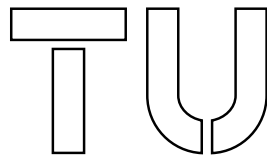


Radiosity Methods Driven by Human Perception

Jan Přikryl



TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

**Radiosity Methods
Driven by
Human Perception**

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften unter der Leitung von

Prof. Werner Purgathofer
E186

Institut für Computergraphik und Algorithmen

eingereicht an der Technischen Universität Wien
Fakultät für Technische Naturwissenschaften und Informatik

von

Ing. Jan Přikryl
Matr.-Nr. 9627154
Beringgasse 25/2/33, A-1170 Wien

Wien, im Mai 2001

I would like to thank Prof. Purgathofer for his comments and ideas regarding this thesis, and for initiating the productive environment at the Institute of Computer Graphics.

Thanks to all colleagues for their ideas and comments, and for the lively atmosphere at the institute. Help and feedback of Prof. Mateu Sbert from the University of Girona in Spain, Dr. Philippe Bekaert and Dr. Karol Myszkowski from Max-Planck Institute for Informatics in Saarbrücken in Germany were essential for completing this work.

Special thanks go to my parents for their kind support during all my studies, and my girlfriend Jana for having born with me for such a long time.

And, of course, Hail to the Master.

Abstract

Despite its popularity among researchers the radiosity method still suffers some disadvantage over other global illumination methods. Usual implementations of the radiosity method use criteria based on radiometric values to drive the computation — to decide about sufficient mesh quality or to estimate the error of the simulation process and to decide when the simulation can be safely terminated.

This is absolutely correct for the case of radiometric simulation, when the user is interested in actual values of radiometric quantities. On the other hand, the radiosity method is very often used just to generate pictures for the human observer and those pictures are not required to be the results of correct physical simulations, they just have to look the same.

The results of research on human visual performance and visual signal processing can be built into the image synthesis algorithm itself under some circumstances and guarantee that no effort will be spent on computing changes that are only marginally important for the human observer.

In the area of image processing, perceptual error metrics are used for image comparison and image coding that enable to better predict the differences between two images as opposed to the perceptually inappropriate and widely used mean-squared error metrics.

Tone reproduction operators known from image synthesis make it possible to map a bright scale of image luminances onto a narrow scale of CRT luminances in such a way that the perceived CRT image produces the same mental image as the original image.

Perceptually-driven radiosity algorithms exist, which use various methods to control the optimum density of the finite-element mesh defining the scene that is being rendered, to include only visible discontinuity lines into this mesh, and to predict the convergence of the method.

We will describe an hierarchical extension to the Monte Carlo radiosity that keeps the accuracy of the solution high only in the area immediately visible from the point of observation.

We will also present a comparison of different perceptual and radiometric termination criteria for a view-independent version of Monte Carlo radiosity.

Contents

1	Introduction	1
2	Human Visual System	3
2.1	Anatomy	4
2.1.1	The Cornea	4
2.1.2	The Lens	4
2.1.3	The Humours	5
2.1.4	The Iris	5
2.1.5	The Retina	5
2.1.6	The Optic Nerve	8
2.1.7	Visual Adaptation	8
2.2	Radiometry and Photometry	9
2.2.1	Radiance and Luminance	12
2.3	Psychophysics	13
2.3.1	Visual Acuity	14
2.3.2	Contrast Sensitivity	18
2.3.3	Masking	20
2.3.4	Visibility Thresholds	21
2.3.5	Brightness	23
2.4	Colorimetry	23
2.4.1	CIEXYZ	25
2.4.2	CIELAB and CIELUV	25
2.5	Measuring Colour Differences	27
2.5.1	The CIE Colour Difference Formula	28
2.5.2	The CIE94 Colour Difference Formula	28
2.6	Summary	29
3	Perceptual Approaches	30
3.1	Introduction	30
3.2	Tone-Mapping	31
3.2.1	Tumblin & Rushmeier’s Mapping	32

3.2.2	Ward's Mapping	34
3.2.3	Proposals of Ferwerda, Pattanaik <i>et al.</i>	36
3.2.4	Visibility Preserving Mapping of Ward-Larson <i>et al.</i>	37
3.2.5	Time-Dependent Adaptation	37
3.2.6	Tone-Mapping for Interactive Walkthroughs	38
3.2.7	Minimum Loss Methods	38
3.2.8	Low Curvature Image Simplifier	38
3.2.9	Retinex Theory	39
3.3	Perceptual Difference Metrics	39
3.3.1	Visible Differences Predictor	40
3.3.2	Sarnoff Visual Discrimination Model	40
3.3.3	Metric Based on the Normalisation Model of Early Vision	42
3.3.4	Comparison Experiments of Rushmeier <i>et al.</i>	42
3.3.5	Wavelet Comparison Metrics	44
3.3.6	Boker's Experiments	44
3.3.7	Stochastic Metric by Neumann <i>et al.</i>	45
3.3.8	Colour Image Fidelity Metrics	45
3.4	Summary	45
4	Monte Carlo Radiosity	47
4.1	Introduction	47
4.2	Monte Carlo Radiosity Methods	49
4.2.1	Stochastic Relaxation Radiosity	50
4.3	Hierarchical Refinement	53
4.3.1	Hierarchical Refinement in Monte Carlo Radiosity	53
4.4	Random Walk Methods	54
4.5	Summary	54
5	Importance-Driven Radiosity	55
5.1	Importance-Driven Radiosity	55
5.2	Importance-Driven Stochastic Jacobi Radiosity	56
5.3	Importance-Driven Hierarchical Stochastic Jacobi Radiosity	57
5.4	Results	59
5.5	Summary	61
6	Perceptually-Driven Radiosity	65
6.1	Principles	66
6.1.1	Assumptions of Perceptually-Driven Computation	66
6.1.2	Adaptation Luminance Estimation	67
6.1.3	Tone-Mapping in a Radiosity Algorithm	67
6.2	State of the Art	69

6.3	Object-Space methods	70
6.3.1	Perceptually-Driven Element Refinement	70
6.3.2	Perceptually-Driven Hierarchical Radiosity	70
6.3.3	Discontinuity Culling	75
6.4	Screen-Space Methods	76
6.4.1	Image-Space Refinement Criterion	76
6.4.2	Applications of VDP to Radiosity	81
6.5	Summary	83
7	Perceptually-Driven Termination	84
7.1	Convergence of the Stochastic Jacobi Method	84
7.2	Perceptually-Driven Termination Criterion	87
7.2.1	Confidence Intervals	87
7.2.2	1/k-Heuristics	92
7.3	Results	94
7.4	Summary	101
8	Conclusion	103
	Bibliography	105
A	Curriculum Vitae	116
B	List of Publications	119

Chapter 1

Introduction

Despite its popularity among researchers the radiosity method still suffers some disadvantage over other global illumination methods. Besides the fact that the original method allows only for solving the global illumination of environments consisting of purely diffuse surfaces, the method is rather computationally demanding. In the search for possible speed-up techniques one of the possibilities is to take also the characteristic features of the human visual system into account. Being aware of how the human visual perception works, one may compute the radiosity solution to lower accuracy in terms of physically based error metrics but being sure that the physically correct solution won't bring any improvements in the image that would be visible for the human observer.

Usual implementations of the radiosity method use criteria based on radiometric values to drive the computation — to decide about sufficient mesh quality or to estimate the error of the simulation process and to decide when the simulation can be safely terminated. This is absolutely correct for the case of radiometric simulation, when the user is interested in actual values of radiometric quantities. On the other hand, the radiosity method is very often used just to generate pictures for the human observer and those pictures are not required to be the results of correct physical simulations, they just have to look the same.

A simple motto for the topic of this thesis may sound: “If the human cannot see something there is no point in computing and displaying it.”

The advantage of producing “just” a visually satisfying solution that is indistinguishable from the physically correct one is that such a solution is believed to be cheaper to compute. Exploiting the fact that current display devices can not by large reproduce the real world range of luminances, or knowing what difference in contrast or colour can be noticed by a human under given viewing conditions, the time needed for a radiosity simulation can be decreased. However, also the contrary may be the case: In situations when the radiometric criteria would permit a visible computational error, perceptually-based methods naturally take longer to

deliver their visually acceptable results.

How shall we find what will be visible and what not? As we will see, this question has been asked by many researchers working in the field of psychology and physiology in the past years. The results of research on visual performance and visual signal processing can be built into the image synthesis algorithm itself under some circumstances and guarantee that no effort will be spent on computing changes that are only marginally important for the human observer.

It has to be noted that in many cases the relationships that provide us with simple models of vision were devised in laboratory conditions and do not necessarily apply in complex environments as those we want to display. One of the aims of visual psychology is still to better understand the way in which the visual information arriving at retina is perceived at all. The visual stimuli used in the research are mostly simple patterns as dots, crosses, lines or gratings of pure colour, and not complex photo-realistic images or photographs. Despite that, many successful applications of these results for the purpose of image coding or image synthesis have proved that even the knowledge gained from simple experiments may significantly help in improving performance of image synthesis algorithms.

The thesis is organised as follows: In the next chapter we provide a brief introduction to the human visual system and to different visual mechanisms that can be used to evaluate the visual appearance of a rendered image. In Chapter 3 we overview the current achievements in the field of perceptual approaches in image synthesis and image processing, concentrating on perceptually-based tone-mapping and image-comparison methods. The next Chapter 4 brings a brief introduction into Monte-Carlo radiosity. Following this introduction, we talk about hierarchical importance-driven radiosity in Chapter 5, which is the first main contribution of this thesis. Then we move towards perceptually-driven radiosity methods. In Chapter 6 we introduce different approaches that perceptually-driven radiosity methods take in order to achieve the necessary visual accuracy of the solution and present an overview of the existing perceptually-driven radiosity algorithms. Finally, Chapter 7 explains our further contributions — several perceptually-based termination criteria for Monte-Carlo radiosity. Chapter 8 sums up the thesis and gives some ideas for future development of perceptually-driven image synthesis algorithms.

Chapter 2

Human Visual System

The human visual system has a highly parallel, very specialized architecture. It has an information transfer rate of around 2 Gigabits per second. Visual stimuli arriving at our eyes are subject to heavy neural processing. During this processing, information as location, depth, texture, color, shape, movement, or pattern is being extracted. It is estimated that over 50% of the brain is involved with visual perception.

The physiology of the retina and connective arrangements of nerves is fairly well understood. As some knowledge about the anatomical and physiological structure of the eye is necessary to understand how the human visual system performs under different conditions, we will first briefly overview the anatomy of the eye.

It is natural that the measured performance of the human visual system differs among persons tested. It is virtually impossible for two different observers to reach a complete consensus over the description of given visual sensations. We describe these sensations by their colour, brightness, or contrast, but these quantities are very subjective. If one wants to model the human visual system behaviour, one needs a standardised type of human observer, that makes it possible to measure the light in such a way that the result would be as close as possible to the observations of the average person exposed to the same radiation.

Methods are needed that describe quantitatively the relationships between simulation and response for specified conditions of viewing. These methods form a basis for the science branch of *visual psychophysics* [4]. Very often, visual psychophysics experimentally determines different kinds of detection thresholds. The knowledge gained from monochromatic psychophysics experiments allows us to extend our models to take also colour into account and predict additional phenomena as, for example, the perception of colour differences.

2.1 Anatomy

Our visual perceptions are strongly influenced by the anatomical structure of the eye and some knowledge of the anatomy and physiology of the eye is necessary to understand how it performs as an optical device. The anatomical structure of the human eye is schematically given in Figure 2.1.

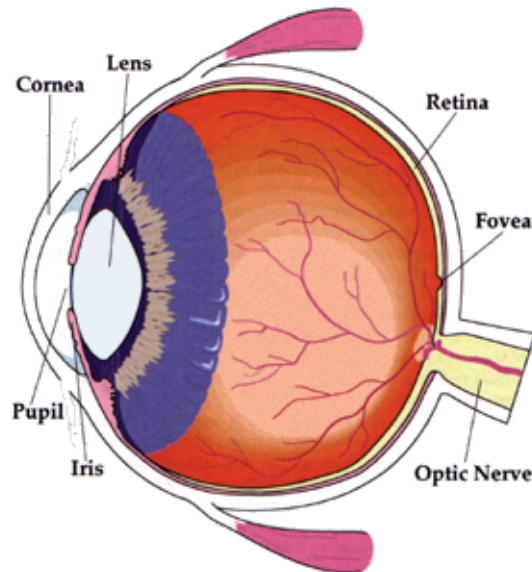


Figure 2.1: Schematic diagram of the human eye. After Fairchild [27].

2.1.1 The Cornea

The cornea is the transparent outer cover at the front of the eye. The interface air/cornea represents the largest change in refraction index found in the eye's optical system. This makes cornea one of the most significant image-forming elements of the eye. Refractive errors as myopia (nearsightedness), hyperopia (farsightedness) or astigmatism may be attributed to imperfections in the shape of cornea.

2.1.2 The Lens

The lens is a flexible, layered structure, the shape of which is controlled by the ciliary muscles. When we gaze at a distant object, the lens becomes flatter, decreasing its optical power in order to focus on the far away object. Gazing at a

nearby object, the lens becomes more “fat” increasing its optical power and thus allowing us to focus on the near object.

The lens varies in the index of refraction — the index is higher in the center of the lens and lower at the edges. This feature serves to reduce some of the aberrations that may normally be present in a simple optical system.

As we age, the internal structure of the lens changes, resulting in loss of flexibility. Finally at some point in time the lens has lost its flexibility and observers can no longer focus on near objects. What’s even more important, concurrently with the hardening of the lens an increase in its optical density occurs and the lens absorbs or scatters short wavelength energy. The effect is most noticeable when performing critical colour matching or comparing colour matching of older and younger observers: For purpleish objects, older people tend to report the colour being significantly redder than the colour reported by younger observers.

2.1.3 The Humours

The volume between the cornea and the lens is filled with a fluid that is essentially water (aqueous humour). The inner structure of the eye is filled with a fluid that has a higher viscosity (vitreous humour). Both humours are in the state of higher pressure (relative to the outer air pressure) to ensure that the flexible eyeball keeps its shape. The flexibility of an entire eyeball serves to increase the resistance to injury — it’s much easier to break something rigid than something that gives way under impact.

2.1.4 The Iris

The iris is the spincter muscle that controls pupil size defining thus the level of illumination on the retina. The pupil is the hollow in the middle of the iris through which the light passes.

Pupil size is largely determined by the level of illumination, but there exist also non-visual phenomena that may influence it. It is therefore difficult to accurately predict pupil size from the prevailing illumination. In practical situations the pupil diameter varies from about 3mm to about 7mm. This results in approximately five-fold change in the retinal illuminance.

2.1.5 The Retina

The optical image formed by the eye is projected onto a thin layer of cells located at the back of the eye — the retina. The retina incorporates photoreceptors (the visual system’s photosensitive cells) and is also responsible for initial visual signal processing and transmission to the visual cortex of our brain.

A schematic cross-section of the retina is illustrated in Figure 2.2. We can see that the visual signal preprocessing that takes place in the retina is already quite complex: multiple photoreceptors are being connected to multiple bipolar cells and the output of bipolar cells is passed to many different ganglion cells that form the optic nerve. In addition, the structure contains two horizontal layers — horizontal cells divert the output of photoreceptors to different bipolar cells and amacrine cells that interconnect the outputs of bipolar cells to different inputs of ganglion cells.

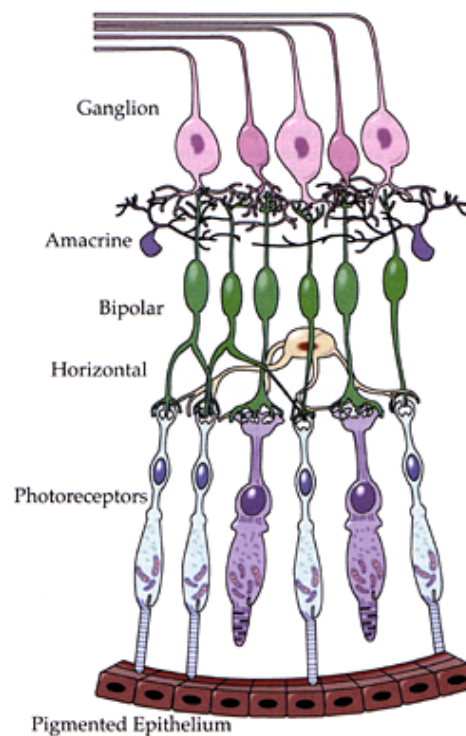


Figure 2.2: Schematic diagram of the human retina structure. After Fairchild [27].

This network of retinal cells serves as a sophisticated neural computer that suppresses that part of the visual information that is not important for visual processing and passes a highly compressed and frequency-encoded information further to the visual cortex. A schematic view of how we suppose retinal signal processing works is shown in Figure 2.3.

The area on the retina where we have the best spatial and colour vision is called the fovea. We tend to move our eyes and head in such a way that the objects of our immediate attention are being projected onto this retinal region. The fovea covers

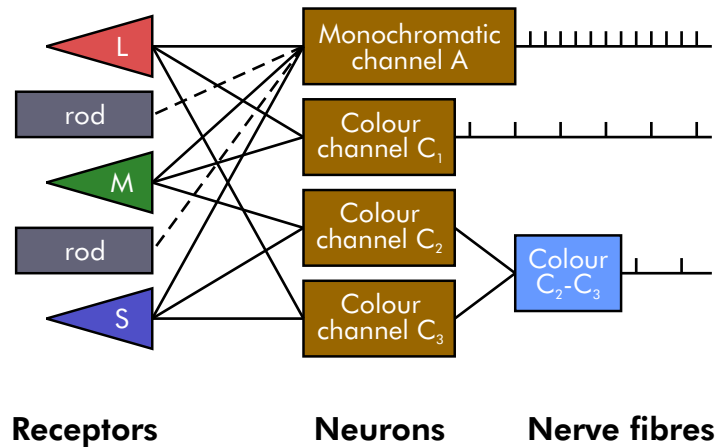


Figure 2.3: Early retinal signal processing and transformation into opponent colour space. The three neural pathways to the right transport frequency-encoded information. The amount of achromatic data (channel A) being transported is significantly higher than amount of the colour information (channels C₁ and C₂ – C₃). After Boynton [16] and Hunt [46].

an area that subtends approximately 2° of visual angle.

Photoreceptors

The retinal photoreceptors are of two basic types that differ by their visual functions: *rods* are responsible for vision at low luminance levels (less than 1cd/m^2) while *cones* serve vision at high luminance levels. Rods and cones owe their names to their actual shape — rods are long and slender, while cones are shorter and of conical shape.

Rods. There is only one type of rod photoreceptor with a peak spectral responsivity at about 510 nm. That implies that the vision at low luminance levels is monochromatic as only one type of photoreceptors is sensitive to light at these luminance levels.

Cones. There are three types of cone photoreceptors with their peak spectral responsivities placed throughout the visible spectrum. The three types of cones are usually referred to as L, M, and S cones (for “long-”, “middle-”, and “short-wavelength”). The three types of cones serve colour vision.

2.1.6 The Optic Nerve

The optic nerve is the nerve made of output of the ganglion cells of the retina. It serves as a communication channel transporting the preprocessed visual stimuli to the lateral region for further processing.

Since the optic nerve takes up the space that would be normally covered by photoreceptors there is a small area in each eye in which no visual stimulation can occur. This area is called the blind spot. The blind spot generally comes unnoticed, the reason being partially its location in an area that is on opposite sides of the visual field for both eyes and partially filling in the most probable visual information by our brain at the place where the optic nerve leaves the retina.

2.1.7 Visual Adaptation

The human visual system operates over 12 orders of luminance magnitude but single photoreceptors only have a dynamic range of approximately three orders of magnitude. Hence, the existence of the two receptor types alone is not sufficient to cope with the high luminance range that our visual system has to be able to process. The response of every receptor has a maximum amplitude and stimulation by higher luminances beyond this maximum value cannot further increase the response of the receptor. The effect is that with rising levels of retinal illumination the receptor's possible response range becomes more and more limited so that the eye slowly becomes blinded.

Additional mechanisms are therefore needed for the adaptation to this broad dynamic luminance range — the human visual system achieves the adaptation using multiplicative and subtractive control mechanisms [124].

Multiplicative Control Mechanisms

The multiplicative control mechanisms scale down the input to a neuron, preventing thus that the maximum response be reached too early. There are three multiplicative control mechanisms.

Pupil size. As it is well known, the pupil becomes smaller with rising illumination such that less light reaches the retina. But the pupil size can only cause luminance change in the range of one log unit.

Depletion of photopigments. At luminance levels where the cone vision prevails photopigments are bleached faster than they can be recovered. Therefore less photons can be absorbed and the receptor becomes less stimulated. In fact, for the cones pigment depletion stops further increases in light absorption already when they have reached half of their maximum response, such that they can operate up to the light damage limit.

Postreceptor gain changes. Also in cells which process the output of photoreceptors (the horizontal, bipolar, and ganglion cells) multiplicative mechanisms can be found.

Subtractive Control Mechanisms

Another possibility for the neuron gain control is to subtract a fixed percentage of the signal from the adapting field. One example are the horizontal cells — if a substantial amount of cones gets stimulated so is the horizontal cell. This leads in turn to an inhibition of all cones connected to the horizontal cell. This inhibition causes the input from the adapting field to be reduced by a certain amount.

2.2 Radiometry and Photometry

The eye, working together with the visual cortex of the brain, constitutes a very powerful sensory organ. We describe visual sensations by their colour, brightness, or contrast, but these sensations are very subjective. Generally, it is virtually impossible for two different observers to reach a complete consensus over the description of given visual sensations.

While *radiometry* studies the physical properties of visible radiation, the aim of *photometry* is to measure the properties of light in such a way that the result would be as close as possible to the observations of a standardised human observer exposed to the same radiation.

We distinguish three different types of human vision:

- *scotopic vision*, where the eye is adapted to very low illumination levels where rod vision prevails,
- *photopic vision*, where the eye is adapted to high illumination levels where cone vision prevails,
- *mesopic vision*, when the vision shifts from scotopic to photopic type or vice versa.

The human eye is not equally sensitive to all different spectral frequencies and its sensitivity also depends on the prevalent vision type (adaptation). The ratio of the perceived optical stimulus to the incident radiant power as a function of wavelength is given by *relative spectral sensitivity*. The functions describing the relative spectral sensitivity as a function of wavelength are called *luminous efficacy* (or *luminous efficiency*) functions and they have been standardised by the International Commission on Illumination (CIE, Commission Internationale de L'Eclairage) for photopic [$V(\lambda)$] and scotopic [$V'(\lambda)$] conditions. The values of

the relative spectral sensitivity given by $V(\lambda)$ and $V'(\lambda)$ are scaled so that the maximal relative spectral sensitivity is equal to one. Unfortunately, despite very thorough studies, the spectral sensitivity in the mesopic range is not yet completely understood and, according to current knowledge, it not possible to devise a single luminous efficacy function for mesopic vision that would work well in all mesopic conditions.

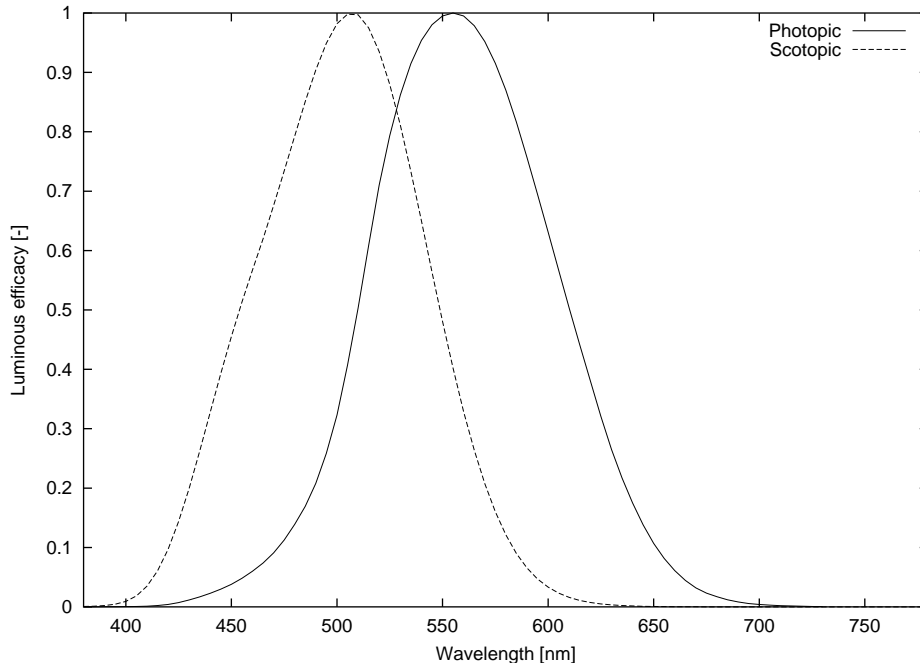


Figure 2.4: Photopic $V(\lambda)$ (1924) and scotopic $V'(\lambda)$ (1951) CIE luminous efficacy functions.

Figure 2.4 shows the original CIE photopic and scotopic luminous efficacy functions. As it is well known that the photopic function seriously underestimates vision in the blue part of the visible spectrum, several modifications of the original 1924 CIE curves exist that address this problem. These modified photopic efficacy functions are shown in Figure 2.5.

Given the luminous efficacy function, the photometric quantity X_v can be obtained from the corresponding spectra of a radiometric quantity X_e using

$$X_v = K_m \int X_e(\lambda) V(\lambda) d\lambda \quad (2.1)$$

for photopic vision and

$$X'_v = K'_m \int X_e(\lambda) V'(\lambda) d\lambda \quad (2.2)$$

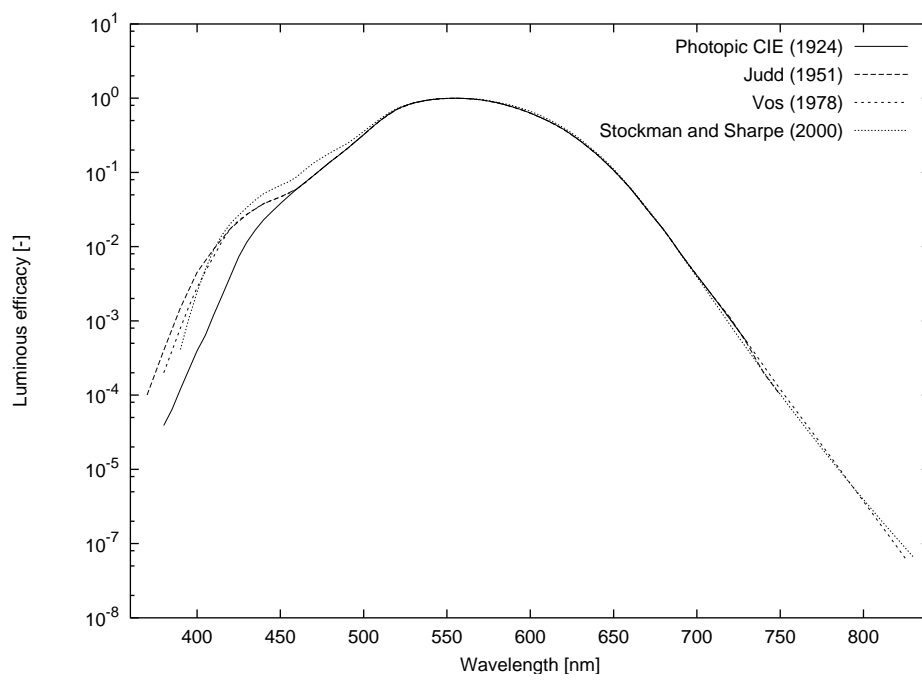


Figure 2.5: Original CIE photopic luminous efficacy function $V(\lambda)$, modifications by Judd (1951), Voss (1978), and a new version of photopic luminous efficacy curve derived from data measured by Stockman and Sharpe (2000) [112].

for scotopic vision, where K_m and K'_m are scaling constants that determine the maximum photopic and scotopic luminous efficacy.

In current practice, almost all the photometric quantities are still measured in the units of photopic vision even at low luminance levels except for special measurements for scientific purposes [23, p. 58].

In computer graphics tristimulus values representing the relation among three primary colour components are usually used to specify colour. Two equal tristimulus values may, however, correspond to different colour spectra and may therefore behave differently when converted using (2.1) or (2.2). This implies that it is impossible to reconstruct the original colour spectrum provided only the tristimulus information is given. As in this case the crucial information about the actual spectrum of the stimuli we are trying to convert is lost anyway, the usual approach (see Ward's remark in the Radiance Digest [127]) is to pretend that the tristimulus values correspond to an equal energy spectrum and to use a conversion coefficient based on this assumption.

2.2.1 Radiance and Luminance

Consider the radiation from a single-point light source emitting into a differential solid angle $d\Omega$ (see Figure 2.6). The *radiance* at a given point in a given direction is the radiant flux per unit solid angle per unit projected area that is perpendicular to $d\Omega$:

$$L_e = \frac{d^2\Phi}{d\Omega dA \cos\theta}, \quad (2.3)$$

where dA is the differential area containing the point in concern, $d\Omega$ is the differential solid angle containing the direction to the point, and θ is the angle between the normal to the elemental area and the direction to the light source.

The photopic *luminance* as a photometric equivalent of radiance, is defined as [128, pp. 256–259]

$$L_v = K_m \int L_e(\lambda)V(\lambda)d\lambda, \quad (2.4)$$

where $L_e(\lambda)$ is the component of a spectral radiance distribution L_e at wavelength λ , $V(\lambda)$ is the value of the CIE photopic luminous efficacy function at this frequency and L_v is the corresponding luminance of the spectral radiance stimulus. $K_m = 683 \text{ lm/W}$ is a scaling constant that determines the maximum photopic luminous efficacy.

Due to the aforementioned fact that it is impossible to reconstruct the corresponding spectral radiance distribution L_e from a tristimulus radiance $L_{e,\text{tri}}$, we will compute the luminance value from the tristimulus radiance as

$$L_v = \sigma L_{e,\text{tri}}. \quad (2.5)$$

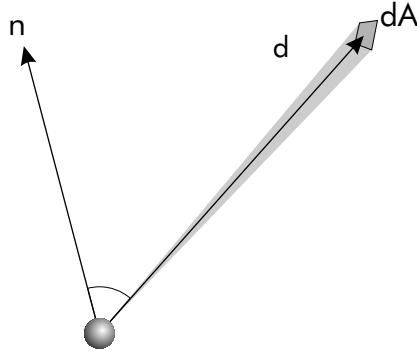


Figure 2.6: Geometry of radiance definition.

