

# 1

## Introdução

O cálculo de correspondências entre imagens estéreo é um problema clássico da área de Visão Computacional e que já é estudado há muito tempo. Ele tem atraído grande interesse de pesquisadores, especialmente pela quantidade de aplicações em computação gráfica e processamento de imagens. Ao longo dos anos muitos algoritmos foram propostos para resolver o problema. Entretanto é uma tarefa bastante árdua desenvolver um algoritmo que apresente bons resultados sem comprometer o desempenho do método.

Dadas duas imagens capturadas a partir de diferentes pontos de vista (no sistema humano, se referem aos olhos esquerdo e direito), o objetivo de um algoritmo de estéreo é identificar pontos em ambas as imagens que se referem ao mesmo ponto na cena. A ideia é determinar a diferença de localização de pontos correspondentes em ambas as imagens.

Em geral os algoritmos de estéreo consideram que as imagens de entrada são retificadas e dessa forma a diferença de localização é calculada apenas pelo deslocamento em  $x$  de um pixel na imagem direita em relação à imagem esquerda. Esse deslocamento é chamado de *disparidade*. Dessa forma a saída de um algoritmo de estéreo consiste em um mapa de disparidades, que é um mapa indicando para cada pixel a diferença de localização deste na imagem direita em relação à imagem esquerda. Através do mapa de disparidades podemos utilizar regras de triangulação para obter o mapa de profundidade. Uma vez que as correspondências entre pixels sejam conhecidas as coordenadas do mundo de cada ponto da imagem podem ser reconstruídas, a menos de um fator de escala que pode ser conhecido pela calibração do sistema de câmeras. A Figura 1.1 mostra duas imagens (esquerda e direita) e o mapa de disparidades da imagem direita em relação à imagem esquerda.

A determinação de um mapa de disparidades acurado continua sendo uma tarefa difícil na prática. Isso acontece devido a três problemas principais enfrentados pelos algoritmos de estéreo e que precisam ser levados em consideração quando a acurácia é essencial, a saber:

- regiões de pouca textura: nesse caso há várias possibilidades de correspondência possíveis;



Figura 1.1: Imagens esquerda e direita e o respectivo mapa de disparidades (01).

- regiões de oclusão: pixels em regiões de oclusão são aqueles que estão presentes em uma imagem mas não estão presentes na outra, o que pode levar a correspondências incorretas se não forem devidamente tratados;
- regiões de descontinuidade: trata-se das bordas dos objetos da cena, onde há mudanças bruscas de profundidade/disparidade.

A reconstrução 3D de uma cena é motivada por diversas aplicações em visão computacional. Por exemplo, Zitnick et al. apresentam em (22) um framework que implementa a renderização de vídeos de alta qualidade usando interpolação de imagens com uma representação em camadas. A Figura 1.2 mostra um diagrama indicando as etapas envolvidas neste processo. Palomo, em seu projeto de mestrado (12), explorou a etapa de renderização com o objetivo de permitir que a mesma seja feita em tempo real, propondo uma implementação em GPU para tal fim.

