

# PHOTOSCAN X COLMAP: COMPARAÇÃO ENTRE DEMS GERADOS POR STRUCTURE FROM MOTION–MULTI-VIEW STEREO

Francesco Barale<sup>1</sup>, Carlos H. Grohmann<sup>2</sup>

<sup>1</sup>Instituto de Geociências (IGc) - Universidade de São Paulo (USP), São Paulo, 05508-080, Brasil, francesco.barale@usp.br;

<sup>2</sup>Instituto de Energia e Ambiente (IEE) - USP, São Paulo, 05508-010, Brasil, guano@usp.br;

Todos autores: Spatial Analysis And Modelling Lab (SPAMLab) – IEE-USP

## RESUMO

O objetivo deste trabalho é avaliar Modelos Digitais de Elevação (MDEs) gerados por dois programas de processamento de imagens por Structure from Motion–Multi-View Stereo (SfM-MVS), um líder de mercado, proprietário, AgiSoft PhotoScan, e outro Livre e de Código Aberto, COLMAP. Os dados foram processados e geraram uma nuvem de pontos distinta para cada software. O MDE gerado com a nuvem de pontos processada pelo Photoscan foi considerado de melhor qualidade por estar visualmente mais próxima da realidade que o MDE do COLMAP. É preciso salientar o fato de que o COLMAP não utilizou a GPU do computador, aumentando muito o tempo de processamento, nem a informação de posição das imagens, resultando em uma nuvem de pontos densa com um bom nível de detalhe, mas sem referência espacial.

**Palavras-chave** – Structure from Motion, Photoscan, COLMAP, DEM, GPU.

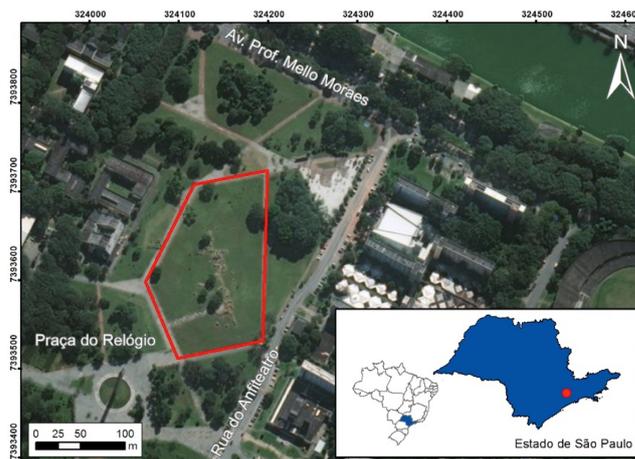
## ABSTRACT

*The objective of this work is to evaluate Digital Elevation Models (DEMs) generated by two Structure of Motion–Multi-View Stereo image processing programs, a market leader, proprietary, AgiSoft PhotoScan, and a Free and Open Source alternative, COLMAP. The data were processed and generated a distinct point cloud for each software. The DEM obtained with the point cloud generated by Photoscan was considered to be of higher quality due its better visual representation of the reconstructed scene than the DEM computed from the COLMAP point cloud. It is important to note that COLMAP did not use the computer's GPU, increasing processing time, nor the position information available from the images, resulting in a dense cloud with good level of detail, but without any spatial reference.*

**Key words** – Structure from Motion, Photoscan, COLMAP, DEM, GPU.

## 1. INTRODUÇÃO

Structure from Motion–Multi-View Stereo (SfM-MVS) é uma técnica de imageamento de intervalo (*range imaging*) estudada no campo de Visão Computacional e Percepção Visual, e trata da reconstituição de uma imagem tridimensional a partir de uma série de imagens obtidas por um sensor em movimento [1, 2]. Essa técnica vem ganhando espaço nas Geociências pela sua facilidade de uso



**Figura 1:** Localização da Praça do Relógio, Campus da Capital da USP. Coordenadas UTM, zona 23 (Sul), WGS84.

e baixo custo, uma vez que pode-se trabalhar com câmeras fotográficas simples e sem necessidade de rigor geométrico na aquisição das imagens [2].

