

Cálculo do Fator-de-Forma exato entre Áreas Diferencial e Finita Usando CSG

por

Isaac Moreira Barreto

Dissertação apresentada ao
Mestrado em Ciência da Computação
Universidade Federal do Ceará

Fortaleza, Ceará, Brasil 2008.

Orientador: Creto Augusto Vidal, PhD

©Isaac Moreira Barreto, 2008

Resumo

Os métodos de Ray-Tracing e Radiosidade são os principais representantes dos métodos existentes para resolver o problema de iluminação global. Em ambos os métodos se faz necessário saber a taxa de transferência de energia luminosa entre duas áreas. Essa taxa de transferência, chamada de fator-de-forma, é um dos pontos principais no método de Radiosidade e vem sendo usado cada vez com mais frequência em métodos de Ray-Tracing com fontes luminosas de área finita. Existem vários métodos para o cálculo do fator-de-forma, a maioria deles são aproximativos por uma questão de desempenho. Porém, em casos específicos, o trabalho extra para calcular o valor exato do fator-de-forma pode melhorar o desempenho global do método. Em geral, nesses casos, o esforço necessário para se obter uma aproximação aceitável do valor do fator-de-forma supera o esforço necessário para calcular o valor exato em si. Além disso, existem situações, tais como a renderização nas áreas de fronteiras de sombras, em que uma alta precisão é mais importante do que um ganho no desempenho. Nessas situações, é desejável que o método tenha ao seu dispor uma maneira de calcular o valor exato do fator-de-forma. Neste trabalho é apresentado um método para calcular o fator-de-forma exato entre uma área finita e uma área diferencial que utiliza de técnicas CSG para identificar as áreas ocluídas do polígono emissor.

Palavras-chave: Fator-de-Forma, Ray-Tracing, Radiosidade, Iluminação Global, Renderização.

Abstract

The Ray-Tracing and Radiosity methods are the main representatives of the methods that solve the global illumination problem. In both methods it is necessary to know the energy transfer ratio between two areas. This ratio, called form factor, is one of the key concepts in Radiosity methods and is being more frequently used in Ray-Tracing methods with finite area light sources. There are many methods for the computation of the form factor, most of them are approximative due to a matter of performance, but, in some specific cases, the extra computational effort needed to compute the exact value of the form factor can improve the overall performance of the illumination method. In general, in these cases, the computational effort needed to obtain an acceptable approximation of the form factor outweighs the effort necessary to compute the exact value. Furthermore there are situations, for example, shadow boundary shading, in which a high precision is far more important than a performance gain. In this work we present a method to compute the exact form factor between a finite area and a differential area which uses CSG techniques to identify the occluded areas of the source.

Key words: Form Factor, Ray-Tracing, Radiosity, Global Illumination, Rendering.

Sumário

1	Introdução	1
1.1	Motivação	2
1.2	Objetivos	2
1.3	Organização do trabalho	3
2	Trabalhos Relacionados	4
2.1	Introdução	4
2.2	Métodos Aproximativos	5
2.2.1	Amostragem de Área	6
2.2.2	Amostragem na Semi-Esfera	8
2.3	Métodos Exatos	10
2.3.1	Malha de Descontinuidade	11
2.3.2	Outros Métodos	13
2.4	Considerações Finais	15
3	Descrição do Método	16
3.1	Introdução	16
3.2	Teste de Visibilidade	17

3.3	Projeção e Montagem da CSG	18
3.4	Conversão CSG para B-REP	22
3.5	Cálculo do Fator-de-Forma	27
3.6	Considerações Finais	30
4	Resultados Experimentais	31
4.1	Introdução	31
4.2	Emissão	32
4.3	Emissão com Obstrução	33
4.4	Sombras Complexas	34
4.5	Oclusão por Pequenas Áreas	35
4.6	Oclusão por Polígono Vazado	36
4.7	Considerações Finais	38
5	Conclusões	39
A	Background	46
A.1	Radiometria	46
A.2	Radiosidade Clássica	50
A.3	Form Factors	53

Lista de Tabelas

3.1	Combinações de classificações de intervalos	25
4.1	Resultados quantitativos da cena de Emissão.	32
4.2	Resultados quantitativos da cena de Emissão com Obstrução.	33
4.3	Resultados quantitativos da cena de Sombras Complexas.	35
4.4	Resultados quantitativos da cena de Oclusão por Pequenas Áreas.	36
4.5	Resultados quantitativos da cena de Oclusão por Polígono Vazado.	37

Lista de Figuras

2.1	Fator-de-Forma entre dA_i e um disco arbitrário.	5
2.2	Geometria do fator-de-forma proposto por Wallace et al. [27].	6
2.3	Analogia de Nusselt.	8
2.4	Semi-cubo.	9
2.5	Geometria da fator-de-forma entre dA_i e um polígono.	11
2.6	Malha de descontinuidade gerada devido à oclusão [12].	12
3.1	Frustrum formado entre dA_i e A_j com um patch k interposto.	17
3.2	Projeção do <i>patch</i> k no plano suporte de A_j	18
3.3	Sistema de coordenadas usado na conversão bidimensional.	19
3.4	Árvores CSG com nodos de união. À esquerda, um polígono representado por uma árvore desbalanceada. À direita, o mesmo polígono representado por uma árvore balanceada.	21
3.5	Uma árvore CSG e a representação dos polígonos que a formam. As arestas em vermelho não são arestas candidatas.	23
3.6	Classificação de duas arestas contra um polígono.	24
3.7	Instância de entrada do Algoritmo 1.	27
3.8	Áreas positivas e negativas de um polígono.	28

4.1	Renderização da cena de Emissão.	32
4.2	Renderização da cena de Emissão com Obstrução.	33
4.3	Renderização da cena de Sombras Complexas.	34
4.4	Renderização da cena de Oclusão por Pequenas Áreas.	36
4.5	Renderização da cena de Oclusão por Polígono Vazado.	37
A.1	Geometria da radiância.	47
A.2	Irradiância e Radiosidade.	48
A.3	Geometria da BRDF.	49
A.4	Geometria do Fator-de-Forma entre dA_i e dA_j	54

Introdução

O cálculo da taxa de energia radiante que é transferida entre duas superfícies é um problema importante que é abordado nas aplicações de renderização que utilizam modelos de iluminação global, como aplicações de Radiosidade e *Ray-Tracing*.

Esse problema foi abordado primeiramente em trabalhos sobre a transferência de energia térmica em ambientes fechados [22]. Goral et al. [10] e Nishita & Nakamae [20] foram os primeiros a utilizar essas idéias no contexto de iluminação.

Desde então, o cálculo do fator-de-forma, nome com o qual foi batizada a taxa de transferência de energia radiante entre duas superfícies, tem sido um tópico de pesquisa ativo na área de iluminação global.

O valor exato do fator-de-forma entre duas superfícies arbitrárias pode ser muito difícil de obter devido à inúmeras complicações, como oclusões causadas por superfícies opacas ou translúcidas. Por esse motivo, muitos dos métodos existentes para o cálculo do fator-de-forma obtêm na verdade um valor aproximado.

1.1. Motivação

As pesquisas feitas acerca do cálculo do fator-de-forma têm em geral uma característica em comum: tentar encontrar o equilíbrio entre desempenho e precisão. Como mencionado anteriormente, o cálculo do fator-de-forma entre duas superfícies arbitrarias não é trivial e, em muitos casos, achar o valor exato pode não ser aconselhável devido ao impacto considerável no desempenho do método.

Porém, em casos específicos, o trabalho extra para calcular o valor exato do fator-de-forma pode melhorar o desempenho global do método. Em geral, nesses casos, o esforço necessário para se obter uma aproximação aceitável do valor do fator-de-forma supera o esforço necessário para calcular o valor exato em si. Além disso, existem situações, tais como a renderização nas áreas de fronteiras de sombras, em que uma alta precisão é mais importante do que um ganho no desempenho. Nessas situações, é desejável que o método tenha ao seu dispor uma maneira de calcular o valor exato do fator-de-forma.

O fator-de-forma entre uma área poligonal finita e uma área diferencial é amplamente utilizado em abordagens de Radiosidade Progressiva assim como em métodos de *Ray-Tracing*.

1.2. Objetivos

O objetivo deste trabalho é desenvolver um método para o cálculo do valor exato do fator-de-forma entre uma área poligonal finita e uma área diferencial que obtenha um desempenho aceitável quando comparado com métodos existentes para o mesmo cálculo do fator-de-forma.

1.3. Organização do trabalho

O restante do trabalho está organizado da seguinte forma: estudo das soluções propostas por outros trabalhos, Capítulo 2; apresentação detalhada da solução desenvolvida por este trabalho, Capítulo 3; apresentação dos testes comparativos feitos com as soluções existentes para validar a solução proposta, Capítulo 4; as conclusões do trabalho e as sugestões para melhorias e trabalhos futuros são apresentadas, Capítulo 5.

Trabalhos Relacionados

2.1. Introdução

Neste capítulo são apresentados os métodos mais utilizados para o cálculo do fator-de-forma seguindo uma classificação baseada primariamente na precisão da solução e, em seguida, na técnica utilizada. De acordo com a precisão da solução, os métodos podem ser classificados como aproximativos ou exatos. Os métodos aproximativos dividem-se entre métodos que utilizam amostragem de área e métodos que utilizam amostragem na semi-esfera. Por sua vez, os métodos exatos são, em sua grande maioria, métodos que utilizam malha de descontinuidade.

Este trabalho propõe um novo método para o cálculo do fator-de-forma entre uma área diferencial e uma área finita (Equação A.20). Portanto, os métodos apresentados neste capítulo estão restritos, em sua maioria, ao cálculo desse fator-de-forma particular. Com o intuito de facilitar a compreensão dos métodos apresentados a seguir, a equação do fator-de-forma entre uma área diferencial dA_i e uma área finita A_j é repetida aqui:

$$F_{dA_i-A_j} = \int_{A_j} \frac{\cos\theta_i \cos\theta_j}{\pi r^2} V(dA_i, dA_j) dA_j. \quad (2.1)$$

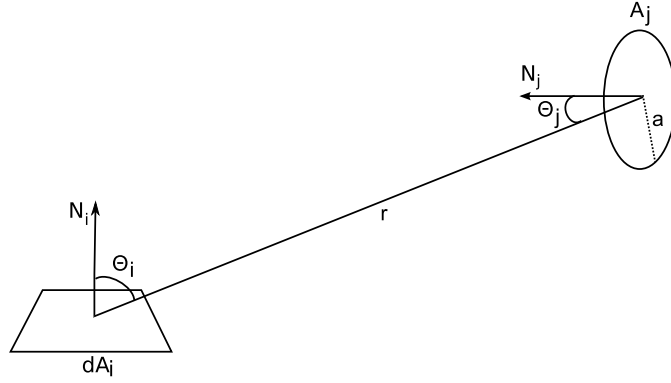


Figura 2.1. Fator-de-Forma entre dA_i e um disco arbitrário.

2.2. Métodos Aproximativos

Os métodos aproximativos para o cálculo do fator-de-forma utilizam-se de artifícios para aproximar o valor da integral na equação 2.1. A diferença entre esses métodos é a técnica utilizada para obter essa aproximação, logo é possível fazer uma subclassificação dos métodos aproximativos em: métodos que utilizam amostragem em área e métodos que utilizam amostragem na semi-esfera.

Contudo, o método aproximativo mais simples para calcular o valor do fator-de-forma em questão não utiliza nenhuma das duas abordagens mencionadas acima. Ele consiste em aproximar A_j à uma forma geométrica simples para a qual o fator-de-forma seja conhecido, como por exemplo um disco (vide Figura 2.1). No caso específico do disco, que é a forma geométrica mais comumente utilizada nessa abordagem, a fórmula para o cálculo do fator-de-forma entre dA_i e A_j é dada por

$$F_{dA_i-A_j_{disco}} = \frac{a^2 \cos\theta_i \cos\theta_j}{r^2 + a^2}. \quad (2.2)$$

No entanto, esse método de cálculo do fator-de-forma é inadequado para a maioria das situações devido à dois motivos principais: 1) O disco nem sempre aproxima bem um polígono e, nessas situações, o erro do fator-de-forma é inaceitável; e 2) O fator de

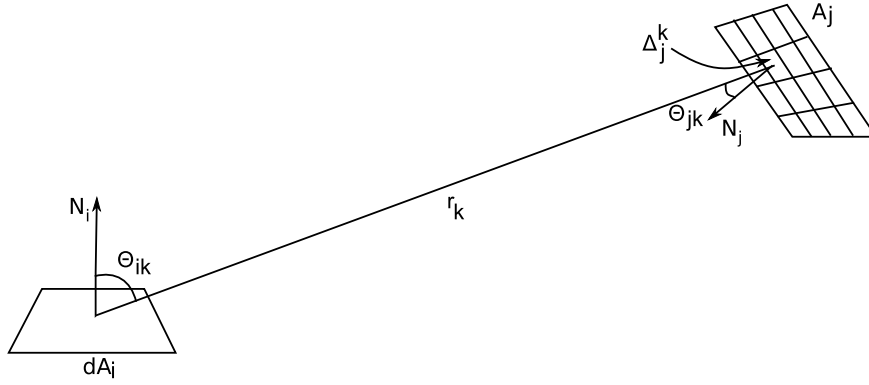


Figura 2.2. Geometria do fator-de-forma proposto por Wallace et al. [27].

visibilidade não é levado em consideração.

