

1

Introdução

Esta dissertação é sobre a geração de imagens realistas, então, antes de levantarmos as questões que afetam o grau de realismo de uma imagem, precisamos definir o que é *fotorrealismo*.

Freqüentemente, em um espírito semelhante ao teste de Turing para inteligência artificial (Turing 50), definimos uma imagem como sendo fotorrealista se um observador humano não for capaz de dizer se a imagem foi gerada no computador ou se foi capturada com uma câmera do mundo “real”. Embora razoável, essa definição restringe a classe das imagens que gostaríamos de poder chamar de fotorrealistas. Por exemplo, imagens de um dinossauro extinto há 65 milhões de anos perseguindo humanos nos dias de hoje certamente não passariam nesse teste, embora a maioria das pessoas considere extremamente realistas as imagens dos dinossauros nos filmes de Spielberg. Uma tentativa de definição que englobasse situações como essa certamente nos levaria para bem longe de nossos objetivos. Então vamos deixar em aberto essa noção mais ampla de fotorrealismo e manter o sentido mais restrito presente em nossa definição original para servir como guia em nossa busca pela geração de imagens realistas.

Com um problema definido podemos seguir em frente tentando resolvê-lo. A computação gráfica certamente não é a primeira a enfrentar esse problema. Durante a Renascença, artistas, ao que parece começando com Brunelleschi, descobriram as leis matemáticas da perspectiva que descrevem como os objetos parecem diminuir de tamanho à medida que se distanciam de nós. Aplicando essas regras na realização de suas pinturas, conseguiram criar a ilusão de profundidade com uma exatidão nunca antes vista. Artistas renascentistas exploravam os mistérios da natureza como, acima de tudo, um meio de adquirir conhecimento sobre o mundo visível que mais tarde poderia ser usado em sua arte.

Em nosso caso, podemos adotar uma atitude semelhante. Podemos estudar como acontece a formação de imagens (como uma fotografia) no mundo real para, em seguida, tentar sistematizar esse conhecimento e simular o processo no computador. Por mundo “real”, questões metafísicas à parte,

queremos dizer a descrição do mundo segundo teorias e leis científicas. Em particular, estamos interessados em teorias que expliquem a nossa capacidade de ver o mundo que nos cerca, o que nos leva naturalmente a teorias que descrevem a natureza da luz.

1.1

A natureza da luz

Os filósofos gregos já buscavam explicações para o sentido da visão. Empédocles, que viveu aproximadamente entre 490 e 430 A.C., especulou que a visão era resultado de emanações vindas de uma chama interna aos nossos olhos, criações da divina Afrodite. Como isso não explicava porque não somos capazes de ver à noite, postulou que essas emanações iam também na outra direção, dos objetos para nossos olhos, e a visão era resultado de uma interação das duas. Em um dos diálogos de Platão, *Mênon*, vemos Sócrates ensinando as idéias de Empédocles de que há certas efluências das coisas e poros de vários tamanhos correspondentes aos tamanhos das efluências, e que cor é a efluência das formas correspondendo aos tamanhos dos poros dos olhos. Os demais sentidos também podiam ser explicados dessa forma.

Com a decadência das civilizações grega e romana, o Ocidente viu poucos progressos na ciência durante os séculos que se seguiram. Nesse período, as maiores conquistas científicas vieram do mundo árabe. Nascido por volta de 965 D.C., o cientista¹ persa Ibn al-Haytham (Alhazen pela versão latinizada de seu nome) fez várias descobertas importantes no campo da ótica. Entre as suas inúmeras investigações, que incluem a formulação completa da lei da reflexão, conhecida em parte desde os tempos de Euclides, estudos sobre espelhos esféricos, cilíndricos e cônicos, e a menção do uso de uma câmera escura, devemos destacar a primeira explicação correta para a visão. Concluiu que a visão era resultado da luz, para ele minúsculas partículas viajando em linha reta, originada de uma fonte, como o Sol ou uma fogueira, refletindo nos objetos e atingindo nossos olhos.

