

03063  
7  
Reje



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**

Unidad Académica de los Ciclos Profesional y Posgrado  
del Colegio de Ciencias y Humanidades

Instituto de Investigaciones en Matemáticas  
Aplicadas y Sistemas

# Radiofidad

Tesis para la obtención del grado de  
Maestro en Ciencias de la Computación

presentada por

**Sergio Santamaría Sánchez**

TESIS CON  
FALLA DE ORIGEN

Julio de 1994



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Quisiera agradecer a mis profesores, compañeros y personal de la maestría,  
muy en especial a mi directora de tesis**

**Ana Luisa Solís,**

**y mis sinodales**

**Hanna Oktaba, Homero Vladimir Ríos**

**Amparo López Gaona y María Garza.**

**También quisiera agradecer a mi familia por el apoyo incondicional que  
me han brindado en todo momento.**

# CONTENIDO

<b>INTRODUCCION</b>	<b>1</b>
<b>CAPITULO I</b>	
<b>CONCEPTOS BASICOS</b>	<b>3</b>
1.1 Introducción	3
1.2 Caracterización de la energía de la luz	3
1.3 Angulos Sólidos	5
1.4 Radiancia	6
1.5 Irradiancia	8
1.6 Intensidad	8
1.7 Caracterización de las Funciones de Reflexión	9
1.7.1 Función de Distribución de la Reflexión Bidireccional	9
1.8 Reflectancia	11
1.9 Ecuación de "Rendering"	14
1.10 Iluminación Directa o Local	14
1.11 Iluminación Indirecta o Global	15
<b>CAPITULO II</b>	
<b>RADIOSIDAD</b>	<b>19</b>
2.1 Introducción	19
2.2 Discretización de la Ecuación de Rendering	20
2.3 Solución del Sistema de Ecuaciones	22
2.4 Factores de Forma	23
2.4.1 Hemisferio	25
2.5 Otras Posibilidades para el Cálculo de los Factores de Forma.	26
2.5.1 Muestreo Adaptativo del Semiespacio	26
2.5.2 Factores de Forma Utilizando "Ray Tracing"	29
2.5.2.1 Factores de Forma desde un Area Diferencial a todos los "Patches"	29
2.5.2.2 Factores de Forma desde un Punto a un "Patch"	30
2.6 Refinamiento Progresivo	32
2.6.1 Acelerando la Convergencia	34
2.6.2 Término Ambiental	35
2.7 Radiosidad Incremental	36

	Contenido
2.7.1 Cambiando los Atributos de la Superficie	37
2.7.2 Cambiando la Geometría de la Escena	38
2.7.3 Factores de Forma Incrementales	40
2.7.4 Colas de Cambios en la Geometría	40
<b>CAPITULO III</b>	
<b>ESPECULARIDAD Y TECNICAS HIBRIDAS</b>	<b>43</b>
3.1 Introducción	43
3.2 Métodos de Radiosidad Direccionales Independientes del Observador	44
3.2.1 Radiosidad Direccional	44
3.2.2 Refinamiento Iterativo de Radiosidades Direccionales	45
3.3 Métodos Dependientes del Observador.	45
3.3.1 Técnicas de "Ray Tracing" y Monte Carlo	45
3.4 Técnicas Híbridas	46
3.4.1 Método de Dos Pasos	46
3.4.2 Método de Dos Pasos Extendido	47
3.4.2.1 Utilización de la Ecuación de Rendering	48
3.4.2.2 Primer Paso: Iluminación Difusa	49
3.4.2.3 Segundo Paso: Distribución Direccional	50
3.4.2.4 Caso General de Reflexión Especular	50
3.4.2.5 Generación de los Factores de Forma Extendidos	51
<b>CAPITULO IV</b>	
<b>CARACTERISTICAS DEL SISTEMA</b>	<b>52</b>
4.1 Introducción	52
4.2 Plataforma de Desarrollo	52
4.3 Modelo de Datos	53
4.3.1 Subestructuración de Patches y Elementos	53
4.3.2 Tipos de Datos	55
4.4 Esquema del Sistema	57
4.4.1 Inicialización del sistema	58
4.4.2 Proceso del Refinamiento Progresivo	61
4.4.3 Proceso Interactivo de Visualización de la Escena	65
4.5 Resultados	67
<b>CONCLUSIONES</b>	<b>80</b>
<b>BIBLIOGRAFIA</b>	<b>81</b>

# INTRODUCCION

La graficación por computadora particularmente la síntesis de imagen ha evolucionado rápidamente, de un enfoque puramente empírico a uno de carácter analítico, introduciendo los avances obtenidos de la investigación de los procesos físicos que ocurren en la transformación de la energía luminosa y su interacción con diversos materiales, así como los procesos de percepción mediante los cuales es posible que el ojo humano capte estas interacciones.

La *síntesis de imagen* es un área de la graficación por computadora que se avoca a la tarea de desarrollar algoritmos que permitan imprimir mayor realismo a las imágenes creadas por computadora; es aquí donde entran las técnicas de iluminación, proporcionando distintos grados de simulación y realismo. En la realidad, un modelo de iluminación completo resulta ser demasiado complicado y su evaluación sería imposible de determinar de forma exacta, por lo que se recurre a disminuir su complejidad, aproximando la simulación de sólo una parte de los procesos.

Históricamente, los modelos de iluminación empezaron por simular únicamente las reflexiones difusas de las superficies, y sólo se consideraban las fuentes de luz directas, es decir aquellos objetos que emiten energía y cuya energía llega directamente a la superficie. Un ejemplo de este tipo de reflexión es el que ocurre en los materiales opacos como el yeso. Este modelo aún cuando crea la sensación de volumen tiene una apariencia apagada. A esto le siguió la introducción de la reflexión especular en la simulación. Esta reflexión es de tipo direccional y depende de la posición del observador; este fenómeno se encuentra en materiales brillantes por ejemplo metales. Al igual que en el modelo que le precedió, este modelo asume que las fuentes de luz son puntuales y directas y en caso de que se manejen sombras, estas son muy marcadas. Con este modelo se realizó el volumen, pero aún la apariencia era "irreal"; irreal en el sentido de que pareciera que los objetos estuvieran aislados, como si estuvieran flotando en el espacio y no existieran otros objetos a su alrededor. Después se añadieron los efectos de refracción, tanto difusa como especular, completando la simulación, pero aún con la imposición de considerar sólo las fuentes de iluminación directas. Con este tipo de modelo es posible crear imágenes

de muy alta calidad cuando los objetos se encuentran alejados uno de otro, ya que este caso se aproxima a la suposición del aislamiento de los objetos. Por otra parte, cuando existen objetos cercanos, la imagen generada no concuerda con la percepción cotidiana de objetos en tales circunstancias; esto es debido a que únicamente se están considerando los emisores de luz, sin embargo, todas las superficies del ambiente al reflejar parte de la energía que reciben de los emisores naturales, se convierten a su vez en emisores, y cuando se encuentran cerca de otro objeto afectan de forma perceptible su iluminación, estas son las *interreflexiones* que no se consideraban en el modelo de iluminación.

Resulta que estas interreflexiones entre superficies son muy relevantes en el cálculo de un modelo de iluminación más acorde a la realidad, pero como todo fenómeno natural, su evaluación sobrepasa fácilmente las capacidades de cálculo de una computadora debido a las técnicas empleadas para evaluar el modelo de iluminación, que hasta ese momento se empleaban, principalmente *Ray Tracing*.

Por lo tanto, era necesario modificar la técnica de evaluación del modelo de iluminación, para que fuera factible modelar estas interreflexiones, tan importantes sobre todo en ambientes cerrados. Es aquí donde surge el método de radiosidad, que además de tener un enfoque básicamente analítico, permite el cálculo de las interreflexiones a un costo computacional no tan excesivo como en las técnicas previas.

La *radiosidad* que es el tema de este trabajo, se puede ver en dos contextos distintos. Primero como una cantidad física relacionada con la energía en interacción entre diversos objetos, y segundo como un modelo de iluminación global basado en la conservación de la energía, que permite crear un algoritmo para simular la interacción de la luz entre objetos en una escena y con esta simulación crear una imagen de muy alta calidad.

El objetivo de este trabajo es crear un sistema que emplee la radiosidad para el cálculo de la iluminación en un ambiente, aprovechando las ventajas del método, y que con los valores obtenidos se generen tanto imágenes estáticas del ambiente iluminado como animaciones interactivas del ambiente, tal como es visto por un observador virtual controlado por el usuario. Para esto el trabajo se dividió en cuatro partes: la primera parte es una introducción a los conceptos físicos empleados para fundamentar el método; la segunda parte es la descripción del método, así como algunas mejoras y modificaciones que han surgido desde su aparición; en la tercera parte, se ven algunas extensiones al método de *radiosidad* y el acoplamiento con la técnica de *ray tracing*, para la creación de un modelo completo; por último en la cuarta parte se describe el sistema desarrollado en detalle y se dan los resultados de su aplicación.

# CAPITULO I

## CONCEPTOS BASICOS

### 1.1 Introducción

La investigación del proceso de generación de imágenes ha progresado hasta el punto que ahora es posible crear imágenes de apariencia muy real. Este progreso ha sido dirigido en los últimos años por un mejor entendimiento de los procesos físicos de iluminación y la física de materiales, así como de los aspectos psicofísicos de la percepción humana.

El problema de la creación de imágenes de tipo real basadas en física, se puede ver como el problema de la simulación de la interacción de la luz con los materiales en un ambiente. Para poder hacer uso de esto, a continuación se darán las ideas básicas desde el punto de vista teórico y conceptual.

### 1.2 Caracterización de la energía de la luz.

La luz es una forma de radiación electromagnética, que a su vez es una onda senoidal formada por campos eléctricos y magnéticos perpendiculares. Las radiaciones electromagnéticas pueden existir en cualquier longitud de onda, clasificándolas de largas a cortas se tienen las ondas de radio, microondas, infrarrojas, luz visible, ultravioleta, rayos X, rayos gamma y rayos cósmicos. El hecho de que la luz haya sido estudiada extensivamente por la física, le da a la graficación por computadora el beneficio de incorporar la teoría y los algoritmos de otras disciplinas, en particular de la física de la transferencia de calor.

Una fuente de luz pura, tal como la producida por un láser, consiste de luz en una sola frecuencia o longitud de onda, sin embargo en el mundo real, la luz casi siempre existe

como una mezcla de diferentes longitudes de onda. Otra propiedad de la luz láser es su coherencia, es decir que la onda mantiene su fase cuando ésta se propaga, a diferencia de la luz natural que es incoherente.

La radiación electromagnética puede también ser polarizada. Esto se refiere a la orientación preferencial de los vectores de los campos eléctrico y magnético, perpendiculares a la dirección de propagación. La polarización de la radiación puede ser un parámetro importante afectando la reflexión de la luz en una superficie.

La *radiometría* es la ciencia de la medición física de la energía electromagnética en cualquier longitud de onda. Una medida radiométrica es expresada en el sistema internacional de unidades por energía o potencia, Joules o Watts respectivamente. Es posible medir la cantidad de luz en cada longitud de onda usando un espectro radiómetro, obteniendo una gráfica de las medidas del espectro de la fuente.

La *fotometría* por otra parte es la medición psicofísica de la sensación visual producida por el espectro electromagnético. Nuestros ojos son sensibles sólo a una parte del espectro, aquella comprendida entre el ultravioleta (380 nanómetros) y el infrarrojo (770 nanómetros). Una de las características más prominentes de la luz formada con mezclas de diferentes longitudes de onda es que aparecen con diferentes colores, luminosidad y brillantez.

Las propiedades radiométricas de un espectro en particular pueden ser convertidas a cantidades fotométricas, usando la curva de eficiencia relativa de luminosidad del espectro. Esta curva da la luminosidad relativa de cada longitud de onda visible. La respuesta del ojo a un espectro es la suma, o integral, de sus respuestas a cada banda espectral. Esto a su vez es igual a la cantidad de energía en esta longitud de onda por su luminosidad relativa:

$$\int_{380}^{770} V(\lambda)S(\lambda)d\lambda$$

donde:

$V$  es la eficiencia relativa  
 $S$  es la energía espectral

La unidad fotométrica para la potencia luminosa es el *lumen* y la unidad para energía luminosa es el *Talbot*. Nuestro ojo es más sensible a la luz amarillo-verdosa con una longitud de aproximadamente 555 nm, que tiene una luminosidad de 684 lumens por watt. La luz de cualquier otra longitud de onda y por lo tanto de cualquier mezcla de luz, dará menos lumens por watt.

Cuando se habla de luz, usualmente potencia y energía se pueden usar indistintamente, esto es debido a que la velocidad de la luz es tan rápida que ésta alcanza su equilibrio inmediatamente. Hay situaciones en las cuales el término energía debe ser usado en vez de potencia. Por ejemplo, la respuesta de una película fotográfica es proporcional a la energía total recibida, y la integral de potencia sobre tiempo es llamada la exposición.

Un principio importante que debe ser observado por cada sistema físico es la conservación de la energía. Esto se aplica a 2 niveles, uno macro o global y otro micro o local.

- En el nivel global, la potencia total puesta en el sistema por las fuentes de luz, debe ser igual a la potencia absorbida por las superficies. En esta situación la energía es conservada.
- En el nivel local, la energía que fluye hacia adentro de una región del espacio o sobre un elemento de superficie debe ser igual a la energía que fluye hacia afuera. Debemos considerar cualquier cambio en el flujo de luz localmente para asegurar que la energía es conservada. Así, la cantidad de luz absorbida, reflejada y transmitida nunca debe ser mayor que la cantidad de luz incidente.

Para poder considerar la conservación de la energía, es necesario cuantificar la energía incidente para cada evento de iluminación presente, tomando como evento de iluminación cada superficie que emita energía no importando su naturaleza. Al cuantificar la energía incidente de un evento en una superficie, invariablemente se encontrará con la cuantificación del porcentaje del área, sobre el hemisferio de iluminación, cubierta por el evento, y una forma de establecerla es mediante el empleo del ángulo sólido que sustenta el evento.

### 1.3 Ángulos Sólidos

Una dirección en  $R^3$  está definida por un vector  $\omega$ . Cuando este vector es unitario, este puede ser representado por un punto en la esfera unitaria, los puntos en la esfera a su vez pueden ser representados por dos ángulos: el ángulo formado con el cenit  $\theta$ , y el ángulo respecto al eje X,  $\phi$ . Las direcciones  $\omega$  y las coordenadas esféricas pueden ser usadas indistintamente.

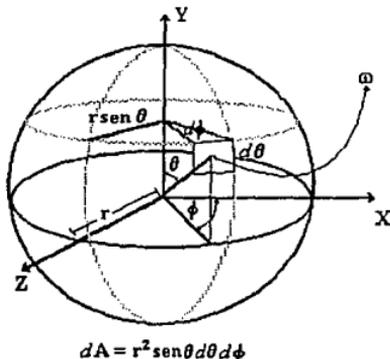


Figura 1.1. Coordenadas esféricas y diferenciales de ángulos sustentados por un área.

Una gran ventaja de pensar las direcciones como puntos en la esfera viene cuando pensamos en distribuciones diferenciales de direcciones. Una distribución de

direcciones puede ser representada por una pequeña región en la esfera unitaria. El área de un elemento de superficie diferencial en una esfera de radio  $r$  es:

$$dA = (r \cdot d\theta)(r \cdot \text{sen}\theta \cdot d\phi) = r^2 \text{sen}\theta d\theta d\phi \quad (1.1)$$

Aquí  $r \cdot d\theta$  es la longitud del arco latitudinal generado entre  $\theta$  y  $\theta + d\theta$ , similarmente  $r \cdot \text{sen}\theta \cdot d\phi$  es la longitud del arco longitudinal generado entre  $\phi$  y  $\phi + d\phi$  (figura 1.1). El producto de estas dos longitudes es el área diferencial de este pedazo de esfera.

Esta derivación usa la definición del ángulo en radianes; el ángulo definido por un arco de longitud  $l$  es igual a  $l / r$ .

El círculo sustenta un ángulo de  $2\pi$  radianes porque la circunferencia del círculo es  $2\pi r$ . Usando una idea similar, podemos definir un ángulo sólido. El *ángulo sólido* sustentado por una área esférica  $a$  es igual a  $(a / r^2)$ . Esta cantidad es la medida del ángulo en esterradianes (radianes cuadrados). Una esfera tiene un área total de  $4\pi r^2$ , así que una esfera tiene  $4\pi$  esterradianes.

Un diferencial de ángulo sólido, el cual indicamos por  $d\omega$ , está dado por

$$d\omega = \frac{dA}{r^2} = \text{sen}\theta d\theta d\phi \quad (1.2)$$

Es conveniente pensar en el diferencial de ángulo sólido como un vector  $d\omega$ . La dirección de  $d\omega$  es la dirección de el punto en la esfera, y la longitud de  $d\omega$ , es igual al tamaño de el diferencial de ángulo sólido en esta dirección.

#### 1.4 Radiancia

La teoría de transporte puede ser aplicada al transporte de luz si se considera a la luz como fotones. Sin embargo, en un sistema de generación de imágenes no se considera la naturaleza cuántica de la luz, así que cuando se habla de transporte de luz, o sea la forma en que se distribuye la luz, se refiere a la potencia  $\Phi$ , o energía radiada por unidad de tiempo, en vez de partículas. Claro que estas cantidades están relacionadas. La energía radiada por unidad de volumen es simplemente la densidad de fotones en el volumen por la energía de cada fotón ( $hf$  donde  $h$  es la constante de Planck y  $f$  es la frecuencia).

$$\frac{d\Phi}{dV} = n(\vec{x}, \vec{v})hf \quad (1.3)$$

donde:  $n(x, v)$  es la cantidad de fotones por unidad de volumen con un vector de velocidad ( $v$ ) dado para un punto ( $x$ ).

Usando el hecho de que

$$n(\vec{x}, \vec{v})d^3v = n(\vec{x}, \vec{\omega})c^2d\omega \quad (1.4)$$

para cambiar de la representación de velocidad a la representación angular de direcciones  $n(x, \omega)$ , donde  $c$  es la velocidad de la luz [Nicodemus77].

$$L(\vec{x}, \vec{\omega}) = n(\vec{x}, \vec{\omega})hfc^3 \quad (1.5)$$

Donde el término radiométrico para esta forma de la densidad del flujo angular es la *radiancia*  $L(x, \omega)$ . Siendo ésta una de las cantidades más importantes en la síntesis de imagen, así que se dará una definición más precisa.

La *radiancia* es la potencia por unidad de área proyectada perpendicular al rayo por unidad de ángulo sólido en la dirección del rayo.

La distribución de la radiancia caracteriza completamente a la distribución de la luz en una escena. Todas las demás cantidades pueden ser calculadas a partir de esta, por ejemplo, el flujo total en un rayo cuya sección transversal tiene área  $dA$  y ángulo sólido  $d\omega$  es:

$$d\Phi = L(\vec{x}, \vec{\omega})d\vec{\omega} \cdot d\vec{A} = L(\vec{x}, \vec{\omega})\cos\theta d\omega dA \quad (1.6)$$

En la expresión anterior aparece el término  $\cos\theta d\omega$ , el cual es llamado el *ángulo sólido proyectado*. Este puede ser pensado como la proyección de un área diferencial sobre un esfera en el plano del ecuador de la esfera. Por lo que el ángulo sólido proyectado de todo el hemisferio ( $\Omega$ ) es  $\pi$ .

$$\begin{aligned} \int_{\Omega} \cos\theta d\omega &= \int_0^{2\pi} \int_0^{\pi} \cos\theta \sin\theta d\theta d\phi \quad (1.7) \\ &= - \int_0^{2\pi} \int_0^{\pi} \cos\theta d \cos\theta \\ &= -2\pi \left. \frac{\cos^2\theta}{2} \right|_0^{\pi/2} \\ &= \pi \end{aligned}$$

Para entender la importancia de la radiancia, hay que considerar las siguientes dos propiedades:

1. La radiancia en la dirección de un rayo de luz se mantiene constante conforme éste se propaga a lo largo del rayo (asumiendo que no hay pérdida debido a absorción o dispersión). Esto es una consecuencia de la conservación de la energía dentro de un delgado hilo de luz. Como una consecuencia de esta propiedad, la radiancia es la cantidad numérica que se debe asociar con un rayo en un "ray tracer".

2. La respuesta de un sensor es proporcional a la radiancia de la superficie visible al sensor. Como una consecuencia de este hecho, la radiancia de una superficie al ojo es la cantidad que debe mandarse al dispositivo de despliegue.

### 1.5 Irradiancia

Otra cantidad importante es la densidad de flujo total incidente dentro de la superficie con una orientación fija ( $E$ ). Esto puede ser calculado integrando la radiancia incidente sobre el hemisferio.

$$E = \int_{\Omega} L \cos\theta d\omega \quad (1.8)$$

La potencia total que llega al elemento de superficie es entonces  $E dA$ .

En radiometría se llama irradiancia a esta cantidad y en fotometría es llamada luminiscencia y es medida en luxes. La distribución de irradiancia en una superficie frecuentemente es trazada usando una gráfica de curvas de nivel.

### 1.6 Intensidad

La radiancia es una forma muy útil para caracterizar el transporte de luz entre elementos de superficie. Desafortunadamente, es difícil describir con radiancia la distribución de energía en una fuente de luz puntual, debido a la singularidad del punto en la fuente. Afortunadamente, es muy fácil caracterizar la distribución de energía si se introduce otra cantidad, "la intensidad luminosa o radiante".

Hay que notar que este uso del término "intensidad" es muy diferente del que comúnmente se utiliza en la comunidad de graficación por computadora. Aún más confusiones surgen del hecho de que "intensidad" es usado para indicar diferentes cantidades de transporte en distintas áreas de trabajo. La intensidad radiante es bastante similar a la usada en la óptica geométrica.

La distribución de la energía desde una fuente de luz puntual se expande al exterior desde el centro. Para un haz pequeño definido por un ángulo sólido en una dirección dada, el flujo en el haz es definido como

$$d\Phi = I(\omega) d\omega \quad (1.9)$$

donde  $I$  es la intensidad radiante de la fuente de luz puntual y tiene unidades de flujo por unidad de ángulo sólido.

La intensidad radiante en una dirección dada es igual a la irradiancia en un punto sobre la esfera unitaria centrada en la fuente. En la óptica geométrica, la intensidad es la potencia por unidad de área ( en vez de por unidad de ángulo sólido). En el caso de un frente de onda esférico (es decir cuando la distribución se hace de forma radial) emanando desde una fuente puntual, la definición de la óptica geométrica es básicamente la misma que la definición radiométrica. Sin embargo la definición de la óptica geométrica es ligeramente más general. El frente de onda esférico emanando de una fuente puntual, puede reflejar o refractar cuando éste interactúa con otra superficie, entonces es distorsionado y se dificulta el seguimiento del ángulo sólido

sustentado por el frente distorsionado. Pero es fácil seguir el área diferencial del frente. De hecho lo que pasa es que la intensidad se vuelve proporcional a la curvatura gaussiana del frente de onda [Mitchell92].

Para una fuente de luz puntual isotrópica la intensidad radiante es:

$$I = \frac{\Phi}{4\pi} \quad (1.10)$$

Claro que una fuente puntual puede actuar como una fuente dirigida y radiar luz en cantidades diferentes para diferentes direcciones. La energía total emitida ( $\Phi$ ) en este caso es

$$\Phi = \int_{\Omega} I(\omega) d\omega \quad (1.11)$$

La irradiancia ( $E$ ) en una superficie diferencial debida a una fuente de luz puntual simple puede ser calculada obteniendo el ángulo sólido sustentado por el elemento de superficie desde el punto de vista de la fuente de luz.

$$E = I \frac{d\omega}{dA} = \frac{\Phi}{4\pi} \cdot \frac{\cos\theta}{|x - x_s|^2} \quad (1.12)$$

donde  $|x - x_s|$  es la distancia del punto a el elemento de superficie.