2.2.1. Amostragem de Área

A característica comum entre todos os métodos de amostragem de área é que eles tentam minimizar os erros causados pela aproximação da geometria e do fator de visibilidade, dividindo a área A_j em áreas menores.

O primeiro método a utilizar essa técnica foi proposto por Wallace et al.[27] em 1989. Nesse método A_j é subdividida em um conjunto de n áreas menores, ΔA_j^k , para que em seguida, $F_{dA_i-A_j}$ possa ser calculado como um somatório dos fatores-de-forma entre dA_i e n discos que aproximam a geometria das áreas ΔA_j^k .

O fator de visibilidade de cada um dos termos do somatório é aproximado verificando a visibilidade entre dA_i e o centro de cada uma das áreas ΔA_j^k . Assim como ocorre com a maioria dos métodos de amostragem de área, a precisão do fator de visibilidade geral depende do nível de subdivisão de A_j .

A geometria da solução proposta pode ser verificada na Figura 2.2. Embora o trabalho de Wallace deduza uma equação para o fator-de-forma $F_{A_j-dA_i}$, o fator-de-forma $F_{dA_i-A_j}$ pode ser obtido através da propriedade de reciprocidade dos fatores-de-forma. Assim,

$$F_{dA_i-A_j} = \frac{A_j}{dA_i} F_{A_j-dA_i} = \frac{A_j}{n} \sum_{k=1}^n V(dA_i, \Delta A_j^k) \frac{\cos\theta_{ik} \cos\theta_{jk}}{\pi r_k^2 + \frac{A_j}{n}} \quad (2.3)$$

A aproximação das ΔA_j^k áreas por discos é falha nos casos em que a distância r_k é pequena quando comparada com a área de ΔA_j^k e, nesses casos, podem surgir artefatos na iluminação, como *aliasing*. Além disso, a aproximação do fator de visibilidade feita ao verificar a visibilidade entre dA_i e os centros das áreas ΔA_j^k áreas pode gerar fronteiras de sombras não suaves.

Em [27], o problema causado pela violação do critério de proximidade é abordado através de um método de subdivisão adaptativa de A_j no qual, enquanto o valor da área de ΔA_j^k for grande quando comparado com a distância r_k , a área ΔA_j^k é recursivamente subdividida até um limite previamente estabelecido. Tampieri [26] apresentou um estudo sobre a subdivisão adaptativa proposta por Wallace et al. [27].

Em 1993, Pietrek [21] apresentou um método similar ao proposto por Wallace, mas que calcula analiticamente o fator-de-forma entre dA_i e a área ΔA_j^k que viola o critério de aproximação utilizando a equação 2.4.

Wang et al. [16] propõem um método de subdivisão da malha e cálculo do fator-de-forma no contexto de radiosidade progressiva. O cálculo do fator-de-forma é semelhante ao proposto por Pietrek, mas difere deste no critério de subdivisão de A_j . Além de considerar o critério de proximidade, o método proposto por Wang tenta minimizar os erros de visibilidade utilizando um índice de visibilidade [24] entre dA_i e ΔA_j^k . A área ΔA_j^k é recursivamente subdividida caso não seja completamente visível ou ocluída a partir de dA_i , ou seja, caso o índice de visibilidade entre as duas áreas seja diferente de 0 ou 1.

Os métodos apresentados nesta subseção são os principais representantes dos métodos que utilizam amostragem de área. Na próxima subseção, são apresentados os principais métodos que utilizam amostragem na semi-esfera centrada em dA_i .

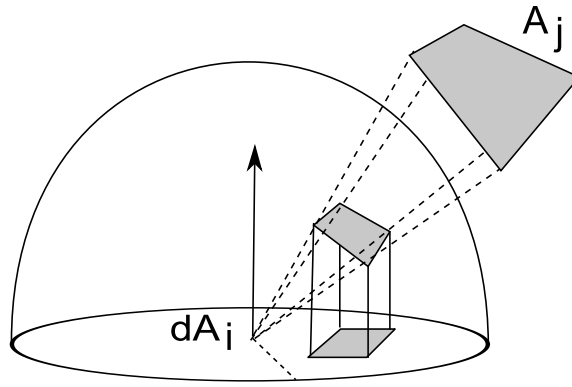


Figura 2.3. Analogia de Nusselt.

2.2.2. Amostragem na Semi-Esfera

Os métodos baseados em amostragem na semi-esfera são úteis quando é necessário calcular o fator-de-forma entre dA_i e todas as superfícies da cena. Uma das vantagens desses métodos é que o cálculo do fator de visibilidade entre dA_i e A_j está implícito no processo.

A analogia de Nusselt (Figura 2.3) está na base da técnica utilizada por esses métodos. Segundo essa analogia, o fator-de-forma entre dA_i e A_j pode ser encontrado do seguinte modo: 1) uma semi-esfera de raio unitário é posicionada com o centro em dA_i ; 2) A_j é projetada na semi-esfera; 3) a projeção de A_j na semi-esfera é projetada ortogonalmente na base da semi-esfera; 4) a fração entre a área duplamente projetada de A_j e a área da base da semi-esfera é o fator-de-forma entre dA_i e A_j .

Em 1985, Cohen et al. [3] propuseram um método inovador para o cálculo dos fatores-de-forma, o método do *Hemi-cube*. Esse método utiliza a analogia de Nusselt como passo para mostrar que áreas que têm a mesma área projetada em qualquer superfície intermediária possuem o mesmo fator-de-forma.

Usando esse conhecimento, Cohen substituiu a semi-esfera por um semi-cubo centrado em dA_i e discretizado em pequenas células (Figura 2.4), chamadas Δ -áreas. O fator-de-

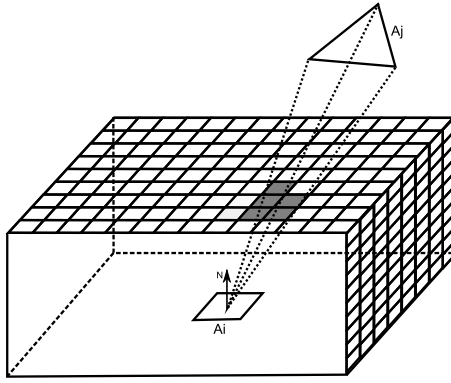


Figura 2.4. *Semi-cubo.*

forma entre dA_i e cada uma dessas Δ -áreas é previamente conhecido e de fácil cálculo, devido à geometria restrita do semi-cubo. Portanto, utilizando-se do semi-cubo, o problema de calcular o fator-de-forma $F_{dA_i-A_j}$ é reduzido a identificar as células do semi-cubo em que A_j é projetado, e somar seus Δ -fatores-de-forma.

Considerando-se uma câmera posicionada no centro de dA_i , cada face do semi-cubo representa um plano de projeção em perspectiva, portanto, o problema de visibilidade pode ser resolvido por algoritmos de superfícies escondidas. O método *Z-buffer* [9] é bastante conhecido por sua simplicidade e eficácia, além de ser implementado em placas gráficas.

O método *Hemi-cube*, apesar das vantagens descritas acima, também possui algumas desvantagens. Primeiro, a precisão do método depende diretamente do nível de discretização das superfícies do semi-cubo, logo, para ser capaz de capturar mudanças sutis no fator-de-forma, como as causadas por pequenos oclusores, o número de Δ -áreas precisa ser muito grande, o que causa um grande impacto no desempenho do método. Segundo, o método sofre do problema comum à todos os algoritmos aproximativos: ele é falho quando o critério de proximidade é violado. Um estudo detalhado dos defeitos desse método é feito por Baum et al. [1].

Em seu trabalho, Baum e seus co-autores [1] propõem uma solução para os problemas

inerentes ao *Hemi-cube*. Baum sugere que o *Hemi-cube* seja mantido nos casos em que os critérios de proximidade não sejam violados, caso contrário, o *Hemi-cube* é mantido apenas para o cálculo do fator de visibilidade. O fator-de-forma é calculado entre dA_i e as subáreas de A_j que são projetadas em alguma Δ -área do semi-cubo, utilizando o método analítico demonstrado na equação 2.4.

Em 1989, Sillion et al. [23] desenvolveram um método que resolve o problema da inflexibilidade do semi-cubo no que diz respeito à subdivisão uniforme de suas faces. Sillion substitui o semi-cubo por um simples plano que situa-se imediatamente acima de dA_i . O método é capaz de subdividir o plano em níveis mais elevados de acordo com o tamanho das áreas a serem testadas.

Malley [18] propôs um método que utiliza Ray-tracing de Monte Carlo. Inicialmente o método gera um conjunto de pontos randômicos situados na base da semi-esfera ao redor de dA_i . Esses pontos são considerados projeções ortogonais de pontos da superfície da semi-esfera sobre o plano de dA_i . Em seguida, lançam-se raios a partir de dA_i passando por cada ponto marcado na superfície da semi-esfera e as intersecções desses raios com os elementos da cena são calculadas. Para determinar o fator-de-forma entre dA_i e A_j , o número de raios que interceptam A_j é dividido pelo número total de raios lançados.

2.3. Métodos Exatos

Os métodos exatos para o cálculo de fatores-de-forma entre dA_i e A_j não utilizam nenhuma aproximação na geometria de A_j e, em geral, calculam o fator de visibilidade exato entre dA_i e A_j . Muitas vezes esse fator de visibilidade não é explicitamente calculado, mas faz parte implicitamente do método de cálculo do fator-de-forma.

A fórmula analítica para o fator-de-forma entre uma área diferencial dA_i e um polígono é particularmente importante, uma vez que a maioria das cenas são compostas de modelos

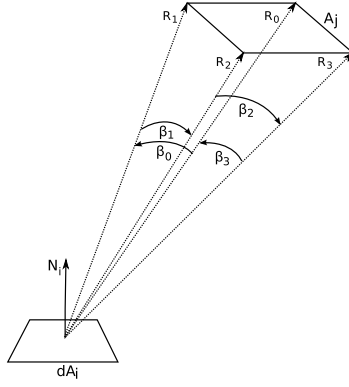


Figura 2.5. Geometria da fator-de-forma entre dA_i e um polígono.

poligonais. A equação para esse fator-de-forma é escrita como [27]

$$F_{dA_i-A_j} = \frac{1}{2\pi} \sum_{k=1}^n \beta_k N_i \cdot \overline{(R_k \times R_{(k+1)\%n})} \quad (2.4)$$

onde: n é o número de arestas de A_j ; β_k é o ângulo entre R_k e $R_{(k+1)\%n}$; N_i é a normal em dA_i ; R_k é o vetor entre dA_i e o k -ésimo vértice de A_j . A Figura 2.5 mostra a geometria envolvida nesse fator-de-forma.

Essa fórmula calcula o fator-de-forma exato entre dA_i e A_j sem considerar oclusões. Contudo, com o algoritmo apropriado de visibilidade, ela pode ser utilizada para calcular o fator-de-forma exato entre dA_i e A_j .

Nas próximas subseções, são apresentados métodos que calculam o fator-de-forma exato entre dA_i e A_j levando em consideração a oclusão. A Seção 2.3.1 apresenta métodos que são empregados, em geral, em técnicas que utilizam malhas de descontinuidade. A Seção 2.3.2 apresenta outros métodos independentes.

