Universidade Federal de Pernambuco
Centro de Informática

Graduação em Ciência da Computação

# Reconstrução 3D com auxílio de segmentações múltiplas

Rafael Francisco Cavalcanti Campos Gouveia
Proposta de Trabalho de Graduação

Orientador: Sílvio Barros Melo

Recife
Setembro de 2016

## Resumo

Reconstrução 3D no contexto de Visão Computacional sempre foi uma tarefa computacionalmente complexa. Aplicações baseadas nesse tipo de técnica vêm ganhando destaque na indústria nos anos recentes. Navegação autônoma para carros auto-dirigíveis, drones, robôs; óculos de realidade aumentada são exemplos dessas aplicações. Técnicas usando câmeras estéreo são bastante pesquisadas como uma solução barata e mais acessíveis quando comparado com sensores de luz. Este trabalho tem como objetivo apresentar uma nova técnica para uso de câmeras estéreo utilizando múltiplas segmentações das imagens, com objetivo de aumentar a precisão na construção de mapas de disparidade. Em conclusão, o trabalho deve validar experimentalmente o método proposto, inclusive com respeito à sua eficácia, apresentando suas vantagens e desvantagens com relação a técnicas existentes.



**Palavras-chave:** reconstrução 3D, câmeras estéreo, segmentação de imagens.

## Abstract

3D reconstruction in Computer Vision has always been a computationally intensive task. Applications based on this technique have grown substantially in the industry in recent years. Autonomous navigation for cars, drones and robots; augmented reality glasses and many others are examples of such kind of applications. Techniques using stereo cameras are widely solicited for being cheaper and more accessible than light sensors, for instance. This work`s goal is to present a novel technique based on stereo cameras to improve any base disparity maps on a set of image segmentations. In conclusion, it should present experiments that can show how effective the method will perform, with its advantages and disadvantages.



**Keywords:** 3D reconstruction, stereo vision, image segmentation

# Sumário

1. Introdução 1

2. Objetivo 3

3. Cronograma 4

4. Possíveis Avaliadores 5

Assinaturas 6

Referências Bibliográficas 7

# 1. Introdução

Nos últimos anos em Visão Computacional a reconstrução 3D com uso de câmeras *estéreo* vem sendo muito pesquisada na academia. O uso de câmeras se mostra como uma alternativa mais barata quando comparada com uso de sensores de luz. Diversas aplicações fazem uso dessa tecnologia, desde navegação autônoma de carros, *drones*[4] e robôs; à aplicações de realidade aumentada[5] que recentemente vem ganhando destaque. No geral, essas aplicações precisam compreender o ambiente em que estão situados para tomar decisões, seja para desviar de obstáculos ou interagir com objetos reais.

As câmeras além de terem a capacidade coletar mais informações como cor, não tem a limitação de distância que sensores de luz apresentam. Sensores típicos conseguem ver em torno setenta metros e uma resolução inferior à de imagens. As principais desvantagens das câmeras são que na ausência de luz, elas perdem a capacidade de coletar as imagens, e o poder de processamento necessário para fazer a reconstrução.

O conceito de câmeras estéreo foi idealizado com o objetivo de imitar a visão humana e consiste em um conjunto de duas ou mais câmeras separadas por uma distância qualquer. Essas câmeras devem estar alinhadas em um mesmo eixo, e miradas em uma mesma direção. Com essa configuração há uma garantia que dado um pixel da foto da esquerda na linha *r*, sua imagem na foto da direita estará na mesma linha *r,* com a diferença de um deslocamento *d*.

Para a técnica de reconstrução com câmeras estéreo funcionar com perfeição é preciso mapear cada pixel da imagem da câmera da esquerda para seu respectivo pixel da imagem da direita sem erros. Mesmo na configuração estéreo, em que o espaço de busca é reduzido para apenas uma linha da imagem, padrões como cores sólidas e texturas com características repetitivas dificultam o mapeamento. Além disso, mais desafios aparecem com o requisito de aplicações que possam rodar em tempo real com recursos limitados. Um bom exemplo é o de navegação para drones mencionado acima, o drone não pode carregar muito peso, logo deve carregar um hardware limitado, e obviamente precisa de execução em tempo real.

