

1

Introdução

Este capítulo apresenta o problema mais geral da composição (*compositing*) em HDR e posiciona o tipo de abordagem adotado na dissertação. Depois, uma estrutura geral para a técnica de iluminação HDR é apresentada, seguida da proposta para o caso de renderização em tempo real. Uma descrição detalhada sobre imagens HDR e suas áreas de interesse também é apresentada. Por fim, apresenta-se a organização desta dissertação.

1.1

O Problema da composição em HDR

O objetivo principal da computação gráfica é a geração de imagens por computador. Imagens geradas por computador podem ser criadas através da aquisição e processamento de fotografias ou através da geração de imagens sintéticas com renderizadores. O processamento digital de fotografias permite a retirada de imperfeições e o ajuste de pequenos detalhes em fotografias. Um sem número de aplicações, como Adobe Photoshop [4], se utilizam de tais algoritmos para permitir que o consumidor médio tenha acesso a efeitos anteriormente apenas possíveis a grandes estúdios de edição de imagens.

A utilização de imagens geradas por computador se torna cada vez mais comum em diferentes áreas, tais como filmes de animação, jogos eletrônicos e sistemas de visualização geográficos. A iluminação tem um papel importante em tais aplicações. A interação da luz com os objetos que formam o mundo real, permite que se possa enxergar tais objetos e que se possa distinguir as diferentes propriedades de cada objeto. Para se corretamente simular o comportamento visualizado no mundo real, é preciso entender o complexo comportamento da luz ao interagir com um ambiente e as diferentes propriedades dos materiais na maneira como absorvem, transportam e refletem a luz incidente.

Existem duas grandes categorias de algoritmos para renderização de imagens em computador, algoritmos de renderização não tempo real (*non-real time rendering*) e algoritmos tempo-real. Cada uma dessas categorias apresenta requisitos diferenciados de qualidade de imagem e tempo total para a geração de imagens. Algoritmos não em tempo-real necessitam gerar imagens com excelente qualidade final, sem a necessidade de permitir interação em tempo real de um usuário. O tempo para a geração da imagem varia entre alguns segundos até algumas horas, dependendo da complexidade da cena. Os algoritmos em tempo real necessitam gerar imagens em um tempo muito curto, de forma a permitir a interação do usuário para processar alguma eventual mudança na cena e renderizá-la novamente.

Os modelos de luzes tipicamente modelados sinteticamente são as luzes pontuais e luzes direcionais. Esses modelos, porém, apresentam falhas ao se tentar modelar ambientes com condições de iluminação complexas, situação muito comum em cenas do mundo real.

Diferentes objetos do mundo real apresentam comportamentos díspares na forma como interagem com a luz. Alguns objetos absorvem grande parte da luz e refletem o restante em todas as direções, de forma difusa. Tais objetos possuem um comportamento predominantemente difuso. Outros objetos possuem a capacidade de refletir a energia recebida em uma particular direção praticamente com a mesma intensidade da fonte luminosa.

A representação correta de um sem número de fontes luminosas pontuais, luzes de superfície e direcionais presentes em um ambiente é complexa e computacionalmente muito cara. Em [21] é apresentado um método, baseado em imagem, que possibilita a codificação de um conjunto de fontes luminosas. Esse método, denominado mapeamento de ambiente ou mapeamento de reflexão ¹, não tenta modelar cada luz encontrada no ambiente e utiliza apenas a técnica de mapeamento de texturas para a renderização final. Esse método é muito utilizado em aplicações em tempo real, pois permite que objetos apresentem reflexos realistas, outrora apenas disponíveis para algoritmos de renderização não em tempo real como *ray-tracing* [23].

¹Rigorosamente falando, embora essas duas denominações sejam livremente usadas como equivalentes, o mapeamento de reflexão tem um significado mais específico. Diz-se que uma textura de mapa de reflexão é gerada quando as propriedades do material da superfície do objeto são usadas para modificar um mapa de ambiente [21]. Desta forma, objetos difusos podem ser renderizados. Normalmente se usam técnicas de pré filtragem, como [34] e [38]

Mapeamento de reflexão pode tanto ser usado para mapas de ambiente renderizados sinteticamente² como para fotografias omnidirecionais reais. Na linha de fotografias, esta técnica foi considerada por Lance Williams [24], em 1983, e desenvolvida independentemente por Gene Miller³ e Robert Hoffman [11] em 1984. Os trabalhos desses autores introduziram a técnica de Mapeamento de Esfera (*Sphere Mapping*), independente do mapa ser sintético ou capturado por fotografia. O mapeamento de esfera se tornou a primeira técnica de mapeamento de reflexão suportada por hardware gráfico. Nesta técnica, um mapa de reflexão fotográfico pode ser gerado fotografando-se uma esfera reflexiva (espelhada) e de dimensões reduzidas colocada no ambiente real. A imagem circular oriunda da fotografia dessa esfera reflexiva é denominada de sonda de luz (*light probe*), porque essa imagem captura a situação de iluminação no local da esfera espelhada. Rigorosamente falando, a sonda de luz é válida apenas para uma posição fixa da câmera (embora ela represente todo o ambiente) e pode apresentar problemas de singularidade para o seu uso em diferentes direções de vista. A Figura 1.1 ilustra a captura de uma mapa de reflexão por fotografia e a renderização de um objeto sintético.



