Rostocker Informatik-Berichte (1994) 15

Illumination along convex edges

Nico Günther

Abstract

Analytical illumination calculation along convex edges or within concave free-form surfaces is a serious problem. Often the edges are too dark (full matrix method) or too bright concerning Monte-Carlo raycasting. For these cases different solutions are provided. A proof is given, that the widely used Monte-Carlo raycasting using constant portions is not an analytical illumination technique. However, using distance dependent portions will yield very bright edges. Comparing the diffuse method with the glossy one it is shown, that this behaviour is caused by the flux redistribution around the patches' normals.

Introduction

In the last few years several investigations were made to generalize the models suitable to analytical illumination. The *radiosity* technique was extended to glossy, specular and transparent materials. Including scattering effects caused by dust, smoke, steam etc. and illumination calculation for free form surfaces is possible. Nevertheless these developments do not affect practice too much. This is caused by two obstacles: At first the preparation of the geometrical input data produced by a modelling (CAD) system leads to a number of problems. Some steps to get a solid geometry must be done by hand. These problems and solutions are described in detail in [Baum91]. Second, the presentation of photorealistic computer graphics often requires the service of a 'rendering expert', who can correct some parameters. The main 'correction potential' can be found along convex edges. These edges are often under- or overilluminated depending on the rendering technique (see title page of Cohen and Wallace book 'Radiosity And Realistic Image Synthesis'). It is shown, that the main reasons for these failures are the 'constant-portion-technique' and the distribution of the reflected flux around the normals. The following can be analogously applied to concave free-form surfaces [1]

Underlighting and full matrix method

Commonly the geometry is tesselated into patches, and all computation is done for the vertices of the patches to use the hardware Gouraud shading without an additional extrapolation step. Unfortunately a number of vertices lies at the border of the polygons and hence along an edge. Figure 1 (left frame) illustrates this. The point F_1 gets no light from any point of the adjacent

[1]The theoretical basis for all explanations can be found in [Nic93], which is available as a preprint.

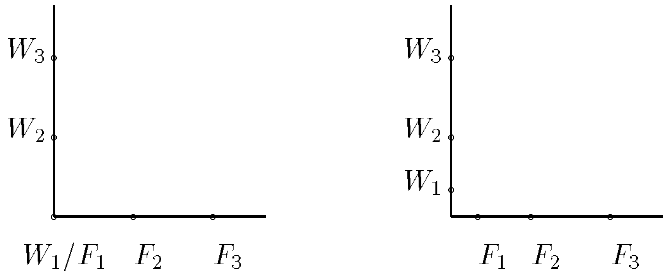

Figure 1: To avoid dark edges the first points on the wall and at the floor, W_1 and F_1, are shiftet a small distance to the centre of the polygon. This is done for both shooting and gathering.

wall W, because $cos(\phi_i)$ is zero. Therefore all 'form-factors' F_1W_i are zero and F_1 must appear a little bit darker. The same occurs for W_1; here all matrix elements W_1F_i are zero. Hence dark stripes will occur on both sides along the edge. In addition to this the color bleeding along the edges is vague and the whole image appears a little bit dark. An easy and elegant solution to this problem can be found in [Kok92]. The 'edge'points F_1 and W_1 are shifted a small distance towards the centre of the corresponding polygon as shown in the right frame of figure 1. This is done for both emitter and target points. The corresponding matrix elements F_1W_i and W_1F_i are bigger than zero now. Since the 'form-factors' depend on the square distance , and the distance between the shifted edge points is very small, these 'new' matrix elements have a big influence on the resulting image. The whole image gets brighter, the color bleeding is noticable and the dark edges disappear. The color images belonging to this article can be reached via anonymous ftp from ftp.uni-rostock.de in pub/graphics/edge_pictures.tar.Z. The pictures 2A to 2C show the differences between these two methods. Picture 2C is computed using the offset technique, whereas picture 2A is computed without shifting.

Monte-Carlo raycasting

Due to the storage problems full-matrix solutions are not capable to compute the lighting of complex scenes. Commonly a Monte-Carlo raycasting technique is used. At Eurographics91 Peter Shirley gave a mathematical proof, that the complexity of such an algorithm is O(N) (see [Shir91]). He made the assumptions, that the ray distribution is proportional to $cos(\phi)$ and all portions are constant. Then he was able to compute, how many rays per patch must be shot to get the desired (overall) image quality. Testing different distribution strategies we found, that the 'constant-portion-technique' is not suitable to produce radiosity-like images (see [Linke93]). Assume a point-like emitter E and two targets perpendicular to the outgoing ray $\vec{R}$. The distance from E to target 2 is as twice as big as the distance from E to target 1. Both targets have exactly the same area a^2, whereby a is the length of every target side. We can project this in the two-dimensional space as seen in figure 2. Under the assumption of constant portions target 1 should be hit twice the times as target 2 to simulate the r^2-dependency of analytical illumination in two-dimensional space[2]. That means, the angle γ must be as twice as

[2]In three-dimensional space target 1 must be hit four times as often as target 2.