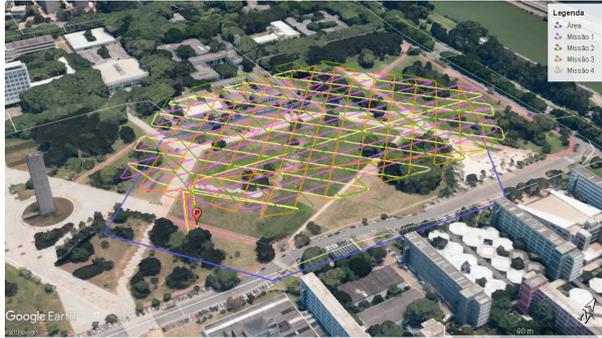
A diferença fundamental entre SfM e a fotogrametria tradicional está no fato de que em SfM a geometria da cena, a posição das câmeras e suas orientações são resolvidas automaticamente sem a necessidade de especificar *a priori* uma série de dados com coordenadas 3D conhecidas. Esses dados são resolvidos simultaneamente através de um procedimento iterativo altamente redundante de ajuste de conjunto (*bundle adjustment*) fundamentado em uma base de dados extraída automaticamente de um conjunto de imagens sobrepostas. O método tem melhor funcionamento em conjuntos de imagens com grandes áreas de sobreposição que capturem toda a estrutura tridimensional da cena vista de uma variada ordem de posições, ou como o nome sugere, de um sensor em movimento [1].

O objetivo deste trabalho é avaliar Modelos Digitais de Elevação (MDEs) gerados por dois programas de processamento de imagens por SfM-MVS, um líder de mercado, proprietário, AgiSoft PhotoScan, e outro Livre e de Código Aberto, COLMAP.

## 2. MATERIAIS E MÉTODOS

### 2.1. Área de Estudo

O projeto teve como área de estudo a Praça do Relógio da Universidade de São Paulo - USP, no município de São Paulo (Fig. 1). A escolha da área levou em consideração a facilidade de acesso e a proximidade com o IEE-USP.



**Figura 2: Vista 3D das missões executadas pelo VANT para aquisição das imagens.**

A área imageada compreende um polígono de 21.260 m<sup>2</sup>, com cinco vértices, localizado na região centro norte da Praça do Relógio. O relevo é constituído por uma colina baixa com amplitude de aproximadamente 1,50 m, cobertura vegetal rasteira e árvores dispersas de porte médio. Na região central ocorre acúmulo de blocos rochosos.

## 2.2. Sobrevoos

O levantamento aéreo foi realizado no dia 12/03/2018, com duração de voo de 2h e 45min no período da manhã. O equipamento utilizado foi um VANT DJI Phantom 4 Pro com câmera de 20 MP e lente de 8,8 mm. A atividade contou com quatro missões com percursos e ângulos de câmera distintos, obtendo um total de 812 imagens. Para o planejamento das missões, utilizou-se do aplicativo MapPilot [3], que permite programar o voo com altura constante sobre o nível de um MDE SRTM, garantindo maior consistência do tamanho do pixel das fotografias. A Fig. 2 apresenta uma visada em perspectiva com todas as linhas de voo e indicação do ponto de partida (P), e a Tabela 1 detalha os parâmetros de cada missão.

## 2.3. Processamento dos Dados

Os dados foram processados em um computador com sistema Windows 10 Pro (64 bits), processador Intel Core i7-7700 4.2 GHz, memória RAM 64GB e placa de vídeo Geforce GTX 1080Ti com 11GB de memória RAM. A reconstrução por SfM-MVS foi realizada nos softwares Agisoft Photoscan Pro 1.4.4 [4] e COLMAP 3.5 [5]. A interpolação das nuvens de pontos e análise dos MDEs foi realizada com GRASS-GIS 7.2 [6, 7].

Para georreferenciamento, apenas as informações do GPS de navegação a bordo do UAV foram utilizadas, disponíveis através dos metadados das imagens.