O próximo passo significativo viria somente no século XVII, um período fascinante para a ciência em geral e, em particular, para a ótica. Foi a invenção do telescópio no início do século que permitiu que Galileu, aprimorando a construção do instrumento, descobrisse que havia outros corpos que, como pequenos planetas, orbitavam Júpiter. Como imaginara Copérnico e contradizendo o modelo Ptolomaico amplamente aceito na época, a Terra não era o centro do universo.

¹E aqui podemos usar o sentido moderno da palavra, já que seu trabalho mais importante, *Kitab al-Manazir*, sobre ótica, deixa claro que suas investigações serão baseadas em evidências experimentais em vez de somente em teorias abstratas.

Johannes Kepler, que estudara os trabalhos de Alhazen, além de seu trabalho mais famoso sobre o movimento planetário, fez também grandes contribuições para a ótica no início do século, incluindo as primeiras explicações corretas para o funcionamento da câmera escura e do olho humano, onde uma imagem invertida era formada na retina. Também descobriu a reflexão total interna e a aproximação para pequenos ângulos da lei da refração.

Lei da refração que já havia sido descoberta em sua forma completa anos antes pelo inglês Thomas Harriot. Entretanto, Harriot não publicou sua descoberta. Em 1621, 20 anos depois, Willebrord Snell redescobriu a lei que, ironicamente, leva hoje o seu nome, apesar de só ter sido publicada pela primeira vez por René Descartes, em *La Dioptrique*, incluída em seu *Discurso sobre o Método*, de 1637. Descartes começa *La Dioptrique* com o discurso *De la lumière*, onde propõe uma explicação para a natureza da luz em cima da qual baseia a lei da refração, exposta no discurso seguinte. Descartes, que se recusava a aceitar a existência do vácuo, descreve a luz como nada mais que um certo movimento que passa para nossos olhos através do ar e outros corpos transparentes, assim como a resistência de uma bengala é transmitida instantaneamente para as mãos de um cego ao atingir um objeto. Mais tarde, descartando a explicação de Descartes, Pierre de Fermat derivou a lei da refração partindo de seu *princípio do tempo mínimo*, segundo o qual a luz move-se de um ponto a outro percorrendo o caminho onde leva menos tempo.

O século XVII também veria o surgimento de um longo conflito entre duas teorias que tentavam explicar o comportamento da luz: a teoria corpuscular e a ondulatória. No mesmo ano da morte de Galileu, em 1642, nasceu Isaac Newton. Segundo Newton, a luz era composta de pequenas partículas e, com seus famosos experimentos sobre dispersão, concluiu que a luz branca não era pura, como se acreditava na época, mas composta pela combinação das diferentes cores.

Ao mesmo tempo surgia outra explicação para a natureza da luz. Efeitos do fenômeno da difração foram observados pela primeira vez por Grimaldi e, mais tarde, pelo inglês Robert Hooke. Hooke tentou explicá-los propondo que a luz era um movimento vibratório rápido do meio onde se propagava com altíssima velocidade e que cada pulso de vibração de um corpo luminoso gerava uma esfera. Nascia a teoria ondulatória da luz. No continente, Christiann Huygens foi capaz de derivar as leis da reflexão e da refração com base nessa teoria. Embora a teoria corpuscular tivesse dificuldades em explicar a difração, talvez pela falta de evidências experimentais mais contundentes, ou, mais provavelmente, pela reputação de Newton, a teoria ondulatória foi negligenciada por grande parte dos cientista no século XVIII, com exceção

notável do matemático Leonhard Euler.