2.3.1. Malha de Descontinuidade

Uma malha de descontinuidade representando uma superfície A_j resolve todas as curvas críticas, ou descontinuidades, da função em questão ao longo da superfície. No con-

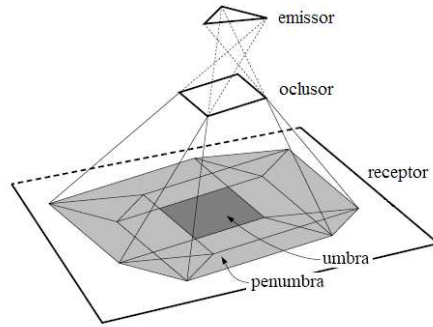


Figura 2.6. Malha de descontinuidade gerada devido à oclusão [12].

texto de radiosidade a função considerada é a função de iluminação e as curvas críticas dessa função são causadas por eventos visuais relacionados às oclusões (Figura 2.6). Cada curva crítica é representada nessa malha como uma ou mais arestas que são indexadas pelo seu grau de descontinuidade. A principal vantagem dessa família de métodos é que uma vez que as descontinuidades ocorrem nas arestas da malha, a função de iluminação dentro de cada elemento é bem comportada e pode, na maioria dos casos, ser capturada com uma quantidade pequena de amostras.

Lischinski et al. [17] apresentaram no seu trabalho sobre malhas de descontinuidade um método para o cálculo exato do fator-de-forma. Nesse trabalho, as malhas de descontinuidade, que são primariamente utilizadas para a subdivisão da malha, também são utilizadas para identificar as áreas visíveis e ocluídas das áreas emissoras através de um algoritmo de *clipping* semelhante ao utilizado em [19].

Em 1994, Drettakis et al. [6] apresentaram uma estrutura de dados, denominada por eles de *backprojection*, que auxilia no cálculo do fator-de-forma. Essa estrutura aproveita métodos utilizados em algoritmos de malhas de descontinuidade para construir um mapa de visibilidade entre dA_i e os elementos emissores da cena. Uma instância dessa estrutura de dados entre dA_i e um elemento emissor da cena, digamos A_j , contém qualquer informação de visibilidade entre eles, inclusive as partes visíveis de A_j . Essas partes podem

ser usadas em conjunto com a Equação 2.4 para calcular o fator-de-forma exato.

O esqueleto de visibilidade é uma estrutura de dados introduzida por Durand et al. [7] com o mesmo conceito da estrutura proposta por Drettakis. Porém, essa estrutura não é restrita à visibilidades entre pontos e polígonos ou entre pontos e elementos emissores de energia. Qualquer informação de visibilidade desejada, como a porcentagem de visibilidade entre dois polígonos quaisquer, está codificada nessa estrutura e pode ser recuperada trivialmente.

Durand et al. [8] utilizam o esqueleto de visibilidade como uma estrutura auxiliar no seu método de Radiosidade. A estrutura é usada em duas ocasiões: 1) na definição das fronteiras de sombra para a geração da malha de descontinuidade; 2) no cálculo dos fatores-de-forma. Assim como a estrutura *backprojection*, o esqueleto de visibilidade pode ser utilizado para identificar as áreas visíveis de um polígono quando visto de dA_i .

Outros métodos utilizam suas próprias variações dos métodos apresentados nesta seção (o leitor é direcionado a [5, 13, 15] para maiores detalhes).

2.3.2. Outros Métodos

Em 1999, Hart et al. [11] desenvolveram um método bastante similar ao método proposto por Drettakis et al. [6]. No seu trabalho, Hart utiliza a coerência entre as *scanlines* utilizadas no *clipping* dos polígonos, para desenvolver uma estrutura de dados, o mapa de oclusores. Essa estrutura de dados é utilizada para evitar que polígonos não oclusores sejam considerados durante o *clipping*. A diminuição do número de polígonos envolvidos afeta diretamente o desempenho dos algoritmos de *clipping* e, conseqüentemente, o desempenho do algoritmo proposto.

Geralmente, os métodos exatos para o cálculo de $F_{dA_i-A_j}$ utilizam abordagens baseadas nas arestas de A_j (Equação 2.4). Stark et al. [25] utilizam uma nova abordagem para solucionar o problema. No seu trabalho, uma nova derivação da Equação 2.4 é feita de

forma que $F_{dA_i-A_j}$ possa ser calculado em função dos vértices de A_j . Assim, no lugar de identificar as áreas visíveis de A_j , o problema é reduzido à identificação dos vértices visíveis.

Por último, Charneau et al. [2] utilizam o espaço Plücker como uma ferramenta para calcular $F_{dA_i-A_j}$. O espaço Plücker fornece uma maneira de representar conjuntos de linhas direcionadas no R^3 como vetores no R^6 . Essa técnica, em conjunto com técnicas de CSG, é utilizada para identificar as áreas visíveis de A_j . Uma vez identificadas essas áreas, o método procede como a maioria dos métodos exatos, utilizando a Equação 2.4 entre dA_i e as áreas visíveis de A_j , para calcular o $F_{dA_i-A_j}$ final.

2.4. Considerações Finais

Neste capítulo foram apresentados os métodos aproximativos e exatos mais utilizados para o cálculo do fator-de-forma. Em geral, os métodos exatos diferem dos métodos aproximativos por identificarem precisamente as áreas visíveis de um polígono emissor, e assim podem utilizar a fórmula analítica (Equação 2.4) para calcular o valor do fator-de-forma.

Os métodos aproximativos são mais adequados em situações em que uma aproximação adequada do valor exato do fator-de-forma possa ser atingida de maneira simples, ou seja, sem necessidade de muito esforço extra. Por sua vez, em situações em que o valor do fator-de-forma precise ter uma precisão muito alta os métodos exatos têm um melhor desempenho.

Descrição do Método

3.1. Introdução

Neste capítulo, é apresentado em detalhes o método proposto para o cálculo do fator-de-forma entre um área diferencial dA_i e uma área finita A_j . O método utiliza conceitos de modelagem de sólidos, como *B-REP* (representação por fronteira) e *CSG* (geometria de construção de sólidos), além de algumas técnicas de projeção e identificação de polígonos oclusores, comumente utilizadas em algoritmos de sombras.

A apresentação do método segue uma ordem estabelecida de forma a facilitar a compreensão geral, assim, o restante do capítulo está organizado da seguinte forma: teste de visibilidade, que tem como propósito classificar a relação de visibilidade entre dA_i e A_j , assim como identificar os polígonos oclusores, Seção 3.2; descrição do método utilizado para a projeção dos polígonos, assim como definição da árvore *CSG* utilizada, Seção 3.3; descrição detalhada do método de conversão entre *CSG* e *B-REP*, Seção 3.4; definição do método utilizado para o cálculo do fator-de-forma, Seção 3.5; conclusão e considerações finais, Seção 3.6.

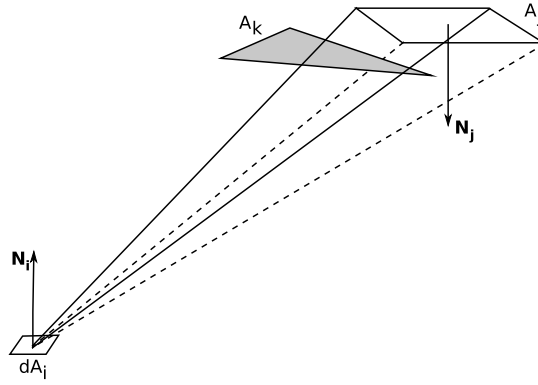


Figura 3.1. Frustrum formado entre dA_i e A_j com um patch k interposto.

3.2. Teste de Visibilidade

O primeiro passo a ser executado é o teste de visibilidade. Dadas uma área diferencial dA_i e uma área finita A_j , o teste de visibilidade entre essas duas áreas responde à seguinte pergunta: "Qual a visibilidade de A_j ao ser vista de dA_i ?". O teste de visibilidade retorna um dos três possíveis resultados: nenhuma, parcial ou total.

No contexto de Radiação, caso dA_i esteja atrás de A_j , ou vice-versa, a visibilidade entre essas duas áreas é nenhuma, portanto, o primeiro passo é identificar as posições relativas entre dA_i e A_j . Como mencionado no Capítulo 1, a cena consiste de polígonos simples (polígonos cujas arestas não se interceptam) e não interpostos (polígonos que não se interceptam). Assim, para efeitos de determinação das posições relativas entre dA_i e A_j , é possível, sem perda de generalidade, considerar A_j como uma área diferencial, dA_j , formada pelo centro de A_j e a sua normal.

O plano de suporte de dA_i , Ω_{dA_i} , em conjunto com sua normal, definem um semi-espaço, e, para que possa existir visibilidade entre as áreas, dA_j deve estar contida nesse semi-espaço. A recíproca também deve ser verdade. Para verificar se um ponto P pertence a um semi-espaço gerado por um plano Ω e uma normal \vec{N}_Ω é utilizada a equação a seguir

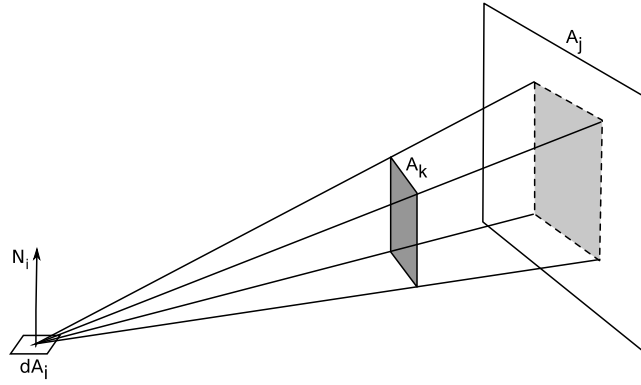


Figura 3.2. *Projeção do patch k no plano suporte de A_j.*

$$\vec{N}_\Omega \cdot (\vec{P} - \vec{P}_\Omega) > 0. \quad (3.1)$$

onde P_Ω é um ponto que se encontra no plano Ω .

Uma vez identificado que pode existir uma relação de visibilidade diferente de zero entre dA_i e A_j , é necessário descobrir se a visibilidade de fato existe, e, caso exista, se ela é total ou parcial. Com esse intuito, cada *patch k* da cena é testado contra o *frustum* formado por dA_i e A_j (Figura 3.1). A idéia é utilizar a equação 3.1 para testar cada vértice de k com os semi-espacos formado por dA_i , A_j e os 4 planos que formam o frustum. Caso um *patch* oclua totalmente A_j o teste de visibilidade termina e retorna visibilidade zero. Caso contrário, os *patches* que causam oclusão parcial são guardados em uma lista T . Ao final do teste, é retornado visibilidade total, caso o tamanho de T seja zero, e visibilidade parcial caso contrário.

3.3. Projeção e Montagem da CSG

As próximas etapas do método consistem na projeção dos polígonos contidos na lista T no plano de suporte de A_j e na montagem da árvore CSG.

Para a projeção de um *patch k*, o algoritmo procede de maneira simples: cada vértice

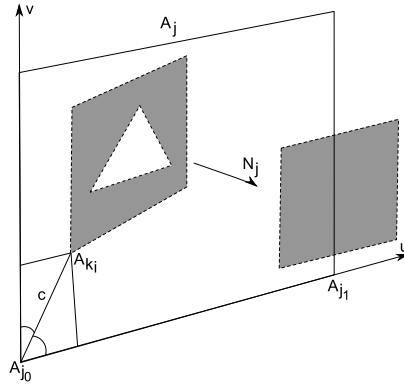


Figura 3.3. Sistema de coordenadas usado na conversão bidimensional.

de k é visitado em sentido anti-horário e um raio é lançado partindo de dA_i em direção ao vértice visitado. A intersecção entre esse raio e o plano suporte de A_j gera um vértice do polígono projetado (Figura 3.2). Ao final da projeção de todos os vértices de k , a lista dos vértices projetados, na mesma ordem em que foram percorridos, formam o polígono projetado Ψ_k .

Se o raio for representado de forma paramétrica, a intersecção entre o raio e um plano Ω pode ser encontrada pela seguinte equação

$$t = \frac{(N_\Omega \cdot P_\Omega) - (N_\Omega \cdot O)}{N_\Omega \cdot D} \quad (3.2)$$

onde N_Ω é o vetor unitário normal ao plano, P_Ω é o vetor-posição de um ponto pertencente ao plano, O é o vetor-posição da origem do raio, \vec{D} é a direção do raio e t é o parâmetro de interpolação do raio.

Após a projeção de todos os polígonos contidos na lista T , é gerada uma nova lista de polígonos planares que têm em comum o mesmo plano de suporte, o plano de A_j . Esses polígonos, por estarem contidos no mesmo plano de suporte, podem ser convertidos para uma representação bi-dimensional, utilizando um sistema de coordenadas comum baseado nas coordenadas dos vértices de A_j (Figura 3.3), como mostrado a seguir:

$$\mathbf{u} = A_{j_0} \xrightarrow{\quad} A_{j_1} \quad (3.3)$$

$$\mathbf{v} = \mathbf{N}_{A_j} \times \mathbf{u} \quad (3.4)$$

A conversão dos vértices de um polígono Ψ_k para esse sistema de coordenadas, gera um novo polígono Φ_k , e cada vértice desse polígono pode ser encontrado da seguinte forma:

$$\Phi_{k_i} = \langle \mathbf{p} \cdot \mathbf{u}, \mathbf{p} \cdot \mathbf{v} \rangle \quad (3.5)$$

onde \mathbf{p} é o vetor entre A_{j_0} e o vértice Ψ_{k_i} .

