

5

Escultura do espaço através de refinamento adaptativo

Os métodos de escultura do espaço apresentados no segundo capítulo trabalham sobre uma representação do espaço de reconstrução determinada por uma *subdivisão espacial regular*. Embora seja simples, e permita a utilização de uma grande variedade de algoritmos, esta forma de estruturação do espaço, normalmente representada por uma matriz de voxels, possui algumas desvantagens que em certos casos podem tornar bastante difícil a aplicação do método.

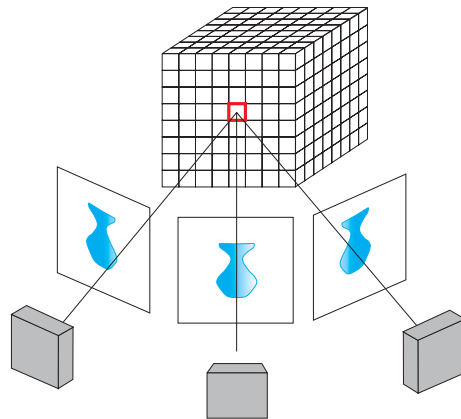


Figura 5.1: Espaço de reconstrução representado por voxels.

Em primeiro lugar, a representação do espaço de reconstrução através de um conjunto de voxels requer a utilização de um espaço de armazenamento cujo tamanho é proporcional à resolução com a qual se deseja reconstruir a cena, o que pode inviabilizar a utilização de métodos de escultura do espaço quando são necessárias reconstruções de alta qualidade. Além disso, nestes casos, é preciso avaliar a foto-consistência de um número gigantesco de elementos, tornando o processamento extremamente caro e conseqüentemente, lento.

Muitas das vezes, podemos classificar regiões extensas no espaço da cena de uma só vez, bastando para isso, tomar decisões com base nas propriedades fotométricas locais de tais regiões, as quais são determinadas

pelas câmeras e pelas imagens de entrada. Por exemplo, em muitos casos, grandes regiões vazias no espaço da cena podem ser facilmente detectadas verificando-se se a sua projeção nas imagens corresponde a um conjunto de pixels de fundo, não sendo necessária a avaliação da foto-consistência individual de um conjunto enorme de pequenos elementos estruturais em seu interior.

Obviamente, para que possamos efetuar processamentos sobre regiões extensas, como mencionamos acima, é necessário saber muito bem quais são estas regiões e como elas estão organizadas, o que infelizmente só é possível quando já tivermos conhecimento sobre como os elementos da cena se distribuem no espaço. Por outro lado, podemos descobrir as propriedades do espaço da cena de forma progressiva, desde que estruturamos o espaço de reconstrução convenientemente e adotemos mecanismos capazes de aproveitar tal estrutura.

Com base nesta idéia, iremos propor neste capítulo um método de escultura do espaço por refinamento adaptativo no qual o espaço de reconstrução é representado através de uma subdivisão espacial adaptativa. Mostraremos que tal método é capaz de realizar o processo de reconstrução de maneira mais eficiente, empregando esforços proporcionais às propriedades locais da cena, as quais são descobertas à medida em que a reconstrução é realizada.

5.1 Motivação

Sabemos que, quanto mais conhecimento tivermos sobre a estrutura do espaço no qual um determinado problema está definido, mais facilmente poderemos traçar estratégias capazes de tomar decisões que levem a soluções eficientes. Um exemplo típico são os problemas de busca em um conjunto de dados espaciais. Quanto melhor estruturados estiverem estes dados, levando em consideração as propriedades do espaço a que pertencem e as próprias relações de vizinhança existentes entre si, mais eficientes serão as operações capazes de identificar elementos com as características desejadas. Além disso, mais facilmente poderemos efetuar modificações locais sobre ela como, por exemplo, a remoção e a inserção de dados.

Os métodos de escultura do espaço estão interessados essencialmente na determinação dos elementos foto-consistentes na superfície dos objetos de interesse no espaço da cena. Conseqüentemente, podemos de certo modo, compreendê-los como métodos que procuram encontrar, em um determinado

espaço de busca, um conjunto de elementos que satisfazem determinadas propriedades. Conseqüentemente, quanto mais informações tivermos sobre esse espaço de busca, isto é, sobre o espaço de reconstrução, mais facilmente poderemos encontrar os elementos relevantes e conseqüentemente efetuar a reconstrução da cena.

É bastante evidente que a estruturação obtida através de uma matriz de voxels apenas estabelece uma organização conveniente do espaço, não incorporando em si mesma nenhuma informação sobre as propriedades locais associadas às regiões da cena. O máximo que os métodos de escultura do espaço conseguem extrair de tais estruturas, com o objetivo de auxiliar o processo de reconstrução, são as relações de ordem, e em alguns casos, de vizinhança, herdadas diretamente das propriedades geométricas e topológicas intrínsecas a uma partição uniforme de um subconjunto do espaço 3D euclidiano.

Se conseguíssemos estruturar o espaço de reconstrução de maneira que pudéssemos empregar esforços computacionais proporcionais as características locais das regiões do espaço em questão, inferidas a partir das informações codificadas nas imagens de entrada, então poderíamos atingir um grau de eficiência maior que o alcançado através de um mecanismo de busca exaustiva em uma partição uniforme. Na verdade, isto pode ser feito através do que chamamos de *método adaptativo*.

Um *método adaptativo* é um método capaz de se ajustar de acordo com as propriedades locais do espaço associado à definição do problema, quer seja através de mudanças de estratégia, quer seja através do emprego de esforços e recursos proporcionais às características locais das regiões pertencentes a tal espaço.

A primeira impressão que temos é a de que é fundamental o conhecimento prévio da organização e das características locais do espaço de definição do problema para que possamos utilizar métodos adaptativos.

Com relação ao problema de reconstrução de cenas através de escultura do espaço, podemos argumentar que, a princípio, não temos nenhum conhecimento sobre as propriedades do espaço de reconstrução, o que nos impede de estruturá-lo de modo mais inteligente, tornando impossível, desta forma, a realização de um processo de busca mais eficiente. Segundo este raciocínio, estamos limitados à estruturação do espaço de busca através de uma partição uniforme e a realização de um processo de busca exaustiva.

Entretanto, podemos observar que a estrutura da cena pode ser determinada progressivamente através de um processo de *refinamento* no qual a reconstrução é realizada inicialmente sobre uma representação bastante

grosseira do espaço de reconstrução inicial, que por sua vez é refinado em certas regiões à medida em que as propriedades do espaço que contém a cena são determinadas com base nos critérios de foto-consistência.