Figura 1.1: Experiência pioneira de Gene Miller com mapa de reflexão fotográfico: (a) ambiente real; (b) bola reflexiva sendo fotografada; (c) cachorro renderizado com o mapa de reflexão. Extraído de G.Miller [1]

O uso mais simples do mapeamento de esfera (Figura 1.1-c) não resolve o problema geral de compor objetos sintéticos em fotografias. A questão da composição (*compositing*) é um problema computacional complexo de iluminação global com várias dificuldades, tais como:

1. Sondas de luz obtidas por câmeras normais não capturam a variação completa da iluminação incidente e, portanto, não conseguem representar o ambiente real.

²Esta técnica foi inventada por Blinn e Newell [21] em 1976

³Gene Miller fotografou uma bola de árvore de natal para capturar um ambiente real externo, em uma experiência de 1982, mas que não foi relatada oficialmente na literatura [5]

2. A iluminação indireta deve considerar ambas as reflexões especular e difusa.
3. Todos os aspectos da iluminação que variam espacialmente devem ser considerados, tais como iluminação não-uniforme, interreflexão e auto-sombreamento.
4. Os objetos sintéticos devem projetar sombras na cena.
5. O objeto sintético e/ou a cena podem ser dinâmicos.
6. A aplicação pode exigir a geração de sondas de luz em tempo real (um programa de TV ao vivo, por exemplo) e/ou a renderização em tempo real (um jogo, por exemplo), o que torna impossível considerar todos os efeitos de iluminação global

O primeiro problema acima mencionado pode ser solucionado pelo uso de câmeras HDR (High Dynamic Range) ou pela aplicação da técnica de geração de imagens HDR a partir de imagens LDR (Low Dynamic Range) proposta por Paul Debevec [52]⁴.

A Figura 1.2 mostra uma câmera HDR de escaneamento panorâmico que gera automaticamente uma sonda de luz sem o uso de esfera reflexiva, num processo de ajustes e escaneamento um tanto lento (5 a 10 minutos, dependendo das condições ambientais). O preço deste tipo de câmera é elevado (da ordem de US\$25000, com lentes).

A Figura 1.3 ilustra a técnica proposta por Debevec [52] para gerar sondas de luz HDR, que usa uma câmera normal e uma esfera reflexiva (o que requer um software de costura de imagens - stitching - para cobrir 360). A técnica de Debevec [52] consiste em fotografar uma mesma cena com várias aberturas diferentes, a partir das quais é gerado um mapa HDR. Uma galeria de sondas de luz é disponibilizada por Paul Debevec em [6]. Os formatos HDR mais comuns são o RADIANCE (.hdr) [27] e o Floating-Point TIFF (.tif). As sondas de luz (esféricas) são normalmente convertidas para a forma de ambiente de cubo em cruz (típica de mapas de ambientes cúbicos).

Paul Debevec [53] também propôs uma técnica excelente para a composição (*compositing*) de objetos sintéticos e estáticos em fotografia (não em tempo real) a partir de sondas de luz HDR, que considera os problemas 2 a 4 acima mencionados (Figura 1.4). Nesta técnica, a sonda de luz HDR

⁴A idéia de usar uma coleção de imagens que somente diferem em exposição não é nova. Isto foi primeiramente colocado por Charles Wyckoff em 1961 e posteriormente tratada por Steve Mann e Rosalind Picard no MIT Media Lab [30], em 1994



Figura 1.2: Câmera HDR de escaneamento panorâmico SpheroCamHDR (26 F-stops, resolução 10.600 pixels horizontal por 5.300 pixels vertical). Imagem e informações retiradas de www.spheron.com [2]

(HDR light probe) é colocada próxima à posição onde devem ficar os objetos sintéticos⁵.

As questões sobre ambientes dinâmicos (problema 5 acima mencionado) envolvem combinações de objetos e cenas em movimento (e.g. um objeto dinâmico em uma cena estática⁶, um objeto estático em uma cena dinâmica, ou ambos o objeto e a cena dinâmicos). No caso de aplicações não em tempo real com cenas de fundo em movimento, as sondas de luz podem ser geradas a partir de câmeras comuns de vídeo LDR, com técnicas de distorção de imagem (image warping) [39]. Normalmente, neste caso, usam-se clusters para fazer a renderização dos objetos (independente dos objetos sintéticos estarem parados ou em movimento). No caso de

⁵A esfera reflexiva (ou a câmera panorâmica HDR) pode ficar em uma posição diferente da do objeto sintético, desde que o objeto não seja muito reflexivo e/ou quando não se quer a consideração de aspectos espaciais de iluminação do tipo interreflexão e auto-sombreamento. A rigor, entretanto, um objeto mais extenso (ou um objeto pequeno em movimento) exigiria mais de um ponto de sonda de luz - o que ainda estaria muito distante da situação ideal (e ainda sem tecnologia para realizar) de se ter o volume de irradiância (que descreveria a irradiância de todos os pontos em uma sala em todas as direções)

