

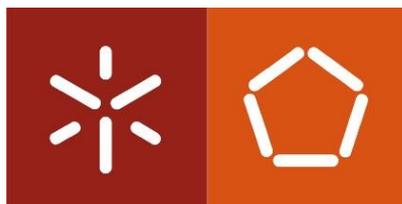


**Universidade do Minho**  
Escola de Engenharia

Márcio Adriano Mendes Contina Inês

**Digitalização 3D por meio de  
Fotogrametria: uma análise comparativa**

Novembro de 2022



**Universidade do Minho**

Escola de Engenharia

Márcio Adriano Mendes Contina Inês

**Digitalização 3D por meio de  
fotogrametria: uma análise comparativa**

Dissertação de Mestrado

Trabalho realizado sob a orientação do Professor:

**Nuno Miguel Faria Araújo**

e coorientação do Professor:

**Adrián José Riquelme Guill**

Novembro de 2022

## **Direitos de autor e condições de utilização do trabalho por terceiros**

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos. Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada. Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



**Atribuição**  
**CC BY**

<https://creativecommons.org/licenses/by/4.0/>

## **Agradecimentos**

Em primeiro lugar, agradeço à Deus, toda honra e glória, pela bênção e proteção durante todo o percurso da minha vida, no qual agora me permite terminar mais um capítulo da minha história.

Ao meu pai, à minha mãe, a minha irmã e especialmente ao meu avô que infelizmente não está entre nós para se orgulhar de mais uma conquista, por toda a ajuda financeira durante esses anos e acima disso por todo o amor, conselho e confiança depositada em mim.

Ao professor orientador, Nuno Araújo, por todos os ensinamentos, toda a disponibilidade, humildade, paciência e ajuda mostrada durante a realização desta dissertação.

Ao professor coorientador, Adrián Guill, agradeço a sua disponibilidade e apoio na realização do modelo no Metashape, que foi fundamental para o estudo realizado.

Ao Eng. Hélder Torres pelo empréstimo do drone utilizado no levantamento aerofotográfico.

A todos os meus colegas e amigos, que sempre estiveram comigo, por partilhar essa luta, por todas as conversas, conselhos e momentos de diversão.

Agradeço a todos os meus familiares pelos incentivos e carinho demonstrado durante esta fase.

A todo o Departamento de Engenharia Civil da Universidade do Minho e todos os docentes com quem eu tive o gosto e a oportunidade de trabalhar, um especial obrigado.

A todos, um muito obrigado.

## **Declaração de integridade**

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração. Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

## **Digitalização 3D por meio de fotogrametria: uma análise comparativa**

### **Resumo**

A reconstrução 3D de objetos ou cena usando um scanner a laser ou um software de modelagem 3D geralmente é caro e requer um alto nível de especialização. O programa de modelação 3D baseado em fotogrametria, por outro lado, oferece uma alternativa mais barata, que pode lidar com essa tarefa com relativa facilidade. Também existe programas gratuitos e de código aberto, com potencial para fornecer dados de qualidade para diferentes fins pretendidos. No entanto, o discurso acadêmico contemporâneo raramente apresenta listas de recursos baseadas em pesquisas ou uma inspeção crítica de potenciais “pipelines” de produção, nem normalmente fornece direção e orientação para não especialistas interessados em aprender, desenvolver e compartilhar conteúdo 3D com um orçamento restrito.

Dentro deste contexto, a presente dissertação tem como objetivo desenvolver um estudo de um conjunto de programas gratuitos com base em seus recursos oferecidos, fluxo de trabalho, tempo de processamento 3D e precisão. Primeiramente dois conjuntos de dados (provetes de granito e fachada do edifício da Associação Académica da Universidade do Minho) são utilizados de modo a explorar as potencialidades dos diferentes programas gratuitos, e em seguida o nosso caso de estudo (maciço rochoso, no qual se insere o Estádio Municipal de Braga) utilizado para comparar e avaliar os programas com base nas nuvens de pontos que eles produziram.

A distância absoluta para realizar a comparação com os dados de verdade do terreno produzidos por um programa comercial (Metashape) foi usado e medido com o software CloudCompare. As reconstruções 3D geradas a partir de programas gratuitos produzem resultados promissores, com precisão significativa e são fáceis de usar. Acreditamos que esta dissertação ajudará usuários não especialistas a entender a fotogrametria e selecionar o programa mais adequado para a produção de modelos 3D baseados em imagens a baixo custo para fins de visualização, apresentação e exploração de várias outras potencialidades com base nas nuvens de pontos.

**Palavras-chave:** análise comparativa, fotogrametria, nuvem de pontos, reconstrução 3D.

## **3D scanning through photogrammetry: a comparative analysis**

### **Abstract**

3D reconstruction of objects or scenes using a laser scanner or 3D modeling software is often expensive and requires a high level of expertise. The photogrammetry-based 3D modeling program, on the other hand, offers a cheaper alternative that can handle this task with relative ease. There are also free and open-source programs with the potential to provide quality data for different intended purposes. However, contemporary academic discourse rarely presents research-based resource lists or a critical inspection of potential production pipelines, nor does it typically provide direction and guidance for non-experts interested in learning, developing, and sharing 3D content on a tight budget.

Within this context, the present dissertation aims to develop a study of a set of free programs based on their offered features, workflow, 3D processing time and accuracy. Firstly, two sets of data (granite test piece and facade of the building of the Academic Association of the University of Minho) are used in order to explore the potential of the different free programs, and then our case study (rock massif, in which insert the Braga Municipal Stadium) used to compare and evaluate the programs based on the point clouds they produced.

The absolute distance to perform the comparison with real terrain data produced by a commercial program (Metashape) was used and measured with CloudCompare software. 3D reconstructions generated from free programs produce promising results with significant accuracy and are easy to use. We believe that this dissertation will help non-expert users to understand photogrammetry and select the most suitable program for producing image-based 3D models at low cost for visualization, presentation and exploration of various other potentialities based on point clouds.

**Keywords:** comparative analysis, photogrammetry, point cloud, 3D reconstruction.

## Índice

Direitos de autor e condições de utilização do trabalho por terceiros.....	ii
Agradecimentos.....	iii
Declaração de integridade .....	iv
Resumo.....	v
Abstract.....	vi
Índice.....	vii
Lista de figuras.....	ix
Lista de tabelas .....	xi
Lista de siglas .....	xii
Introdução.....	1
Enquadramento e motivação .....	1
Objetivos.....	3
Estrutura e organização da dissertação .....	3
1. Estado do conhecimento .....	5
1.1. Fotogrametria.....	5
1.2. Introdução SFM-MVS .....	7
1.3. Fluxo de Trabalho SFM-MVS .....	7
1.3.1. Detecção de recursos .....	9
1.3.2. Correspondência de ponto-chave .....	14
1.3.3. Identificação de correspondências consistentes.....	15
1.3.4. “Structure from Motion” .....	16
1.3.5. Escala e georreferenciação .....	19
1.3.6. Otimização do alinhamento de imagem.....	20
1.3.7. Agrupamento para MVS.....	20
1.3.8. Algoritmos de correspondências de imagem MVS.....	21
1.4. Sensores.....	24
1.5. Vantagens e Desvantagens do SfM-MVS.....	28
2. Materiais e Métodos .....	30
2.1. Software SFM.....	31
2.1.1. VisualSfM .....	31

2.1.2.	Regard3D .....	32
2.1.3.	Colmap .....	34
2.1.4.	Meshroom .....	35
2.1.5.	Autodesk Recap Photo (2021) .....	37
2.1.6.	Agisoft Metashape .....	39
2.2.	Caso do Estudo .....	40
2.2.1.	Enquadramento Geográfico e Descrição do Talude .....	44
2.2.2.	Levantamento Aerofotográfico .....	46
2.2.3.	Georreferenciação .....	47
2.2.4.	Modelação .....	48
2.2.5.	Extração de elementos (CloudCompare) .....	49
3.	Resultados e Discussão .....	55
3.1.	Provetes de Granito .....	55
3.2.	Fachada da sede da Associação de Estudantes da Universidade do Minho .....	59
3.3.	Maciço Rochoso .....	60
3.3.1.	Análise Comparativa (Metashape) .....	65
3.4.	Experiências do usuário .....	69
4.	Aplicação Prática .....	71
4.1.	Volume .....	71
4.2.	Curva de níveis .....	72
5.	Conclusão .....	77
6.	Bibliografia .....	79
7.	Anexos .....	86

## Lista de figuras

<b>Figura 1.</b> Tipos de VANT (Carrivick et al. 2016).....	7
<b>Figura 2.</b> Fluxo de trabalho (adaptado de Carrivick et al. 2016).....	8
<b>Figura 3.</b> Desafios enfrentados pelos algoritmos de correspondência de recursos.....	10
<b>Figura 4.</b> Regiões elípticas detetadas independentemente em cada ponto de vista (Mikolajczyk et al. 2005).....	11
<b>Figura 5.</b> Etapas de SIFT (Lowe, 2004).....	12
<b>Figura 6.</b> Comparação de correspondência de pontos-chave (Carrivick et al. 2016) .....	13
<b>Figura 7.</b> Demonstração da matriz de descritores $2 \times 2$ calculada a partir de um conjunto de amostras $8 \times 8$ (Lowe, 2004) .....	14
<b>Figura 8.</b> O algoritmo CMVS leva imagens $\{I_i\}$ , pontos SfM $\{P_i\}$ e suas informações de visibilidade associadas $\{V_i\}$ para produzir clusters de imagens sobrepostas $\{C_i\}$ (adaptado de Furukawa et al. 2010) .....	21
<b>Figura 9.</b> Reconstrução do mapa de profundidade (fonte: Carrivick et al. 2016) .....	22
<b>Figura 10.</b> Exemplo do filtro de consistência de visibilidade usado em PMVS (adaptado de Furukawa et al. 2010).....	23
<b>Figura 11.</b> Esquema abertura e relação com a luz (Mendonça, 2017).....	25
<b>Figura 12.</b> Esquema profundidade de campo (Mendonça, 2017) .....	26
<b>Figura 13.</b> Distorções mais comuns (Mendonça, 2017).....	27
<b>Figura 14.</b> Fluxo de trabalho VisualSfM (adaptado de <a href="http://ccwu.me/vsfm/doc.html">http://ccwu.me/vsfm/doc.html</a> , visitado em 15 março 2022) .....	32
<b>Figura 15.</b> Fluxo de trabalho Regard3D (adaptado de Rahaman e Champion, 2019) .....	33
<b>Figura 16.</b> Fluxo de trabalho Colmap (adaptado de Rahaman e Champion, 2019) .....	35
<b>Figura 17.</b> Fluxo de trabalho Meshroom .....	36
<b>Figura 18.</b> Fluxo de trabalho Autodesk Recap .....	38
<b>Figura 19.</b> Processo de georreferenciação do modelo 3D Autodesk Recap .....	38
<b>Figura 20.</b> Reconstrução do modelo 3D Autodesk Recap .....	38
<b>Figura 21.</b> Fluxo de trabalho Metashape (adaptado de Rahaman e Champion, 2019) .....	39
<b>Figura 22.</b> Provede de granito .....	40
<b>Figura 23.</b> Provede de granito (base com texto).....	41
<b>Figura 24.</b> Provede de granito (base lisa) .....	43
<b>Figura 25.</b> Provede de granito e a correspondente máscara.....	43