## 1.7 Caracterización de las Funciones de Reflexión

La reflexión es definida como el proceso por el cual la luz incidente sale de la superficie por el mismo lado. La transmisión, absorción, efectos de polarización y espectrales, fluorescencia y fosforescencia son todos importantes de considerar al desarrollar un modelo de interacción de la luz con los materiales. En vez de revisar modelos concretos en esta parte se considerarán nomenclatura y propiedades generales que son satisfechas por todas las funciones de reflexión.

### 1.7.1 Función de Distribución de la Reflexión Bidireccional

Al considerar la luz incidente en una superficie desde un ángulo sólido diferencial pequeño en la dirección  $\omega_i$ , se ha observado que la cantidad de luz reflejada en otra dirección  $\omega_r$ , es proporcional a la irradiancia incidente desde  $\omega_i$ . Esto es:

$$dL_r(\omega_r) \propto dE_i(\omega_i)$$

Esto establece que si se incrementa la energía de luz incidente por unidad de área, entonces la energía de la luz reflejada es incrementada. De la geometría de la situación, se puede incrementar la irradiancia incidente incrementando ya sea el ángulo sólido sustentado por la fuente, el ángulo de incidencia, o la densidad de energía de la luz.

La constante de proporcionalidad es llamada la función de distribución de la reflexión bidireccional o FDRB

$$f_r(\bar{\omega}_i \rightarrow \bar{\omega}_r) \equiv \frac{L_r(\bar{\omega}_r)}{L_i(\bar{\omega}_i) \cos \theta_i d\omega_i} \quad (1.13)$$

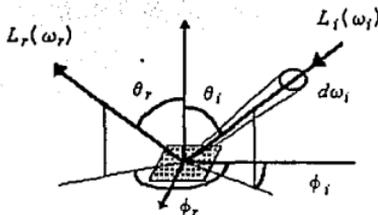


Figura 1.2. Geometría de la reflectancia bidireccional.

Más precisamente, la FDRB es definida como la razón de la radiancia reflejada en la dirección  $\omega_r$  y la irradiancia desde la dirección incidente  $\omega_i$  que la produce [Cohen92]. La FDRB es bidireccional porque depende de dos direcciones. Frecuentemente la dependencia de los cuatro ángulos se hace explícita y la FDRB es escrita como  $f_r(\theta_i, \phi_i; \theta_r, \phi_r)$ . La FDRB es una función de distribución porque es estrictamente positiva, pudiendo tomar valores entre cero e infinito.

Ahora bien, si se agrega luz desde una sola dirección, entonces para otras direcciones incidentes no tiene efecto alguno sobre la cantidad de luz reflejada desde esas direcciones. Así la reflexión se comporta linealmente y de hecho, la cantidad total de luz reflejada por una superficie en una dirección dada está definida por la integral hemisférica sobre todas las posibles direcciones incidentes. Esto nos lleva a la ecuación de reflectancia:

$$L_r(\bar{\omega}_r) = \int_{\Omega} f_r(\bar{\omega}_i \rightarrow \bar{\omega}_r) L_i(\bar{\omega}_i) \cos \theta_i d\omega_i \quad (1.14)$$

Aún cuando se podría pensar que ésta es la forma más general de la ecuación de reflectancia, no lo es. En algunas situaciones, la luz penetra una capa de la superficie, rebota en esta capa y después sale por otro punto. La generalización de la función de reflectancia para manejar esta situación es llamada función de distribución de la reflexión bidireccional dispersada en la superficie (FDRBDS) [Cohen92].

Las FDRB tienen varias propiedades interesantes:

1. Si la FDRB está basada en leyes físicas, entonces esta no es alterada si las direcciones incidente y reflejada son intercambiadas. Esto es

$$f_r(\bar{\omega}_i \rightarrow \bar{\omega}_r) = f_r(\bar{\omega}_r \rightarrow \bar{\omega}_i) \quad (1.15)$$

Esto es llamado el principio de reciprocidad de Helmholtz, y es equivalente a decir que si un fotón se mueve en una trayectoria, éste se moverá en la misma trayectoria si su dirección es invertida.

2. La FDRB es en general anisotrópica. Esto es, si se fijan las direcciones incidente y reflejada y se rota la superficie de interacción alrededor de su normal, el porcentaje de la luz reflejada puede cambiar. Unos ejemplo de materiales anisotrópicos son el aluminio lijado o las telas.

No obstante, hay muchos materiales que son lisos y su reflectividad no depende de la orientación de la superficie. Por lo que su función de reflexión no cambia si la superficie es rotada.

Como un ejemplo de FDRB, considérese un espejo perfecto. Primero, considerando la geometría de la reflexión para un espejo, el ángulo de reflexión es igual al ángulo de incidencia y el vector reflejado está en el mismo plano que el determinado por el vector incidente y el vector normal de la superficie. Esto implica que:

$$\begin{aligned}\theta_r &= \theta_i \\ \phi_r &= \phi_i \pm \pi\end{aligned}\quad (1.16)$$

Segundo, considerando la radiometría de la reflexión. Para un espejo, la radiancia reflejada es exactamente igual a la radiancia incidente.

$$L_r(\theta_r, \phi_r) = L_i(\theta_i, \phi_i \pm \pi) \quad (1.17)$$

Este hecho físico puede ser expresado matemáticamente con una FDRB que contiene funciones delta de Dirac, donde  $f_{r,e}$  es la FDRB del caso especular (espejo)

$$f_{r,e} = \frac{\delta(\cos\theta_i - \cos\theta_r)}{\cos\theta_i} \delta(\phi_i - (\phi_r \pm \pi)) \quad (1.18)$$

La función delta tiene las siguientes propiedades:

$$\delta(x) = 0 \Rightarrow x \neq 0 \quad (1.19)$$

$$2. \int_{-\infty}^{\infty} \delta(x) dx = 1 \quad (1.20)$$

$$3. \int_{-\infty}^{\infty} \delta(x-y) f(x) dx = f(y) \quad (1.21)$$

Para comprobar que esto da la radiancia reflejada correcta se realiza la integración hemisférica.

$$L_r(\theta_r, \phi_r) = \int_{\Omega_i} \frac{\delta(\cos\theta_i - \cos\theta_r)}{\cos\theta_i} \delta(\phi_i - (\phi_r \pm \pi)) L_i(\theta_i, \phi_i) \cos\theta_i d\cos\theta_i d\phi_i \quad (1.22)$$

## 1.8 Reflectancia

El hecho de que la función delta puede ser interpretada como un pico infinitesimalmente delgado e infinitamente alto con área uno, esto implica que la FDRB aunque es siempre positiva, puede ser infinita. Comúnmente es más intuitivo trabajar

con cantidades acotadas entre 0 y 1. Esta cantidad es llamada la *reflectancia bicónica* o simplemente *reflectancia* [Cohen92].

Considérese la razón del flujo reflejado a el flujo incidente (ec.1.23), como el flujo reflejado debe ser menor que el flujo incidente que le dió origen, la reflectancia debe ser siempre menor que 1.

$$\begin{aligned} \frac{d\Phi_r}{d\Phi_i} &= \frac{\int_{\Omega_r} L_r(\omega_r) \cos\theta_r d\omega_r}{\int_{\Omega_i} L_i(\omega_i) \cos\theta_i d\omega_i} \\ &= \frac{\int_{\Omega_r, \Omega_i} f_r(\omega_i \rightarrow \omega_r) L_i(\omega_i) \cos\theta_i d\omega_i \cos\theta_r d\omega_r}{\int_{\Omega_i} L_i(\omega_i) \cos\theta_i d\omega_i} \end{aligned} \quad (1.23)$$

Desafortunadamente, esta cantidad depende de la distribución de luz que llega  $L_i$ . Sin embargo si asumimos que  $L_i$  es uniforme e isotrópica, entonces es posible sacarla de la integral tanto del denominador como del numerador y cancelarla. Así entonces se llega a la relación entre reflectancia y FDRB que forma la definición de reflectancia ( $\rho$ ).

$$\rho(\bar{\omega}_i \rightarrow \bar{\omega}_r) \equiv \frac{\int_{\Omega_r, \Omega_i} f_r(\omega_i \rightarrow \omega_r) \cos\theta_i d\omega_i \cos\theta_r d\omega_r}{\int_{\Omega_i} \cos\theta_i d\omega_i} \quad (1.24)$$

La reflectancia contiene una doble integral sobre las direcciones incidente y reflejada. Sin embargo, hasta el momento no se han especificado los límites de integración. Se puede escoger un ángulo sólido diferencial, un ángulo sólido finito o el hemisferio entero. De la elección que se haga para las direcciones incidente y reflejada se tienen 9 reflectancias diferentes.

	$\omega$	$\Delta\omega$	$2\pi$	
$\omega$	$\rho(\omega, \rightarrow \omega_r)$	$\rho(\omega, \rightarrow \Delta\omega_r)$	$\rho(\omega, \rightarrow 2\pi)$	(1.25)
$\Delta\omega$	$\rho(\Delta\omega_i, \rightarrow \omega_r)$	$\rho(\Delta\omega_i, \rightarrow \Delta\omega_r)$	$\rho(\Delta\omega_i, \rightarrow 2\pi)$	
$2\pi$	$\rho(2\pi \rightarrow \omega_r)$	$\rho(2\pi \rightarrow \Delta\omega_r)$	$\rho(2\pi \rightarrow 2\pi)$	

Los nombres de estas reflectancias son formados combinando las palabras: direccional (para ángulos sólidos diferenciales), cónica (para ángulos sólidos finitos) y hemisférica (para un ángulo sólido igual a todo el hemisferio). Así  $\rho(\omega_i \rightarrow \omega_r)$ ,  $\rho(\Delta\omega_i \rightarrow \Delta\omega_r)$  y  $\rho(2\pi \rightarrow 2\pi)$  son referidas como las reflectancias bidireccional, bicónica y bihemisférica. Quizá la función de reflectancia más interesante es la reflectancia hemisférica-direccional  $\rho(\omega_i \rightarrow 2\pi)$ . Esta es la cantidad de luz esparcida en el hemisferio completo desde una sola dirección incidente. Esta cantidad debe ser menor que 1,

debido a la razón de flujos. Si la FDRB obedece el principio de reciprocidad, entonces la reflectancia hemisférica-direccional es igual a la reflectancia direccional-hemisférica. Esto significa que la cantidad total de luz reflejada dentro de todo el hemisferio desde un rayo incidente es igual a la cantidad de luz reflejada dentro del mismo rayo bajo una iluminación uniforme desde todo el hemisferio.

Otra cantidad que es frecuentemente reportada cuando se miden las propiedades reflectivas de los materiales, es la *reflectancia normal*. Esta es la reflectancia bidireccional haciendo las direcciones incidente y reflejada normales a la superficie. Esta es comúnmente medida, porque es la dirección de máxima reflectividad, es fácil de medir y es suficiente para parametrizar las propiedades reflectivas de un material (ya que se asume que se pueden modelar teóricamente las propiedades direccionales).

Para ilustrar la relación entre la FDRB y la reflectancia, considérese el caso de reflectancia difusa ideal. La reflectancia difusa es modelada asumiendo que la luz es igualmente dispersada en todas direcciones, independientemente de la dirección de incidencia. En otras palabras, la FDRB es constante. Así,

$$\begin{aligned} L_{r,d}(\omega_r) &= \int_{\Omega_i} f_{r,d} L_i(\omega_i) \cos\theta_i d\omega_i & (1.26) \\ &= f_{r,d} \int_{\Omega_i} L_i(\omega_i) \cos\theta_i d\omega_i \\ &= f_{r,d} E_i \end{aligned}$$

Esto lleva a las siguientes dos conclusiones:

1. El valor de la radiancia reflejada es proporcional a la irradiancia incidente.
2. Como ni  $f_{r,d}$  ni  $E_i$  dependen de  $\omega_r$ , la radiancia reflejada es una constante y de hecho la misma en todas direcciones. Esto es cierto independientemente de la distribución de la luz que llega a la superficie.

Pero hay que asegurar que la energía se conserva. Esto se hace forzando a la reflectancia hemisférica-hemisférica a ser menor que 1.

$$\begin{aligned} \rho_d(2\pi \rightarrow 2\pi) &= \frac{\Phi_{r,d}}{\Phi_i} = \frac{\int_{\Omega_r} L_{r,d}(\omega_r) \cos\theta_r d\omega_r}{\int_{\Omega_i} L_i(\omega_i) \cos\theta_i d\omega_i} & (1.27) \\ &= \frac{L_{r,d} \int_{\Omega_r} \cos\theta_r d\omega_r}{E_i} = \frac{\pi L_{r,d}}{E_i} = \pi f_{r,d} \end{aligned}$$

Esto lleva inmediatamente a la conclusión de que si la FDRB es una constante, entonces la reflectancia es también constante. Y más importante aún, esta relación se puede usar para parametrizar la FDRB en términos de la reflectancia:  $f_{r,d} = \rho_d \pi$ .

Como la radiancia que sale de la superficie es constante, la irradiancia de salida

$$E_{r,d} = \pi L_{r,d} \quad (1.28)$$

está relacionada a la irradiancia incidente por la siguiente ecuación:

$$\rho_d = \frac{E_{r,d}}{E_i} \quad (1.29)$$

la que establece que para la reflexión difusa, la reflectancia es igual a la irradiancia que sale dividida por la irradiancia incidente. La irradiancia que sale es lo que se conoce en transferencia de calor y graficación por computadora como *radiosidad*.

## 1.9 Ecuación de "Rendering"

La ecuación de reflectancia (1.24) permite calcular la distribución de luz reflejada a partir de la distribución de luz incidente y la FDRB del material, quedando por especificar la distribución de la luz incidente, esto es llamado comúnmente el modelo de iluminación. El primer y más sencillo caso a considerar es la iluminación directa desde las fuentes de luz. En este caso hay normalmente un pequeño número de luces puntuales o distantes, y asumimos que no hay sombras. Como el modelo no considera el ambiente como un todo y sólo depende de las propiedades individuales de las fuentes y de la superficie que es iluminada, se le llama *modelo de iluminación local*. El segundo caso y considerablemente más complejo es la iluminación indirecta. En este caso la luz puede venir de cualquier superficie en el ambiente, y es muy importante considerar las sombras. A este modelo que considera la escena completa, se le llama *modelo de iluminación global*.

### 1.10 Iluminación Directa o Local

Es fácil incorporar la iluminación directa desde las fuentes de luz puntuales en el modelo de reflexión previo.

$$L_r(\omega_r) = \int_{\Omega_i} f_r(\omega_i \rightarrow \omega_r) L_i(\omega_i) \cos\theta_i d\omega_i \quad (1.30)$$

Recordando que se había derivado la irradiancia de una fuente de luz puntual como

$$E = \frac{\Phi \cos\theta}{4\pi |x - x_f|^2} \quad (1.31)$$

si se asume que la dirección a la fuente de luz está dada por  $\omega_f$ , entonces se puede expresar la radiancia desde una fuente de luz puntual usando una función delta

$$L_i(\omega_r) = \frac{\Phi}{4\pi|x-x_f|^2} \delta(\cos\theta_r - \cos\theta_f) \delta(\phi_r - \phi_f) \quad (1.32)$$

Sustituyendo ésta dentro de la ecuación de reflectancia, se tiene

$$\begin{aligned} L_r(\omega_r) &= \int f_r(\omega_i \rightarrow \omega_r) L_i(\omega_i) \cos\theta_i d\omega_i \\ &= \frac{\Phi}{4\pi|x-x_f|^2} \int f_r(\omega_r, \omega_f) \cos\theta_f \end{aligned} \quad (1.33)$$

Así, si se tienen  $n$  fuentes de luz, entonces la integral hemisférica se colapsa a una suma sobre  $n$  fuentes. Este es el modelo usado normalmente por las implantaciones por hardware de las estaciones de trabajo.

Es fácil extender este modelo a fuentes de luz con distribuciones direccionales arbitrarias, así como fuentes de luz distantes. En principio también se pueden usar fuentes de luz lineales y de área, aunque esto implica la integración de la función de reflectancia sobre las posibles direcciones incidentes desde la fuente de luz.

### 1.11 Iluminación Indirecta o Global

El primer paso en el uso de un modelo global es relacionar el ambiente de iluminación en una superficie a la distribución de luz reflejada desde otra superficie. Esto requiere que se haga la dependencia espacial de la radiancia explícitamente y que se considere la oclusión.

Usando el hecho de que la radiancia es invariante a lo largo de un rayo, se puede escribir la radiancia incidente en  $x'$  debida a la radiancia desde  $x$ .

$$L_i(x', \omega') = L_r(x, \omega_r) V(x, x') \quad (1.34)$$

donde

$$\omega_r = -\omega' = \frac{x-x'}{|x-x'|}$$

La función  $V(x, x')$  es una función de visibilidad. Es 1 si  $x$  y  $x'$  son mutuamente visibles y 0 en otro caso.

Regresando a la función de reflectancia (1.24), el siguiente paso es cambiar la integral hemisférica sobre todas las direcciones incidentes a una integral de área sobre todas las otras superficies en el ambiente. Esto se hace fácilmente relacionando el ángulo sólido sustentado por la fuente al área de la superficie proyectada.

Haciendo el producto punto para formar el ángulo sólido proyectado se tiene

$$d\omega' \cdot dA' = \pi G(x, x') dA \quad (1.35)$$

donde

$$G(x, x') = G(x', x) = \frac{\cos\theta_i \cos\theta_r}{\pi |x - x'|^2} \quad (1.36)$$

$GdA$  es el factor de forma diferencial.

$$\left[ d\omega_i = \frac{\cos\theta_i dA}{|x - x'|^2} \right] \quad (1.37)$$

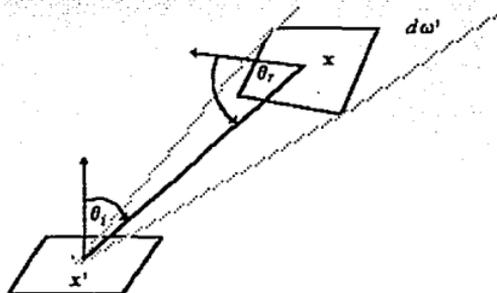


Figura 1.3. Geometría del factor de forma  $G$  entre dos puntos

Haciendo estas sustituciones en la ecuación de reflectancia queda la siguiente ecuación integral.

$$L(x', \omega') = \int_x f_r(x') L(x, \omega) G(x, x') V(x, x') dA$$

En esta ecuación se pueden quitar los subíndices denotando direcciones incidentes o reflejadas ya que sólo involucra radiancias y direcciones reflejadas.

Esta es la conocida como la ecuación de rendering, introducida por Kajiyi [Kajiyi86], aunque la notación es un poco diferente.

Kajiyi también introduce una nueva cantidad de intensidad,  $I(x \rightarrow x')$ , la intensidad de transporte de dos puntos de  $x$  a  $x'$ . Esta cantidad de intensidad es una función de la posición de la superficie, y no involucra ángulos sólidos, lo que a veces es conveniente. La intensidad de transporte de dos puntos es definida por la ecuación:

$$I(x \rightarrow x') dA dA' = \pi L(x, \omega) V(x, x') G(x, x') dA dA' \quad (1.38)$$

Este es el flujo que fluye por el rayo que conecta  $dA$  con  $dA'$ . La ecuación de rendering que se tenía se puede pasar a esta forma multiplicando ambos lados por  $G(x', x) dA' dA$ . Esto lleva a la ecuación

$$I(x' \rightarrow x'') = V(x', x'') G(x', x'') \int_x f_r(x \rightarrow x' \rightarrow x'') I(x \rightarrow x') dA \quad (1.39)$$

Esta ecuación da la cantidad de luz que fluye de  $x$  a  $x'$  y es reflejada a  $x''$  (fig.1.4). Así esta es llamada la *ecuación de transporte de múltiples puntos*.

$$f_r(x \rightarrow x' \rightarrow x'') = f_r(x', \omega_i' \rightarrow \omega_i'') \quad (1.40)$$

y es solo una reparametrización de la FDRB.

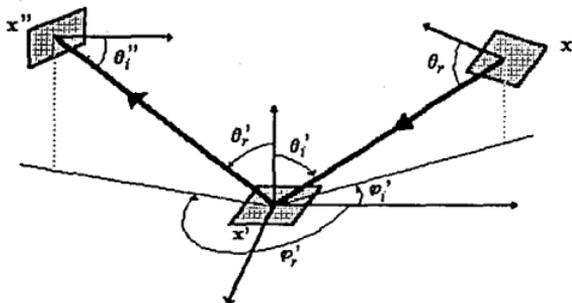


Figura 1.4. Geometría en la reflexión indirecta.

Hay un paso final que se debe tomar antes de tener la ecuación de rendering completa, y es tomar en cuenta todos los otros modos de transporte de luz en una superficie. En un ambiente que consiste solo de superficies opacas (difusas) la única fuente de luz es debida a la emisión de la superficie.

$$I(x' \rightarrow x'') = V(x', x'') \left[ I_e(x' \rightarrow x'') + \int_x f_r(x \rightarrow x' \rightarrow x'') I(x \rightarrow x') dA \right] \quad (1.41)$$

donde  $I_e$  es la intensidad de los puntos de la luz emitida.

En el método de radiosidad se asume que todas las superficies en el ambiente son reflectores difusos. Así la FDRB es independiente de las direcciones de llegada y salida, pero aún depende de la posición.

Finalmente se puede escribir la ecuación de rendering como:

$$\begin{aligned} L(x' \rightarrow x'') &= L_e(x' \rightarrow x'') + \pi f_r(x') \int_x L(x \rightarrow x') G(x \rightarrow x') V(x \rightarrow x') dA \quad (1.42) \\ &= L_e(x' \rightarrow x'') + \rho(x') \int_x L(x \rightarrow x') G(x \rightarrow x') V(x \rightarrow x') dA \end{aligned}$$

Con las suposiciones de radiosidad se puede simplificar esta ecuación aún más, recordando que la radiancia ( $L$ ) de una superficie difusa no depende de la dirección, y

de hecho es igual a la irradiancia de salida dividida por  $\pi$ . Multiplicando la ecuación anterior por  $\pi$  llegamos a la *ecuación de rendering* bajo las suposiciones de radiosidad.

$$B(x') = B_e(x') + \rho(x') \int_x B(x) G(x, x') V(x, x') dA \quad (1.43)$$

donde  $B$  es llamada la *radiosidad*, que es un término de la transferencia de calor y  $B_e$  es la radiancia emisiva.

La ecuación de rendering es un ejemplo de una ecuación integral. La ecuación integral no solo expresa la conservación de la energía localmente, sino que se aplica en cualquier lugar del espacio. La principal característica de una ecuación integral es que la cantidad que se desea calcular ( en este caso la radiancia o radiosidad ) aparece tanto en el lado izquierdo como bajo una integral en el lado derecho. Lo complicado de la solución es porque la integral no puede ser calculada hasta conocer su solución.

# CAPITULO II

## RADIOSIDAD

### 2.1 Introducción

La aplicación de las investigaciones en física y óptica a la graficación por computadora ha constituido una evolución en los modelos de iluminación. Aún cuando en la práctica estos son muy diferentes a sus correspondientes modelos teóricos, los modelos usados hasta ahora caen en una de tres clasificaciones generales (Santamaría90): *empíricos*, *transicionales* y *analíticos*. Los primeros modelos fueron empíricos en naturaleza y son evaluados en una geometría que no correspondía a la original, porque estos se aplican sobre el espacio proyectado en el plano de visión.

Los modelos transicionales usan trabajos previos de física y óptica para mejorar los modelos empíricos y evaluados usando geometría euclidiana en el espacio original, por lo que su iluminación es geoméricamente correcta.

Por su parte los modelos analíticos, hacen del modelo de iluminación el impulsor en la aplicación de las técnicas de equilibrio de energía en la graficación por computadora, además aparte de mantener la geometría original, se debe modelar el movimiento de la energía luminosa a través del ambiente para proporcionar la información requerida para evaluar el modelo de iluminación.

La aplicación de un modelo de iluminación lleva a uno de dos métodos:

Uno que comienza en el ojo y considera las superficies que son visibles desde esta posición. Para cada superficie visible se debe averiguar para cada uno de sus pixels, cual es la información requerida para evaluar el color en esta superficie. Esta

información requerida obliga a considerar otras superficies, y este proceso continúa recursivamente hasta que se satisface algún criterio para no pedir más información adicional. Un ejemplo de esto es *Ray Tracing* [Whitted80].