⁶A esfera reflexiva (ou a câmera panorâmica HDR) pode ficar em uma posição diferente da do objeto sintético, desde que o objeto não seja muito reflexivo e/ou quando não se quer a consideração de aspectos espaciais de iluminação do tipo interreflexão e auto-sombreamento. A rigor, entretanto, um objeto mais extenso (ou um objeto pequeno em movimento) exigiria mais de um ponto de sonda de luz - o que ainda estaria muito distante da situação ideal (e ainda sem tecnologia para realizar) de se ter o volume de irradiância (que descreveria a irradiância de todos os pontos em uma sala em todas as direções).

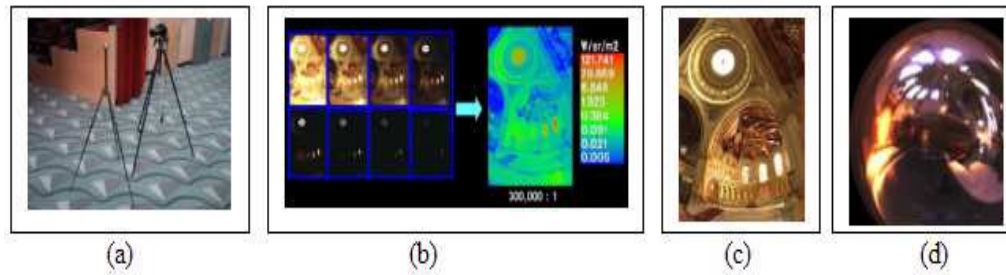


Figura 1.3: Técnica de geração de sonda de luz de Debevec [53]. (a) câmera comum com esfera reflexiva (bolas industriais cromadas de 5 cm ou bolas ornamentais de jardim de 15-30cm, encontradas no varejo); (b) geração de mapa de radiância a partir de fotografias com aberturas diferentes da Grace Cathedral, São Francisco (c); (d) sonda de luz da Grace Cathedral, com variação dinâmica de 200000:1. Material extraído de Paul Debevec [3]



Figura 1.4: Composição de objetos sintéticos em fotografias usando sondas de luz HDR. (a) modelo geométrico do objeto sintético; (b) iluminação do objeto com fundo de floresta; (c) iluminação do objeto com fundo de museu . As sondas usadas estão indicadas no canto superior esquerdo. Material extraído de [3].

cenas de fundo dinâmicas em tempo real (problema 6 acima mencionado), há a necessidade de geração de sondas de luz HDR em tempo real (real-time lighprobe). Estas sondas de luz podem ser obtidas diretamente a partir de sensores de vídeo HDR usados em vigilância (e.g. HDRC da IMS CHIPS, D2500 da Digital Pixel System e Autobrite da SMal Camera Technologies)⁷. Uma outra solução pode ser o uso de câmeras NTSC de alta sensibilidade com uma lente fish-eye (e.g. CV-S3300 da JAI Corp., www.jai.com, com o conversor fisheye FC-E8 da Nikon). Neste caso,

⁷Também podem ser usadas as câmeras cinematográficas de alta definição que entregam imagens cruas e sem compressão (e.g. Viper FilmStream da Thomson/Grass Valley). Entretanto, não há referência sobre isto na literatura (provavelmente devido ao custo proibitivo destas câmeras). Já o uso de sensores de vídeo HDR tem sido mencionado na literatura [46]. Deve-se, entretanto, sempre verificar se o sensor tem capacidade de cor (visto que a maioria dos sensores usados em vigilância é monocromático)

independente dos objetos sintéticos estarem parados ou em movimento, a renderização tem sempre que ser feita em tempo real. Esta renderização pode ser feita por clusters e/ou programação de pixel shading da placa gráfica. Quando um dos dois elementos (objeto sintético ou cena de fundo) é dinâmico em tempo real, tem-se a situação de composição em tempo real⁸, que também é denominada de Realidade Misturada Interativa (Interactive Mixed Reality). A pior situação da composição em tempo real é quando ambos o objeto e a cena são dinâmicos em tempo real. Não há referências na literatura com soluções para esta situação, pelo menos até onde o autor desta dissertação conhece. Na literatura, há alguns trabalhos de composição em tempo real para cenas dinâmicas e objetos estáticos. Atualmente, um dos melhores resultados é o de Pomi et al. [40] que usa sondas de luz HDR em tempo real e um algoritmo de iluminação global em tempo real (raytracing, no caso), mas que apresenta baixas taxas de quadros por segundo (2.9 fps, para um objeto estático de 200000 triângulos, usando um cluster de renderização de 24 nós dual AMD Athlon MP 1800+). O trabalho de Havran et al. [46] captura melhor as sondas de luz HDR em tempo-real, mas não trata diretamente do problema de composição, ficando apenas na questão de reiluminação (*relighting*)⁹ de objetos estáticos sem interação com outros objetos em cena. Os trabalhos sobre composição em tempo real acima citados, apesar da ótima qualidade de imagem [40, 46] tendem a ter desempenho muito baixo em termos de fps. Desta maneira, a presente dissertação procura investigar soluções para a composição em tempo real com o foco em taxas elevadas de quadros por segundo e na programação das placas gráficas. Nesta dissertação, entretanto, o autor considera a cena de fundo estática e analisa apenas a situação onde os objetos sintéticos são interativos (típicos de jogos 3D). Em outras palavras, a presente dissertação trata apenas da renderização em tempo real usando mapas HDR. Para garantir taxas elevadas de fps, o trabalho desta dissertação não emprega técnicas de iluminação global. Procurando buscar uma visão mais organizada da questão de composição, o presente autor propõe que os métodos de composição em tempo real sejam classificados em uma das duas classes a seguir:

- Métodos no domínio espacial (angular). A adaptação da técnica de Paul Debevec [41] para tempo real é uma abordagem importante nesta linha. Entretanto esta adaptação não está oficialmente descrita

⁸Não confundir com a composição em tempo real oferecida pelos sistemas de finalização e efeitos especiais (e.g. Avid)

⁹Reiluminação trata do uso de uma nova situação de iluminação de um objeto ou pessoa

na literatura, embora haja referências a este tipo de desenvolvimento [38].

- Métodos no domínio da frequência. Os principais trabalhos são os baseados em harmônicos esféricos [34] e os baseados em wavelets não-lineares [40]. Estes métodos requerem etapas pesadas de pré-computação. Os métodos baseados em wavelets não-lineares apresentam melhores resultados do que os baseados em harmônicos esféricos. Estes resultados melhores se referem a sombras (incluindo auto-sombreamento) e outros efeitos de transporte de luz de todas as frequências (*caustics*, por exemplo). Entretanto, [40] não trata diretamente do problema geral de composição em tempo-real e ainda tem vários pontos em aberto

A presente dissertação trata de métodos no domínio espacial (angular). Mais especificamente, o presente trabalho tenta adaptar a técnica em [53] para o caso de renderização em tempo real com mapas HDR.

1.2

Iluminação Baseada em Imagem

As técnicas de renderização baseadas em mapas HDR é conhecida como IBI - Iluminação Baseada em Imagens (IBL - Image-Based Lighting) ou simplesmente Iluminação HDR (HDR Lighting). Uma receita desta técnica pode ser encontrada em [7]. É importante observar que Iluminação Baseada em Imagens trata de modelos baseados em medidas de irradiância e é conceitualmente diferente de Modelagem Baseada em Imagem (Image-Based Modeling). Na presente dissertação, o autor propõe a seguinte estrutura para enquadrar a sua proposta para renderização de objetos sintéticos com mapas HDR: a hipótese simplificadora de 3 grupos de objetos (como proposta em [53]) e um processo geral para Iluminação HDR. A hipótese simplificadora considera os seguintes 3 grupos de objetos:

1. A cena distante (que irradia luz na direção dos outros grupos, mas que ignora a luz refletida de volta).
2. A cena local (com geometria conhecida e características estimadas de refletância).
3. Objetos sintéticos (com geometria e características de refletância completamente conhecidas) que devem interagir com a cena local.

O processo geral da técnica de Iluminação HDR pode ser esquematizado conforme a Figura 1.5, onde o Módulo de Iluminação gera a imagem global I_G que é composta com a imagem de fundo I_{fundo} , gerando a imagem final I_{HDR} . A imagem final I_{HDR} ou é exibida diretamente em um monitor HDR ou sofre uma conversão para LDR, por algoritmos de Mapeamento de Tons (Tone Mapping), para que possa ser exibida em monitores usuais¹⁰. A imagem de fundo I_{fundo} é a fotografia HDR do fundo (cena local + cena distante).