The actual value of the scaling coefficient σ depends on the spectral distribution of the radiance of the light sources and on surface characteristics. Typical approximation values that are used in Ward's rendering package RADIANCE [125] are given in Table 2.1. All the images in this thesis have been computed with $\sigma = 179 \text{ lm/W}$.

light source type	σ [lm/W]
artificial light	179
daylight (D65)	203
incandescent (A)	160
sun (B)	208

Table 2.1: Luminous efficacy factors σ used in RADIANCE.

2.3 Psychophysics

Results of different experiments in *visual psychophysics* help us in developing methods that describe quantitatively the relationships between stimulation and response for given viewing conditions. Psychophysics itself is a science that tries to build a descriptive bridge between a physical stimulus and the corresponding mental response.

The classical form of visual psychophysics uses the human visual system only as a null instrument assessing whether two stimuli matched or not.

2.3.1 Visual Acuity

Visual acuity is the capability of the human vision to discriminate among the fine details. It describes how well we are able to detect, resolve, or recognize small stimuli or the difference among different stimuli. The most common test of acuity is performed by reading the *Snellen chart* used in the opthalmological practice.

Detection of a target is a less stringent criterion generally than resolution of the elements of the target. That is the reason why the acuity thresholds for resolution tend to be higher than those for detection. As the illumination of our environment decreases, our ability to distinguish fine detail decreases as well. This fact has been demonstrated by Shlaer in the late 1930s [99].

Figure 2.7 shows the results of another visual acuity experiment conducted by Thomas [115]. In this case, the relation between the true retinal illuminance and visual acuity has been measured. The functional fit to this data has the form

$$\omega = \frac{2.6 E^{0.55}}{E^{0.55} + 65^{0.55}}, \quad (2.6)$$

where ω is the visual acuity expressed in cycles of stimuli change per visual degree and E denotes the retinal illuminance in trolands.

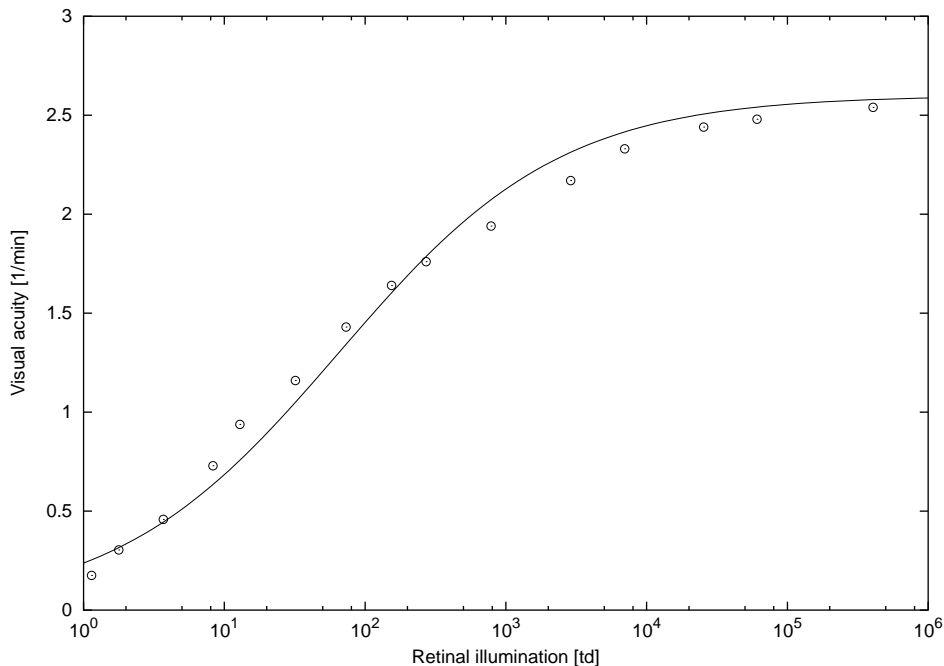


Figure 2.7: Visual acuity data measured by Thomas [115] and the corresponding functional fit.

One *troland* [td] of retinal illuminance is produced by an extended source of 1 cd/m^2 seen through a pupil of 1 mm^2 . Knowing the retinal illuminance, the corresponding luminance L in candelas per meter squared can be computed as

$$L[\text{cd/m}^2] = \frac{E[\text{td}]}{A[\text{mm}^2]}, \quad (2.7)$$

where A is the pupil area in millimeters squared. Neglecting the diversities between particular observers, the pupil diameter d can be approximated as a function of surroundings luminance L . This relationship has been measured by many scientists under different conditions. We will use the formula recommended by Wyszecki and Stiles [128, p. 106],

$$\log d = 0.8558 - 0.000401 (\log L + 8.6)^3. \quad (2.8)$$

The pupil diameter in (2.8) is given in millimeters. Expressing

$$\log A = \log \frac{\pi}{4} + 2 \log d$$

we can combine (2.7) and (2.8) into

$$\log E = 1.607 + \log L - 0.000802 (\log L + 8.6)^3. \quad (2.9)$$

The equation (2.9) can be inverted and in this way a function describing luminance of surroundings L as a function of the retinal illuminance E can be obtained. However, the inverse function has a very complicated form and our analysis has shown that it is unlikely to be evaluated in a numerically stable manner.

Therefore we propose to either to solve (2.9) numerically for L using a standard zero finder or to use the following approximation, which we have found as a functional fit to selected numerical solutions of (2.9) (see also Figure 2.8):

$$\log L = 0.004936 (\log E)^3 + 0.03072 (\log E)^2 + 1.131 \log E - 1.304. \quad (2.10)$$

Table 2.2 shows the visual acuity data from Shaler's paper [99], recomputed from the original retinal illuminance units (trolands) into the standard luminance units of candela per meter squared.

The data from Table 2.2 make it possible for image synthesis algorithms to simulate the decrease in acuity with decrease of illumination in the scene. A model based on this data was used by Ferwerda *et al.* [30] in their tone-mapping algorithm. A similar type of acuity model has been also employed by Ramasubramanian *et al.* [85] to drive the sampling accuracy for a stochastic ray-tracer.

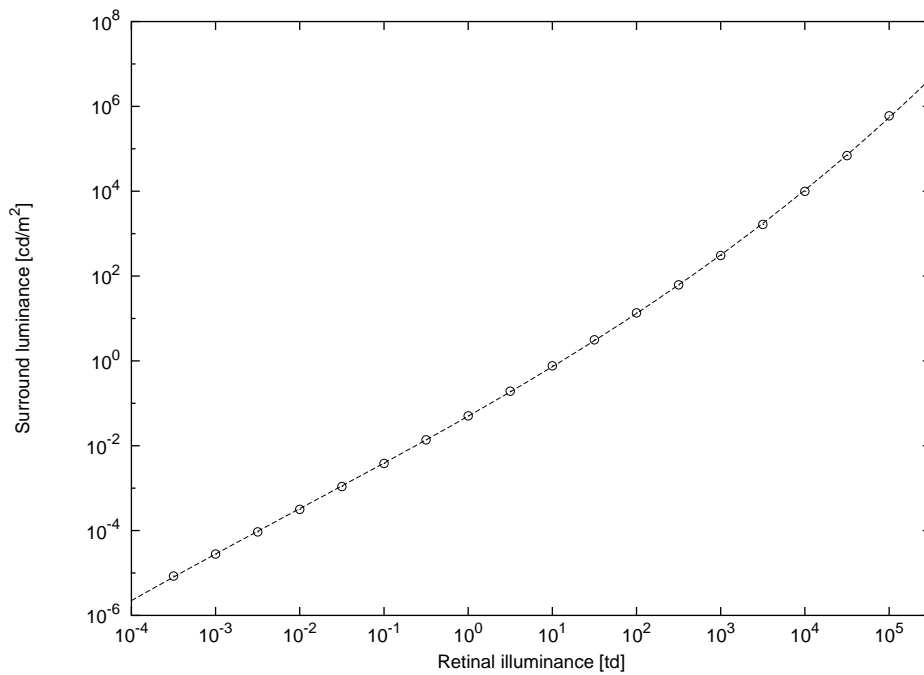


Figure 2.8: Relation between retinal illuminance and surround luminance. Plotted points were obtained by numerically solving (2.9) for L . The functional fit follows (2.10).

L_{wa} (cd/m ²)	ω (cpd)
3.16×10^{-4}	2.1
1.00×10^{-3}	2.9
3.16×10^{-3}	4.1
1.00×10^{-2}	5.5
3.16×10^{-2}	9.0
1.00×10^{-1}	16.3
3.16×10^{-1}	23.8
1.00	32.5
3.16	38.5
1.00×10^1	43.1
3.16×10^1	46.0
1.00×10^2	48.0
3.16×10^2	48.8
1.00×10^3	50.0

Table 2.2: Tabulated visual acuity data measured by Shlaer [99]. Left column: the values of adaptation luminance L_{wa} , right column: maximum recognizable stimulus frequency ω in cycles per visual degree.

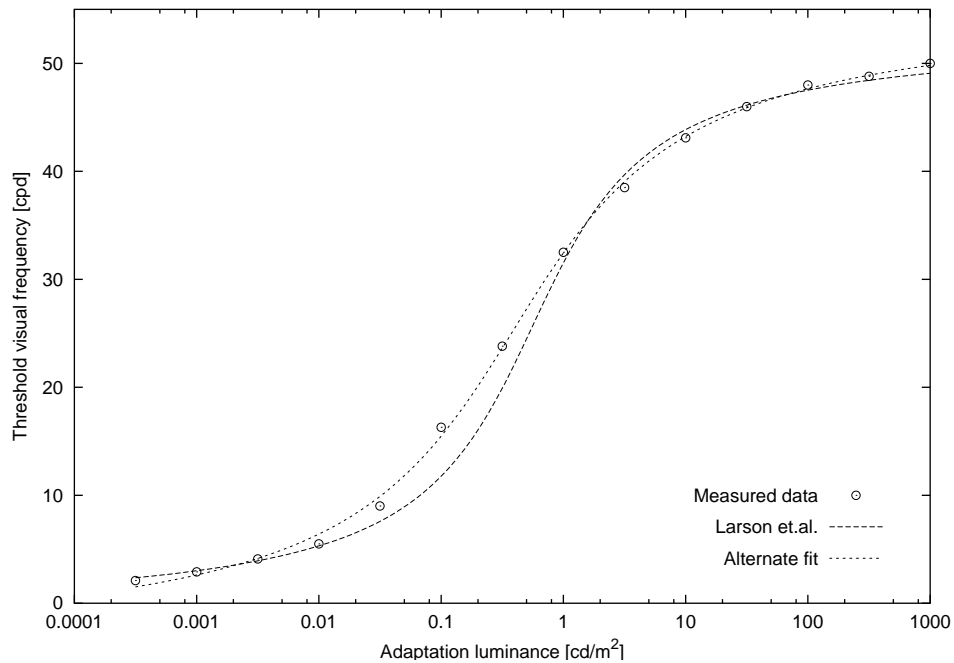


Figure 2.9: Shlaer visual acuity data [99] and corresponding functional fits.

Durand and Dorsey [26] use a piecewise linear approximation of the Shlaer data, while Larson *et al.* [57] approximate them with

$$\omega = 17.25 \arctan(1.4 \cdot \log(L_{wa}) + 0.35) + 25.72 . \quad (2.11)$$

The result can be seen in Figure 2.9. We have found that a slight discrepancy with the measured data around adaptation level of 0.1 cd/m^2 may be improved using an alternate functional fit of the form

$$\omega = 19.2 \arctan(0.953 \cdot \log(L_{wa}) + 0.384) + 25.4 . \quad (2.12)$$

2.3.2 Contrast Sensitivity

It is well known that the human visual system is not able to resolve arbitrary details in the perceived image. The visibility limit for low contrast patterns can be predicted using a measure called *contrast sensitivity* [22, 88], which is defined as inverse of the contrast threshold for a given spatial frequency. This measure tells us how faded or washed out image details can be before they become indistinguishable from a uniform field. It is a function of the size of image features, or the spatial frequency of visual stimuli produced by these features. The *contrast sensitivity function* (CSF) plots the contrast sensitivity for all spatial frequencies and given viewing conditions. Two popular CSFs are plotted in Fig. 2.10.

The contrast sensitivity values were determined as follows: At the given contrast level, a sine-waved stripe pattern with the given spatial frequency was presented to the human observer. When the stripes were very thin, i.e. the spatial frequency of the stimuli was very high (above 60 cycles per visual degree), the test subject was not able to distinguish between particular stripes. As the stripe frequency dropped, the threshold contrast above which the stripes were distinguishable dropped as well. However, after achieving a peak value of approximately 4–8 cycles per visual degree, depending on the adaptation level, the threshold contrast started to grow again. An example figure of *Campbell-Robson Contrast Sensitivity Chart* [74] illustrating this phenomenon is depicted in Fig. 2.11.

Mannos and Sakrison [62] studied the effects of contrast sensitivity on image coding. In their paper a proposal of the functional fit to the experimental data is presented. Their CSF has the form of

$$\tilde{A}(f_r) = (0.05 + 0.2964f_r) \cdot \exp[-(0.114f_r)^{1.1}] \quad (2.13)$$

and predicts the perceptual sensitivity to a visual stimulus of spatial frequency f_r . This CSF form has been proven to be applicable for image comparison [88, 34].

Also popular is the adaptation level dependent CSF proposed by Daly [22], which forms a part of his perceptually-driven image comparison method. This

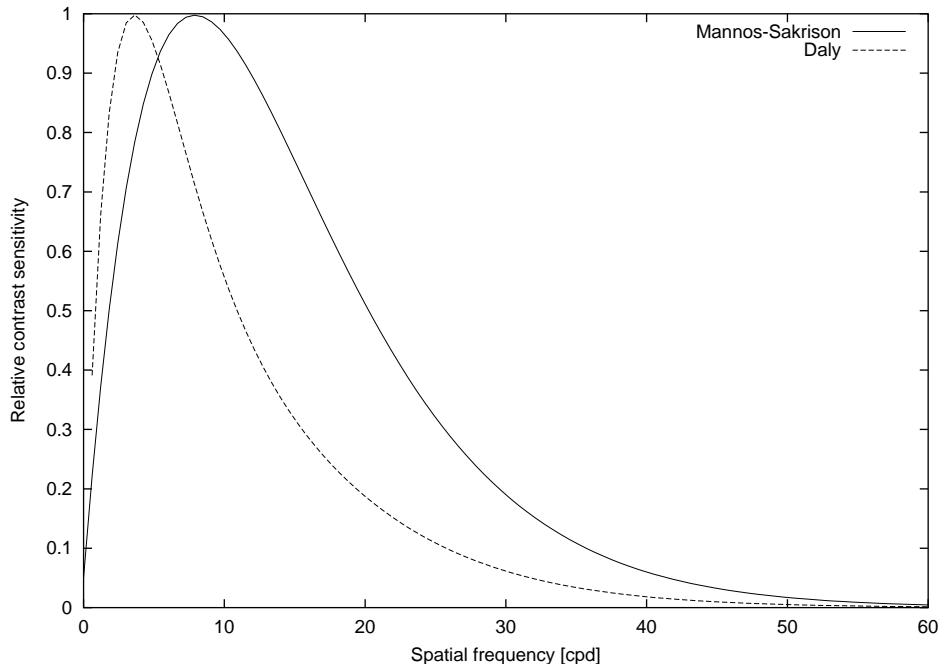


Figure 2.10: Normalised Mannos and Sakrison’s contrast sensitivity function [62] and normalised Daly’s contrast sensitivity function [22] computed for adaptation level 50 cd/m^2 .

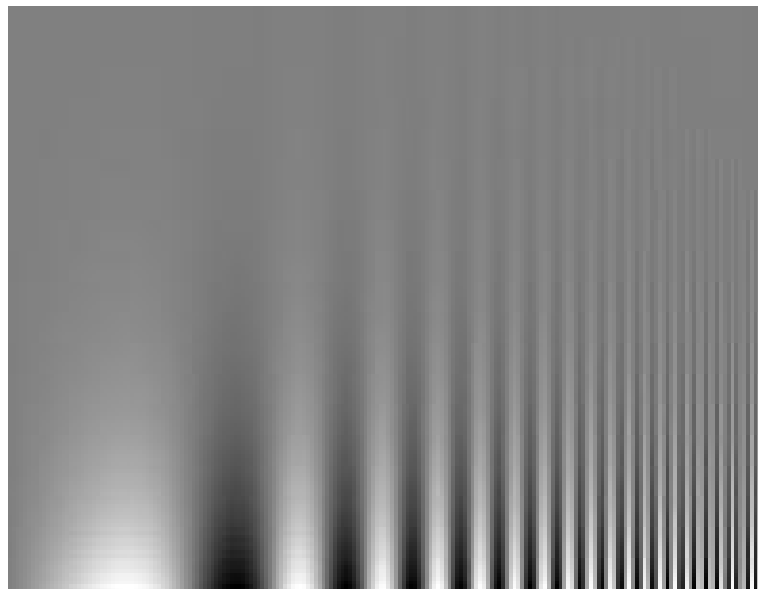


Figure 2.11: Campbell-Robson contrast sensitivity chart.

function models the sensitivity S as a function of radial spatial frequency ρ in cycles per degree, orientation θ in degrees, adaptation luminance L_a in cd/m^2 , image size i^2 in visual degrees, lens accommodation due to distance d in meters, and eccentricity e in degrees, as

$$S(\rho, \theta, l, i^2, d, e) = P \cdot \min [S_1(\rho_{ae\theta}, l, i^2), S_1(\rho, l, i^2)]. \quad (2.14)$$

where

$$\begin{aligned} \rho_{ae\theta} &= \frac{\rho}{r_a r_e r_\theta}, \\ r_a &= 0.856d^{0.14}, \\ r_e &= \frac{1}{1 + 0.24e}, \\ r_\theta &= 0.11 \cos 4\theta + 0.89. \end{aligned}$$

The auxiliary sensitivity function S_1 is given as

$$\begin{aligned} S_1(\rho, l, i^2) &= \left[(3.23(\rho^2 i^2)^{-0.3})^5 + 1 \right]^{0.2} \times \\ &\times a_l \exp(-b_l) \sqrt{1 + 0.06 \exp(b_l \varepsilon \rho)}, \end{aligned} \quad (2.15)$$

where

$$\begin{aligned} a_l &= 0.801 \varepsilon \rho \left(\frac{1.7}{l} \right)^{-0.2}, \\ b_l &= 0.3 \varepsilon \rho \left(\frac{101}{l} \right)^{0.15}. \end{aligned}$$

2.3.3 Masking

Masking is a robust perceptual phenomenon that has been extensively studied for more than one hundred years. It refers to the situation when presence of one pattern in the observed picture prevents the observer to see another pattern that is also present in the picture. The effect of masking may be a decrease in brightness, errors in recognition, or a failure to detect. The strength of the masking effect depends on the frequency spectrum, contrast and orientation of both patterns. An example of masking effect may be seen in Figure 2.12.

We note here that the effect of one stimulus on the detectability of another one does not necessarily have to result in a decrease in detectability. Under certain circumstances a low contrast masker may increase the detectability of a signal [58].

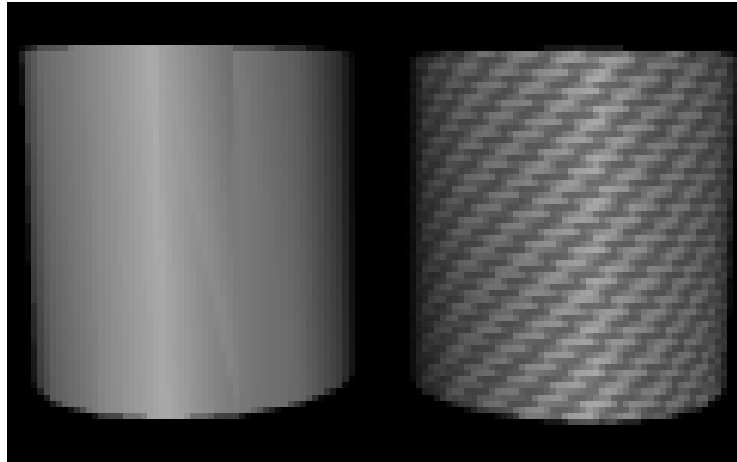


Figure 2.12: Example of visual masking. The imperfection in the cylinder illumination is masked out by the applied texture.

2.3.4 Visibility Thresholds

The term *visibility* can be used to describe the visibility of an object bounded by a contrast border or the visibility of a border itself. For a test target displayed with luminance L_D on a background of luminance L , the contrast C is defined as

$$C = \left| \frac{L_D - L}{L} \right|. \quad (2.16)$$

The contrast can be reduced to the threshold of visibility by reducing the difference between L_D and L while L is being kept constant until the target disappears. Such a contrast is called *contrast threshold* and depends on L . The expression of the threshold contrast as a function of luminance has been approved by CIE to take the form

$$\bar{C} = 0.05936 \left[\left(\frac{1.639}{L} \right)^{0.4} \right]^{2.5}. \quad (2.17)$$

Figure 2.13 shows the standard contrast threshold curve endorsed by CIE together with the original data measured by Blackwell and Blackwell [12].

Similarly to the contrast threshold experiments, the threshold luminance difference L_D can be measured. Plotting the detection threshold luminance against the corresponding background luminance gives us the so-called *threshold-versus-intensity* (often denoted *t.v.i*) function. Figure 2.14 shows the respective *t.v.i* functions for cones and rods.

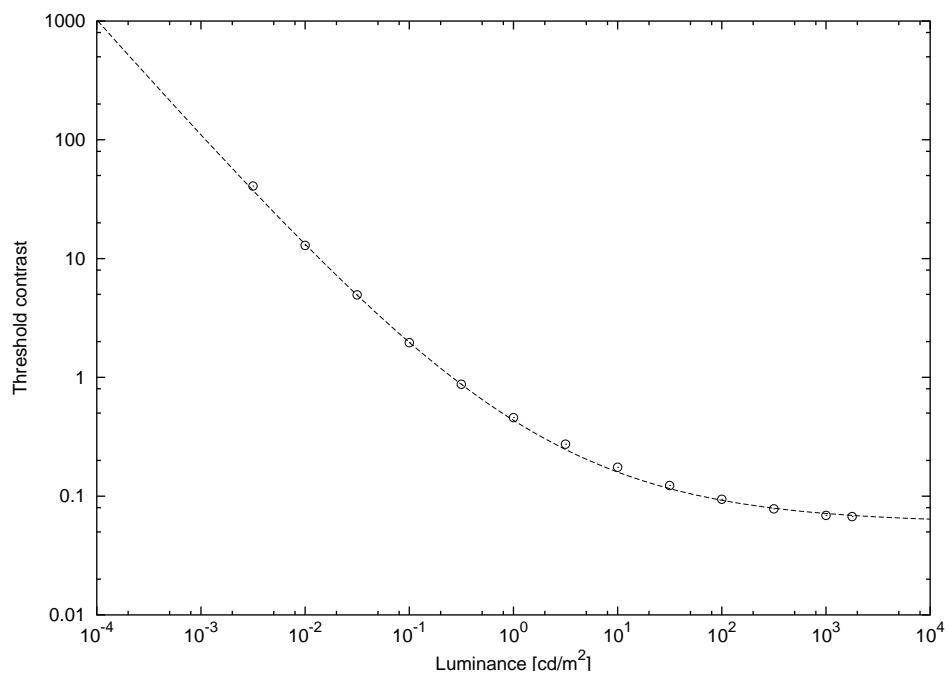


Figure 2.13: Effect of luminance on the contrast threshold. After Blackwell and Blackwell [12] and Fry [32].

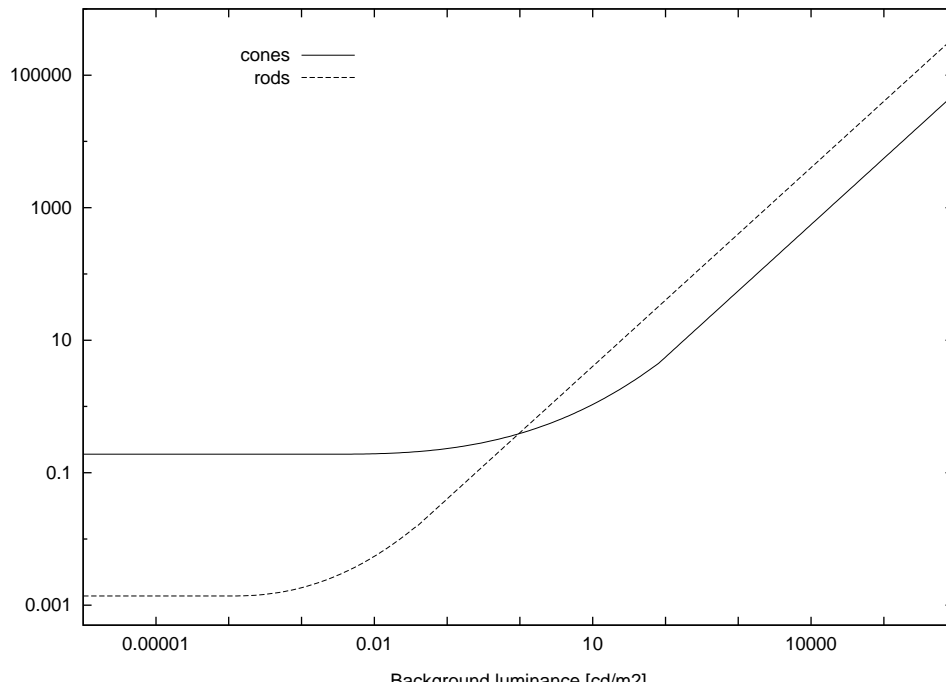


Figure 2.14: Threshold-versus-intensity functions. After Ferwerda *et al.* [30].

As we can see, over a wide middle range of background luminance the function is linear, corresponding to Weber's law [128, 27] $\Delta L/L = \text{const}$. Rods and cones show very similar patterns of response. The visual system threshold sensitivity can be described by the envelope of the rod and cone t.v.i curves.

2.3.5 Brightness

The principal physiological sensation of brightness is linked to the luminance of the source we are looking at — we can say that the human eye compares luminances [23, p. 46].

2.4 Colorimetry

Colorimetry studies the properties of human colour perception — we can also say that colorimetry measures colour. Our visual colour sensations are measured using colour matching experiments, where a human observer tries to match targets of different colour.

The properties of human colour matching are defined by the spectral respon-

sivities of our three cone types. The reason is that once the light energy is absorbed by the cones, the spectral substance of the original signal is lost. This implies that if the output signals from the three cone types are equal for two stimuli, they must match in colour when seen in the same conditions.

If we know the spectral power distributions of the two stimuli, $\Phi_1(\lambda)$ and $\Phi_2(\lambda)$, and each of cone responsivities, $L(\lambda)$, $M(\lambda)$, and $S(\lambda)$, the condition of a visual match can be mathematically described as simultaneous fulfillment of the following three equations:

$$\int \Phi_1(\lambda)L(\lambda)d\lambda = \int \Phi_2(\lambda)L(\lambda)d\lambda, \quad (2.18)$$

$$\int \Phi_1(\lambda)M(\lambda)d\lambda = \int \Phi_2(\lambda)M(\lambda)d\lambda, \quad (2.19)$$

$$\int \Phi_1(\lambda)S(\lambda)d\lambda = \int \Phi_2(\lambda)S(\lambda)d\lambda. \quad (2.20)$$

This condition suggests that there exist stimuli of different spectral distributions that are perceived as being of equal colours. Such stimuli are called *metamers*.

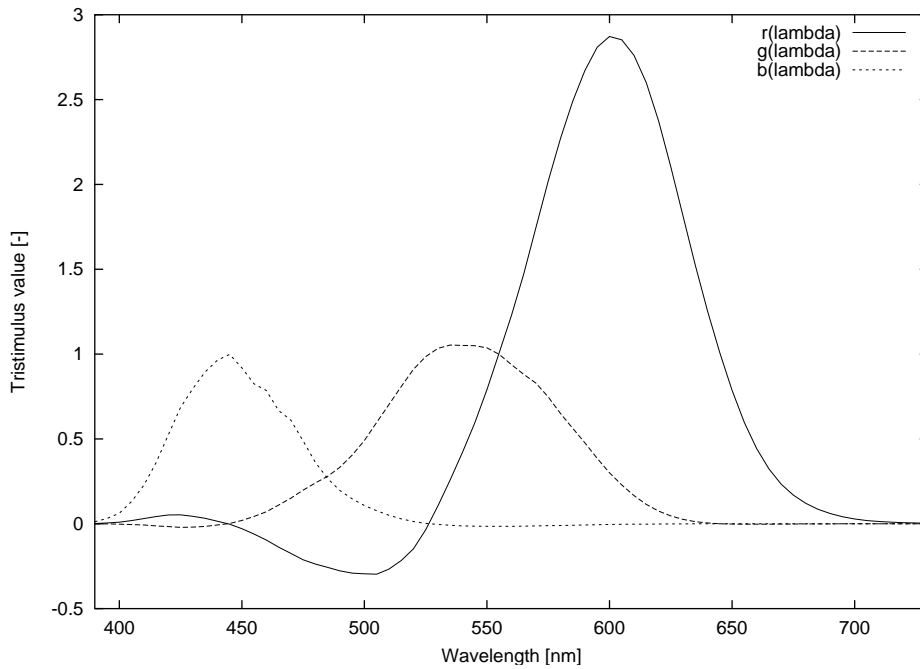


Figure 2.15: CIE RGB colour matching functions $\bar{r}(\lambda)$, $\bar{g}(\lambda)$, and $\bar{b}(\lambda)$ based on the later experiments of Stiles and Burch [128, Table I(5.5.3)].

After having introduced the luminous efficacy function $V(\lambda)$, CIE turned its attention to the development of a system that would specify when two metameric

stimuli match in colour for an average observer. The first attempt resulted in an RGB system with monochromatic primaries at 435.8 nm, 546.1 nm, and 700.0 nm (see Fig.2.15). In an effort to eliminate the negative values of the RGB colour matching functions and to incorporate the photometric luminous efficacy curve into the proposed colorimetry system, an imaginary set of primaries has been constructed. These primaries form the CIEXYZ colour system.

2.4.1 CIEXYZ

The CIEXYZ colour system consists of two imaginary primaries, X and Z, that do not produce any luminance response and carry only the colour information, leaving the luminance response completely to the third primary, Y. The CIEXYZ tristimulus is computed from the spectral stimuli $\Phi(\lambda)$ as

$$X = k \int \Phi(\lambda) \bar{x}(\lambda) d\lambda, \quad (2.21)$$

$$Y = k \int \Phi(\lambda) \bar{y}(\lambda) d\lambda, \quad (2.22)$$

$$Z = k \int \Phi(\lambda) \bar{z}(\lambda) d\lambda. \quad (2.23)$$

The value of the normalization constant, $k = K_m = 683 \text{ lm/W}$, is chosen so that the whole system of colorimetry is compatible with the CIE photometry system.