Desta forma, a estrutura do espaço de busca, isto é, do espaço formado pelos elementos que devem ser classificados, é determinada concomitantemente com a própria reconstrução da cena, tornando o processo muito mais ágil, já que deste modo podemos empregar esforços proporcionais às características locais da cena em questão.

Antes de analisarmos os requisitos necessários ao desenvolvimento de um método de escultura do espaço por refinamento adaptativo, iremos descrever, de forma bem sucinta, alguns dos trabalhos que abordaram previamente este aspecto do problema e quais as dificuldades por eles encontradas.

5.2

Alguns trabalhos anteriores

Até onde temos conhecimento, apenas três trabalhos abordaram, em diferentes níveis de profundidade, o desenvolvimento de métodos para reconstrução de cenas a partir de imagens por refinamento adaptativo: o trabalho de Szeliski [29], o trabalho de Prock [52] e o trabalho de Kutulakos [65].

Na verdade, apenas Prock [52] realmente propôs um método de escultura do espaço por refinamento adaptativo. O método de reconstrução de Szeliski [29], apesar de propor um esquema de refinamento adaptativo, não se baseia em critérios de foto-consistência, utilizando apenas informações provenientes de silhuetas, o que não o caracteriza, desta forma, como um verdadeiro método de escultura do espaço. Kutulakos [65], por sua vez, não chegou definitivamente a desenvolver um método de refinamento adaptativo, apesar de estabelecer as bases para isso, ao propor a avaliação da foto-consistência de elementos não pontuais através de *shuffle transforms*. De fato, em seu trabalho ele afirma que a aplicação de *shuffle transforms* no desenvolvimento de tais métodos ainda é um tópico em aberto, o qual necessita ser investigado com maior profundidade.

A utilização de estratégias de refinamento adaptativo na reconstrução de cenas parece ter surgido juntamente com o trabalho de Szeliski, o qual, na época, apresentou resultados bastante surpreendentes. No entanto, o que parecia ser uma idéia bastante promissora, acabou sendo praticamente

esquecida quando os métodos baseados em foto-consistência começaram a ser adotados.

Uma possível explicação para a não utilização desta técnica está no fato de que a teoria sobre forma baseada em foto-consistência está fundamentada sobre o conceito matemático de ponto, enquanto que os métodos de escultura do espaço por refinamento adaptativo envolvem a avaliação de foto-consistência sobre elementos que dificilmente podem ser aproximados através de elementos pontuais. De fato, a avaliação da foto-consistência de elementos de natureza não pontuais é um problema bastante difícil, cuja compreensão é fundamental para que possamos chegar a uma boa modelagem para o problema de reconstrução volumétrica de cenas através de refinamento adaptativo.

Apesar das dificuldades descritas acima, Prock propôs um método de coloração de voxels por refinamento adaptativo que utiliza uma coloração obtida em baixa resolução como ponto de partida para a obtenção de colorações em resoluções mais altas. Sem entrar em muitos detalhes, o método funciona da seguinte forma: efetua-se uma coloração de voxels sobre um espaço de reconstrução em baixa resolução e em seguida, seleciona-se os voxels que foram considerados foto-consistentes. Estes voxels são subdivididos em oito novos voxels sobre os quais é efetuado um novo passo de coloração de voxels. O processo é aplicado recursivamente até que tenha-se alcançado a resolução desejada.

Prock imediatamente percebeu que esta estratégia não levava a resultados corretos, já que permitia que regiões contendo fragmentos de superfície foto-consistente pudessem ser eliminadas durante o processamento nos níveis de resolução mais grosseiros. Para contornar este problema, que está diretamente relacionado à dificuldade de se definir foto-consistência em regiões extensas do espaço da cena, Prock utilizou uma estratégia de busca local para evitar a perda de voxels consistentes durante as etapas realizadas em baixa resolução. Sua idéia consiste em subdividir e processar, em uma resolução mais alta, não somente os voxels considerados consistentes em um determinado nível de resolução, mas também todos os voxels que estiverem em uma vizinhança unitária dos voxels escolhidos (Figura 5.2).

Aparentemente, para simplificar este processo de busca, Prock não utilizou uma estruturação adaptativa do espaço da cena. Simplesmente, a cada novo nível de refinamento, todo espaço é discretizado uniformemente com uma resolução maior, embora somente os elementos selecionados no nível anterior sejam considerados na nova etapa de coloração de voxels.

Apesar de produzir resultados muito bons, o método proposto por

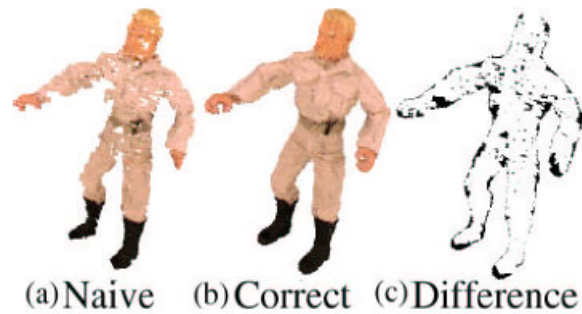


Figura 5.2: Coloração através de refinamento. Figura obtida de [52]. (a) - reconstrução com perdas; (b) - reconstrução correta; (c) - diferença

Prock nos obriga a iniciar o processo de reconstrução com uma discretização do espaço fina o suficiente para que o processo de coloração seja capaz de capturar ao menos um voxel no interior de cada componente conexa correspondente a cada objeto de interesse na cena. Denominamos tais voxels de *sementes* da reconstrução. Caso algumas destas sementes sejam perdidas logo no início do processo, o objeto correspondente não aparecerá no final da reconstrução.

Iremos agora descrever o método de escultura do espaço através de refinamento que propomos neste trabalho, começando pela descrição do esquema utilizado para representar o espaço de reconstrução. De forma surpreendente, a escolha deste esquema nos dará pistas sobre como tratar a questão sobre a avaliação da foto-consistência de elementos não pontuais.

5.3 Representação adaptativa do espaço de reconstrução

Um dos pontos chaves para o desenvolvimento de um método de escultura do espaço por refinamento adaptativo está na escolha do esquema utilizado para representar o espaço de reconstrução. Para que possa permitir que a reconstrução seja realizada de forma incremental e adaptativa, ao mesmo tempo em que preserve as vantagens associadas aos métodos de escultura do espaço, o esquema de representação utilizado deve possuir as seguintes propriedades:

Poder de expressão: o esquema de representação deve ser capaz de representar a cena reconstruída com grau de precisão arbitrário; além disso, deve ser o mais genérico possível, não impondo restrições sobre a topologia

da cena a ser reconstruída.

