jean-Philippe Farrugia and Bernard Péroche. A progressive rendering algorithm using an adaptive perceptually based image metric. *Comput. Graph. Forum*, 23(3):605–614, 2004.

[37] J. Ferwerda. Hi-fi rendering. *ACM Siggraph Eurographics campfire on perceptually adaptive graphics. http://isg.cs.tcd.ie/campfire/jimferwerda2.html*, 2001.

[38] Philip W. Fink, Patrick S. Foo, and William H. Warren. Obstacle avoidance during walking in real and virtual environments. *ACM Trans. Appl. Percept.*, 4(1):2, 2007.

[39] R. Flannery, K.A. Walles. How does schema theory apply to real versus virtual memories? *Cyberspychology and Behavior*, 6(2):151–159, 2003.

[40] Harald Frenz, Markus Lappe, Marina Kolesnik, and Thomas Bührmann. Estimation of travel distance from visual motion in virtual environments. *ACM Trans. Appl. Percept.*, 4(1):3, 2007.

[41] J. M. Gardiner and A. Richardson-klavehn. Remembering and knowing. *In: Tulving, E. and Craik, F. I. M., eds.Handbook of Memory*, 1992.

[42] S. Gibson and R.J. Hubbold. Efficient hierarchical refinement and clustering for radiosity in complex environments. *Computer Graphics Forum*, 15(5):297–310, 1996.

[43] A. Gilchrist, S. Delman, and A. Jacobsen. The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception and Psychophysics*, 33(5):425–436, 1983.

[44] A. L. Gilchrist. The perception of surface blacks and whites. *Scientific American*, 240(3):88–97, March 1979.

[45] Alan Gilchrist and Alan Jacobsen. Perception of lightness and illumination in a world of one reflectance. *Perception*, 13:5–19, 1984.

[46] Baining Guo. Progressive radiance evaluation using directional coherence maps. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 255–266, New York, NY, USA, 1998. ACM.

[47] Jörg Haber, Karol Myszkowski, Hitoshi Yamauchi, and Hans-Peter Seidel. Perceptually guided corrective splatting. *Computer Graphics Forum*, 20(3), 2001.

[48] Christopher G. Healey and James T. Enns. Building perceptual textures to visualize multidimensional datasets. In *Proceedings Visualization '98*, pages 111–118, Research Triangle Park, North Carolina, 1998.

[49] Christopher G. Healey and James T. Enns. Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(2):145–167, 1999.

[50] David Hedley, Adam Worrall, and Derek Paddon. Selective culling of discontinuity lines. In Julie Dorsey and Philipp Slusallek, editors, *Rendering Techniques '97 (Proceedings of the Eighth Eurographics Workshop on Rendering)*, pages 69–80. Springer Wien, 1997. ISBN 3-211-83001-4.

[51] R. Held, A. Efsathiouand, and M Greene. Adaptation to displaced and delayed visual feedback from the hand. *Journal of Experimental Psychology*, 72(6):887–891, 1966.

[52] L. Huang and H. Pashler. A boolean map theory of visual attention. *Psychological Review*, 114(3):599–631, 2007.

[53] L. Huang, A. Treisman, and H. Pashler. Characterizing the limits of human visual awareness. *Science*, 317:823–825, 2007.

[54] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, 1998.

[55] B. Julész. *Foundations of Cyclopean Perception.* University of Chicago Press, Chicago, Illinois, 1971.

[56] B. Julész. Experiments in the visual perception of texture. *Scientific American*, 232:34–43, 1975.

[57] B. Julész. A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 41:131–138, 1981.

[58] B. Julész, E. N. Gilbert, and L. A. Shepp. Inability of humans to discriminate between visual textures that agree in second-order statistics—revisited. *Perception*, 2:391–405, 1973.

[59] B. Julész, E. N. Gilbert, and J. D. Victor. Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31:137–140, 1978.

[60] J. Y. Jung, B. D. Adelstein, and S. R. Ellis. Discriminability of prediction artifacts in a time delayed virtual environment. *In Proc. of the 44th Annual Human Factors and Ergonomics Society meeting*, pages 499–502, 2000.

[61] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Lightness perception in tone reproduction for high dynamic range images. In *The European Association for Computer Graphics 26th Annual Conference EUROGRAPHICS 2005*, volume 24 of *Computer Graphics Forum*, pages xx–xx, Dublin, Ireland, 2005. Blackwell.

[62] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel. Computational model of lightness perception in high dynamic range imaging. In Bernice E. Rogowitz, Thrasyvoulos N. Pappas, and Scott J. Daly, editors, *Human Vision and Electronic Imaging XI, IS&T SPIE's 18th Annual Symposium on Electronic Imaging (2006)*, volume xxxx, pages xxx–xxx, 2006.