# 3. RESULTADOS E DISCUSSÕES

## 3.1. Nuvens de Pontos

O processamento dos dados para geração da nuvem densa de pontos seguiu o fluxo de passos de cada programa, sempre com definições para qualidade alta (*high*) do modelo final. O Photoscan processou o conjunto de dados em 12h e gerou uma nuvem densa com  $\sim 147 \times 10^6$  pontos.

**Tabela 1: Parâmetros das missões realizadas. O ângulo da câmera tem como referência o plano horizontal.**

Missão	Alt. Voo	Azimute	Ang. Câmera	Recobrim.
1	40 m	30°	-90°	70%-70%
2	40 m	300°	-90°	70%-70%
3	40 m	330°	-75°	70%-70%
4	40 m	250°	-75°	70%-70%

O COLMAP levou cerca de 32 horas para o processamento e gerou uma nuvem densa com  $\sim 66,5 \times 10^6$  pontos. É importante notar que o Photoscan utiliza a GPU do computador, quando existente, e que o COLMAP, apesar de indicar a utilização também, aparentemente não o fez, tendo em vista o elevado tempo de processamento.

Ao final do processo de nuvens de pontos foi observado o tamanho de cada pasta criada para armazenar os dados gerados. A pasta com os dados do Photoscan que inicialmente continha apenas as imagens obtidas pelo VANT (cerca de 6,7 GB) gerou 2,95 GB a mais de dados, que ao final do processo foram compactados automaticamente. A pasta com os dados do COLMAP que inicialmente continha apenas as imagens obtidas pelo VANT gerou 308 GB a mais de dados que ao final do processo não foram compactados.

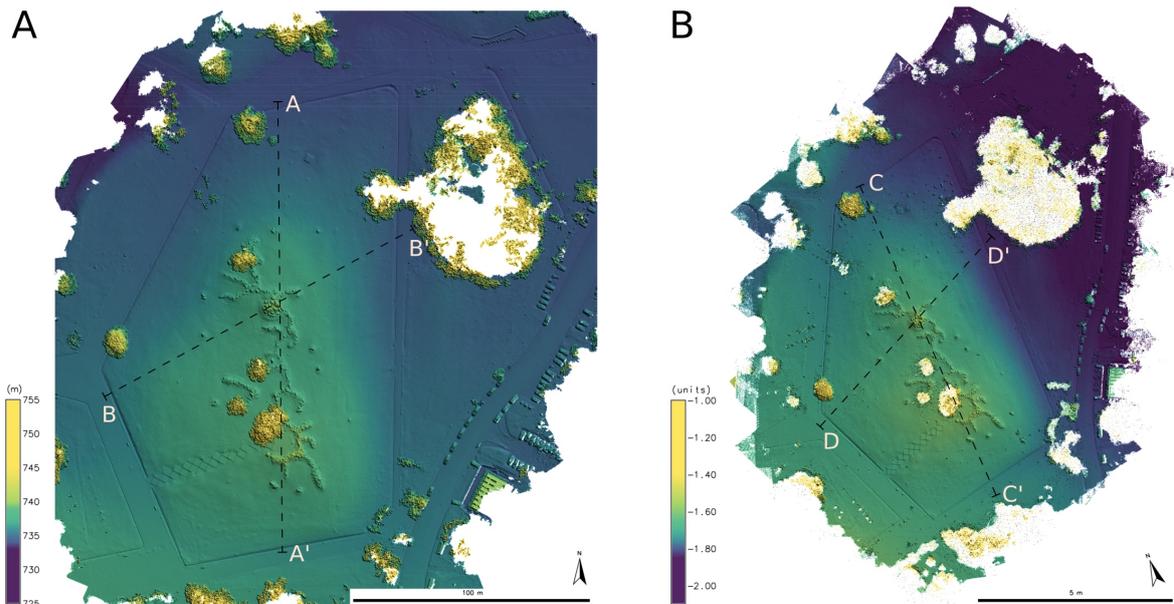
Notou-se que a cada processo executado pelo COLMAP, há duplicação dos arquivos das imagens e uma nova pasta é gerada, contendo grande volume de dados. Também foi observado que ao gerar a nuvem de pontos densa, as 805 imagens utilizadas que inicialmente tinham um tamanho médio de 5,5 MB, passaram por dois processos de reamostragem que aumentaram seu tamanho até 10,5 MB.