Uma vez que todos os polígonos contidos em T tenham sido projetados e convertidos para o sistema de coordenadas bi-dimensional, o algoritmo passa ao próximo passo que consiste em montar a árvore CSG que representa a união dos polígonos projetados. O método usado neste trabalho para a representação CSG de polígonos bi-dimensionais utiliza como primitivas polígonos simples. Como as cenas tratadas neste trabalho são compostas somente de polígonos planares, simples e não-interpostos, nenhum tipo de tratamento é necessário para verificar a validade dos polígonos bi-dimensionais formados (Φ_k), assim, eles podem ser utilizados diretamente como primitivas da árvore CSG.

A primeira etapa na montagem da árvore CSG consiste em criar uma sub-árvore que representa o(s) polígono(s) formado(s) pela projeção dos polígonos contidos em T no plano de A_j . Uma forma ingênua de criar essa sub-árvore é demonstrada no lado esquerdo da Figura 3.4. Assim, as primitivas são sempre adicionadas à direita do nó de união atual. Porém, como estabelecido nas teorias da representação CSG, métodos que operam nas árvores, como um método de conversão CSG \rightarrow B-REP, têm maior desempenho se a árvore estiver balanceada, como demonstrado no lado direito da Figura 3.4.

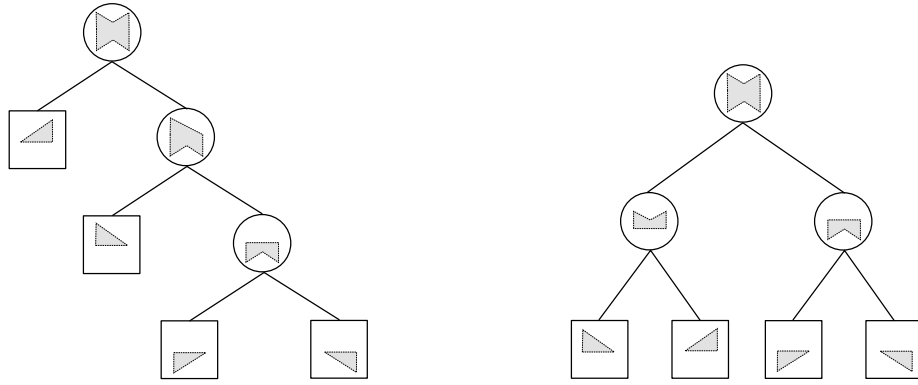


Figura 3.4. Árvores CSG com nodos de união. À esquerda, um polígono representado por uma árvore desbalanceada. À direita, o mesmo polígono representado por uma árvore balanceada.

Após gerada a sub-árvore de união dos polígonos Φ_{k_i} , a árvore final é gerada adicionando um novo nó na árvore que tem como filho esquerdo o polígono A_j , após ser convertido na sua forma bi-dimensional, e como filho direito a sub-árvore de união gerada.

A última etapa consiste em decidir o tipo de operação da raiz da árvore. Duas operações podem ser escolhidas, e elas definem o significado do conjunto de polígonos representado pela árvore. A primeira opção é a operação de interseção, nesse caso, o conjunto de polígonos resultante representa as áreas de A_j ocluídas pelos polígonos contidos em T quando vistos a partir de dA_i . Se, por sua vez, for utilizada a operação de diferença, o conjunto de polígonos resultante representa as áreas visíveis de A_j quando vistos a partir de dA_i , após as oclusões geradas pelos polígonos contidos em T .

Ambas as escolhas para operação booleana da raiz da árvore são válidas e elas somente mudam o método de cálculo do fator-de-forma usado pelo algoritmo na sua etapa final. Neste trabalho, é utilizada a operação de interseção, assim, a árvore CSG representa as áreas ocluídas de A_j .

3.4. Conversão CSG para B-REP

A próxima fase do algoritmo consiste em converter a representação CSG das áreas ocluídas de A_j para uma representação de fronteiras. Isso se faz necessário devido à natureza do cálculo do fator-de-forma utilizado (Equação 2.4) que precisa conhecer as fronteiras das áreas ocluídas de A_j , incluindo as fronteiras dos seus possíveis buracos.

A representação por fronteiras de um polígono bi-dimensional, Φ , consiste de: 1) uma fronteira externa, composta por uma lista de vértices ordenados em sentido anti-horário, $\Phi_0 \dots \Phi_k$; 2) uma lista de representações de fronteiras de outros polígonos bidimensionais, esses polígonos são os buracos de Φ .

A classificação de uma reta paramétrica, R , com relação a um polígono Φ consiste de: 1) uma lista de parâmetros, $t_0 \dots t_k$, que representam intervalos em R ; 2) uma lista de classificações, $c_0 \dots c_{k-1}$, que identificam a relação entre o intervalo e o polígono.

As possíveis classificações de um intervalo com relação a um polígono são três:

1. *IN*. O intervalo encontra-se no interior do polígono.
2. *OUT*. O intervalo encontra-se no exterior do polígono.
3. *ON*. O intervalo faz parte da fronteira externa do polígono.

A conversão de uma árvore CSG, denominada Λ , em uma representação por fronteiras, gera uma lista de polígonos bi-dimensionais. Os passos para encontrar essa lista são explicados a seguir.

O primeiro passo é adquirir as arestas candidatas, as quais podem ou não conter intervalos pertencentes à fronteira externa de algum dos polígonos gerados. A maneira mais simples de adquirir esse conjunto de arestas é fazer com que ele seja igual à união de todas as arestas que compõem alguma primitiva de Λ . Como definido na Seção 3.3,

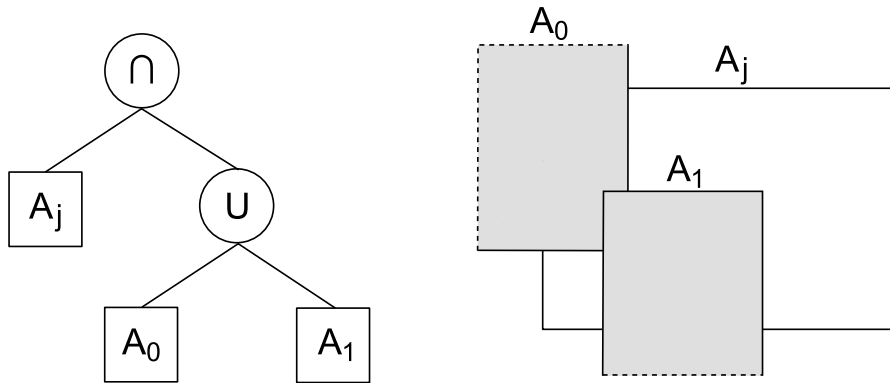


Figura 3.5. Uma árvore CSG e a representação dos polígonos que a formam. As arestas em vermelho não são arestas candidatas.

as primitivas de Λ são polígonos simples e sem buracos, portanto, adquirir o conjunto de arestas candidatas é trivial.

Outra forma de adquirir as arestas candidatas é utilizar o conhecimento da estrutura da árvore CSG utilizada neste método. Como o método está interessado em adquirir as áreas ocluídas de A_j , e para isso utiliza a operação de intersecção no nó raiz da árvore, as únicas arestas que interessam são aquelas que estão parcialmente na fronteira ou no interior de A_j , assim, as arestas que estão claramente no exterior de A_j podem ser descartadas (Ver Figura 3.5).

Uma vez adquirido o conjunto de arestas candidatas, é necessário gerar a classificação de cada aresta com a árvore CSG. Essa classificação é feita recursivamente, descendo até as primitivas da árvore e depois combinando o resultado das classificações de acordo com a operação booleana do nó superior. A classificação de uma aresta com relação a um polígono foi definida anteriormente, porém, é necessário especificar como adquirir os intervalos e como classificar cada um deles.

Os intervalos são adquiridos calculando as interseções entre a aresta que está sendo classificada e cada uma das arestas que compõem o polígono. Os pontos, ou parâmetros, de intersecção são guardados numa lista que é posteriormente ordenada. É interessante

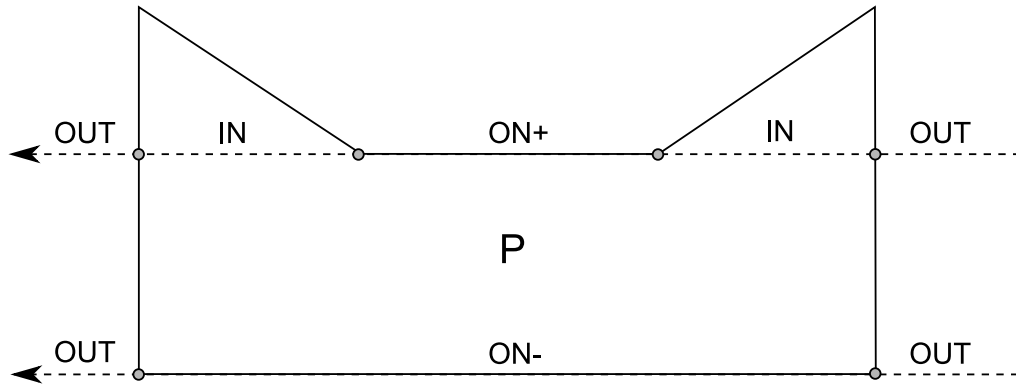


Figura 3.6. *Classificação de duas arestas contra um polígono.*

observar que tanto o ponto inicial como o ponto final da aresta já se iniciam nessa lista.

Após adquiridos os k pontos que definem os intervalos, é necessário classificar os $k - 1$ intervalos. Essa classificação pode ser gerada utilizando qualquer algoritmo de ponto em polígono que consiga lidar com polígonos convexos e côncavos. Neste trabalho é utilizado o método de ponto em polígono apresentado por Weiler [28]. O ponto a ser testado pode ser qualquer um, exceto os pontos extremos do intervalo, por exemplo, o ponto central. A classificação do intervalo terá então a mesma classificação do ponto em relação ao polígono: interno(IN), externo(OUT) ou na fronteira(ON).

É necessário ressaltar que a classificação ON precisa de mais uma informação: a orientação do interior do polígono relativa ao intervalo orientado. Se o interior do polígono estiver à esquerda do intervalo, então a classificação será $ON+$, caso contrário será $ON-$. Essa diferença na classificação de intervalos na fronteira do polígono é necessária para evitar ambiguidades ao combinar duas classificações. A Figura 3.6 ilustra um polígono P e a classificação de duas arestas contra ele.

Uma vez realizadas as classificações de uma aresta em relação às primitivas, essas classificações são combinadas de acordo com o tipo de operação booleana do nó superior a elas. As combinações de classificações são mostradas na Tabela 3.1. O caso particular ON/ON tem o seu resultado definido pelos sinais das classificações. Assim, caso os sinais

Union	IN	OUT	ON	Intersection	IN	OUT	ON
IN	IN	IN	IN	IN	IN	OUT	ON
OUT	IN	OUT	ON	OUT	OUT	OUT	OUT
ON	IN	ON	ON/IN	ON	ON	OUT	ON/OUT

Tabela 3.1. *Combinações de classificações de intervalos*

sejam iguais, o resultado é o primeiro valor definido na tabela, caso contrário, é o segundo valor.

Após classificar todas as arestas contra Λ , é gerado um conjunto de classificações de arestas. Cada classificação nesse conjunto, pode conter intervalos que pertencem à fronteira externa de um dos polígonos finais, ou à fronteira externa de um de seus buracos. Agora, o problema se reduz a gerar os polígonos finais com base na informação sobre a fronteira dos mesmos que é dada pelos intervalos ON das classificações.

Para gerar os polígonos finais, todas as classificações são percorridas, e, toda vez que um intervalo ON é encontrado, seus vértices de extremidade são adicionados a uma lista de vértices sem que haja duplicação de vértices. Essa lista de vértices, após preenchida, deve ser ordenada segundo uma das coordenadas. A escolha da coordenada utilizada para a ordenação não afeta o resultado final do algoritmo. Além da lista de vértices, é mantida uma lista de ligações entre os vértices, que também é atualizada ao encontrar um intervalo ON .