<b>Figura 26.</b> Fachada da sede da Associação de Estudantes da Universidade do Minho (Pólo de Guimarães) .....	44
<b>Figura 27.</b> Localização do talude (área vermelha marcada) (adaptado do Google Earth) .....	45
<b>Figura 28.</b> Imagem adquirida da frente de estudo (Talude) .....	45
<b>Figura 29.</b> VANT utilizado para levantamento aerofotogramétrico .....	46
<b>Figura 30.</b> Ponto de controlo GCP1 .....	47
<b>Figura 31.</b> Localização dos cinco pontos de controle .....	48
<b>Figura 32.</b> Tratamento da nuvem densa 3D .....	50
<b>Figura 33.</b> Nuvem densa 3D editada .....	50
<b>Figura 34.</b> Coordenadas do ponto de controle GCP1 .....	51
<b>Figura 35.</b> Distância entre GCP1 e GCP2 .....	51
<b>Figura 36.</b> Georreferenciação das nuvens densa 3D .....	52
<b>Figura 37.</b> Sobreposição das nuvens densa 3D .....	52
<b>Figura 38.</b> Cálculo da distância entre nuvens densa 3D .....	53
<b>Figura 39.</b> Distância calculada entre nuvens densa 3D .....	54
<b>Figura 40.</b> Nuvem de pontos densa (Provete) .....	57
<b>Figura 41.</b> Nuvem de pontos densa (base de folha de jornal) .....	57
<b>Figura 42.</b> Nuvem de pontos densa (base lisa) .....	58
<b>Figura 43.</b> Nuvem de pontos densa (Fachada da sede da Associação de Estudantes da Universidade do Minho - Colmap) .....	59
<b>Figura 44.</b> Nuvem de pontos densa (Maciço Rochoso) .....	64
<b>Figura 45.</b> Cálculo da distância entre nuvens densa 3D .....	65
<b>Figura 46.</b> Distância calculada entre nuvens densa 3D .....	66
<b>Figura 47.</b> Mapa de calor de distâncias (esquerda) e histograma (direita) .....	67
<b>Figura 48.</b> Histograma Geral - Distância Absoluta .....	68
<b>Figura 49.</b> Determinação da caixa delimitadora .....	71
<b>Figura 50.</b> Determinação do volume da nuvem densa 3D .....	72
<b>Figura 51.</b> Determinação da curva de nível - passo 1 .....	73
<b>Figura 52.</b> Determinação da curva de nível - passo 2 .....	73
<b>Figura 53.</b> Determinação da curva de nível - passo 3 .....	74
<b>Figura 54.</b> Curva de nível .....	76
<b>Figura 55.</b> Nuvem de pontos densa .....	86
<b>Figura 56.</b> "Pipeline" coordenadas GPS Colmap .....	87

## Lista de tabelas

<b>Tabela 1.</b> Especificações do sensor Iphone 11Pro.....	41
<b>Tabela 2.</b> Especificações do sensor do Redmi 9.....	42
<b>Tabela 3.</b> Condições da experiência realizada .....	44
<b>Tabela 4.</b> Especificações do sensor do VANT .....	46
<b>Tabela 5.</b> Informação das coordenadas dos cinco GCPs.....	48
<b>Tabela 6.</b> Resultados obtidos na modelação do provete de granito .....	55
<b>Tabela 7.</b> Comparação das nuvens de pontos densa antes e após o tratamento .....	60
<b>Tabela 8.</b> Resultados obtidos na modelação do talude .....	61
<b>Tabela 9.</b> Volume das nuvens de pontos.....	72

## **Lista de siglas**

DSLR - Digital Single-Lens Reflex Camera  
EXIF – EXchangeable Image File  
GCP – Ground Control Points  
GLOH – Gradient Laction-Orientation Histogram  
GPS – Global Positioning System  
GPU – Graphics Processing Unit  
GUI - Graphical User Interfaces  
ISPRS - International Society for Photogrammetry and Remote Sensing  
LiDAR - Light Detection And Ranging  
PMVS - Patch-based Multi-View Stereo  
RAM – Random Access Memory  
RANSAC – RANdom SAmples Consensus  
SANT – Sistema Aéreo Não Tripulado  
SfM-MVS - Structure from Motion – Multi-View Stereo  
SIFT - Scale-Invariant Feature Transform  
VANT – Veículo Aéreo Não Tripulado  
VPR – Veículo Pilotados Remotamente

## **Introdução**

### **Enquadramento e motivação**

A digitalização 3D está crescendo rapidamente em popularidade em todo o mundo porque está se tornando mais acessível e, talvez ainda mais importante, mais fácil de usar para pessoas que não são especialistas. Uma vez que o custo do equipamento está caindo enquanto a qualidade está aumentando, está se tornando uma tecnologia de escolha para documentando em muitas áreas diferentes, como a cultura património, arqueologia, engenharia civil, medicina e multimídia (Reljić e Dunđer, 2019).

Por estes motivos, têm sido exploradas técnicas de fabricação de modelos 3D, já que a modelagem 3D tradicional requer tempo, habilidade e trabalho manual. Dentre as técnicas atuais, Tannús et al. (2019) destacam o varrimento 3D e a fotogrametria.

De acordo com a American Society of Photogrammetry, a fotogrametria pode ser definida como: “a arte, ciência e tecnologia de obtenção de informação confiável sobre objetos físicos e o meio ambiente por meio de processos de gravação, medição e interpretação de imagens fotográficas e padrões de energia eletromagnética radiante e outras fontes.”

O varrimento 3D é um método de varredura de alta acurácia e precisão que permite a captura automática de grande quantidade de dados em curto espaço de tempo. Esse recurso baseia-se na análise e medição tridimensional, podendo utilizar scanners a laser ou luz estruturada, sensores de profundidade e aplicativos de telemóveis para a geração de um modelo digital de um objeto (Tannús et al., 2019).

Segundo Tannús et al. (2019), as técnicas possuem diferenças entre si, tais quais o nível de precisão, a resolução do modelo 3D produzido, o custo do equipamento e o tempo que leva para ser realizado cada procedimento. As duas técnicas de reprodução tridimensional estão sendo cada vez mais utilizadas e inseridas no mercado, possuindo uma vasta área de aplicações, desde a utilização para a elaboração de próteses médicas, até o registo e documentação de artefactos e construções.

O estudo realizado por Reljić e Dunđer (2019) evidencia que existem muitas tecnologias para a digitalização 3D, podendo ser classificados como:

Contato: digitalização de objetos pequenos e simples (por exemplo: moedas). A varredura é criada usando uma sonda fina que toca o objeto físico em muitas posições predefinidas. O modelo 3D resultante é um conjunto de pontos virtuais nas mesmas posições em que a sonda tocou o objeto.

Ativo sem contato: os dispositivos emitem luz e medem o tempo necessário para retornar ao dispositivo e, dessa forma reconstroem a distância do scanner para todos os pontos do objeto. Este grupo consiste em scanners a laser, LiDAR (Light Detection And Ranging) e scanners de luz estruturados. Esses dispositivos são bastante caros, mas precisos.

Passivo sem contato: dispositivos que recebem luz (como câmaras fotográficas e de vídeo) reconstróem um objeto 3D usando as fotografias feitas de diferentes ângulos ao redor desse objeto. Essa abordagem é a mais simples de usar, mas depende de vários fatores: qualidade de seu sensor, condições de iluminação, é necessário um grande número de fotografias, experiência do operador, ferramentas de software que podem converter fotografias em modelos 3D. Esses dispositivos são muito móveis para que possam ser usados em diferentes condições. Esta abordagem inclui a fotogrametria.

Dos métodos apresentados, o presente trabalho tem como enfoque a fotogrametria. É uma técnica poderosa para digitalização 3D que tem evoluído exponencialmente com a tecnologia, sendo capaz de capturar objetos de grandes dimensões tal como edifícios ou montanhas que seriam difíceis de digitalizar com outras técnicas. Além disso, a fotogrametria é muito acessível, requerendo apenas uma câmara fotográfica e um programa informático capaz de processar as imagens obtidas para criar a nuvem de pontos tridimensionais. A qualidade dos equipamentos de captação de imagens e a capacidade de processamento da informação captada são os dois pontos base da fotogrametria (Osmana e Tahar, 2016). Existem diversos programas para gerar um modelo 3D que se baseiam na funcionalidade “Structure from Motion” (SfM) podendo ser de uso gratuito ou comercial. “Structure from Motion” é uma técnica de levantamento topográfico que surgiu a partir dos avanços na visão computacional e na fotogrametria tradicional. Ele pode produzir nuvens de pontos tridimensionais (3D) densas e de alta qualidade por um custo financeiro mínimo. Os geocientistas podem, portanto, gerar dados tridimensionais (3D) comparáveis em escala e precisão aos da varredura a laser terrestre ou detecção e alcance de luz no ar (LiDAR), simplesmente usando seu PC de mesa padrão e uma câmara de mão. (Carrivick et al., 2016).