A nuvem de pontos gerada pelo COLMAP ainda precisou de um tratamento adicional, pois os dados apresentavam uma inversão no valor das coordenadas Z e uma rotação de 180° das coordenadas X e Y. O ajuste das coordenadas foi realizado no software CloudCompare 2.9.1 [8]. As nuvens densas geradas pelo Photoscan e pelo COLMAP (rotacionada) foram salvas em formato XYZ ASCII. Enquanto que no Photoscan os dados puderam ser exportados com coordenadas UTM, no COLMAP essa opção não é possível, e os dados são exportados sem referência espacial, apenas com valores XYZ arbitrários usados internamente para a reconstrução do modelo.

## 3.2. Modelos Digitais de Elevação

As nuvens de pontos foram importadas para o GRASS-GIS com o comando `r.in.xyz`, que permite importar uma grande quantidade de pontos vetoriais diretamente como um *raster*, ao executar uma operação matemática definida pelo usuário (média, mínimo, máximo etc) para os pontos que estiverem dentro da área de cada célula (*pixel*) do *raster* a ser gerado. Utilizou-se uma operação de média dos valores de elevação dos pontos com células de 10 cm (ou o valor aproximado equivalente, no caso dos dados oriundos do COLMAP).

Os MDEs obtidos são apresentados na Fig. 3. Observa-se que ambos modelos apresentam áreas sem informação, identificadas nas figuras pelas regiões em branco, e correspondentes à áreas de vegetação mais densa e de difícil reconstrução.



**Figura 3: MDEs obtidos pelo processamento dos dados gerados pelo Photoscan (A) e COLMAP (B).**

O MDE gerado a partir da nuvem de pontos do Photoscan (MDE-Photoscan) apresenta um refinamento nas regiões geométricas, praticamente sem ruído. Essa observação pode ser feita na região sudeste da praça, onde há certa simetria dos elementos paisagísticos, como rochas e calçamento.

O MDE gerado a partir dos dados do COLMAP (MDE-COLMAP) apresenta uma quantidade significativa de ruídos e falta de composição, principalmente nos limites da área. Entretanto, observa-se que localmente há maior detalhamento no modelo, tal como em áreas com pequenas árvores no meio do calçamento da praça e pequenas porções de grama próximas a um quadrado de concreto e ao seu redor, na região norte da praça.

Dois pontos relevantes com relação ao MDE-COLMAP devem ser ressaltados. Um deles se refere a escala de elevação do MDE que apresenta valores negativos. O outro ponto se refere a escala de distâncias não condizente com a metragem real da Praça do Relógio, que resultaram no valor de 5 “metros” (unidades arbitrárias) para um trecho de aproximadamente 100 metros. Não foi possível encontrar uma justificativa para esses problemas e nem uma forma de resolvê-los.

Para uma análise visual da qualidade dos modelos, foram elaborados perfis de elevação para os MDEs, cruzando a área da praça em duas direções (Fig. 4).

Os perfis de elevação do MDE-Photoscan estão condizentes com a escala gráfica e com a topografia do terreno. No perfil A-A' o ponto mais alto condiz com os blocos de rochas no centro da praça. No perfil B-B' o pico inicial corresponde a uma pequena árvore, seguida de pequenas elevações até atingir as mesmas rochas do perfil A-A' onde os perfis se cruzam.

Os perfis de elevação do MDE-COLMAP ilustram os problemas com relação as escalas vertical e horizontal dos dados, apresentadas em “unidades de medida”, e não em metros. Apesar da diferença das unidades, os perfis são relativamente similares aos do MDE-Photoscan, com exceção da área correspondente aos blocos de rocha no perfil C-C',

que exibe valores abaixo do nível do solo.