Concisão: a representação deve ser feita de forma concisa aproveitando a coerência espacial existente entre as regiões do espaço no qual a cena está definida.

Capacidade de representar relações de ordem: esta propriedade é útil para o desenvolvimento de eficientes algoritmos de escultura do espaço. Vimos anteriormente que o processo de escultura do espaço pode ser desempenhado de forma muito mais fácil quando existem relações de ordem bem definidas entre os elementos que compõem o espaço de reconstrução. Logo, é importante que o esquema permita que tais relações sejam extraídas de maneira simples e eficiente, sem um aumento da complexidade dos algoritmos baseados em varreduras por planos.

Dinamismo e adaptabilidade: o esquema de representação deve ser dinâmico, de maneira que possa acomodar novas informações que são inferidas no decorrer do processo de escultura do espaço. Além disso, os elementos estruturais utilizados na representação devem se adaptar, da melhor forma possível, às características do modelo que procuram representar.

Os esquemas de representação do espaço que satisfazem com maior naturalidade as propriedades descritas acima são os esquemas baseados em *subdivisão espacial adaptativa*. Tais esquemas se baseiam na propriedade fundamental de que o número de elementos necessários para representar um sólido, através de uma subdivisão do espaço, é proporcional à sua área e não ao seu volume. Logo, enquanto os esquemas baseados em enumeração exaustiva utilizam um número de elementos proporcional a r^3 , onde r é a resolução utilizada, os esquemas baseados em subdivisão adaptativa utilizam um número de elementos proporcional a r^2 .

Computacionalmente, é comum representar os esquemas de subdivisão espacial adaptativa através de estruturas espaciais hierárquicas como, por exemplo, *octrees*, *bintrees*, *k-d trees*, etc. Essas estruturas são capazes de representar uma região do espaço adaptativamente, com base em um processo de subdivisão seletiva, segundo o qual são refinadas somente as subregiões estritamente necessárias. Além disso, por serem dinâmicas e de fácil atualização, são capazes de capturar a organização do espaço de forma progressiva, tornando possível o processo de reconstrução por refinamento.

Por outro lado, devido ao fato de permitirem configurações mais genéricas, tais estruturas não são tão eficazes quanto as estruturas baseadas em particionamentos regulares, na representação das relações de ordem existentes entre os elementos primitivos utilizados na descrição do espaço.

Nestas últimas, uma simples função, definida sobre os índices associados aos seus elementos primitivos, é capaz de nos informar imediatamente quando um determinado elemento está à frente de um outro, segundo uma certa direção. Da mesma forma, nas estruturas espaciais hierárquicas, as relações de vizinhança entre os elementos não são tão explícitas quanto nas estruturas determinadas por partições uniformes.

Tanto as relações de ordem quanto de vizinhança podem ser extraídas a partir das estruturas hierárquicas através de processos de busca, que apesar de não serem tão eficientes quanto uma simples operação sobre índices, não são extremamente custosas, sendo normalmente descritas através de operações de percorrimento em árvores. Veremos mais adiante que tais processos de busca serão fundamentais para que possamos determinar quais elementos da estrutura devem estar ativos em cada uma das iterações do nosso algoritmo de escultura do espaço por refinamento adaptativo.

Em nosso trabalho, representaremos o espaço de reconstrução através de uma octree, que apesar de ser uma estrutura simples, é poderosa o suficiente para atender os requisitos necessários para o desenvolvimento de um método por refinamento adaptativo.

As *octrees* são estruturas espaciais hierárquicas que se baseiam em uma subdivisão recursiva do espaço de interesse em oito octantes, a qual é representada através de uma árvore 8-ária (Figura 5.3). Quando uma região do espaço é representada através de uma octree, cada um dos oito octantes iniciais pode ser classificado em *cheio*, *vazio* ou *parcialmente cheio*, dependendo da sua interseção com o objeto em questão. Os octantes parcialmente cheios devem ser subdivididos em novos oito octantes, os quais são classificados recursivamente da mesma forma que os 8 primeiros. O processo termina apenas quando todos os octantes são considerados homogêneos ou até que um grau de subdivisão máximo, pré-estabelecido, tenha sido alcançado. Na estrutura resultante, os octantes parcialmente cheios compõem os nós internos da árvore, enquanto que os octantes cheios e vazios determinam as suas folhas.

Através do uso de uma estrutura octree, é possível desenvolver um método de escultura do espaço capaz de concentrar esforços nas regiões próximas à superfície dos objetos de interesse, a qual por sua vez, é o alvo principal dos métodos de escultura do espaço, apesar de toda sua conotação volumétrica.

Desta forma, grandes regiões vazias ou grandes espaços pertencentes ao interior do objeto são determinadas logo nos estágios iniciais do processo de reconstrução, sendo representadas através de células grandes de

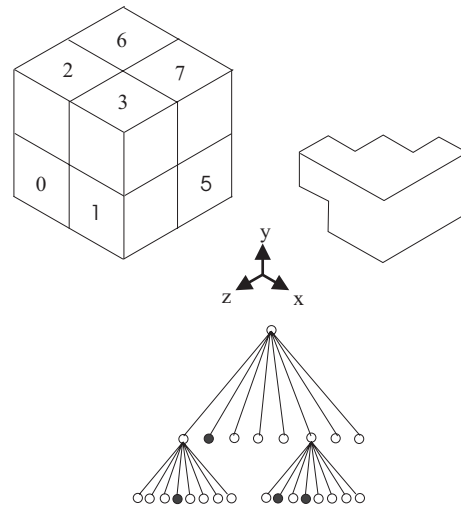


Figura 5.3: Representação através de octrees

baixa resolução. Por sua vez, as regiões correspondentes a fragmentos de superfícies foto-consistentes tendem a ser identificadas em estágios posteriores e representadas através de células de maior resolução.

É possível que outras estruturas espaciais sejam capazes de permitir um maior grau de adaptabilidade e concisão, como por exemplo as árvores baseadas em subdivisão binária (*bintrees*) e *árvores k-d*. Todavia, deixaremos a investigação da representação do espaço de reconstrução através de tais estruturas para trabalhos futuros.

Veremos agora como estas estruturas se encaixam no esquema de refinamento adaptativo proposto e quais as suas implicações sobre o método.