O modelo ondulatório só voltaria a receber atenção no início do século XIX, quando, perante os membros da Royal Society, Thomas Young demonstrou evidências do fenômeno da interferência que, na época, não podiam ser explicados pelo modelo corpuscular. Desconhecendo os trabalhos de Young, Augustin Jean Fresnel, na França, retomou o interesse pela teoria ondulatória, fornecendo um tratamento matemático rigoroso para explicar o fenômeno da difração. Fresnel também foi capaz de explicar a propagação retilínea da luz com base nesse modelo, falha que tinha sido uma das razões mais fortes de Newton para recusar a teoria ondulatória. O golpe de misericórdia na teoria corpuscular veio em 1850, quando, contradizendo as previsões da teoria de Newton, Jean Foucault conduziu experimentos que o levaram a concluir que a velocidade da luz nos líquidos era menor que no ar. Os últimos seguidores de Newton tiveram que ceder.

Mas a consagração final da teoria ondulatória ainda estava por vir. No século XIX, em paralelo à ótica se desenvolvia a eletricidade e o magnetismo. James Clerk Maxwell, em 1864, estendeu e resumiu em um conjunto de equações todo o conhecimento empírico que havia sido acumulado sobre eletromagnetismo. Partindo dessas equações, de maneira puramente teórica foi capaz de prever que o campo eletromagnético poderia se propagar na forma de ondas transversais. A velocidade de propagação era expressa por constantes que dependiam somente de propriedades elétricas e magnéticas do meio. Substituindo os valores conhecidos na época na expressão para a velocidade da onda chegou a um resultado surpreendentemente próximo à velocidade da luz. Então concluiu que, como a velocidade era tão próxima à da luz, tínhamos fortes razões para crer que a luz era uma onda desse tipo. A verificação experimental da previsão de Maxwell viria anos mais tarde quando, em 1888, o físico alemão Heinrich Rudolf Hertz foi capaz de gerar e detectar ondas eletromagnéticas.

No entanto, a teoria eletromagnética da luz ainda encontrava algumas dificuldades. A principal era uma aparente contradição entre a previsão da teoria de Maxwell de que a luz deveria se deslocar em uma velocidade fixa com o princípio presente na mecânica de Galileu e Newton de que não existem referenciais inerciais privilegiados para descrevermos as leis da natureza. Em 1905, com a introdução de sua teoria especial da relatividade, Einstein foi capaz de fornecer uma explicação para essa aparente contradição abandonando a idéia de tempo absoluto herdada da mecânica de Newton.

Uma segunda dificuldade da teoria eletromagnética da luz era em explicar a emissão de elétrons por um metal quando exposto à luz. Esse fenômeno,

o chamado efeito fotoelétrico, fora observado pela primeira vez também por Hertz. Uma explicação para o efeito fotoelétrico viria também de Einstein, em 1905, mesmo ano em que propôs sua teoria especial da relatividade. Baseando sua explicação no conceito de quantização introduzido por Planck para explicar a distribuição da radiação emitida por um corpo negro, Einstein propôs que, de certo modo, tínhamos que retomar a idéia de Newton de que a luz se comportava como um feixe de partículas, chamadas fótons, e que a cada fóton estava associada uma energia $E = hf$, onde h é a constante de Planck e f a frequência da radiação.

Essa natureza dual da luz, que ora se comportava como uma onda e ora como uma corrente de fótons, foi formulada apropriadamente nos anos que seguiram, com o desenvolvimento da física quântica, mais notavelmente, por físicos como Bohr, Heisenberg, Schrödinger, Dirac e Feynman.

1.2

Modelos para a luz

O histórico que acabamos de fazer nos fornece basicamente 3 modelos para representar a luz: a ótica geométrica, o modelo ondulatório e a ótica quântica. Embora o modelo quântico seja o mais geral e com apenas ele sejamos capazes de explicar todos os fenômenos capturados pelos outros dois modelos, é útil termos a liberdade de poder alternar entre os 3 níveis de abstração e escolher o modelo mais adequado para cada problema que estamos querendo resolver. E, mais ainda, como estamos apenas interessados no resultado final da simulação, estamos livres para combinar à vontade em nossos modelos características das 3 teorias. Essa liberdade também nos permite impor comportamentos explicados por um modelo mais completo a uma simulação baseada em uma teoria mais simples. Por exemplo, podemos estudar as propriedades de reflexão de um certo material com base na teoria eletromagnética para a luz e impor o comportamento previsto em uma simulação que ocorre em um nível de abstração mais alto baseada na ótica geométrica.