No histograma do MDE-Photoscan ocorrem dois picos de células entre 730 m e 740 m (Fig. 5A). Os valores entre 732 e 734 metros representam a área mais plana da praça. O pico menor, em aproximadamente 737 m, representa as rochas no centro da praça. As pequenas irregularidades observadas entre 738 m e 755 m são condizentes com o dossel das árvores da praça.

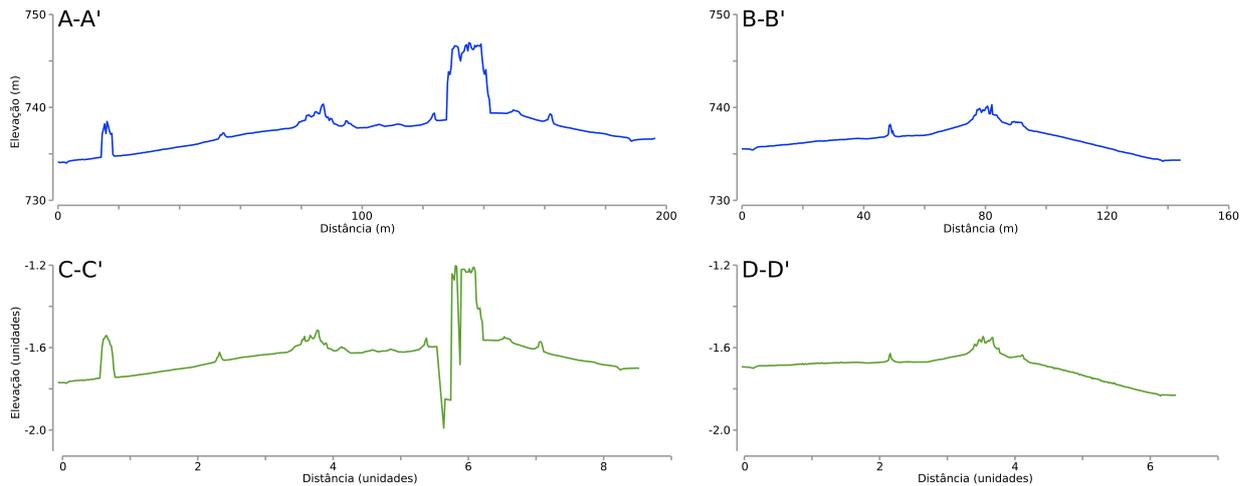
O histograma do MDE-COLMAP (Fig. 5B) apresenta um comportamento geral similar ao do MDE-Photoscan, com um pico principal em  $\sim -1,8$  unidades, correspondente à área plana da praça, e um pico secundário em  $\sim -1,6$  unidades, para as rochas na região central. O pico secundário neste caso é mais pronunciado do que no histograma do MDE-Photoscan.

#### 4. CONCLUSÕES

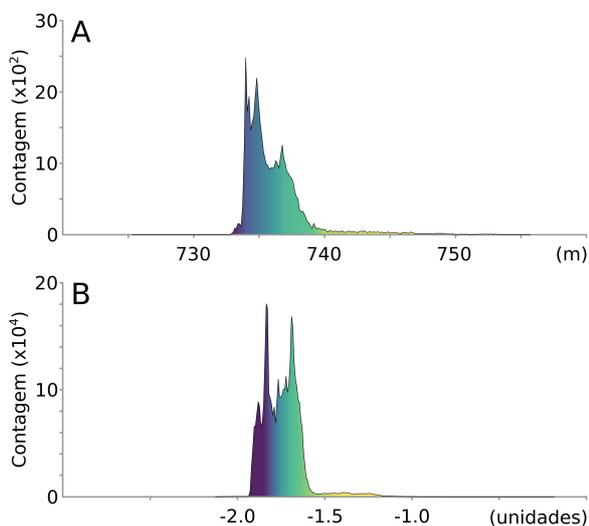
Este trabalho buscou realizar uma comparação entre MDEs derivados de nuvens de pontos geradas pelos programas AgiSoft Photoscan e COLMAP, com um conjunto de imagens obtidas por VANT na área da Praça do Relógio, Campus da Capital da USP.

