



UNIVERSIDADE DO ESTADO DO  
RIO DE JANEIRO



---

INSTITUTO POLITÉCNICO  
GRADUAÇÃO EM ENGENHARIA  
DE COMPUTAÇÃO

Pedro Felipe Pena Barata

Técnicas de reconstrução 3D - ????

Nova Friburgo

2017



## UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO



---

### INSTITUTO POLITÉCNICO GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

Pedro Felipe Pena Barata

#### Técnicas de reconstrução 3D - ?????

Trabalho de Conclusão de Curso apresentado, como requisito parcial para obtenção do título de Graduado em Engenharia de Computação, ao Departamento de Modelagem Computacional do Instituto Politécnico, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Ricardo Fabbri

Nova Friburgo

2017

Pedro Felipe Pena Barata

**Técnicas de reconstrução 3D - ????**

Trabalho de Conclusão de Curso apresentado, como requisito parcial para obtenção do título de Graduado em Engenharia de Computação, ao Departamento de Modelagem Computacional do Instituto Politécnico, da Universidade do Estado do Rio de Janeiro.

Aprovada em 02 de 09 de 2017.

Banca Examinadora:

---

Prof. Dr. Ricardo Fabbri (Orientador)  
Instituto Politécnico – UERJ

---

Banca1  
Instituto Politécnico – UERJ

---

Banca2  
Instituto Politécnico – UERJ

Nova Friburgo  
2017

## **AGRADECIMENTOS**



## RESUMO

BARATA, Pedro Felipe Pena. *Técnicas de reconstrução 3D - ????*. 2017. 20 f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Computação) – Departamento de Modelagem Computacional, Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2017.

A partir dos anos 2000, a área de reconstrução 3D vem sido amplamente explorada. No início, sensores de alcance, tanto aéreos quanto terrestres, eram empregados em diferentes aplicações, devido à facilidade de manuseio e ao baixo custo. Porém, constantes melhorias na tecnologia, sobretudo, nos *hardwares* e *softwares* no âmbito da reconstrução, fizeram com que hoje, quase duas décadas depois, novas técnicas surgissem.

Muitos cientistas que utilizavam a fotogrametria converteram seus esforços na área dos sensores à laser. Pois além de executarem uma reconstrução mais rápida, possuem uma altíssima acurácia, compensando seu alto custo inicial. Isto dificultou e desacelerou o processo de descoberta de novos algoritmos e métodos na área da fotogrametria.

Hoje em dia, graças à esse avanço, a fotogrametria, aliada a novos algoritmos, como o *Structure of Motion (SfM)*, pontos em comum e de combinação de imagens, por exemplo, consegue competir com scanners à laser e sensores de alcance.

ABORDAR O SFM, SIFT E CMVS (POUCO) !!! ????

Com uma combinação de algoritmos, com o *SfM*, junto com o *SIFT* () e o *CMVS* (), é possível gerar uma reconstrução satisfatória apenas utilizando uma câmera de um *smartphone*.

O trabalho foi estruturado da seguinte maneira: previamente apresentam-se os objetivos do projeto, destacando suas funcionalidades e metas, a seguir divide-se em capítulos; O Capítulo 1, que introduz o funcionamento de cada algoritmo e técnica empregada, apresentando e debatendo, comparativamente pontos à favor e contra; O Capítulo 2 é dedicado à ferramenta gráfica utilizada para a obtenção dos resultados (VisualSfM). Finalmente, apresentamos os resultados e conclusões do trabalho, bem como sugestões para implementações e trabalhos futuros.

ABORDAR OS "TODO'S" DO HANGOUTS

Palavras-chave: Reconstrução densa. Nuvem de pontos. *SfM*. ???.

## ABSTRACT

BARATA, Pedro Felipe Pena. . 2017. **20** f. Trabalho de Conclusão de Curso (Graduação em Engenharia de Computação) – Departamento de Modelagem Computacional, Instituto Politécnico, Universidade do Estado do Rio de Janeiro, Nova Friburgo, 2017.