The colour matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ are plotted in Figure 2.16. We remark that $\bar{y}(\lambda) = V(\lambda)$ by definition.

2.4.2 CIELAB and CIELUV

The purpose of the CIELAB and CIELUV colour spaces was a creation of colour spaces that would provide uniform practices for the measurement of colour differences, which is a task that the standard CIEXYZ colour system is unsuitable for. The nonuniformity factor of CIEXYZ is approximately 80:1, which means that the same colour differences in CIEXYZ may be sometimes perceived as being 80 times different, dependent on a colour pair being compared. The perceptual nonuniformity of CIELAB and CIELUV is only about 6:1.

Calculating the CIELAB coordinates, we must begin with two CIEXYZ tristimulus values: the value of the stimulus, (X, Y, Z) , and the value of the reference white (X_n, Y_n, Z_n) . The reference white is used to normalize the stimulus CIEXYZ coordinate to those of the white colour.

The adapted signals are then subjected to a compressive nonlinearity designed to model the compressive response typically found between physical energy measurements and perceptual responses. In CIELAB computation, this compressive

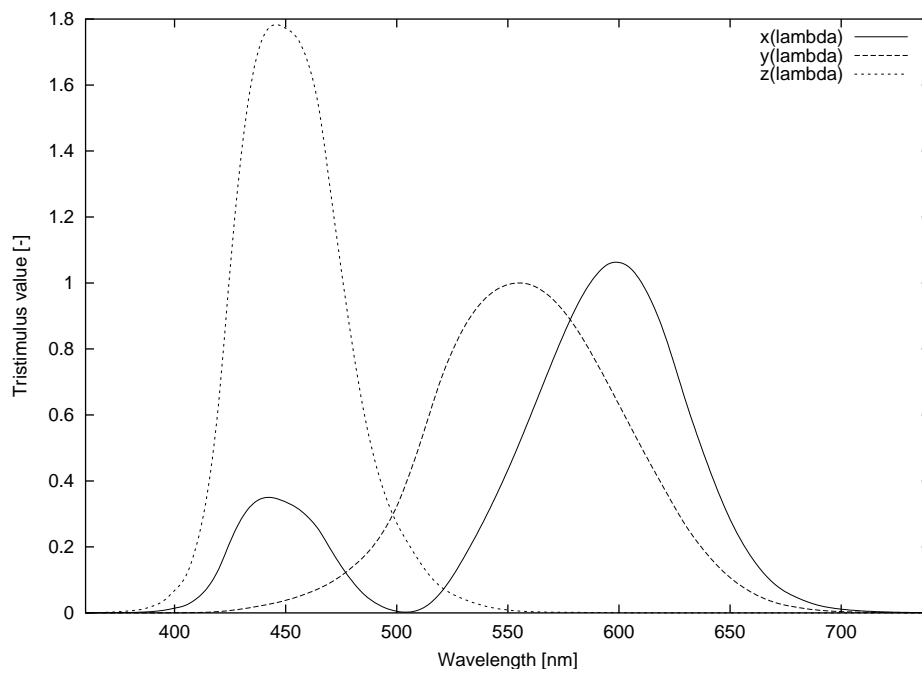


Figure 2.16: Colour matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, and $\bar{z}(\lambda)$ of the CIE standard colorimetry observer. After Wyszecki and Stiles [128].

nonlinearity is represented by a cube root function except for very low stimuli values, where an alternative formula is used that overcomes the design flaw of the original CIELAB equations.

The three resulting “compressed” signals are then combined into three response dimensions corresponding to the light-dark, red-green, and yellow-blue responses of the opponent-colours theory of colour vision.

The full transformation of a CIEXYZ stimulus into CIELAB colour space is given as

$$L^* = 116 f\left(\frac{Y}{Y_n}\right) - 16, \quad (2.24)$$

$$a^* = 500 \left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right], \quad (2.25)$$

$$b^* = 200 \left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right], \quad (2.26)$$

where the scaling function $f(x)$ is defined as

$$f(x) = \begin{cases} \sqrt[3]{x} & \text{if } x > 0.008856 \\ 7.787x + 0.1379 & \text{otherwise} \end{cases}$$

The CIELUV colour space uses the same lightness coordinate L^* as CIELAB, but the colour coordinates are based on the CIE chromaticity diagram. The CIELUV colour coordinates are

$$u^* = 13 L^* [u' - u'_n], \quad (2.27)$$

$$v^* = 13 L^* [v' - v'_n], \quad (2.28)$$

where u' and v' are the CIE chromaticity coordinates of the stimulus and u'_n and v'_n are the CIE chromaticity coordinates of the reference white colour, computed as

$$u' = \frac{4X}{X + 15Y + 3Z},$$

$$v' = \frac{9Y}{X + 15Y + 3Z}.$$

2.5 Measuring Colour Differences

The CIELAB and CIELUV colour spaces were both recommended in 1976 by CIE as an interim solution to the problem of colour-difference specification of reflecting samples. Since then, CIELAB has become almost universally used for colour

difference measurement [27]. Given the fact that CIELUV performs extremely poor when predicting certain colour differences and perceptually corresponding colours [1, 86], there seems to be no reason of using CIELUV over CIELAB [27, Sec. 3.8 and 10.5] nowadays.

2.5.1 The CIE Colour Difference Formula

Colour differences are measured in the CIELAB colour space as the Euclidean distance between the coordinates of the two stimuli. The colour difference formula adopted in 1976 by CIE defines the formula for approximately perceptually uniform colour difference as

$$\Delta E_{ab}^* = \sqrt{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}}. \quad (2.29)$$

The same applies also for the CIELUV colour system:

$$\Delta E_{uv}^* = \sqrt{\Delta L^{*2} + \Delta u^{*2} + \Delta v^{*2}}. \quad (2.30)$$

2.5.2 The CIE94 Colour Difference Formula

To improve the nonuniformity of colour difference measurements provided by (2.29), CIE has evaluated different modifications to the original CIE colour difference formula and recommended a new CIE94 colour difference equation, which gives the colour difference of two stimuli in CIELAB colour space as

$$\begin{aligned} \Delta E_{94}^* &= \sqrt{\left(\frac{\Delta L^*}{k_L S_L}\right)^2 + \left(\frac{\Delta C_{ab}^*}{k_C S_C}\right)^2 + \left(\frac{\Delta H^*}{k_H S_H}\right)^2}, \quad (2.31) \\ S_L &= 1, \\ S_C &= 1 + 0.045 C_{ab}^*, \\ S_H &= 1 + 0.015 C_{ab}^*. \end{aligned}$$

The parametric factors k_L , k_C , and k_H are used to adjust the relative weighting of the formula components for various viewing conditions. C_{ab}^* in (2.31) is called chroma and can be computed as

$$C_{ab}^* = \sqrt{a^{*2} + b^{*2}}. \quad (2.32)$$

H_{ab}^* is hue correlate which can be computed as

$$H_{ab}^* = \sqrt{\Delta E_{ab}^{*2} - \Delta L^{*2} - \Delta C_{ab}^{*2}}. \quad (2.33)$$

Together with the new colour difference formula definition, CIE established a set of reference conditions for the use of the CIE94 colour difference formula. These conditions are listed in Table 2.3.

Illumination	CIE illuminant D65
Illuminance	1000 lux
Observer	Normal colour vision
Background	Uniform, achromatic, $L^* = 50$
Viewing mode	Object
Sample size	Greater than 4° visual angle
Sample separation	Direct edge contact
Sample colour difference magnitude	0 to 5 CIELAB units
Sample structure	No visually apparent pattern or nonuniformity

Table 2.3: Reference conditions for the use of CIE94 colour difference formula.

2.6 Summary

The aim of this chapter was to introduce the reader to the fundamentals of the human visual system and to present the concepts that will be used later when talking about perceptually-based methods used in computer graphics.

We started with a brief explanation of the human eye anatomy and talked about the process by which the eye adapts to the broad range of incoming luminances. We noted that the perception of light and colour is a very subjective process and cannot be measured in physical quantities. The necessary quantification of the human visual sensations is provided by an alternate system of photometric quantities that describe the effects of light on the human observer.

We talked about several visual phenomena studied by psychophysics, namely how our visual acuity changes with luminance of the surroundings, about the relation of contrast sensitivity and visual frequency of the perceived visual stimuli, about masking that prevents some patterns in an image to be noticed, and about visibility thresholds. We have seen how the CIE colorimetric system is defined and described the definition of perceptually-uniform CIELAB and CIELUV colour spaces and the CIE colour difference formulae.

Even if this chapter is rather self-contained, its material will be often used in the remaining parts of the thesis.

Chapter 3

Perceptual Approaches in Computer Graphics and Image Processing

In other areas of computer graphics and computer vision, as image processing and coding or image reproduction, human perception-aware approaches have already been used to drive the comparison of the images, to predict the distortion of compressed images or to transform a high range of image luminances onto a narrow contrast range of the display material. As those methods form a good basis for algorithms that can be used in perceptually-driven radiosity approaches, this chapter provides a brief overview of existing algorithms and techniques.

3.1 Introduction

In the early days of computer science, the choices of CRT phosphors and of screen refresh rates were driven not just by technical and manufacturing constraints but also by detailed studies of phenomena as critical flicker fusion frequency and contrast sensitivity. The introduction of colour into display technology was backed up by usability studies that aimed at showing that using colour improves the operator performance. Compression algorithms for image and video data exist that take into account the discriminability of different colours by the human visual system. Performance of digital video compression algorithms is evaluated by mathematical models that mimic the visual paths of human brain and deliver prediction about visibility of different artifacts caused by the lossy video compression.

In the area of image processing, perceptual error metrics are used for image comparison and image coding that enable us to better predict differences between two images as opposed to the perceptually inappropriate and widely used mean-squared error metrics [19]. All the metrics are based on the fact that if some image imperfection is physically important it does not imply that it will disturb

the human observer. As these metrics are usually based on some approximate model of first stages of the human vision, the comparisons in perceptual space yield principally better results than classical comparisons of radiometric values stored in the image (Daly [22], Lubin [61], Teo and Heeger [114], Boker [14], Rushmeier *et al.* [88], Gaddipatti *et al.* [34]).

Tone reproduction operators known from image synthesis make it possible to map a bright scale of image luminances onto a narrow scale of CRT luminances in such a way that the perceived CRT image can be thought of as producing the same mental image as the original image (Tumblin and Rushmeier [119], Chiu *et al.* [20], Schlick [96], Ward [126], Larson *et al.* [57], Neumann *et al.* [69, 70], Tumblin and Turk [120]).

Modelling features of the human visual system allows us to simulate the effects of visual masking and light adaptation in dependence on current lighting conditions, and use the simulation results to determine the acceptable error that may be introduced by our computation but stays unnoticed by the observer (Ferberda, Pattanaik *et al.* [30, 31, 75]).

3.2 Tone-Mapping

The physical accuracy in rendering itself does not yet guarantee that the displayed images will have a realistic visual appearance [30]. Nowadays, typical CRTs can only display luminances in the range 1 to 100 cd/m² and typical print reproduction materials (film slides, paper) even less, whereas the range of luminances found in the real world can vary between 10⁻⁵ to 10⁵ cd/m² [35, 119] or even 10⁻⁶ to 10⁸ cd/m² [30]. Due to human visual system adaptation we can achieve that the subjective appearance of the real world scene can be reproduced on a display device, even if the luminances themselves cannot. This adaptation process is imitated by *tone-mapping operators*.

The aim of tone reproduction operators is to compute intensities for display on a device such as a CRT, so that when those intensities are viewed by an observer, the mental image they form is as close as possible to that of a real world scene.

The process of transforming radiosity B to the *display luminance* L_d is outlined in Figure 3.1. First, radiosity has to be transformed to radiance, which in turn has to be transformed to *real-world luminance* L_w . This transformation includes integrating the spectral representation of B_λ with corresponding CIE visual sensitivity function $V(\lambda)$. In case we are working with tristimulus values, some approximation — as discussed in Section 2.2 — has to be used. Then, a tone-mapping operator is applied that converts L_w to the appropriate display input that produces the luminance L_d on the display device.

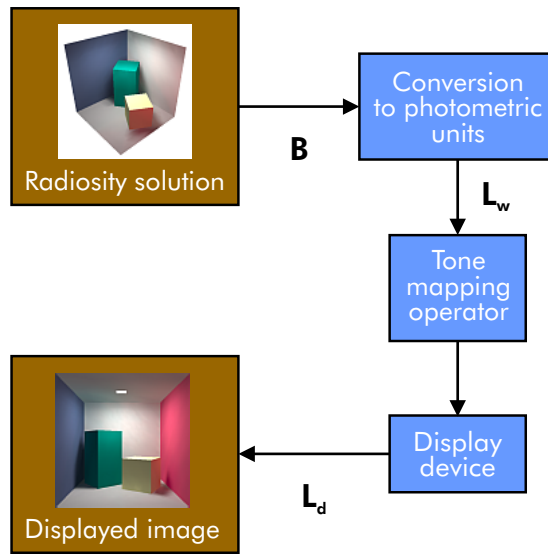


Figure 3.1: Transforming radiosity solution to display luminance

In the next sections we will briefly present the most widely known perceptually-based tone reproduction operators.

3.2.1 Tumblin & Rushmeier's Mapping

Tumblin and Rushmeier's model [118, 119] uses results obtained by Stevens and Stevens [110, 111] regarding the brightness associated with a luminance at a particular adaptation level. The tone reproduction is considered for gray-scale images only, since in this environment the spectral radiance is uniform for all wavelengths and the luminance is therefore just a constant times the uniform spectral radiance. Another reason for this limitation is that for each gray input value the display device gives a unique output luminance, which is not true for colour systems where many different colours may have the same luminance.

In Figure 3.2 we can see that the operator itself consists of three transformations: First, the real world luminance is transformed to the true real world brightness using Stevens' power law [109]. Then, using inverse display observer and inverse display device transformations the value of display input that generates the desired real world response is obtained.

The original mapping of a real world luminance value L_w to the corresponding

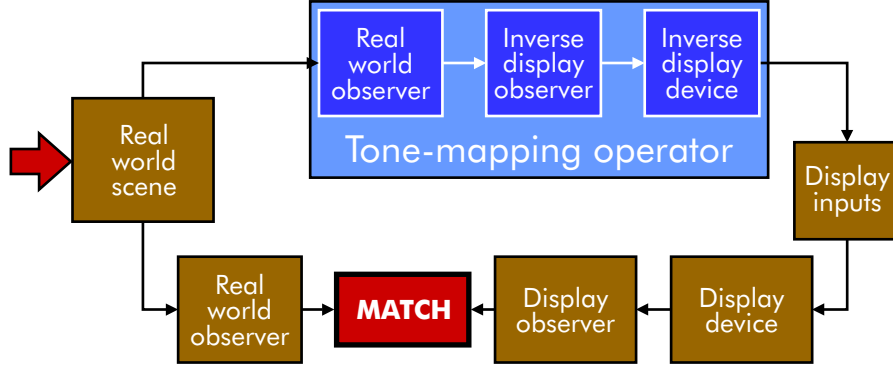


Figure 3.2: Tone-mapping algorithm. Radiances that are to be reproduced on display have to be scaled by the inverse transformation of real world observer, and the inverse transformation of the display observer. One has to account for nonlinearities in display device as well.

display value n from interval $\langle 0, 1 \rangle$ is described as

$$n = \left[\left(\frac{L_w \left(\frac{\alpha_w}{\alpha_d} \right)}{L_{d,\max}} \right) \cdot 10^{\left(\frac{\beta_w - \beta_d}{\alpha_d} \right)} - \left(\frac{1}{C_{\max}} \right) \right]^{\left(\frac{1}{\gamma_d} \right)}, \quad (3.1)$$

where α_w and β_w are constants expressing the real world observer adaptation, α_d and β_d are constants expressing the display observer adaptation, $L_{d,\max}$ is the maximum displayable luminance of the display device, C_{\max} is the maximum achievable contrast of the display device and γ_d is the gamma correction value.

For a given adaptation luminance L_{ax} the adaptation coefficients α_x and β_x can be written as

$$\alpha_x = 0.4 \log_{10}(L_{ax}) + 2.92, \quad (3.2)$$

$$\beta_x = -0.4 [\log_{10}(L_{ax})]^2 - 2.58 \log_{10}(L_{ax}) + 2.02, \quad (3.3)$$

where $x \in \{w, d\}$, and L_{aw} , α_w , and β_w are used for the real world observer whereas L_{ad} , α_d , and β_d are used for the display observer.

The displayed and the real world luminances are bound together by the formula

$$L_d = \zeta(L_w) = L_w^{\frac{\alpha_w}{\alpha_d}} \cdot 10^{\frac{\beta_w - \beta_d}{\alpha_d}}, \quad (3.4)$$

where ζ is called the tone-mapping function.

Gibson and Hubbold [35] report poor performance of the Tumblin and Rushmeier model in mesopic and scotopic conditions — their observations show that

the images appear much lighter than they should. This may be caused by the fact that the original operator inverts contrast for low adaptation levels [117].

3.2.2 Ward's Mapping

Another simple model was proposed by Ward [126] who tried to find a mapping function that would exhibit similar behaviour as the mapping of Tumblin and Rushmeier [119], but that would be less computationally demanding. His model attempts to match the just noticeable differences on the display device and in the real world and to produce a linear mapping function. He tries to find a constant of proportionality between display luminance and world luminance that would yield a display with roughly the same contrast visibility as the actual scene. This means that Ward's tone-mapping operator tries to preserve the perceived contrast, as opposed to Tumblin-Rushmeier's tone-mapping operator that preserves perceived brightness of an image.

Ward's operator consists of a linear formula

$$L_d = mL_w, \quad (3.5)$$

where L_d is the luminance to be displayed and L_w is the real world luminance of the image pixel. As the goal of the operator is to preserve the perceived contrast, the scale factor m binds together the minimum discernible luminance changes, i.e. the just noticeable luminance differences Δ_{JND} at the display and world adaptation levels L_{da} and L_{wa} :

$$\Delta_{\text{JND}}(L_{\text{da}}) = m\Delta_{\text{JND}}(L_{\text{wa}}). \quad (3.6)$$

The luminance mapping according to (3.6) will map the differences that are just visible in the real world to the differences that are just visible on our display. In order to determine a good approximation of m for the given picture and usual viewing conditions, we have to supply two luminance values: the world adaptation luminance L_{wa} and the maximum display luminance $L_{\text{d,max}}$. The display adaptation luminance L_{da} is estimated from the maximum display luminance as

$$L_{\text{da}} = \frac{1}{2}L_{\text{d,max}} \quad (3.7)$$

and the scale factor m can be computed as

$$m = \frac{1}{L_{\text{d,max}}} \left[\frac{1.219 + \left(\frac{L_{\text{d,max}}}{2}\right)^{0.4}}{1.219 + L_{\text{wa}}^{0.4}} \right]^{2.5}. \quad (3.8)$$

An example of image output from Ward's and Tumblin and Rushmeiers' tone-mapping operators is presented in Figure 3.3.



Figure 3.3: Images produced by the two most frequently used tone-mapping operators: top – after Ward [126], bottom – after Tumblin *et al.* [117] (revised version of Tumblin and Rushmeiers’ operator).

3.2.3 Proposals of Ferwerda, Pattanaik *et al.*

Ferwerda *et al.* [30] proposed another tone reproduction technique that is also based on the concept of matching just noticeable differences for a variety of adaptation levels. Their model captures the changes in threshold visibility, colour appearance, visual acuity (sometimes called contrast sensitivity) and time-course of adaptation of the human visual system.

As their work is based on experimental threshold detection data, they based their model on Ward's tone-mapping operator. Their extension consists of taking into account not only the cone characteristics (as Ward did) but also the rod threshold data. Moreover, their model also counts for the whole range of viewing conditions from scotopic over mesopic to photopic level. Using the adaptation dependent visual acuity function they are able to remove those image details that cannot be noticed by the human observer at a given adaptation level. Also, using the data measured for the time-course of the light and dark adaptation they can reproduce images so that they correspond to what the human observer would have seen when entering a dark or bright environment.

In the followup to this work [31], the authors study the effects of texture masking on masking out different artifacts in the resulting image. Their model consists of four stages similar to that of common models of early phases of human vision process: They first transform the spectral radiances into responses in a colour space, producing a colour representation of the image. In the next stage this colour representation is decomposed into pattern representations that account for different spatial frequency and orientation triggered detection mechanisms of the human visual system. Then an appropriate masking function is applied to each of these detection mechanisms to account for the effect of visual masking. In the last stage the results of masking are compared by the detection method to determine whether the input artifacts would be visible or whether the masking behaviour of the human visual system will mask them out and the human observer will not notice them at all.

Later, Pattanaik *et al.* [75] designed tone-mapping operator based on a multi-scale model of adaptation and spatial vision, and colour appearance. The model allows the operator to address the two major problems in realistic tone reproduction: wide absolute range and high dynamic range scenes can be displayed, and the displayed images match our perceptions of the scenes at both threshold and supra-threshold levels to the degree possible given a particular display device. The computational model uses spatial decomposition into image pyramids [18] for band-limited contrast processing and opponent-colour processing with non-linear transducer functions for achieving the proper colour appearance reproduction.

3.2.4 Visibility Preserving Mapping of Ward-Larson *et al.*

Ward-Larson *et al.* [57] present a technique that combines advantages of the previous ideas allowing us for correct display of scenes with wide dynamic ranges for different adaptation levels. The mapping operator is uniform and tries to preserve the visibility of objects in the image.

The task is achieved by using more than a single view adaptation level, which is in fact coherent with the mechanism of human visual system adaptation. The authors noticed that image luminance levels tend to be clusterised rather than to be uniformly distributed throughout the whole dynamic range of the picture. They state that as long as the bright areas are displayed with higher luminances than the dim areas, the absolute difference in the luminance is not so important.

Ward-Larson *et al.* first compute a luminance histogram and cumulative distribution function of all local adaptation luminances found in the image and discover clusters of similar adaptation levels. The adaptation luminances are computed for commonly accepted 1° field of view.

The histogram is then adjusted in order to minimise the visible contrast distortions. The clusters are mapped to the display values preserving local contrast of the cluster. The method uses also knowledge about veiling luminance (i.e. glare effects), colour sensitivity and visual acuity to count for imperfections of human vision.

3.2.5 Time-Dependent Adaptation

Recently, the attention of the tone-mapping research in the computer graphics community has turned to simulating the process of time course of visual adaptation, enabling us to simulate effects that are observed when the adapting luminance changes abruptly (i.e. entering a darkened room from a sunlit garden).

Pattanaik *et al.* [80] proposed a fast visual adaptation model that extends Hunt's model of static colour vision [46] with a time-dependent visual adaptation component. The model consists of the same building blocks as the original proposal of Tumblin and Rushmeier [119] depicted on Figure 3.2. First, the incoming colour signal is compressed with the time-dependent extension of Hunt's static adaptation model, and a very simple colour appearance model is applied to the adapted luminance value. These values, together with other inputs accounting for the correct visual range description, are then fed into an inverse appearance and adaptation model that produces RGB triplets suitable for display in the selected display device.

Durand and Dorsey [26] use a two-pass rendering algorithm capable of time-dependent tone-mapping, rendering flares, and simulation of the loss of visual acuity. In the first pass, their algorithm estimates the correct adaptation level of

the image. This value is then used in the second pass that performs the actual tone-mapping. The operator used in the paper is an updated version of the visual adaptation model of Ferwerda *et al.* [30] improved in the mesopic range, adding flares, after-images and using hardware convolution to simulate the loss of visual acuity in low light.

3.2.6 Tone-Mapping for Interactive Walkthroughs

Scheel *et al.* [95] present a tone-mapping extension for interactive walkthroughs of radiosity scenes that is based on the work of Ward and colleagues [57, 126]. They coarse sample the luminance in the view field using a ray-casting step and use an approach similar to incident light metering [69] to create a small texture that contains the tone-mapped luminances corresponding to the above mentioned tone-mapping operators. When displaying a vertex, the display luminance of the vertex is obtained from the precomputed texture using a clever mapping from real world luminance to texture coordinates. Thus the radiosity mesh can be rendered with hardware acceleration, avoiding the costly update of vertex colours after every frame.

3.2.7 Minimum Loss Methods

Neumann *et al.* [70] present an interesting and simple extension to the linear mapping principle. In their mapping they look for such a colour interval in the image histogram that can be linearly mapped with the given clipping contrast in a way that the minimum information is lost from the input data due to clipping. They present two slightly different approaches to this method: In the first approach they try to minimise the number of colour histogram bins that are clipped, in the second one they try to minimise the number of image pixels affected by clipping.

3.2.8 Low Curvature Image Simplifier

Classical tone-mapping algorithms do not perform very well when dealing with a vast luminance range that significantly exceeds the limited range of the luminance display. As a result, tiny details in the dark and bright regions of the image are usually lost in favor of mimicking the whole input range on the output device.

The low curvature image simplifier algorithm of Tumblin and Turk [120] is an attempt to mathematically mimic an artistic technique for rendering high contrast scenes — the authors noted that when drawing or painting, the artists capture visual appearance with a “coarse-to-fine” sequence of boundaries and shading. In this way, an artist has the freedom to drastically compress the contrast of large features and add the fine details with little or no attenuation to ensure they are

visible in the final image. The algorithm is not perceptually based in the sense that it does not determine whether the fine details that are being preserved would be indeed visible to a human observer.

3.2.9 Retinex Theory

Land's retinex theory of colour vision [56] estimates scene reflectances from the ratios of pixel intensities to their local intensity averages. Jobson, Rahmann, and Woodell [48, 50, 121, 122] used the retinex model to devise a full-colour local scaling and contrast reduction method for image processing. As the single-scale retinex model cannot simultaneously provide good dynamic-range compression and achieve acceptable tone reproduction, a multi-scale version has been proposed [49] that does not suffer from these problems. Marini and Rizzi [63] extended the original retinex implementation with mechanisms that make it possible to better account for colour constancy and achieve also good results when preserving the effects of simultaneous contrast.

3.3 Perceptual Difference Metrics

Nowadays, the mean-squared and root mean-squared error of the colour values are still widely used as convergence metrics for radiosity computations. The fact that these error metrics may produce totally misleading results when judging the similarity of two images perceived by the human observer has been widely discussed in the image evaluation literature in the past years [34, 36, 62, 88]. However, the only work from the field of global illumination that does not ignore this fact is the recent paper by Myszkowski [65].

Chiu *et al.* [20] state that the error metrics employed by rendering algorithms should pay less attention to low frequency errors (as errors caused by indirect lighting), because the human eye is not sensitive to those frequencies. This statement is probably also inspired by the work of Mannos and Sakrison [62], who studied the distortion measures in the scope of image encoding.

To develop an image comparison metric that would provide results corresponding to the results obtained from human observers is a very difficult task. There are however approaches that are already able to provide us with results that are meaningful to some extent. This section will briefly discuss the perceptually driven comparison metrics that are used in computer graphics today.

3.3.1 Visible Differences Predictor

The *visible differences predictor* (VDP) introduced by Daly [22] is an algorithm for describing the human visual response. Its goal is to determine the degree to which the physical differences between two images become visible to the human observer. The block structure of the algorithm is depicted in Figure 3.4.

The VDP is a relative metric — it does not describe the absolute image quality, but rather it describes the visibility of differences between two input images. The algorithm consists of three major components: a calibration component used to transform the input to values understood by the second component which is a model of the human visual system (HVS). The difference of HVS responses is then visualised by the difference visualisation component. The output of the VDP is an image map containing the probabilities of HVS detecting the differences between the input images for every pixel.

The calibration block obtains the image data and calibrates them so that they have any meaning to the subsequent HVS model block. The calibration process takes a number of input parameters describing the conditions for which the VDP will be computed. The parameters include the viewing distance of the human observer, the pixel spacing, and necessary values for the display mapping.

The HVS model concentrates on the lower-order visual system processing, that is, on the visual cortex. The model addresses three main sensitivity variations of the human visual system: the dependence of sensitivity on the illumination level, on the spatial frequency of visual stimuli, and on the signal content itself.

The variations in sensitivity as a function of luminance level are primarily due to light-adaptive properties of the retina — they are often referred to as the *amplitude nonlinearity* of the human visual system. The variations as a function of spatial frequency are expressed by the contrast sensitivity function (see Section 2.3.2). The dependence of sensitivity on the signal contents is due to the postreceptoral neural circuitry and is usually described as masking (see also [31]).

The difference visualisation block allows for two different visualisation techniques, displaying the difference probabilities either as a gray scale value on a uniform field of gray or displaying the probabilities in colour in the context of the reference image — in this case it is easier to judge the correspondence between the predicted differences and the differences actually observed between the two input images.

3.3.2 Sarnoff Visual Discrimination Model

Another model of the human visual system has been developed at Sarnoff Corporation by Lubin [61]. The model was meant as a tool serving purposes of design and evaluation of imaging systems. This visual discrimination model (VDM) fo-

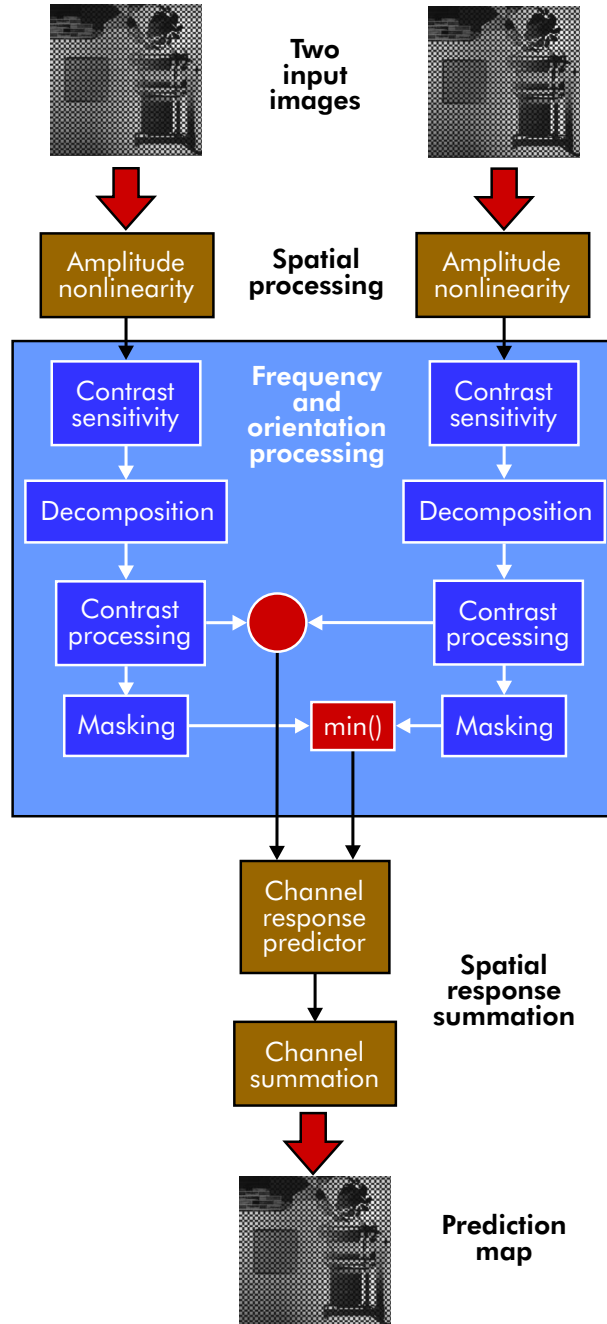


Figure 3.4: The block structure of the visible differences predictor.

cuses more attention on modelling the physiology of the visual pathway. It operates direct in the spatial domain and does a careful simulation of different optical mechanisms in human vision.