O estudo realizado por Rahaman e Champion (2019) apresentaram um estudo comparativo de quatro softwares gratuitos de reconstrução 3D (VisualSfM, Colmap, Regard3D e Python Photogrammetry-try Toolbox), selecionados com base no preço, facilidade de uso e instalação, formato de saída e, mais importante, o tempo de processamento necessário adequado ao ponto de vista do consumidor doméstico. Schöning e Heidemann (2015) realizaram um estudo de comparação, focando na precisão dos modelos reconstruídos obtidos com diferentes softwares, nomeadamente PhotoScan, 123D Catch, VisualSfM, ARC 3D e Ground Truth. Também foi analisado o tempo de computação necessário para cada software. Setiyadi *et al.* (2021) realizaram um estudo apresentando uma breve comparação de três softwares diferentes, ou seja, 3DF Zephyr, Agisoft Metashape e Regard 3D, para fotogrametria de curto alcance e reconstrução do modelo 3D aplicado ao pé humano.

O maciço a utilizar no estudo localizado em Braga, no qual se insere o Estádio Municipal do SCB, foi analisado por Riquelme *et al.* (2020). No estudo foi utilizado métodos tradicionais e ainda foi feita uma análise de um modelo digital capturado pela técnica SfM, identificando posteriormente as vantagens e desvantagens de ambos os métodos.

## **Objetivos**

Nesta dissertação pretende-se explorar a potencialidade de vários programas SfM de uso gratuito para digitalização 3D, com enfoque na análise estrutural das descontinuidades de maciços rochosos.

O levantamento fotogramétrico do maciço será feito com recurso a drone e em seguida é determinado as nuvens de pontos por recurso aos programas SfM em análise. É realizada uma análise comparativa identificativa das limitações/potencialidades entre várias programas SfM e identificação do programa mais apropriado para a análise estrutural de descontinuidades em maciços rochosos.

## **Estrutura e organização da dissertação**

A estruturação utilizada na presente dissertação é dividida em 5 capítulos, bibliografia e anexos com o seguinte formato:

No capítulo 1 é realizado o enquadramento e motivação do estudo, os objetivos definidos na realização dos trabalhos, a apresentação das tecnologias da digitalização 3D com destaque principal para a fotogrametria baseado na abordagem SFM, e por fim a estrutura e organização da dissertação.

No capítulo 2 apresentam-se a introdução a fotogrametria e a sua classificação, e é realizada a introdução da abordagem “Structure from Motion – Multi-View Stereo” (SfM-MVS) e o fluxo de trabalho. Ainda é feito um estudo sobre os sensores das câmaras digitais e as vantagens e desvantagens da abordagem SFM-MVS.

No capítulo 3 apresentam-se os materiais e métodos utilizados na realização dos trabalhos da presente dissertação. São apresentados os softwares SFM utilizados na digitalização 3D, os métodos intermédios baseado nos testes efetuados com os softwares para a criação das nuvens de pontos e o levantamento aerofotográfico, georreferenciação e modelação do nosso caso de estudo (talude). Ainda é apresentado a metodologia de comparação das nuvens de pontos baseado na sobreposição das nuvens com recurso ao software CloudCompare, a determinação das curvas de níveis e o cálculo do volume.

No capítulo 4 são apresentados os resultados finais dos diferentes ensaios e a digitalização do talude bem como a sua discussão e comparação dos modelos gerados nos softwares VisualSfM, Regard3D, Colmap, Meshroom e Autodesk Recap (versão académico) com o modelo de referência gerado no software comercial Metashape através do software CloudCompare. Também é apresentado as curvas de níveis dos modelos e o respetivo volume das nuvens de pontos.

No capítulo 5 é feita uma conclusão com base nos resultados obtidos dos trabalhos desenvolvidos e apresentam-se sugestões para a continuação deste trabalho.

Na parte final apresentam-se as referências bibliográficas e anexos.

## **1. Estado do conhecimento**

### **1.1. Fotogrametria**

“International Society for Photogrammetry and Remote Sensing” (ISPRS) publicou um documento guia para fornecer uma base comum para o uso e definição internacional de terminologias fotogramétricas. Ele define a fotogrametria como a ciência de “derivar medições 3D precisas e confiáveis de imagens”.

Existem muitas situações na vida ou na ciência em que devemos medir coordenadas, distâncias, áreas ou volumes. Em muitos casos, os métodos de medição dependem do tipo dos objetos. Há situações em que as técnicas fotogramétricas podem ser utilizadas como alternativa ou em que a fotogrametria é a única forma possível de medir (Linder, 2009).

Passando por um longo desenvolvimento histórico, a fotogrametria transformou-se da fotogrametria analítica para a fotogrametria analógica e, desde meados dos anos 90 até a data atual, a fotogrametria é conhecida por estar em sua era digital (Granshaw, 2016). Em particular, o desenvolvimento de triangulação digital semiautomática e algoritmos de extração de relevo baseados em imagens, de software livre e de código aberto, e o rápido avanço tecnológico e custo cada vez menor de computadores desktop de última geração melhoraram significativamente a qualidade da fotogrametria (Carrivick et al., 2016).

De acordo com Granshaw (2016), a fotogrametria pode ser classificada como aérea e terrestre. A fotogrametria terrestre (ou curto alcance) usa uma câmara estacionária no/perto do solo para capturar elementos em uma perspectiva mais próxima. A inclinação e outras especificações da câmara são todas controladas (Kar 2014).

A fotogrametria aérea (ou topográfica) utiliza um sensor montado num veículo aéreo (aeronaves, VANT, drones) para capturar várias fotografias sobrepostas da superfície terrestre e características topográficas. As câmaras encaixadas numa máquina que voa tiram fotos e, com o sistema, essas fotos são usadas para gerar medições. Nesse caso, pelo menos duas fotos do mesmo objeto ou superfície devem ser tiradas para comparação estatística. Aviões especialmente projetados são usados para esse tipo de fotografia. A velocidade da câmara é regulada de acordo com a velocidade do avião. Além disso, a altura da aeronave em relação ao terreno é conhecida. O estereógrafo é usado para criar o contorno base da superfície que está sendo fotografada (Kar, 2014).

Segundo Carrivick et al. (2016), avanços rápidos em tecnologia de microchip, Sistema de Posicionamento Global (GPS) e unidade de movimento inercial (IMU) viram uma proliferação de plataformas aéreas de custo relativamente baixo no mercado nos últimos 5 anos. Muitos termos diferentes são usados para descrever essas plataformas, sendo os mais comuns veículos aéreos não

tripulados (VANT), sistemas aéreos não tripulados (SANT), veículos pilotados remotamente (VPR) ou simplesmente drones. Os termos VANT, VPR e "drone" tendem a se referir à própria aeronave, enquanto o SANT descreve mais comumente toda a configuração, incluindo a unidade de controle remoto e o link de dados sem fio. Aqui, usaremos o termo VANT para descrever todas as plataformas aéreas remotas sendo operadas ou controladas por um usuário terrestre. Dado o recente aumento de operadores de VANT, muitos países introduziram regulamentos que regem quando e onde os VANT podem voar.

Existem quatro tipos principais de VANT (fig. 1):

**Aeronave autoprovelsada de asa fixa.** Estes são altamente eficientes (em termos de energia) e têm capacidade de longo alcance, facilitando levantamentos de grandes áreas. Dado seu tamanho e rigidez, eles tendem a ser uma das plataformas de aquisição mais estáveis. No entanto, seu grande tamanho (em relação a outras abordagens) muitas vezes exige uma pista de decolagem e pouso, o que pode não ser viável em terrenos remotos e/ou acidentados.

**Sistemas de rotor duplo de decolagem e pouso vertical (por exemplo, heli).** Estes são sistemas altamente flexíveis que podem ser implantados em praticamente qualquer terreno. Eles têm alcance médio (restringido na maioria dos casos pela vida útil da bateria), não podem ser implantados em condições tempestuosas se for necessária estabilidade e, dada a energia necessária para levantar e manter a plataforma em alta altitude, o tempo de vôo geralmente é curto.

**Dirigíveis.** Essas naves relativamente grandes não são frequentemente utilizadas para capturar imagens SfM-MVS por causa de seus altos custos de operação, embora seu tempo de voo e alcance sejam longos, tornando possíveis pesquisas de área ampla.

**Multicópteros.** A maioria dos entusiastas do voo possui um multicóptero por causa de sua facilidade de operação e alta flexibilidade e estabilidade na maioria das condições climáticas. Com o tempo e o desenvolvimento tecnológico, os custos de compra e manutenção estão diminuindo, e comunidades on-line inteiras são dedicadas a apoiar desenvolvimentos e inovações de multicópteros. No lado negativo, essas aeronaves têm alcance e tempo de voo limitados – a duração típica da bateria pode ser da ordem de aproximadamente 20 minutos. Os multicópteros são provavelmente a plataforma mais popular para coletar imagens SfM-MVS e ajudaram na produção dos maiores modelos 3D SfM-MVS até hoje