El otro método empieza en las fuentes de luz y sigue la energía de la luz conforme ésta se mueve a través del ambiente. Cada una de las superficies es revisada respecto a las fuentes y emisores de luz en el ambiente, averiguando cómo la luz es reflejada o transmitida en esta superficie. Después de esto cada superficie iluminada se convierte en un emisor de luz. Una vez más, cada superficie es revisada y el proceso continúa hasta que el equilibrio es alcanzado. Un ejemplo de esto es *Radiosidad*.

El primer método crea un mapeo, dependiente del observador, de la luz que se mueve en el ambiente desde las superficies seleccionadas por el ojo. Así que un mapeo nuevo se debe crear para cada posición del observador. El segundo método crea un mapeo, independiente del observador, de la luz moviéndose en la escena desde las fuentes de luz hacia todas las superficies, pudiéndose usar este mismo mapeo para cualquier posición del observador.

## 2.2 Discretización de la Ecuación de Rendering

La ecuación de rendering que se vió previamente (1.41) define un problema dimensional infinito. Una solución completa requeriría de los valores de radiancias de un número infinito de puntos sobre un número infinito de direcciones. Para hacer factible este problema, el número de valores desconocidos debe ser reducido a algún número finito con el cual se pueda construir una aproximación a la solución. Una forma de aproximar la radiancia entre porciones discretas, o "*patches*", de una superficie, es con una función polinomial determinada por un conjunto pequeño de incógnitas. Esta es la base para la técnica de *radiosidad*.

Las técnicas de radiosidad tradicionales hacen una suposición importante acerca de la direccionalidad de la radiancia sobre la superficie, asumiendo que todas las superficies son reflectores y/o emisores difusos lambertianos (es decir, que la distribución de la luz es igual en todas las direcciones). Esto permite que en un solo valor escalar se represente la radiancia para una longitud de onda en cualquier punto, ya que la radiancia difusa es por definición igual en todas direcciones.

En adelante se asumirá radiancia monocromática sin pérdida de generalidad, en vista de que si no fuera este el caso, sólo se tendría que aplicar lo mismo a las longitudes de onda de interés (aún cuando éstas no estén en el espectro visible).

La suposición de reflectancia difusa permite reescribir la ecuación de rendering (1.41) con un término de reflectancia difusa. Por otra parte se puede cambiar la terminología de radiancia a radiosidad, definiendo la radiosidad como la potencia hemisférica total o irradiancia saliendo de un punto en una superficie. Esto requiere integrar los términos de radiancia por el ángulo sólido proyectado sobre el hemisferio. Al integrar el ángulo sólido proyectado sobre el hemisferio dá un factor de  $\pi$ ,

$$\int_{\Omega} \cos\theta d\omega = \pi \quad (2.1)$$

así la radiancia,  $L$ , es relacionada a la radiosidad,  $B$ , como:

$$B = \pi L \quad (2.2)$$

la radiancia emisiva,  $L_e$ , es relacionada a la emisión,  $E$ , como:

$$E = \pi L_e \quad (2.3)$$

y la reflectancia bidireccional, FDRB, a la reflectancia,  $\rho$ , como:

$$\rho = \pi FDRB \quad (2.4)$$

Así, la ecuación de rendering para ambientes difusos (1.43) es transformada a:

$$B(x) = E(x) + \rho(x) \int_{\Omega} B(x') \frac{\cos\theta \cos\theta'}{\pi r^2} V(x, x') dx' \quad (2.5)$$

o

$$B(x) = E(x) + \rho(x) \int_x B(x') G(x, x') dA \quad (2.6)$$

Hay que notar que  $G(x, x')$  ahora incluye el término de visibilidad,  $V$ .

Pero aún cuando la radiancia a sido reducida a un solo valor escalar en cada punto de la superficie, el problema sigue siendo de dimensión infinita. Este puede ser reducido a uno de dimensión finita discretizando la superficie. Supóngase que las superficies han sido divididas en un conjunto de subregiones, llamadas "patches" (porción o pedazo). Cada uno de estos patches contiene un valor de radiosidad localizado en el centro del patch. En este caso

$B_i$  = valor de radiosidad en el patch  $i$

$\rho_i$  = reflectividad en el patch  $i$

$E_i$  = valor de emisión en el patch  $i$

$A_i$  = área del patch  $i$

$r_{i,j}$  = distancia de  $dA_i$  a  $dA_j$ , anteriormente  $|x - x'|^2$

donde  $F_{i,j}$  es llamado el *factor de forma*, y representa la fracción de energía que sale del patch  $i$  y llega al patch  $j$ . Realizando la doble integral

$$F_{i,j} = \frac{1}{A_i} \iint_{A_i, A_j} G(dA_i, dA_j) dA_i dA_j \quad (2.7)$$

Por 2.6 y 2.7 y discretizando, resulta el sistema de ecuaciones lineales simultáneas:

$$B_i A_i = E_i A_i + \rho_i \sum_j B_j F_{i,j} A_j \quad (2.8)$$

Una interpretación física posible de la ecuación anterior es que la radiosidad de un patch depende de cualquier luz que es emitida directamente o reflejada. Esta luz reflejada, depende por su parte de la luz que sale de cada patch en el ambiente. Una fracción de la luz que sale de cada superficie puede llegar a la superficie en cuestión y ser reflejada de nuevo al ambiente. Esta fracción depende de la relación geométrica entre las superficies, el factor de forma, y la reflectividad del patch. Esta interrelación entre la radiosidad de un patch y las radiosidades de los restantes es expresada en la ecuación anterior.

Hay una relación de reciprocidad clara entre los factores de forma y esta es:

$$F_{j,i}A_j = F_{i,j}A_i \quad (2.9)$$

Así, dividiendo 2.8 por  $A_i$ , resulta en la clásica ecuación de radiosidad [Cohen85]

$$B_i = E_i + \rho_i \sum_j B_j F_{j,i} \quad (2.10)$$

o en su forma matricial

$$\begin{bmatrix} 1 - \rho_1 F_{1,1} & -\rho_1 F_{1,2} & -\rho_1 F_{1,3} & \cdots & -\rho_1 F_{1,N} \\ -\rho_2 F_{2,1} & 1 - \rho_2 F_{2,2} & -\rho_2 F_{2,3} & \cdots & -\rho_2 F_{2,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\rho_{N-1} F_{N-1,1} & -\rho_{N-1} F_{N-1,2} & \cdots & \cdots & 1 - \rho_{N-1} F_{N-1,N} \\ -\rho_N F_{N,1} & -\rho_N F_{N,2} & \cdots & \cdots & 1 - \rho_N F_{N,N} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_{N-1} \\ B_N \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_{N-1} \\ E_N \end{bmatrix} \quad (2.11)$$

Es importante notar en la matriz anterior que  $\rho$  es el único término con dependencia de la longitud de onda, ya que los factores de forma son elementos de carácter puramente geométrico, por lo que si se está trabajando con color RGB, es necesario establecer un sistema por cada una de las tres longitudes de onda involucradas.

### 2.3 Solución del Sistema de Ecuaciones

El sistema de ecuaciones anterior puede ser resuelto con la mayoría de los algoritmos de resolución de ecuaciones lineales, sin embargo hay que notar que de la propiedades especiales de esta matriz es posible encontrar una solución más eficiente.

En particular, la suma de los factores de forma desde cualquier patch a todos los demás, debe ser por definición igual a 1, y la reflectividad es menor que 1, así la suma de los términos, fuera de la diagonal, de cualquier renglón es menor que el de la diagonal, la matriz por lo tanto es siempre diagonal dominante. Esto significa que un método de solución iterativo como Gauss-Seidel está garantizado a converger a la solución [Hornbeck75].

En cada paso, el proceso de Gauss-Seidel involucra el producto interior de un renglón de la matriz con la aproximación actual a las radiosidades. Si la aproximación inicial es igual a los valores de emisión de los patches, entonces cada paso puede ser pensado

como la proyección de la iluminación de la aproximación actual sobre un patch, resultando en la actualización de la radiosidad para este patch.

#### 2.4 Factores de Forma

El factor de forma entre los patches define la fracción de energía que sale de un patch y llega al otro. Los factores de forma han probado ser la parte más costosa computacionalmente de la radiosidad en ambientes complejos. El factor de forma es una función de la relación geométrica entre los patches, por lo tanto no depende de la posición del observador o de las reflectancias de la superficie.

Nuevamente esta es la relación entre los patches:

$$F_{i,j} = \frac{1}{A_i} \iint_{A_i} \iint_{A_j} G(dA_i, dA_j) dA_i dA_j \quad (2.7)$$

donde  $G$  es la relación geométrica entre los patches  $A_i, A_j$ .

El integrando representa el factor de forma entre áreas diferenciales. Este depende de la distancia entre ellas, y su orientación entre sí. También sigue el decaimiento por  $r^2$  conforme la distancia crece, y es proporcional al área proyectada de un área vista desde la otra, relacionándolas al coseno del ángulo entre la normal de la superficie y la línea que las conecta. Esta relación puede ser expresada como:

$$\frac{\cos\theta_i \cos\theta_j}{\pi r_{i,j}^2} \quad (2.12)$$

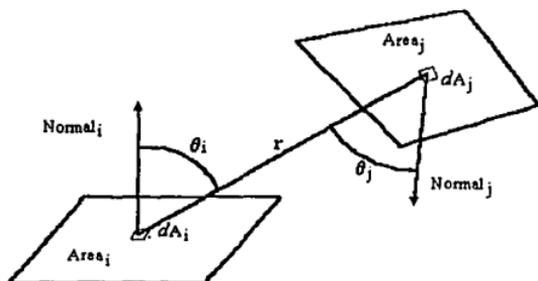


Figura 2.1. Geometría del Factor de Forma entre dos áreas

El factor de forma puede ser encontrado integrando sobre el área  $j$  y tomando el área promedio sobre el área  $i$ . Así, el factor de forma puede ser expresado como una doble integral sobre los dos patches. La ecuación queda entonces como

$$F_{i,j} = \frac{1}{A_i} \iint_{A_i} \iint_{A_j} \frac{\cos\theta_i \cos\theta_j}{\pi r_{i,j}^2} dA_j dA_i \quad (2.13)$$

La ecuación anterior supone que las áreas pueden ser siempre vistas entre sí. Esta no es siempre una suposición válida en ambientes generales. Así que se debe agregar al integrando un término de visibilidad  $V_{ij}$ , de  $dA_i$  a  $dA_j$  tomando valores 0 o 1.

$$F_{i,j} = \frac{1}{A_i} \int \int_{A_i, A_j} \frac{\cos\theta_i \cos\theta_j}{\pi r_{i,j}^2} V_{i,j} dA_j dA_i \quad (2.14)$$

En general, no hay una solución analítica para esta integral. Sin embargo para ciertas formas y orientaciones se pueden encontrar fórmulas en los apéndices de libros de transferencia de calor radiado, y para formas y relaciones más complejas, se han desarrollado aproximaciones numéricas y analógicas.

Una aproximación analógica es la de Nusselt [Cohen92] que establece que, la integral interior de la ecuación anterior representa el factor de forma desde un área diferencial a un patch finito. Esta cantidad puede ser determinada rodeando el área con un hemisferio orientado respecto a la normal del área diferencial (fig.2.2). Entonces un patch finito puede ser proyectado radialmente dentro del hemisferio y después ortogonalmente sobre la base del hemisferio. La fracción del área de la base cubierta por su proyección será igual al factor de forma.

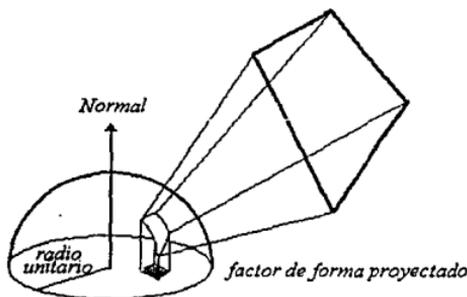


Figura 2.2. Hemisferio de iluminación empleado para la obtención de los factores de forma.

El funcionamiento de esta analogía es el siguiente. La proyección sobre el hemisferio considera los términos  $\cos\theta_j$  y  $1/r^2$ , así como el término  $V_{ij}$ . La proyección sobre la base considera el término  $\cos\theta_i$  y el término  $\pi$  del denominador es el área del círculo unitario.

Un método numérico para resolver la integral puede ser obtenido convirtiendo la integral doble de área en una integral doble de contorno, al aplicar el teorema de Stokes. Desafortunadamente, este método es muy costoso computacionalmente y no considera que en un ambiente complejo que contenga superficies ocultas, estas modifican el contorno de un patch.

### 2.4.1 Hemicubo.

Del hecho de que cualquier patch que cubra la misma área proyectada sobre el hemisferio tiene el mismo factor de forma, es decir el mismo ángulo sólido, podemos ver que cualquier superficie intermedia puede ser usada para proyectar los patches sin cambiar los valores de las proyecciones correspondientes. Esto nos lleva al uso del hemicubo, [Cohen85] en vez del hemisferio.

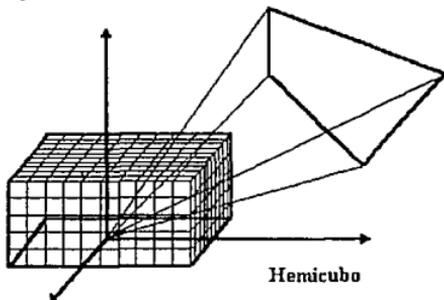


Figura 2.3. Hemicubo empleado para el cálculo de los factores de forma.

Un hemicubo (medio cubo) es colocado sobre una área diferencial, con los lados subdivididos en pequeñas celdas. Cada celda define una dirección particular y un ángulo sólido. Así un *factor de forma delta* puede ser asociado con cada una de estas celdas basándose en su tamaño y orientación. Aunque los factores de forma de las celdas son independientes del tamaño del hemicubo imaginario, es más fácil pensar en este como un hemicubo unitario al igual que se hacía con las semiesfera unitaria. Los factores de forma delta pueden ser precalculados y almacenados en una tabla para ser usados por todos los elementos.

Cada cara del hemicubo representa una vista de la proyección definida por una pirámide con ángulo de visión de  $90^\circ$ , pero con la peculiaridad de que las caras laterales sólo ocupan la mitad de arriba y la de abajo se mantiene oscurecida, este tipo de vista es muy usada en graficación, así que se puede tomar ventaja de toda la tecnología desarrollada para este tipo de proyecciones. Las superficies ocultas deben ser consideradas en la proyección de los patches sobre las cinco superficies del hemicubo, así que se usa la técnica de *Z-buffer* que es estándar en el cálculo de las superficies ocultas, con una pequeña modificación, en vez de guardar la intensidad de la superficie en cada celda, se guarda un identificador que refiere al patch proyectado en esa celda.

Una vez que la escena ha sido proyectada, el identificador contenido en cada celda representa cual patch es visible en esa dirección. Los factores de forma de cada patch son determinados simplemente sumando los factores de forma delta precalculados para cada una de las celdas conteniendo el identificador del patch.

El algoritmo del hemicubo, aunque muy eficiente, tiene dos limitantes principales. Primero, como sólo la integral interior de la ecuación de los factores de forma es

calculada, pueden ocurrir serias inexactitudes si el tamaño del patch es relativamente grande respecto a la distancia entre los patches. Esto puede ser resuelto repitiendo el algoritmo del hemicubo en puntos distintos y promediando. Otra posibilidad es subdividir la escena en patches más pequeños, pero esta solución fallará nuevamente cuando los patches estén muy cerca; segundo, debido a que el hemicubo divide el hemisferio dentro de pequeños ángulos sólidos discretos y espaciados regularmente, se pueden presentar problemas de "aliasing". Los patches pequeños pueden quedar entre los espacios de la discretización o ser muestreados sólo por una o dos celdas causando problemas de exactitud cuando los factores de forma son determinados. Estos problemas pueden ser aminorados incrementando la resolución del hemicubo, pero eventualmente otros métodos son necesarios conforme la complejidad de la escena se incrementa.

## 2.5 Otras Posibilidades para el Cálculo de los Factores de Forma.

Se han sugerido o implantado otros métodos para el cálculo de los factores de forma. Un método similar en muchos aspectos al del hemicubo fué descrito por Nishita y Nakamae en el cual convierten la integral de área en series de integrales de línea [Nishita85]. Baum sugiere un método híbrido que combine métodos tanto numéricos como analíticos [Baum89], otra opción es usar el ray tracing.

Muestrear la escena con ray tracing desde un área diferencial ha sido usado con gran éxito en la evaluación de la integral de los factores de forma, existiendo varios métodos, variando principalmente en como seleccionar las direcciones de muestreo.

Este tipo de métodos superan el aliasing inherente en el algoritmo del hemicubo, reemplazando el z-buffer con un método más flexible para determinar la visibilidad. Ray Casting proporciona este método [Sillion89]. Debido a que los rayos son seleccionados independientemente, estos pueden ser tomados en cualquier distribución de direcciones, proporcionando así una base de muestreo no uniforme, adaptativo y estocástico. Además, ray casting maneja una amplia variedad de tipos de superficie. La desventaja de ray casting como un método de determinación de visibilidad es que aún no hay implementaciones por hardware. El hemicubo, por otra parte puede tomar ventaja del z-buffer disponible por hardware en muchas estaciones de trabajo para acelerar el cálculo de los factores de forma.

### 2.5.1 Muestreo Adaptativo del Semiespacio

El cálculo de los factores de forma requiere de un muestreo de los ángulos sólidos visibles desde un elemento dado. Por lo tanto se debe analizar una mitad completa del espacio sobre el plano tangente en este punto, para poder tomar en cuenta todas las posibles transferencias de luz.

El muestreo adaptativo [Sillion89] es un método en donde a diferencia del hemicubo se proyecta una sola vez, en un plano paralelo a la tangente en el punto dado. Entonces se considera sólo un área restringida en este plano, ignorando así todas las partes de semiespacio que se proyectan fueran del área seleccionada. Esta aproximación es hecha posible por la dependencia angular de los factores de forma.

En el cálculo de los factores de forma se debe integrar una función que depende del coseno del ángulo entre la dirección de visión y la normal de la superficie. Destaca el hecho de que la contribución de las direcciones cercanas a la tangente es mucho más pequeña que la de una sola dirección casi perpendicular. De forma más precisa, la energía esparcida por el cono diferencial es

$$\Delta P = P_0 \sin 2\theta d\theta \quad (2.15)$$

donde  $P_0$  es la potencia total radiada.

Es posible estimar una cota superior de la parte de energía que es ignorada cuando analizamos la energía que es emitida difusamente a través de un área cuadrada de tamaño  $2D$ , centrada en un plano de proyección a una distancia  $H$  de la superficie emisora (con  $H$  mucho menor que  $D$ ). De esto tenemos que la pérdida de energía es menor que la energía radiada en las direcciones con

$$\theta \in \left[ \frac{\pi}{2} - \epsilon, \frac{\pi}{2} \right]$$

si  $\tan \epsilon = H/D$ , entonces la fracción de energía no considerada es tal que:

$$\frac{\Delta P}{P_0} < \epsilon \cdot \sin 2\epsilon \quad (2.16)$$

Como  $\epsilon$  es pequeño comparado con 1, se puede escribir que

$$\tan \epsilon = \epsilon \quad \text{y} \quad \sin 2\epsilon = 2\epsilon$$

así que tenemos:

$$\frac{\Delta P}{P_0} = 2\epsilon^2 = 2 \left( \frac{H}{D} \right)^2 \quad (2.17)$$

Si por ejemplo, se decide que el error de 1% es aceptable, se calcula el valor de la razón  $D/H$ :

$$\frac{D}{H} = \epsilon^{-1} = \sqrt{\frac{2P_0}{\Delta P}} = \sqrt{200} = 14 \quad (2.18)$$

La estimación anterior sirve para una distribución lambertiana de luz emitida.

Ahora hay que analizar la proyección de la escena en el plano que corresponde a la "pantalla" de proyección. Debido a la distorsión inducida por la proyección en perspectiva, sería poco razonable muestrear uniformemente el interior del cuadrado. Esto produciría un sobremuestreo de las zonas externas para obtener una resolución suficiente en el centro. Así que se dividirá la pantalla en *proxeles* (elementos de proyección) de tamaño variable, cada uno contribuyendo una cantidad parecida a los factores de forma [Sillion89].

El factor de forma asociado desde el origen con un área rectangular (fig.2.4), acotada por  $x_1, x_2, y_1, y_2$  está dado por

$$\Delta F_{x_1, x_2, y_1, y_2} = \int_{x_1}^{x_2} \int_{y_1}^{y_2} \frac{H^2}{(x^2 + y^2 + H^2)^2} dy dx \quad (2.19)$$

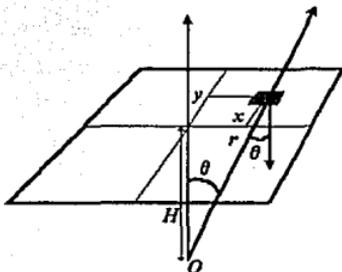


Figura 2.4. Geometría para el cálculo de los factores de forma sobre el plano de proyección.

Queremos encontrar una sucesión de enteros  $(x_i)_{i=0..N}$ , con  $N$  una resolución fija tal que

$$x_0 = 0, \quad x_N = D \quad \text{y} \quad \forall_{i,j} \Delta F_{x_i, x_{i+1}, x_j, x_{j+1}} \approx \frac{1}{N^2} \quad (2.20)$$

Prácticamente debido a la simetría radial del integrando, el requerimiento anterior es imposible de obtener. Pero se eligen los valores tales que todos los  $\Delta F$  a lo largo de los ejes tengan aproximadamente el mismo valor, estimando numéricamente la integral.

Se obtiene así una partición del plano en regiones rectangulares, por líneas paralelas a los ejes, cuyas posiciones  $(x_i)$  solo necesitan calcularse una vez, para una resolución dada. Estos son almacenados en un archivo y serán usados como coordenadas de los proxeles. Ahora para cada patch de las superficies en la escena, se quiere analizar la proyección del ambiente, en esta malla rectangular y asociar un patch a cada proxel.

Para poder emplear la coherencia espacial de la proyección, se usa un esquema de subdivisión adaptativa para analizar la imagen en la pantalla, tal como lo hizo Warnock [Warnock69].

El principio de este algoritmo es analizar la imagen proyectada en una región rectangular del plano. Si el contenido de la ventana es "suficientemente simple", es decir, si el problema de visibilidad está resuelto, el algoritmo se detiene. En cualquier otro caso, la ventana es subdividida y el proceso es aplicado a las subventanas.

Una vez que el contenido de la ventana ha sido identificado como el elemento correspondiente, se agrega a su factor de forma la contribución de la ventana. Las contribuciones de los distintos proxeles, son precalculadas y almacenadas en una tabla, ya que no todos son iguales, pero como una ventana podría (y de hecho debía de) contener muchos proxeles, y no se desea sumar todas las contribuciones de los proxeles dentro de la ventana, se adopta la siguiente convención: En el lugar de la

contribución del proxel  $(i,j)$  se guarda la suma de las contribuciones de los proxeles  $(p,q)$  tales que  $p \leq i$  y  $q \leq j$ . De esta forma, las contribuciones de cualquier ventana pueden ser estimadas en tiempo constante, con sólo una suma y dos restas: Si las integrales son almacenadas en un arreglo bidimensional  $T$ , tenemos:

$$\Delta F_{x,x,y,y} = T[j,l] - T[j,k] - T[i,l] + T[j,l] \quad (2.21)$$

## 2.5.2 Factores de Forma Utilizando "Ray Tracing"

A continuación se verán dos métodos de muestreo de la integral de los factores de forma utilizando *Ray Casting*. El primer método es útil cuando en un momento dado se desea determinar los factores de forma desde un área diferencial particular a todos los patches vistos desde esta, que es el caso manejado por el hemicubo. El segundo método es apropiado cuando se requiere el factor de forma de solo un patch.