Empiric comparisons [59] show that both VDP and VDM deliver very similar results but that the Sarnoff VDM gives better image difference maps and requires less calibration. The VDM is also faster than VDP at the cost of significantly higher memory consumption. A simplified VDM forms a base of adaptive sampling algorithm by Bolin and Meyer [15].

3.3.3 Metric Based on the Normalisation Model of Early Vision

Teo and Heeger [114] present a perceptual distortion measure based on the so-called *normalisation model* of early phases of human vision. Their model fits the empirical measurements of the neuron response properties in the primary visual cortex and the psychophysics of the spatial pattern detection.

They use a set of linear sensors that are tuned to different spatial orientations and frequencies as a model for early stages of the human vision. The result of the combined sensor response is squared and transformed to normalised contrast values — this yields a response that is close to the response of neurons in the visual cortex. The simulated neuron responses to both images are then compared using the usual mean-squared error mechanism.

3.3.4 Comparison Experiments of Rushmeier *et al.*

Rushmeier *et al.* [88] stated the following rules defining properties of a perceptually-based image difference metric M :

1. If we are comparing the image A with itself, the predicted difference has to be zero, $M(A, A) = 0$.
2. The difference between two images A, B must not depend on their order in the comparison, $M(A, B) = M(B, A)$.
3. $M(A, C)/M(A, B) \gg 1$ for A and B appearing similar and A and C appearing different.
4. $M(A, C)/M(A, B) \approx 1$ for A, B and C appearing similar to one another.
5. $M(A, B)/M(C, D) \approx 1$ for the difference between A and B appearing similar to the difference between C and D .

Two of the three metrics proposed in [88] gave promising results. Both of the metrics transform the image luminances with fast Fourier transform (FFT) into the frequency space and use CSF in image luminances to account for the eye sensitivity on luminance variations.

The first promising model has been derived from the visual fidelity criterion of Mannos and Sakrison [62]. The image luminances are first normalised by the image mean luminance. A cubed root is then applied to the normalised values in order to account for nonlinearity of the human perception. In the next step, a FFT of the resulting values is computed transforming the image data into the spatial frequency space. The FFT result is filtered with the contrast sensitivity function (2.13). The pixel-based mean-squared error difference between the resulting filtered values is then used to obtain the metric value. Mathematically, we can write

$$\begin{aligned} f_r &= \sqrt{u^2 + v^2}, \\ A_M(f_r) &= 2.6(0.0192 + 0.114f_r) \exp[-(0.114f_r)^{1.1}], \\ g_{uv} &= f_{uv} \left(\sqrt[3]{\frac{L_{ij}}{L_{\text{mean}}}} \right) A_M(f_r), \\ M(X, Y) &= \frac{1}{N} \sum_{u,v} (g_{X,uv} - g_{Y,uv})^2, \end{aligned}$$

with $A_M(f_r)$ being the CSF for an *unspecified* adaptation level and f_{uv} being the FFT of the given argument.

The second model is inspired by the first part of Daly's VDP [22] and is similar to the previous one. However, the image luminances are not normalised — instead of that, a pixel-based transformation of luminances that counts for both adaptation and nonlinearity of human perception is used. Then, as in the previous model, a FFT of the transformed values is computed and the result is filtered using the adaptation level dependent CSF computed for the typical adaptation level 50 cd/m^2 , as has been proposed by Daly in his paper. The results for both images are then compared using the mean-squared error. Mathematically, we can write

$$\begin{aligned} b_{ij} &= \frac{L_{ij}}{L_{ij} + 12.6(L_{ij})^{0.63}}, \\ f_r &= \sqrt{u^2 + v^2}, \\ A_D(f_r) &= \left(\frac{0.008}{f_r^3} + 1 \right)^{-0.2} 1.42f_r \exp[-0.3f_r] \sqrt{1 + 0.06 \exp[0.3f_r]}, \\ g_{uv} &= f_{uv}(b)A_D(f_r), \\ M(X, Y) &= \frac{1}{n} \sum_{u,v} (g_{X,uv} - g_{Y,uv})^2, \end{aligned}$$

where $A_D(f_r)$ is Daly's CSF precomputed for adaptation level 50 cd/m^2 and f_{uv} denotes again the Fourier transform.

3.3.5 Wavelet Comparison Metrics

A wavelet-based approach has been used by Jacobs *et al.* [47] to query images from an image database based on simple drawings supplied by a user. The system compared wavelet “thumbnails” that were found to be close to the thumbnail of the query image. Besides the multi-resolution approach given by wavelets, the system did not incorporate any further models of human visual system.

Gaddipatti *et al.* [34] use a wavelet-based metric for the decision whether the rendering of two subsequent images has brought some perceptually noticeable difference or not. The metric introduces a simple human visual system model in that the wavelet coefficients are weighted by the contrast sensitivity function modelled after experiments of Mannos and Sakrisson [62]. Since the comparison involves the wavelet transformation, image comparison operates on different “levels of detail” in each step. This fact allows us to avoid the undesired features of the mean-squared error approaches, especially the sensitivity to the high frequency blur and brightness shifts.

Bradley’s *wavelet VDP* [17] predicts perceptual distortion at each pixel of an image that is to be encoded using lossy wavelet compression scheme. Bradley’s model builds upon Daly’s VDP [22], but it modifies the original difference predictor in several ways: the light adaptation preprocessing has been removed, separable wavelet transform has been introduced instead of the cortex transform of VDP, a wavelet CSF is applied directly to wavelet coefficients, and a simplified definition of sub-band contrast is used that allows for direct prediction of noise visibility from wavelet coefficients.

3.3.6 Boker’s Experiments

Boker [14] states that the human visual system performs near to an optimum value for an ideal trichromatic system composed of three linear components. His experiments with an unconstrained four factor maximum model seems to fit significantly better than a three factor unconstrained model, suggesting that a colour metric is better represented in four dimensions than in a three dimensional space.

He mentions that the visual system would attempt to preserve, as much as possible, the covariances present in the distribution of photon energies generated by the product of the illuminant spectra with reflectance spectra from objects present in the environment. The visual system adapts to differing brightness and overall spectral content of illumination sources such that a perception of *colour constancy* is maintained within a wide range of environmental lighting conditions. If colour constancy is to be achieved, the interaction between the mean and variance of each wavelength must be removed in order to preserve an invariant pattern of covariances between wavelengths reflected from objects in the environment. It is this

pattern of covariances that would define the perceived colour of an object.

3.3.7 Stochastic Metric by Neumann *et al.*

Neumann *et al.* [71] present a stochastic approach to image comparison. Their basic idea is to place a limited number of random-sized rectangle pairs in both compared images and to determine the average colour in each rectangle. For every rectangle pair the average colour difference is computed and weighted according to the rectangle size and the contrast sensitivity function. The result of the image comparison is then obtained by combining the CSF weighted colour differences for all tested rectangles. The use of quasi-random numbers makes the final method deterministic in the sense that repeating the computation with a set of fixed input parameters leads to exactly equal comparison results.

3.3.8 Colour Image Fidelity Metrics

Zhang *et al.* [130, 132] proposed another simple model for colour image comparison. The metric is based on S-CIELAB ΔE_{94} [131], an extension to the revised CIE colour difference formula [27] that is not very suitable for image fidelity assessment due to missing spatial extent of the original CIELAB formula. The purpose of S-CIELAB is to remove those image components that cannot be seen by the human eye before actually comparing the images. This is accomplished by converting the images into an opponent colour space and passing every colour channel through a spatial filter that mimics human eye spatial sensitivity for that colour component. Filtered images are then compared using the standard ΔE_{94} .

3.4 Summary

In this Chapter we have overviewed existing perceptually-based methods of image tone reproduction and image comparison that are currently used in computer graphics and image processing communities. These methods make it possible to drive the comparison of the images, to predict the distortion of compressed images or to transform a high range of image luminances onto a narrow contrast range of the display device.

We discussed the fact that the physical accuracy in rendering itself does not yet guarantee that the displayed images will have a realistic visual appearance. This is the place where tone reproduction operators are used to compute intensities for display on a device such as a CRT, so that when those intensities are viewed by an observer, the mental image they form is as close as possible to that of a real world scene.

We have seen that the fact that mean-squared and root mean-squared error metrics may produce totally misleading results when judging the similarity of two images perceived by the human observer has been widely discussed in the image evaluation literature in the past years. We have mentioned several metrics that use models of human visual perception to achieve results that correspond better to those of typical human observer.

The approaches presented in this chapter form a basis for algorithms that are used in perceptually-driven radiosity methods. Before discussing this topic, however, we will deal with Monte Carlo radiosity methods as these are a fundamental tool for the methods developed by the author and described later in this thesis.

Chapter 4

Monte Carlo Radiosity

The importance of generating realistic images from electronically stored scenes has significantly increased during the last few years. A number of methods have been introduced to simulate various effects which improve the realism of computer generated images.

Calculating the overall light propagation within a scene is a very difficult problem. With standard ray tracing algorithms, computing the global illumination of the scene is a very time consuming task since a huge number of rays have to be shot for every image that is to be generated. The radiosity method was invented to facilitate this task. The main idea of the method is to use only purely diffuse surfaces and store illumination values on the surfaces of the objects, as the light is propagated starting at the light sources.

However, even the deterministic radiosity algorithms, which have been used for quite some time, are too slow for calculating global illumination for very complex scenes. This led to the invention of stochastic methods that simulate the photon propagation using a Monte Carlo type algorithm.

This chapter briefly summarizes achievements in the radiosity and Monte Carlo research in the past years. More detailed explanation of Monte Carlo algorithms for radiosity may be found in the theses of Bekaert [5], Sbert [90] and Keller [55]. Good introductory texts on Monte Carlo methods in general are the books of Kalos and Whitlock [51] or Rubinstein [87].

4.1 Introduction

The radiosity method (see [21, 37, 105] for an introduction) belongs to the broad family of finite element methods. In general the method works as follows: The light propagation in the scene is described using a radiometric quantity called radiosity. The scene is represented as a mesh of planar patches, which makes

it possible to approximate the original integral algebraic equation describing the propagation of radiosity in the space by a set of linear equations. The influence of radiosity some patch leaves on the others is described by a set of form factors. The radiosity function representing the light distribution in the scene is then computed by solving the above system of linear equations.

First methods of this type used the Gauss-Seidel or Southwell relaxation to solve the linear system. Early approaches to compute the form factors used Z-buffer-like algorithms for computing the patch-to-patch visibility and suffered from aliasing problems. In these days, these methods have been replaced by more elaborate ones. We present just a short overview.

Using a simple mesh structure, one often computes radiosity transfers that have a very small influence on the final error of the radiosity system solution. Moreover, if the scene is large the number of form factors that have to be computed grows rapidly, slowing down the computation to the very edge of usability. A first step towards improving this situation was a two-level hierarchy proposed together with adaptive mesh refinement techniques. From this algorithm, more sophisticated algorithms of rapid hierarchical radiosity have been developed in the early 90s by Hanrahan *et al.* [41]. Followups to this work include adopting the concepts to wavelet radiosity, parallelization of the method and using clustering strategies to further speed up the computation (Gortler *et al.* [39], Smits *et al.* [107], Bohn and Garman [13], Stuttard *et al.* [113], Arnaldi *et al.* [2], Funkhouser [33], Bekaert and Willems [11]).

In the meantime, a faster Monte Carlo approach to solving integral equations has been applied to radiosity (Pattanaik and Mudur [76], Neumann *et al.* [68, 73]). Further improvements include better sampling techniques (Keller [53], Neumann *et al.* [72]) and hierarchical methods (Tobler *et al.* [116], Bekaert *et al.* [6]).

As the radiosity scene grows in complexity, it is becoming less desirable to compute a radiosity solution for the whole complex scene, since the observer is often interested in the image of his/her immediate surroundings only. For such cases importance-driven radiosity solutions have been proposed that use a patch property dual to radiosity to express the importance of a patch for the observer. Extensions to this method include the combination of progressive refinement and importance radiosity, hierarchical importance-driven radiosity and the combination of Monte Carlo and importance-driven radiosity (Smits *et al.* [108], Bekaert and Willems [8, 9], Neumann *et al.* [67]).

Nowadays, higher order elements and Galerkin approaches to solving radiosity are often used (Zatz [129], Feda [28]). Form factors in hierarchical radiosity are usually computed using Monte Carlo ray-casting (Wallace *et al.* [123], Keller [52]); the Monte Carlo radiosity approach does not even require form factors to be explicitly computed. A fast hybrid algorithm combining a stochastic solution with hardware-accelerated rendering has also been developed (Keller [54]).

In the rest of this chapter we will shortly explain how the Monte Carlo radiosity and hierarchical refinement work.

4.2 Monte Carlo Radiosity Methods

The fundamental idea of the Monte Carlo method is to formulate the solution of a given mathematical problem as a parameter of a carefully chosen random variable. By sampling this variable the parameter yielding the solution of the original problem can be estimated. Typically the parameter we estimate is the mean value of the random variable. The main advantage of Monte Carlo methods is their wide applicability and their conceptual simplicity. The main drawback of using Monte Carlo methods is their slow convergence.

The original radiosity method [38] solves the radiosity system of linear equations

$$B_i = E_i + \rho_i \sum_{j=1}^M F_{ij} B_j, \quad (4.1)$$

where B_i is the radiosity of the receiving element, E_i the self-emitted radiosity of this element, ρ_i is the reflectance of the element, F_{ij} the patch-to-patch form-factor determining the part of the source element radiosity B_j is received at the patch i .

Monte Carlo radiosity is based on particle transport and solves the power form of the system (4.1) that is given by

$$P_i = W_i + \rho_i \sum_{j=1}^M F_{ji} P_j, \quad (4.2)$$

where P_i is the power of a receiving element, W_i is the self-emitted power of this element, and F_{ji} the form-factor determining how much the power P_j radiated from element j contributes to the incoming power of the element i . Equation (4.2) is solved by probabilistic simulation of photon paths traveled by photons leaving light sources in the scene.

There exist two basic classes of Monte Carlo radiosity methods:

- *stochastic relaxation methods* that use Monte Carlo integration to iteratively estimate the coefficients of the system (4.2), and
- *random walk methods* that directly estimate the entire Neumann series expansion of (4.2).

4.2.1 Stochastic Relaxation Radiosity

The basic idea of relaxation methods, such as the Jacobi or Gauss-Seidel iterative methods, Southwell relaxation or conjugate gradient methods is to construct a sequence of approximate solutions to the system of linear equations that converge to the true solution of the system. In the iterative process, the next approximation is constructed by adding a correction term to the current approximation. The way the correction term is chosen makes the difference between the particular relaxation algorithms.

We will now describe the stochastic Jacobi relaxation radiosity algorithm of Bekaert [5, Chap. 6]. The algorithm is a combination of *incremental stochastic Jacobi relaxation* steps and *regular stochastic Jacobi relaxation* steps. It has to be noted that the well-distributed ray-set (WDRS) algorithm of Neumann *et al.* [72] may be slightly more effective than the below described combination, however, this gain has to be paid for by a high price: Since the WDRS has to be able to “remove” some portions of the energy in order to maintain the ray-set optimally distributed, it relies on almost constant photon energy being used in the whole scene in a single iteration. As a result, the WDRS cannot be extended to an importance-driven radiosity algorithm. Also, the process of the well-distributed ray-set construction is more difficult to implement than the combination of regular and incremental stochastic Jacobi method.

Incremental Stochastic Jacobi Method

The incremental stochastic Jacobi radiosity method computes the radiosity P_i of a patch i as the sum of intermediate results obtained in each step k according to

$$\Delta P_i^{(k)} = \rho_i \sum_{j=1}^M F_{ji} \Delta P_j^{(k-1)}.$$

The solution for patch i is then

$$P_i = \sum_k \Delta P_i^{(k)}$$

with $\Delta P_i^{(0)} = W_i$. The implementation of the method is described in Algorithm 1.

Regular Stochastic Jacobi Method

The regular stochastic Jacobi method (see Algorithm 2 and publications of Neumann and colleagues [68, 72, 73]) is in its nature very similar to the incremental stochastic Jacobi. The difference is that in every iteration step the regular Jacobi

Algorithm 1 Incremental stochastic Jacobi iterative method for radiosity.

```

for all patches  $i$  do
  initialise total power  $P_i \leftarrow W_i$ 
  clear unshot power  $\Delta P_i \leftarrow 0$ , and received power  $P_{\text{rcv},i} \leftarrow 0$ 
  update total unshot power  $\Delta P_{\text{total}} \leftarrow \Delta P_{\text{total}} + P_i$ 
end for
while  $\|\Delta \mathbf{P}\| < \varepsilon$  or number of steps lower than maximum do
  choose number of samples  $N$ 
  generate a uniform random number  $\xi \in \langle 0, 1 \rangle$ 
  initialise  $N_{\text{prev}} \leftarrow 0$  and  $q \leftarrow 0$ 
  for all patches  $i$  do
     $q_i \leftarrow \Delta P_i / \Delta P_{\text{total}}$ 
     $q \leftarrow q + q_i$ 
     $N_i \leftarrow \lfloor q \cdot N + \xi \rfloor - N_{\text{prev}}$ 
    for all  $N_i$  samples from patch  $i$  do
      uniformly sample a random point  $\mathbf{x}$  on patch  $i$ 
      sample cosine-distributed direction  $\theta$  at  $\mathbf{x}$ 
      determine patch  $j$  containing the nearest intersection point with the
      scene surfaces of the ray originating at  $\mathbf{x}$  with direction  $\theta$ 
       $P_{\text{rcv},j} \leftarrow P_{\text{rcv},j} + \Delta P_{\text{total}} \cdot \rho_j / N$ 
    end for
     $N_{\text{prev}} \leftarrow N_{\text{prev}} + N_i$ 
  end for
  clear total unshot power  $P_{\text{total}} \leftarrow 0$ 
  for all patches  $i$  do
    increment total power  $P_i \leftarrow P_i + P_{\text{rcv},i}$ 
    replace unshot power  $\Delta P_i \leftarrow P_{\text{rcv},i}$ 
    clear received power  $P_{\text{rcv},i} \leftarrow 0$ 
    update total unshot power  $\Delta P_{\text{total}} \leftarrow \Delta P_{\text{total}} + \Delta P_i$ 
  end for
  display image using  $P_i$ 
end while

```

propagates the total power of the system and that the result replaces the previous power approximation rather than being added to it. The regular Jacobi iteration therefore computes the new power $P_i^{(k)}$ of a patch i as

$$P_i^{(k)} = W_i + \rho_i \sum_{j=1}^M F_{ji} P_j^{(k-1)}$$

Algorithm 2 Regular stochastic Jacobi iterative method for radiosity.

```

for all patches  $i$  do
  initialise power  $P_i \Leftarrow W_i$ , and received power  $P_{\text{rcv},i} \Leftarrow 0$ 
  update total power  $P_{\text{total}} \Leftarrow P_{\text{total}} + P_i$ 
end for
while converged or number of steps lower than maximum do
  choose number of samples  $N$ 
  generate a uniform random number  $\xi \in \langle 0, 1 \rangle$ 
  initialise  $N_{\text{prev}} \Leftarrow 0$  and  $q \Leftarrow 0$ 
  for all patches  $i$  do
     $q_i \Leftarrow P_i / P_{\text{total}}$ 
     $q \Leftarrow q + q_i$ 
     $N_i \Leftarrow \lfloor q \cdot N + \xi \rfloor - N_{\text{prev}}$ 
    for all  $N_i$  samples from patch  $i$  do
      uniformly sample a random point  $\mathbf{x}$  on patch  $i$ 
      sample cosine-distributed direction  $\theta$  at  $\mathbf{x}$ 
      determine patch  $j$  containing the nearest intersection point with the
      scene surfaces of the ray originating at  $\mathbf{x}$  with direction  $\theta$ 
       $P_{\text{rcv},j} \Leftarrow P_{\text{rcv},j} + \rho_j P_i / (q_i \cdot N)$ 
    end for
     $N_{\text{prev}} \Leftarrow N_{\text{prev}} + N_i$ 
  end for
  clear total power  $P_{\text{total}} \Leftarrow 0$ 
  for all patches  $i$  do
    replace power  $P_i \Leftarrow W_i + P_{\text{rcv},i}$ 
    clear received power  $P_{\text{rcv},i} \Leftarrow 0$ 
    update total power  $\Delta P_{\text{total}} \Leftarrow \Delta P_{\text{total}} + P_i$ 
  end for
  display image using  $P_i$ 
end while

```

Stochastic Jacobi Radiosity Method

The stochastic Jacobi radiosity method is a combination of the incremental and regular stochastic Jacobi relaxation radiosity methods. The particular steps of stochastic Jacobi radiosity method are outlined in Algorithm 3. More detailed discussion of the topic is available in Bekaert [5].

Algorithm 3 Stochastic Jacobi radiosity method.

```

repeat
  perform a single incremental Jacobi iteration
until  $\|\Delta\mathbf{P}^{(k)}\| < \epsilon$ 
repeat
  perform a single regular Jacobi iteration
until  $\|\Delta\mathbf{P}_{\text{total}}^{(k)}\| < \epsilon$  or number of steps exceeds maximum

```

4.3 Hierarchical Refinement

Multi-resolution energy transport has been introduced to radiosity by Hanrahan *et al.* [41]. The hierarchical radiosity algorithm makes it possible to effectively compute all the radiosity transport in the scene at a given accuracy level — some elements may interact at higher levels of the hierarchy, some of them have to be subdivided in order to fulfill the accuracy conditions. The authors limit an overall error of the computation by allowing the energy to be transported only over such links from shooter j to receiver i that fulfill the condition

$$B_j F_{ij,\text{est}} < B_\epsilon. \quad (4.3)$$

Here, $F_{ij,\text{est}}$ is a cheap form-factor estimate that gives an upper bound of the actual form-factor and B_ϵ is the user-supplied error threshold that determines the accuracy of the solution.

4.3.1 Hierarchical Refinement in Monte Carlo Radiosity

Monte Carlo radiosity approaches have lacked a suitable hierarchical refinement approach for some period of time. Heckbert [43] and Tobler *et al.* [116] proposed element subdivision schemes for continuous random-walk algorithms that used adaptive photon maps. The deficiency of these methods was the need to discard the current power stored at an element as the element was refined. The authors of the latter scheme reported 25% of recorded photon hits are typically discarded later due to element refinement.

The WDRS algorithm [72] was extended to work with the hierarchical element subdivision by Bekaert *et al.* [6]. The method is a variant of stochastic Jacobi approach, which makes it possible to evaluate the power-equivalent of the refinement condition (4.3) for every particle hitting a receiving surface with decent computational cost. An interaction between a source j and a receiver i will be refined if

$$P_j \frac{A_i}{A_j} F_{ij, \text{est}} > P_\epsilon, \quad (4.4)$$

where A_i and A_j are surface areas of the receiver and source patch.

4.4 Random Walk Methods

Contrary to stochastic relaxation methods, the solution of linear systems by random walks is a well-covered topic in literature (see theses of Sbert [90] and Bekaert [5] for more discussion on this topic). Random-walk methods were proposed for solving various global illumination problems [76, 98]. The application of random walk to radiosity has been studied by Sbert [90, 91, 92, 93].

Importance-driven extensions to continuous random walks that concentrate most of the particle paths to the region of interest have been discussed by Pattanaik and Mudur [77, 78, 79]. A hierarchical extension to the random-walk radiosity has been developed by Heckbert [43] and Tobler *et al.* [116].

4.5 Summary

Monte Carlo radiosity algorithms are — as well as their deterministic counterparts — based on algorithms used to compute radiative energy transport. Their advantage is that they quickly deliver solutions in which higher order inter-reflections are visible. Unfortunately, as with all Monte Carlo methods, the variance of the solution drops slowly and the results suffer from noise.

In this chapter we have briefly explained the stochastic Jacobi radiosity, a combination of stochastic Jacobi incremental and regular radiosity methods. We have also talked about how the construction of the element hierarchy can be controlled in classical and Monte Carlo hierarchical radiosity methods. We noted that while hierarchical extensions to random walk radiosity exist, they are bound to be very ineffective in terms of convergence as part of the already existing solution has to be discarded after a subdivision takes place.

Chapter 5

Importance-Driven Radiosity

Radiosity algorithms generally attempt to compute radiosity to a uniform precision throughout the whole environment. This results in globally over-solved and locally under-solved radiosity solutions for most scenes [108].

If we allow low accuracy of the solution in those parts of the scene, that are not directly visible and that do not influence the visible parts too much, we can spend more computational effort on the parts that are directly visible. In this way we can save a lot of computational time when we are interested in illumination of only a part of a complex scene.

The principle of the method is well-known: During the course of radiosity system computation, we are computing a second quantity called *visual importance*. This quantity expresses the influence the radiosity of a particular mesh element has on the solution in the visible part of the scene.

In this chapter a hierarchical extension to the stochastic Jacobi radiosity is presented. The method combines the importance-driven Monte-Carlo radiosity and the hierarchical refinement of scene mesh elements. This allows us to compute a faster converging solution in the visible surroundings of the observer and in the same time to influence the size of mesh elements in such a way that the discretisation error in these areas is decreased.

5.1 Importance-Driven Radiosity

The original idea of hierarchical refinement in radiosity [41] was extended by Smits *et al.* [108] by incorporating the view importance into the hierarchical refinement criterion (see Sec. 4.3) in order to decrease the number of subdivisions in invisible areas. In this approach, a link is considered acceptable only if

$$B_j \rho_i \Psi_i F_{ij, \text{est}} < B_\epsilon \Psi_{\text{max}}, \quad (5.1)$$

where Ψ_i is the radiosity-like form of the view importance at the receiver and Ψ_{\max} is a correction factor that corresponds to the maximum visible radiosity-like importance. As a consequence, the hierarchical subdivision in visible areas is comparable to that of the non-importance-driven method while in invisible areas the level of subdivision is significantly reduced.

Another possibility of using the view importance for the hierarchical radiosity solutions has been shown by Bekaert and Willems [8]. They used the view importance to order the shooters in the course of shooting iterations so that light sources having the most influence at the region of interest are processed first. This approach was combined with that of Smits *et al.* in [9].

5.2 Importance-Driven Stochastic Jacobi Radiosity

The importance-driven extension to stochastic Jacobi radiosity has been proposed by Bekaert [5] and is partially based on the former work of Neumann *et al.* [67]. The method makes use of *visual importance* to concentrate the computational effort to those parts of the scene that are visible at the given moment. The visual importance of an element can be computed using an equation that is adjoint to (4.2),

$$I_i = V_i + \sum_{j=1}^M F_{ji} \rho_j I_j. \quad (5.2)$$

Here, I_i is the power-like importance of the i -th element, V_i stands for its initial directly received importance. The value of V_i can be determined using the usual approaches for the computation of point-to-patch form-factors [21, 37, 105].

An importance-driven iteration generally consists of two steps: First, the iteration that propagates importance is performed. The new importance values are then used to influence the propagation of power in the second step.

The power-like importance is propagated in the same way as power is propagated in the stochastic Jacobi radiosity method. As the importance I_i is an incident quantity, it has to be multiplied by the surface reflectance ρ_i prior to shooting, which results in this form of element sampling probabilities:

$$q_{i,\text{imp}} = \frac{\rho_i I_i}{\sum_{k=1}^M \rho_k I_k}. \quad (5.3)$$

The element sampling probabilities $q_{i,\text{pow}}$ for power propagation now depend not only on the power of element i but also on its view importance,

$$q_{i,\text{pow}} = \frac{P_i I_i / A_i}{\sum_{k=1}^M P_k I_k / A_k}. \quad (5.4)$$

As I_i in (5.2) is the power-like form of importance, it had to be converted to radiosity-like importance by dividing by the element area A_i to properly scale the values of P_i .

It can be shown that using the total importance value for determining the element sampling probability is not an optimal choice [5]. In many situations we can do better by using just the received importance,

$$q_{i,\text{pow}} = \frac{P_i(I_i - V_i)/A_i}{\sum_{k=1}^M P_k(I_k - V_k)/A_k}. \quad (5.5)$$

Until now we have been describing an importance-driven radiosity scheme that uses separate steps for importance transport (stochastic Jacobi radiosity with element sampling probabilities corresponding to (5.3)) and for power transport (sampling probabilities given by (5.5)). However, it is possible to combine both steps into a single one in which power and importance are propagated simultaneously. The element sampling probability is then a combination of both the above mentioned equations:

$$q_{i,\text{comb}} = (1 - \alpha) \frac{P_i(I_i - V_i)/A_i}{\sum_{k=1}^M P_k(I_k - V_k)/A_k} + \alpha \frac{\rho_i I_i}{\sum_{k=1}^M \rho_k I_k}. \quad (5.6)$$

The value of α , $0 \leq \alpha \leq 1$ determines how much computational effort in the combined power-importance iteration is spent on importance propagation. A value of $\alpha = 0.1$ is often used in praxis.

The use of importance influences the size of the elementary power quanta that are being shot from a particular element. Knowing that a contribution of a single particle in the regular stochastic Jacobi step is

$$w_{i,\text{pow}} = \frac{P_i}{Nq_{i,\text{pow}}}$$

we can see that higher importance would cause more particles carrying lower energy quanta to be shot, effectively resulting in lower variance in regions with higher importance.