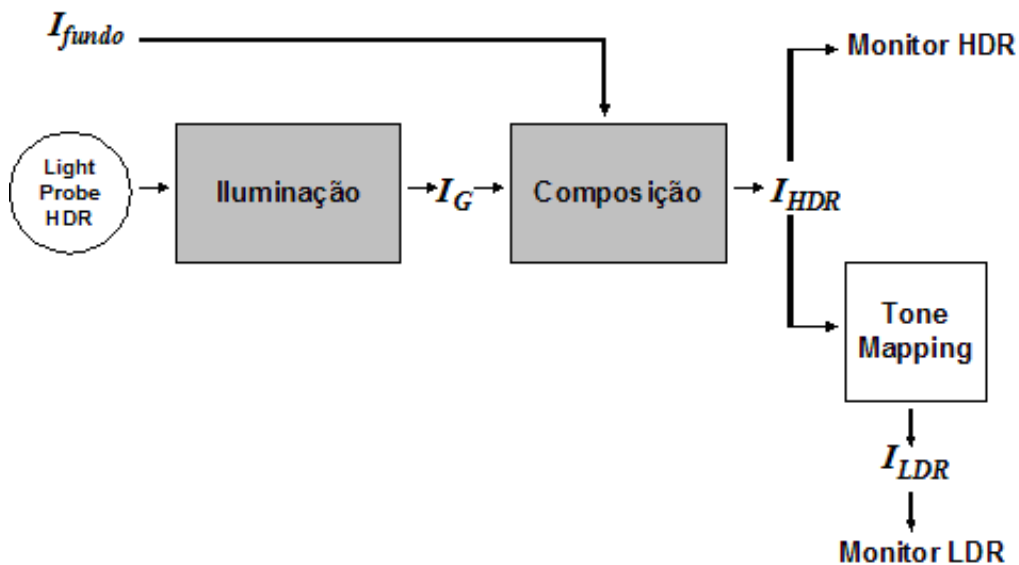


Figura 1.5: Processo geral da técnica de Iluminação Baseada em Imagem

No caso de renderização não em tempo real descrito em [53], o Módulo de Iluminação da Figura 1.5 considera os seguintes passos:

- Constrói um modelo geométrico grosseiro da cena distante. A versão mais simples é a de uma caixa envolvendo o ambiente. Esta cena é, então, mapeada em textura HDR a partir da sonda de luz HDR (Figura 1.6): T_{HDR}
- Gera o modelo aproximado da cena local, estimando as características de refletância¹¹
- Um algoritmo de iluminação global¹² é aplicado aos objetos sintéticos e à cena local, tendo a textura T_{HDR} como iluminação. Nesta

¹⁰Os monitores trabalham com 8 bits por canal de cor, o que não é suficiente para capturar a variação completa da iluminação incidente

¹¹Um método para estimar as características de refletância da cena local pode ser encontrado em [53], item 7

¹²RADIANCE [32] e Mental Ray [8] são alguns dos sistemas disponíveis

renderização, a cena distante não considera a luz refletida de volta (no sistema RADIANCE [32], este comportamento exclusivamente emissor pode ser especificado com a propriedade de material "glow"). A saída desta renderização é a imagem I_G

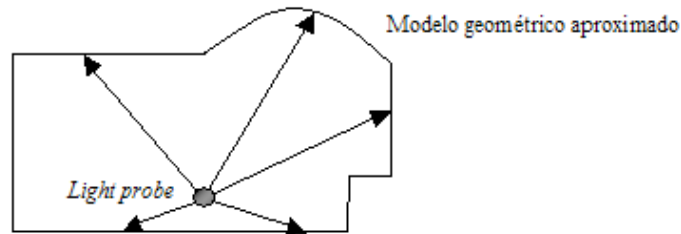


Figura 1.6: Geração da textura HDR da cena distante para ser usada na solução de iluminação global

Ainda neste caso [53], a composição de I_G com I_{fundo} é feita com uma técnica de renderização diferencial. No caso de renderização em tempo real proposto nesta dissertação, há apenas o grupo de objetos sintéticos e a cena distante. Neste caso, o Módulo de Iluminação compreende os seguintes passos:

- A cena distante é diretamente representada pela sonda de luz HDR no formato de ambiente de cubo em cruz (como já costuma estar disponível em galerias públicas de sondas de luz [6]). Esta textura (THDR) pode ser corrigida para o caso de objetos sintéticos difusos. Na presente dissertação, no caso de objetos difusos, T_{HDR} é modificada por um filtro de convolução inspirado em [27]
- O objeto sintético é renderizado com T_{HDR} e a imagem resultante I_G é armazenada no render target da placa. No caso de objetos difusos, o ponto do mapa T_{HDR} é definido diretamente pelo prolongamento da normal (a posição da câmera não importa). Ainda neste passo, é feita uma filtragem da textura (usando pixel shading) para se ter uma qualidade melhor da imagem final¹³.

Quanto ao Módulo de Composição, têm-se os seguintes passos:

- Gera um render target menor (1/4) para facilitar as operações de efeitos especiais e composição final: I_{Gmenor} ;

¹³A consideração aproximada de sombras pode ser feita neste passo. Uma proposta é apresentada no Capítulo 5 (Oclusão de Ambiente)

- Encontra as partes mais iluminadas da cena, usando a luminância dos pixels. Esta operação é direta, no caso de placas gráficas recentes que têm render target com maior precisão
- Na imagem mascarada, aplicam-se os filtros para efeitos especiais (tais como glare e halo). É importante observar que estas filtragens são dependentes da aplicação e não há como as placas incorporarem. A imagem resultante é I_{filtro} .
- As seguintes imagens são somadas: I_G e I_{filtro} escalada para o tamanho original do *render target*

Por fim, o Módulo de Mapeamento de Tons é aplicado, ou feito diretamente pela placa de acordo com o formato de *render target* utilizado [9].

1.3 Imagens High Dynamic Range

Dynamic Range representa a razão entre o valor de brilho mais intenso e o menos intenso encontrados em um ambiente. Ambientes do mundo real apresentam uma grande variação de intensidade entre áreas muito iluminadas e pouco iluminadas. Como exemplo a iluminação proveniente do Sol ao meio dia pode ser 100 milhões de vezes maior que a iluminação das estrelas à noite.