[63] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, October 1997.

[64] William B. Lathrop and Mary K. Kaiser. Perceived orientation in physical and virtual environments: changes in perceived orientation as a function of idiothetic information available. *Presence: Teleoper. Virtual Environ.*, 11(1):19–32, 2002.

[65] Guillaume Lavoué. A roughness measure for 3d mesh visual masking. In *APGV '07: Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 57–60, New York, NY, USA, 2007. ACM.

[66] A. Liu, S. Tharp, L. Lai, French, and L. Stark. Some of what one needs to know about using head-mounted displays to improve teleoperator performance. *IEEE Transactions on Robotics and Automation*, 9(5):638–648, 1995.

[67] G. Liu, E. Austen, K. Booth, B. Fischer, M. Rempel, and J. Enns. Multiple object tracking is based on scene, not retinal, coordinates. *Journal of Experimental Psychology: Human Perception and Performance*, 31(2):235–247, 2005.

[68] G. R. Loftus and N. H. Mackworth. Cognitive determinants of fixation location during picture viewing. *Educational Psychology: Human Perception and Performance*, 4(4):565–572, 1978.

[69] A. Mack and I. Rock. *Inattentional Blindness*. MIT Press, Menlo Park, California, 2000.

[70] G. W. Maconcie and L.C. Loschky. Human performance with a gaze linked multi-resolutional display. *In Proceedings of the Advanced Displays and Interactive Displays First Annual Symposium 1997*, pages 25–34, 1997.

[71] K. Mania, A. Robinson, and K. Brandt. The effect of memory schemata on object recognition in virtual environments. *Presence, Teleoperators and Virtual Environments*, 14(5):606–615, 2005.

[72] K. Mania, T. Troscianko, R. Hawkes, and A. Chalmers. Fidelity metrics for virtual environment simulations based on human judgments of spatial

memory awareness states. *Presence, Teleoperators and Virtual Environments*, 12(3):296–310, 2003.

[73] Katerina Mania, Bernard D. Adelstein, Stephen R. Ellis, and Michael I. Hill. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 39–47, New York, NY, USA, 2004. ACM.

[74] Katerina Mania, Shahrul Badariah, Matthew Coxon, and Phil Watten. Cognitive transfer of spatial awareness states from immersive virtual environments to reality. *ACM Trans. Appl. Percept.*, 7(2):1–14, 2010.

[75] Katerina Mania, Dave Wooldridge, Matthew Coxon, and Andrew Robinson. The effect of visual and interaction fidelity on spatial cognition in immersive virtual environments. *IEEE Trans. Vis. Comput. Graph.*, 12(3):396–404, 2006.

[76] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. In *APGV '05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization*, pages 87–94, New York, NY, USA, 2005. ACM.

[77] G Marmitt and A. Duchowski. Modeling visual attention in vr: measuring the accuracy of predicted scanpaths. *Proceedings of Eurographics 2002, Short Presentations*, pages 217–226, 2002.

[78] Jeffrey McCandless, Stephen R. Ellis, and Bernard D. Adelstein. Localization of a time-delayed, monocular virtual object superimposed on a real environment. *Presence*, 9(1):15–24, 2000.

[79] B. H. McCormick, T. A. DeFanti, and M. D. Brown. Visualization in scientific computing. *Computer Graphics*, 21(6):1–14, 1987.

[80] A. McNamara. Evaluating image quality metrics vs. human evaluation. ACM SIGGRAPH 2000 Sketches Program, 2000.

[81] Ann McNamara. Visual perception in realistic image synthesis. *Comput. Graph. Forum*, 20(4):211–224, 2001.

[82] Ann McNamara. Visual perception in realistic image synthesis. *Comput. Graph. Forum*, 20(4):211–224, 2001.

[83] Michael Meehan, Sharif Razzaque, Mary C. Whitton, and Frederick P. Brooks, Jr. Effect of latency on presence in stressful virtual environments. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, page 141, Washington, DC, USA, 2003. IEEE Computer Society.

[84] G. W. Meyer and A. Liu. Color spatial acuity control of a screen subdivision image synthesis algorithm. *Human Vision, Visual Processing, and Digital Display*, 1666(3):387–399, 1992.

[85] Don P. Mitchell. Generating antialiased images at low sampling densities. *Computer Graphics*, 21(4):65–72, July 1987.

[86] Yann Morvan and Carol O'Sullivan. A perceptual approach to trimming unstructured lumigraphs. In *APGV '07: Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 61–68, New York, NY, USA, 2007. ACM.