5.3 Importance-Driven Hierarchical Stochastic Jacobi Radiosity

When looking for a fast Monte-Carlo radiosity method that could be extended to a hierarchical importance-driven method, we first tried to use the well-distributed ray set method [72]. However, the WDRS heavily relies on the fact that the elementary power quantum remains approximately constant for all the patches in the

scene. If the elementary quanta differ too much, the method does not converge. Our initial approach [83] was therefore based on the stochastic ray radiosity [73] and its importance-driven extension [67]. Inbetween, an equally fast stochastic Jacobi variant of the radiosity method has been developed by Bekaert [5] that overcame the deficiencies of WDRS (see Section 4.2.1). The results presented in the following text are therefore based on the stochastic Jacobi variant.

Our new method extends the hierarchical stochastic Jacobi radiosity to a full importance-driven method of the type introduced in Section 5.2 and it extends the hierarchical refinement to take importance into account as discussed in Section 4.3.

The algorithm works either in two separate importance and power passes or in a single pass, where importance and power are propagated simultaneously.

As we work with power and power-like importance values, the hierarchical refinement criterion (4.4) for a source element j and a receiver element i has to be modified to take power-like importance into account, yielding the criterion

$$P_j \frac{1}{A_j} \rho_i I_i F_{ij,est} > P_\epsilon I_{max}, \quad (5.7)$$

where I_{max} is the correction factor that has the same function as ψ_{max} in (5.1). The resulting importance-based hierarchical refinement oracle is presented in Algorithm 4.

Algorithm 4 Importance-driven subdivision oracle.

Require: Source element j and receiver element i

Estimate the receiver-to-source form-factor $F_{ij,est}$

$P_{est} \leftarrow P_j \frac{1}{A_j} \rho_i I_i F_{ij,est}$

if $P_{est} > P_\epsilon$ **then**

 Refine link ji

else

 Link ji is precise enough

end if

As the scene is now composed of an element hierarchy, importance values can be stored either in top level patches or in leaf elements of the hierarchy. We do not allow for element subdivision due to importance transfer as importance serves only as an additional information for the radiosity algorithm, and even if it is computed with a higher discretisation error it still serves its purpose well. The subdivision due to importance transport would be necessary only if we were interested in the exact importance values on the mesh.

5.4 Results

We have implemented the method outlined in the previous Section and evaluated its convergence. As the method is importance based, we compared the resulting image quality using perceptual image comparison after Mannos and Sakrison [62] as described by Rushmeier *et al.* [88].

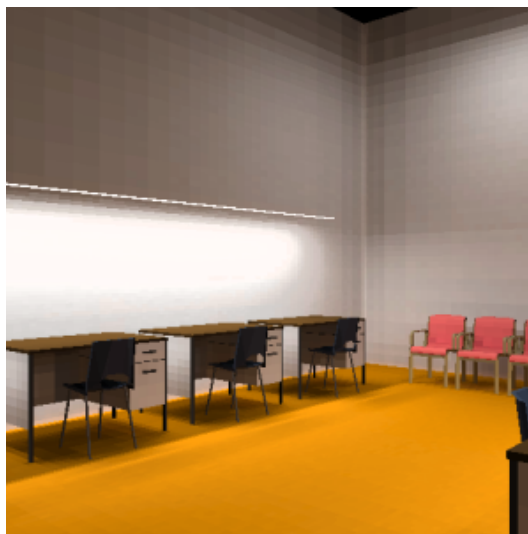


Figure 5.1: Reference solution.

The reference solution shown in Figure 5.1 was obtained by running 100 iterations (640 steps) of hierarchical well-distributed ray set radiosity [6] using 433×10^6 particles and subdividing the initial 23182 patches into 308018 elements.

Figure 5.2 presents the result of the importance-driven hierarchical method with combined power-importance propagation (upper image) and the classical hierarchical method (lower image) for 1.1×10^6 particles traced. The perceptual distance to the reference solution of solutions generated using importance and without importance is depicted in Figure 5.3.

Figure 5.4 shows the indirect importance solution and radiosity in a maze scene after 30 iterations. An interesting observation is that the variance in unimportant regions is not as extremely high as with the non-hierarchical importance-driven method. The reason for this phenomenon is the importance-driven subdivision scheme: the accuracy of the solution in non-important regions is still lower, but this time the inaccuracy shows up more as the discretisation error than as the variance.



Figure 5.2: Top: Importance-driven solution after 100 iterations with 1.1×10^6 particles (48702 elements). Bottom: Corresponding hierarchical solution without use of importance (324606 elements).

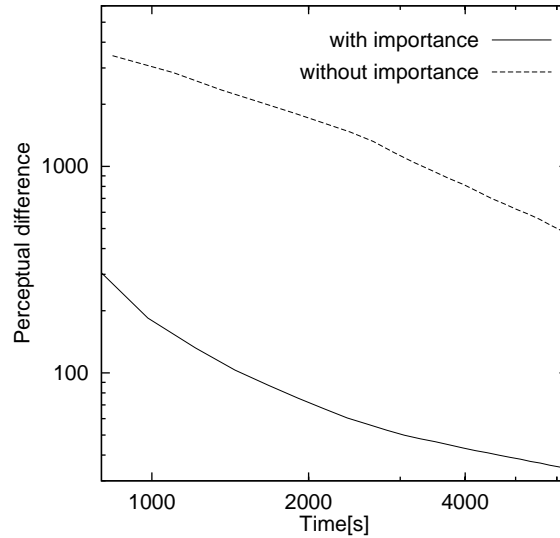


Figure 5.3: Image-based comparison of importance-driven and classical hierarchical stochastic ray radiosity.

Finally we were interested in the influence of different minor improvements of the importance-driven method on its convergence. The questions that we wanted to answer were:

- How large is the influence of hierarchical importance storage on the solution quality?
- Does the combined power-importance iteration scheme outperform the separated power and importance iterations?

The results of the tests are shown in Figures 5.5 and 5.6. We can see that there is just minimal difference in convergence speed for the separate and the combined importance propagation scheme, even if the combined scheme scores a bit better. An interesting observation is that the hierarchical importance storage combined with the separate propagation scheme is even better than the combined propagation scheme. All other improvements exhibit roughly the same convergence rate.

5.5 Summary

In this chapter we presented a hierarchical importance-driven stochastic Jacobi radiosity algorithm which consists of a combination of the importance-driven Monte-Carlo radiosity approach and of the hierarchical stochastic Jacobi method.



Figure 5.4: From top to bottom: Importance solution, corresponding importance-driven hierarchical radiosity solution and importance driven solution on fixed size elements for the maze scene without furniture.

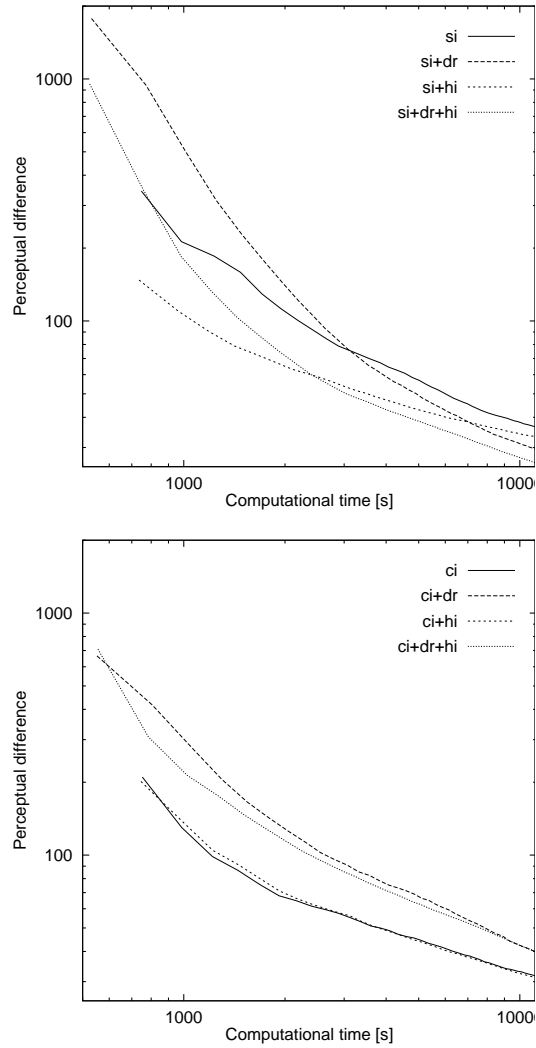


Figure 5.5: Influence of different improvements in the algorithm on its overall convergence. Method description: *si* – separate importance and power propagation, *ci* – combined power-importance iterations, *dr* – no refinement during the warm-up phase, *hi* – hierarchical importance.

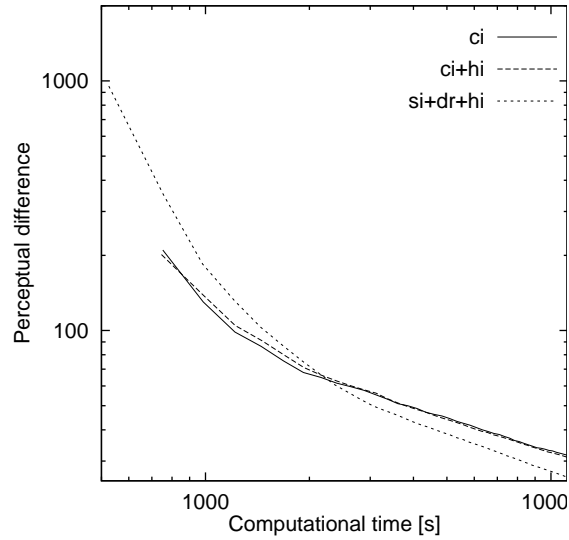


Figure 5.6: Best methods selected from Figure 5.5 (the keys are the same as before).

As one can expect, the method we presented indeed converges faster than its classical counterpart. For our tests scenes, the importance driven solution needed only 15% of the memory that was occupied by the elements produced by the non-importance-driven method. Our further experiments showed that the fastest convergence could be achieved by using either the separate importance propagation scheme combined with deferred element refinement and hierarchical importance storage, or using the combined power-importance propagation scheme.

As we could see in our measurements, it probably does not pay off to store the importance in the hierarchy: the convergence rate of the method using hierarchical importance was in the best case approximately equal to the convergence rate of the method that stores importance values only at top-level elements.

Chapter 6

Perceptually-Driven Radiosity

The radiosity method uses criteria based on radiometric values to drive the computation — to decide about sufficient mesh quality or to estimate the error of the simulation process. This is absolutely correct for the case of radiometric simulation, when the user is interested in actual values of radiometric quantities. On the other hand, the radiosity method is very often used just to generate images for the human observer and those pictures do not require to be correct physical simulations, they just have to look the same.

Therefore, it seems beneficiary to take the behaviour of the human visual system into account when producing radiosity images and, using this knowledge combined with the knowledge about the process of transforming the scene luminances for display, to develop criteria in a perceptually mapped space in order to decide what actually causes a visible error in the radiosity output and what can be safely ignored. Exploiting the fact that current display devices can not by large reproduce the real world range of luminances, or knowing which difference in contrast or colour can be noticed by a human under given viewing conditions, we hope to be able to reasonably decrease the time needed for a radiosity simulation. However, we have to be also aware of the fact that under certain circumstances the perceptually-driven solution will require more computational effort than a radiosity solution that is computed to some larger physical error.

This chapter presents a common methodology used in perceptually-driven radiosity methods. We will discuss typical building blocks of a perceptually-driven radiosity algorithm and an inherent problem of all perceptually-driven computations: a priori estimation of the correct visual adaptation level.

6.1 Principles

By ensuring that only those features of the radiosity solution that are perceptually important are computed to great accuracy, we hope to be able to produce visually acceptable images in a more efficient manner and thus faster.

Because the mechanism by which we perceive the radiosity values is nonlinear, any metric used inside the radiosity system that is based solely on radiometric comparisons cannot guarantee that additional errors will not be introduced during the display process. Differences in radiosity values that pass the refinement test (and therefore cause subdivision) may in fact be undetectable after the display transformation has been performed. The same is true in the opposite direction: we cannot guarantee that radiosity changes that lay below some threshold value will not become visible under certain conditions.

In case that we are more concentrated on the resulting images than on the actual radiometric values obtained during the global illumination simulation, one way of improving the existing radiosity methods is to *compute only the necessary things*, that means to concentrate our computational effort on places where the errors are visible and disturbing. To achieve the maximal effect, the existing algorithms use

- perceptual approaches to hierarchy subdivision that include perceptual error metrics as a part of the subdivision oracle,
- perceptual convergence metrics instead of radiometric metrics based on the mean-squared error.

6.1.1 Assumptions of Perceptually-Driven Computation

Perceptually-driven radiosity computations are usually based on the following three assumptions:

First, computations will be carried out in a well-defined colour system (usually CIEXYZ) and not in RGB colour space, as RGB is not precisely defined and actual colour values depend on the used CRT phosphors or inks.

Second, the algorithm uses a perceptually-based tone-mapping operator to map the real world radiosity to display luminances before doing any further evaluation. This makes it possible to directly evaluate changes in the displayed radiosity values.

Last, displayed radiosity changes are computed using the CIE colour difference formulae (2.29) and (2.30), usually using CIELUV as the approximation of a perceptually uniform colour space despite the negative experiences mentioned in Section 2.5. Due to the very limiting specifications of reference viewing conditions that have to be met to compute meaningful results using the CIE94 colour

difference formula (2.31), this formula is not used to compute visible differences in perceptually-driven radiosity algorithms.

6.1.2 Adaptation Luminance Estimation

In order to be able to use any perceptually-based approach, we have to know the lighting conditions under that visual tasks are being performed. Information about the adaptation luminance is crucial for estimating performance of the human visual system. However, at the very beginning of the computation this information is unknown. We therefore need some means of estimating the adaptation luminance before the computation starts.

Several possibilities to address this problem are available:

- Gibson and Hubbard [35] assumed that the lighting of the scene is approximately uniform. This allows for using an ambient term as an estimator for the average irradiance the human observer encounters and computing the adaptation luminance as a logarithmic area weighted average of estimated surface luminances, adjusted so that brightnesses are kept around the Stevens' brightness constancy level [110, 111] of 8.4 dB.
- A possibility to improve the performance of the original ambient term-based approach would be to use a nonuniform ambient term that would be computed separately for different parts of the scene. The structure of these separate ambient cells may be based on hierarchical space subdivision structures that are normally used to accelerate ray-casting computations.
- When solving the illumination using a Monte Carlo radiosity algorithm, it would be also possible to compute a cheap estimation of the solution with some variant of the first shot method. For a view-dependent radiosity algorithm, the estimate of the adaptation luminance may be then computed directly from the radiosity of the visible scene patches.

Figure 6.1 shows a radiosity scene displayed with Ward's tone-mapping operator for two different methods of adaptation luminance estimation.

6.1.3 Tone-Mapping in a Radiosity Algorithm

In the perceptually-driven radiosity algorithms that will be discussed in this chapter, tone-mapping operators are used to perform any radiosity-based decisions not on the real world values, but rather on the displayed ones. This makes it possible to account only for those changes that are going to be visible when the radiosity

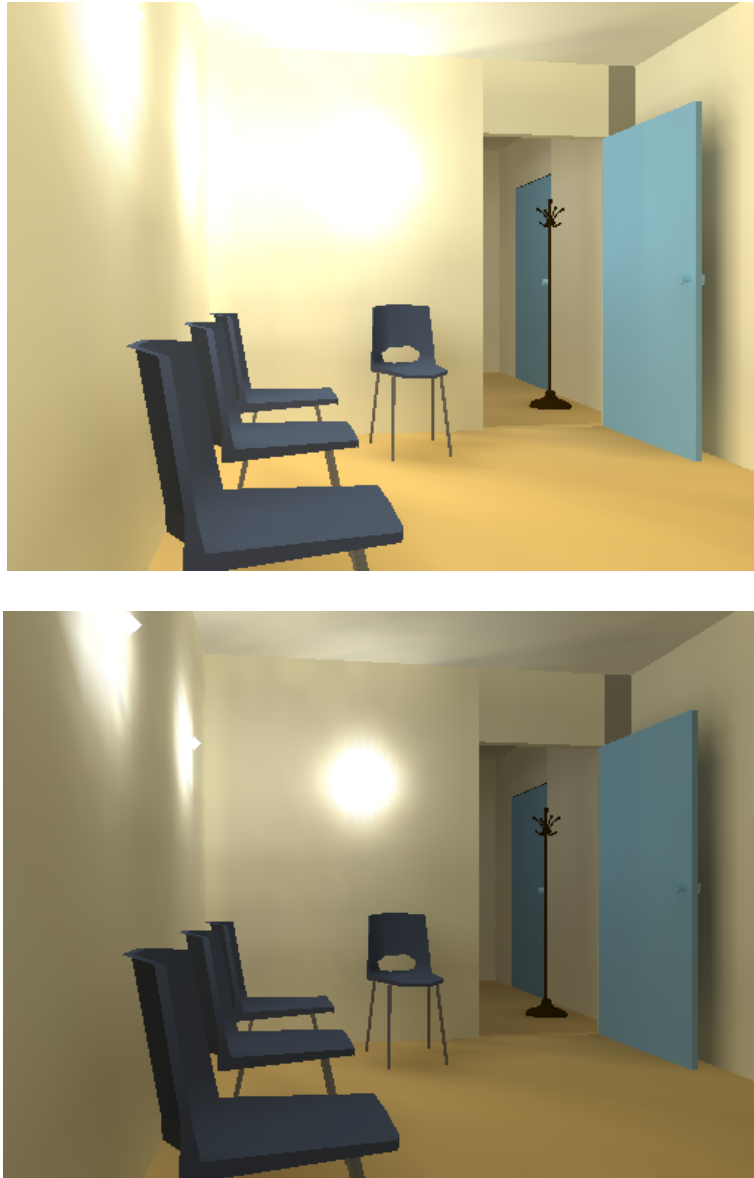


Figure 6.1: Hospital scene displayed using different adaptation estimations. Top: Average ambient term according to Gibson and Hubbard [35]. Bottom: The adaptation estimate computed directly from the image data.

solution will be displayed. A more detailed overview of this field may be found, for example, in our report [84].

In Section 3.2 we noted that the physical accuracy in rendering itself does not yet guarantee that the displayed image will have a realistic visual appearance. Having computed an image containing luminance values that appear in the real world, the last step of the image synthesis process is to use a tone-mapping operator to display this image as perceptually correct as possible using a device that is unable to reproduce the broad real world luminance scale.

If a tone-mapping operator is to be embedded into a radiosity algorithm, a problem mentioned in Section 6.1.2 arises: In order to perform the tone-mapping, we have to determine the luminance in the modelled environment that the observer's visual system will adapt to. This real world adaptation luminance L_{wa} influences what the observer will actually see in the scene. The usual approach to a priori estimation of the real world adaptation luminance is that of Gibson and Hubbard [35] mentioned in previous section.

Modern tone-mapping approaches are rather complex and usually process the whole image at once in order to better model the behaviour of the human visual system. This implies that not every tone-mapping operator is suitable for embedding directly into a radiosity method — only the mappings that are able of independent processing of luminance values (Tumblin and Rushmeier [119], Ward [126], and Tumblin's revised mapping [117]) can be used directly, whereas the model of Ferwerda *et al.* [30] can be used only partially, as the color discrimination process consists in fact of low-pass image filtering. In all cases, the tone-mapping algorithms embedded into radiosity methods ignore any issues related to gamut mapping.

6.2 State of the Art

If we exclude importance-driven radiosity methods that are sometimes classified as perceptually-based, applications of human perception-aware techniques to the radiosity systems are still rather rare. Nevertheless, several papers have appeared recently that use human perception-aware techniques to guide the computations. The techniques described in them can be subdivided into two groups.

The first group of perceptually driven radiosity approaches works directly in the *object space*. The algorithms belonging to this group estimate the perceptual transformation from luminance to display intensity before the radiosity simulation starts. This makes it possible to use perceptually driven oracle functions in hierarchical radiosity or to cull perceptually unimportant discontinuity lines (Myszkowski *et al.* [66], Gibson and Hubbard [35], Hedley *et al.* [45]).

The second group of algorithms works in the *image space*. The two algorithms

presented in this group compare some features present in images resulting from the radiosity simulation. This information is then used to drive the hierarchical refinement of patches or to estimate the convergence of the method (Martin *et al.* [64], Myszkowski [65]).

The rest of this chapter is devoted to a more detailed overview of these two groups of approaches.

6.3 Object-Space methods

6.3.1 Perceptually-Driven Element Refinement

Myszkowski *et al.* [66] used variable luminance thresholds in their refinement oracle to decrease the number of generated mesh elements while maintaining the same visual quality of the solution. The refinement criterion presented in the paper tested luminance at several points along the edges of an element. In cases when the luminance difference at these points was found to be higher than a threshold chosen in accordance to the Stevens' brightness function [111], the element was subdivided using a special non-uniform subdivision scheme. As a result, total time of the lighting simulation was shortened, and subtle shading details could be revealed where other algorithms fail.

6.3.2 Perceptually-Driven Hierarchical Radiosity

Gibson and Hubbard [35] present an improvement to the hierarchical radiosity method that uses perceptually based measures to control the generation of view independent radiosity solutions.

In their paper they describe a new oracle that stops patch refinement once the difference between successive levels of elements becomes perceptually unnoticeable. The authors also show how the perceived importance of any potential shadow falling across a receiving element can be determined. This importance is then used to control the number of rays that are cast during visibility computations and allows to significantly reduce the total number of rays required for a form factor solution without significant loss in visible image quality. They also discuss how a method similar to the refinement oracle can be used to join elements of the radiosity mesh when the differences between elements at successive levels become unnoticeable.

The improvements are achieved using an a priori estimate of the real world adaptation luminance, making it possible to incorporate a tone-mapping algorithm into the simulation process. Then, using Tumblin-Rushmeier's tone reproduction

operator [119] in perceptually uniform CIELUV colour space to transform luminance values to display colour values, those aspects of the displayed solution that are visually important can be determined. As the above mentioned model does not perform well under scotopic and mesopic conditions — which are actually the viewing conditions in usual computer laboratories — it would be desirable to use the more general tone reproduction operators (such as that of Ferwerda *et al.* [30] or Larson *et al.* [57]) to represent images at low levels of illumination more accurately. However, as we mentioned in Section 6.1.3, these models require typically information about the whole image that is being generated and are therefore unsuitable for view-independent radiosity solutions.

In the following paragraphs we will discuss the three perceptually driven parts of Gibson and Hubbard’s algorithm in more detail.

Adaptive Refinement

The underlying hierarchical radiosity method used by the authors works with CIEXYZ colour values. However, this colour space is not perceptually uniform — colours which have the same numerical distance do not have to be perceptually equidistant. This means that the comparison of two colour differences in the CIEXYZ colour space may suggest both differences are the same, even if one would be perceived as smaller than the other. For performing colour comparisons, it is therefore better to transform the colour values from CIEXYZ to CIELAB or CIELUV colour spaces, which are more perceptually uniform (CIELUV has a nonuniformity ratio of approximately 6:1, but the nonuniformity ratio of CIEXYZ space is 80:1).

For an adaptive refinement test, one needs to construct a numerical measure that bounds the perceived difference in intensity between successive levels of element subdivision. There are two different cases that have to be distinguished: a receiving element being fully or partially visible to a source element.

When a receiving element is *fully visible* to a planar convex light source, the luminance function over that element will be unimodal and continuous in all derivatives. Consider a point ν on a receiving element e , with luminance L_ν and diffuse reflectivity ρ_e , and a source patch s with luminance L_s . The perceived colour at ν after the radiosity from s has been transferred will be [35]

$$C_\nu = \zeta(L_\nu + \rho_e [F_{\nu s} L_s + L_{\text{amb}}]), \quad (6.1)$$

where $F_{\nu s}$ is the unoccluded point to area form factor between ν and s , and ζ is the tone-mapping function. The inclusion of the ambient correction term L_{amb} is necessary to account for the effect of later iterations — when not used, the first few iterations would be transferring light to completely dark receivers, and the perceptual importance of the transfers would be artificially high, which would

trigger excessive refinement. The ambient luminance L_{amb} is used during element refinement and shadow testing and it is recalculated after every iteration.

The subdivision oracle for a triangle patch and linear basis functions can evaluate (6.1) at the element vertices and edge midpoints to obtain a tone-mapped colour C_v for six points v , $v \in \{A, B, \dots, F\}$ of the triangle (see Figure 6.2). If the linear element at level i was rendered (i.e. triangle ABC) instead of four linear elements at level $i + 1$ (i.e. triangles ADF , BED , CFE , and DEF), the colours C_{iD} , C_{iE} , and C_{iF} at positions D , E and F would be equivalent to linearly interpolating the corner luminances

$$C_{iD} = \frac{C_A + C_B}{2}, \quad (6.2)$$

$$C_{iE} = \frac{C_B + C_C}{2}, \quad (6.3)$$

$$C_{iF} = \frac{C_A + C_C}{2}. \quad (6.4)$$

If C_D , C_{iD} , C_E , C_{iE} , C_F , and C_{iF} are represented in CIELUV colour space, we can calculate the perceived difference ΔE_{uv}^* for each colour pair a and b as

$$\Delta E_{uv}^* = \sqrt{(L_a^* - L_b^*)^2 + (u_a^* - u_b^*)^2 + (v_a^* - v_b^*)^2}. \quad (6.5)$$

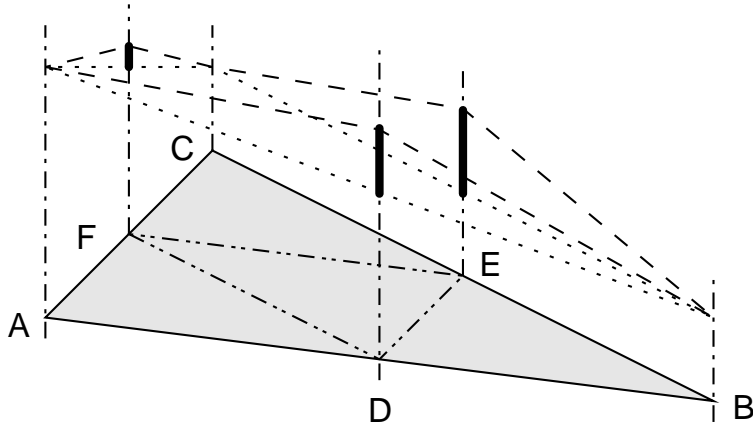


Figure 6.2: Determining the perceived difference in luminance between successive levels of element refinement. The dotted line shows the luminance representation for the parent element, the dashed line shows the luminance represented by the successive level of elements. After Gibson and Hubbard [35].

The largest of the six ΔE_{uv}^* values gives a numerical bound on the perceived difference in luminance obtained with and without the extra subdivision. This

bound is then tested against a refinement threshold, and the element is subdivided if the threshold is exceeded.

Selecting an appropriate subdivision threshold for a hierarchical radiosity algorithm is a difficult task. The perceptual radiosity method allows for easy selection of the threshold by setting it equal to the just noticeable difference in perceived luminance, i.e. setting it to the minimum value of ΔE_{uv}^* that will be visible to the user at the given adaptation level. The authors report a threshold of $\Delta E_{uv}^* = 2$ suited the computation well. As $\Delta E_{uv}^* = 100$ for a difference between reference black and reference white in the CIELUV colour system, this value corresponds to the commonly used 2% threshold visibility level.

If the source element is only *partially visible* to the receiver, the luminance function over the receiving element can exhibit many changes in continuity. In addition to this, calculating the exact portion of the source that is visible from the receiving element is an expensive operation. Since the method employs a ray cast estimate of visibility (see below), the refinement test for these situations needs to be more conservative than that used when the source is totally visible to the receiver. Once partial visibility is detected the perceptual difference between the vertex m of the receiver with minimum luminance

$$C_m = \zeta(L_m + \rho_e L_{amb}) \quad (6.6)$$

and the vertex n with the maximum luminance plus the radiosity transferred from the source patch s

$$C_n = \zeta(L_n + \rho_e [F_{ns} L_s + L_{amb}]) \quad (6.7)$$

is computed. The notation is the same as for (6.1). This assumes that vertex m is occluded from the source and hence does not receive any radiosity, and vertex n is totally visible. Then, both C_m and C_n are transformed to the CIELUV colour space, and ΔE_{uv}^* is calculated and tested against the subdivision threshold.

Shadow Testing

When testing for occlusion between a source and a receiver using ray casting, great savings in computation time can be made by identifying the visual effect any single source to receiver interaction has, and using that to reduce the number of rays where possible.

Computing patch-to-patch form factors using Monte Carlo integration [123] involves casting a number of rays between the two patches, intersecting each ray with any potentially blocking objects, and counting the number of rays that are found to connect unoccluded parts of the patches. Visibility detection is often the most computationally intensive phase of the form factor computation. Reducing the number of rays would reduce the computation time, but it can also introduce

unpleasant artifacts in the radiosity solution, such as aliasing or noise caused by point sampling errors. If the perceptual impact of light being transferred from the source patch to the receiving one could be determined, also the perceptual impact of any shadow cast over the receiver could be determined. This information can then be used to select an appropriate number of rays to use for the visibility testing. Shadows that are not likely to be noticed could be tested for very cheaply, and more effort focused towards those that are more visible.

Gibson and Hubbard use the colours at a vertex before and after light from a source is considered, C_{old} and C_{new} , to estimate the importance of a possible shadow:

$$C_{\text{old}} = \zeta(L_e + \rho_e L_{\text{amb}}), \quad (6.8)$$

$$C_{\text{new}} = \zeta(L_e + \rho_e [F_{vs} V_{vs} L_s + L_{\text{amb}}]). \quad (6.9)$$

The notation is the same as for (6.1). The results have to be converted to CIELUV colour space in order to determine the perceived difference, ΔE_{uv}^* between these quantities. This difference gives a measure of the perceptual importance of any potential shadow boundary caused by the light from the source arriving at the vertex. Also in this case the ambient correction has to be included in the transformation in order to estimate the importance of the shadow taking in account the radiosity received in later iterations.

Having obtained the value of ΔE_{uv}^* , the next step is to determine the number of rays that shall be cast between the elements in the course of the form factor computation. The method described by Gibson and Hubbard works with lower and upper error thresholds, $\Delta E_{\text{uv},\text{min}}^*$ and $\Delta E_{\text{uv},\text{max}}^*$, that again depend on just noticeable difference. Provided $\Delta E_{\text{uv}}^* < \Delta E_{\text{uv},\text{min}}^*$, the test for shadows consists of casting a single ray between the receiver and source. For changes above the upper threshold $\Delta E_{\text{uv},\text{max}}^*$, a user specified number of rays, N , is used for shadow testing. For changes inbetween, the number of rays linearly dependent on ΔE_{uv}^* given by

$$1 + \left\lceil (N - 1) \cdot \frac{\Delta E_{\text{uv}}^* - \Delta E_{\text{uv},\text{min}}^*}{\Delta E_{\text{uv},\text{max}}^* - \Delta E_{\text{uv},\text{min}}^*} \right\rceil \quad (6.10)$$

is used.