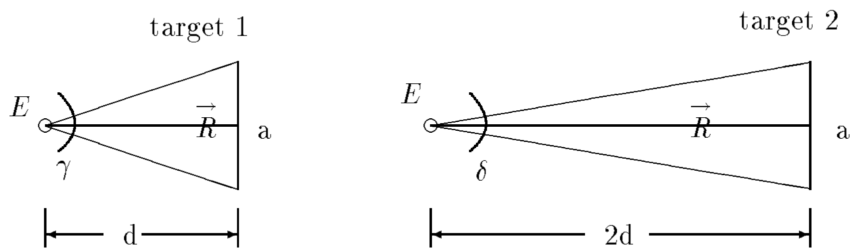

Figure 2: Geometry where target 1 should be hit twice as often as target 2

big as the angle δ ($\gamma \stackrel{?}{=} 2\delta$).

$$\gamma = 2 * arctan(\frac{a}{2d}) \qquad \delta = 2 * arctan(\frac{a}{4d})$$

$$arctan(\frac{a}{2d}) \stackrel{?}{=} 2 * arctan(\frac{a}{4d}) \tag{1}$$

The angles γ and δ are very small. Hence the illumination model of the emitter does not cause any flux density deviation within the frustum. Assuming this, equation(1) is correct if and only if the function $y = arctan(x)$ can be approximated by a linear function. Looking at the graph of the arctan-function it is evident, that this approximation may be done for big arguments x only. This forces a contradiction! For physically based illumination the distance between emitter and target squared must be large compared with the area of the target. Using the 'constant-portion-technique' the distance must be small. These two constraints are diametrically opposed. **As a result we get:** Any analytical illumination technique, i.e. a technique strongly based on physical laws and requirements, is not permitted to use constant portions. On the other hand the use of constant portions creates a new **empirical illumination technique**, which is not radiosity.

Picture 3 is computed using constant portions and a ray distribution according to the diffuse $cos(\phi)$. Comparing this with the result of the correct full-matrix solution (see picture 2C) it appears that:

- The lamp is not in the middle of the ceiling. Hence the yellow wall must be brighter at its right half as seen in picture 2C. In picture 3 the brightest area is shifted into the middle of the yellow wall. The same situation occurs for the green wall.

- The invisible wall right behind the observer is yellow too. In picture 2C the color bleeding from this wall and from the red wall onto the floor is noticeable. We can see color bleeding from the reddish wall to the ceiling too. The raycasting picture 3 shows a lack of all color bleeding effects. And especially this color bleeding is what radiosity became famous for! (remember the wooden sculpture used to show the advantages of radiosity over raytracing, see [CoWa93])

- The computation time grew up tremendously (50 minutes vs. 72 seconds by 3800 triangles). Every time a constant small portion is distributed into the environment; although it is theoretically possible to distribute a bigger one. A bigger portion is allowed, if the distance between emitter and target is small. The portion used in the picture is around ten up minus four. A bigger (constant) portion causes a spotted effect on the resulting image.

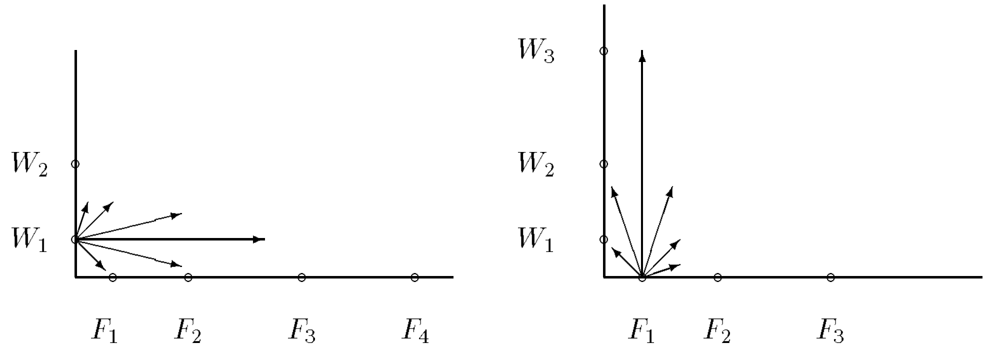

Figure 3: The biggest part of flux emitted from source-point W_1 will land on the nearest neighbour F_1. If F_1 shoots its flux into the environment, the biggest part will land on its nearest neighbour W_1. Hence a large amount of flux is shot back and forth between these two points. This results in bright edges.