O próximo passo consiste em encontrar os ciclos formados pelos vértices e suas ligações. Cada um desses ciclos é uma fronteira externa de um polígono, que pode ou não ser um buraco de outro polígono previamente gerado. Uma busca em profundidade é utilizada para encontrar esses ciclos. O método mantém uma lista de polígonos criados, uma lista temporária de vértices e um marcador do vértice sendo examinado. O Algoritmo 3.4 mostra o pseudo-código usado para gerar os polígonos finais.

A lista de vértices, γ , é usada temporariamente para manter os ciclos encontrados.

Entrada: Lista de vértices ordenada α , Lista de ligações β
Saída : Lista de polígonos K

- 1 Lista de vértices temporária γ ;
- 2 **Enquanto** α não for vazia **faça**
- 3 | limpe γ ;
- 4 | $x =$ primeiro elemento de α ;
- 5 | remova x de α ;
- 6 | adicione x em γ ;
- 7 | **Enquanto** existir vértice ligado a x em α **faça**
- 8 | | $x =$ vértice ligado à x ;
- 9 | | remova x de α ;
- 10 | | adicione x em γ ;
- 11 | **Fim**
- 12 | **Se** K não for vazio **então**
- 13 | | AdicionaPoligonoBuraco(K , γ);
- 14 | **senão**
- 15 | | AdicionaPoligono(K , α);
- 16 | **Fim**
- 17 **Fim**

Algoritmo 1. Criação dos polígonos representados pela fronteira.

Quando o ciclo é fechado, γ é adicionada como um novo polígono em K , ou como um buraco do último polígono inserido em K . Essa decisão é realizada pela função *AdicionaPoligonoBuraco*, a qual verifica se um ponto qualquer em γ é interno ao último polígono inserido em K , ou a um de seus buracos, utilizando um algoritmo de ponto em polígono. Caso o ponto seja interno, γ é adicionado com um buraco do polígono mais baixo na hierarquia à qual ele é interno. Caso contrário, γ é adicionada como fronteira externa de um novo polígono em K .

O motivo pelo qual somente o último polígono em K precisa ser verificado é a ordenação prévia dos vértices em α . O algoritmo sempre começa um novo ciclo com o primeiro vértice de α , ou seja, com o vértice de menor coordenada. Assim, se existe inclusão, ela só pode ocorrer com o último polígono criado, uma vez que os polígonos não se interpõem entre si. A Figura 3.7 exemplifica uma instância de entrada, demonstrando

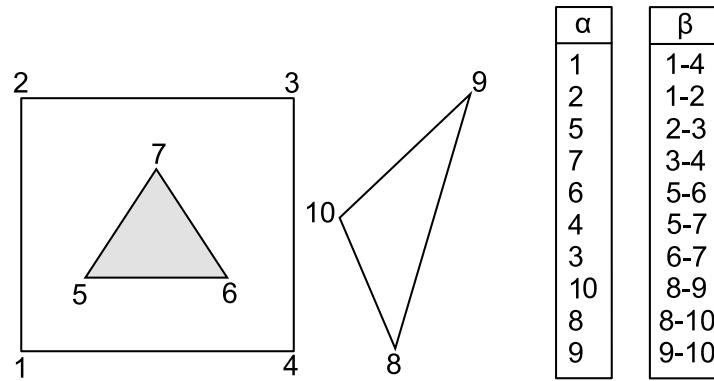


Figura 3.7. *Instância de entrada do Algoritmo 1.*

as listas α e β .

3.5. Cálculo do Fator-de-Forma

O passo final do método consiste no cálculo do fator-de-forma em si. O método utiliza as informações sobre as áreas ocluídas e visíveis de A_j encontradas anteriormente para um cálculo exato do fator-de-forma entre dA_i e A_j .

O fator-de-forma é a porcentagem de energia que é transferida entre duas áreas, como é explicado em detalhes no Apêndice A.3. Assim, o fator-de-forma entre dA_i e A_j pode ser visto como a soma das porcentagens de energia oriundas de dA_i que chegam nas partes visíveis de A_j , ou também, como a porcentagem da energia que deixa dA_i e chega em A_j , sem considerar oclusão, menos a porcentagem da energia que atinge as áreas ocluídas de A_j . Essa diferença na forma de encarar o fator-de-forma especifica como a árvore CSG deve ser montada. Como mencionado na Seção 3.3, a escolha feita é a de descobrir as áreas ocluídas de A_j , o que acarreta em um cálculo do fator-de-forma usando a segunda abordagem.

Com acesso às informações das áreas ocluídas de A_j , calculadas pelo método descrito na seção anterior, o cálculo do fator-de-forma é feito de maneira simples.

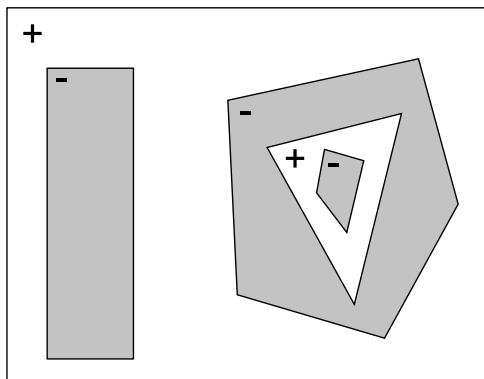


Figura 3.8. Áreas positivas e negativas de um polígono.

É importante salientar que, dado um polígono bidimensional, Φ , representado por suas fronteiras, é necessário definir a orientação, ou sinal, de suas áreas. O conjunto de áreas positivas de Φ é composto pela área definida pela sua fronteira externa e as áreas dos buracos de nível par na hierarquia. O conjunto de áreas negativas é composto pelas áreas dos buracos de nível ímpar na hierarquia. Para efeitos práticos deste trabalho, o que isso significa é que, para cada polígono Φ_i em K , as áreas positivas são áreas ocluídas de A_j , enquanto que as áreas negativas são as áreas visíveis. Mais uma vez, é necessário lembrar que a forma como as áreas positivas e negativas são empregadas depende exclusivamente do método usado para calcular o fator-de-forma. Caso o método escolhido seja o primeiro apresentado anteriormente, o significado das áreas positivas e negativas seria o oposto. A Figura 3.8 ilustra um exemplo das áreas positivas e negativas de um polígono.

Antes de serem utilizadas para o cálculo do fator-de-forma, todos os polígonos de K , com suas respectivas áreas positivas e negativas, precisam ser convertidos de volta à sua forma tridimensional. A conversão para a forma tridimensional utiliza o mesmo sistema de coordenadas utilizado na conversão bidimensional (equações 3.3 e 3.4) e o método é o inverso do apresentado na Equação 3.5.

Assim, o fator-de-forma entre dA_i e A_j é igual ao fator-de-forma entre dA_i e A_j sem considerar oclusões, subtraindo o fator-de-forma entre dA_i e as áreas positivas de cada

polígono de K , e por fim, somando o fator-de-forma entre dA_i e as áreas negativas de cada polígono de K . A fórmula para o cálculo desse fator de forma é dada por

$$F_{dA_i-A_j}^{exato} = F_{dA_i-A_j} - \left(\sum F_{dA_i-A_K^+} \right) + \left(\sum F_{dA_i-A_K^-} \right) \quad (3.6)$$

onde $F_{dA_i-A_j}^{exato}$ é o fator-de-forma exato entre dA_i e A_j levando em consideração oclusões; $F_{dA_i-A_j}$ é o fator-de-forma exato analítico entre dA_i e A_j sem considerar oclusões, calculado usando a Equação 2.4; A_K^+ e A_K^- são, respectivamente, as áreas negativas e positivas contidas em K ; e por fim, $F_{dA_i-A_K^+}$ e $F_{dA_i-A_K^-}$ são calculados usando a Equação 2.4.

A Equação 3.6 calcula o fator-de-forma exato entre dA_i e A_j , usando conhecimentos das áreas visíveis e ocluídas de A_j , e trata perfeitamente as oclusões complexas, como a demonstrada na Figura 3.8. O método do cálculo do fator-de-forma apresentado nesta seção já foi utilizado em outros trabalhos, como demonstrado no Capítulo 2, a diferença está na forma de calcular as áreas visíveis e ocluídas de A_j .

3.6. Considerações Finais

Neste capítulo foi apresentado um método para o cálculo exato do fator-de-forma entre dA_i e A_j . O método apresentado utiliza conhecimentos de *CSG* para identificar as áreas visíveis e ocluídas de A_j . Algumas considerações acerca do método são feitas a seguir.

O teste de visibilidade, apresentado na Seção 3.2, pode ser substituído por quaisquer outros métodos existentes que resolvam o problema da classificação de visibilidade. Embora nenhum estudo sobre a eficácia desses métodos seja feito neste trabalho, o teste de visibilidade tem um impacto considerável no desempenho geral do método, e um futuro estudo sobre o desempenho dos métodos existentes e a adequação deles ao método proposto pode se mostrar frutífero.

O método utilizado para a conversão entre *CSG* e *B-REP* é facilmente paralelizável. A classificação de uma aresta é completamente independente da classificação de outra aresta, assim, as classificações podem ocorrer em paralelo, reduzindo drasticamente o tempo de execução do método em ambientes com vários núcleos.

No próximo capítulo, são apresentados os resultados adquiridos utilizando o método aqui apresentado.

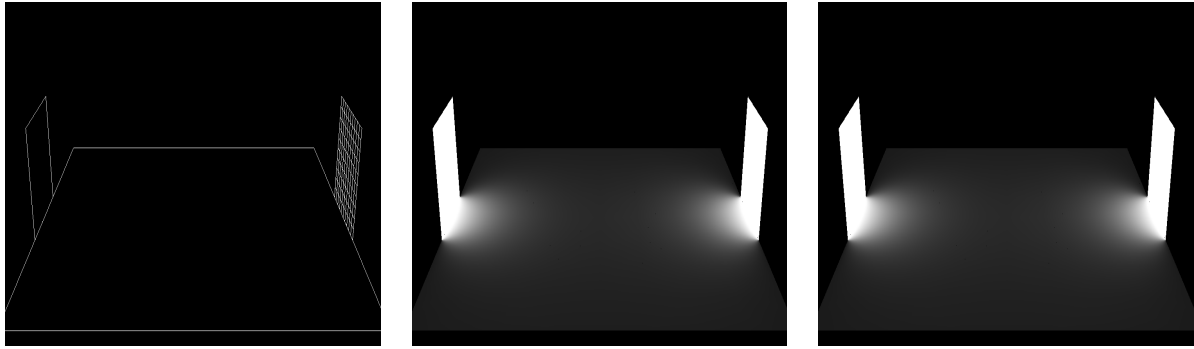
Resultados Experimentais

4.1. Introdução

Neste capítulo são apresentados os resultados obtidos com o método apresentado neste trabalho (Capítulo 3), aplicado no contexto de Radiosidade Progressiva. O método de amostragem em área apresentado por Wang et al. [16] é o utilizado como método de comparação.

Os dois métodos foram implementados em C++, utilizando o OpenGL para renderização, em um PC com 2.5Ghz Intel Core Duo, 1GByte de memória e uma placa de vídeo ATI Radeon X600XT com 128MByte de memória. A comparação entre os dois métodos é feita tanto qualitativamente, por análise visual, como quantitativamente, comparando o tempo para atingir a solução, bem como a quantidade de elementos e vértices gerados.

No restante deste capítulo, é feita uma discussão sobre vários testes que foram elaborados para avaliar a metodologia proposta neste trabalho. Associado a cada teste foi elaborado um cenário, com finalidade de avaliar aspectos importantes do cálculo do fator-de-forma.