5.4

Escultura de cenas em espaços representados através de octrees

Os primeiros métodos de reconstrução volumétrica a utilizar octrees para representar os espaços de reconstrução foram os métodos de reconstrução a partir de silhuetas. Citamos como exemplos os métodos propostos nos trabalhos de Chien[18], Potmesil [23], Veenstra [20], Srivastava[27], e principalmente Szeliski [29].

Neste trabalho, de forma similar aos trabalhos mencionados acima, apresentamos um método de reconstrução a partir de imagens que trabalha sobre espaços volumétricos representados através de octrees, o qual, no entanto, utiliza para reconstruir a cena de interesse, tanto informações

fotométricas, quanto informações provenientes da segmentação do objeto de interesse nas imagens (Figura 5.4).

É bastante comum que os algoritmos desenvolvidos para solucionar um determinado problema sejam influenciados pela escolha da estrutura de dados utilizada para representar o espaço no qual tal problema é definido. Por esta razão, o método que propomos possui muitos pontos em comum com os esquemas utilizados para construção de representações através de octrees.

A idéia fundamental do método consiste em descrever inicialmente o espaço de reconstrução através de uma célula, correspondente à raiz de uma octree, cujo volume é especificado de maneira que englobe a cena que desejamos reconstruir. Em seguida, procuramos classificá-la segundo a sua foto-consistência em relação às imagens nas quais se encontra visível, o que nem sempre é possível, por motivos que descrevemos posteriormente. Se conseguirmos classificar a célula inicial, quer seja como foto-consistente, quer seja como não-foto-consistente, então podemos encerrar o processo de reconstrução, bastando apenas atribuir as informações fotométricas apropriadas à célula em questão. Por outro lado, quando não for possível classificá-la com precisão, o que ocorre com grande probabilidade no início do processo, precisamos subdividi-la uniformemente em oito novas células as quais devem sofrer um novo processo de classificação. Este procedimento é então repetido sucessivamente até que toda a região do espaço seja classificada adequadamente, conforme sua foto-consistência em relação às imagens de entradas, ou até que o nível máximo de subdivisão da árvore tenha sido atingido.

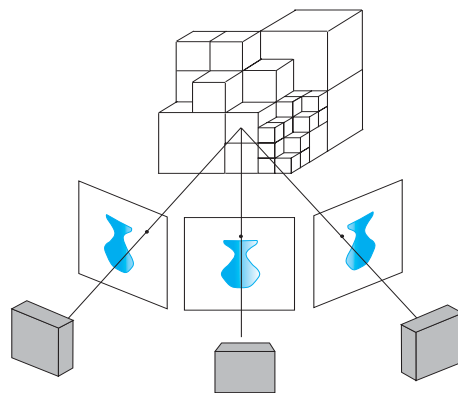


Figura 5.4: Coloração em uma octree

A aplicação deste método de reconstrução requer o tratamento das seguintes questões que serão abordadas nas seções seguintes.

- Como classificar uma célula com base em sua foto-consistência.
- Como efetuar a classificação do espaço de reconstrução sistematicamente.
- Como atribuir as informações fotométricas às células classificadas.

5.4.1

Como classificar uma célula da octree

Como vimos no capítulo 2, os métodos de escultura do espaço tradicionais procuram classificar o conjunto de voxels representando uma sub-região de interesse do espaço 3D em duas categorias: *voxels foto-consistentes* e *voxels não-foto-consistentes*. Analogamente, o método de escultura do espaço por refinamento adaptativo que propomos deve ser capaz de classificar as células da octree em células foto-consistentes e células não-foto-consistentes.

O grande desafio que devemos enfrentar, ao propor um método de reconstrução que trabalha sobre um espaço representado através de uma octree, é o de como avaliar a foto-consistência nas regiões do espaço correspondentes a cada célula sem precisar avaliar um conjunto muito grande de pontos em seu interior.

Antes de tomar qualquer decisão, precisamos primeiramente analisar se as classes de foto-consistência utilizadas nos métodos tradicionais são capazes de representar todas as circunstâncias que surgem em um processo de escultura do espaço por refinamento.

Devemos nos lembrar que, no caso geral, a detecção de não-foto-consistência em um um pequeno conjunto de amostras no interior de uma região extensa não nos permite afirmar com certeza absoluta que toda a região é inconsistente, pois existe a possibilidade de que haja um outro conjunto de amostras capaz de satisfazer as restrições determinadas pelas imagens nas quais elas se encontram visíveis. Além disso, a região a ser avaliada pode conter uma combinação de partes que correspondem a regiões vazias com partes que correspondem a regiões do objeto, tornando ainda mais difícil o processo de classificação.

Ambas as questões não eram tão relevantes quando utilizávamos uma representação do espaço da cena através de voxels de dimensões desprezíveis, pois nestes casos a teoria sobre foto-consistência se comporta de maneira adequada.

Como descrevemos anteriormente, apenas Kutulakos propôs um critério, baseado nas *shuffle transforms*, capaz de classificar regiões extensas

apropriadamente. Por outro lado, o critério por ele utilizado é extremamente conservativo, uma vez que requer que a projeção de uma região do espaço nas imagens de entrada não contenha cores similares em uma vizinhança em torno de seus centros. Além disso, a determinação do tamanho adequado para as regiões de vizinhança, dentro do contexto de métodos por refinamento adaptativo, é um problema que ainda se encontra totalmente em aberto.

Ao invés de buscarmos critérios complexos para determinação da foto-consistência em regiões não pontuais, propomos uma abordagem mais simples, a qual, no entanto, é capaz de manter o aspecto conservativo dos métodos de escultura do espaço. Simplesmente criamos um novo grupo de classificação para as células da octree: o grupo das células *indefinidas*, as quais contemplam aquelas que não podem ser classificadas nem como foto-consistentes nem como não-foto-consistentes.

Conseqüentemente, as células da octree passam a ser classificadas da seguinte forma:

- células foto-consistentes.
- células não-foto-consistentes.
- células indefinidas.

Em nosso esquema de classificação, o conceito de célula foto-consistente é análogo ao conceito de foto-consistência existente na teoria desenvolvida por Seitz e Kutulakos. Por outro lado, restringimos o conceito de célula não-foto-consistente apenas às células de mais alta resolução que falharam no teste de foto-consistência ou então às células, neste caso de qualquer nível, que se projetam completamente em regiões de fundo.