[87] N. Mourkoussis, F. Rivera, T. Troscianko, T. Dixon, R. Hawkes, and K. Mania. Quantifying fidelity for virtual environment simulations employing memory schema assumptions. *ACM Transactions on Applied Perception*, 2010.

[88] K. Myszkowski. The visible differences predictor: Applications to global illumination problems. In G. Drettakis and N. Max, editors, *Rendering Techniques '98 (Proceedings of Eurographics Rendering Workshop '98)*, pages 233–236, New York, NY, 1998. Springer Wien.

[89] K. Myszkowski, A. B. Khodulev, and E. A. Kopylov. Validating global illumination algorithms and software. *In Visual Proceedings, Technical Sketch at ACM Siggraph 1997*, 1997.

[90] U. Neisser. The control of information pickup in selective looking. In A. D. Pick, editor, *Perception and its Development: A Tribute to Eleanor J. Gibson*, pages 201–219. Lawrence Erlbaum and Associates, Hilsdale, New Jersey, 1979.

[91] K. Nemire, R. H. Jacoby, and S. R. Ellis. Simulation fidelity of a virtual environment display. *Human Factors*, 36(1):1994, 1994.

[92] Carol O'Sullivan, Sarah Howlett, Yann Morvan, Rachel McDonnell, and Keith O'Conor. Perceptually Adaptive Graphics. In Christophe Schlick and Werner Purgathofer, editors, *STAR-Proceedings of Eurographics 2004*, number STAR-6 in State of the Art Reports, pages 141–164. INRIA and the Eurographics Association, 2004.

[93] A. J. Parkin, J. M. Gardiner, and R. Rosser. Functional aspects of recollective experience in face recognition. *Consciousness and Cognition*, 4(4):387–398, 1995.

[94] S. N. Pattanaik, J. A. Ferwerda, D. A. Greenberg, and M. D. Fairchild. A multiscale model of adaptation and spatial vision for realistic imaging. In *Computer Graphics (ACM SIGGRAPH '98 Proceedings)*, pages 287–298, 1998.

[95] Bobby Program Chair-Bodenheimer, Carol Program Chair-O'Sullivan, Katerina Conference Chair-Mania, and Bernhard Conference Chair-Riecke. Apgv '09: Proceedings of the 6th symposium on applied perception in graphics and visualization. *Chania, Crete, Greece*, 2009.

[96] Sarah Program Chair-Creem-Regehr and Karol Program Chair-Myszkowski. Apgv '08: Proceedings of the 5th symposium on applied perception in graphics and visualization. *Los Angeles, California*, 2008.

[97] P. Rademacher, J. Lengyel, E. Cutrell, and T. Whitted. Measuring the perception of visual realism in images, 2001.

[98] Ganesh Ramanarayanan, Kavita Bala, and James A. Ferwerda. Perception of complex aggregates. In *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*, pages 1–10, New York, NY, USA, 2008. ACM.

[99] Ganesh Ramanarayanan, James Ferwerda, Bruce Walter, and Kavita Bala. Visual equivalence: towards a new standard for image fidelity. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 76, New York, NY, USA, 2007. ACM.

[100] Mahesh Ramasubramanian, Sumanta N. Pattanaik, and Donald P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 73–82, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[101] J. E. Raymond, K. L. Shapiro, and K. M. Arnell. Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception & Performance*, 18(3):849–860, 1992.

[102] Matthew J. P. Regan, Gavin S. P. Miller, Steven M. Rubin, and Chris Kogelnik. A real-time low-latency hardware light-field renderer. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 287–290, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[103] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 267–276, New York, NY, USA, 2002. ACM Press.

[104] R. A. Rensink, J. K. O'Regan, and J. J. Clark. To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, 8:368–373, 1997.

[105] Ronald A. Rensink. Seeing, sensing, and scrutinizing. *Vision Research*, 40(10-12):1469–1487, 2000.

[106] James C. Rodger and Roger A. Browse. Choosing rendering parameters for effective communication of 3d shape. *IEEE Computer Graphics and Applications*, 20(2):20–28, 2000.

[107] Thomas B. Sheridan. Remote manipulative control with transmission delay. *IEEE Transactions on Human Factors in Electronics*, 4(1):25–29, 1963.

[108] Thomas B. Sheridan. Musings on telepresence and virtual presence. *Presence*, 1(1):120–125, 1992.

[109] D. J. Simons and R. A. Rensink. Change blindness: Past, present, and future. *Trends in Cognitive Science*, 9(1):16–20, 2005.

[110] Daniel J. Simons. Current approaches to change blindness. *Visual Cognition*, 7(1/2/3):1–15, 2000.

[111] P. H. Smith and J. Van Rosendale. Data and visualization corridors report on the 1998 CVD workshop series (sponsored by DOE and NSF). Technical Report CACR-164, Center for Advanced Computing Research, California Institute of Technology, 1998.