This model of shadow importance can also be improved, as it does not take in account that the just noticeable differences vary according to adaptation level and viewer distance [22, 30].

All visibility computations were accelerated using shaft culling [40]. As the number of iterations increases and the distribution of light approaches convergence, the visual importance of each transfer is reduced and perceptually adaptive visibility becomes even more beneficial.

Mesh Optimization

The final application of perceptual awareness discussed by Gibson and Hubbold is the process of mesh optimization. They noticed that the resulting mesh from a progressive radiosity solution will be far from optimal, since element refinement triggered by some source may become unnecessary when also the light arriving at the receiving element during later iterations contributed to its radiosity. The authors show how perceptual measures can be used to detect where mesh refinement has become unnecessary, and how to apply this knowledge either as a post-process, or during the course of the radiosity computation in order to conserve memory occupied by the radiosity mesh.

In order to reduce the number of elements in a mesh a posteriori, one can compare the tone-mapped intensity representations of the parent and children elements in the hierarchy in the same way as it was done during adaptive refinement in cases of full visibility. When the difference ΔE_{uv}^* between the two representations is less than the *just noticeable difference*, the leaf elements can be removed without visual impact on the image.

The authors state that they have found it beneficial to perform this optimization after every primary light source patch, except the first one, has distributed its light, and then once again at the end of the simulation.

When a final gather is being used to render a high quality image, similar techniques to those proposed here could be used to reduce the number of shadow rays required to recompute visibility terms.

6.3.3 Discontinuity Culling

In order to improve the visual quality of the radiosity simulation, discontinuity meshing is often used to improve the meshing quality in areas where discontinuities in the radiosity distribution occur [44, 60]. However, this method usually produces a too dense mesh of elements, which slows down further computations and may induce various numerical problems. Diverse discontinuity culling techniques are therefore used that make it possible to discard those discontinuity mesh edges that do not improve the solution very much. Gibson and Hubbold [35] proposed using the knowledge about the perceived importance of light from a source element that arrives at a receiver to decide whether or not the element should be subdivided along the discontinuity lines.

At the same time, Hedley *et al.* [45] proposed a new approach to culling the discontinuity lines. They use a perception-based metric to determine which discontinuities are important and which can be safely ignored. They noted that even if the discontinuity line itself may not necessarily be visible, it may exert a perceptible influence on the mesh due to improving the general triangulation or stopping

artifacts like shadow or light leaks.

In their approach they sample the illumination information along the discontinuity line and also in a small distance at either side of the discontinuity. They record radiance values before and after the current light source has contributed energy at several pairs of points lying aside the discontinuity line. Then, for every sample point they compute differences between tone-mapped colour values in the perceptually uniform CIELUV colour space. If they find such a sample point where the difference exceeds a given threshold, they assume the discontinuity line makes visible difference to the mesh and therefore it should be included. Other discontinuity lines are then culled and the algorithm proceeds with the next shooting patch.

The perceptual metric for discontinuity culling presented in the paper uses either Tumblin-Rushmeier's operator [119] or the linear mapping introduced by Ward [126]. The world adaptation level L_{wa} is computed either with an approach similar to that of Gibson and Hubbard [35] (see also Section 6.3.2) or — for non-closed environments, where the energy loss during the radiosity computations would result in an overestimate of the world adaptation level — it can be specified manually.

6.4 Screen-Space Methods

6.4.1 Image-Space Refinement Criterion

Martin *et al.* [64] present a refinement criterion for the hierarchical radiosity method, which tries to improve image quality taking into account the smoothness of the solution based on pixel intensity values instead of energy ones, and visibility changes along the surfaces for high gradient detection (sharp shadows).

Similarly to the observations of Gibson and Hubbard, the authors noticed that most existing oracle functions are based on the computation of radiometric magnitudes in object space — such as form-factors and energy values. These oracle functions do not take the image space features such as the pixel intensities into account. The observations suggest that some refinements in the energy transfer will have a little noticeable effect on the result of the radiosity simulation. Martin *et al.* therefore propose a refinement criterion that progressively improves image quality working directly on the rendered Gouraud shaded image displaying the result of the radiosity computation.

Martin *et al.* propose two different oracles for receiver and shooter refinement. We will discuss them in more detail in the following two paragraphs.

Receiver Oracle

The receiver oracle takes into account the smoothness of the Gouraud approximation and the possible visibility artifacts. Beside the shading constraint, the authors also assume that the radiosity system works with triangular patches.

In such a case, the luminance value L at a given pixel X is obtained by linear interpolation of the triangle vertex luminances for the triangle that covers the pixel X . The vertex intensity $I_i^k(\mathbf{T})$ for the k -th basis function and some vertex \mathbf{T} that belongs to patch i is computed by applying a tone-mapping function ζ to the luminance value $L_i^k(\mathbf{T})$ at the vertex:

$$I_i^k(\mathbf{T}) = \zeta(L_i^k(\mathbf{T})). \quad (6.11)$$

In the paper, the tone-mapping operator of Tumblin and Rushmeier [119] has been used.

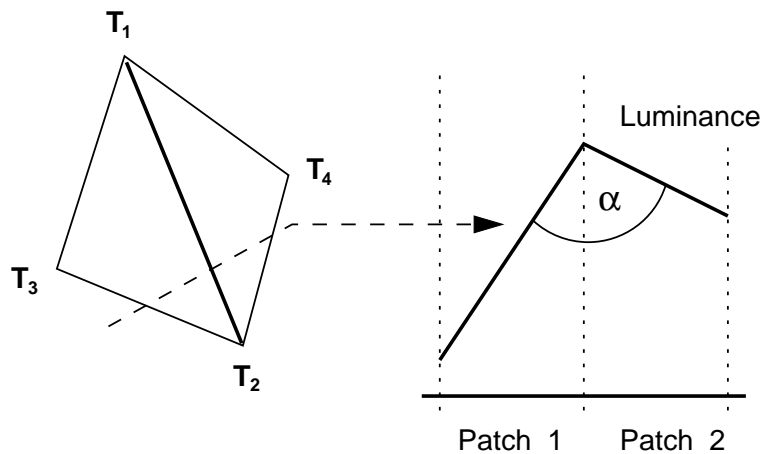


Figure 6.3: The slope difference. The angle α measures the discontinuity in Gouraud shading. After Martin *et al.* [64].

The subdivision oracle supposes that the transition in the Gouraud shading at the shared edge between two neighbouring patches of a same surface is continuous. This means that the change in the intensity along a scan line across two coplanar adjacent patches is expected to be the same at the right side and at the left side of the separating edge. As the reader can see in Figure 6.3, the angle including the two linear intensity distributions at any point on the separating edge can be used as a measure of the shading discontinuity. If the surface is curved, the angle value depends on the geometrical angle containing both patches as well.

As the proposed subdivision criterion is based on continuity in Gouraud shading, another question appears: Is it possible that Gouraud shading across two

adjacent mesh elements is continuous also in cases when the receiver should be subdivided? Unfortunately the general answer is yes — this can really happen. The possible cases are depicted in Figure 6.4.

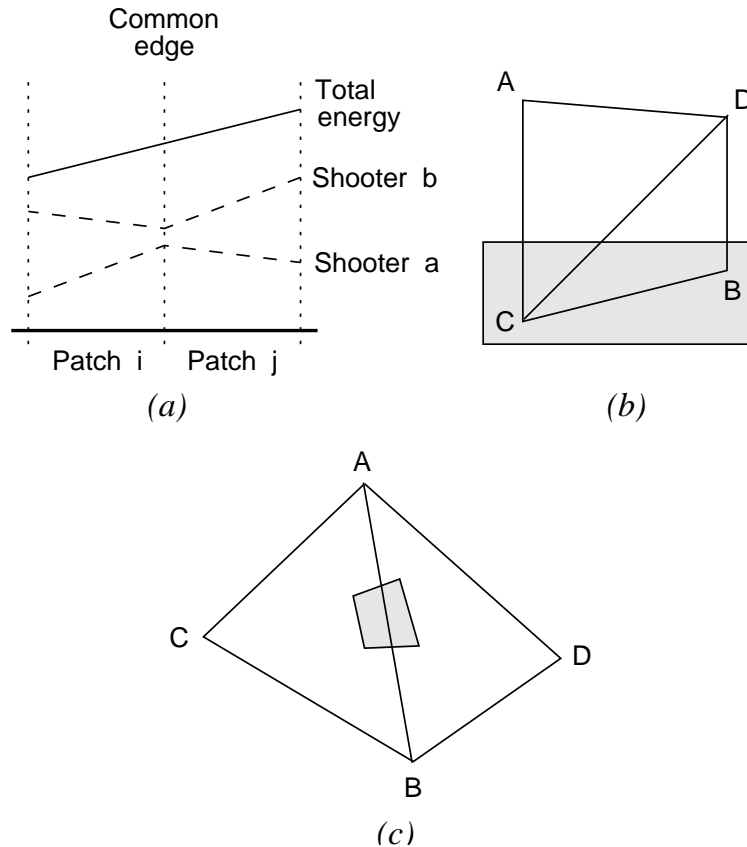


Figure 6.4: Different cases of Gouraud shading failing to identify of refinement candidates. After Martin *et al.* [64].

The first difficulty is that many different patches acting as shooters contribute to a given radiosity value. These contributions may compensate each other when they are added up. Thus, nonlinear contributions of each patch separately may give a globally linear radiosity distribution. This is illustrated in Figure 6.4(a), where two shooters *a* and *b* are contributing to receiver patches *i* and *j*. The radiosity function resulting from adding the two shooter contributions seem to be continuous across the edge separating the two receivers. However, each isolated shooter contribution shows a high slope change across the shared edge. The contributions are opposite in sign though, and they compensate each other when they are added. Therefore, the oracle has to test the interpolation continuity for each shooter contribution separately. As a consequence, not all shooters interacting

with a given patch may cause discontinuity and thus the receiver patch refinement should be realized only for the shooting patches which cause it.

A second aspect to be considered is illustrated in Figure 6.4(b): A shadow is projected over two adjacent patches in such a way that vertices A and D are outside the shadow area and vertices B and C are inside. If the radiosities at A and D are approximately the same and the radiosities at B and C also happen to be nearly the same, the Gouraud shading will show a smooth transition between the patches and fail in detecting that they should be subdivided. Fortunately, this case can easily be detected, as it corresponds to a discontinuity in the visibility along the edge. Therefore, the oracle has to take into account the difference of visibility of the edge vertices towards the shooting patches leading possibly to a subdivision.

Finally, as with other radiosity methods, the initial mesh used by the method of Martin *et al.* must have a minimum density in order to avoid precision cases such as illustrated in Figure 6.4(c). The authors state that the initial mesh must be at least so fine that the maximum patch size is less than the minimum projection size of the smallest object in any scene surface. In practice, however, the initial meshing size will be fixed to some uniform value.

Taking the above mentioned cases into account, the oracle function proposed by Martin *et al.* tests the Gouraud shading continuity at the edges between adjacent receiving elements and checks the visibility of the edge vertices towards the shooting element. The oracle decides whether it is necessary to subdivide the receiver element and if so, the radiosity algorithm subdivides it and establishes new links. As the radiosity is represented by linear basis functions, new links are created between the shooter and the new sampling points.

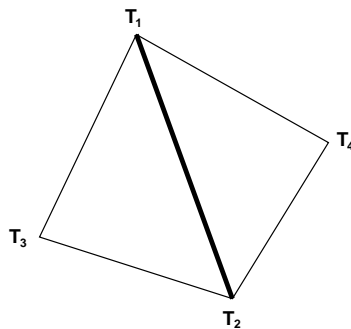


Figure 6.5: Vertex geometry for computing the slope estimator. After Martin *et al.* [64].

The following receiver oracle is proposed in [64]. The oracle implements ideas outlined above and uses view importance to influence the hierarchical refinement

process in addition. The method searches all the edges over each mesh and performs the following three steps for every shooter patch k that contributes to a pair of adjacent patches at the given hierarchy level:

1. For the edge shared by the pair of adjacent patches, the visibility difference ΔV_k between its vertices \mathbf{T}_1 and \mathbf{T}_2 and with respect to the shooter k is computed as

$$\Delta V_k = |V_{k,\mathbf{T}_1} - V_{k,\mathbf{T}_2}|, \quad (6.12)$$

where $\Delta V_k \in \langle 0, 1 \rangle$ and $\Delta V_k = 0$ if there is no change in the visibility between \mathbf{T}_1 and \mathbf{T}_2 and shooting element k . The value $\Delta V_k = 1$ indicates that one point of the pair $\mathbf{T}_1, \mathbf{T}_2$ is completely occluded from the shooter k , while the other is visible from the whole shooting element.

2. The value of V_k is tested against a given visibility threshold ϵ_{vis} in order to decide if the visibility change is small enough to consider the change of slope as a valid measure of smoothness:

- (a) If $\Delta V_k < \epsilon_{\text{vis}}$, the change in the slope itself can be used as a measure of smoothness. As this oracle is applied many times during the algorithm, an easy but non-conservative estimation of the slope change is used.

Since the method considers values in image space, the slope change is computed using pixel intensity values instead of radiosity ones. The mapping to the pixel intensities is nonlinear, which means that also the contribution of the other shooters must be taken into account in order to estimate the real change of slope produced by shooter k . The greater the energy of the other shooters the smaller the slope change.

Using the geometry shown in Figure 6.5, the slope change is computed as

$$\Delta I_k = \frac{1}{2} | \zeta(L_1 + L_{\min}) + \zeta(L_2 + L_{\min}) - \zeta(L_3 + L_{\min}) + \zeta(L_4 + L_{\min}) |, \quad (6.13)$$

where L_{\min} is the minimum luminance of the vertex luminances of the two adjacent patches L_1, \dots, L_4 without considering the contribution of the shooting patch k . Note that L_1 and L_2 are the luminances of the separating edge vertices.

- (b) If $\Delta V_k \geq \epsilon_{\text{vis}}$, a more conservative estimation of the slope change is used based on the difference between the maximum and the minimum pixel intensity values of the four vertices, again taking into account the contribution of other shooters

$$\Delta I_k = \max_{j=1,\dots,4} \zeta(L_{\min} - L_j) - \min_{j=1,\dots,4} \zeta(L_{\min} - L_j). \quad (6.14)$$

3. The estimator ΔI_k is tested against the slope threshold ϵ_{slope} , weighted by the view importances of the two adjacent patches, Y_{left} and Y_{right} .

If $\Delta I_k \cdot (Y_{\text{left}} + Y_{\text{right}}) \geq \epsilon_{\text{slope}}$, both the adjacent patches are subdivided and the links from shooter k are created on the lower level of hierarchy.

Shooter Oracle

The above described criterion relies on accurate radiosity values at the vertices. If the vertex radiosities are not computed with enough accuracy, the oracle will fail.

When a link is initially established between a shooter and a vertex (associated to a basis function at a given level), it is therefore necessary to decide if the energy transfer at that level is accurate enough to guarantee a good approximation of the radiosity value at the vertex. If the accuracy of the energy transfer is low, the link should be refined.

This leads to another oracle that decides if a link at a given level is accurate enough. The authors use an energy based oracle that compares the energy carried by the link with a given threshold. If the transferred energy exceeds the threshold, the interaction has to be refined. As the input of the oracle, the maximum of the three products of shooter coefficients by the transport coefficients, weighted by the importance of the receiver triangle, is used.

According to the paper, the current implementation results show that the above presented criteria significantly reduce the number of links that are needed for a given accuracy while keeping reasonable memory requirements of the method.

6.4.2 Applications of VDP to Radiosity

Myszkowski [65] recently studied the applicability of Daly's VDP [22] (see Section 3.3.1) to the global illumination process. His aim was to develop a suitable image-space criterion to judge about convergence of the global-illumination solution, which would help to predict when to stop the simulation process. He also experiments with the influence of texture masking on a non-uniform adaptive subdivision algorithm.

Perceptual Convergence

Using the VDP, quantitative measures of differences between two images are generated. In addition, image regions where such differences will be noticed by the human observer are identified. If the images contain variance in indirect lighting, absolute difference metrics (as MSE) would immediately report those images as

being different. The VDP responds more selectively to this phenomenon, taking into account the local contrast, spatial frequency of the variance, and visual masking.

Terminating Criteria for Image Synthesis

If the goal of the image synthesis process is the perceptual accuracy of the resulting image and one does not have to care about the actual physical accuracy, the global illumination computation should be stopped in the moment when the resulting image quality becomes indistinguishable from that of the fully converged solution measured by some norm of the radiosity values. Myszkowski states that approaches like mean-squared error norm do not suite this task very well, as they predict rather large differences between images tone-mapped with Tumblin-Rushmeier's operator. When discussing the features of this norm, he also objects that the mean-squared error predicts the global error for the whole scene while, in fact, the local error may be much higher at some places.

The author again proposes using the VDP in the image space as the termination criterion and tries to determine which images in the sequence of gradually converging images should be compared in order to reliably predict the actual convergence. In his experiments, he finds the VDP comparison between images obtained in time τ and 0.5τ of the simulation to predict the convergence well.

Adaptive Mesh Subdivision

Myszkowski remarks that research on perceptually-aware meshing strategies has up to now been almost exclusively limited to using the tone-mapping operator of Tumblin and Rushmeier [35, 45, 64] to predict the influence of subdividing the current mesh element. He therefore proposes a three-step process that uses VDP to predict the masking effect triggered after the patch has been textured. The approach is based on an older method of the author [66], which uses nonuniform adaptive subdivision and is as follows:

- Step 1:** Choose candidates according to the traditional non-perceptual approach. Generate edges.
- Step 2:** Update the lighting recomputing only the lighting situation for the candidates from step 1. Compute VDP between the image from the last iteration step and the current image.
- Step 3:** Restore edges that do not produce visible differences between both images.

The author reports that this approach works well for areas where illumination changes only gradually, but in the proximity of illumination discontinuities the

mesh reduction is rather poor.

6.5 Summary

We have seen that applications of human perception-aware techniques to the radiosity systems are still rather rare. Just recently several papers appeared that use human perception-aware techniques to guide the radiosity computations. The current perceptually-driven radiosity methods can be subdivided into two groups.

The first group of perceptually driven radiosity approaches works directly in the object space on values stored within the finite element mesh. Algorithms belonging to this group estimate the perceptual transformation from luminance to display intensity before the radiosity simulation starts. This makes it possible to use perceptually driven oracle functions in hierarchical radiosity or to cull perceptually unimportant discontinuity lines.

The second group of algorithms works in the image space. Algorithms from this group compare some features present in images resulting from the radiosity simulation. This information is then used to drive the hierarchical refinement of patches or to estimate the convergence of the method.

Chapter 7

Perceptually-Driven Termination

In the previous chapter, we learned that the perceptually-based termination criterion presented by Myszkowski requires an image to be generated and a pair of images to be evaluated after every iteration step of the radiosity algorithm, which is appropriate for view-dependent computations only. When computing a view-independent radiosity solution, Myszkowski suggests using a set of representative views and generating images corresponding to every view after every iteration. The corresponding image pairs are then evaluated in order to predict the convergence. This approach has one obvious disadvantage, mentioned by the author himself — the process of the representative views generation is not trivial. Obviously, this set should be as small as possible, but still producing enough information in order to predict the global perceptual convergence correctly. We feel that it may often be impossible to fulfill both these conditions at the same time.

In this chapter we will describe several other approaches developed by the author that do not require any images to be generated and work directly on radiosity values of a stochastic radiosity method with constant elements.

7.1 Convergence of the Stochastic Jacobi Method

There exist several theorems that explain how the error of the average of N independent samples from an estimator \hat{S} of the quantity S reduces as the number of samples N increases. The error bounds provided by these theorems are not deterministic but probabilistic. These theorems provide us with a certain level of *confidence* that the expected value $E[\hat{S}]$ does not exceed the error bound.

A strong — and the only one we are going to mention here — statement about the accuracy of a Monte Carlo computation is given by the *central limit theorem of probability*. This theorem states that the expected value $E[\hat{S}]$ taken as the average of N independent samples of any random variable \hat{S} with finite variance $\text{Var}[\hat{S}]$

is asymptotically distributed according to the normal (Gaussian) distribution with variance $\text{Var}[\hat{S}]/N$ and standard error $\sqrt{\text{Var}[\hat{S}]/N}$ as $N \rightarrow \infty$ and that in the limit for large N we have

$$P \left\{ a \sqrt{\frac{\text{Var}[\hat{S}]}{N}} \leq E[\hat{S}] - S \leq b \sqrt{\frac{\text{Var}[\hat{S}]}{N}} \right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt. \quad (7.1)$$

This relationship suggests that for an estimate computed using N samples, the value $E[\hat{S}]$ will stay within one standard error 68.3 % of the time, within two standard errors 95.4 % of the time and within three standard errors 99.7 % of the time.

Knowing that after the k -th iteration of the stochastic Jacobi method introduced in Section 4.2 every element has been computed with the error

$$\varepsilon_i \approx 3 \sqrt{\frac{\text{Var}[\hat{P}_i^{(k)}]}{N_i}},$$

such a total number of samples in one iteration N can be estimated that keeps the total error $\varepsilon \leq \varepsilon_i$ for all i with probability 99.7 %. An example evaluation of this approach is presented in Figure 7.1. This graph shows how the maximum relative luminance error in the Monte Carlo radiosity computation,

$$\varepsilon \approx \max_i \left\{ \frac{3}{L_i} \sqrt{\frac{\text{Var}[\hat{L}_i^{(k)}]}{k}} \right\},$$

evolves for the corridor scene from the RENDERPARK distribution (see Figure 7.4).

We remark that Bekaert [5, Sec. 6.4.4] proved that the above mentioned total number of samples N in a regular stochastic Jacobi iteration is asymptotically

$$N \approx \frac{9P_{\text{total}}^{(k)}}{\varepsilon^2} \cdot \max_i \left[\frac{\rho_i(B_i^{(k+1)} - E_i)}{A_i} \right]. \quad (7.2)$$

The formula (7.2) provides us with a possibly pessimistic upper bound on the number of rays that have to be shot in a single iteration. As we cannot know a priori how many rays will hit the patch in concern, we have to assume that actually each of the N samples contributes to every patch in the scene — in case the ray misses the patch the contribution would be recorded as zero. Moreover, presence of very small patches (that can be grouped into clusters) would cause the number of sample rays to reach unnecessarily high quantities.

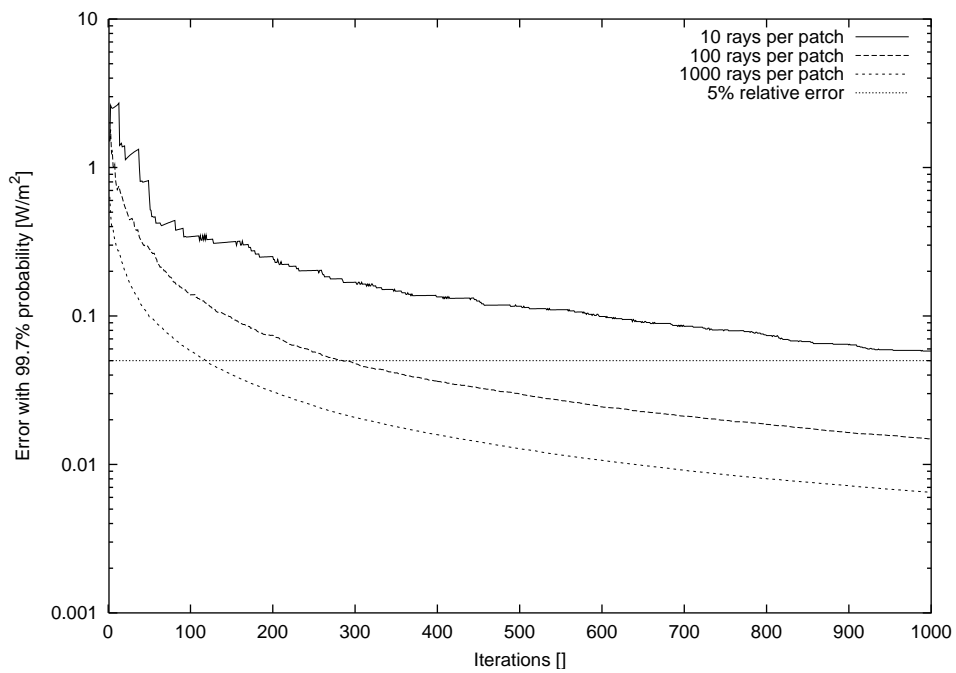


Figure 7.1: Maximum relative luminance error in a Monte Carlo radiosity simulation, estimated with 99.7% probability.

7.2 Perceptually-Driven Termination Criterion

In the previous Section we have seen that the convergence criteria based on the element power variance require a significant amount of work before one can be reasonably sure that the computation has converged. In this Section we will present several possibilities for improving the convergence detection using perceptually-based approaches.

7.2.1 Confidence Intervals

The power carried by a patch i that is being computed during the process of a Monte Carlo simulation is an average value of a certain number of power samples arriving at the patch. As a result, if the number of samples N is high enough, the power of patch i will gain values that are statistically distributed according to normal (Gaussian) distribution. For such values it is quite easy to find a so-called *confidence interval*, that means such an interval around the estimated value that — with the given probability (confidence) — contains all successive power values. The size of this interval depends on the variance of the N power samples.

A similar approach has been used by Bekaert to derive (7.2). As we will see, the main differences in our case are the use of a sample-based variance estimate and more importantly the transformation of the convergence criterion into the perceptually uniform space.

Single Element

In final effect we are interested in the variance of element's brightness I_i corresponding to its power P_i . A confidence interval for power or radiosity of the i -th element can be constructed without many problems [5, Chap. 6] as closed-form expressions for the variance $\text{Var}[\hat{P}_i^{(k)}]$ for both stochastic Jacobi relaxation methods exist. Unfortunately, the transformation $I_i = \mathcal{T}(P_i)$ is nonlinear and quite complicated as it includes conversion to luminance, tone-mapping, gamma and gamut mapping, and transformation to lightness. As a result, a simple closed form solution for expressing the lightness variance $\text{Var}[I_i^{(k)}]$ is not available.

We will therefore estimate the lightness variance using the standard formula for variance,

$$\text{Var}[I_i^{(k)}] = \frac{1}{k-1} \left[\sum_{j=0}^k (I_i^{(j)})^2 - k \left(\frac{1}{k} \sum_{j=0}^k I_i^{(j)} \right)^2 \right]. \quad (7.3)$$

The value of $\text{Var}[I_i^{(k)}]$ can be computed incrementally keeping track of $m_i = m_i +$

$(I_i^{(k)})^2$ and $n_i = n_i + I_i^{(k)}$ for every patch:

$$\text{Var}[I_i^{(k)}] = \frac{1}{k-1} \left[m^{(k)} - \frac{1}{k} (n^{(k)})^2 \right]. \quad (7.4)$$

Having estimated the lightness I_i of patch i with variance $\text{Var}[I_i^{(k)}]$, the confidence $C_i^{(k)}$ that

$$I_i \in \left[I_i^{(k)} - n \sqrt{\frac{\text{Var}[I_i^{(k)}]}{k}}, I_i^{(k)} + n \sqrt{\frac{\text{Var}[I_i^{(k)}]}{k}} \right] \quad (7.5)$$

is given by

$$C_i^{(k)} = \text{erf} \left(\frac{n}{\sqrt{2}} \right), \quad (7.6)$$

where erf is the *error function* defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

The higher the value of n in (7.5) and (7.6), the higher confidence $C_i^{(k)}$ we have that the value of I_i would really stay within the interval defined by (7.5).

Let us now determine the value of n that would assure us that the future changes in brightness I_i will be invisible. That requires these changes always being lower than the just noticeable difference Δ_{JND} :

$$I_i \in [I_i^{(k)} - \Delta_{\text{JND}}, I_i^{(k)} + \Delta_{\text{JND}}]. \quad (7.7)$$

Comparing (7.5) and (7.7) we can express n as

$$n = \Delta_{\text{JND}} \sqrt{\frac{k}{\text{Var}[I_i^{(k)}]}}, \quad (7.8)$$

and we can compute the confidence that the lightness will not visibly differ from the current computed value of $I_i^{(k)}$ using

$$C_i^{(k)} = \text{erf} \left(\Delta_{\text{JND}} \sqrt{\frac{k}{2\text{Var}[I_i^{(k)}]}} \right). \quad (7.9)$$

The value of $C_i^{(k)}$ from (7.9) tells us the probability of the perceptual convergence on the i -th element with the lightness variance $\text{Var}[I_i^{(k)}]$.

Global Confidence

The minimum confidence over the whole scene,

$$C_{\min}^{(k)} = \min_i \{C_i^{(k)}\} = \operatorname{erf} \left(\Delta_{\text{JND}} \sqrt{\frac{k}{2 \max_i \{\operatorname{Var}[I_i^{(k)}]\}}} \right), \quad (7.10)$$

can be then used as a convergence criterion: we can stop the computation as soon as the worst convergence confidence $C_{\min}^{(k)}$ has reached some threshold confidence level. Using the threshold confidence C_T we can define an upper bound on unnoticeable variance for given Δ_{JND} . Using (7.9) we get

$$\frac{\operatorname{Var}[I_i^{(k)}]}{k} \leq \frac{\Delta_{\text{JND}}^2}{2 [\operatorname{erf}^{-1}(C_T)]^2} = \frac{\Delta_{\text{JND}}^2}{\kappa}, \quad (7.11)$$

where $\kappa = 2 [\operatorname{erf}^{-1}(C_T)]^2$. A typical example how the minimum confidence evolved during the course of a Monte Carlo radiosity simulation is shown in Figure 7.2. Values of κ for different confidence levels are listed in Table 7.1.

C_T	κ
0.9000	2.70554
0.9500	3.84146
0.9900	6.63490
0.9950	7.87944
0.9990	10.8276
0.9999	15.1367

Table 7.1: Values of factor κ from (7.11) corresponding to different threshold confidence levels C_T .

Termination

Inequality (7.11) can be used as a termination criterion: If for every element i of the scene mesh after some iteration k the values of $\operatorname{Var}[I_i^{(k)}]$ are below $\Delta_{\text{JND}}^2/\kappa$, the computation has converged with probability given by the confidence C . The resulting algorithm is outlined in Algorithm 5. Figure 7.3 shows different confidence-based variance limits for the “corridor” scene from the RENDERPARK distribution (see Figure 7.4).