As células indefinidas correspondem às células cuja foto-consistência foi invalidada pela análise, mas que de certa forma, não podemos afirmar com certeza absoluta que não há em seu interior nenhum fragmento de superfície capaz de satisfazer as restrições impostas pelas imagens. Isto pode ocorrer como conseqüência dos seguintes eventos: primeiro, quando a célula contém uma combinação de pontos que pertencem ao espaço vazio com pontos que pertencem ao objeto; segundo, quando o critério de foto-consistência não pode afirmar com precisão se a célula é foto-consistente ou não. O último caso acontece quando os elementos utilizados no procedimento de avaliação da foto-consistência não são suficientemente representativos ou então, quando o valor de foto-consistência obtido é muito próximo ao limiar de separação das classes.

Quando uma célula indefinida é detectada, devemos refiná-la, isto é, subdividi-la em oito novas células de forma que estes novos elementos gerados possam ser processados novamente em uma nova iteração do algoritmo de escultura do espaço em um nível de refinamento mais elevado.

Através desta estratégia não precisamos nos preocupar em utilizar um conjunto de amostras representativas para a célula que desejamos classificar, já que nos casos de não-foto-consistência, com exceção do caso em que a célula corresponde a uma região vazia, sempre efetuamos uma subdivisão da mesma até que a amostragem utilizada seja representativa o suficiente para que possamos efetuar uma classificação fidedigna. Em nosso método, a representatividade só é garantida quando chegamos ao nível de resolução máximo.

Por esta razão, podemos classificar as células da octree da mesma forma que classificamos voxels no método não adaptativo, isto é, através da avaliação da foto-consistência de um conjunto de amostras em uma região planar determinada pela interseção entre as células com um plano passando pelo seus centros e com a mesma orientação da direção de varredura.

5.4.2

Classificação do espaço de reconstrução

Como vimos anteriormente, a estratégia baseada em varredura de planos no espaço da cena é uma das mais eficientes para determinar a foto-consistência em um conjunto de regiões planares com uma mesma orientação. No entanto, para que possamos adotar esta mesma solução, para o caso de regiões planares associadas às células de uma octree, precisamos levar em consideração os seguintes problemas: primeiro, as células são geradas progressivamente, o que nos impede de utilizar um único passo do algoritmo de varredura de planos; segundo, o conceito de camadas, fundamental para o algoritmo de varredura de planos, não é muito bem definido em uma octree, uma vez que as células tem tamanho diferentes e se distribuem pelos diferentes níveis da estrutura (Figura 5.5).

Ambos os problemas podem ser resolvidos de maneira simples através da aplicação de um passo do algoritmo de varredura de planos para cada nível de refinamento. Logo, diferentemente dos métodos tradicionalmente utilizados para obtenção de representações através de octrees, não exploramos a estrutura em profundidade, mas sim em largura, o que significa que todos os nós de um determinado nível são processados antes de processarmos os nós do nível de resolução seguinte.

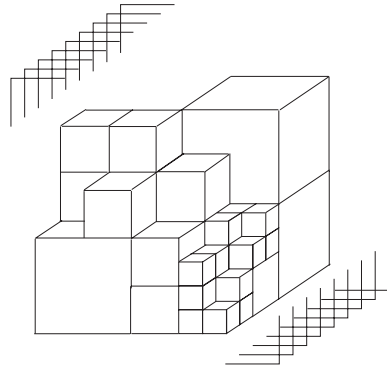


Figura 5.5: Varredura de planos em uma octree

O número de planos varridos em uma determinada aplicação do algoritmo de varredura de planos é igual ao número de camadas definidas na estrutura utilizada para representar o espaço de reconstrução. Como não é possível identificar camadas em uma octree, a não ser em casos muito particulares, determinamos o número de planos varridos através do número de camadas existente em uma partição regular com resolução igual à resolução do nível corrente na octree. Por exemplo, o número de planos no nível zero da octree é igual a um, pois existe uma única camada em uma partição regular equivalente; no nível 1, o número planos é igual a dois, já que existem duas camadas na partição regular equivalente, e assim por diante (Figura 5.6). Desta forma, o número de planos em uma aplicação do algoritmo de varredura de planos depende do nível de resolução corrente do processo, sendo então proporcional a 2^κ , onde κ é o nível da octree.

Como nas octrees não existem estruturas equivalentes às camadas, como existem nas representações baseadas em subdivisão regular, precisamos especificar um mecanismo para determinar quais são as células ativas durante uma iteração correspondente a determinação da foto-consistência em um determinado plano de referência em um certo nível de resolução. A solução encontrada também é bastante simples e consiste em percorrer a árvore em profundidade, determinando quais são as células que são interceptadas pelo plano de referência corrente que ainda não foram avaliadas (obviamente, isto exclui as células que já foram checadas nos níveis anteriores)(Figura 5.7).

Assim como nos métodos de escultura do espaço não adaptativos, a realização do algoritmo de varredura de planos requer a definição das seguintes etapas: registro das informações no espaço da cena, tratamento da visibilidade, determinação da foto-consistência e colorização dos elementos estruturais. A maioria destes procedimentos requer algumas modificações e

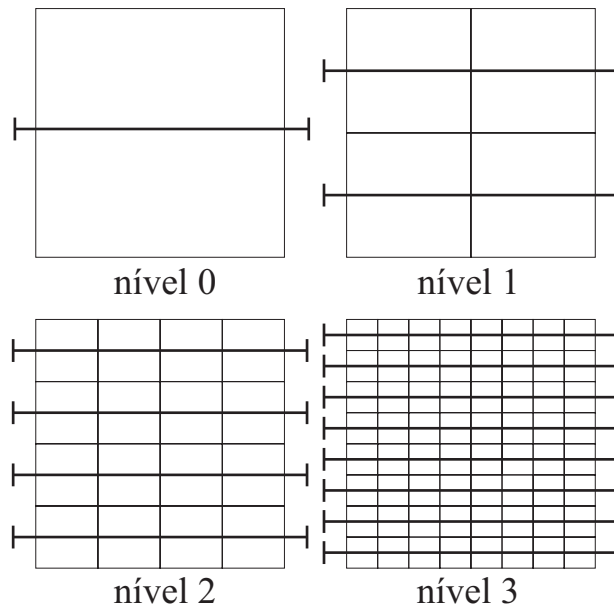


Figura 5.6: Número de planos em cada nível de resolução

adaptações para que possamos adotá-los no método de escultura do espaço adaptativo, as quais serão vistas em seguida.

Registro

O registro das informações no espaço da cena pode ser realizado de modo similar ao que é realizado no método de escultura do espaço não adaptativo, isto é, através do mapeamento projetivo das imagens de entrada sobre os planos de referência utilizados em cada iteração do algoritmo de varredura de planos em cada nível de refinamento. No entanto, uma atenção especial deve ser tomada em relação à resolução com a qual cada uma das imagens projetadas devem ser geradas.