Ambiente	Iluminação
Estrelas	10^{-3}
Lua	10^{-1}
Ambiente interno iluminado	10^2
Sol	10^5

Tabela 1.1: Luminância para diferentes ambientes de iluminação.

Atualmente, diferentes aplicações que se utilizam de imagens digitais estão disponíveis. Tais imagens podem ser criadas através do processo convencional de escaneamento de fotografias, máquinas fotográficas digitais, ou geradas por computador. A busca por criação de imagens com maior resolução, impulsionada pelo aparecimento e barateamento das máquinas fotográficas digitais (máquinas com cinco megapixels são facilmente disponíveis para o consumidor comum), é um passo significativo na geração de imagens com melhor qualidade.

Apesar do aumento de qualidade com imagens de maior resolução, uma grande mudança na abordagem em relação a imagens digitais está ocorrendo

nos dias de hoje. Tal mudança diz respeito à representação dos pixels em uma imagem e o *Dynamic Range* suportado. Os dispositivos tradicionais de aquisição de imagens (filmes fotográficos ou *charge coupled devices* (CCD) em máquinas digitais) possuem domínio de valores menos abrangente, geralmente de apenas oito bits por canal (256 valores de intensidade). Com tal limitação os valores de brilho dos pixels armazenados são saturados em 255, mesmo que possuam intensidade maior. Outro problema mais notável decorre do fato que, devido a menor precisão, os valores dos pixels são aproximados para valores inteiros, de forma que pequenas variações de intensidade são perdidas. Em geral, imagens tradicionais geralmente apresentam regiões pouco expostas, onde os pixels aparecem escuros e com pouca variação entre si, ou regiões super expostas, onde os pixels com maior radiância tem valor saturado.

Tal problema é comum entre fotógrafos, que muitas vezes necessitam decidir qual é a área mais relevante em uma fotografia e ajustar a exposição para que uma particular região seja bem amostrada e os resultados sejam satisfatórios. Com tal procedimento, porém, outras áreas sofrem de sub ou super exposição.

É importante salientar que o *Dynamic Range* apresentado nas imagens não é o mesmo que a quantidade de cores que um dispositivo pode exibir. Com oito bits para cada canal de cor R, G, B é possível representar 2^{32} cores. Muitos fabricantes de dispositivos de visualização e softwares para edição de imagens se referem a tal termo como "Milhões de cores". Apesar de parecer um numero impressionante, 255 valores para cada canal de cor é insuficiente para correta representação de muitas cenas.

Imagens *High Dynamic Range* permitem que diferentes áreas da imagem, com diferentes níveis de exposição sejam amostradas corretamente. Dessa forma, não é necessário a escolha de uma área específica para uma melhor amostragem.

As Figuras 1.7 e 1.8 mostram uma imagem *Low Dynamic Range* e *High Dynamic Range* respectivamente. A versão HDR consegue representar com detalhes tanto regiões claras quanto escuras. Tal imagem foi capturada, processada e armazenada utilizando técnicas que serão descritas no Capítulo 2. Na imagem tradicional, é facilmente perceptível que as regiões mais escuras foram sub-expostas. É perceptível na imagem HDR que tanto a região da janela quanto o interior do quarto podem ser observados em detalhe.

De forma semelhante, imagens HDR são mais apropriadas para imagens com condições de iluminação adversas, mesmo não possuindo áreas



Figura 1.7: Imagem convencional.

com grande variação entre si. A imagem dois ilustra tal fato: As fotografias foram tiradas em um ambiente aberto e com pouca variação entre as áreas da imagem, porém devido ao ambiente escuro a imagem LDR perde detalhes e aparenta ser muito mais escura que a versão HDR.

A informação codificada em imagens HDR corresponde aos valores de luminância observados na cena. Tal fato difere da abordagem tradicional utilizada em fotografia com filme ou digital, onde cada pixel representa apenas uma cor a ser exibida em um dispositivo de visualização. Devido a esse fato, imagens HDR são conhecidas como *Scene-refered* enquanto imagens tradicionais são conhecidas como *Output-refered*.

Para apresentar essa característica, imagens LDR sofrem uma série de mapeamentos não lineares para que sejam exibidas de forma mais agradável ao olho humano. Tais mapeamentos não lineares aliados à falta de precisão na representação dos valores impedem que essas imagens sejam representações fiéis das condições de iluminação em uma cena.

Na outra ponta do problema, se encontra a exibição de imagens HDR. A visão humana é capaz de se adaptar a diferentes condições de iluminação, cerca de 10 ordens de magnitude [33]. Ao olhar para uma cena, o olho humano é capaz de distinguir cerca de cinco ordens de magnitude



Figura 1.8: Imagem HDR.

simultaneamente. Dispositivos tradicionais de exibição, os monitores de tubos de raios catódicos (CRT) são capazes de reproduzir duas ordens de magnitude para variação de intensidade luminosa. Tal limitação está ligada ao fato de que o fósforo não pode ser excitado acima de um determinado limite. Por esta razão, conversores digital-analógico de 8 bits são suficientes para a geração dos sinais analógicos formadores da imagem. De forma semelhante, os monitores de cristal líquido (LCD) sofrem dessa limitação, porém limitados pela intensidade da luz emissora (*backlight*), e apesar de serem mais brilhantes que monitores CRT, não o são por ordens de magnitude.