### 2.5.2.1 Factores de Forma desde un Area Diferencial a todos los "Patches"

La ecuación para el factor de forma de un área diferencial  $i$  a un patch finito  $j$  es

$$F_{d_i, A_j} = \int_{A_j} \frac{\cos \theta_i \cos \theta_j}{\pi r_{i,j}^2} V_{i,j} dA, \quad (2.22)$$

donde esta integral es evaluada para todos los patches.

Para realizar la integración numérica, la función es evaluada para áreas discretas y los resultados son sumados. En el algoritmo del hemicubo la función es evaluada en las direcciones y para las delta-áreas determinadas por el hemicubo. Cada dirección muestreada representa una contribución al factor de forma del patch visible en esa dirección.

En un método de seguimiento de rayos, las direcciones pueden ser eligidas más flexiblemente que en el caso del hemicubo. La forma de elegir estas direcciones puede modificar la eficiencia final, por ejemplo la forma inmediata de elegir la distribución de direcciones es haciendo una distribución uniforme sobre el hemisferio del área diferencial. El inconveniente de esta distribución es que mientras cada una de las muestras requiere del mismo esfuerzo, las muestras no contribuyen al resultado de la misma proporción. Idealmente, el muestreo se debe concentrar en aquellas direcciones que hacen las mayores contribuciones al resultado.

Una mejor elección es distribuir las muestras de forma no uniforme sobre el hemisferio, con una densidad proporcional al valor de  $\cos \theta_j$ , haciendo así que cada muestra contribuya igualmente al resultado.

Malley [Malley88] presenta un método que genera la distribución de direcciones deseada. Este método puede ser explicado en base a la analogía de Nusselt [Cohen92]. Recordando que en esta interpretación, el patch es proyectado sobre el hemisferio encima del área diferencial y después la proyección es nuevamente

proyectada ortogonalmente sobre la base del hemisferio, la fracción del área del círculo cubierta por esta proyección es el factor de forma que corresponde al patch.

De esta forma, si se genera un conjunto de puntos esparcidos uniformemente sobre el círculo base del hemisferio (fig. 2.5) y se proyectan cada uno de estos puntos de forma ortogonal sobre el hemisferio, entonces se obtendrán las direcciones deseadas. Las direcciones de los rayos son dirigidas radialmente a través de estos puntos proyectados. Los rayos son lanzados para cada uno de los puntos muestra en el círculo.

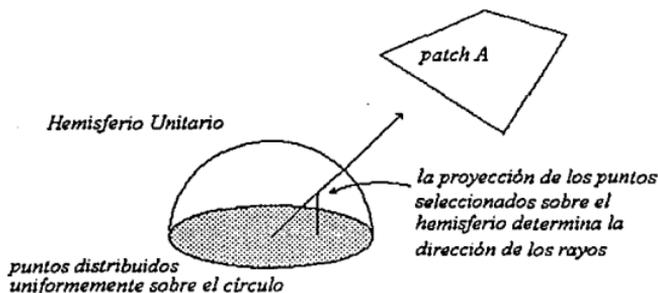


Figura 2.5. Esquema para la distribución de direcciones de los rayos

Al lanzar estos rayos se almacena el número de veces que cada patch en la escena es intersectado por un rayo, obteniéndose los factores de forma de dividir el número de veces que un patch es intersectado por el número total de rayos lanzados. Regresando otra vez a la analogía de Nusselt, el número total de rayos lanzados estima el área del círculo y la fracción de los rayos que tocan un patch son el área relativa de la proyección del patch sobre el círculo.

Como este método muestrea el hemisferio completo, entonces calcula los factores de forma de todos los patches desde el área diferencial, tal como lo hace el hemicubo. La principal ventaja sobre el hemicubo es que el uso de ray casting permite seleccionar las direcciones de muestreo. Una de las principales razones para utilizar ray casting, es que los rayos pueden ser trazados recursivamente en vez de detenerse en la primer superficie que tocan, permitiendo de esta manera incluir los efectos de las superficies especulares en los factores de forma entre patches difusos.

### 2.5.2.2 Factores de Forma desde un Punto a un "Patch"

En algunos casos no es necesario calcular todos los factores de forma, de un patch a los demás, a un mismo tiempo, sino que es suficiente conocer el factor entre un punto y un patch. Obviamente, si se calculan todos los factores de forma se obtendría el que se desea, pero esto significaría un desperdicio de tiempo de cómputo, porque los demás factores no serían tomados en cuenta. Idealmente sólo se deben lanzar los rayos que tocan al patch en cuestión.

Una forma alternativa a la integración numérica de la ecuación es subdividir directamente el patch  $j$ , y así aproximar la integral como la suma de los factores de forma calculados para regiones de área  $\Delta A_j$  más pequeñas.

$$F_{dA_1-A_1} = \sum_{i=1}^n \frac{\cos\theta_{1i} \cos\theta_{2i}}{\pi r_i^2} \Delta A_i \quad (2.23)$$

En esencia, el patch fuente es separado en pequeños pedazos que son tratados como área diferenciales, y de hecho cada uno de los términos de la sumatoria es el factor de forma de un área diferencial a otra área diferencial. Este cálculo es realizado lanzando un rayo desde el punto a cada una de las áreas diferenciales, y sumando las contribuciones de aquellas áreas que son visibles.

Es necesario que la suposición de áreas diferenciales sea válida, de otra forma el resultado obtenido no será correcto. Esto requiere que  $\Delta A$  sea mucho menor que  $r^2$ . Una forma de asegurar esto es subdividir el patch adaptativamente hasta que las áreas resultantes cumplan este criterio.

Una alternativa es tratar explícitamente cada área como un área finita. Por ejemplo, cada área diferencial puede pensarse como un disco, y sustituir la ecuación para el factor de forma de un punto a un disco en la sumatoria. Esta aproximación puede en muchos casos producir resultados aceptables, aún para subdivisiones limitadas y puede ser útil cuando se requiere una solución rápida.

El factor de forma de un área diferencial a un disco de área  $A_2$ , directamente opuesto es:

$$F_{dA_1-A_2} = \frac{A_2}{\pi r^2 + A_2} \quad (2.24)$$

El efecto de la orientación diferencial de la fuente y el área que recibe pueden ser aproximados incluyendo los cosenos de los ángulos entre la normal en cada superficie y la dirección entre las superficies.

$$F_{dA_1-A_2} = \frac{A_2 \cos\theta_1 \cos\theta_2}{\pi r^2 + A_2} \quad (2.25)$$

El factor de forma desde un área diferencial a un patch de área  $A_2$  subdividida en  $n$  áreas aproximadas por discos de área  $A_2/n$  está dado por:

$$F_{dA_1-A_2} = \frac{A_2}{n} \sum_{i=1}^n \frac{A_2 \cos\theta_{1i} \cos\theta_{2i}}{\pi r_i^2 + A_2} \quad (2.26)$$

Esta ecuación asume una subdivisión uniforme del patch. Sin embargo, la subdivisión uniforme no es adecuada en algunos casos. La aproximación de los discos fallará en el caso en que la distancia desde el disco al receptor es pequeña relativa al tamaño del patch. En este caso se requiere de una subdivisión adaptativa.

Una dificultad adicional con la subdivisión uniforme es que, como un rayo es lanzado para cada una de las subdivisiones, el área del patch será tratada como varias luces puntuales en las pruebas de visibilidad. Como resultado de esto, la frontera de la sombra puede verse como una serie de sombras traslapadas bien delimitadas, en vez de una suave penumbra.

Ambos problemas pueden ser corregidos subdividiendo adaptivamente el patch, esta subdivisión puede hacerse de forma recursiva hasta que la subdivisión alcance un cierto criterio.

## 2.6 Refinamiento Progresivo

El método estándar de radiosidad, que requiere la solución de un gran número de ecuaciones simultáneas, puede crear escenas simuladas de calidad fotorealista, pero el tiempo de cómputo es del orden de  $O(n^2)$ , donde  $n$  es el número de patches, lo cual es demasiado tiempo como para poder interactuar con la escena en tiempo real.

En 1988 se introdujo un nuevo método de refinamiento progresivo para la solución del cálculo de la radiosidad [Chen88]. El procedimiento implica distribuir la luz desde cada fuente, en vez de recolectar la luz de cada patch (fig.2.6). Además reduce ampliamente el tiempo de cálculo y los requerimientos de almacenamiento. Otra ventaja importante es que las imágenes son inmediatamente generadas y refinadas progresivamente.

En el algoritmo de radiosidad convencional, se usa el método iterativo de Gauss-Siedel para obtener la solución del sistema de ecuaciones simultáneas. Este método converge a la solución resolviendo el sistema de ecuaciones un renglón a la vez.

Una interpretación física de este proceso es útil. La evaluación del  $i$ -ésimo renglón de las ecuaciones proporciona una estimación de la radiosidad,  $B_i$ , del patch  $i$  basado en el estado actual de las radiosidades de los demás patches.



Figura 2.6. Esquemas de distribución de la radiosidad

$$\begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} = \begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} + \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{bmatrix} \begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} \quad \left[ \begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} = \begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} + \begin{bmatrix} x \\ x \\ x \\ x \\ x \end{bmatrix} \right]$$

para todo  $j$

$$B_i = E_i + \rho_i \sum_{j=1}^n B_j F_{ji}$$

$$B_j = B_j + B_j (\rho_j F_{ji})$$

$$\text{donde } F_{ji} = \frac{F_{ij} A_i}{A_j}$$

En cierto sentido la luz que sale de un patch  $i$  es determinada por la recolección de la luz del resto del ambiente. Un solo término de la sumatoria determina la contribución a la radiosidad del patch  $i$  desde patch  $j$ .

$$B_i \text{ debido a } B_j = \rho_i B_j F_{ji} \quad (2.27)$$

Si la solución pudiera ser desplegada durante el proceso de cálculo, al fin del cálculo de cada renglón, se sabría el efecto de todos los patches  $j$  sobre el patch  $i$ , y la radiosidad  $B_i$  podría ser actualizada para el patch  $i$ . Al presenciar el cambio dinámico de la escena, se vería a la imagen cambiando secuencialmente en intensidad, un patch a la vez en el mismo orden que los renglones de las ecuaciones.

Desde un punto de vista de refinamiento progresivo, sería más adecuado si la imagen completa fuera actualizada en vez de solo un patch a la vez. Esto es posible de realizar determinando la contribución hecha por el patch  $i$  a todos los demás. Esta es una solución iterativa por columnas para las ecuaciones de radiosidad.

La interpretación física de esta solución por columnas es que la distribución de la luz del  $i$ -ésimo patch en la escena, modifica las radiosidades de todos los patches en la escena.

La contribución de la radiosidad desde el patch que distribuye,  $i$ , a todos los demás patches que reciben,  $j$ , es:

$$B_j \text{ debido a } B_i = \rho_j B_i F_{ji} \quad (2.28)$$

La relación de reciprocidad descrita anteriormente proporciona la base para calcular eficientemente los factores de forma de las columnas.

Así, sustituyendo para  $F_{ji}$ , la ecuación se convierte en:

$$B_j \text{ debido a } B_i = \rho_j B_i F_{ij} \frac{A_i}{A_j} \quad (2.29)$$

y esto es válido para todos los patches.

Se debe notar que una ventaja importante de usar la solución por renglones es la eficiencia obtenida en el cálculo de los factores de forma. Usando el hemicubo, se

obtiene todo un renglón de factores de forma por cada posición de hemisferio calculada. Es decir, cuando el hemisferio es colocado sobre el patch  $i$ , los factores de forma  $F_{ij}$  son calculados para todos los patches  $j$ . Usando la relación de reciprocidad, se puede obtener la misma eficiencia de cálculo en los factores de forma por columna  $F_{ji}$  con sólo el cálculo adicional de multiplicar por la razón  $A_i/A_j$ .

Usando la solución por columnas ahora cada paso consiste de realizar una operación de hemisferio sobre un patch, determinando los factores de forma por renglón, transformarlos a factores por columna, y calcular las contribuciones de la radiosidad desde el patch a las radiosidades de los otros patches.

Si la solución es desplegada durante el proceso de cálculo, al final del cálculo de cada columna, las radiosidades son actualizadas para todos los patches, observándose que la imagen completa aumenta su intensidad cada vez que una nueva solución de columna es obtenida. La magnitud del incremento depende de la cantidad de energía  $B_i A_i$ , distribuida por el patch.

### 2.6.1 Acelerando la Convergencia

Además de la convergencia de este método, sería deseable que la aproximación a la solución final fuera alcanzada lo más rápido posible. Esto proporcionaría el beneficio de obtener imágenes de alta calidad a velocidades interactivas, cosa que aún no es posible con las estaciones de trabajo actuales, así que el objetivo es tratar de converger a la solución exacta de la iluminación global en la forma más rápida posible de tal manera que las imágenes sean generadas casi inmediatamente.

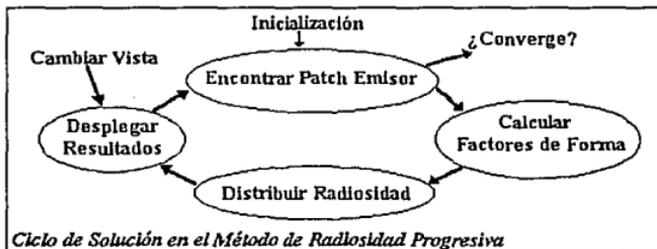


Figura 2.7. Algoritmo para el método de radiosidad progresiva

Teóricamente la solución no converge hasta que toda la energía reflejada o emitida es esparcida en el ambiente. Por esta razón la radiosidad final  $B_j$  del patch dado consiste de la suma de las contribuciones de todos los otros patches. El valor final de esta suma sería obtenido más rápidamente si las mayores contribuciones fueran agregadas al principio. Estas contribuciones tienden a venir de los patches que radian la mayor energía, es decir, aquellos que tienen los productos  $B_i A_i$  mayores. Así de forma intuitiva, los patches que radian más energía son los que tienen un mayor efecto en la iluminación del ambiente y deben ser procesados primero.

La implementación del algoritmo se sigue lanzando la energía desde el patch que tenga la mayor energía por lanzar (fig.2.7). La mayoría de las fuentes de luz son lanzadas automáticamente en un principio, iluminando bien gran parte del ambiente. El siguiente conjunto de patches procesados, de acuerdo a este criterio de selección del patch emisor, serían aquellos patches que reciben la mayoría de la luz de las fuentes, y así en adelante.

En el transcurso de esta solución iterativa el patch  $i$  puede ser repetido varias veces conforme la solución converge. Cada vez que la estimación de la radiosidad  $B_i$  del patch  $i$  será más exacta. Sin embargo, el ambiente ya contiene la contribución de la estimación previa de  $B_i$ , así que sólo la diferencia  $\Delta B_i$  entre las estimaciones previa y actual de la radiosidad debe ser considerada.  $\Delta B_i$  representa la radiosidad sin lanzar.

Este lanzamiento de energía desde un patch se continúa hasta que la solución converge dentro de una cierta tolerancia. Cada paso intermedio obtiene simultáneamente la solución para muchos patches, proporcionando resultados que pueden ser desplegados conforme la solución avanza.

### 2.6.2 Término Ambiental

Al usar el procedimiento descrito previamente, las imágenes intermedias irán desde un ambiente oscuro, continuamente aumentando su brillantez, a una escena totalmente iluminada incluyendo todas las interreflexiones difusas entre superficies. La iluminación de la escena durante los primeros pasos de la solución será inadecuada, particularmente en las regiones que no reciben iluminación directa, ya que la iluminación global aún no es representada exactamente.

En los primeros modelos de iluminación, el efecto de la iluminación global fue aproximado agregando un término ambiental arbitrario. Un uso similar se puede hacer aquí, pero su valor para un determinado momento de la solución estará basado en las estimaciones actuales de las radiosidades de todos los patches y la reflectividad del ambiente.

El término ambiental es agregado solo para efectos de despliegue y no debe ser tomado en cuenta en la solución. La contribución del término ambiental decrece conforme la solución avanza, proporcionando una imagen aceptable casi inmediatamente y que no obstruye el progreso de la solución.

Una primera aproximación razonable a los factores de forma se puede hacer sin ningún conocimiento previo acerca de la visibilidad o las relaciones geométricas entre patches. El factor de forma de cualquier patch  $i$  al patch  $j$ , puede ser aproximado como la fracción del área total del ambiente tomada por el patch  $j$ . Y como en los factores de forma correctos la suma total es igual a la unidad, así:

$$F_{i,j} = \frac{A_j}{\sum_{j=1}^n A_j} \quad (2.30)$$

La reflectividad promedio del ambiente puede ser calculada como el promedio del área pesada con las reflectividades de los parches:

$$\rho_{media} = \frac{\sum_{i=1}^n \rho_i A_i}{\sum_{i=1}^n A_i} \quad (2.31)$$

Para cualquier unidad de energía mandada en el ambiente, una parte de esta será reflejada en forma promedio, y algo de esta reflejada otra vez, etc. Así, un factor de interreflexión general  $R$  es simplemente la suma geométrica:

$$R = 1 + \rho_{media} + \rho_{media}^2 + \rho_{media}^3 + \dots = \frac{1}{1 - \rho_{media}} \quad (2.32)$$

De estas suposiciones se deriva el término ambiental de radiosidad. Este es simplemente el promedio de la radiosidad que aún no ha sido lanzada via los factores de forma por el factor de reflexión.

$$Ambiental = R \sum_{j=1}^n \Delta B_j F_{j,i} \quad (2.33)$$

Así en cualquier momento del cálculo, la estimación de la radiosidad de cada patch puede ser mejorada agregando la contribución de la radiosidad ambiental. Si  $B_i$  es la radiosidad del patch  $i$  debido a la radiosidad recibida desde otros patches, una estimación mejorada está dada por:

$$B'_i = B_i + Ambiental \quad (2.34)$$

Esta estimación de  $B'_i$  es usada solo para propósitos de despliegue donde la contribución ambiental no es agregada a  $\Delta B_i$  y por lo tanto no es lanzada durante la solución. Conforme la solución avanza el promedio de energía sin lanzar disminuye y así el término ambiental disminuye de igual forma. El valor de  $B_i$  y  $B'_i$  convergen y la imagen ambiental inicial se transforma a una estimación más exacta de la iluminación global dada por las ecuaciones de radiosidad.

## 2.7 Radiosidad Incremental

Los métodos de radiosidad tradicional calculan la iluminación de una escena independientemente de la posición del observador, no obstante, si cualquier parte de la geometría de la escena cambia, el método debe empezar desde el principio. En el método de radiosidad progresiva aún cuando se obtienen imágenes muy rápidamente, esta técnica requiere que la escena sea estática, por lo que no puede existir una interacción en el modelado de la escena, en caso de que la geometría o las características reflectivas o emisivas cambien, entonces el proceso debe ser reiniciado.

El método de radiosidad incremental [Chen90] es una extensión del algoritmo de radiosidad progresiva, el algoritmo (fig. 2.8) prosigue con la solución progresiva normal. Al principio de cada ciclo de solución, el algoritmo monitorea cualquier solicitud de cambio en la escena. Cuando un cambio es solicitado, la solución será interrumpida temporalmente para procesar la solicitud y después continuará en cuanto la solicitud sea procesada. Los cambios posibles incluyen el cambio en los atributos de las superficies, como reflectancia o emisión, y el cambio en la geometría de la escena.

Dentro de las extensiones al método progresivo se incluyen algoritmos para ajustar incrementalmente los valores de radiosidad después de que un atributo o la geometría cambian. También se introduce un método para almacenar los cambios de la geometría en una cola de tal forma que cambios múltiples puedan ser hechos en sucesión, sin esperar la convergencia del método incremental.

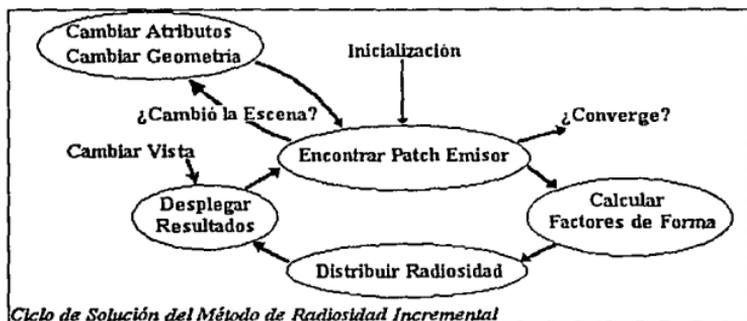


Figura 2.8. Algoritmo de Radiosidad Incremental.

### 2.7.1 Cambiando los Atributos de la Superficie

Cuando los atributos de un patch son cambiados, la *radiosidad incremental* del patch (la diferencia en radiosidades como resultado del cambio del atributo) es calculada primero. Entonces la radiosidad incremental es distribuida en la misma forma que como se distribuyó la radiosidad sin lanzar en el proceso de refinamiento progresivo. Los cambios son propagados a través de la escena y eventualmente resultarán en la solución correcta. Por ejemplo, apagar una fuente de luz resulta en el hecho de lanzar una iluminación negativa desde las fuentes de luz. La luz negativa se propagará a toda la escena y sustraerá toda la radiosidad que había emitido.

La radiosidad incremental es calculada restando la radiosidad anterior (antes del cambio del atributo) de la nueva radiosidad (después del cambio). La radiosidad anterior del patch está dada por:

$$B_i = E_i + \rho_i \sum_{j=1}^n B_j F_{ji} \quad (2.35)$$

Sean  $E_i'$  y  $\rho_i'$  la nueva emisión y reflectancia del patch  $i$ . La nueva radiosidad del patch  $i$  es entonces:

$$B_i' = E_i' + \rho_i' \sum_{j=1}^n B_j F_{ji} \quad (2.36)$$

La radiosidad incremental del patch  $i$ ,  $\Delta B_i$  es obtenida restando las ecuaciones anteriores:

$$\Delta B_i = E_i' - E_i + (\rho_i' - \rho_i) \sum_{j=1}^n B_j F_{ji} \quad (2.37)$$

y esto se convierte en:

$$\Delta B_i = E_i' - E_i + \frac{(\rho_i' - \rho_i)(B_i - E_i)}{\rho_i} \quad (2.38)$$

Hay que notar que el valor de  $\Delta B_i$  es indefinido si  $\rho_i$  es cero. Por lo tanto, la reflectancia de un patch no puede ser cero si ésta se va a cambiar a un valor distinto de cero. Pero esto no es ningún problema porque los materiales reales rara vez tienen reflectancia cero (es decir, la mayoría de las superficies negras reflejan aún una pequeña cantidad de luz).

El paso anterior puede ser fácilmente agregado al método original de refinamiento progresivo. Una vez que la radiosidad incremental del patch cambiante es calculada, esta es agregada a la radiosidad sin lanzar del patch. El método de solución normal puede continuar después de esto. La única modificación que se necesita hacer al método original es que realice la búsqueda del patch emisor basado en el valor absoluto de la radiosidad sin lanzar, para poder lanzar la energía negativa.

### 2.7.2 Cambiando la Geometría de la Escena

A diferencia de los cambios de atributos, los cambios en la geometría modificarán algunos factores de forma. Por lo tanto, todos los patches que hayan radiado energía en la escena necesitan relanzar sus radiosidades basadas en los cambios de los factores de forma. Cuando se hace un cambio en la geometría, se visita cada patch que ha radiado energía. Para cada uno de los patches, es necesario primero remover su contribución en el ambiente basándose en la geometría anterior, y entonces relanzar la radiosidad con la nueva geometría. Después de que todos los patches son procesados, la solución iterativa normal continua propagando las radiosidades y eventualmente llega al resultado correcto. Una vez que el ciclo normal es retomado, solo la nueva geometría es necesitada, ya que todos los patches que habían radiado energía han redistribuido sus radiosidades.

El proceso anterior se puede ilustrar con el siguiente ejemplo. Inicialmente, un plano es iluminado por una fuente de luz (fig.2.9). Cuando se agrega una esfera en la escena, la luz es apagada primero, la esfera es agregada y después la luz se prende nuevamente. El efecto total es que las radiosidades en el área de sombra son restadas del plano y agregadas a la esfera. Como resultado la energía se conserva. Debido a que la radiosidad del plano es modificada, este necesita relanzar su radiosidad también. El proceso completo es iterado hasta su convergencia.

