

Diffraction Shader

Bachelorarbeit

der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Michael Single

2014

Leiter der Arbeit:
Prof. Dr. Matthias Zwicker
Institut für Informatik und angewandte Mathematik

Abstract

Inhaltsverzeichnis

1	Introduction	1
1.1	Motivation	1
1.2	Goals	2
1.3	Previous work	2
1.4	Overview	3
2	Theoretical Background	4
2.1	Basics in Modeling Light in Computer Graphics	4
2.1.1	Radiometry	4
2.1.2	Spectral Energy	4
2.1.3	Spectral Power	5
2.1.4	Spectral Irradiance	5
2.1.5	Spectral Radiance	6
2.1.6	BRDF	6
2.1.7	Colorspace	6
2.1.8	Spectral Rendering	6
2.2	Wave Theory for Light and Diffraction	6
2.3	Stam's BRDF formulation	6
A	Appendix	9
A.1	Schlick's approximation	9
A.2	Spherical Coordinates	9
A.3	Tangent Space	10
	List of Tables	11
	List of Figures	11
	List of Algorithms	12
	Bibliography	13

Kapitel 1

Introduction

1.1 Motivation

In Nature, coloring mostly comes from the inherent colors of materials but sometimes colorization has a pure physical origin such as the effect diffraction or interference of light. Both phenomenon are causing the so called structural coloration, which is the production of color through the interaction of visible light with micrioscopically structured surfaces. Color production is due to wave interference with quasiperiodic structures whose periodicity leads to interaction with visible light. Therefore we perceive color when the different wavelengths composing white light are selectively interfered with by matter (absorbed, reflected, refracted, scattered, or diffracted) on their way to our eyes, or when a non-white distribution of light has been emitted. In animals, such as feathers of birds and the scales of butterflies, interference is created by a range of photonic mechanisms, including diffraction grating, selective mirrors, photonic crystals. The connection between microscopic structures and coloration has been observed by Robert Hooke in the early seventeenth century. The discovery of the wave nature of light led to the conclusion that the cause for the coloration lies in wave interference.

In the field of computer graphics, many researchers have been attempted rendering of structural colors by formulating a the bidirectional reflectance distribution function (BRDF) for this purpose. But most of the techniques so far, however, are either too slow for interactive rendering or rely on simplifying assumption, like modeling light as rays, to achieve real-time performance, which are not able capturing the essence of diffraction at all.



(a) Elaphe Guttata Snake



(b) Xenopeltis Snake

Abbildung 1.1: Effect of diffraction on snake sheds for different species

1.2 Goals

The purpose of this thesis is to simulate realistically by rendering structural colors caused by the effect of diffraction on different biological structures in realtime. We focus on structural colors generated by diffraction gratings, in particular our approach applies to surfaces with quasiperiodic structures at the nanometer scale that can be represented as heightfields. such structures are found on the sehs of snakes, wings of butterflies or the bodies of various insects. we restrict ourself and focus on different snake skins sheds which are acquired nanoscaled heightfields using atomic force microscopy.

In oder to achieve our rendering purpose we will rely J. Stam's formulation of a BRDF which basically describes the effect of diffraction on a given surface assuming one knows the hightfield of this surface and will further extend this. Appart from Stam's approach, which models the heightfield as a probabilistic superposition of bumps and proceeds to derive an analytical expression for the BRDF, our BRDF representation takes the heightfield from explicit measurement. I.E. in our case, those heightfields are small patches of the microstructured surfaces (in nano-scale) taken by AFM of snake skin patches provided by our collaborators in Geneva.. So this approach is closer to real truth, since we use measured surfaces instead of statistical surface profile.

Therefore, this work can be considered as an extension of J. Stam's derivations for the case one is provided by a explicit height field on a quasiperiodic structure.

Real time performance is achieved with a representation of the formula as a power series over a variable related to the viewing and lighting directions. Values closely related to the coefficients in that power series are precomputed.

