



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. 30 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) iii ?????

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. 30 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 - Etapa de pré-processamento de vídeo: regiões de interesse (a, c) e correção de gama automática opcional (d) otimizamos esse problema, mas não é necessário no sistema atual. O sinal de onda está concentrado nos primeiros 40/255 níveis de cinza (16%) (b). Veja os materiais suplementares para os vídeos.	23
Figura 2 - Detector de borda de terceira ordem aplicado a um quadro de vídeo. Observe que a informação bruta é puramente localizada nas bordas, com muitos ruídos detectados em artefatos de compressão. O quadro é ajustado em gama (superior), mas nossas experiências mostraram que o detector de borda é robusto o suficiente para funcionar diretamente sem ajuste de gama.	24
Figura 3 - Filtros de união e quebra de contorno geométrico: exemplos de efeitos.	25
Figura 4 - Resultados da amostra de estágios de aprendizagem de máquinas supervisionados.	26
Figura 5 - Os melhores resultados de curva agrupada não supervisionados para reconstrução em 3D e fotogrametria. Os resultados exibem uma alta precisão, mas a cobertura pode ser melhorada (dados faltantes). Veja os vídeos nos materiais suplementares.	27
Figura 6 - Resultados finais de alta cobertura em um quadro de um dos quatro vídeos. Veja o material suplementar para os vídeos.	28

SUMÁRIO

	INTRODUÇÃO	9
0.1	Introdução e Justificativa	9
0.2	Objetivos	11
0.3	Organização deste manuscrito	12
1	PONTOS DE INTERESSE	13
2	HPMVS	14
3	MVE	15
4	KINECT	16
5	VISUALSFM	17
6	EXPERIMENTOS	18
6.1	Arquitetura de Implementação	18
6.2	Processamento inicial	19
6.3	Processamento de Imagem	19
6.4	Detecção de borda	20
6.5	Extração do contorno inferior	21
6.6	Extração de contorno baseada em aprendizagem	22
6.7	Resultados principais	22
	CONCLUSÃO	29

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 Steet 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 drasticamente 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, tipicamente usando-se projetores infra-vermelho, laser ou câmeras ToF (time of flight), como no caso dos dispositivos Kinect, figura ??.

Diversos scanners a laser ou ToF que há alguns anos possuíam um custo elevado, hoje são acessíveis. Muitos deles são baseados nos sensores Kinect, e há software berto disponível na internet para realizar a calibragem externa e reconstrução 3D usando-se tais dispositivos, incluindo a fusão de mapas de profundidade para obtenção de reconstruções 3D com precisão razoável, mesmo que aquém dos sistemas de alto custo, figura ??. No entanto, comparando-se com câmeras convencionais, o hardware ainda é especializado e o custo, apesar de baixo, é uma ordem de magnitude maior que o de câmeras convencionais.

A preservação de patrimônio tem sido realizada tradicionalmente com scanners dedicados de alto custo, como no projeto David ??.

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 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?

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 ??.

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 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 suavizada e empobrecida por métodos convencionais. Algumas esculturas apresentam pouca textura, quase sem nenhum padrão de textura/musgo. Seria de grande interesse avaliar 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

COLOCAR A ORGANIZACAO AQUI iii

1 PONTOS DE INTERESSE

2 HPMVS

3 MVE

4 KINECT

5 VISUALSFM

6 EXPERIMENTOS

CONCLUSÃO