Keywords:

## LISTA DE FIGURAS

Figura 1 - A reconstrução usando-se apenas imagens, sem controle de aquisição, como em um vídeo de um smartphone filmado em torno do objeto, fornece uma nuvem de pontos, que pode ser densificada (??????????), ou atribuída de curvas (??????????), de forma a preservar a resolução em áreas de alto conteúdo informativo. Tais representações estão sendo atualmente unificadas na pesquisa da área. Este projeto propõe explorar os limites da reconstrução 3D usando-se apenas imagens, no contexto de preservação de patrimônio. . . . .	10
Figura 2 - Kinnects de primeira geração (a) consistindo de câmeras e projetores infra-vermelho (b) e de segunda geração, consistindo de tecnologia ToF (c). Ambos os kinnects são largamente utilizados para escaneamento em tempo real, formando a base de scanners manuais (d), porém nem sempre são úteis para preservação detalhada de patrimônio. Um dos objetivos deste projeto é explorar os limites desta tecnologia. . . . .	11
Figura 3 - A reconstrução usando-se Kinnect (de primeira ou segunda geração) usando software atual de super-resolução (c) fornece precisão similar a um sistema estéreo de média resolução, inferior um sistema a laser de alta qualidade (d) porém de baixo custo e muito mais versátil devido ao sistema de aquisição manual e a software amplamente utilizado e desenvolvido(????). . . . .	12
Figura 4 - Protótipo do scanner a laser de triangulação. O objeto a ser escaneado é uma réplica em tamanho real de um sarcófago egípcio (a). O scanner foi reconfigurado para escanear objetos maiores, pois a escultura possui 517 centímetros (b), o da cabeça também sofreu uma reconfiguração, este scanner gira em 90 graus, que faz o laser rotacionar, da posição horizontal para a vertical e também roda em torno da cabeça como um todo (c). Para a reconstrução, o primeiro passo foi alinhar cerca de 100 scans em diversas posições, após isso, utilizado um alinhamento automático em pares dos scans, utilizando um algoritmo modificado de iterações de pontos próximos (ICP - <i>iterated-closest-points</i> ). Após isso, faz-se um processo de relaxação global a fim de minimizar erros de alinhamento por toda a estátua. Depois de alinhados, usa-se o algoritmo de profundidade volumétrica de processamento de imagens (VRIP - <i>volumetric range image processing</i> - do Brian Curless) (d). . .	12
Figura 5 - Algumas esculturas do Jardim do Nêgo . . . . .	13

## SUMÁRIO

	<b>INTRODUÇÃO</b>	9
0.1	<b>Introdução e Justificativa</b>	9
0.1.1	<u>O Jardim do Nêgo, Nova Friburgo</u>	12
0.2	<b>Objetivos</b>	13
0.3	<b>Organização deste manuscrito</b>	14
1	<b>PONTOS DE INTERESSE</b>	15
2	<b>RECONSTRUÇÃO DENSA</b>	16
2.1	<b>Introdução</b>	16
2.2	<b>hpmvs</b>	16
2.2.1	<u>falar sobre o hpmvs...</u>	16
2.3	<b>mve</b>	16
2.3.1	<u>falar sobre o mve...</u>	16
3	<b>KINECT</b>	17
4	<b>VISUALSFM</b>	18
5	<b>EXPERIMENTOS</b>	19
	<b>CONCLUSÃO</b>	20

# INTRODUÇÃO

## 0.1 Introdução e Justificativa

A reconstrução 3D de cenas gerais a partir de múltiplos pontos de vista usando-se câmeras convencionais, sem aquisição controlada, é um dos grandes objetivos de pesquisa em visão computacional, ambicioso até mesmo para os dias de hoje. Aplicações incluem a reconstrução de modelos 3D para uso em videogames (??), filmes (??), arqueologia, arquitetura, modelagem 3D urbana (*e.g.*, Google Streetview); técnicas de *match-moving* em cinematografia para fusão de conteúdo virtual e filmagem real (??), a organização de uma coleção de fotografias com relação a uma cena (*e.g.*, o sistema *Phototourism* (??) e a funcionalidade *Look Around* do Google Panoramio e Street View), manipulação robótica, e a metrologia a partir de câmeras na indústria automobilística e metal-mecânica.

Os desafios estão ligados às escolhas de grande escala de representações adequadas e de técnicas que possam modelar simultaneamente com materiais drásticamente diferentes (*e.g.*, não-Lambertianos), modelos geométricos (*e.g.*, variedades curvilíneas gerais, descontinuidades, texturas, deformações, em escalas diferentes), tipos de regiões (com ou sem textura), condições de iluminação variadas, sombras, fortes diferenças de perspectivas, desbalanceamento devido a excesso de detalhes em partes menos importantes, número arbitrário de objetos e câmeras não-calibradas.