The contribution is that this approach is more broadly applicable than the previous work. Although the previously published formula theoretically has this much flexibility already, there is a novel contribution in demonstrating how such generality can be leveraged in practical implementation

1.3 Previous work

stam, hooke, see our paper, see stams paper, see own research.

Robert Hooke = observed connection between microscopic structures and colorisation wave nature of light led to conclusion that the cause for the colouration lies in wave interference.

previous

In computer graphics literature, Stam was the first to develop reflection models based on wave optics called diffraction shaders, that can produce colorful diffraction effects. His approach is based on a far field approximation of the Kirchhoff integral. He shows that for surfaces represented as nanoscale heightfields it is possible to derive their BRDF as the Fourier transformation of a function of the heightfield. Nevertheless, this formulation is not immediately useful for efficient rendering of measured complex nanostructures since this would require the on-the-fly evaluation of and integration over Fourier transforms of the heightfield that depend on the light and viewing geometry. In his derivations, Stam models heightfields as probabilistic superpositions of bumps forming periodic like structures. This provides him an analytical identity for this class of heightfields. However, biological nanostructures are way more complex and do not lend themselves to this simplified statistical model.

follow ups

1.4 Overview

The remainder of this thesis is organized as follows: due to the fact that this thesis has a rather advanced mathematical complexity the first part of chapter 2 introduces some important definitions which are required in order to be able to follow the derivations in the last third of chapter 2. Before starting the derivations, a brief summary of J. Stam's Paper about diffraction shaders is provided since this whole thesis is based on his BRDF representation. Our derivations itself are listed step-wise, whereas there is a final representation provided by the end of chapter 2. Chapter 3 addresses the practical part of this thesis, the implementation of our diffraction model, explaining all precomputation steps and how rendering is performed in our developed framework for this thesis. Chapter 4 gives some further insight about diffraction by explaining the topic about diffraction grating in depth. Furthermore, within this chapter we evaluate the qualitative validity of our BRDF models applied on different surface gratings by computing their reflectance and comparing this to the grating equation under similar conditions. Chapter 5 presents our rendered results, first the so called BRDF maps for all our gratings and shading approaches under various shading parameters and then the actual renderings on a snake mesh. Chapter 6 contains the conclusion of this thesis which starts by a review briefly discussing what has been achieved in this thesis and the drawbacks. There are also some words about my personal experience during this thesis.

Kapitel 2

Theoretical Background

2.1 Basics in Modeling Light in Computer Graphics

2.1.1 Radiometry

One purpose of Computer Graphics is to simulate the interaction of light on a surface and how a real-world observer, such as a human eye, will perceive this. These visual sensations of an eye are modeled relying on a virtual camera which captures the emitted light from the surface. The physical basis to measure such reflected light depicts radiometry which is about measuring the electromagnetic radiation transferred from a source to a receiver.

Fundamentally, light is a form of energy propagation, consisting of a large collection of photons, whereat each photon can be considered as a quantum of light that has a position, direction of propagation and a wavelength λ . A photon travels at a certain speed $v = \frac{c}{n}$, that depends only the speed of light c and the refractive index n through which it progresses. Its frequency is defined by $f = \frac{v}{\lambda}$ and its carried amount of energy q , measured in the SI unit Joule, is given by $q = hf = \frac{hv}{\lambda n}$ where h is the Plank's constant. The total energy of a large collection of photons is hence $Q = \sum_i q_i$.

2.1.2 Spectral Energy

It is important to understand that the human eye is not equally sensitive to all wavelength of the spectrum of light and therefore responds differently to specific wavelengths. Remember that our goal is to model the human visual perception. This is why we consider the energy distribution of a light spectrum rather than considering the total energy of a photon collection since then we could weight the distribution according the human visual system. So the question we want to answer is: How is the energy distributed across wavelengths of light?