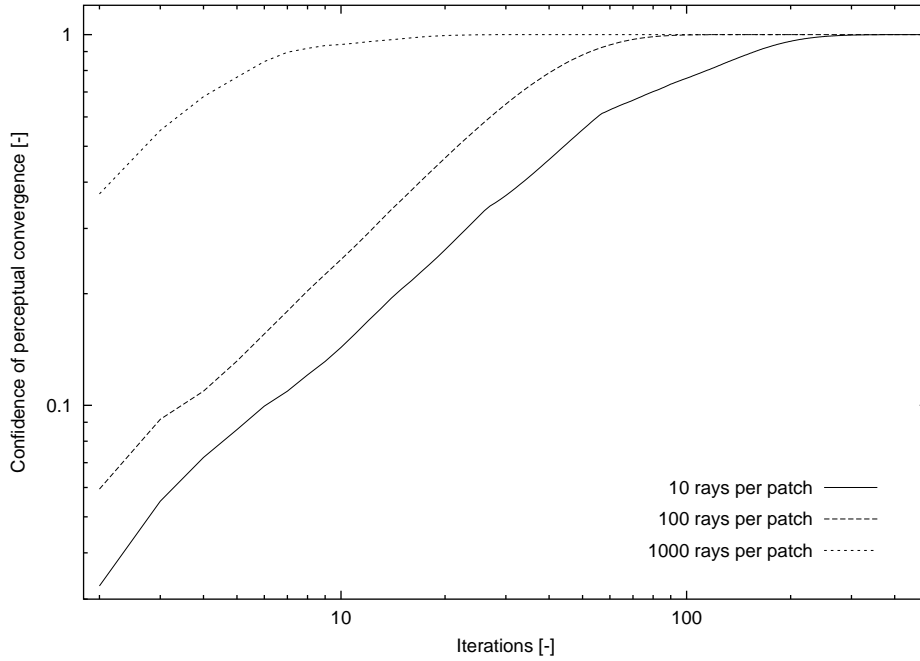


Figure 7.2: Confidence of the perceptual convergence for a Monte Carlo radiosity simulation of the corridor scene in Figure 7.4. Note the logarithmic scale on both axes.

Algorithm 5 Confidence-based perceptual termination.

initialise radiosity computation ($\text{Var}[I_i] \leftarrow 0$, $m_i \leftarrow 0$, $n_i \leftarrow 0$)

repeat

 perform single radiosity iteration

 converged \leftarrow true

for all $i \in \text{element_indices}$ **do**

$$m_i \leftarrow m_i + \left(I_i^{(k)} \right)^2$$

$$n_i \leftarrow n_i + I_i^{(k)}$$

 compute $\text{Var}[I_i^{(k)}]$ using (7.4)

if $\text{Var}[I_i^{(k)}]/k > \Delta_{\text{JND}}^2/\kappa$ **then**

 converged \leftarrow false

end if

end for

until converged = true

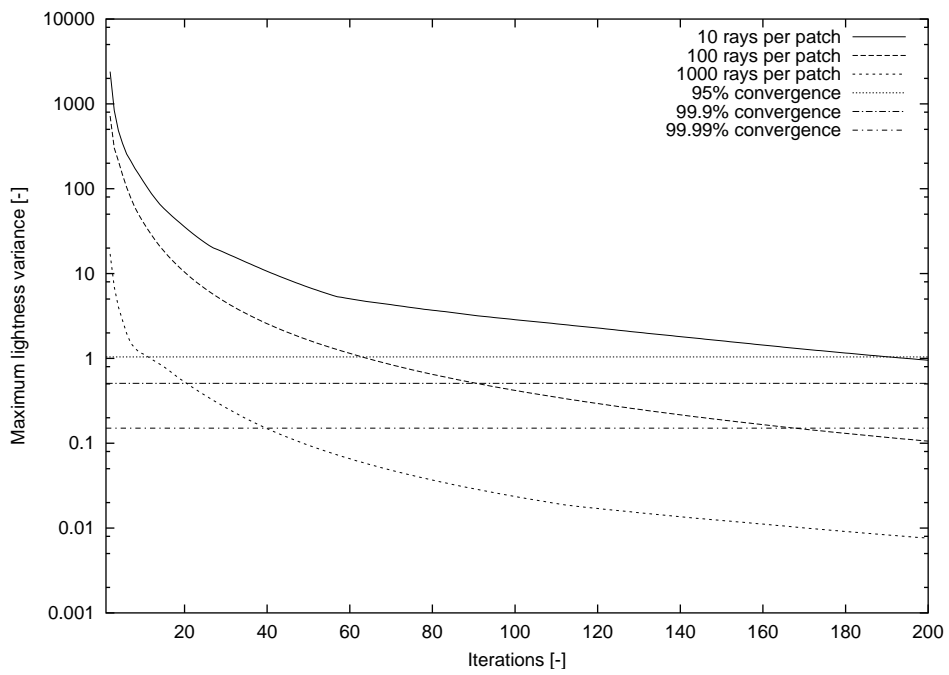


Figure 7.3: Confidence-based limits of lightness variance and values of the maximum lightness variance recorded for the scene from Figure 7.4.

Hierarchical Refinement

Until now we have not considered the case of hierarchical subdivision of scene elements. In a hierarchical algorithm, $\text{Var}[I_i]$ shall be pushed to children elements in the moment when the parent is subdivided — the variance of subelements in that moment is the same as the variance of the parent element. This way we can maintain the correspondence between the variance of a parent element and variance of its children.

When working with hierarchy, the termination condition (7.11) will be tested only on leaf elements of the hierarchy, as only these elements are used for displaying the scene. That also implies that it suffices to maintain the correct value of $\text{Var}[I_i]$ in hierarchy leaves only.

7.2.2 1/k-Heuristics

During our experiments with the hierarchical well-distributed ray set radiosity algorithm [6] we have noticed that the maximum change in display luminances behaves as $O(1/k)$, where k is the number of iterations (in our case, one *iteration* means shooting the power from the set of all patches in the scene). This led us to the following heuristics.

After every iteration k we determine the maximum change of displayed luminances $\Delta E_{\text{uv,max}}^{*(k)}$ using a tone mapping operator and comparisons in CIELUV colour space as it has been discussed in Section 6.1.3. After that, the least squares functional fit to the collected data is performed in order to arrive at a function $\mathcal{F}(k)$ that predicts the maximum luminance change for given iteration k ,

$$\Delta E_{\text{uv,max}}^{*(k)} \approx \mathcal{F}(k). \quad (7.12)$$

Knowing the maximum unnoticeable luminance difference Δ_{JND} , the last iteration number k_{term} can be predicted using $\mathcal{F}^{-1}(\Delta_{\text{JND}})$. This process is written down in Algorithm 6.

Algorithm 6 Termination criterion.

```

repeat
  perform  $k$ -th iteration
  determine  $\Delta E_{\text{uv,max}}^{*(k)}$ 
  approximate  $\mathcal{F}(k)$  using the known values of  $\Delta E_{\text{uv,max}}^{*(k)}$ 
   $k_{\text{term}} \leftarrow \lceil \mathcal{F}_k^{-1}(\Delta_{\text{JND}}) \rceil$ 
until  $k > k_{\text{term}}$ 

```

The values of $\Delta E_{\text{uv,max}}^{*(k)}$ are taken as the maxima over all elements of the changes in their displayed radiosities. We do not have to compute the vertex

radiosities in every step in this case — when the radiosity over an element is represented using a constant basis function, the radiosity change at a vertex of the element will be the average of the radiosity changes over N elements that coincide at the vertex:

$$\Delta B_v = \frac{1}{N} \sum_{i=1}^N \Delta B_i.$$

This implies that for $\Delta B_i \leq \Delta B_{\max}$ we will always have $\Delta B_v \leq \Delta B_{\max}$ and we can use the maximum element radiosity change as a possibly pessimistic prediction of the maximum vertex radiosity change.

In order to compute the actual change in displayed radiosity for the given element, tone mapped radiosities in CIEXYZ colour space at iterations k and $k - 1$ are computed first, resulting in the values that will be reproduced on the given display device. The display values are then converted to CIELUV colour space and the difference $\Delta E_{\text{uv,max}}^{*(k)}$ is obtained using the colour difference formula (2.30). Algorithm 7 describes this process in more detail.

Algorithm 7 Determining $\Delta E_{\text{uv,max}}^{*(k)}$.

```

 $\Delta E_{\text{uv,max}}^{*(k)} \Leftarrow 0$ 
for all  $i \in \text{element\_indices}$  do
   $B_{\text{d,xyz}}^{(k-1)} \Leftarrow \text{tonemap}[P_i^{(k-1)} / (\pi A_i)]$ 
   $B_{\text{d,xyz}}^{(k)} \Leftarrow \text{tonemap}[P_i^{(k)} / (\pi A_i)]$ 
   $B_{\text{d,luv}}^{(k-1)} \Leftarrow \text{xyz\_to\_luv}(B_{\text{d}}^{(k-1)})$ 
   $B_{\text{d,luv}}^{(k)} \Leftarrow \text{xyz\_to\_luv}(B_{\text{d}}^{(k)})$ 
   $\Delta E_{\text{uv}}^* \Leftarrow \text{luvdiff}(B_{\text{d,luv}}^{(k)}, B_{\text{d,luv}}^{(k-1)})$ 
   $\Delta E_{\text{uv,max}}^{*(k)} \Leftarrow \max\{\Delta E_{\text{uv,max}}^{*(k)}, \Delta E_{\text{uv}}^*\}$ 
end for

```

The function $\mathcal{F}(k)$ introduced in (7.12) is approximated with the power function of the form

$$\Delta E_{\text{uv,max}}^{*(k)} \approx \mathcal{F}(k) = a \cdot k^b, \quad (7.13)$$

where the coefficients a and b are determined using the least squares method:

$$b = \frac{k \sum_n (\ln n \ln \Delta E_{\text{uv,max}}^{*(n)}) - \sum_n \ln n \sum_n \ln \Delta E_{\text{uv,max}}^{*(n)}}{k \sum_n (\ln n)^2 - (\sum_n \ln n)^2},$$

$$a = \frac{\sum_n \ln \Delta E_{\text{uv,max}}^{*(n)} - b \sum_n \ln n}{k}.$$

All the sums are for n going from 2 to k . The sums can be updated in every iteration without need to compute them from scratch every time.

The iteration number k_{term} for which the approximated value $\Delta E_{\text{uv,max}}^{*(k)}$ is believed to stay below the given threshold Δ_{JND} can be then determined as

$$k_{\text{term}} = \left\lceil \exp \left(\frac{\ln \Delta_{\text{JND}} - \ln a}{b} \right) \right\rceil. \quad (7.14)$$

The whole process of determining k_{term} with incremental updates of the sums is described in Algorithm 8.

Algorithm 8 Determining k_{term} .

```

if  $k = 1$  then
     $s_k, s_\Delta, s_{kk}, s_{k\Delta} \leftarrow 0$ 
     $k_{\text{term}} \leftarrow \infty$ 
else
    compute  $\Delta E_{\text{uv,max}}^{*(k)}$ 
     $s_k \leftarrow s_k + \ln k$ 
     $s_\Delta \leftarrow s_\Delta + \ln \Delta E_{\text{uv,max}}^{*(k)}$ 
     $s_{kk} \leftarrow s_{kk} + (\ln k)^2$ 
     $s_{k\Delta} \leftarrow s_{k\Delta} + \ln k \cdot \ln \Delta E_{\text{uv,max}}^{*(k)}$ 
     $b \leftarrow (ks_{k\Delta} - s_k s_\Delta) / (ks_{kk} - (s_k)^2)$ 
     $a \leftarrow \exp [ (s_\Delta - bs_k) / k ]$ 
     $k_{\text{term}} \leftarrow \lceil (\exp [ (\ln \Delta_{\text{JND}} - \ln a) / b ] ) \rceil$ 
end if

```

7.3 Results

In the experiments presented in this section we used $\Delta_{\text{JND}} = 2.0$, which corresponds roughly to a just perceivable luminance difference of 2% measured in laboratory conditions. In normal environments this difference is likely to be higher. The presented results have been obtained using Ward's tone mapping operator [126]. We have found that Tumblin and Rushmeiers' operator [119] produces similar results.

We present our results for two different scenes shown in Figures 7.4 (Scene 1 in what follows) and 7.5 (Scene 2). For the both test scenes we computed three solutions with 10, 100, and 1000 rays shot from every scene patch. In both the presented cases the setup with 10 rays per patch produced solutions with high initial variance of patch radiosities, which resulted in high variance of $\Delta E_{\text{uv,max}}^{*(k)}$.



Figure 7.4: Corridor scene from the RENDERPARK distribution [7].

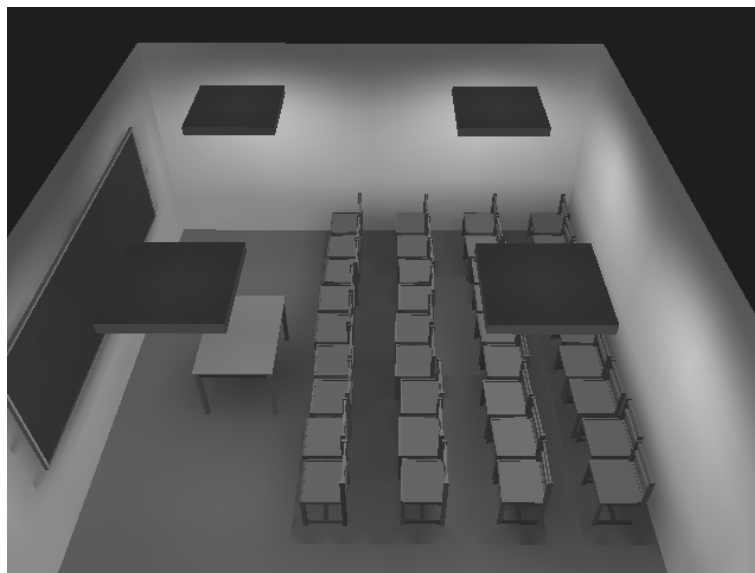


Figure 7.5: Scene #8 from Peter Shirley's package of test scenes for radiosity [97].

We recorded the values of k_{term} , $\Delta E_{\text{uv,max}}^{*(k)}$, and approximation curve coefficients during the computations. The computations were terminated when $k > k_{\text{term}}$. Table 7.2 lists the last iteration numbers for all six setups.

Rays per patch	$k_{\text{term}}(1)$	$k_{\text{term}}(2)$
10	577	118
100	91	17
1000	22	5

Table 7.2: Last iteration numbers for the two test scenes depicted in (1) Figure 7.4, and (2) Figure 7.5.

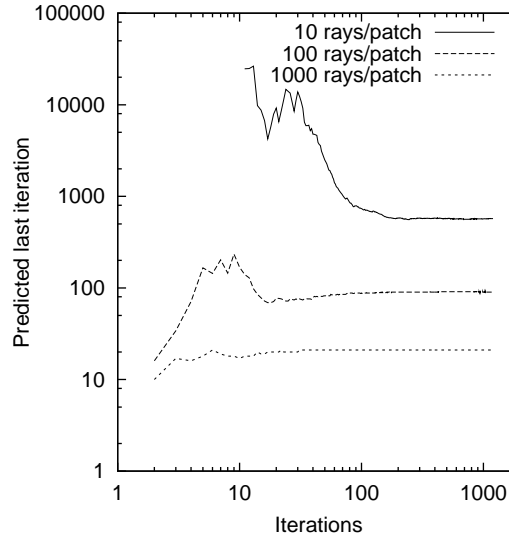


Figure 7.6: Predicted last iteration numbers k_{term} for Scene 1.

In Figures 7.6 and 7.7, the reader can see how the predicted last iteration number k_{term} evolved during the computation for a given fixed number of iterations (1200 for Scene 1, 200 for Scene 2). We can see that after having performed the terminating iteration as listed in Table 7.2, the further predictions of k_{term} do not differ much from the k_{term} value that would be used for the termination of the algorithm in the standard case. The stability of the prediction depends on the variance of the $\Delta E_{\text{uv,max}}^{*(k)}$ data, which in turn depends on the number of rays shot from every patch. Our experience shows that even if after some further iteration the value of k_{term} changes, those changes are minimal (typically up to $k_{\text{term}} - 2$) for the case of

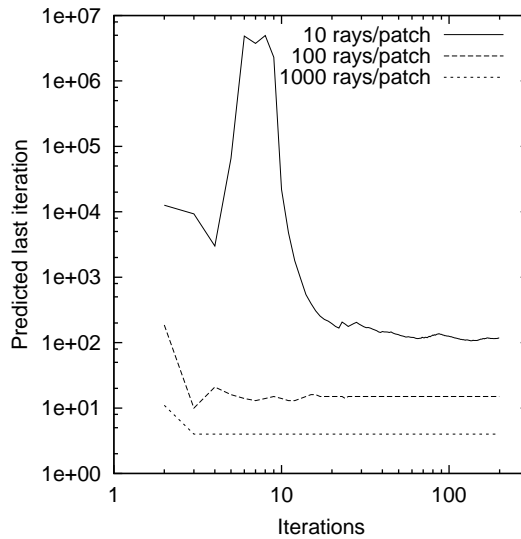


Figure 7.7: Predicted last iteration numbers k_{term} for Scene 2.

a reasonable variance. Unfortunately with the 10 rays per patch setup the prediction may rise up a bit in later iterations (for Scene 1 the highest predicted iteration number is 579 and not 577, for Scene 2 the computation should have been stopped at iteration 119 instead of 118). Figure 7.8 shows how the variance of the error data evolved during the computations of Scene 1.

Figures 7.9 and 7.10 show the measured values of $\Delta E_{\text{uv,max}}^{*(k)}$ and approximating functions $\mathcal{F}(k)$ at the moment of terminating the algorithm for all the six setups (two scenes, three different numbers of rays per patch). We can see that the quality of the prediction increases as the result variation drops off (notice the logarithmic scale on the y-axis). The prediction is quite acceptable for computations with 100 and 1000 rays per patch, but for 10 rays our method tends to overestimate the convergence speed of the algorithm.

Figure 7.11 shows the values of parameters a and b of the approximating function $\mathcal{F}(k)$ determined after every iteration k for Scene 1.

To evaluate the usability of our criterion, we performed a simple test using an approach similar to that proposed by Myszkowski [65] — after every iteration of the radiosity algorithm we generated an image of the scene in display luminances. We then compared the image produced at the k -th iteration with the image obtained at the terminating iteration, i.e. with the image that our algorithm claimed to be the perceptually converged one. Instead of using the *visible differences predictor* as Myszkowski did, we have used the perceptual colour image metrics of Neumann *et al.* [71] to evaluate the quality of the termination criterion. Two im-

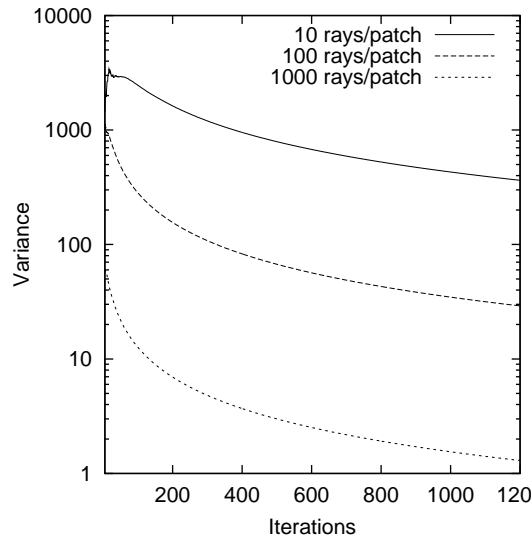


Figure 7.8: Variance of $\Delta E_{uv,\max}^{*(k)}$ for Scene 1.

ages X and Y compared by this metric will be indistinguishable, if the difference reported by the metric lays below the threshold $\delta_t = 1.0$. The metric has been computed for a CRT with diagonal 21 inches, resolution 1280×1024 pixels, and viewing distance 0.5 meter. Compared images were computed with the resolution 640×480 pixels.

The graph in Figure 7.12 shows the reported differences between images generated after every iteration and perceptually converged images for Scene 2. For the predicted last iteration number k_{term} the graph shows that although there are differences for intermediate solutions where $k > k_{\text{term}}$, the differences are always lower than the threshold δ_t and will therefore not be noticed by the human observer. The perceptually converged images have been obtained at iterations listed in Table 7.2. In the graph we can see that zero differences correspond to the images generated at iterations k_{term} . Table 7.3 lists the image differences between the reference image and perceptually converged images for Scene 2. In this case, the reference image was the most precise one obtained during our simulations — 200 iterations of stochastic Jacobi relaxation method with 1000 rays per patch. From the presented table we can see that the difference between the perceptually converged and reference image are always below the discernibility level of the human visual system.

If we confront the results presented in Figure 7.10 with the results of our comparison experiment presented in Figure 7.12, we can notice that according to the image comparison we could stop the computations much sooner than our termi-

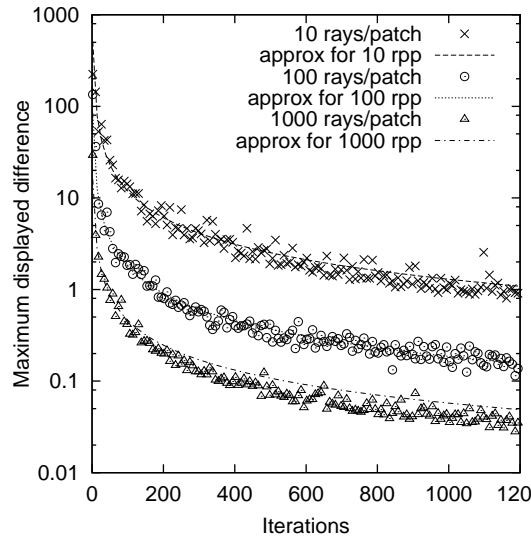


Figure 7.9: Perceptual convergence for Scene 1. Only every 5-th value of $\Delta E_{uv,max}^{*(k)}$ is plotted in order to maintain the graph readable.

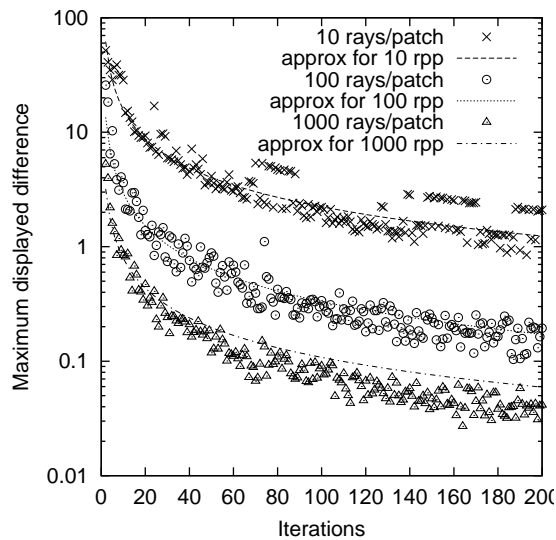


Figure 7.10: Perceptual convergence for Scene 2. Only every 5-th value of $\Delta E_{uv,max}^{*(k)}$ is plotted.

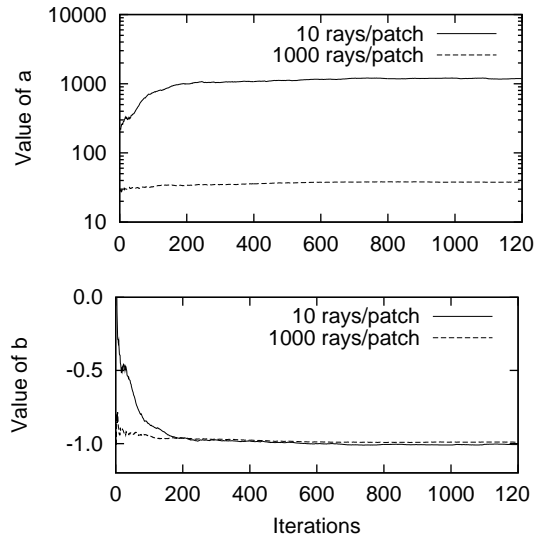


Figure 7.11: Values of regression curve parameters a and b for Scene 1.

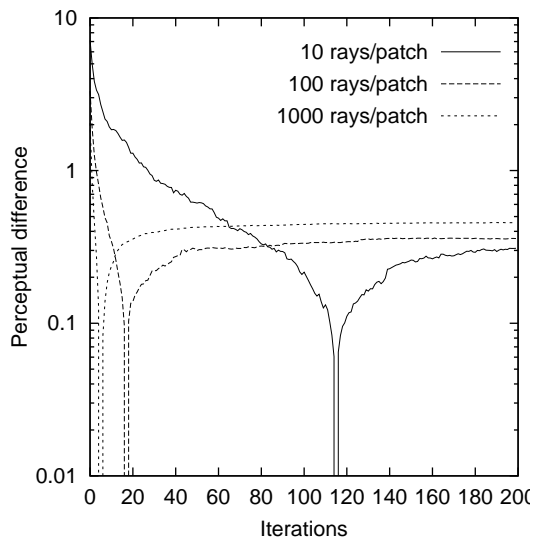


Figure 7.12: Differences between the results of particular iterations and perceptually converged results for Scene 2.

Rays per patch	$\Delta(I_{\text{perc}}, I_{\text{ref}})$
10	0.453917
100	0.370932
1000	0.455763

Table 7.3: Perceptual differences between the perceptually terminated and reference solutions.

nation criterion suggests. The reason for this phenomenon is simple: our metric evaluates the global view independent error, while the image evaluation process takes only the error in the visible part of the scene into account.

As our last result, Table 7.4 compares different termination criteria mentioned in this chapter for the “corridor” scene with a slightly coarser meshing than that in the computations compared in Table 7.2.

Rays per patch	$\epsilon = 5\%$	99.9% confidence	1/k heuristic
10	1253	334	336
100	286	107	62
1000	119	25	18

Table 7.4: Final iterations for the corridor scene (see Figure 7.4) using different termination criteria described in Sections 7.1, 7.2.1, and 7.2.2.

7.4 Summary

In this chapter we have presented a heuristic perceptually-driven termination criterion for stochastic radiosity methods. The criterion uses a perceptual tone mapping operator and the perceptually uniform CIELUV colour space to compute the maximum change in displayed luminance for the given iteration step. The iteration at which the maximum displayed luminance change lays below a given threshold is then predicted using extrapolation of a functional fit to the maximum perceptual difference data.

The criterion certainly has many drawbacks. The main of them is the fact that computations with high initial variance tend to be terminated too soon as the approximation $\mathcal{F}(k)$ tends to underestimate the value of k_{term} . One of the possible solutions to this problem is to e.g. modify the termination threshold based on the variance of the obtained error data, another possibility is to filter the $\Delta E_{\text{uv,max}}^{*(k)}$

values by selecting local maxima over several iterations in order to get an estimate of the upper bound of the error.

A possible improvement of this method would be an extension to a patch variance and covariance analysis approach, terminating the computation when the variance over elements of a patch drops below a given threshold. For hierarchical radiosity, this approach can be then improved with link variance analysis, possibly developing an algorithm for *perceptual error control* allowing for perceptual error limits similar to the radiance error limits presented by Bekaert and Willems [10] for hierarchical radiosity. A good basis for the analysis would probably be provided by the approach of Ramasubramanian *et al.* [85] who predicted visibility of noise in pictures generated by path tracing.

Instead of using the comparison algorithm of Neumann *et al.*, it could also be interesting to compare the results that we obtained with those produced by an image comparison method based on a human visual system model, e.g. with that of Daly [22].

Chapter 8

Conclusion

In this thesis we presented an overview of perceptually-driven graphics algorithms for tone-mapping, image comparison, and computing radiosity solutions. We have then concentrated on Monte Carlo radiosity methods and on their perceptually- and importance-based extensions.

The contributions of this thesis are the combination of the hierarchical refinement and the importance-driven stochastic Jacobi method described in Chapter 5, and the design and the tests of various convergence criteria for view-independent Monte Carlo radiosity methods described in Chapter 7. Besides that, our research led to the birth of perceptual tone-mapping and perceptual image comparison modules for the ART rendering package [3] and to several improvements of the RENDERPARK system [7] regarding tone-mapping and radiosity computations.

The experience with perceptually-driven radiosity methods shows that taking the behaviour of the human perception into account when designing image synthesis algorithms does not work as an implicit time-saver, although this was the hope at the beginning of our research. The reasons are twofold:

First, features of the human visual system cannot be grasped using radiometric quantities. Hence, when using even very pessimistic radiometric error criteria to drive an image synthesis algorithm, it is well possible that the perceptual error criteria would require even additional work in some cases.

Second, evaluating even simple models of the human visual system on a computer is a computationally very demanding task. Such models are typically image-based and they require either a decomposition into several different frequency layers or a transformation into the frequency space. In addition, quite computationally intensive filtering is required to obtain a meaningful result. Therefore it may often be the case that the gain of using perceptually-based error criteria would actually be eliminated by the cost of using the perceptual method itself.

In general, we feel that we have to be very careful in selecting in which case we would use a complete perceptually-based error model and when not. The

hierarchical Monte Carlo radiosity presents a remarkable example in this respect. The subdivision oracle used in this method has to be kept as simple as possible as it is evaluated upon every hit of the receiver surface — therefore only a very simple and fast visual system model can be used for the perceptually-driven refinement in such a case. The approach pioneered by Ramasubramanian *et al.* [85] may provide a lead for the future research: In order to keep the perceptual criteria efficient, one would simulate the visual phenomena just in the “real world”, effectively ignoring the costly and complicated transformation to the display device.

Possible improvements of the importance-driven stochastic Jacobi include the use of linear basis functions, a robust perceptually driven subdivision oracle, and a fast variance estimator that would be used for displaying intermediate patch hierarchy levels for interactive walkthroughs. The importance measure used by the methods is based solely on the visibility from the point of observation — we would probably profit from using a perceptually-based importance quantity instead. Such a quantity would include effects of visual masking and visual response to patch size on the solution.

A possible improvement of the termination criterion would be an extension to a patch variance and covariance analysis approach, as covariance is probably the main source of visible noise in the solution [94]. For hierarchical radiosity, this approach can then be improved with link variance analysis, possibly developing an algorithm for perceptual error control allowing for perceptual error limits similar to the radiance error limits presented by Bekaert and Willems [10] for hierarchical radiosity.

There are still many other possibilities for future improvements of perceptually-based radiosity methods. In order to improve the visual appearance of the solutions, a fast predictor of visible discontinuities in the radiosity function would be desirable. Also a noise visibility predictor based on visual masking would drastically improve the quality of textured radiosity sceneries. The current adaptation level estimator is based on an ambient term and may deliver quite incorrect results in complicated lighting situations. Using a more elaborate method for spatially-varying adaptation level estimation would certainly improve performance of any perceptually-based acceleration routine.