Como é comum na computação gráfica, a discussão nessa dissertação será centrada na ótica geométrica, onde o modelo mental que adotamos é o da luz viajando como um feixe de partículas que se propagam como raios em linha reta e interagem com o meio segundo regras puramente geométricas. Com esse modelo somos capazes de descrever os fenômenos que explicam em grande parte a aparência do mundo que nos cerca: emissão, reflexão e transmissão, difusas ou especulares, e absorção.

1.3

Aplicações

Definido o que é fotorrealismo e como iremos proceder para tentar gerar imagens realistas, resta dizermos o porquê. Não é difícil imaginarmos inúmeras aplicações.

Sem dúvida, uma das principais motivações para o desenvolvimento do fotorrealismo na computação gráfica foi e continua sendo a indústria do cinema. A formulação correta da propagação da luz é importante tanto para a criação por completo de seqüências concebidas somente na imaginação dos diretores, quanto para a criação de elementos virtuais que serão combinados com filmagens reais, onde uma iluminação consistente é fundamental.

Arquitetos podem, por exemplo, verificar como a cor de uma parede ou o posicionamento de uma janela influencia a iluminação de uma sala, examinar a aparência exterior de um prédio ou do interior de um ambiente durante o dia, à tarde, à noite, sob diferentes condições atmosféricas, etc. . . Enfim, têm a seu dispor uma ferramenta que os torna capazes de visualizar seus trabalhos antes de terem sido construídos.

Em situações onde deseja-se criar a ilusão de imersão em um ambiente real, como em simuladores de vôo, por exemplo, uma representação realista do mundo é essencial. Em simuladores, assim como em jogos, a interatividade deve ser priorizada em detrimento do realismo, o que limita consideravelmente a sofisticação dos algoritmos que podem ser usados. Apesar disso, avanços no hardware gráfico vêm reduzindo em um ritmo acelerado essa limitação.

1.4

Histórico do fotorrealismo na computação gráfica

Desde o início da computação gráfica, foi percebido o seu potencial para gerar imagens parecidas com fotografias. Uma das principais motivações era a construção de simuladores de vôo, mas, devido a limitações tecnológicas, a exigência de interatividade limitava consideravelmente a sofisticação das técnicas que podiam ser empregadas. Removida a restrição de interatividade, como na geração de seqüências de animação, por exemplo, os algoritmos passaram a poder gastar mais tempo tentando aumentar o grau de realismo. Mesmo assim, a maior parte das técnicas desenvolvidas eram essencialmente coleções de truques que tentavam obter determinados efeitos, como o modelo de iluminação de Phong (Phong 75), que tentava simular o realce em volta da direção de reflexão especular que observamos em materiais como plástico. Desse período, que cobre principalmente a década de 70, herdamos técnicas como o *shading* de Gouraud (Gouraud 71) e Phong (Phong 75), mapeamento

de textura (Catmull 74) e outras técnicas relacionadas, como *bump mapping* (Blinn 78) e *environment mapping* (Blinn & Newell 76).

Nos anos 80, surgiram dois algoritmos que revolucionariam nos anos seguintes a síntese de imagens: o algoritmo de traçado de raios e o de radiosidade. O primeiro usava técnicas de amostragem pontual e, o segundo, métodos de elementos finitos.