(a) Modelo em *Wireframe*

(b) Método proposto

(c) Amostragem de área

Figura 4.1. *Renderização da cena de Emissão.*

	Patches	Elementos	Vértices	Tempo
Método Proposto	102	4770	5040	62,203s
Wang	102	4767	5037	64,75s

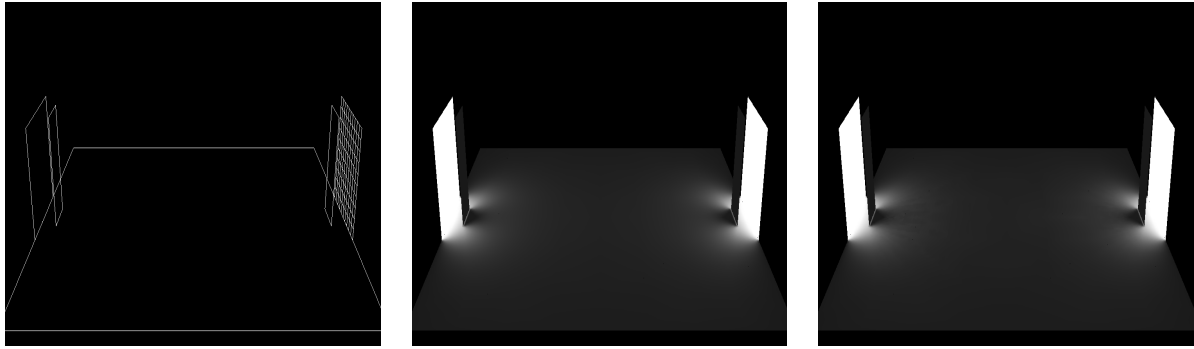
Tabela 4.1. *Resultados quantitativos da cena de Emissão.*

4.2. Emissão

A cena utilizada neste teste (Figura 4.1) é composta por duas fontes de luz retangulares e um polígono base. A luz esquerda é composta por um único polígono, enquanto que a luz direita é composta por cem polígonos que estão dispostos uniformemente e cuja área conjunta é igual à área da luz esquerda (Figura 4.1.(a)).

O propósito desse teste é verificar a capacidade do algoritmo de calcular a iluminação correta independentemente do nível de subdivisão das fontes de luz. Como a cena não contém nenhuma oclusão, algoritmos de amostragem de área devem ter a mesma qualidade visual dos algoritmos exatos, porém, os primeiros tendem a fazer uma subdivisão excessiva das fontes de luz quando a distância entre a área iluminada e a fonte de luz diminui, o que pode causar uma perda de desempenho.

As iluminações geradas pelos dois métodos são idênticas qualitativamente (Figuras 4.1.(b) e 4.1.(c)) e, os desempenhos também são muito similares (Tabela 4.1).



(a) Modelo em *Wireframe*

(b) Método proposto

(c) Amostragem de área

Figura 4.2. *Renderização da cena de Emissão com Obstrução.*

	Patches	Elementos	Vértices	Tempo
Método Proposto	104	3035	3295	147,235s
Wang	104	3656	3988	338,047s

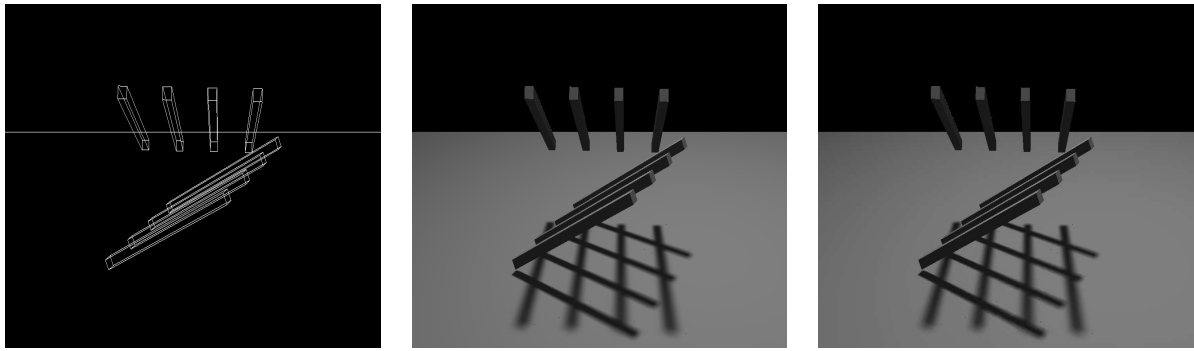
Tabela 4.2. *Resultados quantitativos da cena de Emissão com Obstrução.*

4.3. Emissão com Obstrução

A segunda cena utilizada nos testes (Figura 4.2) é similar à primeira, porém foram adicionados dois polígonos opacos e não refletores que obstruem a iluminação proveniente das fontes luminosas (Figura 4.2.(a)).

O objetivo deste teste é demonstrar como o método proposto consegue capturar as sutilezas na iluminação causadas por oclusões suaves sem a necessidade de subdividir a fonte de luz. O resultado esperado é que as sombras geradas pelos dois polígonos oclusores sejam as mesmas, uma vez que eles têm a mesma área e estão posicionados à mesma distância das duas fontes luminosas.

Como pode ser observado na Figura 4.2.(b), o método proposto gera sombras idênticas para ambos os polígonos oclusores. Por sua vez, o método de amostragem de área não consegue um bom resultado quando interage com oclusões onde o polígono oclusor tem área muito menor que o polígono emissor, como ocorre do lado esquerdo da cena (Figura



(a) Modelo em *Wireframe*

(b) Método proposto

(c) Amostragem de área

Figura 4.3. *Renderização da cena de Sombras Complexas.*

4.2.(c)). Nesses casos, para conseguir um resultado com qualidade similar, o polígono emissor precisa ser subdividido excessivamente, o que acarreta em um tempo de execução elevado, como pode ser notado na Tabela 4.2.

4.4. Sombras Complexas

A próxima cena utilizada nos testes (Figura 4.3) é composta de uma única fonte de luz, oito longos feixes oclusores e um polígono base (Figura 4.3.(a)). A fonte de luz é rotacionada de forma a incidir perpendicularmente nos feixes inferiores. Ademais, a distância de cada feixe ao plano base varia ao longo do seu comprimento.

O objetivo neste teste é verificar a eficácia dos métodos na presença de: 1) oclusões complexas e sobrepostas; e 2) transições entre sombras duras e suaves. O resultado esperado no primeiro caso é uma descontinuidade da iluminação nas áreas sombreadas, semelhante ao que ocorre nas oclusões simples, onde áreas mais escuras são geradas no interior das sombras. Quanto ao segundo caso, espera-se um aumento gradativo das áreas de penumbra e, conseqüentemente, uma diminuição das áreas de umbra.

Os resultados do teste demonstram que os dois métodos conseguem obter imagens

	Patches	Elementos	Vértices	Tempo
Método Proposto	50	18716	19977	99,145s
Wang	50	18644	19929	197,047s

Tabela 4.3. *Resultados quantitativos da cena de Sombras Complexas.*

de qualidade aceitável (Figuras 4.3.(b) e 4.3.(c)), porém os seus desempenhos diferem consideravelmente (Tabela 4.3). O desempenho inferior do método de amostragem é causado pela subdivisão excessiva da fonte de luz, necessária para atingir uma qualidade similar à qualidade do método proposto.

4.5. Oclusão por Pequenas Áreas

Esta cena (Figura 4.4) é composta de uma única fonte luminosa, um polígono base e vinte e cinco polígonos oclusores. A fonte de luz tem a mesma área do polígono base e está posicionada diretamente acima desse polígono. Os polígonos oclusores têm áreas de aproximadamente um centésimo da área do polígono base.

O objetivo desse teste é verificar a capacidade do método proposto em capturar sombras causadas por pequenos oclusores. Esse tipo de sombra tem grande parte de sua área composta pela penumbra, ou seja, a maior parte da área ocluída recebe alguma energia luminosa.

Os resultados renderizados da cena (Figuras 4.4.(b) e 4.4.(c)) mostram que apesar do método de amostragem conseguir uma qualidade de imagem aceitável, ainda é possível identificar artefatos causados pela aproximação do fator-de-forma. Os resultados quantitativos, apresentados na Tabela 4.4, mostram um desempenho consideravelmente superior do método proposto em contraposição ao método de amostragem, o que demonstra que o método é adequado para tratar de oclusões múltiplas.

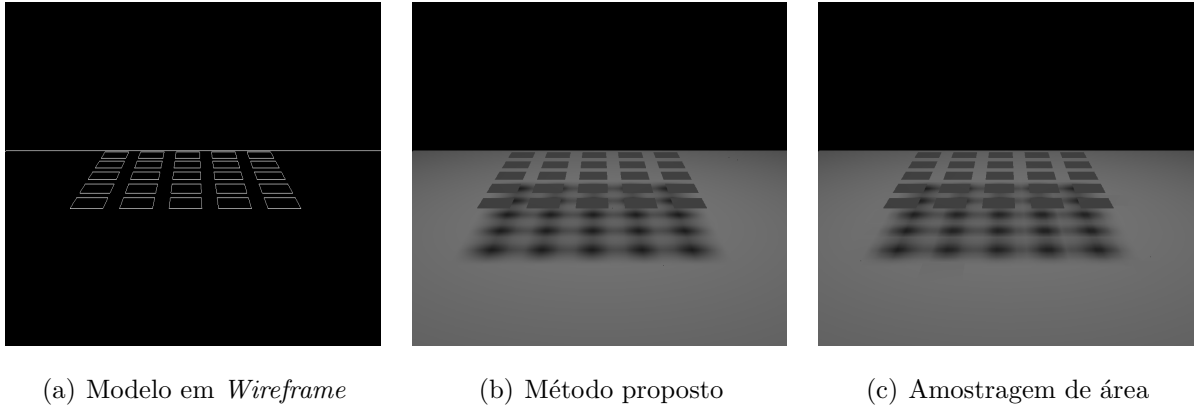


Figura 4.4. *Renderização da cena de Oclusão por Pequenas Áreas.*

	Patches	Elementos	Vértices	Tempo
Método Proposto	27	3684	3959	42,235s
Wang	27	3711	3968	150,391s

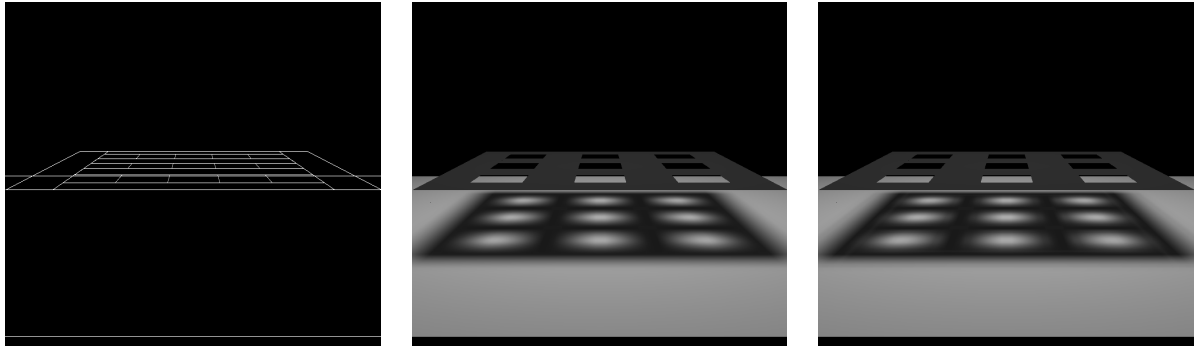
Tabela 4.4. *Resultados quantitativos da cena de Oclusão por Pequenas Áreas.*

4.6. Oclusão por Polígono Vazado

A última cena (Figura 4.5) utilizada nos testes pode ser vista como o oposto da cena anterior. Ela é composta da mesma fonte de luz e polígono base, mas as oclusões são causadas por um grande painel, com aproximadamente metade da área do polígono base, que contém nove buracos, uniformemente distribuídos em sua extensão, que permitem a passagem da luz (Figura 4.5.(a)).

Assim como na cena anterior, o objetivo é testar a capacidade do método de capturar pequenas nuances causadas na transição entre umbra e penumbra, porém, neste caso, as áreas de penumbra circundam áreas completamente iluminadas situadas diretamente abaixo dos buracos no painel.

Os resultados mostram uma qualidade visual semelhante dos dois métodos (Figuras 4.5.(b)), porém, o método proposto obtêm um desempenho superior quantitativamente (Tabela 4.5), o que confirma a capacidade do método em calcular a iluminação correta



(a) Modelo em *Wireframe*

(b) Método proposto

(c) Amostragem de área

Figura 4.5. *Renderização da cena de Oclusão por Polígono Vazado.*

	Patches	Elementos	Vértices	Tempo
Método Proposto	14	3684	3959	42,235s
Wang	14	3711	3968	150,391s

Tabela 4.5. *Resultados quantitativos da cena de Oclusão por Polígono Vazado.*

proveniente de buracos presentes nos polígonos oclusores, uma das grandes dificuldades dos métodos aproximativos.

4.7. Considerações Finais

Neste capítulo foram apresentados os resultados obtidos com o método proposto, comparando-o ao método de amostragem de área, um dos métodos aproximativos mais utilizados em Radiosidade.

Os resultados mostram que o método proposto é adequado para todas as situações consideradas difíceis nos métodos de cálculo do fator-de-forma e, que na maioria dos casos, tem um desempenho consideravelmente superior ao método comparado.

Conclusões

Este trabalho apresentou um método para calcular o fator-de-forma exato entre uma área diferencial e uma área finita poligonal, utilizando técnicas de projeção e modelagem de sólidos.