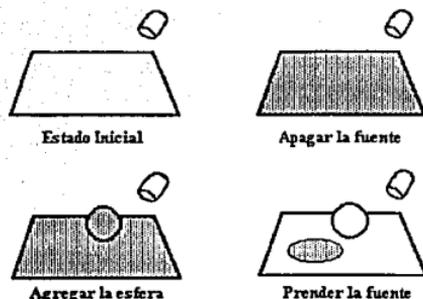


Figura 2.9. Ejemplo de cambios en la geometría de los objetos.

Para cada patch  $j$ , el primer paso es eliminar la contribución desde el patch  $i$ :

$$B_j' = B_j - \rho_j B_i^l F_{ji} \quad (2.39)$$

donde  $B_i^l$  es la radiosidad lanzada del patch  $i$  y es calculada de  $B_i$ , la radiosidad total, y de  $B_i^r$ , la radiosidad sin lanzar del patch  $i$ , como sigue:

$$B_i^l = B_i - B_i^r \quad (2.40)$$

El siguiente paso es relanzar la radiosidad total del patch  $i$  basada en el nuevo factor de forma,  $F_{ji}'$ , calculado de la nueva geometría:

$$B_j'' = B_j' + \rho_j B_i^l F_{ji}' \quad (2.41)$$

De las tres ecuaciones anteriores, se puede obtener la nueva radiosidad del patch  $j$

$$B_j'' = B_j + \rho_j (B_i^l \Delta F_{ji} + B_i^r F_{ji}') \quad (2.42)$$

donde

$$\Delta F_{ji} = F_{ji}' - F_{ji} \quad (2.43)$$

La radiosidad sin lanzar de este patch puede ser obtenida de la misma manera.

El término  $\Delta F_{ji}$  es llamado el factor de forma incremental, el cual es la diferencia en los factores de forma como resultado del cambio de geometría. Como los cambios de modelado afectan sólo áreas locales, los factores de forma incrementales serán en su mayoría cero con sólo algunos términos distintos de cero indicando los cambios. Así la

tarea del proceso incremental es identificar los factores de forma distintos de cero rápidamente y repropagar la radiosidad.

### 2.7.3 Factores de Forma Incrementales

Los factores de forma incrementales puede calcularse eficientemente si la coherencia del modelado es completamente explotada. Claramente, solo un patch que puede ver el cambio de la geometría, tendrá un factor de forma distinto de cero. Más aún solo las áreas cubiertas o descubiertas por el cambio de geometría necesitan evaluarse. Dependiendo del algoritmo para obtener los factores de forma, se puede hacer uso de la coherencia en distintas formas. Si se usa el hemicubo, solo es necesario realizar el cálculo en una parte del hemicubo, suficientemente grande para cubrir las áreas afectadas por el cambio. Si un algoritmo de tipo ray tracing es el empleado, los rayos solo serán lanzados desde áreas afectadas por el cambio. En muchos casos solo es necesario examinar una pequeña parte de la escena y la mayoría de los objetos en la escena pueden ser trivialmente eliminados.

De forma ingenua, los factores de forma incrementales pueden ser calculados encontrando los nuevos factores de forma y los anteriores separadamente. Sin embargo, sólo la parte dinámica de la escena necesita tener la evaluación de sus factores de forma separadamente. Cuando se usa el hemicubo, la parte estática de la escena sólo necesita ser proyectada una vez. El hemicubo es entonces duplicado y la parte dinámica de la escena es dibujada a cada hemicubo.

### 2.7.4 Colas de Cambios en la Geometría

Siempre que hay un cambio en la geometría, todos los patches que han radiado energía necesitan ser procesados de acuerdo a los pasos anteriores antes de que el proceso de solución normal pueda continuar. Además no pueden haber más cambios durante este período. La siguiente figura muestra un ejemplo de como la solución puede fallar si se viola la suposición mencionada. En la figura 2.10.a, una esfera es agregada a una escena que consiste de un plano iluminado por dos fuentes. Una de las fuentes relanza su radiosidad y crea correctamente una sombra de la esfera sobre el plano. En la figura 2.10.b, la esfera es borrada, antes de que la segunda fuente haya podido relanzar su radiosidad. Después de la operación de borrado, las dos fuentes de luz relanzan sus radiosidades otra vez para remover las sombras. Como la segunda fuente de luz nunca creó una sombra en el plano en el primer cambio, esto resulta en una sobre emisión en la supuesta sombra del plano.

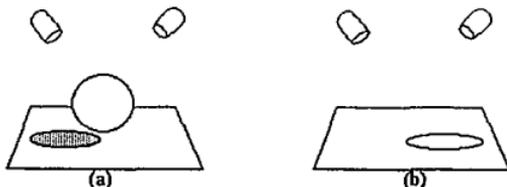


Figura 2.10 Posible caso de error cuando existe un cambio en la geometría.

Una forma de evitar la restricción anterior es mantener para cada patch emisor, una lista de todos los patches a los que éste ha distribuido radiosidad. Antes de cada emisión, el patch checa si el objeto que se desea borrar está en su lista. En el ejemplo anterior, la segunda fuente sólo tiene el plano en su lista y por lo tanto no intentará remover la sombra creada por la esfera. Sin embargo los requerimientos de almacenamiento de tener tal lista para cada patch es de  $O(n^2)$ , siendo  $n$  el número de patches.

Hay una mejor manera de resolver este problema sin imponer los requerimientos de memoria. Una solución es acumular todos los cambios de tal forma que puedan ser procesados en un momento conveniente (es decir, cuando el usuario no está haciendo más cambios). Los cambios geométricos pueden ser almacenados en una cola de geometría. La cola contendrá la historia de todos los cambios geométricos, y cuando son ejecutados en orden describen la actualización de la escena. En este algoritmo sólo existen dos operaciones que pueden agregarse a la cola. Una es "agregar patch  $i$ " y la otra "borrar patch  $i$ ". Estas dos operaciones son capaces de describir cualquier cambio geométrico. Por ejemplo mover un patch es descrito al borrar el patch de la posición anterior y agregarlo en la nueva posición.

Siempre que existe una solicitud de cambio geométrico, el cambio es agregado a la cola y entonces el ciclo de solución normal puede continuar. Cada patch tiene un contador de cola el cual indica el número de cambios geométricos en la cola que han sido considerados por el patch. Por ejemplo, un contador de cola de  $m$  significa que los primeros  $m$  patches han sido procesados por el patch. Los contadores de cola son inicializados a cero y son incrementados cuando son procesados los cambios en la cola. Cuando un patch es seleccionado como el patch emisor, este puede determinar cuales son los patches que ha visto antes (es decir, la escena anterior) a partir de su contador de cola y la cola, además también puede determinar la escena actual del contenido completo de la cola. Un conjunto de factores de forma incrementales desde el patch a los otros patches puede ser calculado a partir de las dos escenas, como ya se ha mencionado.

En la figura 2.11 se ilustra el proceso. Nótese que cuando un patch es borrado, este aún existe en la solución en proceso pero es simplemente invisible. El patch solo puede quitarse de la solución hasta que haya sido procesado por todos los patches. Entonces la cola puede ser reducida eliminando todas las entradas relacionadas con el patch.

Inicialmente, se agregan (+) a la escena dos luces ( $L_1, L_2$ ), un plano ( $P$ ) y una esfera ( $E$ ).  $L_1$  distribuye su radiosidad a todos los patches y su contador es 4, lo que significa que los cuatro cambios en la cola han sido realizados.  $L_2$  distribuyó su radiosidad antes de que  $E$  fuera agregada, por lo tanto su contador es 3. En (2.11.b),  $E$  es borrada (-) de la escena. En (2.11.c),  $L_1$  redistribuye su radiosidad y su contador es actualizado a 5. En (2.11.d),  $L_2$  es el patch emisor, como la cola indica que  $E$  ha sido borrada, el proceso simplemente actualiza el contador a 5 y se tiene el resultado correcto.

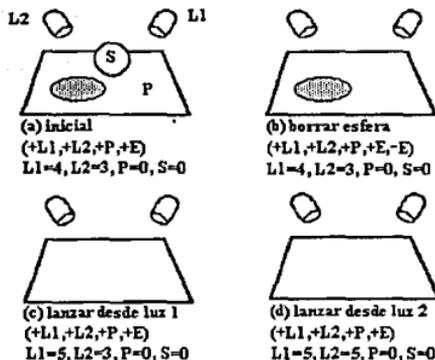


Figura 2.11. Ejemplo del uso de las colas de cambios en la geometría.

Otra ventaja de almacenar los cambios de la geometría en la cola es que los cambios hechos en sucesión no implican mucho cálculo extra, ya que el algoritmo solo se fija en los cambios efectivos. En el ejemplo anterior la esfera puede ser movida muchas veces y sólo importarán la primer y última posición.

# CAPITULO III

## ESPECULARIDAD Y TECNICAS HIBRIDAS

### 3.1 Introducción

El método de radiosidad que se ha estado discutiendo, solo es válido para reflectores difusos (lambertianos) ideales, ya que fue derivado de un problema muy simplificado de transporte de energía. Desafortunadamente una escena difusa se ve muy apagada ante los ojos, sobre todo considerando que el sistema visual humano es un buen detector e intérprete de los altos brillos [Hal89]. En particular los brillos son usados en la detección de curvaturas en las superficies, y cualquier intento de producir imágenes de tipo real debe considerar las reflexiones no difusas.

Cuando la suposición de reflexión difusa ideal es eliminada, entonces se debe usar la forma general de la ecuación de transferencia. Esto requiere del uso de una descripción bidireccional completa de la reflectancia. En el caso general, el equilibrio de energía para un conjunto de superficies radiantes es expresado, como se vió en el la sección 1.10, con una ecuación integral de la forma:

$$L(x, x') = V(x, x') \left[ L_r(x, x') + \int_{\Omega} \rho_{br}(x, x', x'') L(x', x'') dx'' \right] \quad (3.1)$$

- $L(x, x')$  es la radiancia saliendo del punto  $x$  en la dirección del punto  $x'$ , anteriormente denotada por  $L(x \rightarrow x')$ .
- $V(x, x')$  es una función de visibilidad entre  $x$  y  $x'$
- $\Omega$  es la union de todas las superficies que componen la escena.
- $\rho_{br}(x, x', x'')$  es la función de distribución de reflectancia bidireccional (FDRB).

Como se mencionó antes la FDRB es definida como la razón entre la radiancia reflejada en una dirección y la densidad de energía que viene de otra dirección. Esta es por lo tanto, una función de cuatro variables angulares más la longitud de onda y puede tomar formas muy complicadas.

En el caso difuso ideal que se consideró, la suposición de que la emisión y la reflexión son lambertianas, resultó en una constante  $\rho_{hd}$  y por lo tanto en una radiancia uniforme sobre el hemisferio. Si se consideran reflectores no difusos, entonces  $\rho_{hd}$  debe tener una forma más compleja, y la presencia de esta función dentro de la integral hace el problema de dispersión mucho más difícil de resolver.

### 3.2 Métodos de Radiosidad Direccionales Independientes del Observador

Reemplazar la ecuación de transferencia continua por un conjunto de ecuaciones lineales, es la idea básica bajo el desarrollo del método de radiosidad. Esto mismo se puede extender a el caso no difuso, considerando un conjunto adecuado de variables en vez del valor constante que se tenía para la radiosidad de un patch. El principal beneficio de esta modificación es que se puede calcular una solución independiente del observador que incorpora efectos direccionales (especulares).

#### 3.2.1 Radiosidad Direccional

Immel [Immel86] propuso un algoritmo donde el conjunto de variables es obtenido considerando una cantidad fija de direcciones (relacionadas con ángulos sólidos) desde cada superficie, y considerando todas los valores de radiosidades resultantes como las incógnitas del problema. Por lo tanto una variable típica es *la radiancia del patch  $i$  en la dirección  $p$* .

Aunque este método incrementa el número de variables en un factor bastante grande comparado con la formulación tradicional, se puede hacer la observación de que la información de visibilidad de patch a patch que debe ser calculada para establecer las relaciones entre superficies, también se puede usar para obtener información acerca de la irradiancia direccional de cualquier patch.

La información de visibilidad es obtenida primero, proyectando todas las superficies sobre un "cubo global" rodeando el patch dado (similar al hemicubo pero con una orientación fija), y almacenada en buffers de elementos (uno para cada cara del cubo). Las relaciones entre pares de variables son establecidas igualando cada dirección de salida desde un patch a la dirección de llegada de otro patch. Esto es posible debido a que todos los patches comparten la misma orientación. Una solución iterativa del sistema de ecuaciones resultante produce un conjunto completo de radiancias en todas las direcciones del cubo desde cada patch. Para calcular una imagen, la radiancia de una superficie vista a través de un pixel dado es interpolada del conjunto discreto de radiancias en la solución.

El principal inconveniente de esta técnica es por supuesto su demanda exagerada de recursos, tanto de tiempo de cálculo como de requerimientos de almacenamiento para los cubos globales. Además, el uso de un conjunto de radiancias discretas independientes del observador, puede resultar en problemas graves de aliasing, dado

que los efectos especulares típicamente producen cambios rápidos de pixel a pixel en una imagen, lo que no se puede solucionar con la interpolación de valores calculados. El usar patches más pequeños, sólo incrementaría la ya de por sí exagerada carga computacional.

### 3.2.2 Refinamiento Iterativo de Radiosidades Direccionales

Shao *et al.* [Shao88] introdujo una técnica de refinamiento iterativo en la cual los resultados del cálculo de la radiosidad difusa son usados como punto de inicio para una serie de aproximaciones continuamente mejoradas. La idea es similar a la de Immel, en que los buffers de elementos de los hemisferios son almacenados y reusados después para extraer información adicional. Como la orientación del hemisferio depende del patch seleccionado, es difícil establecer relaciones directas entre direcciones discretas. En vez de esto, la información relevante es expresada como la proporción de la energía total saliendo de un patch particular y alcanzando otro patch. Claramente, en el caso de reflectores no difusos, esta cantidad depende de la distribución direccional de la energía recibida. Esta información, junto con los valores de radiosidad para todos los patches dan valores de energía absoluta.

Una variable típica, en este esquema es la proporción del flujo de energía saliendo del patch  $i$  que llega al patch  $j$  (factores de forma). Estos son usados además de las variables de radiosidad normal.

En cada paso de la iteración, los factores de forma asociados a un patch no difuso son recalculados con la información del hemisferio y el conjunto actual de radiosidades. Estos son entonces usados para formar un sistema de ecuaciones completo que es resuelto igual que en el caso difuso, para dar un nuevo y mejorado conjunto de radiosidades.

Del hecho de que las variables del problema no pueden ser directamente expresadas como radiancias para las imágenes finales, y de los problemas de exactitud mencionados en el método anterior, Shao usa un segundo paso, que es dependiente del observador, para hacer el "render" de la imagen. Shao usa un ray tracer, en el cual usa como sombreado difuso la radiosidad calculada previamente.

### 3.3 Métodos Dependientes del Observador.

En vez de intentar extender el paradigma de radiosidad independiente del observador, para manejar reflexiones especulares, se pueden extender los algoritmos dependientes del observador tradicionales para hacer frente a los reflectores difusos.

#### 3.3.1 Técnicas de "Ray Tracing" y Monte Carlo

El ray tracing distribuido, introducido por Cook [Cook84], fué un primer paso en esta dirección: en vez de sólo considerar de forma precisa las leyes de óptica geométrica, las direcciones de los rayos son distribuidas de acuerdo a la FDRB de la superficie. Esto funciona muy bien para superficies altamente direccionales porque los rayos son distribuidos en la práctica sobre un ángulo sólido restringido y un pequeño número de rayos es suficiente.

Kajiya estableció el problema con una ecuación integral (la ecuación de rendering), y usó la técnica de "Monte Carlo Path Tracing" [Kajiya86]. Donde los árboles de rayos tradicionales son reemplazados por trayectorias simples en el ambiente, usando varias distribuciones de probabilidad en cada reflexión para decidir cual dirección seguir. Esta técnica fue exitosamente aplicada y las imágenes resultantes muy convincentes.

Aunque Kajiya [Kajiya86] discute varias técnicas de reducción de varianza, el número de trayectorias necesitadas para reducir el ruido estocástico es muy alto, debido a que no es posible predecir cuales son las direcciones importantes, de llegada de luz a la superficie. Las direcciones de las fuentes de luz por supuesto son buenos candidatos, y son exploradas sistemáticamente en la implementación de Kajiya, pero la iluminación indirecta sigue siendo muy difícil de estimar.

Ward [Ward88] observó que la componente difusa de la luz reflejada por un objeto varía muy lentamente sobre la superficie, y que calcular una iluminación difusa completa en cada pixel en la imagen es una pérdida de recursos. El pudo reducir dramáticamente el tiempo de cálculo tomando los valores de irradiancia difusa e interpolándolos de las muestras cercanas cuando era posible.

### 3.4 Técnicas Híbridas

En resumen, las técnicas dependientes del observador tienen una clara ventaja sobre las independientes del observador para la iluminación especcular, debido a que las distribuciones de las radiancias son siempre calculadas de acuerdo a la precisión necesaria. Por otra parte, el método dependiente del observador es muy mala elección para reflectores difusos, donde el número de rayos necesarios para obtener buenos resultados se vuelve prohibitivo.

Por lo tanto, la opción inmediata es tomar lo mejor de ambos esquemas, dejando a los esquemas basados en radiosidad manejar la parte difusa y a los esquemas de ray tracing la parte especcular.

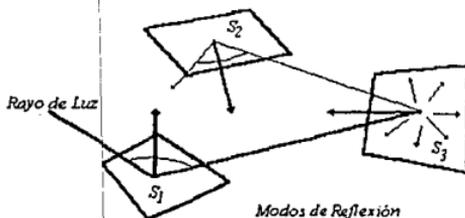


Figura 3.1. Modos de reflexión, en este ejemplo la primera reflexión es especcular, seguida de una reflexión difusa y finalmente una reflexión especcular.

#### 3.4.1 Método de Dos Pasos

Este tipo de técnica fue introducida por Wallace [Wallace87] en un caso específico, la técnica de utilizar dos pasos está basada en la distinción de los modos de reflexión. La esencia de esta técnica es tener un programa de radiosidad que calcule la parte

difusa de la luz y un programa de ray tracing que calcule la parte especular. Desafortunadamente, no se pueden separar completamente el cálculo de las componentes especular y difusa de la luz, porque la luz en si misma no es ni especular ni difusa, estas son cualidades que de hecho solo se aplican a los modos de reflexión de la luz (fig.3.1). En otras palabras, cierta cantidad de luz puede ser reflejada especularmente por una superficie  $S_1$ , y después reflejada difusamente por una superficie  $S_2$ , etc.

### 3.4.2 Método de Dos Pasos Extendido

Este método de dos pasos, permite simular todo tipo de reflexiones, elimina la restricción de que las superficies especulares deben ser espejos planos e incluye refracción entre los efectos de iluminación modelados.

En este método la reflexión es separada en dos modos:

- una reflexión difusa: donde una parte de la luz incidente es reemitida de acuerdo a la ley de Lambert.
- una reflexión (y refracción) especular (direccional): donde otra parte de la luz es reemitida alrededor de las direcciones asociadas con la dirección incidente por la ley de Snell.

En otras palabras, la función de reflectancia es expresada como una suma:

$$\rho(x, x', x'') = \rho_d(x') + \rho_s(x, x', x'') \quad (3.2)$$

$\rho_d$  es el coeficiente de la reflexión difusa en el punto  $x'$ , y  $\rho_s$  es la función de reflexión especular, que depende de la posición del punto  $x$  y  $x'$ . Esta función de reflexión especular vista como función de  $x''$  tiene un pico alrededor de la imagen reflejada del punto  $x$  sobre la superficie en el punto  $x'$  ( y otro pico alrededor de la dirección de refracción).

En este planteamiento se asume que las fuentes de luz son puramente difusas, es decir:

$$L_r(x, x') = L_r(x) \quad (3.3)$$

Bajo esta suposición, se puede reemplazar  $\rho$  por su expresión completa, reescribiendo la ecuación de rendering de Kajiya con el operador de reflexión global:

$$L(x, x') = V(x, x')\beta(x') + \tau L(x, x') \quad (3.4)$$

donde  $\beta$  solo depende de  $x'$ :

$$\beta(x') = L_r(x') + \rho_d(x') \int_{\Omega} L(x', x'') dx'' \quad (3.5)$$

y el operador  $\tau$  de reflexión-refracción especular, es tal que:

$$\tau L(x, x') = V(x, x') \int_{\Omega} \rho_r(x, x', x'') L(x', x'') dx'' \quad (3.6)$$

$\tau$  es un operador lineal, así que despejando la distribución de la luz  $L$  de la distribución obtenida de seguir una reflexión y refracción en todas las superficies de la escena ( $\tau$ ). La nueva ecuación establece las relaciones entre las distribuciones direccional ( $L$ ) e isotrópica ( $\beta$ ), esta puede ser invertida para dar:

$$L = [1 - \tau]^{-1} V \beta = \left[ \sum_{k=0}^{\infty} \tau^k \right] V \beta = S \cdot V \beta \quad (3.7)$$

$$\text{donde } S = \sum_{k=0}^{\infty} \tau^k$$

$S$ , el operador especular global representa el efecto de todas las posibles reflexiones especulares en la distribución  $L$ . Podemos asegurar que la sumatoria infinita converge porque los valores propios de  $\tau$  tienen módulo estrictamente menor que 1. Esto es claro, pues el equilibrio de energía dentro de un ambiente cerrado, establece que la intensidad reflejada es menor que la intensidad incidente, donde la diferencia es absorbida por los distintos materiales.

La distribución  $\beta$  puede ser expresada en la misma forma como:

$$\beta = L_r + DL$$

donde  $D$  es el operador de reflexión difusa, definido por:

$$DL(x') = \rho_r(x') \int_{\Omega} L(x', x'') dx''$$

El operador  $D$  representa el efecto de una reflexión difusa simple (en todas las superficies) en la distribución  $L$ . Este es el operador usado en el método convencional de radiosidad.

Finalmente, reemplazando  $L$  en la ecuación de la distribución isotrópica:

$$\beta = L_r + D \cdot S \cdot V \beta$$

#### 3.4.2.1 Utilización de la Ecuación de Rendering

Ahora se verá como producir un algoritmo de cálculo a partir de las formulaciones anteriores de la ecuación de rendering. Hasta ahora, sólo se ha tratado con ecuaciones integrales. Dividiendo el ambiente en patches de tamaño finito podemos convertir estas integrales en sumatorias sobre los patches. Asumamos por ahora un comportamiento puramente especular, lo que significa que la función  $\rho_r$  es cero excepto en las direcciones exactas de reflexión y refracción. Esta suposición no es

necesaria y sólo es introducida para mayor claridad, ya que permite el uso de ray tracing convencional.

El algoritmo de dos pasos primero estimará la distribución isotrópica  $\beta$ , y después derivará la distribución completa  $L$ , para las direcciones que el observador ve. Es importante notar que ninguno de estos pasos puede ser omitido. El resultado del primer paso es una distribución de la luz donde cada patch actúa como un reflector difuso, aún si la cantidad de energía emitida depende de las interacciones especulares con el ambiente. Debe enfatizarse el hecho de que la discretización de los objetos en patches sólo es necesaria para el cálculo de tipo radiosidad y no para los de tipo ray tracing. Así es posible usar una representación de los objetos más simple y compacta en la parte del proceso que usa ray tracing.

#### 3.4.2.2 Primer Paso: Iluminación Difusa

La ecuación de la distribución isotrópica  $\beta$  da una forma de calcularla. De hecho, esta es una ecuación de radiosidad, como la discutida en el método clásico, con el operador de reflexión difusa siendo reemplazado por  $D'S$ . El método usual de radiosidad resuelve esta ecuación calculando los factores de forma geométricos, que representan la relación entre todos los patches de la escena. Estos factores de forma son usados para construir la matriz del operador de reflexión difusa, entonces esta matriz es invertida numéricamente. Esto sugiere que un método de radiosidad puede proporcionar la distribución  $\beta$ , modificando solamente el cálculo de los factores de forma. De forma más precisa, la noción de factor de forma se extenderá para incluir efectos especulares.