Mesmo que um sistema completo esteja fora do alcance da tecnologia atual, um progresso significativo tem sido atingido nos últimos anos. Por um lado, uma tecnologia operacional tem evoluído, mais recentemente para sistemas de grande escala (????), a partir do desenvolvimento da detecção robusta de *features* (??), o *fitting* robusto e seleção de correspondências baseados em RANSAC, e o desenvolvimento de métodos de geometria projetiva para calibrar duas ou três imagens e progressivamente adicionar imagens e extrair estrutura 3D dessas *features* na forma de nuvens de pontos. Com o código fonte do sistema Bundler (????) liberado por Noah Snavely, e sua subsequente incorporação ao sistema VisualSfM (??), é possível utilizar este sistema para a reconstrução de patrimônio.

No paradigma usando-se apenas imagens convencionais – denominado **reconstrução estéreo multiocular passiva** – a posição das câmeras são estimadas a partir apenas de imagens, usando pontos de interesse, em seguida uma nuvem de pontos é reconstruída ?? . As câmeras podem então ser utilizadas para obter modelos mais detalhados de reconstrução, como algoritmos de densificação (??) e interpolação (??) da nuvem de pontos, bem como demais algoritmos densos de visão estéreo multi-perspectiva/multiocular, como os do grupo de Michel Goesele (??), também com código disponível. Tais algoritmos, no entanto, têm problemas, em particular a reconstrução suaviza partes bem-

delineadas do objeto, e pode conter buracos em áreas homogêneas. Pode-se, portanto, utilizar a reconstrução 3D de curvas do pesquisador proponente (?????????) para auxiliar na reconstrução mais bem-delinada nesses casos problemáticos, bem como para ajudar no problema de escalabilidade quando a reconstrução 3D se torna muito grande. Um segundo paradigma, denominado **reconstrução estéreo multiocular ativa**, tem se tornado viável devido à indústria de videogames, e consiste na utilização de sistemas que alteram o funcionamento de câmeras convencionais, típicamente usando-se projetores infra-vermelho, laser ou câmeras ToF (time of flight), como no caso dos dispositivos Kinect, figura 2.