O método proposto é similar aos métodos exatos existentes(Capítulo 2) no que diz respeito à forma utilizada para o cálculo do fator-de-forma. A diferença está na forma utilizada para identificar as áreas visíveis e ocluídas do polígono emissor que, no método proposto, consiste em utilizar uma combinação entre técnicas de projeção e técnicas de *CSG*, como detalhado no Capítulo 3.

O método proposto foi comparado com o método de amostragem em área e apresentou desempenho comparável a esses métodos que, em geral, apresentam melhor desempenho do que os métodos exatos para o cálculo do fator-de-forma. Um ponto a ser trabalhado no futuro é a comparação do método apresentado neste trabalho com os métodos mais representativos dentre os métodos exatos para o cálculo do fator-de-forma. Além disso, o método proposto pode ser utilizado em uma abordagem híbrida, onde um método aproximativo com melhor desempenho pode ser utilizado quando o valor exato não for importante ou aproximação for boa.

A abordagem do método proposto para a identificação das áreas visíveis e ocluídas dos polígonos com relação aos vértices também pode ser utilizada como uma forma alternativa às utilizadas nos trabalhos que constrõem estruturas de dados de visibilidade, como [6, 7, 8]. Além disso, o método proposto se beneficiaria bastante de uma implementação paralela, visto que os principais módulos do método, os testes de visibilidade e as conversões entre polígonos *CSG* e *B-REP*, ocorrem independentemente, podendo assim ser executados em paralelo.

Em conclusão, este trabalho apresentou com sucesso um método original para o cálculo do fator-de-forma exato entre áreas diferenciais e áreas finitas poligonais, apesar de conter algumas falhas comparativas. Além disso, várias melhorias para o método apresentado foram propostas ao longo do texto, tanto de embasamento comparativo como de desempenho.

Referências Bibliográficas

- [1] D. R. Baum, H. E. Rushmeier, and J. M. Winget. Improving radiosity solutions through the use of analytically determined form-factors. In *SIGGRAPH '89: Proceedings of the 16th annual conference on Computer graphics and interactive techniques*, pages 325–334, New York, NY, USA, 1989. ACM.
- [2] Sylvain Charneau, Lilian Aveneau, and Laurent Fuchs. Exact, robust and efficient full visibility computation in plücker space. *Visual Computer*, 23(9):773–782, 2007.
- [3] Michael F. Cohen and Donald P. Greenberg. The hemi-cube: a radiosity solution for complex environments. In *SIGGRAPH '85: Proceedings of the 12th annual conference on Computer graphics and interactive techniques*, pages 31–40, New York, NY, USA, 1985. ACM.
- [4] Michael F. Cohen, John Wallace, and Pat Hanrahan. *Radiosity and realistic image synthesis*. Academic Press Professional, Inc., San Diego, CA, USA, 1993.
- [5] François Cuny, Laurent Alonso, and Nicolas Holzschuch. A novel approach makes higher order wavelets really efficient for radiosity. *Computer Graphics Forum (Pro-*

- ceedings of Eurographics 2000*), 19(3):C99–C108, sep 2000. (Eurograph.
- [6] George Drettakis and Eugene Fiume. A fast shadow algorithm for area light sources using backprojection. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 223–230, New York, NY, USA, 1994. ACM.
 - [7] Frédo Durand, George Drettakis, and Claude Puech. The visibility skeleton: a powerful and efficient multi-purpose global visibility tool. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 89–100, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.
 - [8] Frédo Durand, George Drettakis, and Claude Puech. Fast and accurate hierarchical radiosity using global visibility. *ACM Transactions on Graphics*, 18(2):128–170, 1999.
 - [9] James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes. *Computer graphics: principles and practice (2nd ed.)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.
 - [10] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaille. Modeling the interaction of light between diffuse surfaces. In *SIGGRAPH '84: Proceedings of the 11th annual conference on Computer graphics and interactive techniques*, pages 213–222, New York, NY, USA, 1984. ACM.
 - [11] David Hart, Philip Dutré, and Donald P. Greenberg. Direct illumination with lazy visibility evaluation. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 147–154, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

- [12] Paul S. Heckbert. Discontinuity meshing for radiosity. In *Third Eurographics Workshop on Rendering*, pages 203–226, 1992.
- [13] Nicolas Holzschuch and Laurent Alonso. Combining higher-order wavelets and discontinuity meshing: a compact representation for radiosity. In *Rendering Techniques 2004 (Eurographics Symposium on Rendering)*, jun 2004.
- [14] James T. Kajiya. The rendering equation. *SIGGRAPH Comput. Graph.*, 20(4):143–150, 1986.
- [15] Janne Kontkanen, Emmanuel Turquin, Nicolas Holzschuch, and François Sillion. Wavelet radiance transport for interactive indirect lighting. In Wolfgang Heidrich and Thomas Akenine-Möller, editors, *Rendering Techniques 2006 (Eurographics Symposium on Rendering)*. Eurographics, jun 2006.
- [16] Wang-Yeh Lee and Jung-Hong Chuang. Effective meshing and form factor calculation for accurate progressive radiosity. In *TPCG '04: Proceedings of the Theory and Practice of Computer Graphics 2004 (TPCG'04)*, pages 144–151, Washington, DC, USA, 2004. IEEE Computer Society.
- [17] Dani Lischinski, Filippo Tampieri, and Donald P. Greenberg. Discontinuity meshing for accurate radiosity. *IEEE Comput. Graph. Appl.*, 12(6):25–39, 1992.
- [18] T. J Malley. A shading method for computer generated images. Master’s thesis, Dept. of Computer Science, University of Utah, 1988.
- [19] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of three-dimensional objects taking account of shadows and interreflection. *SIGGRAPH Comput. Graph.*, 19(3):23–30, 1985.

- [20] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of three-dimensional objects illuminated by sky light. *SIGGRAPH Comput. Graph.*, 20(4):125–132, 1986.
- [21] Georg Pietrek. Fast Calculation of Accurate Formfactors. In *Fourth Eurographics Workshop on Rendering*, number Series EG 93 RW, pages 201–220, Paris, France, 1993.
- [22] R. Siegel and J.R. Howeol. *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corp., Washington DC, 1981.
- [23] F. Sillion and C. Puech. A general two-pass method integrating specular and diffuse reflection. In *SIGGRAPH '89: Proceedings of the 16th annual conference on Computer graphics and interactive techniques*, pages 335–344, New York, NY, USA, 1989. ACM.
- [24] Francois X. Sillion and Claude Puech. *Radiosity and Global Illumination*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1994.
- [25] Michael M. Stark and Richard F. Reisenfeld. Exact illumination in polygonal environments using vertex tracing. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, pages 149–160, London, UK, 2000. Springer-Verlag.
- [26] Filippo Tampieri. Accurate form-factor computation. In *Graphics Gems III*, pages 329–333. Academic Press Professional, Inc., San Diego, CA, USA, 1992.
- [27] J. R. Wallace, K. A. Elmquist, and E. A. Haines. A ray tracing algorithm for progressive radiosity. In *SIGGRAPH '89: Proceedings of the 16th annual conference on Computer graphics and interactive techniques*, pages 315–324, New York, NY, USA, 1989. ACM.

- [28] Kevin Weiler. An incremental angle point in polygon test. In *Graphics gems IV*, pages 16–23. Academic Press Professional, Inc., San Diego, CA, USA, 1994.

Background

As teorias básicas necessárias para a compreensão dos métodos abordados neste trabalho vão ser apresentadas neste apêndice. Primeiro será feita uma breve introdução sobre o campo em que toda a teoria dos métodos de Radiosidade está fundamentada, a Radiometria. Em seguida o método de Radiosidade Clássica, como proposto por Goral et al. [10], será apresentado. E por último, a teoria acerca do cálculo dos fatores-de-forma é demonstrada, tanto para áreas diferenciais como para áreas finitas.

A.1. Radiometria

A medida mais importante utilizada no nosso estudo sobre o comportamento da luz, e de fato em qualquer estudo nessa área, é sem dúvida a radiância. A radiância indica quanto da energia luminosa que deixa uma superfície é recebida por um observador que a observa de um determinado ângulo, ou formalmente:

Definição A.1.1. Radiância mede a quantidade de energia luminosa($d\phi$) que flui ou é emitida por uma determinada área(x) e atinge um ângulo sólido em uma determinada direção($\vec{d\omega}$). A radiância é medida em $\frac{W}{m^2 sr}$.

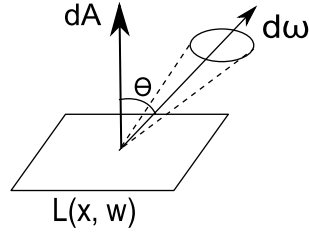


Figura A.1. Geometria da radiância.

A radiância também pode ser definida em termos matemáticos como:

$$L(x, \vec{\omega}) = \frac{d\phi(x)}{\cos\theta d\omega dA} \quad (\text{A.1})$$

É necessário definir duas outras medidas essenciais para este estudo, a irradiância e a radiosidade. Informalmente a irradiância pode ser descrita como a quantidade de energia luminosa que incide em uma superfície. A radiosidade pode ser vista como o oposto da irradiância, ou seja, a quantidade de energia luminosa que deixa uma superfície. As definições formais seguem:

Definição A.1.2. Irradiância é a quantidade de energia luminosa($d\phi$) por unidade de área(dA) que incide em uma superfície em todas as direções($\vec{d\omega}$). Irradiância é medida em $\frac{W}{m^2}$.

Definição A.1.3. Radiosidade é a quantidade de energia luminosa($d\phi$) por unidade de área(dA) que deixa uma superfície em todas as direções($\vec{d\omega}$). Radiosidade é medida em $\frac{W}{m^2}$.

Intuitivamente, pela definição, a irradiância pode ser considerada como a quantidade de energia luminosa proveniente de todas as direções por unidade área. Pela equação A.1 é possível deduzir que:

$$d\phi(x) = L(x, \vec{\omega}) \cos\theta d\omega dA \quad (\text{A.2})$$

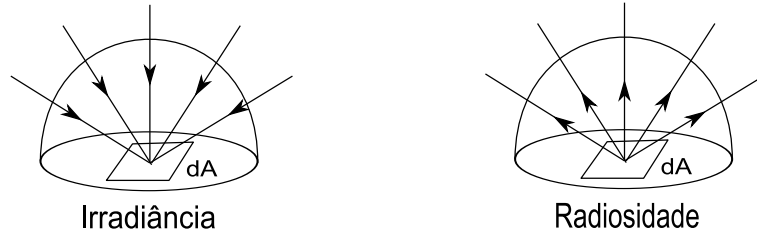


Figura A.2. Irradiância e Radiosidade.

Para obter toda a energia incidente na superfície x é preciso integrar a radiância incidente L_i sobre todas as direções de uma semi-esfera, Ω . Ao ser aplicada essa integral é obtida a equação de energia incidente em x :

$$d\phi_i(x) = \int_{\Omega} L_i(x, \vec{\omega}) \cos\theta d\omega dA \quad (\text{A.3})$$

Utilizando a equação A.3 e a definição A.1.2 sobre irradiância e A.1.3 sobre radiosidade, pode-se derivar a equação da irradiância, e consequentemente da radiosidade:

$$I(x) = \frac{d\phi_i}{dA} = \int_{\Omega} L_i(x, \vec{\omega}) \cos\theta d\omega \quad (\text{A.4})$$

$$B(x) = \frac{d\phi_o}{dA} = \int_{\Omega} L_o(x, \vec{\omega}) \cos\theta d\omega \quad (\text{A.5})$$

Terminadas as definições necessárias, a questão da reflexão da luz em uma superfície e como ela ocorre pode ser considerada. Considere uma luz incidente em uma superfície na direção ω_i . A quantidade de luz refletida na direção ω_r é proporcional à quantidade de energia incidente na direção ω_i . Além disso, a proporção entre a quantidade de luz incidente e refletida, uma vez fixados os ângulos de incidência e reflexão, é constante e independente da quantidade de energia luminosa incidente.