As resoluções das imagens projetadas dependem do número de amostras utilizadas no processo de avaliação de foto-consistência, as quais por sua vez estão diretamente relacionadas ao número de células cortadas pelo plano de referência correspondente. Como o número de células varia de acordo com o nível de resolução com o qual trabalhamos, então, a resolução de uma imagem projetada deve ser proporcional ao nível de refinamento sobre a qual trabalha a etapa de varredura de planos corrente.

Obviamente, podemos gerar todas as imagens projetadas com a resolução máxima. No entanto, isto não parece sensato pois implica na diminuição do número de amostras para cada célula, à medida em que a resolução da representação do espaço aumenta. Uma estratégia mais

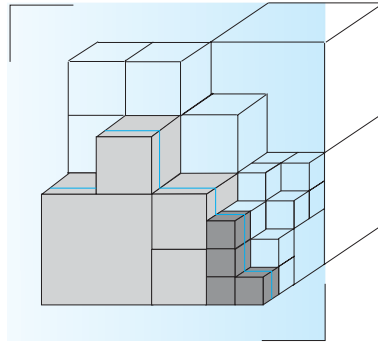


Figura 5.7: Células interceptadas por um plano de varredura. As células em tom escuro são as células ativas, as em tom mais claro são células já avaliadas.

razoável consiste em gerar imagens projetadas com um número de amostras que se mantém constante (ou que aumente), em cada célula, à medida em que aumenta o nível de refinamento da representação do espaço de reconstrução.

Uma justificativa para a não utilização de imagens com máxima resolução, nas etapas iniciais do processo de varredura de planos, é a de que nestes estágios dificilmente haverá correlação das imagens de entrada em uma dada região planar, implicando em um desperdício de esforço computacional causado pela utilização de um número excessivo de amostras. Por este motivo preferimos manter um número de amostras constante para cada célula, em todos os níveis de resolução, o que significa que teremos uma amostragem menos densa no início do processo e mais densa nos estágios de maior refinamento. Um valor razoável, por exemplo, é o de quatro amostras para cada célula (Figura 5.8).

A geração de imagens nas diferentes resoluções pode ser realizada de maneira simples através da variação do tamanho da janela (*viewport*) associada ao canvas (*framebuffer*) no qual as imagens projetadas são produzidas. Como as imagens projetadas são geradas com tamanho inferior ao das imagens originais, com exceção das geradas no último nível de refinamento, é fundamental a utilização do processo de mipmapping durante o mapeamento projetivo, já que tal mapeamento se caracteriza com um mapeamento de redução.

O fato de produzirmos as imagens em resolução proporcional ao nível de refinamento da representação do espaço de reconstrução tem implicações fundamentais sobre a etapa de determinação da foto-consistência das células, principalmente em relação à escolha do limiar de classificação, uma vez que neste caso, as amostras sofrem alterações consideráveis em

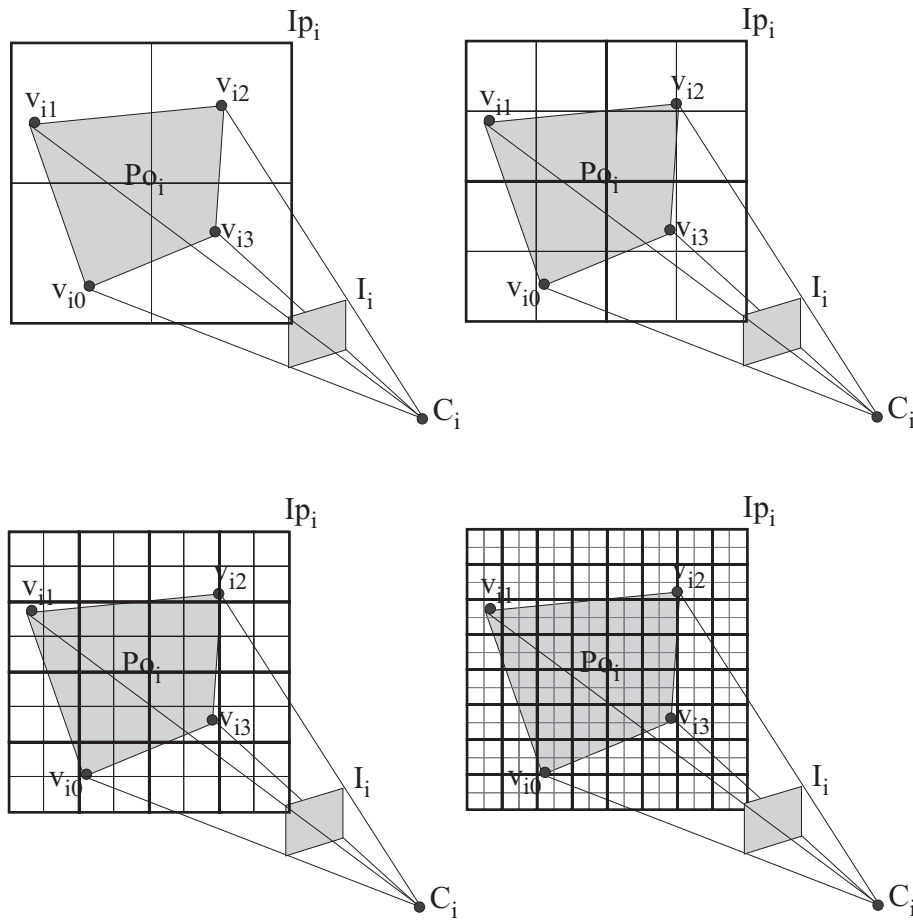


Figura 5.8: Imagens projetadas em diferentes níveis de resolução

função do processo de filtragem. Iremos investigar este aspecto com mais detalhes na seção que aborda a determinação de foto-consistência.

Uso das informações sobre a segmentação dos objetos de interesse nas imagens

A utilização das informações sobre a segmentação dos objetos de interesse nas imagens de entrada é fundamental no método de escultura do espaço adaptativo que propomos. Em nosso método, somente através de uma análise do espaço de reconstrução, com base nas informações sobre segmentação, é que podemos detectar e remover grandes regiões vazias existentes no espaço da cena, visto que a análise de foto-consistência, por si somente, não nos dá garantia sobre a não-foto-consistência de uma determinada região, a menos que estejamos no estágio de refinamento máximo.

Uma outra função importante das informações sobre segmentação é a

de que elas funcionam como um dos mecanismos reguladores do processo de refinamento. Graças à reamostragem das informações de segmentação através de mipmapping, podemos garantir que o espaço é refinado sempre que necessário, independentemente do grau de refinamento inicial do espaço de reconstrução, diferentemente do que ocorre no método de Prock.