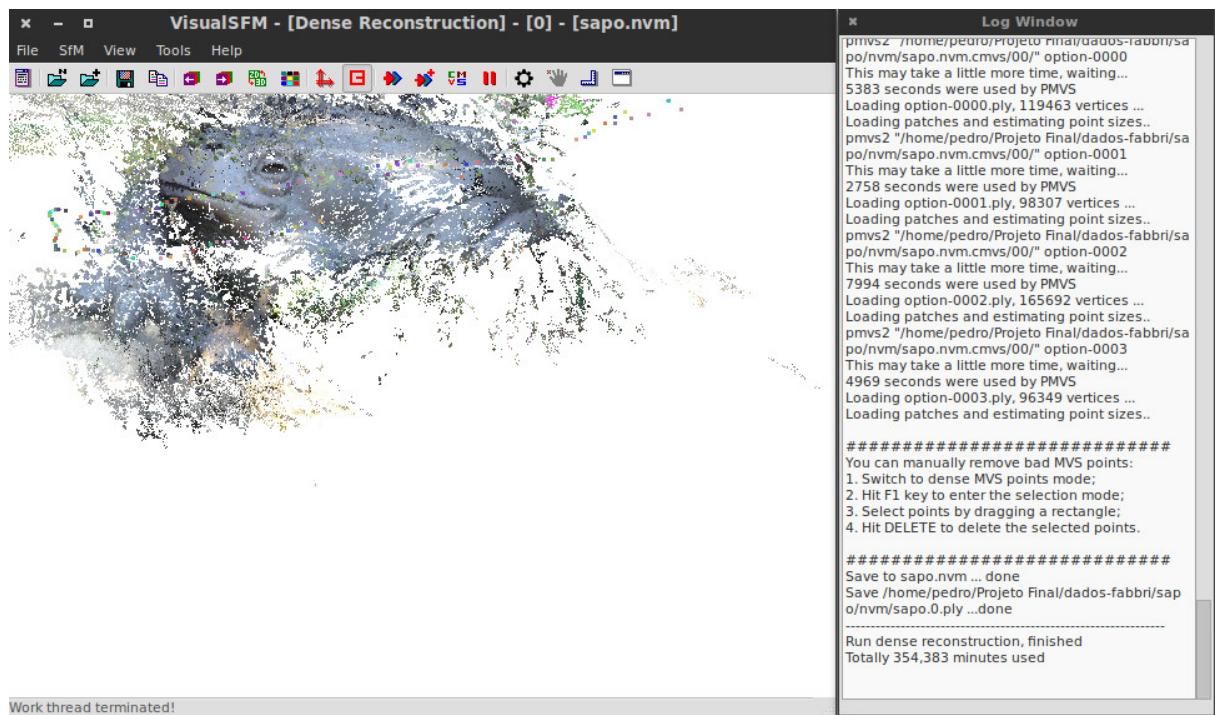


Figura 1 - A reconstrução usando-se apenas imagens, sem controle de aquisição, como em um vídeo de um smartphone filmado em torno do objeto, fornece uma nuvem de pontos, que pode ser densificada (?????????), ou atribuída de curvas (?????????), de forma a preservar a resolução em áreas de alto conteúdo informativo. Tais representações estão sendo atualmente unificadas na pesquisa da área. Este projeto propõe explorar os limites da reconstrução 3D usando-se apenas imagens, no contexto de preservação de patrimônio.

A preservação de patrimônio tem sido realizada tradicionalmente com scanners dedicados de alto custo, como no projeto David 4. O projeto teve início em 1992 e tem como objetivo a utilização de scanners a laser de profundidade (*rangefinder scanners*), aliado com algoritmos que combinam diferentes profundidades e cores da imagem, para realizar uma digitalização da parte externa e da superfície de forma acurada da estátua de David (porém, esse método pode ser utilizado em diferentes objetos no mundo real, como partes de máquinas, artefatos culturais e na indústria de video games, por exemplo). Para

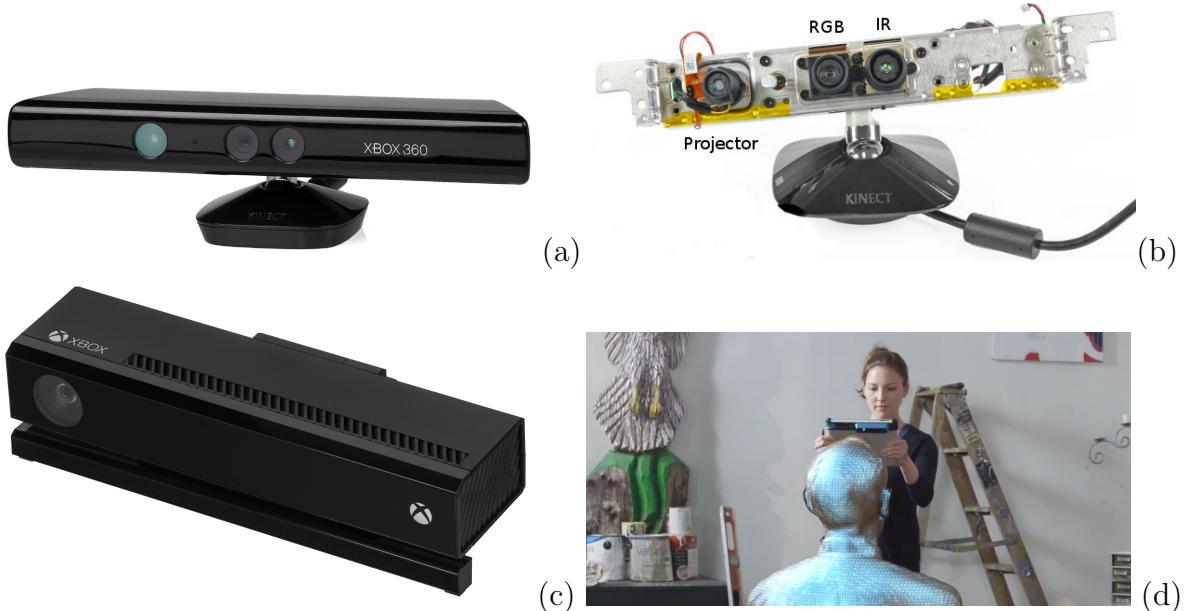


Figura 2 - Kinnects de primeira geração (a) consistindo de câmeras e projetores infra-vermelho (b) e de segunda geração, consistindo de tecnologia ToF (c). Ambos os kinnects são largamente utilizados para escaneamento em tempo real, formando a base de scanners manuais (d), porém nem sempre são úteis para preservação detalhada de patrimônio. Um dos objetivos deste projeto é explorar os limites desta tecnologia.

as partes mais detalhadas, foi utilizado um scanner de menor escala que faz uma pequena triangulação com laser de profundidade.