a) Aeronave autopropulsada  
de asa fixa

b) Heli

c) Multicópteros

**Figura 1.** Tipos de VANT (Carrivick et al. 2016)

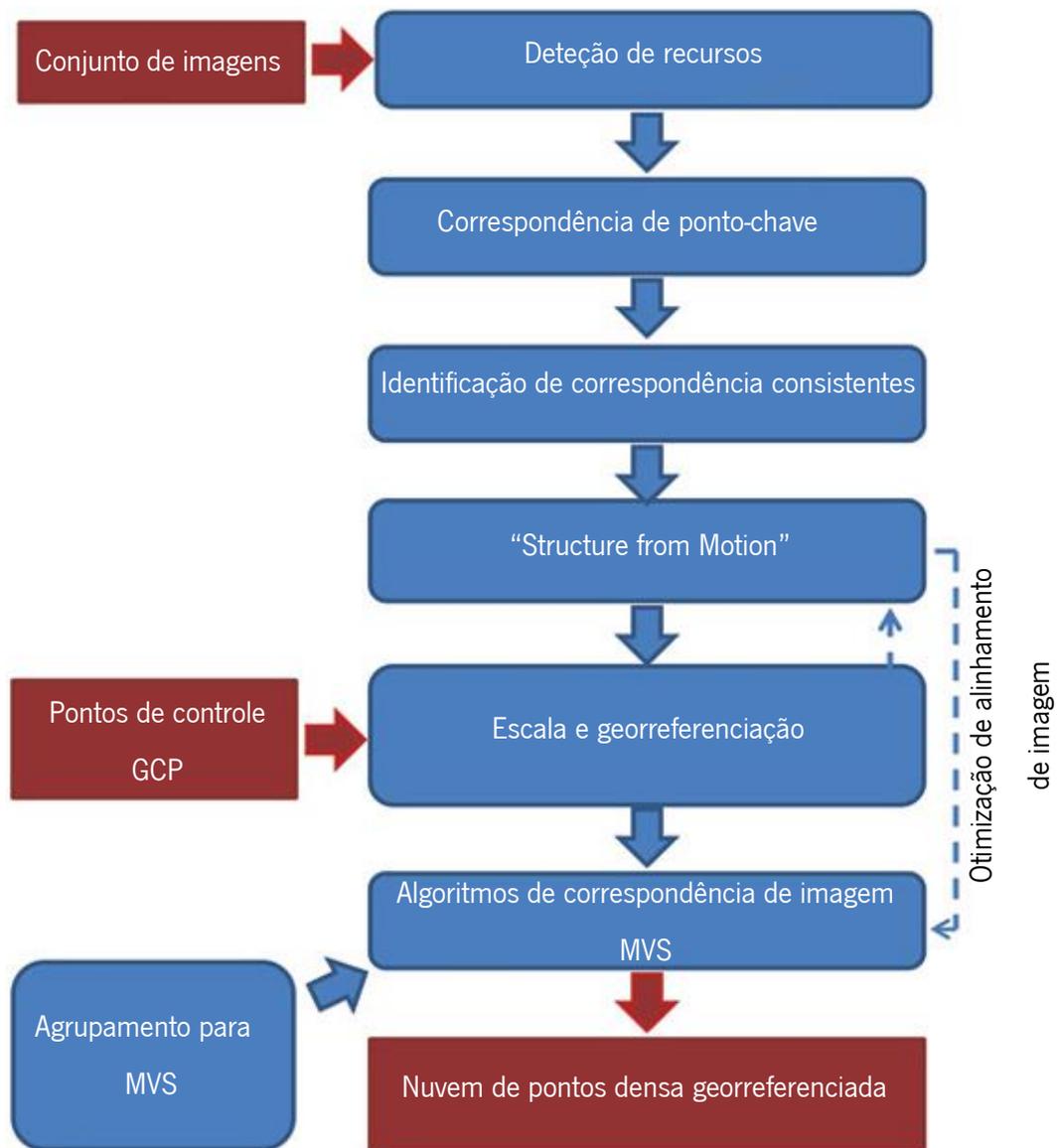
## **1.2. Introdução SFM-MVS**

Desde a década de 1980, o “Structure from Motion” (SfM) tem-se revelado uma ferramenta valiosa na geração de modelos 3D a partir de imagens 2D, especialmente com o desenvolvimento de software com amigáveis “Graphical User Interfaces” (GUI). Na íntegra, o fluxo de trabalho é conhecido como SfM-MVS, por conta dos algoritmos “Multi-View Stereo” (MVS) usados nas etapas finais. O SfM utiliza os algoritmos para identificar recursos correspondentes num conjunto de imagens digitais sobrepostas e calcula a localização e a orientação da câmara a partir das posições diferenciais de vários recursos correspondentes. Tendo como base esses cálculos, utiliza essas imagens sobrepostas para construir um modelo de nuvem de pontos 3D disperso. O modelo 3D obtido pelo método SfM é geralmente aperfeiçoado para uma resolução muito mais fina através do método Multi-View Stereo (MVS), concluindo assim o fluxo de trabalho SfM-MVS.

As principais vantagens do SfM-MVS é o baixo custo em ambos os requisitos de hardware e software, é rápido em comparação com outros levantamentos digitais no campo e é um fluxo de trabalho virtualmente independente da escala espacial. Além disso, o SfM-MVS pode produzir uma densidade/resolução espacial de pontos de pesquisa e precisão de ponto 3D que, em algumas circunstâncias, é comparável à dos modernos scanners a laser terrestres.

## **1.3. Fluxo de Trabalho SFM-MVS**

Esta secção fornece informações mais detalhadas sobre a operação SfM-MVS, tendo como objetivo apresentar os diferentes processos envolvidos na reconstrução da geometria tridimensional (3D) a partir de uma sequência de imagens (fig. 2).



**Figura 2.** Fluxo de trabalho (adaptado de Carrivick et al. 2016)

Segundo Carrivick et al. (2016), os métodos fotogramétricos mais flexíveis surgiram com a fotogrametria digital agora amplamente utilizada, permitindo a derivação de produtos de dados topográficos de alta qualidade em uma ampla gama de escalas espaciais. O SfM-MVS combina princípios fotogramétricos com desenvolvimentos de origem totalmente diferente, provenientes de avanços em algoritmos de visão computacional 3D desde os anos 1980. Com a proliferação da fotografia digital no final da década de 1990 e o aumento da disponibilidade de imagens, as aplicações dos algoritmos SfM-MVS tornaram-se mais evidentes.

Em seguida é descrita cada etapa do fluxo de trabalho típico implementado em muitos sistemas SfM-MVS, desde um conjunto de imagens até uma nuvem de pontos densa e georreferenciada. De acordo

com o Carrivick et al. (2016) os detalhes específicos variam entre os pacotes de software, mas a grande maioria das implementações de geociências compartilham muitos recursos com o fluxo de trabalho:

- Detecção de recursos.
- Correspondência de ponto-chave.
- Identificação de correspondências consistentes.
- “Structure from Motion”
- Escala e georreferenciação.
- Otimização do alinhamento de imagem.
- Agrupamento para MVS
- Algoritmos de correspondência de imagem MVS.

Existem muitas abordagens alternativas baseadas em SfM onde cada implementação de software específica do SfM-MVS é ligeiramente diferente. A seguinte descrição da abordagem SfM-MVS descreve os processos típicos e fluxo de trabalho necessários para reconstruir a geometria da cena 3D a partir de um conjunto de imagens onde os parâmetros de calibração extrínsecos e intrínsecos são desconhecidos (Carrivick et al., 2016).

### **1.3.1. Detecção de recursos**

Nas últimas décadas, vários avanços ligeiramente diferentes na detecção de características foram feitos em paralelo nos campos de processamento digital de imagens e visão computacional. Este último também teve como objetivo recuperar a estrutura 3D das imagens. A questão fundamental que impulsionou o desenvolvimento de ambas as abordagens foi como extrair melhor as descrições dos pontos locais de uma maneira que permitisse a identificação correta das correspondências entre esses pontos, mas que fosse insensível a mudanças na orientação, escala, iluminação ou posição 3D (fig. 3) (Carrivick et al., 2016).



**Figura 3.** Desafios enfrentados pelos algoritmos de correspondência de recursos

(Carrivick et al. 2016)

A primeira etapa da correspondência de imagens envolve a identificação de pontos comuns em várias fotografias diferentes. São esses pontos-chave que permitem que as diferentes imagens sejam combinadas e a geometria da cena reconstruída. As aplicações iniciais da correspondência de pontos-chave foram para rastreamento de movimento estéreo e de curto alcance. No entanto, conforme discutido por Snavely (2008), essas técnicas limitavam-se à identificação de pontos-chave em uma única escala e eram aplicáveis apenas a cenas com imagens tiradas de um ponto de vista semelhante. O desafio é identificar recursos que podem ser rastreados entre imagens tiradas de visões muito diferentes (Baumberg 2000; Matas et al., 2004).

**Pontos característicos** (“sets of pixels”) que são invariantes a mudanças em escala e orientação são necessários para correspondência de linha de base ampla (Mikolajczyk et al., 2005). A forma da região de interesse deve ser capaz de se adaptar para lidar com distorções geométricas no recurso de destino devido a uma mudança de perspectiva entre duas imagens. Esta acomodação da distorção geométrica é demonstrada na Figura 4 onde as regiões elípticas são detetadas independentemente em cada ponto de vista, mas correspondem à mesma região da superfície. Deformações geométricas e fotométricas podem ser normalizadas para obter descrições invariantes de pontos de vista e iluminação do padrão de intensidade na região (Mikolajczyk et al., 2005).

Embora muitos detetores de região e tipos de recursos diferentes estejam disponíveis, o sistema de reconhecimento de objeto de transformação de recurso invariante de escala “scale-invariant feature transform” (SIFT) é amplamente o mais usado (Lowe 1999, 2001, 2004). O SIFT permite que a posição relativa do recurso mude drasticamente com apenas pequenas alterações no descritor. Além disso, o

SIFT é robusto contra mudanças no ponto de vista 3D para superfícies não planares (Lowe 2004). No entanto, SIFT não é totalmente invariante afim, o que seria mais útil para combinar superfícies planas sob grandes mudanças de visão (fig. 4). Lowe (2004) sugeriu que uma combinação de SIFT com outros tipos de recursos forneceria mais correspondências de imagem em diferentes circunstâncias e que provavelmente seria implementada em sistemas futuros. De facto, a conclusão de Mikolajczyk et al. (2005) foi que o uso de vários detetores complementares deve ser utilizado para extrair regiões com propriedades diferentes.



**Figura 4.** Regiões elípticas detetadas independentemente em cada ponto de vista (Mikolajczyk et al. 2005)

Segundo Lowe (2004), resumidamente, o SIFT segue quatro etapas principais:

- 1. Detecção de extremos espaciais.** Esta primeira etapa do SIFT envolve uma identificação eficiente de localizações e escalas que podem ser atribuídas repetidamente ao mesmo objeto a partir de diferentes pontos de vista. É usada uma abordagem de escala espacial que deteta locais que são invariáveis às mudanças de escala, procurando recursos estáveis em uma função contínua de escala. Os extremos locais são então detetados comparando cada ponto de amostra com seus oito vizinhos na imagem atual e nove vizinhos nas escalas acima e abaixo. Lowe (2004) analisou a frequência de amostragem preferida em escala e espaço, sugerindo que a maioria das características estáveis e úteis poderiam ser detetadas com amostragem grosseira.
- 2. Localização de ponto-chave.** O SIFT então executa um ajuste detalhado de uma função 3D quadrática para cada ponto-chave candidato a dados próximos para localização, escala e proporção das curvaturas principais. Um grande número de pontos-chave são normalmente identificados (fig. 5b). São rejeitados pontos que possuem baixo contraste

(removidos da Fig. 5c) ou que estão mal localizados ao longo de uma borda (removidos da Fig. 5d). A densidade dos pontos-chave identificados em uma imagem depende da textura, nitidez e resolução da imagem. Cenas complexas funcionarão melhor, enquanto superfícies relativamente sem características, como neve e areia, provavelmente serão as mais desafiadoras. Para uma ilustração da localização de pontos-chave, a Fig. 6 compara correspondências de pontos-chave para pares de imagens em superfícies de solo e neve.