Por exemplo, vejamos o que ocorre quando representamos o espaço de reconstrução inicial através de uma única célula englobando toda a cena. Neste caso, os mapas de segmentação serão projetados sobre o plano de referência correspondente após terem sido filtrados e subamostrados. Logo, os valores associados à classificação dos elementos na região planar correspondente, em função dos mapas de segmentação, não serão homogêneos, o que levará a determinação de tal célula como indefinida. Em outras palavras, isto significa que foi detectada, na célula, a presença de elementos do fundo da cena mesclados com elementos pertencentes aos objetos de interesse. Conseqüentemente, em função desta classificação tal célula será obrigatoriamente subdivida, o que fará com que o processo de escultura do espaço tenha prosseguimento.

Tratamento da visibilidade

O esquema adotado para a determinação da visibilidade das células da octree é similar ao esquema adotado no algoritmo de escultura do espaço não adaptativo no espaço da cena. Isto significa que precisamos definir mapas de visibilidade, através de mapas de textura, para cada uma das câmeras, os quais são registrados sobre os planos de referência à medida em que as iterações do algoritmo de varredura do espaço são realizadas.

A única diferença importante, que deve ser ressaltada, é a de que precisamos reconstruir os mapas de visibilidade a cada nova etapa do algoritmo de varredura de planos correspondente a cada novo nível de refinamento. Isto é necessário porque as informações no mapa de visibilidade não são mais válidas quando mudamos de um nível para o outro.

Os mapas de visibilidade, definidos por mapas de bits que são atualizados de forma sincronizada com o processo de varredura, apenas refletem a visibilidade dos voxels no plano de referência corrente. Quando mudamos de um nível para o outro, as informações de visibilidade não mais são válidas pois apenas refletem a visibilidade do último plano de varredura do nível de resolução anterior. Como neste novo nível podemos encontrar novos voxels consistentes que se posicionam a frente de tal plano, isto é, do último plano de referência do nível anterior, então não podemos nos guiar por tais

informações. Portanto, precisamos limpar os mapas de visibilidade antes da execução do algoritmo de varredura de planos em um novo nível de refinamento.

Uma consequência imediata deste fato é a de que os mapas de visibilidade são reconstruídos a cada nova execução do algoritmo de varredura de planos em nível de resolução mais refinado. Por esta razão, na etapa de atualização dos mapas de visibilidade, que ocorre após a determinação da foto-consistência das células ativas interceptadas por um determinado plano de referência, precisamos considerar não somente as novas células foto-consistentes, mas também todas as células foto-consistentes que interceptam o plano de referência corrente que foram determinadas nos níveis de refinamento anteriores.

Determinação da foto-consistência

O procedimento utilizado para determinar a foto-consistência das células de uma octree é similar aos procedimentos vistos nos capítulos 4 e 5, utilizados no método de escultura do espaço não adaptativo para determinar a foto-consistência de um conjunto de voxels.

Entretanto, existe uma diferença fundamental causada pelo fato de que, no método de escultura do espaço por refinamento adaptativo, as imagens projetadas são geradas através de um processo de reamostragem baseado em *mipmapping*, o que significa que cada um de seus elementos é, por si próprio, uma combinação de elementos do nível de resolução mais alto da representação. Mais importante ainda é o fato de que o número de elementos em tal combinação depende diretamente da resolução com a qual as imagens são geradas. Obviamente isto tem um impacto direto sobre a forma como a foto-consistência deve ser calculada, pois cada amostra, tomada para uma determinada célula, neste caso, é uma combinação de elementos.

Para que possamos determinar como a foto-consistência deve ser realizada de maneira que leve em consideração este fato, precisamos primeiramente compreender qual é o efeito da operação de filtragem por *mipmapping* sobre a variância das amostras tomadas para um dado elemento, já que é através dela que estimamos o quão próximo suas médias se encontram, o que por sua vez nos permite afirmar através de um teste de comparação com um limiar pré-estabelecido, se as observações são foto-consistentes ou não. Vimos anteriormente que, a grosso modo, a operação de filtragem efetua uma média sobre um conjunto de elementos de textura em uma vizinhança

determinada de modo aproximado, pela região correspondente ao mapeamento de um pixel no espaço da textura. Intuitivamente, tal operação, causa uma redução no desvio das amostras observadas. Logo, para que possamos compensar esta redução é necessário reduzir o limiar que caracteriza a foto-consistência de maneira proporcional à redução observada.

Para ajustar o limiar nos estágios em que trabalhamos com imagens em resolução mais baixa, usamos o fato de que, se X_1, X_2, \dots, X_n são independentes, com mesma variância σ^2 , sua média \bar{X} tem variância igual a σ^2/n . Em nosso caso, consideramos que cada amostra em um dado nível de resolução seja a média de quatro amostras no nível superior. Isto significa que em um dado nível, cada amostra é uma média de $4^{(nimax-ni)}$ amostras, onde $nimax$ é o nível máximo de resolução e ni é o nível corrente. Portanto, a variância calculada para um conjunto de amostras representadas por X_1, X_2, \dots, X_n , em um dado nível ni de fato deve ser igual a $\sigma^2/4^{(nimax-ni)}$, onde σ é o desvio calculado com base nos valores observados.

Por exemplo, no teste utilizado no capítulo 4, o limiar em um dado nível de refinamento ni deve ser proporcional ao desvio padrão correspondente àquele nível. Logo deve ser igual a $limiar/2^{(nimax-ni)}$, onde $limiar$ é o limiar no nível mais alto de resolução.

Na verdade, em nosso caso, não podemos assumir, a princípio, que os elementos que compõem uma determinada amostra são independentes, e de fato não são, já que a convolução, realizada durante a reamostragem, faz com que cada uma delas receba contribuições de amostras vizinhas. Logo, devemos estar conscientes de que tal ajuste é apenas uma aproximação.

Atribuição de cores às células da octree

Esta é uma das questões mais complexas no método de escultura do espaço por refinamento adaptativo, e que precisa de um estudo mais profundo para ser tratada de forma apropriada. Em nossa implementação simplesmente colorimos as células da mesma forma que colorimos voxels de dimensões pequenas, isto é, através de uma média das cores das amostras tomadas para cada elemento pertencente à região planar associada. Entretanto, sabemos que esta solução pode produzir resultados bastante insatisfatórios quando aplicada a alguns casos particulares, como, por exemplo, quando poucas câmeras enxergam uma região com geometria quase planar que, no entanto, possui informações de cor de grande variabilidade. Nestes casos, há uma probabilidade razoável de que tal região seja determinada como foto-consistente nos níveis de resolução mais baixos o que levaria a

uma atribuição de cores inadequada, já que a média de uma região de grande variabilidade não seria representativa o suficiente. Apesar deste problema, na prática os resultados são bastante satisfatórios, e o que é mais importante: tal comportamento não é simples consequência do acaso e sim da própria estrutura intrínseca do método proposto.