Seria de grande interesse explorar os dois paradigmas supracitados para avaliar as possibilidades disponíveis no estado da arte de reconstrução 3D para o escaneamento de baixo custo para a preservação de Patrimônio. O que se pode atingir com apenas uma filmagem de esculturas realizada por um smartphone, sem calibração prévia e *in situ*, ou seja, sem ambiente controlado? Como esta reconstrução se compara nos dias de hoje com a reconstrução realizada por um scanner padrão baseado em Kinect?

### 0.1.1 O Jardim do Nêgo, Nova Friburgo

No caso de Nova Friburgo, há a necessidade redobrada de preservação do patrimônio, em especial devido às chuvas e deslizamentos inerentes à região. O Jardim do Nêgo consiste em grandes esculturas em encostas, cobertas por um tapete de vegetação, as quais desfrutam de grande reconhecimento regional e internacional (??), figura 5.

Idealizado e criado por Geraldo Simplicio (Nêgo), artista cearense que mora no local a mais de 30 anos, ganhou notoriedade por suas esculturas de barro, com traços singulares e técnicas únicas. Hoje, trabalha para reconstruir o Jardim da tragédia de 2011 na região serrana, onde algumas estruturas foram destruídas. Portanto, com o

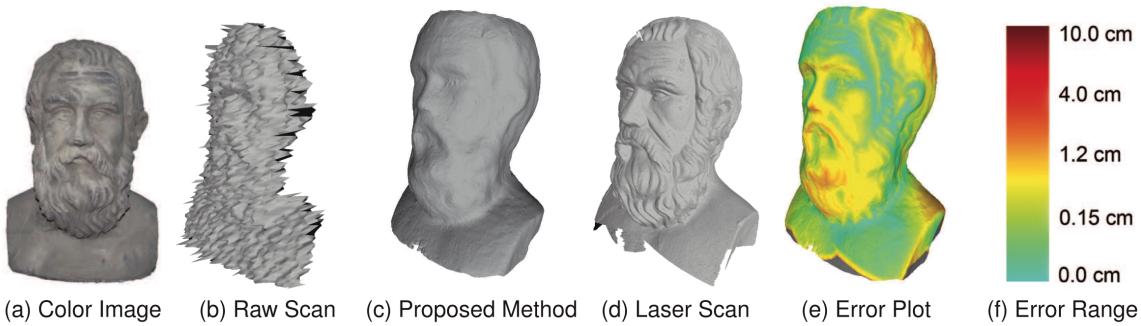


Figura 3 - A reconstrução usando-se Kinect (de primeira ou segunda geração) usando software atual de super-resolução (c) fornece precisão similar a um sistema estéreo de média resolução, inferior um sistema a laser de alta qualidade (d) porém de baixo custo e muito mais versátil devido ao sistema de aquisição manual e a software amplamente utilizado e desenvolvido(????).