Los factores de forma extendidos tienen un significado ligeramente diferente comparado con los usuales.

*$F_{ij}$  es la porción de energía saliendo de un elemento  $i$  y alcanzando el elemento  $j$ , después de cualquier número de reflexiones o refracciones especulares.*

Wallace en su método [Wallace87], también usa factores de forma extendidos, pero estos sólo permiten una reflexión en espejos planos. Los factores de forma extendidos que se manejan en este modelo son más generales porque permiten un número arbitrario de reflexiones especulares, con patches de cualquier geometría. Los cálculos de estos factores de forma extendidos pueden ser derivados de un análisis detallado de la ecuación. Se desea modelar la acción del operador  $D'S$ , que es equivalente a determinar de donde viene la luz recibida en algún punto, después de haber sido aplicado este operador. El operador  $D$  significa que se deben considerar todas las superficies visibles desde aquel punto, como en un cálculo clásico de factores de forma; y el operador  $S$  que para cada uno de estos elementos, se tiene que revisar un árbol de rayos reflejados y refractados (fig. 3.2).

Estos factores de forma extendidos pueden ser usados en un proceso de radiosidad clásico, es decir: se crea la matriz relacionando las distribuciones  $L_e$  y  $\beta$ , multiplicando los factores de forma por los valores de las reflectancias difusas para cada longitud de onda, se resuelve la matriz con algún método iterativo. La matriz construida e invertida

corresponde al operador  $[I-DSV]$ , ya que  $\beta = [I-DSV]^{-1}L_c$ . Es importante notar que los coeficientes de reflexión difusa pueden ser cambiados fácilmente, como en el proceso clásico, ya que no tienen ningún efecto en los factores de forma extendidos. Por el contrario, los coeficientes especulares son usados en el cálculo de los factores de forma extendidos, por lo tanto no pueden ser cambiados sin recalcular estos factores.

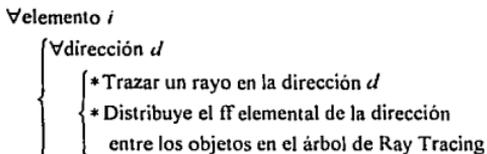


Figura 3.2. Algoritmo para el cálculo del primer paso.

### 3.4.2.3 Segundo Paso: Distribución Direccional

La distribución direccional de la luz  $L$ , debe ser calculada para todas las direcciones del espacio convergiendo en el punto de observación. Se calcula esta distribución utilizando la ecuación (3.7) anteriormente derivada. La distribución  $\beta$  ha sido calculada por el proceso de radiosidad extendido, así que sólo se tienen que evaluar los efectos de aplicar el operador  $S$  a esta distribución. Recuérdese que  $S$  representa los efectos de cualquier cantidad de reflexiones ( o refracciones). Para calcular  $L$  para todas las direcciones que llegan al observador, es suficiente usar un ray tracer clásico desde la posición del observador, con las siguientes modificaciones:

- No son necesarios los rayos de sombra, cosa que lo hace más rápido que el ray tracing convencional. También el tiempo de cálculo no es dependiente del número de fuentes de luz.
- El "modelo de iluminación" es trivial. Simplemente es el valor de la distribución  $\beta$  calculado para el punto de intersección con el rayo.

### 3.4.2.4 Caso General de Reflexión Especular

El operador de reflexión especular global se modeló como un operador de ray tracing, debido a la suposición del comportamiento especular puro impuesta anteriormente.

Es posible usar una función de reflectancia más complicada siempre que haya una forma de calcular el efecto del operador  $S$ . Un algoritmo de ray tracing distribuido podría ser útil para este propósito, tanto en el cálculo de los factores de forma extendidos como en el segundo paso.

El método de dos pasos proporciona un buen equilibrio entre calidad de la imagen y costo de CPU, al usar las respectivas ventajas de ray tracing y radiosidad, para calcular las componentes de la luz. Aunque este método involucra el cálculo de los factores de forma extendidos, que requieren de más tiempo de CPU que los factores de forma difusos, este tiene varias ventajas respecto al método de Shao:

- Es independiente de la distribución de la luz. Si los factores de forma extendidos son almacenados en un archivo, una nueva imagen puede ser generada con nuevas condiciones de iluminación, con solo el costo de la solución de la matriz de radiosidad y el segundo paso.

- No hay necesidad de almacenar los hemicubos para los patches especulares.

- Los patches especulares no tienen que ser finamente subdivididos. Un problema que aparece cuando se estudia la reflexión de la luz en una superficies especular es que, las direcciones de interés varían rápidamente en la superficie. Por lo tanto, todos los patches especulares deben ser finamente subdivididos para que el método de Shao produzca resultados adecuados.

Por otra parte, el "método de la imagen" de Wallace y la técnica de ray tracing mostrada, permiten tratar a los patches especulares como un todo, sin subestructuración posterior. Esta deficiencia de precisión para la reflexión especular es precisamente la razón por la cual Shao usa un segundo paso de ray tracing.

#### 3.4.2.5 Generación de los Factores de Forma Extendidos

Utilizando el algoritmo de muestreo adaptativo del espacio, para el cálculo de los factores de forma, este es capaz de asociar un factor de forma elemental a cualquier región rectangular del plano de proyección, correspondiendo a la emisión difusa hacia esta ventana. Mientras que para la radiosidad tradicional, las regiones son simplemente regiones de visibilidad de patches, en los factores de forma extendidos esto no es cierto, ya que se desea seguir la luz a través de reflexiones especulares. Así, cuando el algoritmo de Warnock detecta una ventana y su patch asociado, se lanzan rayos desde las esquinas de la ventana y se comparan los árboles de rayos generados. Se prueba un criterio de subdivisión, el cual básicamente establece que los primeros niveles deben ser iguales. Si este no es el caso, la ventana es subdividida y son lanzados nuevos rayos. Por lo que el algoritmo planteado en el capítulo anterior solo es modificado por un post-proceso que se debe realizar por cada ventana del algoritmo de Warnock. Una vez que la ventana es suficientemente subdividida, de tal forma que los árboles de ray tracing en las esquinas de la ventana se igualan, el factor de forma elemental asociado con una ventana es distribuido entre los objetos en el árbol.

Con este enfoque una escena puede ser generada en rangos interactivos de tiempo sólo hasta el primer paso. Pero al tratar de usar una solución completa, es decir, al incorporar los efectos del operador S, las limitaciones de la velocidad en el procesamiento se hacen evidentes.

## CAPITULO IV

# CARACTERISTICAS DEL SISTEMA

### 4.1 Introducción

Hasta este momento se han mostrado los fundamentos del método de radiosidad junto con algunas modificaciones que permiten un mejor planteamiento práctico, ahora se dará una descripción del sistema de radiosidad desarrollado con sus características más importantes y algunos detalles de su implantación.

El sistema está dividido en dos programas: *Genilum* (Generador de Iluminación) que a partir de una descripción geométrica y las propiedades reflectivas de una escena, calcula las radiosidades de los patches; y el segundo programa es *Visual* (Visualización de un Recorrido Virtual) que permite recorrer la escena de forma interactiva utilizando las radiosidades obtenidas por *Genilum*.

### 4.2 Plataforma de Desarrollo .

El sistema se desarrolló en las estaciones de trabajo Silicon Graphics, porque además de su gran velocidad de cómputo cuentan con un procesador gráfico, que permite el uso de rutinas gráficas básicas por medio de hardware, que son usadas en ambos programas. Aún cuando el programa *Genilum* usa estas rutinas, se puede prescindir de ellas implantando por software las rutinas necesarias, de tal forma que este programa puede ser ejecutado en cualquier computadora (hasta en una PC), con sólo el costo extra del tiempo de procesamiento. Por otra parte el programa *Visual*, si está íntimamente ligado a las rutinas de hardware, para poder proporcionar la interactividad de la animación al recorrer la escena, y por lo tanto no puede cambiar de plataforma a menos que se brinden servicios similares en otra plataforma.

Otra opción factible y muy plausible es utilizar una supercomputadora como la Cray para el cálculo de las radiosidades y una Silicon Graphics para el despliegue del recorrido interactivo.

### 4.3 Modelo de Datos

Como ya se ha mencionado el programa Genilum recibe como datos de entrada la descripción geométrica de la escena y sus propiedades reflectivas y emisivas, además de las características con las que va a ejecutarse el método de radiosidad.

La geometría de la escena es establecida en base a cuadriláteros los cuales se definen en base a los puntos en el espacio que forman sus vértices. Estos datos se separan en dos listas: una de puntos, en la cual están los vértices de todos los cuadriláteros; y una lista de cuadriláteros, donde la descripción de cada cuadrilátero contiene una lista ordenada de los índices, dentro de la lista de puntos, de los vértices que lo componen, además también contienen la información de reflectancia y emisión del cuadrilátero que se maneja en un modelo de color RGB.

El método de radiosidad que se implantó está basado en el refinamiento progresivo de Eric Chen, utilizando el algoritmo del hemicubo para el cálculo de los factores de forma, entre los datos que se necesitan para aplicar este método de radiosidad, están el umbral de convergencia necesario para detener el proceso iterativo de solución, la resolución que se usará en el hemicubo, un factor de escala de intensidad a ser aplicado sobre toda la escena, un indicador de si se desea agregar el término ambiental en el despliegue de la solución durante el proceso, el tamaño de la ventana de despliegue de la solución y el campo de visión del observador.

Estos datos son almacenados en las estructuras de datos que se definen en el programa para su posterior utilización en el proceso de solución.

Antes de describir estas estructuras de datos hay que considerar un problema respecto a la representación de la geometría de la escena.

#### 4.3.1 Subestructuración de Patches y Elementos

Un problema con el algoritmo de radiosidad es que la iluminación a lo largo de una superficie puede cambiar muy rápidamente debido a las fronteras de las sombras. Para poder captar estos cambios se necesitaría de una gran cantidad de patches, desafortunadamente, la matriz de los factores de forma crece por un factor igual al número de patches al cuadrado. Así que el método podría llegar a ser excesivamente costoso computacionalmente hablando.

Una posible solución es usar una representación jerárquica de la matriz de factores de forma. La idea es reducir el tamaño de la matriz y no el número de elementos.

Cohen propuso una jerarquía a dos niveles [Cohen86]. En el nivel más fino se encuentran los *elementos*. Los elementos son formados de la subdivisión de los *patches*. Y los patches como ya se había mencionado son las regiones en las que se subdivide una superficie. La transferencia de energía se lleva a cabo al nivel de

patches, y como hay un número mucho menor de patches que de elementos, esto resulta en una aceleración significativa en el cálculo.

Esta idea de subestructuración es muy empleada en problemas de elemento finito. La idea consiste en tratar a las fuentes a una resolución menor que a los receptores. La suposición donde se refleja es la idea de representar una fuente como un patch con una radiosidad constante, aún cuando la radiosidad varíe sobre el patch. Esta es una suposición frecuentemente adecuada y que no introduce errores notables. Por otra parte, no es aceptable representar grandes áreas receptoras con un solo elemento de radiosidad constante, pues estas radiosidades son empleadas para el despliegue final de la escena y una resolución insuficiente en la representación de la solución final de la radiosidad produciría una imagen inaceptable.

Al igual que en la formulación original de radiosidad, se asume una radiosidad constante en cada elemento. La radiosidad de un patch está dada como un promedio de las radiosidades de los elementos pesado por sus áreas. Al promediar de esta forma, se asegura que la energía que sale del patch es igual a la suma de las energías saliendo de cada uno de los elementos.

$$B_{\tilde{n}} A_{\tilde{n}} = \sum_c B_{c_i} A_{c_i} \quad (4.1)$$

En una forma similar, los factores de forma de patch a patch son la suma ponderada por las áreas de los factores de forma de los elementos.

$$F_{\tilde{n}, \tilde{p}_j} A_{\tilde{p}_j} = \sum_{\tilde{n}_i} \sum_{p_j} F_{c_i, c_j} A_{c_i} \quad (4.2)$$

Claro que si se calculan los factores de forma de esta manera se requiere de  $m^2$  cálculos si existen  $m$  elementos, porque se deben considerar todos los pares de elementos.

El factor de forma de patch a patch también se puede obtener de la suma de los factores de forma de patch a elemento.

$$F_{\tilde{n}, \tilde{p}_j} A_{\tilde{p}_j} = \sum_{p_i} F_{\tilde{n}, c_i} A_{c_i} \quad (4.3)$$

Aquí el factor de forma de patch a elemento es calculado usando la derivación del factor de forma de un área finita (patch) a un área diferencial (elemento).

Este razonamiento lleva de forma natural al siguiente algoritmo:

1. Calcular las matrices de factores de forma patch-elemento y patch-patch. Los factores de forma patch-elemento son calculados de acuerdo a la mejor técnica disponible y los factores patch-patch a partir de las relaciones anteriores.

2. Resolver el sistema de ecuaciones de patch a patch para las  $n$  radiosidades de los patches usando un algoritmo iterativo de Gauss-Seidel.

3. Calcular las radiosidades de los elementos multiplicando la matriz de factores de forma patch-elemento por las radiosidades de los patches.

Por lo general hay muchos más elementos que patches, así que un método jerárquico es más rápido de calcular que la solución entre todos los elementos.

#### 4.3.2 Tipos de Datos

Las principales estructuras usadas son arreglos y listas indexadas de *objetos* que encapsulan sus atributos. Estos objetos son:

Los parámetros de la radiosidad, que agrupa a todos los objetos o apuntadores a los objetos necesarios para realizar el cálculo de las radiosidades.

```
typedef struct {
    double umbral;           /* umbral de convergencia (% de la energía emitida) */
    unsigned long nCuad;    /* número de cuadriláteros */
    TCuad *cuad;           /* apuntador a cuadriláteros */
    TPunto3f *vertices;    /* vértices de los cuadriláteros */
    TSpectro *pColores;    /* colores de los cuadriláteros */
    unsigned long nPatches; /* número de patches */
    TPatch *patches;      /* apuntador a los patches */
    unsigned long nElementos; /* número de elementos */
    TElemento *elementos; /* apuntador a los elementos */
    unsigned long nPuntos; /* número de vértices de los elementos */
    TPunto3f *puntos;     /* apuntador a los vértices de los elementos */
    TVista vistaDespliegue; /* características del despliegue */
    unsigned short resHemicubo; /* resolución del hemicubo */
    float tamanoEscena;    /* diámetro aprox. de la esfera que acota la escena. */
    float escalasIntensidad; /* usado para magnificar la intensidad de despliegue */
    int agregaAmbienta;    /* cuando agregar la aprox. ambiental en el despl. */
} TParamsRad;
```

El objeto que representa una vista, es decir la definición de las características con las que se establece un sistema de visión de una escena.

```
typedef struct {
    TPunto3f camara;       /* posición de la cámara */
    TPunto3f miraa;       /* punto de interés que se desea ver (centro) */
    TVector3f arriba;     /* vector hacia arriba */
    float advx, advy;     /* ángulo de visión en x, y (en grados) */
    float cerca, lejos;   /* distancia desde la cámara a los planos cerca y lejos */
    unsigned short xRes, yRes; /* resolución de la imagen */
    unsigned long* buffer; /* apuntador a el frame buffer de la imagen */
    long wid;            /* id o apuntador a la ventana asociada */
} TVista;
```

El objeto que representa a un cuadrilátero en la escena (las superficies), contiene sus elementos geométricos y sus propiedades reflectivas y emisivas.

```
typedef struct {
    short verts[4];           /* vértices del cuadrilátero */
    short nivelpatch;        /* nivel de subdivisión del patch */
    short nivelelemento;    /* nivel de subdivisión del elemento */
    float area;              /* Área del cuadrilátero */
    TVector3f normal;        /* normal del cuadrilátero */
    TSpectro* reflectancia;  /* reflectancia difusa del cuadrilátero */
    TSpectro* emision;       /* emisión del cuadrilátero */
    unsigned long elementos; /* apuntador a los elementos del cuadrilátero */
} TCuad;
```

Un Patch es modelado con un objeto que contiene sus propiedades reflectivas y emisivas, atributos relacionados con el hemisferio y la radiosidad que debe distribuir.

```
typedef struct {
    TSpectro* reflectancia; /* reflectancia difusa del patch */
    TSpectro* emision;     /* emisión del patch */
    TPunto3f centro;       /* centro del patch donde se colocará el hemisferio */
    TVector3f normal;      /* normal del patch, para orientación del hemisferio */
    TSpectro RadSinLanzar; /* radiosidad sin lanzar por el patch */
    double area;           /* área del patch */
} TPatch;
```

El objeto que representa un elemento, contiene la información geométrica que lo define, la radiosidad que refleja, un indicador del patch al que pertenece y su color en el modelo RGB que es asociado al despliegue.

```
typedef struct {
    unsigned short nVerts; /* número of vértices del elemento */
    unsigned long verts[4]; /* vértices */
    TVector3f normal;      /* normal del elemento */
    TSpectro rad;          /* radiosidad total del elemento */
    double area;           /* área del patch */
    TPatch* patch;        /* apuntador al patch padre */
    unsigned long color;   /* color del elemento */
} TElemento;
```

Para manejar la transferencia de energía se utiliza un espectro de energía que en este caso sólo emplea tres muestras de las longitudes de onda que forman el espectro.

```
#define kNumeroMuestrasRad 3
```

```
typedef struct {
    double muestras[kNumeroMuestrasRad];
} TSpectro;
```

Estas son las estructuras manejadas durante la solución de las radiosidades. Al terminar la solución, el sistema deja como datos de salida un archivo que contiene la

información necesaria para reestablecer las condiciones de la escena y los valores de radiosidad calculados para cada uno de los vértices que forman los elementos. Estos datos pueden ser tomados por el programa *Visual* y recorrer la escena de forma inmediata sin tener que procesar de nuevo los datos.

#### 4.4 Esquema del Sistema

El esquema seguido por un algoritmo de radiosidad tiene las siguientes etapas:

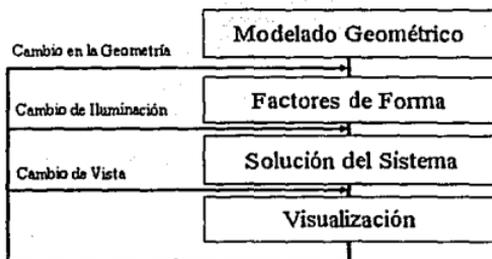


Figura 4.1. Algoritmo de radiosidad.

**Modelado Geométrico**, en esta etapa se crea la forma de la escena y se asignan propiedades reflectivas y emisivas a cada una de las superficies.

**Factores de Forma**, aquí se establecen las relaciones geométricas entre los patches de las superficies, esta es la parte del proceso que consume mayor cantidad de recursos de cómputo.

**Solución del Sistema**, una vez establecidas las interdependencias de los patches para sus radiosidades, se resuelve el sistema de ecuaciones, para cada longitud de onda de interés, por medio de un método iterativo de Gauss-Seidel.

**Visualización**, ya con el sistema resuelto, los valores de las radiosidades calculados son empleados para crear una imagen de la escena iluminada.

En caso de ocurrir modificaciones se repite el proceso desde cierta etapa, dependiendo del tipo de modificaciones efectuadas, por ejemplo, si se modifica la posición del observador o la dirección de observación, entonces sólo hay que repetir la última etapa, pero en si el cambio ocurrió en la geometría, es decir si se cambió de lugar un objeto o se agregó alguno más, entonces hay que regresar a la etapa de los factores de forma.

Sin embargo, cuando se emplea un método de refinamiento progresivo de las radiosidades, las etapas del cálculo de los factores de forma y la solución se intercalan

dentro de una misma etapa que sería el refinamiento de las radiosidades, de hecho la visualización también se puede intercalar en esta misma etapa.

#### 4.4.1 Inicialización del sistema

En el sistema desarrollado se asume que la etapa de modelado se realiza de forma externa al sistema, por lo que el modelado geométrico se toma como parte de los datos de entrada.

El modelado geométrico actualmente se realiza utilizando un software comercial llamado *3DStudio*, el cual tiene facilidades para el modelado de objetos. Con este software se definen los objetos que componen la escena, el modelado se hace en base a cajas no necesariamente regulares, es decir, sus caras y aristas no tienen que ser paralelas. Además de los objetos también se define la posición del observador y la dirección de visión de éste, con las que se visualizará la escena durante el proceso de solución del refinamiento progresivo.

Este software crea un archivo de texto con la descripción geométrica de los cuadriláteros que forman los objetos, este archivo tiene un formato que no es muy adecuado para las necesidades del programa de radiosidad, así que el archivo es procesado y dividido en dos archivos, cada uno conteniendo información necesaria para el programa de radiosidad, uno de los archivos contiene únicamente información geométrica (con extensión .geo) y el otro (con extensión .rad) la características con las que se desea emplear el método de radiosidad progresiva. Esta conversión sólo adecúa los datos al formato necesario, el archivo todavía está incompleto porque no contiene la información acerca de las propiedades reflexivas o emisivas, ni de como cada superficie será estructurada.

**El formato de los archivos .GEO es:**

```
- NUMERO DE PUNTOS
[ LISTA DE LOS PUNTOS EN COORDENADAS (X,Y,Z) ]

- NUMERO DE COLORES
[ LISTA DE COLORES USADOS EN FORMATO : { NUM_COLOR, R, G, B } ]

- NUMERO DE CUADRILÁTEROS
{ LISTA DE CUADRILÁTEROS EN FORMATO :
  ( INDICES_DE_VERTICES,
    SUBDIVISION_PATCH, SUBDIVISION_ELEMENTO,
    COLOR_REFLECTANCIA, COLOR_EMISION
  ) }
```

**El formato de los archivos .RAD es:**

- ÚMBRAL DE CONVERGENCIA PARA LA SOLUCION
- INTENSIDAD DE LAS FUENTES
  
- POSICION DEL OBSERVADOR
- PUNTO QUE SE DESEA VER
- DIRECCION HACIA ARRIBA
- CAMPO DE VISION (PARA X y Y)
- RESOLUCION DE LA IMAGEN

Una vez que se tienen estos archivos sólo hay que agregar los datos no suministrados anteriormente, editando estos archivos y llenando los espacios libres dejados para esos datos.

El programa para el cálculo de las radiosidades tiene entonces, en forma simplificada la siguiente forma:

```
int main()
{
    LeeDatosRad();
    InitParams();
    InicializaRad();
    HazRad();
    Visualiza();
    EscribeResultados()
    LimpiaRad();
}
```

donde el modelado geométrico es sustituido por la lectura de los datos generados previamente. Después se hacen las inicializaciones necesarias antes de lanzar el proceso de solución. Se llama al procedimiento que crea la solución y una vez alcanzada la aproximación a la solución, se puede visualizar la escena y hacer un recorrido interactivo de ella. En cuanto se detiene la visualización se guardan las características de la escena junto con sus radiosidades para una futura utilización. Y finalmente se liberan todos los recursos que se solicitaron al sistema, como memoria y ventanas de despliegue.

La lectura de datos se encarga de asignar los valores definidos en el paso externo de modelado a las estructuras internas del programa, dando de alta dentro del sistema los cuadriláteros que componen la escena.