Dois fatores são cruciais para a obtenção de resultados corretos no processo de colorização utilizado: o emprego da estratégia por refinamento adaptativo, combinada com o uso de informações de segmentação em resoluções incrementais, e o uso de limiares que se adequam aos diferentes níveis de refinamento no teste de checagem da foto-consistência dos elementos do espaço de reconstrução.

O primeiro fator automaticamente favorece a subdivisão dos elementos próximos à superfície do objeto durante o processamento nos estágios de refinamento mais grosseiros. Isto se deve ao fato de que, nestes estágios, os valores provenientes dos mapas de segmentação projetados, observados em tais elementos, são determinados por uma combinação (causada pela operação de filtragem) de valores que correspondem ao fundo e aos objetos de interesse. Logo, é esperado que nos estágios mais grosseiros tais regiões sejam consideradas duvidosas, o que fará com que sejam obrigatoriamente subdividas. Este aspecto justifica tanto o uso de informações de segmentação quanto o emprego de representações em multiresolução das informações utilizadas na avaliação de foto-consistência.

O segundo fator dificulta a classificação de células de baixa resolução como foto-consistentes, pois os limiares são bastante baixos no início do processo, compensando desta forma a diminuição dos desvios causadas pela operação de filtragem. Logo, células de baixa resolução somente são classificadas como foto-consistentes se houver uma correlação muito grande entre as informações registradas, o que dificilmente acontecerá a menos que a variabilidade da região seja muito pequena e a superfície da cena coincida com o plano de varredura.

Obviamente, o procedimento que utilizamos pode falhar em alguns casos patológicos. Por este motivo, seria melhor encontrarmos uma solução que cobrisse todos os casos possíveis. Infelizmente, a validação da determinação de uma célula foto-consistente de baixa resolução envolve o conhecimento sobre a variabilidade nas regiões das imagens registradas na resolução original, o que tornaria o processo extremamente caro se fôssemos obrigados a projetar as imagens sempre na resolução máxima. Uma possível estratégia seria calcular as informações de variabilidade no espaço das imagens e armazená-las através de estruturas compactas como quadrees, que pudessem

ser recuperadas eficientemente. O processo resultante seria dado então por uma combinação de processamento no espaço da cena com processamento no espaço da imagem. Isto é um tema que teremos que investigar futuramente. Abaixo descrevemos o algoritmo para avaliação da foto-consistência das células nos diferentes níveis de refinamento de uma octree (Algoritmo 7) e em seguida apresentamos o algoritmo de escultura do espaço por refinamento adaptativo.

Algoritmo 7 *AvaliaFotoconsistênciaII*(*limiar, nível, rp, Ip, Mvp*)

Determine o conjunto $E_v = \{ev_{st} \in rp \mid \exists n, Mvp_{nst} = 1\}$
se $E_v = \emptyset$ **então**
 retorne CONSISTENTE
fim se
para cada $ev_{st} \in E_v$ **faça**
se $\exists am_{st}^n$ de ev_{st} tal que $Ip_{nst}.a \leq tolfund$ e $nível < NÍVELMÁXIMO$
então
 retorne INDEFINIDA
senão
 retorne INCONSISTENTE
fim se
 $dp_{st} = \frac{1}{n-1} \sum_{l=1}^n (am_{st}^l - \overline{am}_{st})^2$
fim para
 $dpm = \frac{1}{|E_v|} \sum_{s,t} (dp_{st})$
se $dpm < limiar$ **então**
 retorne CONSISTENTE
senão se $nível < NÍVELMÁXIMO$ **então**
 retorne INDEFINIDA
senão
 retorne INCONSISTENTE
fim se

5.4.3

Algoritmo

Algoritmo 8 Escultura do espaço por refinamento adaptativo

$nível \leftarrow octree$

Inicie $nó.célula$ com o volume de reconstrução inicial.

repita

$nível \leftarrow 0$

Limpe os mapas de visibilidade.

para cada plano de varredura π_k , $k = 2^{nível}..0$ **faça**

Determine o conjunto de imagens projetadas I_p na resolução apropriada.

Determine o conjunto de mapas de visibilidade projetados M_{vp} na resolução apropriada.

para cada célula c na octree corrente tal que $(c \cap \pi_k \neq \emptyset)$ e $(c.fconsist = \text{NÃOAVALIADA})$ **faça**

$rp \leftarrow c \cap \pi_k$

$limiar \leftarrow limiarmáximo / 2^{(NÍVELMÁXIMO - nível)}$

$c.fconsist \leftarrow \text{AvaliaFotoconsistênciaII}(limiar, nível, rp, I_p, M_{vp})$

se $c.fconsist = \text{CONSISTENTE}$ **então**

Atribua as cores apropriadas à c .

senão se $c.fconsist = \text{INDEFINIDA}$ **então**

Crie oito nós filhos de $nó$.

Subdivide c em 8 novas células.

Rotule cada nova célula como NÃOAVALIADA .

Atribua as novas células aos respectivos filhos de $nó$.

fim se

fim para

Atualize os mapas de visibilidade.

fim para

$nível \leftarrow nível + 1$

até que que nenhuma célula tenha sido subdividida ou $nível < \text{NÍVELMÁXIMO}$

5.4.4

Observação

É possível tornar o algoritmo acima ainda mais eficiente se o aplicarmos sobre uma estimativa do espaço de reconstrução dada pelo fecho visual representado através de uma octree. Como o fecho visual é utilizado apenas como estimativa, não precisamos determiná-lo através de processamento no espaço da cena, bastando calculá-lo diretamente a partir das imagens de entrada. Um algoritmo bastante eficiente que retorna o fecho visual representado através de uma octree é o algoritmo de Szelisk [29]. Na verdade, em

nossa implementação, adotamos esta estratégia, isto é calculamos o fecho visual no espaço das imagens, através de um algoritmo semelhante ao de Szeliski e então, sobre tal estimativa, aplicamos o algoritmo de escultura do espaço por refinamento adaptativo, o que nos retorna o fecho fotométrico ou *Photo Hull*.