The idea is to make an energy histogram from a given photon collection. For this we have to order all photons by their associated wavelength, discretize wavelength spectrum, count all photons which then will fall in same wavelength-interval, and then, finally, normalize each interval by the total energy Q . This will give us a histogram which tells us the spectral energy Q_λ for a given discrete

λ interval and thus models the so called spectral energy distribution ¹.

2.1.3 Spectral Power

Rendering an image in Computer Graphics corresponds to capturing the color sensation of an illuminated, target scene at a certain point in time. As previously seen, each color is associated by a wavelength and is directly related to a certain amount of energy. In order to determine the color of a to-be-rendered pixel of an image, we have to get a sense of how much light (in terms of energy) passes through the area which the pixel corresponds to. One possibility is to consider the flow of energy $\Phi = \frac{\Delta Q}{\Delta t}$ transferred through this area over a small period of time. This allows us to measure the energy flow through a pixel during a certain amount of time.

In general, power is the estimated rate of energy production for light sources and corresponds to the flux. It is measured in the unit Watts, denoted by Q . Since power is a rate over time, it is well defined even when energy production is varying over time. As with Spectral Energy for rendering, we are really interested in the spectral power $\Phi_\lambda = \frac{Q}{\lambda}$, measured in Watts per nanometer.

2.1.4 Spectral Irradiance

Before we can tell how much light is reflected from a given point on a surface towards the viewing direction of an observer, we first have to know how much light arrives at this point. Since in general a point has no length, area or even volume associated, let us instead consider an infinitesimal area ΔA around a such a point. Then, we can ask ourself how much light falls in such a small area. When further observing this process over a short period in time, this quantity is the spectral irradiance E as illustrated in figure 2.1. Summarized, this quantity tells us how much spectral power is incident on a surface per unit area and mathematically is equal:

$$E = \frac{\Phi_\lambda}{\Delta A} \quad (2.1)$$

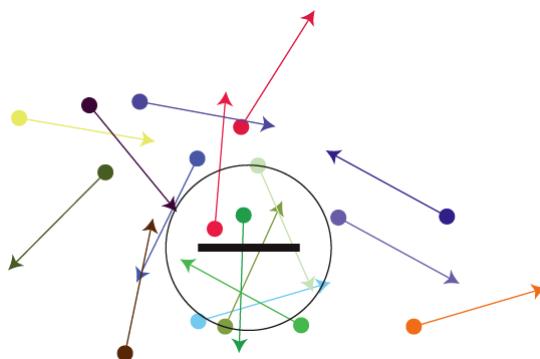


Abbildung 2.1: Irradiance is the summed up radiance over all directions

¹Intensive quantities can be thought of as density functions that tell the density of an extensive quantity at an infinitesimal point.

2.1.5 Spectral Radiance

2.1.6 BRDF

how much light is reflected back from P in a given direction (generally the viewing direction), we first need to know how much light arrives at P from all directions in the hemisphere above the point.

2.1.7 Colorspace

2.1.8 Spectral Rendering

2.2 Wave Theory for Light and Diffraction

2.3 Stam's BRDF formulation

In his paper about Diffraction Shader, J. Stam derives a BRDF which is modeling the effect of diffraction for various analytical anisotropic reflexion models relying on the so called scalar wave theory of diffraction for which a wave is assumed to be a complex valued scalar. It's noteworthy, that Stam's BRDF formulation does not take into account the polarization of the light. Fortunately, light sources like sunlight and light bulbs are unpolarized.

A further assumption in Stam's Paper is, the emanated waves from the source are stationary, which implies the wave is a superposition of independent monochromatic waves. This implies that each wave is associated to a definite wavelength lambda. However, sunlight once again fulfills this fact.

In our simulations we will always assume we have given a directional light source, i.e. sunlight. Hence, Stam's model can be used for our derivations.