Bibliography

- [1] D. H. Alman, R. S. Berns, G. D. Snyder, and W. A. Larson. Performance testing of color-difference metrics using a color tolerance dataset. *COLOR research and applications*, 14:139–151, 1989.
- [2] B. Arnaldi, T. Priol, L. Renambot, and X. Pueyo. Visibility masks for solving complex radiosity computations on multiprocessors. *Parallel Computing*, 23(7):887–897, July 1997.
- [3] Advanced rendering toolkit. Available from <http://www.cg.tuwien.ac.at/research/rendering/ART/>.
- [4] C. J. Bartleson and F. Grum, editors. *Visual Measurements*, volume 5 of *Optical Radiation Measurements*. Academic Press, 1984.
- [5] P. Bekaert. *Hierarchical and Stochastic Algorithms for Radiosity*. PhD thesis, Katholieke Universiteit Leuven, Dec. 1999.
- [6] P. Bekaert, L. Neumann, A. Neumann, M. Sbert, and Y. D. Willems. Hierarchical monte carlo radiosity. In Drettakis and Max [25], pages 259–268.
- [7] P. Bekaert, F. Suykens, P. Dutré, and J. Přikryl. RenderPark - a photorealistic rendering tool. Available from <http://www.cs.kuleuven.ac.be/~graphics/RENDERPARK/>.
- [8] P. Bekaert and Y. D. Willems. A progressive importance-driven rendering algorithm. In E. Ružický, P. Eliáš, and A. Ferko, editors, *Proceedings of the Tenth Spring School on Computer Graphics '94*, pages 58–67, Comenius University, Bratislava, Slovakia, June 1994.
- [9] P. Bekaert and Y. D. Willems. Importance-driven progressive refinement radiosity. In Hanrahan and Purgathofer [42], pages 316–325.
- [10] P. Bekaert and Y. D. Willems. Error control for radiosity. In Pueyo and Schröder [81], pages 153–164.

- [11] P. Bekaert and Y. D. Willems. HIRAD: A hierarchical higher order radiosity implementation. In Purgathofer [82], pages 213–227.
- [12] O. M. Blackwell and H. R. Blackwell. Visual performance data for 156 normal observers of various ages. *Journal of Illumination Engineering Society*, 1:3–13, 1971.
- [13] C.-A. Bohn and R. Garmann. A parallel approach to hierarchical radiosity. In Skala [106], pages 26–35.
- [14] S. M. Boker. A measurement of the adaptation of color vision to the spectral environment. *Psychological Science*, 8(2):130–144, 1997.
- [15] M. R. Bolin and G. W. Meyer. A perceptually based adaptive sampling algorithm. In SIGGRAPH [103].
- [16] R. M. Boynton. *Human Color Vision*. Optical Society of America, 1992.
- [17] A. P. Bradley. A wavelet visible difference predictor. *IEEE Transactions on Image Processing*, 8(5):717–730, May 1999.
- [18] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540, Apr. 1983.
- [19] A. Chalmers, S. Daly, A. McNamara, K. Myszkowski, and T. Troscianko. Image quality metrics, 2000. SIGGRAPH 2000 Course Notes #44.
- [20] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Proceedings of Graphics Interface '93*, pages 245–253, May 1993.
- [21] M. F. Cohen and J. R. Wallace. *Radiosity and Realistic Image Synthesis*. Academic Press Professional, Boston, MA, 1993.
- [22] S. Daly. The visible differences predictor: An algorithm for the assessment of image fidelity. In A. B. Watson, editor, *Digital Images and Human Vision*, chapter 14, pages 179–206. The MIT Press, 1993.
- [23] C. DeCusatis, editor. *Handbook of Applied Photometry*. Optical Society of America and Springer Verlag, 1997.
- [24] J. Dorsey and P. Slusallek, editors. *Rendering Techniques '97 (Proceedings of the Eighth Eurographics Workshop on Rendering)*. Eurographics, Springer-Verlag Wien, 1997.

- [25] G. Drettakis and N. Max, editors. *Rendering Techniques '98 (Proceedings of the Ninth Eurographics Workshop on Rendering)*. Eurographics, Springer-Verlag Wien, 1998.
- [26] F. Durand and J. Dorsey. Interactive tone mapping. In B. Péroche and H. Rushmeier, editors, *Rendering Techniques '00 (Proceedings of the Eleventh Eurographics Workshop on Rendering)*, pages 219–230. Eurographics, Springer Verlag Wien, 2000.
- [27] M. D. Fairchild. *Color Appearance Models*. Addison Wesley Longmann, 1998.
- [28] M. Fedà. A monte carlo approach for galerkin radiosity. *The Visual Computer*, 12(8):390–405, 1996.
- [29] D. W. Fellner and L. Szirmay-Kalos, editors. *Eurographics '97 Conference Proceedings*, volume 16 of *Computer Graphics Forum*. Eurographics, Sept. 1997.
- [30] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual adaptation for realistic image synthesis. In SIGGRAPH [101], pages 249–258.
- [31] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual masking for computer graphics. In SIGGRAPH [102], pages 143–152.
- [32] G. A. Fry. The eye as an detector. In C. J. Bartleson and F. Grum, editors, *Visual Measurements*, volume 5 of *Optical Radiation Measurements*, chapter 3, pages 131–182. Academic Press, 1984.
- [33] T. A. Funkhouser. Coarse-grained parallelism for hierarchical radiosity using group iterative methods. In SIGGRAPH [101], pages 343–352.
- [34] A. Gaddipatti, R. Machiraju, and R. Yagel. Steering image generation with wavelet based perceptual metric. In Fellner and Szirmay-Kalos [29], pages 241–251.
- [35] S. Gibson and R. J. Hubbard. Perceptually driven radiosity. *Computer Graphics Forum*, 16(2):129–140, 1997.
- [36] B. Girod. What's wrong with mean-squared error? In A. B. Watson, editor, *Digital Images and Human Vision*, chapter 15, pages 207–220. The MIT Press, 1993.

- [37] A. S. Glassner. *Principles of Digital Image Synthesis*. Computer Graphics and Geometric Modeling. Morgan Kaufmann, San Francisco, CA, 1995.
- [38] C. M. Goral, K. K. Torrance, D. P. Greenberg, and B. Battaile. Modelling the interaction of light between diffuse surfaces. In *Computer Graphics Proceedings*, Annual Conference Series, pages 213–222. ACM SIGGRAPH, 1984.
- [39] S. J. Gortler, P. Schröder, M. F. Cohen, and P. Hanrahan. Wavelet radiosity. In SIGGRAPH [100], pages 221–230.
- [40] E. A. Haines and J. R. Wallace. Shaft culling for efficient ray-traced radiosity. In Sakas et al. [89], pages 122–138.
- [41] P. Hanrahan, D. Salzman, and L. Aupperle. A rapid hierarchical radiosity algorithm. In *Computer Graphics Proceedings*, Annual Conference Series, pages 197–206. ACM SIGGRAPH, 1991.
- [42] P. M. Hanrahan and W. Purgathofer, editors. *Rendering Techniques '95 (Proceedings of the Sixth Eurographics Workshop on Rendering)*. Eurographics, Springer-Verlag Wien, 1995.
- [43] P. Heckbert. Adaptive radiosity textures for bidirectional ray tracing. In *Computer Graphics Proceedings*, Annual Conference Series, pages 145–154. ACM SIGGRAPH, 1990.
- [44] P. Heckbert. Discontinuity meshing for radiosity. In *Third Eurographics Workshop on Rendering*, pages 203–226, Bristol, UK, May 1992.
- [45] D. Hedley, A. Worall, and D. Paddon. Selective culling of discontinuity lines. In Dorsey and Slusallek [24], pages 69–80.
- [46] R. W. G. Hunt. *Measuring Colour*. Ellis Horwood Series in Applied Science and Industrial Technology. Ellis Horwood Ltd, 2nd edition, 1991.
- [47] C. E. Jacobs, A. Finkelstein, and D. H. Salesin. Fast multiresolution image querying. In *Computer Graphics Proceedings*, Annual Conference Series, pages 277–286. ACM SIGGRAPH, 1995.
- [48] D. J. Jobson, Z. Rahman, and G. A. Woodell. Retinex image processing: Improved fidelity to direct visual observation. In *Proceedings of the IS&T Fourth Colour Imaging Conference: Colour Science, Systems, and Applications*, 1996.

- [49] D. J. Jobson, Z. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6(7):965–976, July 1997.
- [50] D. J. Jobson, Z. Rahman, and G. A. Woodell. Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*, 6(3):451–462, Mar. 1997.
- [51] M. H. Kalos and P. A. Whitlock. *Monte Carlo Methods*, volume I. Basics. John Wiley & Sons, 1986.
- [52] A. Keller. The fast calculation of form factors using low discrepancy sequences. In Purgathofer [82], pages 195–204.
- [53] A. Keller. Quasi-monte carlo radiosity. In Pueyo and Schröder [81], pages 101–110.
- [54] A. Keller. Instant radiosity. In SIGGRAPH [102], pages 49–56.
- [55] A. Keller. *Quasi-Monte Carlo Methods for Photorealistic Image Synthesis*. PhD thesis, Department of Computer Science, University of Kaiserslautern, Germany, 1998.
- [56] E. Land. Recent advances in retinex theory. *Vision Research*, 26:7–21, 1986.
- [57] G. W. Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, Oct. 1997.
- [58] G. E. Legge and J. M. Foley. Contrast masking in human vision. *Journal of the Optical Society of America*, 70(12):1458–1471, Dec. 1980.
- [59] B. Li, G. W. Meyer, and R. V. Klassen. A comparison of two image quality models. In *IS&T/SPIE Conference on Human Vision and Electronic Imaging III*, volume 3299, pages 98–109. SPIE, Jan. 1998.
- [60] D. Lischinski, F. Tampieri, and D. P. Greenberg. Discontinuity meshing for accurate radiosity. *IEEE Computer Graphics and Applications*, 12(6):25–39, Nov. 1992.
- [61] J. Lubin. A visual discrimination model for imaging systems design and evaluation. In E. Peli, editor, *Vision Models for Target Detection and Recognition*, pages 245–283. World Scientific, Singapore, 1995.

- [62] J. L. Mannos and D. J. Sakrison. The effect of a visual fidelity criterion on the encoding of images. *IEEE Transactions of Information Theory*, 20(4):525–536, July 1974.
- [63] D. Marini and A. Rizzi. A computational approach to color adaptation effects. *Image and Vision Computing*, 18(13):1005–1014, Oct. 2000.
- [64] I. Martin, X. Pueyo, and D. Tost. An image-space refinement criterion for linear hierarchical radiosity. In *Graphics Interface '97*, pages 26–36, May 1997.
- [65] K. Myszkowski. The visible differences predictor: application to global illumination problems. In Drettakis and Max [25], pages 223–236.
- [66] K. Myszkowski, A. Wojdala, and K. Wicynski. Non-uniform adaptive meshing for global illumination. *Machine GRAPHICS & VISION*, 3(4):601–609, 1994.
- [67] A. Neumann, L. Neumann, P. Bekaert, Y. D. Willems, and W. Purgathofer. Importance-driven stochastic ray radiosity. In Pueyo and Schröder [81], pages 111–122.
- [68] L. Neumann, M. Feda, M. Kopp, and W. Purgathofer. A new stochastic radiosity method for highly complex scenes. In Sakas et al. [89], pages 195–206.
- [69] L. Neumann, K. Matković, A. Neumann, and W. Purgathofer. Incident light metering in computer graphics. *Computer Graphics Forum*, 17(4):235–248, 1998.
- [70] L. Neumann, K. Matković, and W. Purgathofer. Automatic exposure in computer graphics based on the minimum information loss principle. In *Proceedings of CGI 98*, pages 666–677, Hannover, Germany, June 1998.
- [71] L. Neumann, K. Matković, and W. Purgathofer. Perception based color image difference. In F. N. Ferreira and M. Göbel, editors, *Eurographics '98 Conference Proceedings*, volume 17 of *Computer Graphics Forum*, pages C233–C241. Eurographics, Sept. 1998.
- [72] L. Neumann, A. Neumann, and P. Bekaert. Radiosity with well distributed ray sets. In Fellner and Szirmay-Kalos [29], pages C261–C270.
- [73] L. Neumann, W. Purgathofer, R. F. Tobler, A. Neumann, P. Eliáš, M. Feda, and X. Pueyo. The stochastic ray method for radiosity. In Hanrahan and Purgathofer [42], pages 206–218.

- [74] I. Ohzawa. Campbell-robson contrast sensitivity chart: A new rendition. Available from http://totoro.berkeley.edu/izumi/CSF/A_JG_RobsonCSFchart.html.
- [75] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In SIGGRAPH [103], pages 287–298.
- [76] S. N. Pattanaik and S. P. Mudur. Computation of the global illumination by monte carlo simulation of the particle model of light. In A. Chalmers and D. Paddon, editors, *Third Eurographics Workshop on Rendering*, pages 71–83, Bristol, May 1992. Eurographics.
- [77] S. N. Pattanaik and S. P. Mudur. Efficient potential equation solutions for global illumination computation. *Computers and Graphics*, 17(4):387–396, July 1993.
- [78] S. N. Pattanaik and S. P. Mudur. The potential equation and importance in illumination computations. *Computer Graphics Forum*, 12(2):131–136, June 1993.
- [79] S. N. Pattanaik and S. P. Mudur. Adjoint equations and random walks for illumination computation. *ACM Transactions on Graphics*, 14(1):77–102, Jan. 1995.
- [80] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Computer Graphics Proceedings*, Annual Conference Series, pages 47–54. ACM SIGGRAPH, 2000.
- [81] X. Pueyo and P. Schröder, editors. *Rendering Techniques '96 (Proceedings of the Seventh Eurographics Workshop on Rendering)*. Eurographics, Springer-Verlag Wien, 1996.
- [82] W. Purgathofer, editor. *Proceedings of the Twelfth Spring School on Computer Graphics '96*, Comenius University, Bratislava, Slovakia, June 1996.
- [83] J. Pšikryl, P. Bekaert, and W. Purgathofer. Importance-driven hierarchical stochastic ray radiosity. In V. Skala, editor, *Proceedings of the Eighth International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media '2000*, volume Short Communication Papers, pages 100–106, University of West Bohemia, Pilsen, Czech Republic, Feb. 2000.

- [84] J. Přikryl and W. Purgathofer. State of the art in perceptually-driven radiosity. In A. A. de Sousa and B. Hopgood, editors, *Eurographics 98 State of the Art Reports*, pages 89–104, Sept. 1998. ISSN 1017-4656.
- [85] M. Ramasubramanian, S. N. Pattanaik, and D. P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In SIGGRAPH [104], pages 73–82.
- [86] A. R. Robertson. Historical development of cie recommended color difference equations. *COLOR research and applications*, 15:167–170, 1990.
- [87] R. Y. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley and Sons, 1981.
- [88] H. Rushmeier, G. Ward, C. D. Piatko, P. Sanders, and B. Rust. Comparing real and synthetic images: Some ideas about metrics. In Hanrahan and Purgathofer [42], pages 82–91.
- [89] G. Sakas, P. Shirley, and S. Müller, editors. *Photorealistic Rendering Techniques (Proceedings of the Fifth Eurographics Workshop on Rendering)*. Eurographics, Springer-Verlag Berlin–Heidelberg–New York, 1995.
- [90] M. Sbert. *The use of global random directions to compute radiosity. Global Monte Carlo techniques*. PhD thesis, Departament de LLenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, Nov. 1996.
- [91] M. Sbert. Error and complexity of random walk monte carlo radiosity. *IEEE Transactions on Visualization and Computer Graphics*, 3(1):23–38, Mar. 1997.
- [92] M. Sbert. Optimal source selection in shooting random walk monte carlo radiosity. In Fellner and Szirmay-Kalos [29], pages C301–C308.
- [93] M. Sbert. Random walk radiosity with infinite path length. *Computer & Graphics*, 22(2):161–166, 1998.
- [94] M. Sbert. Personal communication, 2000.
- [95] A. Scheel, M. Stamminger, and H.-P. Seidel. Tone reproduction for interactive walkthroughs. In M. Gross and F. R. A. Hopgood, editors, *Eurographics '2000 Conference Proceedings*, volume 19 of *Computer Graphics Forum*, pages C301–C311. Eurographics, Sept. 2000.

- [96] C. Schlick. Quantization techniques for visualization of high dynamic range pictures. In Sakas et al. [89], pages 7–20.
- [97] P. Shirley. Radiosity test scenes. Available from <http://radsite.lbl.gov/mgf/scene/erw5.tar.Z>.
- [98] P. Shirley. Radiosity via ray tracing. In J. Arvo, editor, *Graphics Gems II*, pages 306–310. Academic Press Professional, Boston, MA, 1991.
- [99] S. Shlaer. The relation between visual acuity and illumination. *Journal of General Physiology*, 21:165–188, 1937.
- [100] *Computer Graphics Proceedings*, Annual Conference Series. ACM SIGGRAPH, 1993.
- [101] *Computer Graphics Proceedings*, Annual Conference Series. ACM SIGGRAPH, 1996.
- [102] *Computer Graphics Proceedings*, Annual Conference Series. ACM SIGGRAPH, 1997.
- [103] *Computer Graphics Proceedings*, Annual Conference Series. ACM SIGGRAPH, 1998.
- [104] *Computer Graphics Proceedings*, Annual Conference Series. ACM SIGGRAPH, 1999.
- [105] F. Sillion and C. Puech. *Radiosity and Global Illumination*. Morgan Kaufmann, San Francisco, CA, 1994.
- [106] V. Skala, editor. *Proceedings of the Winter School of Computer Graphics and CAD Systems '95*, University of West Bohemia, Pilsen, Czech Republic, Feb. 1995.
- [107] B. Smits, J. Arvo, and D. Greenberg. A clustering algorithm for radiosity in complex environments. In *Computer Graphics Proceedings*, Annual Conference Series, pages 435–442. ACM SIGGRAPH, 1994.
- [108] B. E. Smits, J. R. Arvo, and D. H. Salesin. An importance-driven radiosity algorithm. In *Computer Graphics Proceedings*, Annual Conference Series, pages 273–282. ACM SIGGRAPH, 1992.
- [109] S. S. Stevens. To honor Fechner and repeal his law. *Science*, 133:80–86, 1961.

- [110] S. S. Stevens and J. C. Stevens. Brightness function: Parametric effects of adaptation and contrast. *Journal of the Optical Society of America*, 50(11):1139, Nov. 1960.
- [111] S. S. Stevens and J. C. Stevens. Brightness function: Effects of adaptation. *Journal of the Optical Society of America*, 53:375–385, 1963.
- [112] A. Stockman and L. T. Sharpe. Spectral sensitivities of the middle- and long-wavelength sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000.
- [113] D. Stuttard, A. Worrall, D. Paddon, and C. Willis. A parallel radiosity system for large data sets. In Skala [106], pages 421–430.
- [114] P. C. Teo and D. J. Heeger. Perceptual image distortion. In *Human Vision, Visual Processing and Digital Display V, IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology*, 1994.
- [115] J. P. Thomas. Spatial resolution and spatial interaction. In E. C. Carette and M. P. Friedman, editors, *Handbook of Perception*, volume 5, pages 233–264. Academic Press, New York, 1979.
- [116] R. F. Tobler, A. Wilkie, M. Feda, and W. Purgathofer. A hierarchical subdivision algorithm for stochastic radiosity methods. In Dorsey and Slusallek [24], pages 193–204.
- [117] J. Tumblin, J. K. Hodgins, and B. K. Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, Jan. 1999.
- [118] J. Tumblin and H. E. Rushmeier. Tone reproduction for realistic computer generated images (revision). Technical Report GIT-GVU-92-31, Graphics, Visualization & Usability Center, College of Computing, Georgia Institute of Technology, 1992.
- [119] J. Tumblin and H. E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, Nov. 1993.
- [120] J. Tumblin and G. Turk. LCIS: A boundary hierarchy for detail-preserving contrast reduction. In SIGGRAPH [104], pages 83–90.
- [121] Z. ur Rahman, D. J. Jobson, and G. A. Woodell. Multi-scale retinex for color image enhancement. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 3, pages 1003–1006. IEEE, 1996.

- [122] Z. ur Rahman, D. J. Jobson, and G. A. Woodell. A multiscale retinex for color rendition and dynamic range compression. In A. G. Tescher, editor, *Applications of Digital Image Processing XIX*, volume 2847. SPIE, 1996.
- [123] J. R. Wallace, K. A. Elmquist, and E. A. Haines. A ray tracing algorithm for progressive radiosity. In *Computer Graphics Proceedings, Annual Conference Series*, pages 315–324. ACM SIGGRAPH, 1989.
- [124] J. Walraven, C. Enroth-Cugell, D. C. Hood, D. I. A. MacLeod, and J. L. Schnapf. The control of visual sensitivity: Receptor and postreceptor processes. In L. Spillmann and J. S. Werner, editors, *Visual Perception: The Neurophysiological Foundations*, pages 53–102. Academic Press, 1990.
- [125] G. Ward. Radiance rendering package. Available from <http://radsite.lbl.gov/radiance/HOME.html>.
- [126] G. Ward. A contrast-based scalefactor for luminance display. In P. Heckbert, editor, *Graphics Gems IV*, chapter VII.2, pages 415–421. Academic Press, 1994.
- [127] G. Ward-Larson. LUMFACTOR. *Radiance Digest*, 2(0), Dec. 1991.
- [128] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley and Sons, 2 edition, 1982.
- [129] H. R. Zatz. Galerkin radiosity: A higher order solution method for global illumination. In SIGGRAPH [100], pages 213–220.
- [130] X. Zhang, D. A. Silverstein, J. E. Farrel, and B. A. Wandell. Color image quality metric S-CIELAB and its application on halftone texture visibility. In *COMPCON97 Digest of Papers*, pages 44–48. IEEE, 1997.
- [131] X. Zhang and B. Wandell. A spatial extension of CIELAB for digital color image reproduction. In *Society for Information Display 1996 Symposium Digest*, volume 27, pages 731–734, San Diego, 1996.
- [132] X. Zhang and B. A. Wandell. Color image fidelity metrics evaluated using image distortion maps. *Signal Processing*, 70:201–214, 1998.

Appendix A

Curriculum Vitae

Born 1970 in Prague, Czech Republic.

1977 – 1989: High school. Specialization: electrical engineering. Examination for school-leaving certificate at 1989.

1989 – 1994: Studies at Faculty of Electrical Engineering, Czech Technical University, Prague. Specialization: computer science. MSc. (Eng.), diploma thesis: Ray-tracing and animation in the parallel computational environment.

1994 – 1997: PhD. studies at the Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University, Prague. Supervisor: prof. Slavík. Administrator of UNIX machines (IRIX/HP-UX/Linux).

Since 1997: PhD. studies at the Institute of Computer Graphics, Vienna University of Technology, Vienna, Austria. Supervisor: prof. Purgathofer.

Since 2000: Researcher at the Virtual Reality group of Research Center for Virtual Reality and Visualization, Vienna, Austria.

Professional interests

Computer graphics: Photorealistic rendering techniques, mainly radiosity, making use of certain features of human perception to speed up image synthesis, parallelization of ray tracing and radiosity.

User interfaces: X-Window, GTK+, Gnome.

UNIX: Administration, networking, security.

Experience

Language knowledge: English - fluent, German - fluent, Russian - fair.

Operating systems: Unix (IRIX, HP-UX, Linux), MS-DOS/Windows, VMS.

Programming languages: C/C++, Python, Objective-C, Pascal, Fortran, x86 Assembler, Java.

Graphical APIs: GL/OpenGL/Mesa, OpenInventor, PHIGS, GKS.

Funding: FWF 11545 MAT “Perceptually-Driven Radiosity” (1996-1999), ÖAD 1999/17 “Efficient Space Subdivision Methods for Global Illumination”.

Conferences and workshops: WSCG, 1993, 1995–1997, 1999, 2000; First European PVM Users Group Meeting, Rome, Italy, 1994; Spring School of Computer Graphics, Bratislava, Slovakia, 1994; Eurographics Rendering Workshops, 1997–2000; Spring Conference on Computer Graphics, Budmerice, Slovakia, 1998, 1999; Eurographics 1998; Study Day on Physically-Based Rendering, Leuven, Belgium, 1999.

Talks:

- State of the Art in Perceptually-Driven Radiosity. EG’98, Lissabon, Portugal, 1998/09/04.
- Perceptually-Driven Termination For Stochastic Radiosity. WSCG’99, Pilsen, Czech Republic, 1999/02/12.
- What Can Be Improved on Perceptually-Driven Radiosity. Study Day on Physically-Based Rendering, Leuven, Belgium, 1999/12/16.
- Overview of Perceptually-Driven Rendering Algorithms. University of Girona, Spain, 2000/11/14.

Membership: ACM, CSTUG (Czechoslovak TeX Users Group), CZLUG (Czech Linux Users Group).

Projects

RTTONVW

(1994-1996, CTU Prague and University of Paderborn, C)

Parallel ray-tracer based on PVM suite, participated on the initial release of the ray-tracer code, later work mainly on the communication layer.

ART

(since 1996, TU Wien, Objective C)

Distribution ray-tracer, tone-mapping, perceptual image comparison routines, image pyramids, colour appearance models.

RENDERPARK

(since 1997, Katholieke Universiteit Leuven, C/C++)

Tone-mapping, CIE colour systems, radiosity tweaks (hierarchical stochastic Jacobi, perceptual subdivision and form-actor computation for Galerkin radiosity)

GOLEM

(2000, CTU Prague, C++)

Routines for estimating scene complexity measures.

GNU WGET

(registered GNU developer since 2000, C)

Support for different FTP server architectures, occasional code, web pages, and documentation tweaks.

GNU TRANSLATION PROJECT

(member of the Czech team since 1997, GNU gettext)

Translating message catalogues for GNU programs into Czech language.

STUDIERTUBE

(since 2000, TU Wien and VRVis, C++)

Linux port, tracker linearisation, support for Intersense and Polhemus trackers, presentation system based on VRML and OpenInventor.

Appendix B

List of Publications

- [P1] J. Bittner and J. Příkryl. Exact regional visibility using line space partitioning. Technical Report TR-186-2-01-06, Institute of Computer Graphics, Vienna University of Technology, March 2001.
- [P2] R. Berka, M. Brachtl, A. Holeček, M. Novotný, J. Příkryl, P. Slavík, and J. Žára. Three-dimensional information system on the world wide web. In B. Urban, editor, *Multimedia '96*, pages 70–78. Eurographics, Springer-Verlag Wien, 1996.
- [P3] A. Fuhrmann, R. Splechtma, and J. Příkryl. Comprehensive calibration and registration procedures for augmented reality. To appear in *Proceedings of The Fifth Eurographics Workshop on Virtual Environments*, Springer-Verlag Wien, June 2001.
- [P4] A. Fuhrmann, J. Příkryl, R. Tobler, and W. Purgathofer. Interactive content for presentations in virtual reality. Technical Report TR-VRVis-2001-011, VRVis Competence Center, Vienna, April 2001.
- [P5] V. Havran, J. Příkryl, and W. Purgathofer. Statistical comparison of ray-shooting efficiency schemes. Technical Report TR-186-2-00-14, Institute of Computer Graphics, Vienna University of Technology, May 2000.
- [P6] K. Menzel, O. Schmidt, F. Stangenberg, Ch. Hornung, B. Lange, A. Holeček, J. Příkryl, and J. Žára. Distributed rendering techniques using virtual walls. In *Proceedings of First European PVM Users Group Meeting*. Roma, Italy, October 1994.
- [P7] L. Neumann, A. Neumann, J. Příkryl, and W. Purgathofer. The constant radiance term. *Machine GRAPHICS & VISION*, 7(3):535–549, 1998.

- [P8] J. Přikryl, P. Bekaert, and W. Purgathofer. Importance-driven hierarchical stochastic ray radiosity. Technical Report TR-186-2-99-22, Institute of Computer Graphics, Vienna University of Technology, October 1999.
- [P9] J. Přikryl, P. Bekaert, and W. Purgathofer. Importance-driven hierarchical stochastic ray radiosity. In V. Skala, editor, *Proceedings of the Eighth International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media '2000*, volume Short Communication Papers, pages 100–106, University of West Bohemia, Plzeň, Czech Republic, February 2000.
- [P10] J. Přikryl and H. Löffelmann, editors. *Proceedings of the 3rd Central European Seminar on Computer Graphics for Students (CESCG'99)*. Institute of Computer Graphics, Vienna University of Technology, April 1999.
- [P11] J. Přikryl and W. Purgathofer. State of the art in perceptually-driven radiosity. In A. A. de Sousa and B. Hopgood, editors, *Eurographics 98 State of the Art Reports*, pages 89–104, September 1998.
- [P12] J. Přikryl and W. Purgathofer. Perceptually-driven termination for stochastic radiosity. In V. Skala, editor, *Proceedings of the Seventh International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media '99*, pages 418–425, University of West Bohemia, Plzeň, Czech Republic, February 1999.
- [P13] J. Přikryl and W. Purgathofer. Perceptually-driven termination for stochastic radiosity. Technical Report TR-186-2-99-01, Institute of Computer Graphics, Vienna University of Technology, January 1999.
- [P14] J. Přikryl. On mesh preprocessing for radiosity. In N. Thalmann and V. Skala, editors, *Proceedings of The Fifth International Conference in Central Europe on Computer Graphics and Visualization '97*, volume IV, pages 687–690. University of West Bohemia, Plzeň, Czech Republic, February 1997.
- [P15] P. Slavík and J. Přikryl. Dithering as a method for image data compression. In V. Skala, editor, *Proceedings of The Third International Conference in Central Europe on Computer Graphics and Visualisation '95*, volume II, pages 283–288. University of West Bohemia, Plzeň, Czech Republic, February 1995.
- [P16] J. Žára, A. Holeček, and J. Přikryl. Parallelisation of the ray-tracing algorithm. In V. Skala, editor, *Proceedings of Winter School of Computer Graphics and CAD Systems*, pages 113–117. University of West Bohemia, Plzeň, Czech Republic, January 1994.

- [P17] J. Žára, A. Holeček, and J. Příkryl. When the parallel ray-tracer starts to be efficient? In E. Ružický, editor, *Proceedings of 10th Spring School on Computer Graphics*, pages 108–116, Bratislava, Slovakia, June 1994.
- [P18] J. Žára, A. Holeček, J. Příkryl, J. Buriánek, and K. Menzel. Load balancing for parallel raytracer on virtual walls. In V. Skala, editor, *Proceedings of The Third International Conference in Central Europe on Computer Graphics and Visualisation '95*, pages 439–447. University of West Bohemia, Plzeň, Czech Republic, February 1995.