Introduzido por Turner Whitted em 1980 (Whitted 80), o algoritmo de traçado de raios estendia os modelos de iluminação local como o de Phong para também considerar parte da iluminação indireta que atingia um objeto. Passava a considerar não apenas a luz que vinha diretamente das fontes para interagir com um objeto e, em seguida, ser refletida para os olhos de um observador, mas também a luz vinda de uma fonte que, antes de atingir o objeto e ser refletida para um observador sofria um número de reflexões especulares e refrações nos demais objetos da cena. Com o lançamento de um raio ficava fácil também determinar se um ponto estava em sombra em relação a uma fonte de luz pontual. Partindo do olho do observador e gerando recursivamente raios refletidos e refratados, o algoritmo empregava as mesmas técnicas que já eram usadas há séculos na ótica para estudar o comportamento de sistemas óticos compostos por espelhos e lentes.

Em seguida, o algoritmo foi estendido consideravelmente em (Cook et al. 84). Reconhecendo que a intensidade da luz refletida em um ponto sobre uma superfície podia ser descrita por uma integral acima da superfície do produto de uma função de iluminação incidente com uma função que descrevia a reflexão da superfície, aplicaram técnicas de amostragem aleatória para avaliar essa integral. Por exemplo, para a iluminação direta, distribuindo raios na direção de pontos amostrados sobre uma fonte de luz eram capazes de capturar regiões de penumbra. Os raios secundários, que no modelo de Whitted eram gerados somente na direção de reflexão especular, agora podiam ser gerados segundo essa função de reflexão distribuída em torno da direção de reflexão especular. E o mesmo valia para a refração, fornecendo um método para capturar a aparência de materiais translúcidos. Mais ainda, a técnica não se restringia a calcular a função de iluminação. Podia também ser aplicada para distribuir raios ao longo do tempo, o que geraria *motion blur*, podia ser aplicada para se distribuir raios na direção da abertura de uma lente, que, em seguida, seriam refratados segundo uma aproximação para lentes delgadas e causaria o efeito da profundidade de campo. Enfim, reconhecendo que a intensidade de um pixel era resultado de uma série de integrais aninhadas (ao longo do tempo, sobre sua área, sobre a superfície da lente), o método estendia-se de maneira uniforme para obter uma série

de efeitos. No entanto, a falha principal do algoritmo era não ser capaz de capturar a aparência da inter-reflexão entre as superfícies, já que a reflexão e a transmissão eram tratadas apenas como uma distribuição em torno das direções especulares.

Falha que era justamente o problema que o algoritmo de radiosidade se propunha a resolver parcialmente. Partindo do princípio de que a maior parte dos ambientes reflete luz de maneira difusa, em 1984 (Goral et al. 84) apresentou o algoritmo de radiosidade para calcular a troca de energia entre superfícies difusas. Baseado em métodos de elementos finitos que já eram usados em problemas semelhantes na área de transferência de calor, o algoritmo de radiosidade assume que a cena é composta por um grande número de pequenos elementos idealmente difusos e resolve uma equação matricial para determinar a quantidade de luz atingindo cada elemento. Nos anos seguintes, o algoritmo de radiosidade recebeu ampla atenção e foi o foco de muita pesquisa. Inicialmente, a atenção voltou-se principalmente para a aceleração do cálculo da solução do sistema de equações lineares em que se traduzia a formulação do problema. Mais tarde, para a geração adaptativa de malhas que melhor expressassem a distribuição da energia, de modo que fossem capazes de, por exemplo, capturar detalhes que seriam perdidos em uma malha inicial grosseira ou, no limite, para a geração de malhas que fossem capazes de capturar discontinuidades na iluminação, como ocorre na separação entre regiões de sombra. Introduzida originalmente para cenas somente com superfícies difusas, o algoritmo de radiosidade foi mais tarde estendido para também conseguir lidar com modelos de reflexão mais complicados.