For his derivations Stam uses the Kirchhoff integral (ADD REF TO WIKI), which is relating the reflected field to the incoming field. This equation is a formalization of Huygen's well-known principle that states that if one knows the wavefront at a given moment, the wave at a later time can be deduced by considering each point on the first wave as the source of a new disturbance. Mathematically speaking, once the field $\psi_1 = e^{ik\mathbf{x} \cdot \mathbf{s}}$ on the surface is known, the field ψ_2 everywhere else away from the surface can be computed. More precisely, we want to compute the wave ψ_2 equal to the reflection of an incoming planar monochromatic wave $\psi_1 = e^{ik\omega_i * x}$ traveling in the direction ω_i from a surface S to the light source. Formally, this can be written as:

$$\psi_2(\omega_i, \omega_r) = \frac{ik e^{iKR}}{4\pi R} (F(-\omega_i - \omega_r) - (-\omega_i + \omega_r)) \cdot I_1(\omega_i, \omega_r) \quad (2.2)$$

with

$$I_1(\omega_i, \omega_r) = \int_S \hat{\mathbf{n}} e^{ik(-\omega_i - \omega_r) \cdot \mathbf{s}} d\mathbf{s} \quad (2.3)$$

In applied optics, when dealing with scattered waves, one does use differential scattering cross-section rather than defining a BRDF which has the following identity:

$$\sigma^0 = 4\pi \lim_{R \rightarrow \infty} R^2 \frac{\langle |\psi_2|^2 \rangle}{\langle |\psi_1|^2 \rangle} \quad (2.4)$$

where R is the distance from the center of the patch to the receiving point x_p , \hat{n} is the normal of the surface at s and the vectors:

The relationship between the BRDF and the scattering cross section can be shown to be equal to

$$BRDF = \frac{1}{4\pi} \frac{1}{A} \frac{\sigma^0}{\cos(\theta_i) \cos(\theta_r)} \quad (2.5)$$

where θ_i and θ_r are the angles of incident and reflected directions on the surface with the surface normal n . See 2.2.

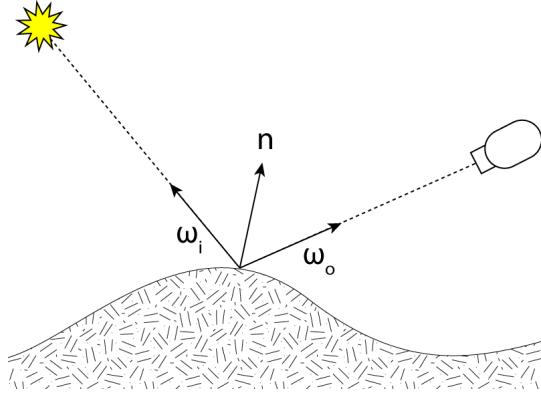


Abbildung 2.2: ω_i points toward the light source, ω_r points toward the camera, n is the surface normal

The components of vector resulting by the difference between these direction vectors: In order to simplify the calculations involved in his vectorized integral equations, Stam considers the components of vector

$$(u, v, w) = -\omega_i - \omega_r \quad (2.6)$$

explicitly and introduces the equation:

$$I(ku, kv) = \int_S \hat{n} e^{ik(u, v, w) \cdot \mathbf{s}} d\mathbf{s} \quad (2.7)$$

which is a first simplification of 2.3. Note that the scalar w is the third component of 2.6 and can be written as $w = -(\cos(\theta_i) + \cos(\theta_r))$ using spherical coordinates. The scalar $k = \frac{2\pi}{\lambda}$ represent the wavenumber.

During his derivations, Stam provides a analytical representation for the Kirchhoff integral assuming that each surface point $s(x, y)$ can be parameterized by $(x, y, h(x, y))$ where h is the height at the position (x, y) on the given (x, y) surface plane. Using the tangent plane approximation for the parameterized surface and plugging it into 2.7 he will end up with:

$$I(ku, kv) = \int \int (-h_x(x, y), -h_y(x, y), 1) e^{ikwh(x, y)} e^{ik(ux+vy)} dx dy \quad (2.8)$$

For further simplification Stam formulates auxillary function which depends on the provided height field:

$$p(x, y) = e^{iwh(x,y)} \quad (2.9)$$

which will allow him to further simplify his equation 2.8 to:

$$\mathbf{I}(ku, kv) = \int \int \frac{1}{ikw} (-p_x, -p_y, ikwp) dx dy \quad (2.10)$$

where he used that $(-h_x(x, y), -h_y(x, y), 1)e^{kwh(x,y)}$ is equal to $\frac{(-p_x, -p_y, ikwp)}{ikw}$ using the definition of the partial derivatives applied to the function 2.9.