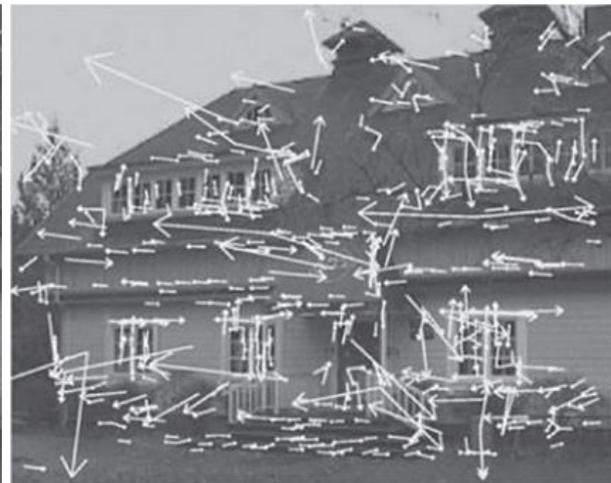
**a)** A imagem original de  $233 \times 189$  pixels



**b)** Identificação de pontos-chaves

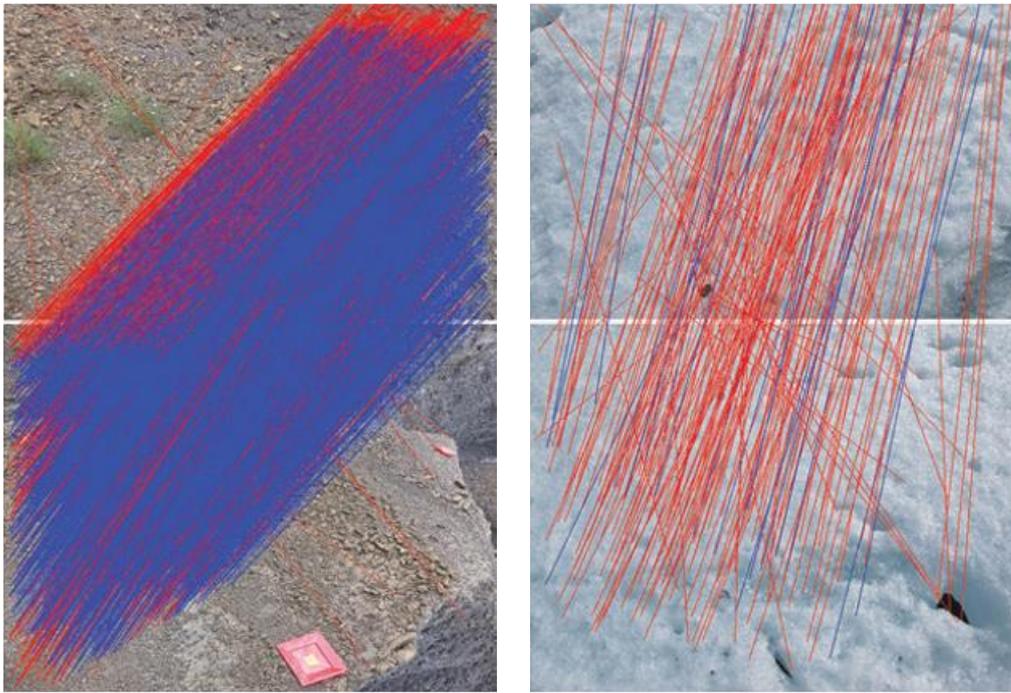


**c)** Remoção de pontos de baixo contraste



**d)** Remoção de pontos mal localizados

**Figura 5.** Etapas de SIFT (Lowe, 2004)



**a)** Superfície de solo

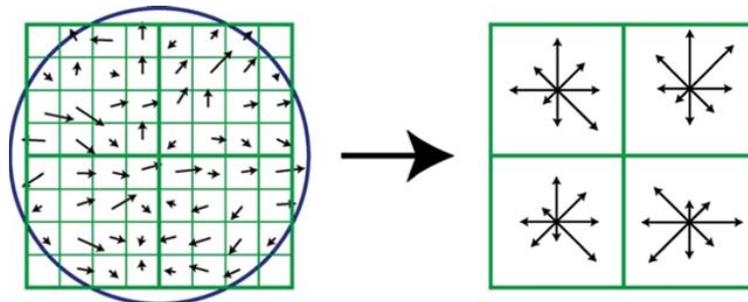
**b)** Superfície de neve

**Figura 6.** Comparação de correspondência de pontos-chave (Carrivick et al. 2016)

**3. Tarefa de orientação.** Uma orientação consistente para cada ponto-chave é atribuída através da análise das direções dominantes dos gradientes de intensidade locais usando a imagem mais próxima da escala do ponto-chave. Onde um segundo pico no histograma de orientação é identificado, um segundo ponto-chave é criado naquele local e escala, mas com uma orientação diferente.

**4. Descritor de ponto-chave.** Em seguida, é necessário um descritor para cada ponto-chave que seja suficientemente distinto, mas o mais invariável possível às mudanças no ponto de vista 3D ou na iluminação. A abordagem do SIFT para obter esse descritor é que ele considera a sensibilidade aos gradientes de intensidade, mas não a localização desses gradientes. As magnitudes e orientações do gradiente são amostradas em torno de cada ponto-chave, giradas em relação à orientação do ponto-chave (fig. 7). Uma função de ponderação gaussiana é aplicada a esses gradientes para evitar grandes gradientes distantes do centro do descritor determinando o descritor específico. Os gradientes são acumulados em histogramas de orientação ao longo de  $4 \times 4$  regiões de amostra. O descritor é, portanto, uma matriz  $4 \times 4$  de histogramas com oito caixas de orientação, cada uma resultando em um vetor de características de 128 elementos para cada ponto-chave. Para evitar efeitos de iluminação, o vetor é normalizado para o comprimento da unidade, corrigindo assim as

alterações de contraste. Além disso, os valores no vetor de recurso unitário são limitados para evitar que grandes magnitudes de gradiente afetem a correspondência e, assim, coloquem maior foco nas orientações do vetor. O descritor resultante mostrou discriminar pontos-chave individuais de grandes bancos de dados.



**Figura 7.** Demonstração da matriz de descritores  $2 \times 2$  calculada a partir de um conjunto de amostras  $8 \times 8$  (Lowe, 2004)

Um estudo comparativo de outros descritores de imagens locais invariantes de visão feito por Mikolajczyk e Schmid (2005) descobriu que o “Gradient Laction-Orientation Histogram” (GLOH), uma variante do SIFT, superou mas apenas por uma pequena margem. O GLOH difere do SIFT porque usa “bins” log-polares em vez de “bins” quadrados para calcular os histogramas de orientação. Para notar, outras variações subsequentes incluem SURF (Bay et al., 2008), ASIFT (Morel e Yu, 2009), BRIEF (Calonder et al., 2010) e LDAHash (Strecha et al., 2012).

### 1.3.2. Correspondência de ponto-chave

Uma vez que os pontos-chave tenham sido localizados em cada imagem, as correspondências entre os pontos-chave em diferentes imagens precisam ser determinadas. No entanto, não há garantia de que qualquer ponto-chave terá uma correspondência em outra imagem. Portanto, métodos para descartar pontos sem boa correspondência são necessários. Trabalhando com os dados de pontos-chave de 128 dimensões do algoritmo SIFT, Lowe (2004) usou a razão da distância euclidiana do vizinho mais próximo com a do segundo mais próximo, especificando um valor mínimo de 0,8. Esse critério de “razão de distância” foi observado para eliminar 90% das correspondências falsas, descartando apenas menos de 5% das correspondências corretas. Também foi encontrado um desempenho melhor do que um limite de distância global, pois a razão de distância especifica que as correspondências corretas devem ser substancialmente mais corretas do que as outras opções, aumentando a probabilidade de uma correspondência confiável. Por outro lado, é improvável que as correspondências espúrias sejam muito

melhores do que a segunda correspondência incorreta mais próxima, dada a alta dimensionalidade do espaço de recursos (Lowe, 2004).

O estudo realizado por Lowe (2004) evidencia que cortar a busca aproximada do vizinho mais próximo (ANN) após verificar apenas os primeiros 200 candidatos ao vizinho mais próximo fornece uma grande economia de tempo (duas ordens de magnitude onde os pontos-chave são  $> 100.000$ ) enquanto perde apenas menos de 5% das correspondências corretas. Portanto, quando acoplado ao critério de razão de distância descrito anteriormente, o algoritmo BBF não precisa fornecer soluções exatas dos casos mais difíceis de pontos-chave com muitos vizinhos próximos, pois a correspondência seria rejeitada pelo critério de razão de distância em qualquer caso. A implementação de pesquisas de vizinhos mais próximos de alta dimensão em “graphics processing unit” (GPU) também demonstrou diminuir os tempos de pesquisa necessários executando seis vezes mais rápido (Bustos et al., 2006).

### **1.3.3. Identificação de correspondências consistentes**

Para ter certeza de que apenas as correspondências corretas permanecem, uma etapa adicional é aplicada para filtrar quaisquer correspondências errôneas restantes. Com vários pontos-chave identificados em um par de imagens da mesma cena (fig. 6), a matriz fundamental (matriz  $F$ ) para o par de imagens é calculada para câmaras não calibradas. Ao especificar a relação entre as duas imagens, a matriz fundamental restringe as localizações dos pontos-chave corretamente identificados em ambas as imagens e pode ser calculada usando o algoritmo de oito pontos (Longuet-Higgins 1981). Este algoritmo usa correspondências de oito pontos (ou mais) em duas visualizações não calibradas e um conjunto de equações lineares para reconstruir uma cena até uma transformação projetiva em que todos os pontos situados em uma única linha permanecerão alinhados dessa maneira. O algoritmo de oito pontos é uma abordagem simples e rápida para calcular a matriz  $F$ , mas é sensível ao ruído na localização especificada dos pontos-chave. A normalização simples dos pontos na imagem antes de resolver o conjunto de equações lineares foi mostrada por Hartley (1997) para produzir uma grande melhoria no uso do algoritmo de oito pontos.