Working on the analytical illumination field we must pay strong attention to the distance-dependency of all portions. Hence the Monte-Carlo raycasting used to produce picture 3 was modified to distribute r^2-dependent portions. The points lying along all edges are shifted according to figure 1. The final image of this approach does not look realistic. All edges are very bright, whereas the inner parts of the polygons are too dark [3]. The cause of this behaviour is illustrated in figure 3. Nearly 50% of all flux emitted from W_1 lands on the floor F. Due to the distance dependency of all portions the biggest part will land on F_1 and is gathered there. If F_1 distributes its flux, it is shot around F_1's normal. Hence nearly 50% will land on the adjacent wall, and the biggest part on the point W_1[4]. This process continues until no flux is left. Hence the majority of the scene light flux is distributed between the edge points. The main problem and reason for the bright edges is the distribution strategy of the random rays. The rays are distributed according to the illumination model of the emitter, i.e. for diffuse materials proportional to $cos(\phi_i)$. And ϕ is the angle between the emitters normal and the current ray direction. That means: The ray distribution is referred to the emitters normal always, no matter where the flux came from. This behaviour is derived out of Lamberts cosine-law and is responsible for the view independence of ideally diffuse illumination.

Let us make a short excursion into glossy surfaces here. The corresponding geometry is computed after Snellius' law of reflection. The incoming ray, the outgoing ray and the surface normal lie in one plane. The angle between the incoming ray and the normal is the same as that between the outgoing ray and the normal (see [Optics],[Feyn63]). It is obvious, that the rays 'carrying' the reflected flux back into the environment must be distributed within a frustum having an opening angle less than 180 degrees to achieve a glossy look. It is obvious too, that the 'reflection club' must be centered around the reflected ray, and **not** around the emitter's normal. Let the illumination model of the emitter be proportional to $cos^n(\phi)$ (Phong model, instead we can take a Microfacet model $e^{-\kappa\phi^2}$ or something completely different), whereby the reflection character is modified by the shiny-exponent n. For all n bigger than one we will get a glossy or - for big n

[3] see picture 4 in the file mentioned above.

[4] The real percentage of flux leaving F_1 landing on W_1 depends on the granularity of the tesselation, on the offset the 'edge- points' are shifted and on the surrounding geometry.

- a specular look. Vice versa this means: Downto $n = 1.05$ the reflection club is centered around the reflected ray, whereas for $n = 1.0$ in all common implementations it is centered around the normal. This does not make sense. Such a small change in the reflection behaviour cannot cause a fundamental change in the reference direction. The reason for this bug is a misinterpretation of Lamberts law. Lambert [5]wrote in 1760:

The 'light strength' of a radiative area is proportional to the cosine of the emitting angle[6].

Lambert did not say anything about the reflection behaviour, in fact he did not mention the word diffuse. It is obvious, that the law in the form above is not a general one. However looking at the time Lambert lived, people had candles and the sun as self-radiative areas only. And these two are diffuse. To ensure the generality of the Lambert law, physicists added two constraints later. First, the material must be diffuse. Second, and this was misunderstood: *If* the intensity along the emitter's normal is I_0, then the intensity viewed from an arbitrary angle is computed after $I_0 \cdot cos(\phi)$. It is assumed, that *in this special case* the normal is the reference direction. It is not postulated, that the surface's normal must be the reference direction always! If the maximal intensity I_0 can be seen along any other direction, than this is the reference direction to the Lambert law and to the ray distribution, respectively.

Hence the solution to the bright edges is easy: **All reflected flux must be distributed around the reflected ray computed after Snellius' law of reflection**[7]. Independently of reflection behaviour this must be done for all surfaces . Therefore only one algorithm is necessary; ideally diffuse, glossy and specular behaviour effects the shape of the 'reflection club' only. The more specular a surface, the more slender the reflection club (the reflection club as introduced in [Nic93] can be compared with the commonly used BRDF).
As mentioned above, the distribution around the surfaces normal is responsible for the view independence of diffuse illumination. Now the reference direction is the reflected ray instead of the normal. Nevertheless diffuse illumination *looks* view independent. This is caused by the 'law of big number'.