Ahora en base a los cuadriláteros y a su descripción de subestructuración se crean los patches que componen los cuadriláteros, y los elementos que componen a los patches.

```
void InitParams()
{
    nPatches=0;
    for (i=numDePoligonos; i--; )
        nPatches += nivelpatchi2;
    nElementos=0; /* calcula número total de elementos */
    for (i=numDePoligonos; i--; )
        nElementos += (nivelelementoi * nivelpatchi)2;
    nPuntos=0; /* calcula el núm. total de vértices de los elementos */
    for (i=numDePoligonos; i--; )
        nPuntos += (nivelelementoi * nivelpatchi + 1)2;
    for (i=0; i<numDePoligonos; i++) /* subdivide el cuarto a patches y elementos */
        ReticulaCuad( poligCuarto[i] );
}
```

En la inicialización del algoritmo de radiosidad progresiva se obtienen los recursos para aplicar el método del hemisferio y para los factores de forma de cada uno de los elementos en la escena. Como en este esquema de solución sólo es necesario un hemisferio y de este sólo se necesita una cara a la vez, los requerimientos de memoria para el hemisferio bajan considerablemente. Además, se asignan los valores iniciales de radiosidad a partir de los cuales empezará el refinamiento progresivo, donde cada patch empieza con su radiosidad y su radiosidad sin lanzar igual a su emisividad, es decir sólo las fuentes de luz se comportarán como tales al principio del método.

```
void InicializaRad(p)
TParamsRad *p;
{
    { establecer las características básicas del hemisferio }

    CalcFactoresSup(mediaRes/2, hemisferio.coeSuperior);
    CalcFactoresLat(mediaRes/2, hemisferio.coeLateral);

    factoresforma = (double*)calloc( nElementos, sizeof(double));

    pp = patches;                                     /* inicializa radiosidades */
    for (i=nPatches; i--; pp++)
        pp->RadSinLanzar = *(pp->emision);

    ep = elementos;
    for (i=nElementos; i--; ep++)
        ep->rad = *(ep->patch->emision);

    energiaTotal = 0;                                 /* calcula la energía total */
    pp = patches;
    for (i=nPatches; i--; pp++)
        for (j=0; j<kNumeroMuestrasRad; j++)
            energiaTotal += pp->emision->muestras[j] * pp->area;

    DespliegaResultados();
}

```

En el procedimiento anterior también se calculan los deltas de los factores de forma, estos como se había mencionado antes, son la fracción de ángulo sólido que representa cada uno de los pixels que componen las caras del hemisferio, y es en base a estos como se construyen los factores de forma. Como estos factores son independientes de la posición del hemisferio, son calculados y almacenados antes de empezar el proceso de solución, además debido a las simetrías en estos delta factores, es posible almacenar sólo la cuarta parte de ellos y los demás pueden ser obtenidos a partir de estos.

Los delta de los factores de forma se separan en dos conjuntos, los que corresponden a la cubierta superior del hemisferio, y los de las caras laterales. La geometría y sus expresiones algebraicas son dadas a continuación.

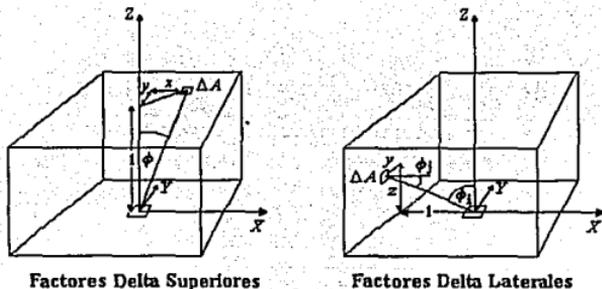


Figura 4.2. Geometría para el cálculo de los factores delta de cada elemento en el hemicubo.

para los deltas de la cara superior

$$r = \sqrt{x^2 + y^2 + 1} \quad \cos\phi_1 = \cos\phi, \quad \cos\phi = \frac{1}{\sqrt{x^2 + y^2 + 1}}$$

$$\Delta F_{arriba} = \frac{\cos\phi_1 \cos\phi}{\pi r^2} \Delta A = \frac{1}{\pi(x^2 + y^2 + 1)^2} \Delta A \quad (4.4)$$

y para los deltas de las caras laterales

$$r = \sqrt{x^2 + y^2 + 1} \quad \cos\phi_1 = \frac{z}{\sqrt{y^2 + z^2 + 1}} \quad \cos\phi = \frac{1}{\sqrt{y^2 + z^2 + 1}}$$

$$\Delta F_{lateral} = \frac{\cos\phi_1 \cos\phi}{\pi r^2} \Delta A = \frac{z}{\pi(y^2 + z^2 + 1)^2} \Delta A \quad (4.5)$$

#### 4.4.2 Proceso del Refinamiento Progressivo

Una vez realizadas las inicializaciones necesarias y que se desplegó la primera aproximación de las radiosidades, el proceso de refinamiento progresivo empieza.

```
void HazRad()
{
    while ( EncuentraPatchEmisor( &patchEmisor ) )
    {
        CalcFactoresForma( patchEmisor );
        DistribuyeRad( patchEmisor );
        DespliegaResultados();
    }
}
```

Como se había visto el método de refinamiento progresivo es un método iterativo el cual prosigue hasta que se cumple algún criterio, en este caso el criterio de

convergencia es que la radiosidad por lanzar máxima en el ambiente sea menor que un cierto umbral que se estableció dentro de los parámetros de radiosidad. Mientras este criterio no sea alcanzado, se busca dentro de todos los parches, aquel que tenga una radiosidad por lanzar mayor a la de los demás, seleccionándolo como el patch emisor para este paso de la iteración.

Una vez que se sabe cual es el patch emisor se calculan los factores de forma desde este patch hacia todos los elementos, utilizando un hemisucubo centrado en el patch y orientado respecto al vector normal del patch.

```
static void CalcFactoresForma(patchEmisor)
unsigned long patchEmisor;
{
    /* asigna la direccion de vision y la normal para cada cara */
    fp = factoresforma; /* limpia los factores de forma */
    for (i=nElementos; i--; fp++)
        *fp = 0.0;

    for (cara=0; cara < 5; cara++)
    {
        hemisucubo.vista.miraa = miraa[cara];
        hemisucubo.vista.arriba = arriba[cara];

        EmpiezaIluminar(); /* dibuja elementos en el hemisucubo */
        for (i=0; i< nElementos; i++)
            DibujaElemento(elementos[i], i);
        FinIluminar(&(hemisucubo.vista)); /* obtiene factores de forma */

        if (cara==0)
            SumaFactores(factoresforma, vista.buffer, hemisucubo.coefSuperior);
        else
            SumaFactores(factoresforma, vista.buffer, hemisucubo.coefLateral);
    }

    ep = elementos; /* calc. factores de forma reciprocos */
    fp = factoresforma;
    for (i=nElementos; i--; ep++, fp++)
        *fp *= ep->area / ep->area;
}
}
```

El primer paso es crear las cinco caras del hemisucubo con las que se realizarán los cálculos de visibilidad y aportación de cada una de las direcciones a los factores de forma.

Una vez que se han inicializado a cero todos los factores de forma de los elementos, se empiezan a revisar cada una de las caras, empezando por la cara superior del hemisucubo, el proceso de revisión es en base a la creación de una imagen sobre la cara del hemisucubo, para esto se preparan las condiciones de visión de la escena, después cada uno de los elementos que la componen es dibujado en la imagen y finalmente se obtienen los valores de los pixels en la imagen.

Con los valores obtenidos de la proyección de la escena en la cara del hemicubo, y junto con los factores de forma delta calculados para la cara que se está considerando, se obtienen los factores de forma para cada uno de los elementos.

Ahora simplemente se obtienen los factores de forma recíprocos que son los que se necesitan para realizar la solución iterativa por distribución de la energía, que es el caso del refinamiento progresivo.

Detallando las características de la creación de las imágenes para un análisis más fino del proceso, se tiene el código simplificado de la rutina:

```
void EmpiezaIluminar()
```

```
{
    { se abre una ventana con las características determinadas para el hemicubo }
    RGBmode();
    zbuffer(TRUE);
    perspective(vista->advy, 1, vista->cerca, vista->lejos);
    shademodel(FLAT);
    ColocaVista( vista->camara, vista->miraa, vista->arriba );
}
```

primero se limpia el display buffer de la imagen que se va a crear, se establecen las condiciones de visión de la escena de acuerdo a la orientación de la cara del hemicubo. Para la creación de la imagen se utiliza una proyección en perspectiva aplicando un método de sombreado constante para los elementos, esto último es muy importante porque la forma de identificar al elemento que cubre algún cierto pixel, es mediante el valor asociado al pixel en el frame buffer.

Para dibujar cada uno de los elementos en esta proyección establecida, se utiliza la posición del elemento dentro del arreglo de elementos, como el color con el que se dibujará en la proyección, y como el sombreado que se emplea en la creación de la imagen es constante, entonces todos los pixels que sean cubiertos por este elemento, contendrán el índice del elemento.

Al finalizar la imagen se hace una copia del frame buffer empleado en la ventana de despliegue para utilizarse en el cálculo de las aportaciones a los factores de forma por parte de los elementos de la escena.

La forma de encontrar las aportaciones es tomar los datos de la imagen creada, para cada uno de estos datos se modifica su factor de forma sumándole el factor delta de forma que corresponde a la cara y posición del dato en la imagen.

Así después de revisar todas las caras y sumar las contribuciones a los factores de forma correspondientes, se tienen los factores de forma del patch a los elementos.

Ahora se tiene que distribuir la energía que no ha lanzado el patch hacia los elementos en base a sus factores de forma.

Como se puede ver en el siguiente código, el procedimiento para distribuir la radiosidad de un patch hacia los elementos de la escena, considera cada uno de los

elementos que tengan un factor de forma distinto de cero, es decir para aquellos que pueden ver al patch emisor, a estos se les añade a sus radiosidades, en cada longitud de onda considerada, la energía emitida modificada por el factor de forma y por la reflectancia del elemento, y a la energía por lanzar del patch al que pertenecen le agrega la energía dependiendo del porcentaje del área representada por el elemento en el patch, y dejando sin energía por lanzar a el patch emisor, puesto que la acaba de distribuir.

```
static void DistribuyeRad(patchEmisor)
{
    sp = patchEmisor;
    ep = elementos;
    fp = factoresforma;
    for (i=nElementos; i--; ep++, fp++) { /* distribuye RadSinLanzar a cada elemento */
        if ((*fp) > 0.0) {
            for (j=0; j<kNumeroMuestrasRad; j++)
                deltaRad.muestras[j] = RadSinLanzar.muestras[j]
                    * (*fp) * reflectancia->muestras[j];
            /* incrementa la radiosidad del elemento y la no lanzada */
            w = ep->area/ep->patch->area;
            for (j=0; j<kNumeroMuestrasRad; j++) {
                ep->rad.muestras[j] += deltaRad.muestras[j];
                ep->patch->RadSinLanzar.muestras[j] += deltaRad.muestras[j] * w;
            }
        }
    }
    RadSinLanzar = negro;
}
}
```

Al final de cada paso iterativo se despliegan los resultados intermedios de ese paso para que el usuario vaya viendo el desarrollo de la solución.

```
static void DespliegaResultados(vista)
{
    ObtenAmbiente(ambiente);
    EmpiezaIluminar(vista, 0);
    ep = elementos;
    for (i=0; i< nElementos; i++, ep++) {
        if (agregaAmbiental) { /* agrega aproximacion ambiental */
            for(k=kNumeroMuestrasRad; k--;)
                s.muestras[k] = ( ep->rad.muestras[k] + (ambiente.muestras[k]*
                    patch->reflectancia->muestras[k]) ) * escalaIntensidad;
        } else {
            for (k=kNumeroMuestrasRad; k--;)
                s.muestras[k] = ep->rad.muestras[k] * escalaIntensidad;
        }
        ep->color = Spectro_a_RGB(&s); /* cuantifica color */
        DibujaElemento(ep, ep->color);
    }
    FinIluminar(vista);
}
}
```

Para desplegar los resultados intermedios se utilizan las mismas rutinas empleadas en la creación de las imágenes del hemisferio, sólo que ahora en vez de emplear el índice

del elemento como el color asociado, se utiliza una cuantificación de la radiosidad del elemento.

Para empezar se calcula el término de iluminación ambiental visto en la sección 2.6.2, el cual es sólo una representación de la energía que aún no ha sido lanzada en el ambiente, como se había mencionado este término sólo tiene efecto en el despliegue de la escena, es por esto que no está incluido en la radiosidad del elemento, además este término es calculado para cada paso de la solución disminuyendo conforme la aproximación converge a la solución. Este valor sólo es agregado a la radiosidad en caso de que haya sido solicitado en los datos de entrada.

Con esto termina cada uno de los pasos del proceso de refinamiento progresivo y en cuanto el patch emisor tenga menor energía que lanzar que el umbral preestablecido, el proceso se detendrá. Obteniendo la solución a las radiosidades que se buscaba para un cierto margen de error.

#### 4.4.3 Proceso Interactivo de Visualización de la Escena

Ahora bien, ya se tiene una solución de radiosidades pero ésta es sólo para elementos de la superficie, los cuales se recubren con el color que representa su radiosidad de forma constante sobre todo el elemento, por lo que la imagen se ve como si estuviera hecha en base a pequeños mosaicos.

En el programa Genilum además se agregó un procedimiento que permite hacer el recorrido interactivo a través de la escena, que normalmente hace el programa Visual este procedimiento se debe remover en caso de estar en una plataforma que no proporcione las rutinas gráficas por medio del hardware, porque como ya se había mencionado, hacer las rutinas por software no permitiría crear las imágenes necesarias en un rango de tiempo interactivo.

Dentro de las características especiales de este procedimiento está el uso de un doble buffer para crear la animación del recorrido por la escena, además del sombreado que se aplica a los elementos de la escena, que es de tipo Gouraud, con esto además de la interacción de movimiento dentro de la escena se logran suavizar los cambios de iluminación en las superficies de la escena. El sombreado de tipo Gouraud hace una interpolación bilineal sobre las intensidades, definidas en los vértices de los elementos dibujados, por lo que es necesario obtener estos valores antes de empezar a crear la imagen de la escena.

```
void Visualiza(params)
{
    RGBmode();
    doublebuffer();
    zbuffer(TRUE);
    perspective(vista->advy, 1, vista->cerca, vista->lejos);
    shademodel(GOURAUD);
    backface(1);

    for(cp = params->cuad, i=0; i<params->nCuad; i++, cp++) {
        ObtenColores(cp, params);
    }
}
```

```

while ( { mientras no se use la tecla ESC } ) {
  czclear(0,zval);
  pushmatrix();
  if (getvaluator(RIGHTMOUSE)) { /* acercarse o alejarse */
    vista->camara.x += desplazamiento * dirV.x ;
    vista->camara.y += desplazamiento * dirV.y ;
    vista->camara.z += desplazamiento * dirV.z ;
  }
  if (getvaluator(LEFTMOUSE)) { /* girar la vista a los lados */
    rotLateral(dirV,dirD,rotacion);
  }
  if (getvaluator(MIDDLEMOUSE)) { /* subir o bajar la vista */
    rotVertical(dirV, vista->arriba, rotacion);
  }
  ColocaVista(vista->camara, vista->miraa, vista->arriba );
  for(cp = cuad, i=0; i<nCuad; i++, cp++) {
    iluminaCuadrilatero( cp, puntos, elementos );
  }
  popmatrix();
  swapbuffers();
}

```

Como las radiosidades fueron calculadas para los elementos y no para sus vértices es necesario obtener los valores adecuados para los vértices teniendo en cuenta que un vértice puede ser compartido por varios elementos, así que habría que tomar en cuenta a todos los elementos que comparten el vértice para obtener su color.

El esquema seguido para obtener estos valores en los vértices es de acuerdo al siguiente esquema:

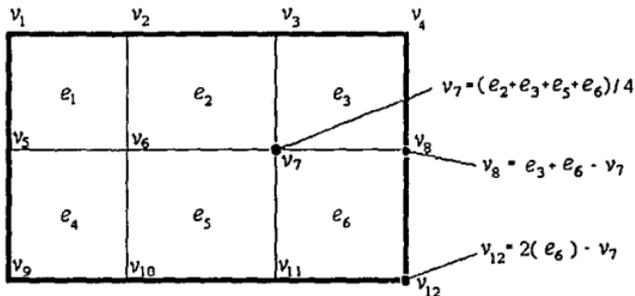


Figura 4.3. Esquema de interpolación de color de los elementos para obtener los valores de los vértices.

tomando el promedio de las radiosidades de los elementos circunvecinos se obtiene la radiosidad para los vértices interiores. Para los vértices en las orillas de la superficie se emplea una extrapolación de los valores de radiosidad.

El proceso de visualización emplea un ciclo en el que los eventos de mouse son leídos e interpretados para ser reflejados en un determinado movimiento dentro de la escena, pudiendo avanzar o retroceder, girar a la derecha o izquierda y alzar o bajar la vista.

El evento leído modifica las condiciones de visión, y una vez establecidas se crea la imagen de la escena mandando a hacer el render de los cuadriláteros, esto significa emplear las radiosidades calculadas para su iluminación.

Para terminar este ciclo se utiliza la tecla de escape, procediendo a almacenar los valores de radiosidad junto con la información geométrica necesaria para reestablecer las condiciones de la escena en el programa Visual.

#### 4.5 Resultados

Los resultados obtenidos con el sistema varían en calidad dependiendo del nivel de subdivisión de los patches y de la aproximación requerida. Por ejemplo, la generación de las radiosidades de una escena muy simple, la cual consiste de un cuarto con dos paralelepípedos y una fuente de iluminación en el centro (imágenes 1 y 2), requirió de los siguientes tiempos de acuerdo a los niveles de subdivisión utilizados y la aproximación solicitada.

**La escena consta de 18 cuadriláteros y se utilizó un hemisferio de 200 píxeles por lado.**

#de Patches	#de Elementos	Umbral ( en % de energía total)		
		0.1	0.01	0.001
18	2096	7 seg.	16 seg.	34 seg.
		6 iters.	16 iters.	34 iters.
94	2096	4 seg.	20 seg.	2 min.
		4 iters.	18 iters.	126 iters.
188	8384	0 seg.	1'18 min.	6'27 min.
		0 iters.	36 iters.	184 iters.

En otra escena compuesta por un cuarto con 6 columnas y 4 fuentes de luz (imágenes 3 y 4), se obtuvo el siguiente tiempo.

**Esta escena consta de 66 cuadriláteros y se utilizó un hemisferio de 300 píxeles por lado.**

#de Patches	#de Elementos	Umbral ( en % de energía total)		
		0.1	0.01	0.001
308	6582			9 min. 184 iters.

Obviamente a mayor cantidad de patches y elementos, la calidad de la imagen generada es mejor. No obstante, si el propósito principal es interactuar con la escena en tiempo real, hay que tratar de limitar un poco el número de elementos, para que la velocidad de despliegue permita hacer un movimiento continuo.

En las imágenes 5, 6 y 7, se muestra otro ejemplo de las radiosidades calculadas con el sistema, en esta escena se aumenta el número de cuadriláteros empleados y se tiene un modelado más elaborado. En las imágenes se pueden percibir el manejo de sombras que tiene el método de radiosidad. Estas sombras tienen una región de penumbra, eliminando las sombras marcadas que se obtienen con otro tipo de métodos.

Finalmente, como un resultado extra, se agregó el mapeo de texturas sobre las superficies, a las que previamente se les calculó sus radiosidades (imágenes 8 y 9). Y para completar el modelo de iluminación, se agregó el uso de ray tracing para el manejo de la iluminación especular, empleando una técnica híbrida (sección 3.3) de dos pasos, empleando las radiosidades como primer paso y ray tracing como segundo paso (imágenes 10 y 11).

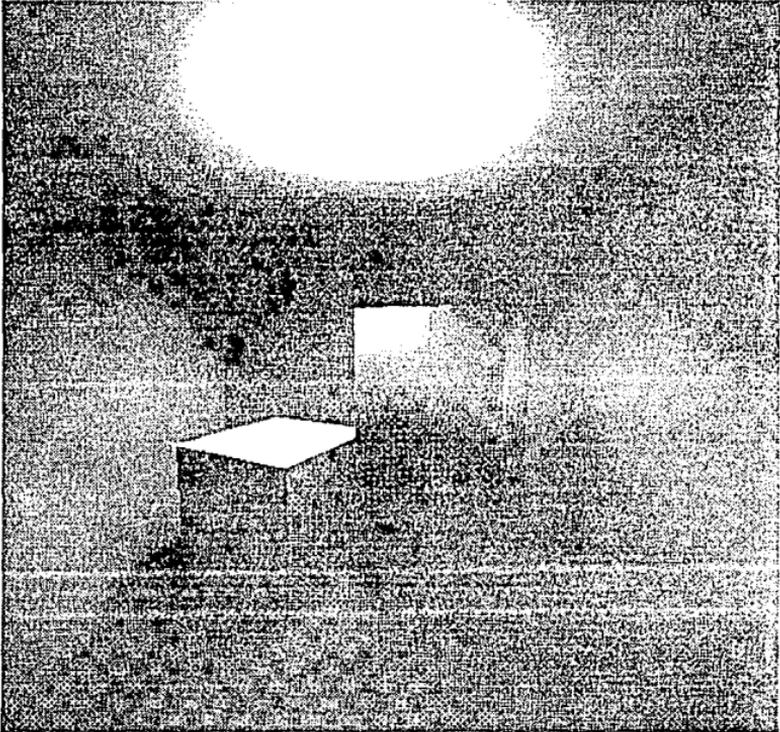


Imagen 1. Escena de un cuarto con dos cajas en el centro, en esta escena es posible observar las sombras suaves, características del método de radiosidad.

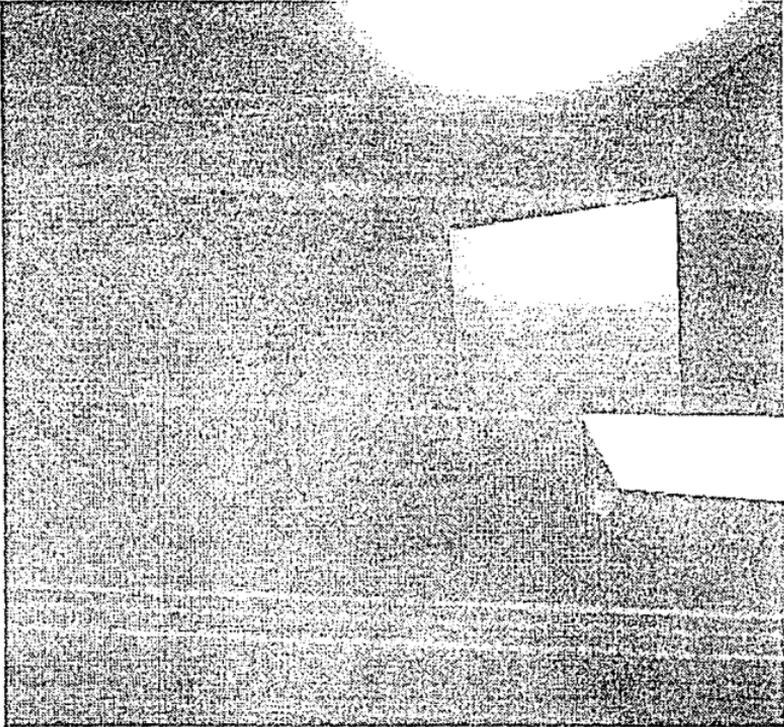


Imagen 2. Otra vista de la misma escena de la imagen anterior.

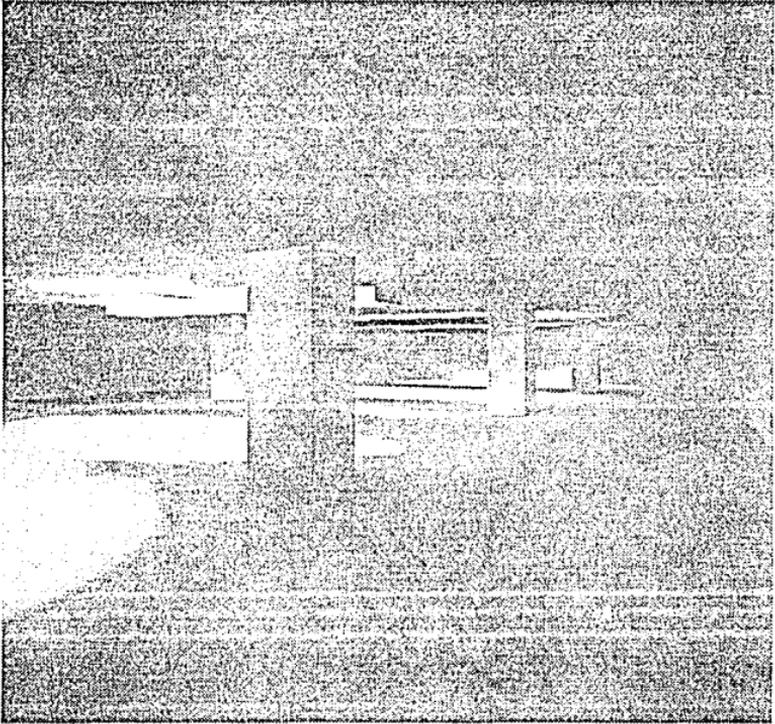


Imagen 3. Ejemplo de otra escena en la cual se tiene una habitación vacía, en la cual sólo se ven las columnas interiores y las lámparas. En esta escena es muy claro el papel de las interreflexiones, ya que las lámparas están dirigidas hacia el piso y éste es el que ilumina el techo.