Apesar da recente introdução de dispositivos de visualização HDR [45, 41], estes ainda são muito caros e não são acessíveis ao consumidor médio. Imagens HDR necessitam de alguma forma de processamento para exibição em dispositivos com apenas oito bits de representação. Tipicamente imagens HDR utilizam uma forma própria para representação correta de seus valores (utilizando representação em ponto flutuante para cada pixel). Existem diferentes algoritmos que possuem como função mapear os valores HDR armazenados em representações que os dispositivos de visualização são capazes de exibir. Esses algoritmos são chamados Tone-Mapping Operators.

Mesmo com a advento e popularização de dispositivos de exibição *High Dynamic Range*, a redução de *Dynamic Range* é uma tarefa importante,

principalmente em relação à mídia impressa. É fato que a aquisição de imagens HDR e seu posterior processamento apresenta resultados melhores que imagens tradicionais, mesmo em dispositivos de exibição LDR. A mídia impressa em papel nunca será HDR, pois isso implicaria a criação de um tipo especial de papel capaz de emitir luz. Dessa forma, o processamento de imagens HDR para exibição em papel será sempre necessário. Imagens *High Dynamic Range* são um dos principais focos de pesquisa em computação gráfica atualmente. Diversas são as vantagens apresentadas por tais imagens em contraste com imagens tradicionais. A maior precisão utilizada para armazenar a intensidade de cada pixel proporciona cenários virtuais renderizados com maior qualidade. O fato de cada pixel possuir valor diretamente proporcional a radiância na cena, permite a utilização de iluminação baseada em imagens com alta precisão, entre outras possibilidades.

As principais áreas de interesse relacionadas com imagens *high dynamic range* são:

– ***Aquisição e geração de imagens HDR***

Diz respeito ao processo fundamental de aquisição de imagens HDR. Existem duas maneiras fundamentais para tal: A geração através de imagens de computador, utilizando-se de algoritmos de renderização e outras técnicas em computação gráfica. Um dos exemplos mais famosos de geração de imagens HDR são as imagens geradas pelo renderizador Radiance [32]. A segunda maneira se utiliza da aquisição de fotografias utilizando-se de câmeras tradicionais (LDR). Em [52], é apresentado um método para criação de mapas de radiância (imagens onde cada pixel é uma representação da radiância original da cena). Esse trabalho foi um dos responsáveis pela nova onda de pesquisas em imagens HDR. O artigo se baseia na construção de mapas de radiância através da recuperação da curva de resposta característica dos dispositivos de captura de imagens.

A técnica se baseia na retirada de diferentes fotografias de cenas estáticas, com uma variação no tempo de exposição para cada fotografia. A partir daí é possível combinar esse conjunto de imagens, recuperar a curva de resposta e gerar uma imagem HDR. Essa técnica é explicada em detalhes no Capítulo 2. O surgimento de dispositivos capazes de capturar diretamente dados HDR está mudando a forma como a aquisição dessas é feita. Atualmente já existem alguns modelos de câmeras HDR[2] no mercado, inclusive para vídeo, como

a FilmStream Viper. Apesar de serem dispositivos impressionantes e gerarem imagens com altíssima qualidade, a maioria desse dispositivos produz dados brutos de imagens a uma taxa impressionante, capazes de acabar rapidamente com o espaço disponível em discos rígidos.

– ***Representação de imagens HDR***

Uma vez que os dados HDR são adquiridos, é necessário armazená-los de alguma forma. Atualmente existem diferentes formatos para imagens HDR. Tal variação é dependente do propósito do formato e das questões de projeto levadas em consideração. As principais questões no projeto de formatos para arquivos de imagens HDR são: O tamanho final dos arquivos resultantes, o *dynamic range* capaz de ser representado, e a precisão utilizada para representar intervalos de valores sucessivos. Os formatos HDR mais utilizados são apresentados no Capítulo 2.

– ***Dispositivos de exibição HDR***

Como dito anteriormente, os monitores padrões atuais servem bem ao propósito de exibir imagens com oito bits para cada canal de cor. Porém, são incapazes de exibir imagens HDR corretamente. Dispositivos HDR serão de fundamental importância na aceitação geral da tecnologia HDR por parte do mercado consumidor. Imagens HDR exibidas em dispositivos HDR apresentam resultado visual superior a imagens HDR processadas e exibidas em dispositivos convencionais. As propostas atuais para dispositivos HDR se baseiam em retro-projetores, semelhantes a projetores de slides, ou tecnologias que substituem os *backlights* encontrados em monitores LCD [45].

– ***Processamento de imagens***

Processamento de imagens é um grande beneficiado de técnicas baseadas em imagens HDR. Um exemplo de tarefa grandemente favorecida pela utilização de imagens HDR é a correção do balanço de regiões brancas. Devido ao fato de vários pixels estarem superexpostos e com seu real valor saturado em 255. Imagens HDR também tornam menos importantes os processos de iluminar cuidadosamente a cena e capturam corretamente cores altamente saturadas.