A princípio este trabalho não irá se preocupar com as limitações de hardware e aplicações de tempo real. O foco será em moldar uma técnica nova que possa melhorar a precisão de um mapa de disparidade previamente calculado. Tipicamente métodos com otimização global tendem a superar métodos locais em precisão. Contudo, métodos locais tendem a ser mais rápidos e necessários para as aplicações que precisam ser executadas em tempo real. Em uma tentativa de buscar o meio termo entre essas duas técnicas, métodos que usam segmentação de imagens começam a ser apresentados [1][3]. As segmentações permitem que otimizações sejam feitas

baseada em regiões com contexto semelhante, reduzindo o escopo quando comparado com métodos de otimização global, mas absorvendo mais informação que métodos locais tradicionais.

Esse trabalho propõe testar uma técnica com uso de várias segmentações para aprimorar um mapa disparidade dado como entrada. Em cima de cada uma das segmentações, planos serão estimados com intuito de corrigir ruídos do mapa inicial. Em seguida as informações dos vários planos serão combinadas a fim de produzir um resultado mais preciso. Outros métodos de pós-processamento ainda podem ser aplicados em cima do resultado combinado para melhorar ainda mais a precisão.

# 2. Objetivo

O objetivo deste trabalho é propor um método novo para correção de mapas disparidade com base em múltiplas segmentações da imagem-base, e um mapa de disparidade previamente calculado. A técnica proposta deve dar como saída um novo mapa de disparidade com precisão melhorada quando comparado à entrada. Além disso, também serão feitos experimentos para avaliar a técnica proposta.

# 3. Cronograma

As atividades serão desenvolvidas conforme o cronograma abaixo.

| Atividades | Março | | | | Abril | | | | Maio | | | | Junho | | | | Julho | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Revisão Bibliográfica** | ■ | ■ | | | | | | | | | | | | | | | | | | |
| **Elaboração da Proposta** | | | ■ | ■ | | | | | | | | | | | | | | | | |
| **Implementação das Técnicas** | | | | | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | | | |
| **Experimento** | | | | | | | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | |
| **Escrita do Relatório Final** | | | | | | | | | | | | ■ | ■ | ■ | | | | | | |
| **Preparação da Defesa** | | | | | | | | | | | | | | | ■ | | | | | |
| **Defesa** | | | | | | | | | | | | | | | | ■ | | | | |

# 4. Possíveis Avaliadores

Para os avaliadores do projeto, vejo como possíveis professores, os ligados a área de processamento de imagem, visão computacional e aprendizagem de máquina. Seguem como sugestão:

- Verônica Teichrieb
- Tsang Ing Ren
- Judith Kelner

# Assinaturas

Rafael Francisco Cavalcanti C. Gouveia
Aluno

# Referências Bibliográficas


[1] R. Gouveia, A. Spyropoulos and P. Mordohai. **Confidence Estimation for Superpixel-based Stereo Matching**, *International Conference on 3D Vision (3DV)*, Lyon, France, 2015.

[2] Jure Zˇbontar, Yann LeCun. **Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches**, Journal of Machine Learning Research 17 (2016) 1-32.

[3] Ang Li, Dapeng Chen, Yuanliu Liu, Zejian Yuan *, **Coordinating Multiple Disparity Proposals for Stereo Computation**,* The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4022-4030

[4] Barry, Andrew J., and Russ Tedrake. **Pushbroom stereo for high-speed navigation in cluttered environments.** *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015.

[5] Markovic, Marko, et al. "Stereovision and augmented reality for closed-loop control of grasping in hand prostheses." *Journal of neural engineering* 11.4 (2014): 046001.