Let $P(x, y)$ denote the Fourier Transform (FT) of $p(x, y)$. Then, the differentiation with respect to x respectively to y in the Fourier domain is equivalent to a multiplication of the Fourier transform by $-iku$ or $-ikv$ respectively. This leads him to the following simplification for 2.8:

$$\mathbf{I}(ku, kv) = \frac{1}{w} P(ku, kv) \cdot (u, v, w) \quad (2.11)$$

Let us consider the term $g = (F(-\omega_i - \omega_r) - (-\omega_i + \omega_r))$, which is a scalar factor of 2.2. The dot product with g and $(-\omega_i - \omega_r)$ is equal $2F(1 + \omega_i \cdot \omega_r)$. Putting this finding and the identity 2.11 into 2.2 he will end up with:

$$\psi_2(\omega_i, \omega_r) = \frac{ike^{iKR}}{4\pi R} \frac{2F(1 + \omega_i \cdot \omega_r)}{w} P(ku, kv) \quad (2.12)$$

By using the identity 2.5, this will lead us to his main finding:

$$BRDF_\lambda(\omega_i, \omega_r) = \frac{k^2 F^2 G}{4\pi^2 Aw^2} \langle |P(ku, kv)|^2 \rangle \quad (2.13)$$

where G is the so called geometry term which is equal:

$$G = \frac{(1 + \omega_i \cdot \omega_r)^2}{\cos(\theta_i)\cos(\theta_r)} \quad (2.14)$$

Anhang A

Appendix

A.1 Schlick's approximation

The Fresnel's equations describe the reflection and transmission of electromagnetic waves at an interface. That is, they give the reflection and transmission coefficients for waves parallel and perpendicular to the plane of incidence. Schlick's approximation is a formula for approximating the contribution of the Fresnel term where the specular reflection coefficient R can be approximated by:

$$R(\theta) = R_0 + (1 - R_0)(1 - \cos \theta)^5 \quad (\text{A.1})$$

and

$$R_0 = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2$$

where θ is the angle between the viewing direction and the half-angle direction, which is halfway between the incident light direction and the viewing direction, hence $\cos \theta = (H \cdot V)$. And n_1, n_2 are the indices of refraction of the two medias at the interface and R_0 is the reflection coefficient for light incoming parallel to the normal (i.e., the value of the Fresnel term when $\theta = 0$ or minimal reflection). In computer graphics, one of the interfaces is usually air, meaning that n_1 very well can be approximated as 1.

A.2 Spherical Coordinates

$$\forall \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \exists r \in [0, \infty) \exists \phi \in [0, 2\pi] \exists \theta \in [0, \pi] \text{ s.t.}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin(\theta) \cos(\phi) \\ r \sin(\theta) \sin(\phi) \\ r \cos(\theta) \end{pmatrix}$$

A.3 Tangent Space

The concept of tangentspace-transformation of tangent space is used in order to convert a point between world and tangent space. GLSL fragment shaders require normals and other vertex primitives declared at each pixel point, which mean that we have one normal vector at each texel and the normal vector axis will vary for every texel.

Think of it as a bumpy surface defined on a flat plane. If those normals were declared in the world space coordinate system, we would have to rotate these normals every time the model is rotated, even when just for a small amount. Since the lights, cameras and other objects are usually defined in world space coordinate system, and therefore, when they are involved in a calculation within the fragment shader, we would have to rotate them as well for every pixel. This would involve almost countless many object to world matrix transformations needed to take place at the pixel level. Therefore, instead doing so, we transform all vertex primitives into tangent space within the vertex shader.