Tarefas mais cotidianas como ajuste de cor, contraste e brilho também são favorecidos com dados HDR. Tais operações podem realizar escalas não lineares entre os valores de pixels originais, de forma que algumas regiões do domínio necessitem de uma precisão maior que as disponíveis em 8 bits por pixel.

– ***Reprodução tonal***

Existem diferentes algoritmos capazes de mapear valores de pixels em imagens HDR para dispositivos LDR. Tais algoritmos podem ser agrupados em diferentes classes de acordo com seu propósito final. Alguns se preocupam na representação fiel do valores de luminância da imagens, outros se preocupam com o aspecto final da imagem exibida no dispositivo LDR.

– ***Renderização de cenários virtuais***

A Renderização de cenários virtuais também sofre com o problema da falta de precisão utilizados para o *colorbuffer*, ou seja, a área de memória reservada ao armazenamento dos valores de cor de cada pixel do *viewport* da cena renderizada.

A arquitetura das placas gráficas modernas satura os valores de cada canal de cor em $[0,1]$. A representação interna dessas cores na placa gráfica é um valor em ponto flutuante correspondente ao domínio $[0,255]$ padrão. Com essa limitação, os cálculos de iluminação perdem precisão em áreas com maior iluminação. O problema é análogo ao encontrado com as imagens digitais. Caso um pixel tenha valor 156.0 (em representação de ponto flutuante) e outro 2.0, ambos serão saturados no valor máximo um. Fica claro que, os pixels do exemplo apresentam uma grande variação de luminância e a saturação de seus valores é um erro considerável.

Ao utilizar superfícies de renderização com formatos que não forcem a saturação de valores, os cálculos de iluminação seriam armazenados com precisão suficiente para sua correta utilização. Após a renderização, processos de mapeamento do domínio de valores da superfície original de renderização para valores do domínio do *colorbuffer* permitem que grande parte das nuances de iluminação da cena sejam mantidos. O pipeline de renderização HDR e os efeitos de pós-processamento mais utilizados são detalhadamente explicados no capítulo 3.

– ***Image Based Lighting***

Image Based Lighting é uma das aplicações mais interessantes de imagens HDR. A computação gráfica se preocupa geralmente com a criação de imagens através da simulação da interação entre uma cena e as luzes que o iluminam [31]. Geralmente primitivas geométricas como triângulos, polígonos e curvas são utilizadas para modelar os

objetos de uma cena. Associados a tais objetos estão os materiais que descrevem o comportamento da interação desses objetos com as diferentes luzes. A última parte fundamental para geração das imagens é a definição das fontes de luz que irão iluminar a cena. Ao associar esses componentes em um algoritmo de renderização que simula a iluminação e produz a imagem final. Imagens são tradicionalmente utilizadas no processo[19] no qual texturas descrevem como as superfícies variam espacialmente. Essas imagens simulam bem o comportamento de materiais difusos, porém materiais que refletem a luz especularmente, possuem *highlights* com praticamente a mesma luminância da fonte de luz. Esse é o caso para superfícies emissivas. Para corretamente simular a luminância proveniente de ambientes codificados em imagens, imagens HDR são necessárias. Geralmente a iluminação codificada nessas imagens é extremamente complexa e difícil de ser atingida com fontes de iluminação artificiais. O Capítulo 4 descreve em detalhes a utilização de mapas de iluminação HDR para renderização de imagens.

Apesar do conceito de imagens HDR ser inerentemente simples, a representação de cenas com os valores de luminância compatíveis com os encontrados no mundo real, a implicação da pesquisa em HDR são profundas e representa um novo paradigma na aquisição, representação e exibição de imagens e sua utilização em computação gráfica.

1.4

Objetivos e Organização da Dissertação

Esta dissertação tem o objetivo de estudar a área de Iluminação Baseada em Imagem e de propor um pipeline de renderização em tempo real para jogos 3D, no caso simples de cenas estáticas, adaptando a técnica não em tempo real apresentada por Paul Debevec em 1998 [53].

A organização da dissertação é como se segue. No Capítulo 1, os conceitos de imagens HDR, composição e iluminação baseada em imagens são apresentados em detalhes. O segundo capítulo se dedica aos conceitos de radiometria e fotometria, fundamentais para o entendimento dos métodos apresentados nos capítulos seguintes. No Capítulo 3 são apresentados os trabalhos anteriores que motivaram essa dissertação, especialmente a criação de imagens HDR através de um conjunto de imagens tradicionais. Neste capítulo também são apresentadas as formas tradicionais de codificação de imagens HDR em arquivos utilizados pelas aplicações.

No Capítulo 4 é apresentado o *pipeline* para renderização de imagens HDR em tempo real e uma técnica para utilização de mapas de iluminação HDR. As técnicas de Iluminação Baseada em Imagem e os conceitos de iluminação de objetos sintéticos são apresentados no Capítulo 5. As conclusões e trabalhos futuros são apresentados no Capítulo 6.