Quanto ao processamento dos dados, nota-se que o COLMAP gera uma grande quantidade de arquivos secundários (308 GB neste caso), muitas vezes com duplicação de dados, enquanto que o Photoscan gerou um volume menor de dados secundários (2,95 GB). Além disso, o COLMAP aparentemente não utilizou a GPU do computador, requerendo pouco mais que o dobro do tempo de processamento do Photoscan.

A nuvem densa gerada pelo Photoscan tem  $\sim 147 \times 10^6$  pontos, um pouco mais que o dobro da quantidade de pontos gerados pelo COLMAP ( $\sim 66,5 \times 10^6$  pontos). Ainda, a nuvem gerada pelo COLMAP não trouxe a referência espacial, mas apenas unidades usadas internamente para a reconstrução do modelo.



**Figura 4: Perfis morfológicos dos MDEs. A-A' e B-B': Photoscan; C-C' e D-D': COLMAP. Localização e orientação dos perfis indicada na Fig. 3.**



**Figura 5: Histogramas dos MDEs: A) Photoscan; B) COLMAP.**

Os MDEs gerados são similares em seu aspecto geral, mas o MDE-Photoscan apresentou menos ruídos que o MDE-COLMAP, principalmente em feições geométricas existente no terreno, enquanto que o MDE-COLMAP apresentou maior nível de detalhamento do que o MDE-Photoscan em alguns locais. Este aspecto pode estar relacionado ao tratamento dos ruídos pelo Photoscan que, em contrapartida, reduz as informações de detalhe.

#### AGRADECIMENTOS

Este projeto tem apoio FAPESP (proc. #2016/06628-0). F.B. é graduando em Geologia pelo IGc-USP e bolsista PIBIC-CNPq; G.P.B.G. é doutorando pelo RMH IGc-USP; R.W.A. é doutorando pelo PROCAM IEE-USP; C.H.Grohmann é

bolsista de Produtividade CNPq (proc. #307647/2015-3). O trabalho foi realizado com apoio parcial da CAPES Brasil - Código de Financiamento 001. SPAMLab: <<https://spamlab.github.io>>.

#### 5. REFERÊNCIAS

- [1] WESTOBY, M. J. et al. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, v. 179, p. 300–314, 2012. ISSN 0169-555X. Available at: <<http://www.sciencedirect.com/science/article/pii/S0169555X12004217>>.
- [2] CARRIVICK, J. L.; SMITH, M. W.; QUINCEY, D. J. *Structure from Motion in the Geosciences*. Wiley, 2016. (Analytical Methods in Earth and Environmental Science). ISBN 9781118895832. Available at: <<https://books.google.com.br/books?id=tZCvDAAAQBAJ>>.
- [3] Drones Made Easy. *Map Pilot for DJI*. 2017. Available at: <<http://www.dronesmadeeasy.com/Articles.asp?ID=254>>.
- [4] AGISOFT. *Agisoft PhotoScan User Manual Professional Edition v. 1.4*. 2018. Available at: <[http://www.agisoft.com/pdf/photoscan-pro\\_1\\_4\\_en.pdf](http://www.agisoft.com/pdf/photoscan-pro_1_4_en.pdf)>.
- [5] SCHÖNBERGER, J. L.; FRAHM, J.-M. Structure-from-Motion Revisited. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016.
- [6] NETELER, M. et al. GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, v. 31, p. 124–130, 2012. ISSN 1364-8152. Available at: <<http://www.sciencedirect.com/science/article/pii/S1364815211002775>>.
- [7] GRASS Development Team. *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.2*. [S.l.], 2017. Available at: <<http://grass.osgeo.org>>.
- [8] GIRARDEAU-MONTAUT, D. *CloudCompare (version 2.9.1 Omnia)*. 2018. Available at: <<https://www.danielgm.net/cc/>>.