To make this point clear an example: Even we would rotate the cube in figure A.1, the tangent space axis will remain aligned with respect to the face. Which practically speaking, will save us from performing many space transformations applied pixel-wise within the fragment shader and instead allows us to perform us the tangenspace transformation of every involved vertex primitive in the vertex-shader.

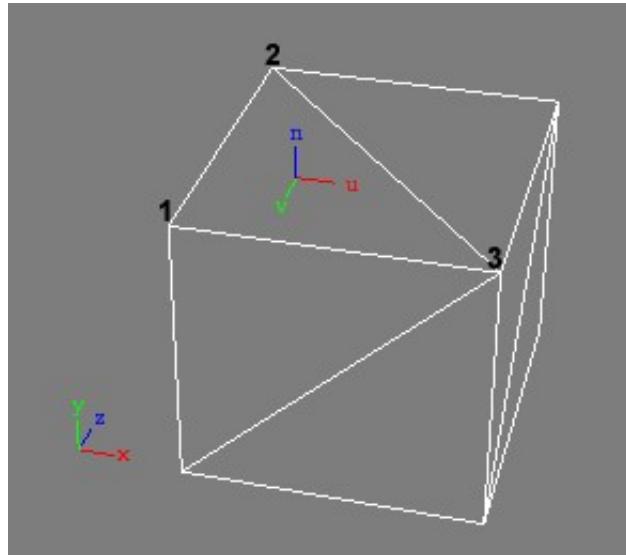


Abbildung A.1: Cube in world space (x, y, z) showing the tangen space (u, v, n) of its face (2, 1, 3)

Tabellenverzeichnis

Abbildungsverzeichnis

1.1	Effect of diffraction on snake sheds for different species	2
2.1	Irradiance is the summed up radiance over all directions	5
2.2	ω_i points toward the light source, ω_r points toward the camera, n is the surface normal	7
A.1	Cube in world space (x, y, z) showing the tangent space (u, v, n) of its face (2,1,3)	10

List of Algorithms

Literaturverzeichnis

- [Bar07] BARTSCH, Hans-Jochen: *Taschenbuch Mathematischer Formeln*. 21th edition. HASNER, 2007. – ISBN 978–3–8348–1232–2
- [DSD14] D. S. DHILLON, et a.: Interactive Diffraction from Biological Nano-structures. In: *XYZ* (2014), January
- [For11] FORSTER, Otto: *Analysis 3*. 6th edition. VIEWEG+TEUBNER, 2011. – ISBN 978–3–8348–1232–2
- [JG04] JUAN GUARDADO, NVIDIA: Simulating Diffraction. In: *GPU Gems* (2004). <https://developer.nvidia.com/content/gpu-gems-chapter-8-simulating-diffraction>
- [PAT09] PAUL A. TIPLER, Gene M.: *Physik für Wissenschaftler und Ingenieure*. 6th edition. Spektrum Verlag, 2009. – ISBN 978–3–8274–1945–3
- [PS09] P. SHIRLEY, S. M.: *Fundamentals of Computer Graphics*. 3rd edition. A K Peters, Ltd, 2009. – ISBN 978–1–56881–469–8
- [RW11] R. WRIGHT, et a.: *OpenGL SuperBible*. 5th edition. Addison-Wesley, 2011. – ISBN 978–0–32–171261–5
- [Sta99] STAM, Jos: Diffraction Shaders. In: *SIGGRPAH 99 Conference Proceedings* (1999), August

Erklärung

gemäss Art. 28 Abs. 2 RSL 05

Name/Vorname:

Matrikelnummer:

Studiengang:

Bachelor

Master

Dissertation

Titel der Arbeit:

.....
.....

LeiterIn der Arbeit:

.....

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe o des Gesetztes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist.

.....
Ort/Datum

.....
Unterschrift