Em 1986, com base na física da propagação da luz, (Kajiya 86) apresentou a equação integral que generalizava o problema que esses dois algoritmos tentavam resolver. No contexto unificado dessa equação, que agora podia ser vista como o problema central da iluminação global, Kajiya descreveu as aproximações que eram feitas por vários algoritmos que haviam surgido na computação gráfica, como os algoritmos de iluminação direta, o de radiosidade e o de traçado de raios. No mesmo artigo, apresentou também uma solução que resolvia de forma completa essa equação integral com base no método de Monte Carlo. O seu algoritmo generalizava as idéias do algoritmo de traçado estocástico de raios de (Cook et al. 84), sendo capaz de levar em consideração as inter-reflexões entre superfícies. Embora capaz de calcular a troca de energia entre superfícies com características de reflexão arbitrárias, a principal desvantagem do algoritmo era exigir que fosse gerado um número muito grande de amostras para se produzir uma imagem com pouco ruído. Isso motivou o desenvolvimento de inúmeras técnicas que buscavam gerar com eficiência

as amostras nos algoritmos de Monte Carlo, como (Veach & Guibas 95) e (Shirley et al. 96). (Lafortune & Willems 93) e (Veach & Guibas 94), independentemente, desenvolveram as idéias do algoritmo de Kajiya para traçar simultaneamente raios partindo do olho e das fontes de luz. Entretanto, mesmo com os melhores algoritmos que surgiram, gerar imagens sem ruído ainda era muito custoso para ser usado em ambientes de produção.

Foram também desenvolvidos algoritmos híbridos que combinavam radiossidade com traçado de raios tentando obter o melhor dos dois mundos. Como os algoritmos de radiossidade eram melhores para calcular interações difusas e os baseados em traçado de raios, para reflexões especulares, uma tentativa natural era combinar os dois métodos. Esses algoritmos normalmente consistem em múltiplas passadas. A primeira resolve as interações difusas com métodos de radiossidade e a segunda, tomando cuidado para não calcular mais de uma vez as mesmas quantidades, calcula as interações especulares levando em consideração as interações difusas já calculadas. (Wallace et al. 87) foi o primeiro método a propor algoritmos desse tipo e muitos outros seguiram, como (Sillion & Puech 89) e (Chen et al. 91).

Como último marco, não podemos deixar de mencionar o surgimento de um algoritmo relativamente simples proposto por Henrik Wann Jensen para acelerar os algoritmos de traçado de raios baseados no método de Monte Carlo. O algoritmo de mapeamento de fótons (Jensen 01), como é chamado, funciona em duas passadas. A primeira emite fótons partindo das fontes de luz, os quais são armazenados em uma estrutura espacial desacoplada da geometria para descrever a distribuição da energia na cena. A segunda passada é essencialmente um traçado estocástico de raios que coleta as informações da estrutura espacial gerada na primeira passada. Com isso, o algoritmo permite que a iluminação indireta seja calculada de maneira eficiente e robusta, além de capturar efeitos como cáusticas, que dificilmente eram exibidas corretamente pela maior parte dos outros algoritmos. O algoritmo foi também adaptado para calcular a aparência de materiais translúcidos, como mármore.

1.5

Organização da dissertação

Nesta dissertação iremos percorrer todo o caminho que envolve a construção de um sistema capaz de gerar imagens realistas segundo a descrição física da luz.

No capítulo 2 iremos descrever a física da propagação e espalhamento da luz, e formular matematicamente o problema do transporte da luz entre superfícies. Estudaremos também como podemos modelar o processo de formação

de imagens.

No capítulo 3 apresentaremos os algoritmos de traçado de raios baseados em técnicas de integração de Monte Carlo para resolver o problema do transporte da luz.

As questões referentes à implementação serão discutidas no capítulo 4, onde iremos apresentar a arquitetura de um sistema capaz de gerar imagens fotorrealistas baseado nas técnicas descritas no capítulo 3.

No capítulo 5 apresentaremos os resultados obtidos com o sistema construído e, no capítulo 6, a conclusão e trabalhos futuros.

O apêndice A contém uma breve introdução à integração de Monte Carlo.