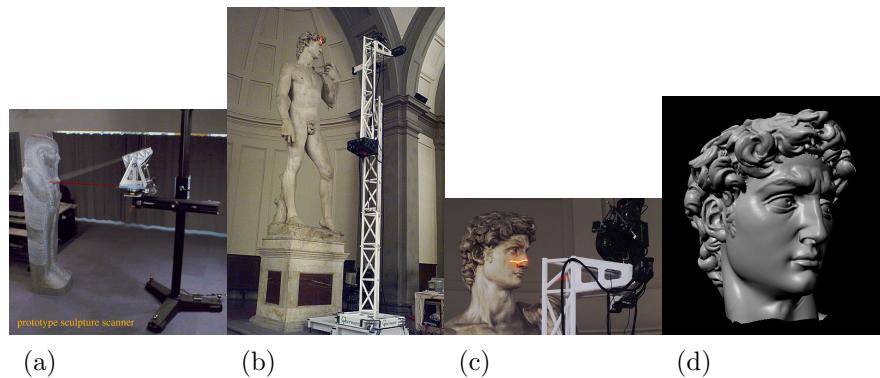


Figura 4 - Protótipo do scanner a laser de triangulação. O objeto a ser escaneado é uma réplica em tamanho real de um sarcófago egípcio (a). O scanner foi reconfigurado para escanear objetos maiores, pois a escultura possui 517 centímetros (b), o da cabeça também sofreu uma reconfiguração, este scanner gira em 90 graus, que faz o laser rotacionar, da posição horizontal para a vertical e também roda em torno da cabeça como um todo (c). Para a reconstrução, o primeiro passo foi alinhar cerca de 100 scans em diversas posicoes, após isso, utilizado um alinhamento automatico em pares dos scans, utilizando um algoritmo modificado de iteracoes de pontos próximos (ICP - *iterated-closest-points*). Após isso, faz-se um processo de relaxação global a fim de minimizar erros de alinhamento por toda a estátua. Depois de alinhados, usa-se o algoritmo de profundidade volumétrica de processamento de imagens (VRIP - *volumetric range image processing* - do Brian Curless) (d).

consentimento do Nêgo, surgiu a motivação desta pesquisa: além de explorar métodos de reconstrução, também tem o objetivo de eternizar um patrimônio que é reconhecido no mundo todo.

A preservação das esculturas do Jardim do Nego se torna um desafio à pesquisa em reconstrução 3D, pois apresentam curvas bem delineadas, que são representadas de maneira



Figura 5 - Algumas esculturas do Jardim do Nêgo

suavizada e empobrecida por métodos convencionais. Algumas esculturas apresentam pouca textura, quase sem nenhum padrão de textura/musgo. Seria de grande interesse availiar o potencial de técnicas atuais de reconstrução 3D geral sem controle de aquisição, as quais têm seu código fonte disponível na internet.

## 0.2 Objetivos

O presente projeto pretende fazer com que o aluno ganhe experiência com técnicas modernas de reconstrução 3D fotogramétrica, no contexto de uma aplicação bem-definida de preservação de patrimônio. A entrada do sistema deverá ser um conjunto de vídeos realizados por câmeras de baixo custo, ou um conjunto de escaneamentos realizados por scanners à mão de baixo custo baseados em Kinect.

O objetivo concreto do aluno será explorar as tecnologias supracitadas para desenvolver um esquema de escaneamento usando software aberto, câmeras e scanners de baixo custo, representando o estado da arte em reconstrução 3D sem restrições de aquisição. Perguntas fundamentais a serem respondidas são: que nível de detalhe, facilidade e precisão se pode obter usando-se apenas imagens e software aberto? É possível utilizar scanners de baixo custo baseados em Kinect com melhorias significativas em termos de qualidade, conveniência ou tempo de processamento? Quais são as restrições desses sistemas? Seria útil na prática uma reconstrução de curvas para auxiliar na reconstrução de nuvem de pontos e de superfícies densas? Onde o estado da arte deve ser avançado de forma a permitir uma solução mais conveniente e completa para a preservação de patrimônio?

O principal objetivo em termos de pesquisa científica será comparar as diferentes abordagens do estado da arte disponíveis para reconstrução 3D e explicitar suas limitações práticas. O aluno deverá, com o entendimento das abordagens, desenvolver um esquema de aquisição de esculturas que permita ampliar os detalhes ou ajudar a resolver os problemas dos métodos. Com a experiência obtida, o aluno estará pronto para desenvolver pesquisa futura na área de reconstrução 3D, com conhecimento de causa para avaliar direções de pesquisa de efetivo e alto impacto na prática.

### 0.3 Organização deste manuscrito

O trabalho foi estruturado da seguinte maneira: previamente introduzimos os métodos baseados em pontos de interesse no Capítulo 1, destacando suas funcionalidades. No Capítulo 2 discutimos e aprofundamos o funcionamento de cada algoritmo de reconstrução densa empregados, apresentando e debatendo, comparativamente, pontos à favor e contra; O Capítulo 3 é apresentada a técnica de reconstrução utilizada nos *Kinects*, da Microsoft; Com isso, temos o Capítulo 4, que é dedicado à ferramenta gráfica utilizada para a obtenção dos resultados (VisualSfM) dos algoritmos de reconstrução densa utilizados. Finalmente, apresentamos os resultados e conclusões do trabalho, bem como sugestões para implementações e trabalhos futuros.

## 1 PONTOS DE INTERESSE

## 2 RECONSTRUÇÃO DENSA

### 2.1 Introdução

### 2.2 hpmvs

#### 2.2.1 falar sobre o hpmvs...

### 2.3 mve

#### 2.3.1 falar sobre o mve...

### 3 KINECT

#### 4 VISUALSFM

## 5 EXPERIMENTOS

## CONCLUSÃO