As matrizes  $F$  candidatas são calculadas ao longo de várias iterações usando o método “Random Sample Consensus” (RANSAC) (Fischler & Bolles 1981). O RANSAC é o método mais popular porque é rápido, preciso e robusto (Sunglok et al. 2009). O método RANSAC assume que todos os pontos-chave podem ser divididos em dois conjuntos: “outliers” e “inliers”. Um ajuste de modelo perfeito ignoraria todos os “outliers” e seria calculado exclusivamente a partir dos “inliers”. Uma amostra aleatória inicial dos pontos-chave é retirada da qual a matriz  $F$  é calculada usando o menor subconjunto possível dos dados, neste caso sete pontos. O erro de cada ponto-chave em relação à estimativa é então calculado e o número

de pontos internos contados. Para definir “inliers”, algum limite deve ser especificado. Snavely et al. (2008) definiram esse limite como 0,6% da dimensão máxima da imagem, normalmente seis pixels para uma imagem de  $1024 \times 768$ . O processo de amostragem é repetido em diferentes subconjuntos iterativamente por um número suficiente de vezes para garantir que haja 95% de chance de que um dos subconjuntos contenha apenas “inliers”. RANSAC então retorna a matriz F com o maior número de “inliers”. O modelo final é então calculado usando apenas o conjunto interno.

Snavely et al. (2008) refina ainda mais esta matriz F executando o algoritmo iterativo Levenberg-Marquardt (LM) no conjunto “inlier”. Todas as correspondências “outlier” adicionais são removidas. No caso de restar apenas um pequeno número de correspondências “inlier” (<20 em Snavely et al. 2008), todas as correspondências são removidas da consideração. Outras abordagens para remover correspondências de pontos-chave geometricamente inconsistentes estão disponíveis e aplicadas em diferentes softwares.

Com os pontos-chave limitados àqueles com correspondências geometricamente consistentes, as ligações entre cada par de imagens podem ser identificadas e organizadas em trilhas: conjuntos conectados de pontos-chave correspondentes através da biblioteca de imagens usadas na reconstrução (Snavely et al. 2008). Um mínimo de dois pontos-chave localizados em três imagens é necessário para uma trilha. Quando o mesmo ponto-chave ocorre duas vezes em uma única imagem, a faixa é considerada inconsistente. Mapas de trilhas consistentes podem então ser feitos, identificando a conectividade de cada imagem. Essas faixas são utilizadas nas etapas a seguir.

#### **1.3.4. “Structure from Motion”**

Geralmente refere-se toda a cadeia de processo aqui descrita como SfM, mas é o processo único de estimar simultaneamente a geometria 3D (ou estrutura) de uma cena e as diferentes poses de câmara (ou seja, movimento) que é mais tecnicamente conhecido como “SfM” (Ullman, 1979). Usando as correspondências de recursos geometricamente corretas identificadas na seção anterior, o SfM visa reconstruir simultaneamente: (i) estrutura de cena 3D, (ii) posições e orientações da câmara (ou seja, estimativa de pose ou calibração extrínseca) e, muitas vezes, (iii) parâmetros intrínsecos de calibração da câmara.

Os parâmetros extrínsecos da câmara de (ii) representam a transformação do corpo rígido entre as coordenadas da cena 3D e o sistema de coordenadas da câmara. Existem muitos modelos de câmaras diferentes para descrever os parâmetros intrínsecos de (iii). O modelo mais comum é uma projeção em perspectiva descrita por um modelo de câmara pinhole; outros incluem projeções afins, projeções ortográficas e modelos “pus-broom”.

Os parâmetros intrínsecos da câmara são definidos por uma matriz triangular superior  $3 \times 3$  conhecida como matriz de calibração da câmara  $K$ :

$$K = \begin{bmatrix} a_u & s & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

onde  $a_u$  e  $a_v$  dimensionam a imagem nas direções  $x$  e  $y$ , respetivamente, e  $s$  representa a inclinação. Sob a suposição de pixéis quadrados  $s = 0$  e  $a_u = a_v = a$  onde  $a$  é considerado a distância focal da lente em unidades da dimensão do pixel. O ponto principal  $(u_0, v_0)$  é definido como a localização no plano da imagem que intercepta o eixo ótico.

Parâmetros intrínsecos adicionais são necessários para modelar aberrações internas (ou seja, parâmetros de distorção radial), assumindo que as câmaras não foram pré-calibradas. A distorção radial faz com que os pontos da imagem sejam deslocados numa direção radial do centro da distorção, que muitas vezes é considerado o ponto principal, e pode ser corrigido com o conhecimento de dois coeficientes da função de distorção ( $k_1$  e  $k_2$ ). O grau em que essa distorção radial é incorporada aos modelos de câmara é variável.

**Ajusto de pacotes** produz “jointly optimal” estrutura 3D e parâmetros de visualização (pose e/ou calibração) (Granshaw, 1980; Triggs et al., 2000). As palavras “jointly optimal” aplicam-se aqui como as estimativas de parâmetros que se aplicam a ambas as estruturas, e as variações de câmara são feitas minimizando o valor de uma função de custo que quantifica o erro de ajuste do modelo. Segundo Tomasi e Kanade (1992), os algoritmos de factoração calculam a pose da câmara e a geometria da cena 3D usando todas as imagens simultaneamente usando uma decomposição de valor singular. No entanto, esses algoritmos exigem que todos os pontos-chave sejam visíveis em todos os quadros, mas o estudo realizado por Szeliski (2011) evidencia que existem maneiras de lidar com essa limitação. Os métodos sequenciais são uma alternativa mais popular e são descritos a seguir.

Os valores dos parâmetros devem ser atribuídos aos valores iniciais antes da otimização dos parâmetros não lineares do ajuste do pacote. Para evitar encontrar soluções não ótimas (mínimos locais) para problemas SfM de grande escala, o processo de reconstrução da cena normalmente começa com um único par de imagens, chamado de “par inicial”. O par inicial deve ter um grande número de correspondências e uma grande linha de base para uma reconstrução esparsa. Uma inicialização dos valores dos parâmetros é necessária antes que a reconstrução da cena possa começar. Quando os parâmetros intrínsecos da câmara são inicialmente desconhecidos, é necessário um método de auto-

calibração. Enquanto Hartley e Zisserman (2003) observam que a auto-calibração pode ser alcançada com três ou mais quadros ou pode recuperar distâncias focais de dois “frames”, fazendo suposições sobre outros parâmetros da câmara, Snavely et al. (2008) conseguem essa inicialização limitando a escolha do par inicial para imagens para as quais as estimativas de distância focal estão disponíveis (por exemplo, de formato de arquivo de imagem intercambiável ou “tags” EXIF). Eles então obtêm os parâmetros restantes da câmara usando o algoritmo de cinco pontos de Nistér (2004), que é um algoritmo de pose relativa calibrada. As trilhas visíveis no par inicial são então trianguladas para obter estimativas iniciais das posições das feições (Hartley e Sturm 1997).

O principal objetivo da inicialização é minimizar o erro entre as projeções de cada faixa e os pontos-chave correspondentes no par inicial. Usando este erro como critério de otimalidade a ser minimizado, o problema de mínimos quadrados não linear resultante (Nocedal & Wright 2006) é resolvido usando um ajuste de pacotes de dois “frames”.

Com o erro de reprojeção entre as duas imagens minimizado, outra câmara é adicionada à otimização (ou várias câmaras). A câmara que contém o maior número de trilhas cujas localizações 3D já estão estimadas é selecionada (ou qualquer câmara com pelo menos 75% das correspondências máximas de pontos-chave). Os parâmetros extrínsecos da câmara para a nova câmara são inicializados usando a técnica de transformação linear direta (Abel-Aziz e Karara, 1971), que usa um conjunto de pontos de controle conhecidos (as localizações 3D conhecidas existentes) e mapeia as coordenadas 2D da nova imagem nas coordenadas 3D do espaço do objeto 3D. Isso pode ser implementado dentro de um procedimento RANSAC e também retorna uma matriz triangular superior  $K$  para valores de inicialização dos parâmetros intrínsecos da câmara (junto com “tags” EXIF). Uma etapa de ajuste de pacote adicional é executada com esta nova imagem, mas, no entanto, apenas os novos parâmetros da câmara e os pontos observados podem ser alterados. Onde os pontos-chave na nova imagem são observados por pelo menos uma outra câmara que já foi adicionada ao modelo, todos os raios existentes para aquele ponto são usados para triangular a posição do ponto. Onde o ângulo máximo de separação entre qualquer par de raios é menor que um limite especificado (por exemplo,  $2^\circ$ ), o novo ponto é rejeitado.

Para melhorar a precisão da solução, é preferível realizar um ajuste global do pacote em todas as câmaras para refinar todo o modelo (Szeliski, 2011). A minimização da função de custo é um processo iterativo, ajustando uma aproximação quadrática local à função de custo em cada iteração ou, onde tais modelos fornecem um ajuste impreciso, usando o método gradiente descendente (ou seja, o algoritmo LM mencionado anteriormente). O LM pode convergir rapidamente a partir de uma ampla gama de inicializações (Hiebert 1981). No entanto, com muitas câmaras e vários parâmetros desconhecidos por câmara, o espaço do parâmetro de ajuste do pacote rapidamente se torna de alta dimensão. Algoritmos de ajuste de feixe esparso (Lourakis & Argyros 2009) reduzem a carga computacional de outra forma

intratável resultante da alta dimensionalidade do problema, levando em conta a falta de interação entre os parâmetros para diferentes câmaras e pontos 3D.

Trilhas discrepantes contendo pontos-chave com alto erro de reprojeção são removidas após cada execução da otimização de ajuste do pacote. Estes podem ser definidos de acordo com a distribuição de probabilidade de todos os erros de reprojeção, limites de erro de pixel específicos ou uma combinação dos dois. As câmaras são adicionadas sequencialmente ao modelo e o processo mencionado anteriormente é repetido. O processo termina quando nenhuma câmara restante contém um número suficiente de pontos 3D reconstruídos para serem adicionados de forma confiável ao modelo.