Results - Conclusions

i The distribution of constant portion within a Monte-Carlo raycasting conflicts with the demands of analytical illumnination. Two solutions are possible: First use this constant-portion-raycasting as a new empirical illumination technique. Second, modify Monte-Carlo raycasting into variable, i.e.'form-factor' dependend portions.

ii Variable portions cause bright edges, if the ray distribution is centered around the patchs normal. Solution: The main emission direction for all reflected flux is Snellius' reflection or refraction direction. Hence the same strategy and algorithm can and must be used for diffuse, glossy and specular materials.

iii The algorithm to solve the illumination problem concerning different material properties is of recursive type. That means, any single flux portion emitted by an active emitter is distributed further according to the 'form-factor' and the main emission direction.

[5]Johann Heinrich Lambert, 1728-1777

[6]Die Lichtstärke einer strahlenden Fläche ist dem Kosinus des Ausstrahlungswinkels proportional.

[7]And in addition to this the transmitted flux, if the material is transparent or translucent, must be distributed around the main transmision direction computed accordingly to Snellius' law of refraction.

The recursion strategy provides us from storing a three-dimensional 'form-factor'-matrix. (remember Immel in '86)

iv Although the reference direction is no longer the patchs normal, diffuse illumination looks view-independent. This is caused by the 'law of big number'.

v The analytical 'variable-portion-technique' saves a lot of computing time compared with the 'constantly' Monte- Carlo raycasting. For complex environments a single portion can be proportional to the 'form-factor', i.e. the 'form- factor' is multiplied with a constant C bigger than one to get the resulting portion. Hence Shirleys observation of the O(N) complexity is still true. Some more investigation is necessary to determine the constant C.

Literature

[Baum89] D.R.Baum, H.E.Rushmeier, J.M.Winget
Improving Radiosity Solutions Through the Use of Analytical Determined Form-Factors
Computer Graphics Vol23(3) July1989

[Baum91] D.R.Baum, S.Mann, K.P.Smith, J.M.Winget
Making Radiosity Usable: Automatic Preprocessing and Meshing techniques for the Generation of Accurate Radiosity Solutions, Computer Graphics, Vol 25(4) July1991

[CoWa93] M.F.Cohen, J.R.Wallace
Radiosity and Realistic Image Synthesis, 1993 by Academic Press,Inc.

[Feyn63] R.P.Feynman, R.B.Leighton, M.Sands
Lectures on Physics, Volume I
Addison-Wesley Publishing Company, Reading, MA 1963

[Heck92] P.Heckbert
Radiosity in Flatland, EG92 in Cambridge

[Kok92] A.J.F.Kok, C.Yilmaz, L.H.J.Bierens
A Two-Pass Radiosity Method for Bezier Patches
in K.Bouatouch, C.Bouville 'Photorealism in Computer Graphics', Berlin 1992

[Linke93] D.Linke
Monte Carlo-Radiosity - Test verschiedener Verteilungsfunktionen
Diplomarbeit Universität Rostock, Mai 1993

[Nic93] N.Günther
Analytical illumination - A physical based approach
preprint Universität Rostock

[Optics] D.Malacara (editor)
Methods Of Experimental Physics, Vol. 26, Physical Optics and Light Measurements, Academic Press Inc. 1988

[Pueyo] X.Pueyo
Diffuse interreflections. Techniques for form-factor computation: a survey
Visual Computer Vol7(4) July1991

[Shir91] P.Shirley
Time Complexity of Monte Carlo Radiosity, EG91 in Wien

[Siegel] R.Siegel, J.R.Howell
Thermal Radiation Heat Transfer

1972,1981 by Hemisphere Publishing Corp., New York

[Sill89] F.Sillion, C.Puech
A General Two-Pass Method Integrating Specular and Diffuse Reflection
Computer Graphics Vol23(3) July1989

[Sill91] F.X.Sillion, J.R.Arvo, S.H.Westin, D.P.Greenberg
A Global Illumination Solution for General Reflectance Distributions
Computer Graphics Vol25(4) July1991

[Wall89] J.R.Wallace, M.F.Cohen, D.P.Greenberg
A Two-Pass Solution To The Rendering Equation:
A Synthesis Of Ray Tracing And Radiosity Methods
Computer Grahpics, Vol 21(4), July1987

contact address:

Nico Günther
Universität Rostock
Fachbereich Informatik
18051 Rostock
GERMANY

email: nic@informatik.uni-rostock.de