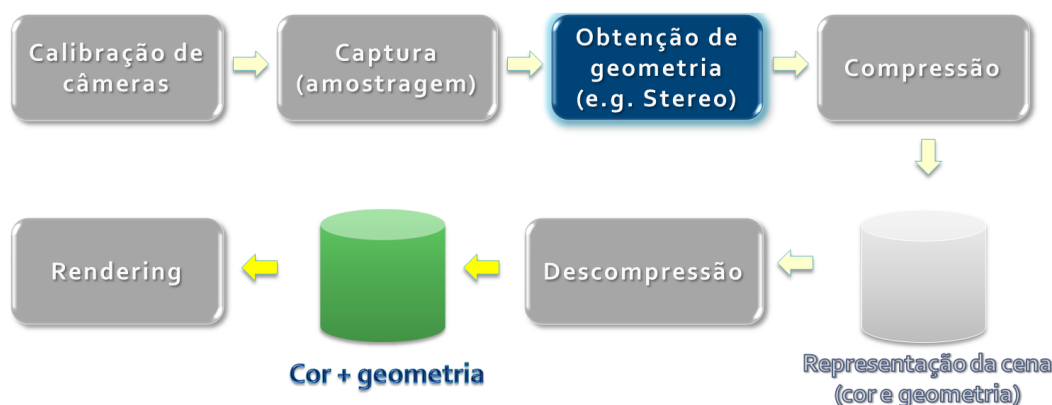


Figura 1.2: Framework do método proposto por Zitnick

Como pode ser visto na Figura 1.2 uma das etapas deste processo é justamente a obtenção da geometria da cena através do cálculo de correspondências de pixels entre duas imagens de referência para a determinação da imagem referente a uma câmera virtual. Nosso trabalho visa desenvolver um algoritmo de estéreo, que entre outras finalidades, possa ser integrado nesse framework.

Em paralelo às técnicas de estéreo, o método de Cortes de Grafo tem se tornado uma ferramenta de otimização bastante poderosa e popular e que tem sido bastante utilizada nas áreas de segmentação de imagens, correspondência por estéreo, restauração de imagens, etc. É utilizado quando o problema em questão pode ser modelado como um problema de minimização de energia. Nesse caso o problema é mapeado em um grafo de maneira que o corte mínimo do grafo minimize a energia total do sistema. Para realizar o cálculo do corte mínimo algoritmos de fluxo máximo (03) em grafos são utilizados. Entretanto a técnica de cortes de grafo pode ser bastante custosa e existem trabalhos que visam reduzir o custo de se utilizar tal algoritmo.

Uma alternativa para melhorar o desempenho de um algoritmo de estéreo consiste em reduzir o tamanho do conjunto de rótulos (valores de disparidades) que são usados pelo algoritmo em cada iteração, conjunto este denominado *espaço de disparidades*, que foi abordada por Veksler (14) e Worby (19). Outra abordagem possível é a técnica de Multi-Resolução, que utiliza o conceito de *Pirâmide Gaussiana*. Trata-se de uma pirâmide de imagens que é construída, na qual cada nível da pirâmide contém uma versão menos amostrada da imagem segundo um fator de escala e um parâmetro de ponderação. Nesse caso uma pirâmide Gaussiana para cada imagem de entrada é construída e o algoritmo de cortes de grafo é utilizado num conjunto restrito de disparidades e em níveis menores, com um número de pixels reduzido.

O propósito desse trabalho é pesquisar e analisar algoritmos de estéreo propostos para o cálculo de correspondências, as técnicas de Cortes de Grafo e Multi-Resolução e desenvolver um novo método de cálculo de correspondências utilizando Cortes de Grafo e Multi-Resolução, de maneira que tenhamos um algoritmo com melhor desempenho que os estudados sem comprometer a acurácia dos resultados produzidos pelos mesmos.

Este trabalho está organizado da seguinte maneira: no Capítulo 2 apresentamos alguns trabalhos que estão diretamente relacionados ao tema e que se mostraram competitivos na solução do problema de estéreo, relacionando suas principais características e evidenciando nossa escolha por essa direção. No Capítulo 3 vamos tratar do problema de minimização de energia e de como este pode ser resolvido utilizando o método de Cortes de Grafo de maneira eficiente. No Capítulo 4 vamos abordar o método de Multi-Resolução e da integração da técnica de Cortes de Grafo a este método visando maior qualidade e melhor desempenho. O Capítulo 5 apresenta os resultados obtidos com a aplicação do método proposto utilizando como entrada os dados disponíveis em (01). Por fim, no Capítulo 6 apresentamos nossas conclusões a respeito do trabalho desenvolvido e sugerimos aprimoramentos para trabalhos futuros.