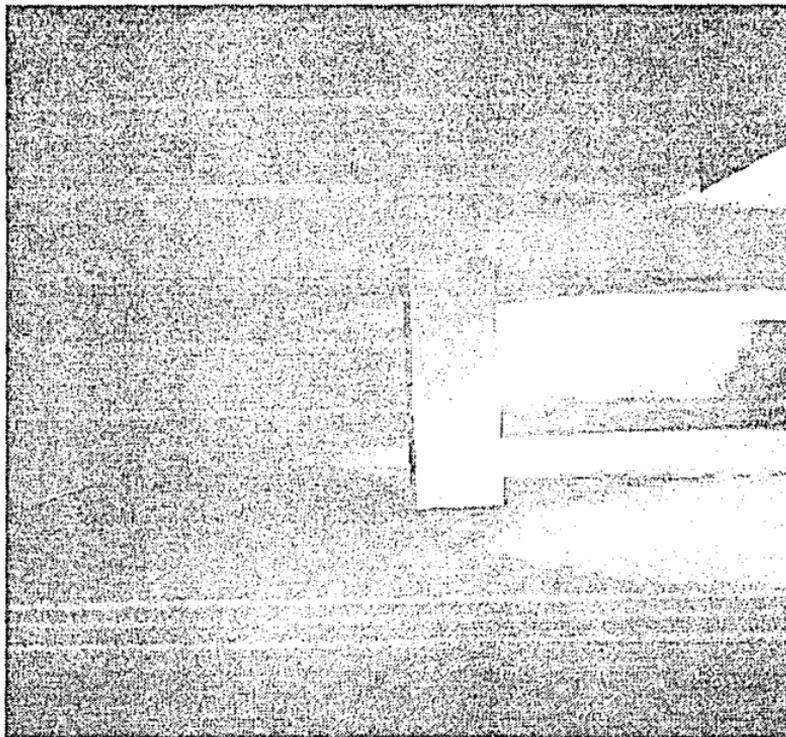


Imagen 4. Otra vista de la escena anterior.

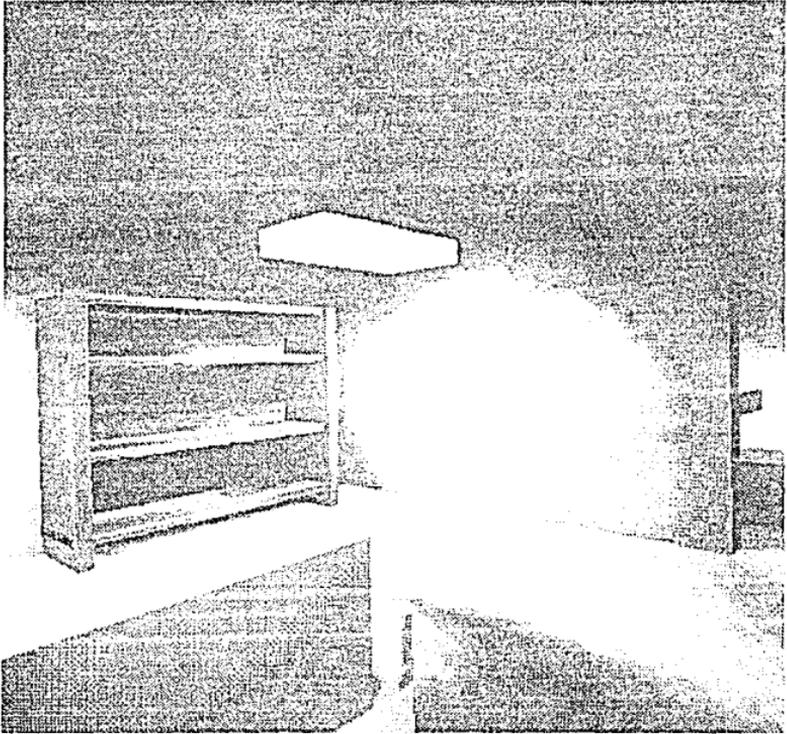


Imagen 5. Escena compuesta por varios cuarto y algunos objetos colocados en éstos, al igual que en las otras imágenes, nótese la forma de las sombras producidas por las fuentes de luz, que no son consideradas como puntuales, como normalmente lo hacen los sistemas gráficos comerciales.

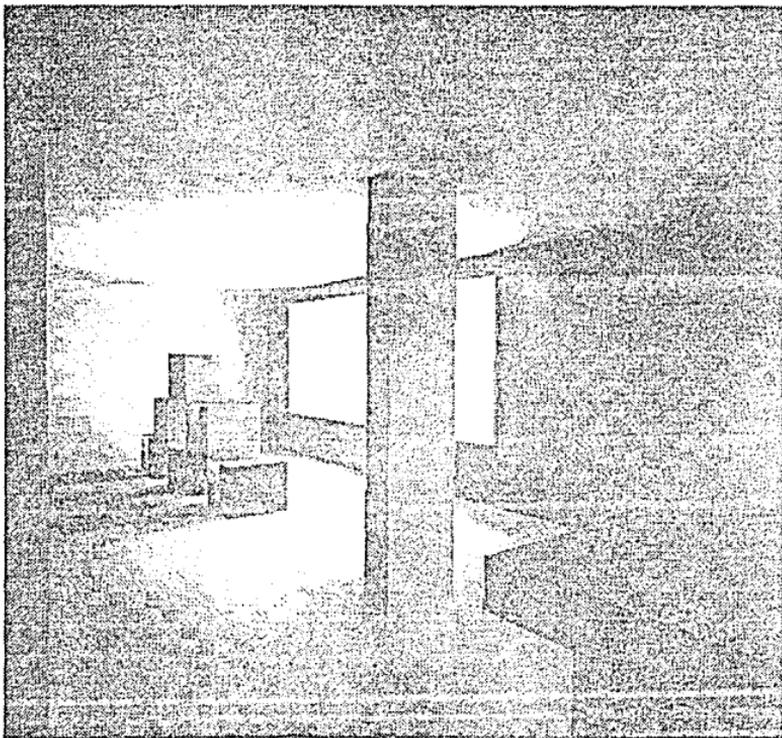


Imagen 6. Otro de los cuartos de la escena anterior, en este caso la fuente de luz se colocó en una de la paredes para observar las sombra proyectadas, así como la iluminación en las zonas donde la luz no llega directamente.

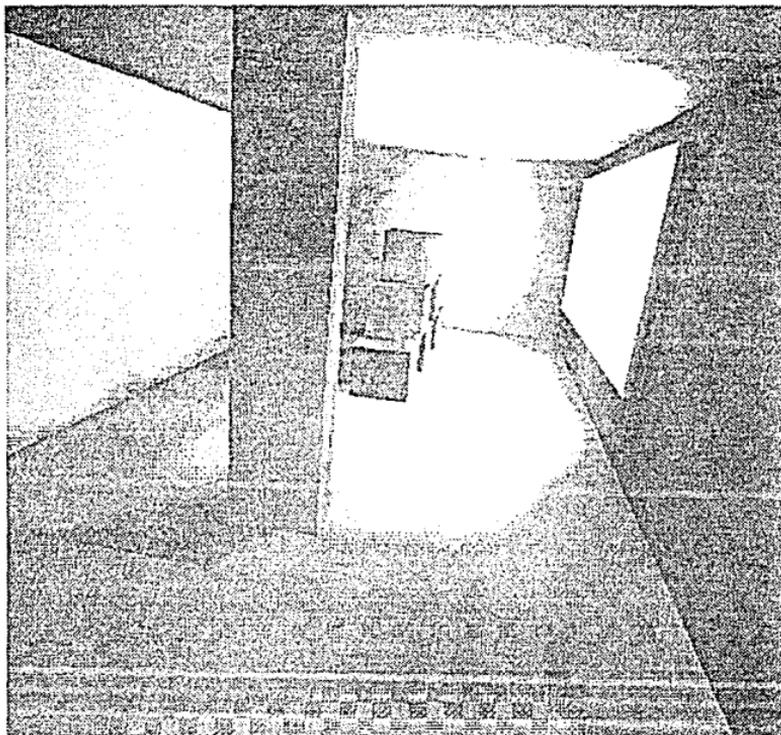


Imagen 7. Otra vista de la escena anterior.

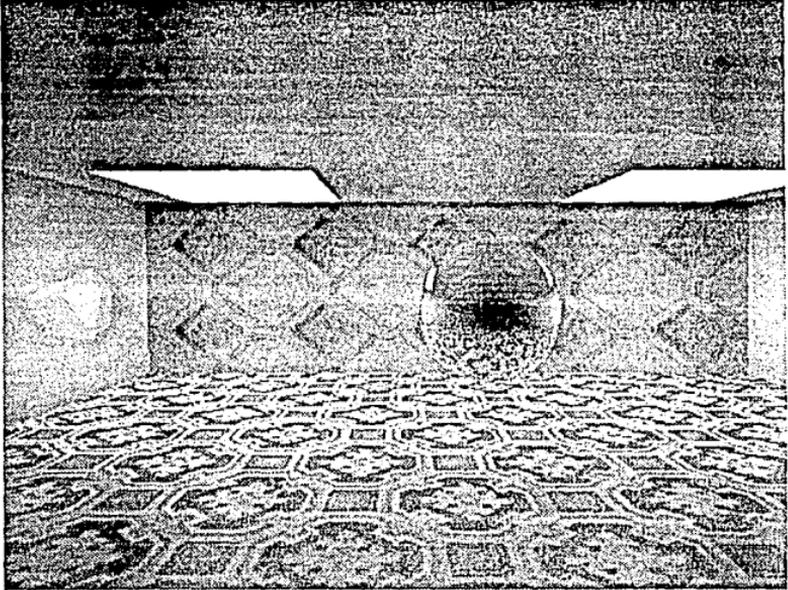


Imagen 8. Mapeo de texturas sobre las superficies de los objetos en la escena. En esta imagen se empleó el método de radiosidad para calcular la iluminación de las superficies y después se mapearon las texturas, además se le agregó una esfera de espejo en el centro del cuarto, para la cual se utilizó el método de ray tracing para obtener su iluminación.

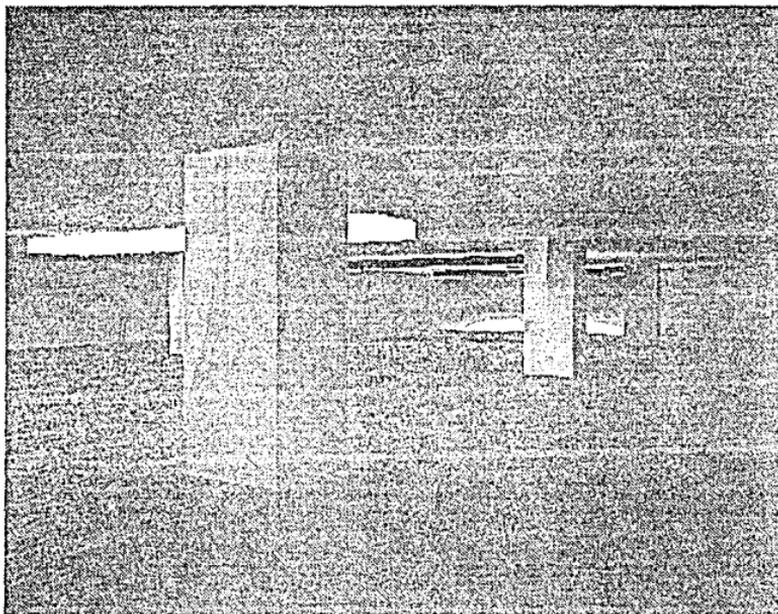


Imagen 9. Esta imagen es la misma escena de las imágenes 3 y 4 pero aplicando mapeo de texturas en el piso y el techo.

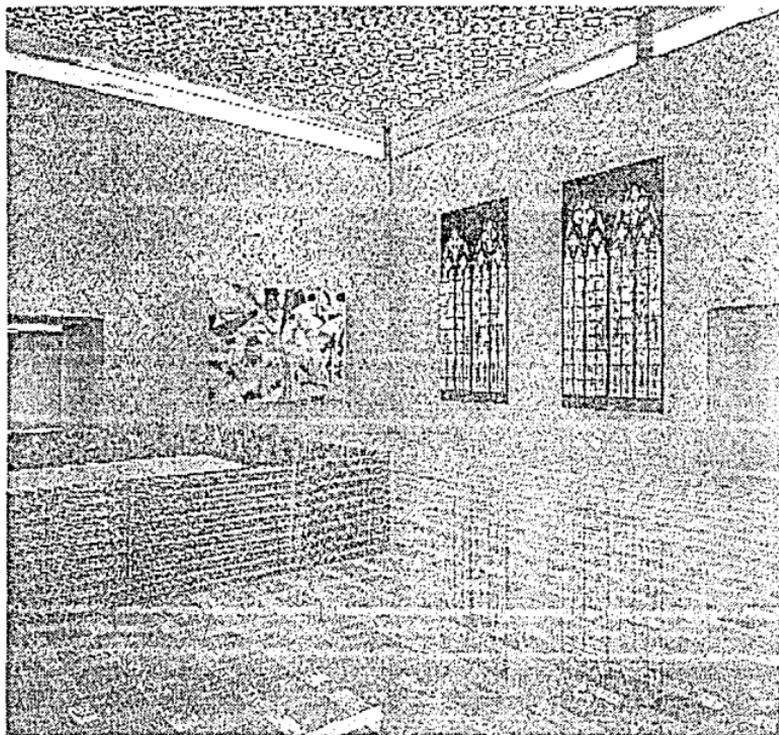


Imagen 10. Escena con un manejo total de la iluminación, para la cual se empleó un método de iluminación híbrido, combinando radiosidad con ray tracing. Con este modelo es posible simular la reflexión de cualquier tipo de material.

**ESTA TESIS NO DEBE  
SALIR DE LA BIBLIOTECA**

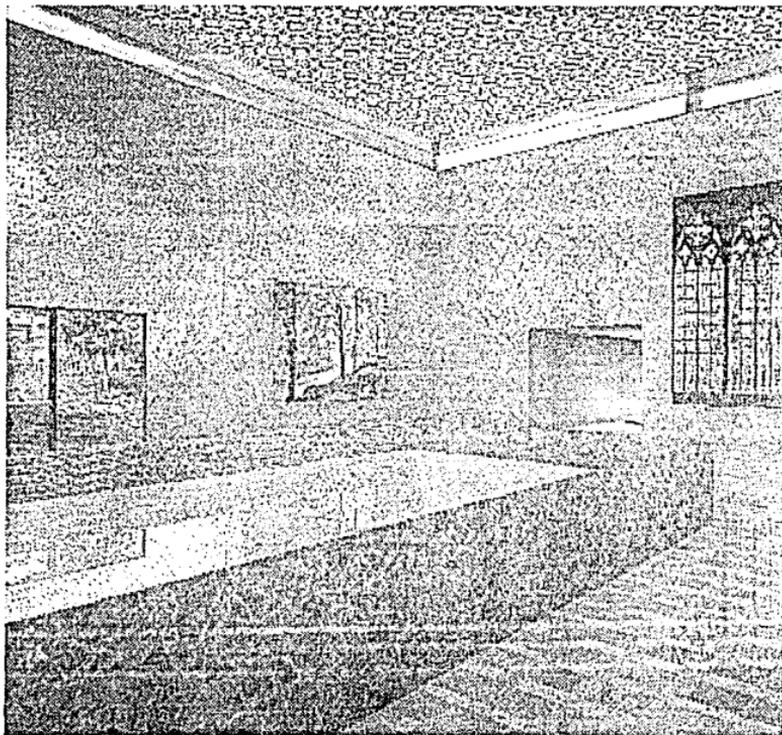


Imagen 11. Otra vista de la escena anterior con las mismas características de iluminación. En ésta se puede observar como los materiales combinan reflexión difusa con especular, por ejemplo en el piso o en la superficie del agua.

## CONCLUSIONES

Con el sistema desarrollado se han logrado crear escenas con un manejo completo de las interreflexiones en un ambiente y se ha explotado el poder gráfico de las estaciones de trabajo en beneficio de crear la interacción con una escena, para cuyo "render" se empleó un modelo global. Esto es algo que no está contemplado en los sistemas comerciales, además de que no existen desarrollos comerciales que incluyan este método de iluminación ni siquiera para la generación de imágenes estáticas.

Los tiempos de generación de imágenes en realidad son rápidos respecto a la implantación de Sumant N. Pattanaik [Pattanaik93] que se pudo comparar, aunque no es muy confiable esta comparación, pues se tiene la ventaja de las rutinas por hardware empleadas en este sistema.

Un resultado importante del sistema es que ha probado ser un método factible de aplicar en procesos de visualización de escenas donde el manejo de las interreflexiones es importante, básicamente en la arquitectura, aunque no por esto se descartan otro tipo de aplicaciones. De hecho en la literatura sobre realidad virtual pueden encontrarse muchos otros ejemplos. Otro campo de aplicación es la visualización científica.

Con este sistema se ha llenado un vacío que existía en el proyecto "Síntesis de Imagen" de la Facultad de Ciencias de la UNAM, completando los elementos para el cálculo de la iluminación global, pero aún quedan muchas cosas por hacer y mejorar, quizás en un futuro próximo sea posible manejar las reflexiones especulares más rápidamente y se alcance la interactividad de imágenes generadas con un modelo global completo.

# BIBLIOGRAFIA

[Baum89]

Daniel R. Baum, Holly E. Rushmeier y James M. Winget.  
"Improving radiosity solutions through the use of analytically determined form-factors"  
ACM SIGGRAPH'89 Conference Proceedings, Vol. 23, No. 3 pp.325-334, July 1989.

[Campbell90]

A.T. Campbell III y Donald S. Fussell.  
"Adaptive mesh generation for global diffuse illumination"  
ACM SIGGRAPH'90 Conference Proceedings, Vol 24, No. 4 pp.155-164, August 1990.

[ChenWu90]

Hong Chen, En-Hua Wu,  
"An efficient radiosity solution for bump texture generation"  
ACM SIGGRAPH'90 Conference Proceedings, Vol 24, No. 4 pp.125-134, August 1990

[Chen88]

Shenchang E. Chen, M.F. Cohen, J.R. Wallace, y D.P.Greenberg  
"A progressive refinement approach to fast radiosity image generation"  
ACM SIGGRAPH'88 Conference Proceedings, Vol 22, No.4 pp.75-84, July 1988

[Chen90]

Shenchang E. Chen,  
"Incremental radiosity: an extension of progressive radiosity to an interactive image synthesis system"  
ACM SIGGRAPH'90 Conference Proceedings, Vol 24, No. 4 pp.135-144, August 1990

[Cohen85]

Michael F. Cohen y Donald P. Greenberg,  
"The hemi-cube a radiosity solution for complex environments"  
ACM SIGGRAPH'85 Conference Proceedings, Vol 19, No. 3 pp.31-40, July 1985

[Cohen86]

Michael F.Cohen, Donald P.Greenberg, David S.Immel y Philip J.Brock.  
"An efficient radiosity approach for realistic image synthesis"  
IEEE Computer Graphics and Applications, Vol. 6 No. 2 pp 26-35, March 1986

- [Cohen88]  
Michael F. Cohen  
"A consumer's and developer's guide to radiosity"  
ACM SIGGRAPH'88, Atlanta, Georgia, August 1988.
- [Cohen92]  
Cohen, Campbell III, Greenberg, Hanrahan, Rushmeier, Sillion y Wallace  
"Radiosity"  
Course Notes, SIGGRAPH'92, July 1992
- [Cook84]  
Robert L. Cook, Thomas Porter y Loren Carpenter  
"Distributed ray tracing"  
ACM SIGGRAPH'84 Conference Proceedings, Vol 18, No. 3 pp.137-145, July 1984
- [Hall89]  
Roy Hall,  
*"Illumination and color in computer generated imagery"*  
Springer Verlag, New York, 1989.
- [Heckbert90]  
Paul S. Heckbert,  
"Adaptive radiosity textures for bidirectional ray tracing"  
ACM SIGGRAPH'90 Conference Proceedings, Vol 24, No. 4 pp.145-154, August 1990
- [Hornbeck75]  
Robert W. Hornbeck,  
*"Numerical Methods"*  
Quantum Publishers, New York, N. Y., 1975
- [Immel86]  
David S. Immel, Michael F. Cohen, y Donald P. Greenberg  
"A radiosity method for non-diffuse environments"  
ACM SIGGRAPH'86 Conference Proceedings, Vol.20, No. 4 pp.133-142, August 1986
- [Kajiya86]  
James T. Kajiya,  
"The rendering equation"  
ACM SIGGRAPH'86 Conference Proceedings, Vol.20, No.4 pp.143-150, August 1986
- [Malley88]  
Thomas J.V. Malley,  
"A shading method for computer generated images"  
Reporte Técnico, Universidad de Utah, Junio 1988
- [Mitchell92]  
Don Mitchell y Pat Hanrahan,  
"Illumination from curved reflectors"  
ACM SIGGRAPH'92 Conference Proceedings, Vol. 26 No.4, July 1992
- [Nicodemus77]  
F.E.Nicodemus, J.C.Richmond, J.J.Hsia, I.W.Ginsberg y T.Lamperts  
"Geometric consideration and nomenclature for reflectance"  
Reporte Técnico, National Bureau of Standards, Washington D.C., 1977

## [Nishita85]

Tomoyuki Nishita y Eihachiro Nakamae  
 "Continuous tone representation of three-dimensional objects taking account of shadows and interreflections"  
 SIGGRAPH'85 Conference Proceedings, Vol.19, No. 3 pp.23-30, July 1985

## [Pattanaik93]

Sumant M. Pattanaik,  
 Graphics & CAD Division, National Center for Software Technology  
 sumant@shakti.ncst.ernet.in

## [Santamaría90]

Sergio Santamaría S.  
 "Modelos matemáticos de iluminación para la síntesis de imagen"  
 Tesis de Licenciatura, Fac. de Ciencias, UNAM. 1990

## [Shao88]

Min-Zhi Shao, Qun-Sheng Peng y You-Dong Liang,  
 "A new radiosity approach by procedural refinements for realistic image synthesis"  
 ACM SIGGRAPH'88 Conference Proceedings, Vol.22, No.4 pp.93-101, August 1988

## [Siegel81]

Robert Siegel y John R. Howell,  
 "Thermal radiation heat transfer"  
 Hemisphere Publishing, 1981

## [Sillion89]

Francois Sillion, y Claude Puech.  
 "A general two-pass method integrating specular and diffuse reflection"  
 ACM SIGGRAPH'89 Conference Proceedings, Vol 23, No. 3 pp.335-344, July 1989

## [Wallace87]

John R. Wallace, M.F. Cohen, y D.P. Greenberg.  
 "A two-pass solution to the rendering equation: a synthesis of ray-tracing and radiosity methods"  
 ACM SIGGRAPH'87 Conference Proceedings, Vol 21 No. 4 pp 311-320, July 1987.

## [Wallace89]

John R. Wallace, Kells A. Elmquist, Eric A. Haines  
 "A Ray-Tracing Algorithm for Progressive Radiosity"  
 ACM SIGGRAPH'89 Conference Proceedings, Vol 23 No.3 pp.315-324, July 1989

## [Ward88]

Gregory J. Ward, Francis M. Rubinstein, Robert D. Clear,  
 "A Ray Tracing Solution for Diffuse Interreflection"  
 ACM SIGGRAPH'88 Conference Proceedings, Vol.22 No.4 pp.85-92, August 1988

## [Warnock69]

John E. Warnock  
 "A hidden surface algorithm for computer generated halftone pictures"  
 Reporte Técnico 4-15, Universidad de Utah, CS Dept., Junio 1969.

## [Whitted80]

Turner Whitted,  
 "An improved illumination model for shaded display"  
 CACM, vol.23 No.6, pp.343-349, June 1980