O processo SfM produz uma nuvem de pontos dispersa e poses de câmara reconstruídas. É esta nuvem de pontos dispersa (uma vez georreferenciada) que é usada em muitas aplicações de SfM, incluindo aquelas nas geociências (Fonstad et al. 2013). Entretanto é necessário processamento adicional para reconstruções de superfície mais detalhadas e de alta qualidade. A maioria dos aplicativos aplicará técnicas de MVS para produzir uma nuvem de pontos muito mais densa (Rosnell & Honkavaara 2012).

### **1.3.5. Escala e georreferenciação**

O SfM-MVS fornece apenas localizações de câmaras relativas e geometria de cena, de modo que a saída da nuvem de pontos é gerada em um sistema de coordenadas arbitrário. Distâncias absolutas entre câmaras ou entre pontos reconstruídos nunca podem ser recuperadas apenas a partir de imagens, independentemente de quantas câmaras ou pontos sejam usados (Szeliski, 2011). O georreferenciamento e o dimensionamento da nuvem de pontos requerem um mínimo de três “Ground Control Points” (GCPs) com coordenadas XYZ para uma transformação de semelhança linear de sete parâmetros, que compreende três parâmetros de tradução global, três parâmetros de rotação e um parâmetro de dimensionamento. Alternativamente, o georreferenciamento e dimensionamento “direto” podem ser realizados a partir de posições de câmara conhecidas derivadas de medições de GPS diferencial cinemático em tempo real (dGPS) e uma unidade de medição inercial (Turner et al., 2014). Um híbrido comum das duas abordagens usa o georreferenciamento direto para fornecer localizações aproximadas da câmara para inicializar o ajuste do pacote e, em seguida, usa GCPs externos para restringir melhor a solução (Rippin et al., 2015).

Como identificar pequenos recursos diretamente em nuvens de pontos pode ser difícil, muitos fluxos de trabalho de software SfM-MVS permitem que o usuário localize o alvo diretamente das imagens (James e Robson, 2012). As coordenadas arbitrárias dos alvos do modelo SfM-MVS são emparelhadas com as coordenadas absolutas dos GCPs e usadas para derivar uma transformação de similaridade. Um número

maior de alvos do que três é recomendado, que é o mínimo absoluto necessário para uma solução única para a transformação (Carrivick et al., 2016).

### **1.3.6. Otimização do alinhamento de imagem**

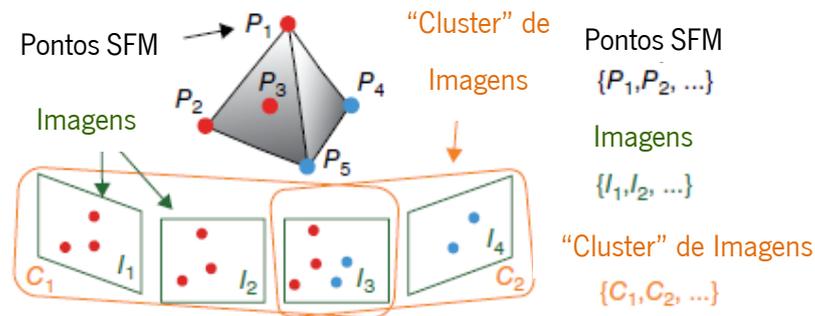
Erros na estimativa dos parâmetros intrínsecos e extrínsecos da câmara decorrentes do processo SfM-MVS podem levar a deformações não lineares do modelo final. A entrada de GCPs na etapa anterior fornece informações adicionais sobre a geometria 3D que pode ser usada para refinar ainda mais os parâmetros da câmara e a geometria da cena reconstruída. As coordenadas conhecidas (e estimativas de erro pontual) fornecem uma fonte adicional de erro na minimização da função de custo não linear durante a etapa de ajuste do pacote. Com essas informações externas incluídas no modelo, o ajuste do pacote pode ser executado novamente para otimizar o alinhamento da imagem à luz dessas novas informações. A distribuição espacial dos GCPs também é crucial para esse processo de otimização, onde os GCPs não cobrem adequadamente a área de interesse e a otimização pode prejudicar a precisão geral do levantamento. Portanto, recomenda-se cautela ao realizar esta etapa. (Carrivick et al., 2016)

### **1.3.7. Agrupamento para MVS**

Antes que as técnicas MVS sejam aplicadas à nuvem de pontos, há uma etapa adicional opcional que pode ser necessária em projetos com grandes conjuntos de imagens. Alguns algoritmos MVS resolvem um mapa de profundidade para cada imagem por vez (usando imagens próximas) e depois mesclam as reconstruções separadas (Micusik e Kosecka, 2009). Isso permite a paralelização, mas à custa de mapas de profundidade ruidosos e altamente redundantes que exigem pós-processamento adicional para limpeza e mistura (Furukawa et al. 2010).

Em contraste, muitos dos algoritmos MVS de melhor desempenho reconstróem a geometria da cena globalmente usando todas as imagens simultaneamente (Pons et al., 2007). Quando o número de imagens aumenta, a carga computacional de tal abordagem aumenta rapidamente e surgem problemas de escalabilidade. Os requisitos de memória de acesso aleatório (RAM) aumentam com o número de imagens usadas na reconstrução e colocam um limite prático no número de imagens que podem ser combinadas simultaneamente. A solução para esse problema de RAM é o agrupamento de imagens, ou seja, a divisão de um projeto grande em partes. Furukawa et al. (2010) detalham uma etapa de pré-processamento conhecida como “clustering views for MVS” (CMVS), que é um método pelo qual o conjunto de imagens é decomposto em clusters de exibição sobrepostos para permitir que reconstruções MVS densas sejam executadas nos clusters separadamente. A nuvem de pontos dispersa gerada a partir de SfM é usada para produzir clusters de imagens sobrepostas de tamanho gerenciável, de modo que

cada ponto 3D seja reconstruído por pelo menos um “cluster”. A ideia básica subjacente à abordagem de agrupamento de imagens é demonstrada na Figura 8.



**Figura 8.** O algoritmo CMVS leva imagens  $\{I_i\}$ , pontos SfM  $\{P_i\}$  e suas informações de visibilidade associadas  $\{V_i\}$  para produzir clusters de imagens sobrepostas  $\{C_i\}$  (adaptado de Furukawa et al. 2010)

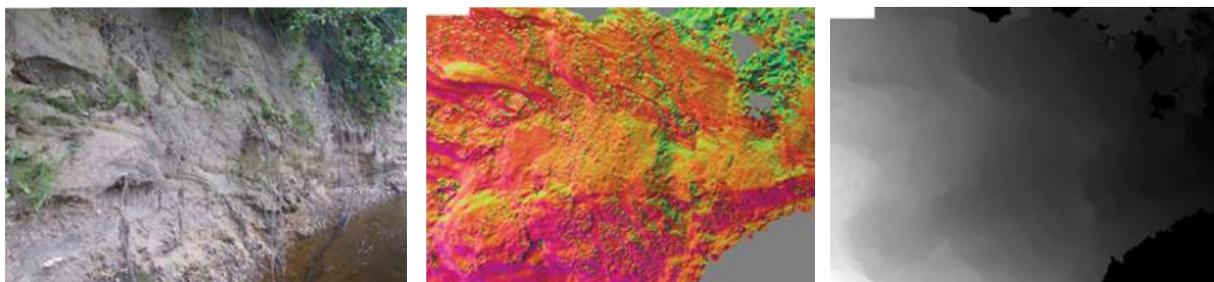
O número total de imagens nos clusters de saída é minimizado e as imagens redundantes são removidas da reconstrução de modo que todos os pontos SfM dentro dela já estejam bem reconstruídos (determinados pelas linhas de base da câmara e taxas de amostragem de pixels) em pelo menos um “cluster” de imagem. Um tamanho máximo de “cluster” de imagem é especificado de forma que cada cluster seja pequeno o suficiente para a reconstrução do MVS. Outros filtros de pontos são aplicados no processo de mistura de “clusters” após a aplicação dos algoritmos MVS, conforme detalhado na seção a seguir, de modo que apenas pontos de qualidade relativamente alta sejam retidos com uma consistência de visibilidade entre “clusters” aplicada (Furukawa et al., 2010).

### 1.3.8. Algoritmos de correspondências de imagem MVS

Uma nuvem de pontos esparsa gerada pelo SfM geralmente é apenas uma etapa intermediária na produção de nuvens de pontos muito mais densas usando o MVS. O objetivo do MVS é fornecer uma reconstrução completa da cena 3D a partir de uma coleção de imagens de parâmetros intrínsecos e extrínsecos da câmara conhecidos. Em comparação com uma nuvem de pontos esparsa gerada por SfM, uma nuvem de pontos densa gerada por MVS mostra um aumento na densidade de pontos de pelo menos duas ordens de magnitude.

Segundo Seitz et al. 2006, existe uma grande variedade de algoritmos MVS, e estes geralmente podem ser divididos em quatro classes:

1. Métodos baseados em “voxel” representam o volume da cena 3D diretamente usando grades de ocupação de “voxel” (Seitz e Dyer, 1999). Esses métodos são relativamente simples, mas são limitados em precisão pela resolução da grade de “voxels” e exigem conhecimento da caixa delimitadora que contém a cena.
2. Os métodos baseados na evolução da superfície usam malhas poligonais deformáveis que são evoluídas iterativamente para minimizar uma função de custo (Furukawa e Ponce, 2009). Esses algoritmos baseados em evolução de superfície requerem uma inicialização (por exemplo, usando um modelo visual de casco) que limita sua aplicabilidade, especialmente em cenas de grande escala (Shen, 2012).
3. Métodos de mistura de mapas de profundidade computam mapas de profundidade individuais para cada imagem que são então combinados em um único modelo 3D (Li et al., 2010). Um mapa de profundidade é uma imagem que representa a distância do ponto de vista aos objetos da cena 3D (fig. 9). Esses algoritmos evitam a necessidade de reamostragem em um domínio 3D e são mais flexíveis para cenas lotadas.
4. Métodos baseados em “patches” representam cenas por coleções de pequenos “patches”, que são simples e eficazes e não requerem inicialização (Lhuillier e Quan, 2005).



a) Margem no Rio Mersey

b) Vetor normal calculado

c) Mapa de profundidade

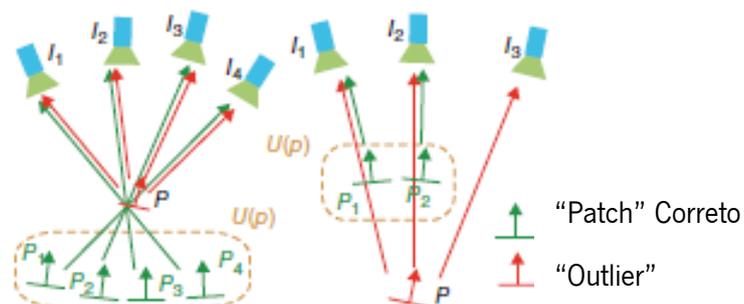
**Figura 9.** Reconstrução do mapa de profundidade (fonte: Carrivick et al. 2016)

Furukawa e Ponce (2010) descrevem um algoritmo MVS baseado em “patch” (PMVS) que é amplamente utilizado e teve um bom desempenho em testes comparando algoritmos MVS (Ahmadabadian et al., 2013). O algoritmo PMVS procede em três etapas principais, que devido ao seu amplo uso são descritas brevemente aqui: (i) recursos de correspondência, (ii) expansão de patches e (iii) filtragem de correspondências incorretas.