[112] Stephen Spencer, editor. *APGV '10: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, New York, NY, USA, 2010. ACM. 106103.

[113] William A. Stokes, James A. Ferwerda, Bruce J. Walter, and Donald P. Greenberg. Perceptual illumination compone high quality global illumination rendering. *ACM Transactions on Graphics*, 23(3):742–749, 2004.

[114] Veronica Sundstedt, Efstathios Stavrakis, Michael Wimmer, and Erik Reinhard. A psychophysical study of fixation behavior in a computer game. In *APGV '08: Proceedings of the 5th symposium on Applied perception in graphics and visualization*, pages 43–50, New York, NY, USA, 2008. ACM.

[115] J. J. Thomas and K. A. Cook. *Illuminating the Path: Research and Development Agenda for Visual Analytics*. IEEE Press, Piscataway, New Jersey, 2005.

[116] A. Treisman. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, 31:156–177, 1985.

[117] A. Treisman. Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception & Performance*, 17(3):652–676, 1991.

[118] A. Treisman and S. Gormican. Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95(1):15–48, 1988.

[119] E. Tulving. *Elements of Episodic Memory*. Oxford : Oxford Science Publications, 1992.

[120] Jack Tumblin and Holly E. Rushmeier. Tone Reproduction for Realistic Images. *IEEE Computer Graphics and Applications*, 13(6):42–48, November 1993.

[121] S. Uno and M. Slater. The sensitivity of presence to collision response. In *VRAIS '97: Proceedings of the 1997 Virtual Reality Annual International Symposium (VRAIS '97)*, page 95, Washington, DC, USA, 1997. IEEE Computer Society.

[122] C.J. van den Branden Lambrecht. *Perceptual models and architectures for video coding applications.* PhD thesis, Ecole Polytechnique Federal de Lausanne, 1996.

[123] David Waller, Earl Hunt, and David Knapp. The transfer of spatial knowledge in virtual environment training. *Presence: Teleoper. Virtual Environ.*, 7(2):129–143, 1998.

[124] B.J. Walter. *Density estimation techniques for global illumination.* Ph.D. thesis, Cornell University, 1998.

[125] Leonard Wanger. The effect of shadow quality on the perception of spatial relationships in computer generated imagery. In *I3D '92: Proceedings of the 1992 symposium on Interactive 3D graphics*, pages 39–42, New York, NY, USA, 1992. ACM.

[126] Colin Ware. *Information Visualization: Perception for Design, 2nd Edition.* Morgan Kaufmann Publishers, Inc., San Francisco, California, 2004.

[127] Benjamin Watson, Neff Walker, Larry F. Hodges, and Martin Reddy. An evaluation of level of detail degradation in head-mounted display peripheries. *Presence*, 6(6):630–637, 1997.

[128] Benjamin Watson, Neff Walker, Peter Woytiuk, and William Ribarsky. Maintaining usability during 3d placement despite delay. In *VR '03: Proceedings of the IEEE Virtual Reality 2003*, page 133, Washington, DC, USA, 2003. IEEE Computer Society.

[129] R. B. Welch, T. T. Blackmon, A. Liu, B. A. Mellers, and L. W. Stark. The effects of pictorial realism, delay of visual feedback and observer interactivity on the subjective sense of presence. *Presence, Teleoperators and Virtual Environments*, 5(3):263–273, 1996.

[130] Betsy Williams, Gayathri Narasimham, Claire Westerman, John Rieser, and Bobby Bodenheimer. Functional similarities in spatial representations between real and virtual environments. *ACM Trans. Appl. Percept.*, 4(2):12, 2007.

[131] Jeremy M. Wolfe. Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review*, 1(2):202–238, 1994.

[132] Jeremy M. Wolfe and Kyle R. Cave. Deploying visual attention: The Guided Search model. In T. Troscianko and A. Blake, editors, *AI and the Eye*, pages 79–103. John Wiley & Sons, Inc., Chichester, United Kingdom, 1989.

[133] Jeremy M. Wolfe, Kyle R. Cave, and Susan L. Franzel. Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance*, 15(3):419–433, 1989.

[134] P Zimmons. *The Influence of Lighting Quality on Presence and task Performance in Virtual Environment, Unpublished PhD thesis.* PhD thesis, University of North Carolina, ftp://ftp.cs.unc.edu/pub/publications/techreports/04-017.pdf, 2005.

[135] Alexandros Zotos, Katerina Mania, and Nicholaos Mourkoussis. A schema-based selective rendering framework. In *APGV '09: Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, pages 85–92, New York, NY, USA, 2009. ACM.