Essa constante de proporcionalidade é chamada de função bi-direcional de distribuição da reflexão, ou BRDF (*Bidirectional reflection distribution function*). Mais especifica-

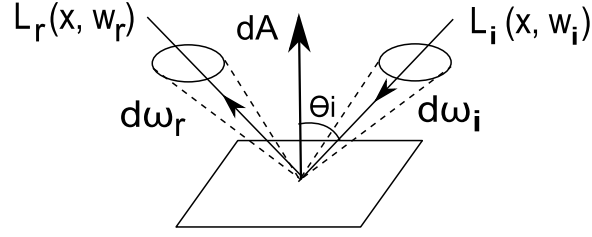


Figura A.3. Geometria da BRDF.

mente a BRDF é definida como a taxa de radiância refletida na direção ω_r devido à quantidade de energia luminosa proveniente de ω_i . Em termos matemáticos a BRDF é definida como:

$$f(\vec{\omega}_i \rightarrow \vec{\omega}_r) = \frac{L_r(\vec{\omega}_r)}{L_i(\vec{\omega}_i) \cos \theta_i d\omega_i} \quad (\text{A.6})$$

Utilizando-se da equação A.6 e de algumas propriedades da BRDF, omitidas aqui por motivos de relevância (O leitor é direcionado à [4] para maiores esclarecimentos sobre a BRDF e suas propriedades), uma das equações mais importantes no estudo de reflexão da luz pode ser derivada, a equação de reflectância:

$$L_r(\vec{\omega}_r) = \int_{\Omega_i} f_r(\vec{\omega}_i \rightarrow \vec{\omega}_r) L_i(\vec{\omega}_i) \cos \theta_i d\omega_i \quad (\text{A.7})$$

No caso específico da reflexão difusa, que é o assunto focado neste trabalho, algumas deduções podem ser feitas. Em ambientes em que todas as superfícies são Lambertianas (emissoras e/ou difusoras perfeitas) a BRDF é constante, independente da direção de incidência e reflexão, assim uma nova medida pode ser usada em seu lugar, a reflectividade, e ela se corresponde com a BRDF da seguinte forma:

$$\rho(x) = f(x) \pi \quad (\text{A.8})$$

Outra dedução importante é que a radiosidade é constante em todas direções obser-

vadas, isso quer dizer que:

$$B(x) = \int_{\Omega} L_o(x, \vec{\omega}) \cos\theta d\omega = L_o(x, \vec{\omega})\pi \quad (\text{A.9})$$

Essas duas deduções são importantes para que a função de renderização possa ser reescrita no caso dos métodos de Radiosidade, onde somente superfícies Lambertianas são consideradas. A equação de renderização proposta inicialmente por Kajiya [14] pode ser escrita da seguinte forma:

$$L(x, \vec{\omega}_r) = L_e(\vec{\omega}_r) + \int_S f_r(\vec{\omega}_i \rightarrow \vec{\omega}_r) L(x, \vec{\omega}_i) G(x, x') V(x, x') dA' \quad (\text{A.10})$$

Utilizando-se as deduções demonstradas nas equações A.8 e A.9 os termos podem ser substituídos para alcançar a equação de renderização específica para métodos de radiosidade:

$$B(x) = E(x) + \rho(x) \int_S B(x') \frac{G(x, x') V(x, x') dA'}{\pi} \quad (\text{A.11})$$

A próxima seção irá mostrar como essa equação de renderização de radiosidade foi discretizada por Goral et al. [10] para atingir um resultado em tempo hábil.

A.2. Radiosidade Clássica

Os métodos de Radiosidade resolvem uma subclasse do problema de iluminação global. Nesse subproblema as superfícies que compõem o ambiente são superfícies Lambertianas, como visto na seção passada. Superfícies Lambertianas são superfícies perfeitamente difusas, ou seja, a intensidade da luz que emana de uma área diferencial da superfície em uma dada direção é proporcional ao cosseno do ângulo dessa direção com a normal da superfície, além disso, a iluminação de uma superfície Lambertiana é constante, indepen-

dente do ângulo de observação.

A solução desse subproblema de iluminação é encontrar a radiosidade de cada ponto das superfícies do ambiente, ou seja, resolver a equação de renderização de radiosidade (equação A.11) para cada ponto. A radiosidade de uma área diferencial dA_i depende de sua própria emissividade, caso a superfície em que dA_i se encontra seja emissora, da reflectividade da superfície de dA_i e da energia luminosa que chega em dA_i proveniente de outras superfícies da cena. Levando isso em consideração, o valor de radiosidade de uma área dA_i pode ser encontrado pela equação:

$$B_{dA_i} = E_{dA_i} + \rho_{dA_i} \int_{dA_j} H_{dA_j} \quad (\text{A.12})$$

onde B_{dA_i} é a radiosidade da área diferencial dA_i , E_{dA_i} é a emissividade da área diferencial dA_i , ρ_{dA_i} é a reflectância da área diferencial dA_i , H_{dA_j} é a energia luminosa proveniente de dA_j que chega em dA_i .

Mesmo sendo um subproblema do problema original de iluminação, a sua solução ainda é muito complexa. É impraticável calcular a radiosidade de cada ponto das superfícies do ambiente, portanto uma forma de simplificar mais ainda esse subproblema é necessária.

A solução proposta por Goral et al.[10] foi discretizar as superfícies do ambiente em componentes chamadas *patches*, e assumir que a radiosidade em todos os pontos dessas componentes é constante. Dessa forma, ao invés de calcular o valor da função de radiosidade em um domínio contínuo, o problema é reduzido à calcular o valor da função sobre um domínio discreto.

Nessa nova instância do problema, a energia que chega em uma superfície j , denotada por $H(j)$, é calculada somando-se as contribuições das outras $n-1$ superfícies do ambiente com a própria contribuição de j , caso j veja a si mesma. A energia que deixa a superfície j é denotada por $B(j)$, ou radiosidade da superfície j , e é o que queremos encontrar para

cada superfície do ambiente.

A radiosidade de uma superfície j pode ser encontrada somando-se a energia que a superfície emite naturalmente, no caso dela ser uma fonte de luz, com a energia refletida proveniente das outras superfícies da cena. Para calcular essa energia refletida, são necessárias duas informações: a reflectividade de cada superfície, esse é um dado de entrada e deve ser provido ao método; e a fração da energia que sai de cada superfície e chega na superfície j , essa informação deve ser calculada para cada combinação de superfícies do ambiente e é denotada por Fator-de-Forma. O cálculo do Fator-de-Forma é um problema em si e será abordado em outras seções deste trabalho.

Portanto o cálculo da radiosidade da superfície j pode ser dado por:

$$B(j) = E(j) + \rho(j)H(j) \quad (\text{A.13})$$

onde $B(j)$ é a radiosidade da superfície j , $E(j)$ é a emissividade de j , $\rho(j)$ é a reflectividade de j e $H(j)$ é a energia incidente em j proveniente de todas as outras superfícies da cena.

A energia incidente em j , $H(j)$, pode ser encontrada ao somar-se a contribuição de cada superfície da cena, o que gera a seguinte equação:

$$H(j) = \sum_i B(i)F(i, j) \quad (\text{A.14})$$

onde $H(j)$ é a energia incidente em j , $B(i)$ é a radiosidade da superfície i , $F(i, j)$ é a fração de energia que sai de i e chega em j .

Combinando as duas funções temos a equação de radiosidade de uma superfície j :

$$B(j) = E(j) + \rho(j)H(j) \sum_i B(i)F(i, j) \quad (\text{A.15})$$

Essa equação existe para cada superfície do ambiente, gerando um conjunto de n

equações lineares com n valores, $B(j)$, desconhecidos. Representando esse sistema de equações em uma forma matricial nós temos:

$$\begin{pmatrix} 1 & -\rho_1 F_{12} & \cdots & -\rho_1 F_{1n} \\ -\rho_2 F_{21} & 1 & \cdots & -\rho_2 F_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -\rho_n F_{n1} & -\rho_n F_{n2} & \cdots & -\rho_n F_{nn} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{pmatrix} = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{pmatrix}$$

Como esse método assume que cada superfície tem uma radiosidade constante, o tamanho das superfícies deve ser pequeno o suficiente para simular as nuances de iluminação que ocorrem em um ambiente real. Assim sendo as superfícies iniciais do ambiente são subdivididas em elementos menores com as mesmas propriedades das superfícies originais. No método proposto por Goral et al [10] o número de elementos em que cada superfície é dividida é especificado como parâmetro de entrada do método.

Esse é o método de Radiosidade Clássica como foi proposto inicialmente, ele ainda é até hoje, apesar de não levar em conta as reflexões especulares entre os objetos, uma das famílias de métodos mais bem sucedidas para simular a iluminação de ambientes reais.

A.3. Form Factors

Nesta subseção vamos abordar o problema de determinar os fatores-de-forma em um contexto teórico. Essa abordagem teórica é necessária para compreensão correta dos vários métodos utilizados na prática para determinar os fatores-de-forma.

Uma definição formal do que exatamente significa o fator-de-forma é mostrada a seguir:

Definição A.3.1. O fator-de-forma entre duas superfícies é um valor que caracteriza a troca de energia luminosa entre elas, mais especificamente, ele caracteriza a fração de energia que deixa uma superfície e é recebida pela outra.

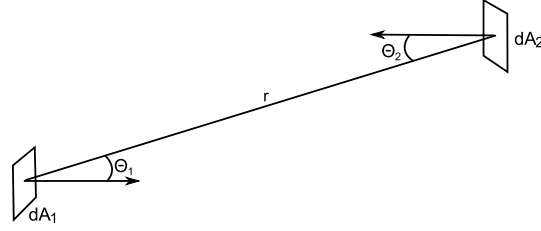


Figura A.4. Geometria do Fator-de-Forma entre dA_i e dA_j .

Vamos considerar agora o problema de determinar o fator-de-forma entre duas áreas diferenciais dA_i e dA_j (veja figura A.4). Pela definição de fator-de-forma, esse valor é encontrado ao se dividir a quantidade de energia proveniente de dA_i que atinge dA_j pela energia total que deixa dA_i , logo:

$$F_{dA_i-dA_j} = \frac{\text{Energia que atinge } dA_j \text{ proveniente de } dA_i}{\text{Energia Total que deixa } dA_i} \quad (\text{A.16})$$

De acordo com a equação A.2 a energia que deixa dA_i e atinge dA_j é:

$$d\phi(dA_i, d\vec{\omega}_{ij}) = L(dA_i, d\vec{\omega}_{ij}) \cos\theta_i d\omega_{ij} dA \quad (\text{A.17})$$

E a energia total que deixa dA_i pode ser calculada usando a equação A.9:

$$B(dA_i) = L(dA_i, d\vec{\omega}_{ij}) \pi \quad (\text{A.18})$$

Portanto o fator-de-forma entre dA_i e dA_j é encontrado ao substituímos as equações A.17 e A.18 na equação A.16. A equação final gerada, após ser feita também a substituição do valor de $d\omega_{ij}$, é:

$$F_{dA_i-dA_j} = \frac{L(dA_i, d\vec{\omega}_{ij}) \cos\theta_i d\omega_{ij} dA}{L(dA_i, d\vec{\omega}_{ij}) \pi} = \frac{\cos\theta_i \cos\theta_j dA_j}{\pi r^2} \quad (\text{A.19})$$

Esta equação, apesar de importante, não provê uma informação muito relevante nos

métodos de Radiosidade, isso devido ao fato de que a maioria dos métodos necessita do Fator-de-Forma entre duas áreas finitas ou entre uma área finita e uma diferencial.

Porém, tomando a equação A.19 como base é possível deduzir a equação desses dois Fatores-de-forma. O fator-de-forma entre uma área diferencial dA_i e uma área finita A_j pode ser encontrado ao integrarmos a equação A.19 sobre A_j :

$$F_{dA_i-A_j} = \int_{A_j} \frac{\cos\theta_i \cos\theta_j dA_j}{\pi r^2} \quad (\text{A.20})$$

O fator-de-forma entre duas áreas finitas A_i e A_j pode ser calculado integrando a equação A.20 sobre A_i e fazendo a média pela área de A_i . É necessário perceber que o motivo pelo qual isso é possível é por que estamos tratando com superfícies Lambertianas, ou seja, a quantidade de energia que deixa A_i é independente da direção $d\omega$ e é constante em toda a extensão de A_i . A equação para o Fator-de-Forma entre áreas finitas é então:

$$F_{A_i-A_j} = \frac{1}{A_i} \int_{A_i} \int_{A_j} \frac{\cos\theta_i \cos\theta_j}{\pi r^2} dA_j dA_i \quad (\text{A.21})$$

O cálculo dos fatores-de-forma é o núcleo de todo método de Radiosidade, esses cálculos geram os maiores impactos nos métodos, sejam eles de performance ou de precisão. As equações deduzidas aqui não podem ser usadas na prática por questões de performance, uma vez que elas integram em um domínio contínuo, e por essa razão vários métodos foram desenvolvidos para calcular os Fatores-de-Forma de uma maneira mais eficiente.