Primeiro, na etapa de correspondência, os operadores de diferença de Gaussian e Harris são usados para detectar recursos de canto e “blob”, que são combinados em várias imagens. A consistência fotométrica local é avaliada por correlação cruzada normalizada. Devido à falta de regularização na etapa de geração do “patch”, a reconstrução densa do PMVS depende de informações de textura confiáveis que podem resultar em lacunas na nuvem de pontos densa em superfícies de textura pobre.

Em segundo lugar, na etapa de expansão, a partir dessas correspondências iniciais (manchas esparsas), os “pixels” vizinhos nas imagens onde a mancha é projetada são considerados para expansão. Isso cria as manchas densas, expandindo a reconstrução. A expansão não ocorre onde células de imagem vizinhas já estão reconstruídas ou onde ocorrem descontinuidades de profundidade.

Terceiro, na etapa de filtragem, as restrições de visibilidade são usadas para filtrar correspondências incorretas levando em conta a oclusão nos modelos. A consistência da visibilidade global é reforçada pela filtragem de “patches” discrepantes, conforme demonstrado na figura 10. Filtros adicionais são aplicados para remover manchas visíveis em apenas um pequeno número de mapas de profundidade (Furukawa e Ponce, 2010).



**Figura 10.** Exemplo do filtro de consistência de visibilidade usado em PMVS (adaptado de Furukawa et al. 2010)

No PMVS, o processo de expansão e filtragem é repetido várias vezes (normalmente três). Outros métodos baseados em “patches” substituem esse processo iterativo de expansão e filtragem por procedimentos de expansão gananciosos. Em todos os casos, o resultado é uma nuvem de pontos densa

que possui uma densidade de pontos semelhante aos dados do scanner a laser terrestre (Lhuillier e Quan, 2005).

O estudo realizado por Carrivick et al. (2016) evidencia que os desenvolvimentos contínuos em muitas das etapas de processamento (por exemplo, correspondência de recursos, algoritmos de ajuste de pacote e algoritmos MVS) sugerem que mais refinamentos e melhorias serão feitos à medida que as velocidades de processamento do computador aumentam, os requisitos de memória aleatória reduzem e as densidades e a precisão dos pontos aumentam. Os usuários finais devem estar cientes dos parâmetros arbitrários empregados no fluxo de trabalho SfM-MVS, que têm um grande efeito na velocidade de processamento de dados e na densidade e precisão da nuvem de pontos resultante. À medida que a tecnologia amadurece e os profissionais exigem uma maior capacidade de ajustar esses parâmetros, é provável que a situação mude no futuro. Atualmente, os pacotes de código aberto geralmente oferecem ao usuário maior transparência quanto ao fluxo de trabalho específico implementado do que o software disponível comercialmente.

#### **1.4. Sensores**

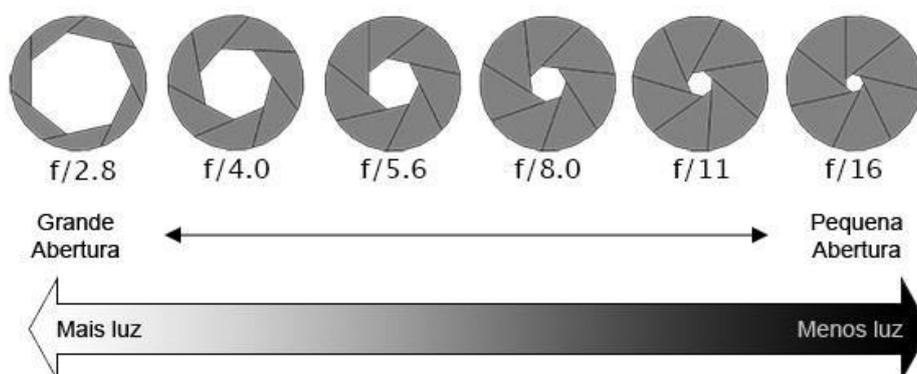
As imagens necessárias para SfM-MVS podem ser adquiridas de praticamente qualquer sensor. As câmaras compactas prontas para uso (Eisenbeiss & Zhand 2006), câmaras de smartphones (Klein & Murray 2009) e câmaras de trilha dominam, e onde a maioria dos usuários de SfM-MVS se encontra, não apenas devido à facilidade e custo-benefício, mas também porque a qualidade dos produtos 3D derivados muitas vezes pode competir com aqueles adquiridos de sensores mais caros (Thoeni et al. 2014).

O objetivo final é obter muitas fotografias bem expostas do recurso de interesse em resolução suficiente para que o algoritmo de correspondência seja capaz de funcionar de forma eficaz. A resolução ideal é, portanto, determinada pela geometria de visualização (distância do recurso), o tamanho do recurso (em relação a essa distância), as condições de iluminação e o número de imagens que serão adquiridas (muita imagem em baixa resolução pode ser preferível a menos imagens em boa resolução). É claro que a resolução da imagem, o contraste e a nitidez do pixel tendem a melhorar com o custo da câmara ou sensor, mas muitos dos recursos aprimorados de uma “Digital Single Lens Reflex” (DSLR) de ponta (por exemplo, resolução de imagem 30–40 megapixéis) tornam-se redundantes se arquivos grandes tamanhos estendem os tempos de processamento além de um nível aceitável. Nesses casos, muitos usuários optam por degradar a resolução de suas imagens antes do processamento, percebendo que uma grande economia no tempo de processamento pode levar a apenas uma leve degradação na qualidade da nuvem de pontos. Em baixas altitudes, a maioria das unidades prontas para uso são

capazes de atingir uma resolução de solo de centímetros a decímetros e ter estabilidade suficiente para produzir imagens de alta qualidade adequadas para derivação de nuvens de pontos (Shortis et al., 2006). As câmaras DSLR e de distância focal fixa oferecem as imagens mais apropriadas para SfM-MVS devido à sua alta qualidade de imagem (James & Robson 2012). A geometria interna (ou modelo da câmara) é mais fácil e comumente replicada para lentes grande angulares (equivalente a cerca de 35 mm em uma SLR tradicional), enquanto aquelas com lentes mais longas (ou seja, cerca de três ou mais vezes a distância diagonal do sensor) ou mesmo lentes olho de peixe requerem modelos ou algoritmos sob medida (Micusik & Pajdla 2006). Por esse motivo, embora sejam um dos sensores mais populares para captura de dados, as câmaras GoPro demonstraram produzir dados 3D comparativamente maus (Thoenig et al., 2014). Dito isto, as comparações entre nuvens de pontos geradas a partir de uma variedade de sensores mostraram que câmaras DSLR caras também não garantem resultados de alta qualidade, com câmaras compactas prontas para uso superando unidades mais caras em vários casos (Thoenig et al., 2014). A qualidade da estratégia de captura de dados pode, portanto, ser tão importante quanto a qualidade do sensor de imagem (Micheletti et al., 2014).

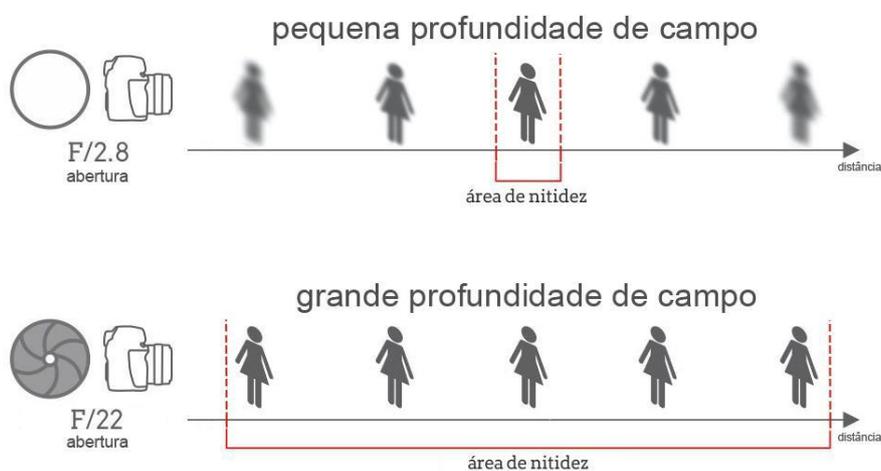
Mendonça 2017 analisou algumas características e definições ajustáveis das máquinas digitais que influenciam as fotografias e por consequência a fotogrametria digital. As características e definições abordadas são: Abertura, profundidade de campo, sensibilidade à luz (ASA, DIN e ISO), velocidade, lente e deformações, número de megapixéis.

**Abertura:** é representada por um rácio entre a abertura (**f**) e distância de foco (**n**) (**f/n**: f/2, f/8, f/16). Sendo que **f** (abertura) representa a abertura do diafragma que se encontra no interior do corpo da máquina digital. A figura 11 demonstra a relação entre a abertura, sua designação e a luz que esta deixa passar para o sensor da câmara digital.



**Figura 11.** Esquema abertura e relação com a luz (Mendonça, 2017)

**Profundidade de campo:** é inversamente proporcional à abertura. Quanto menor a abertura maior será a profundidade de campo. A profundidade de campo é o conjunto de distâncias em que os pixels se distinguem e formam uma imagem nítida. A Figura 12 representa um esquema explicativo da profundidade de campo, o seu efeito e a relação com a abertura representada na Figura 11.

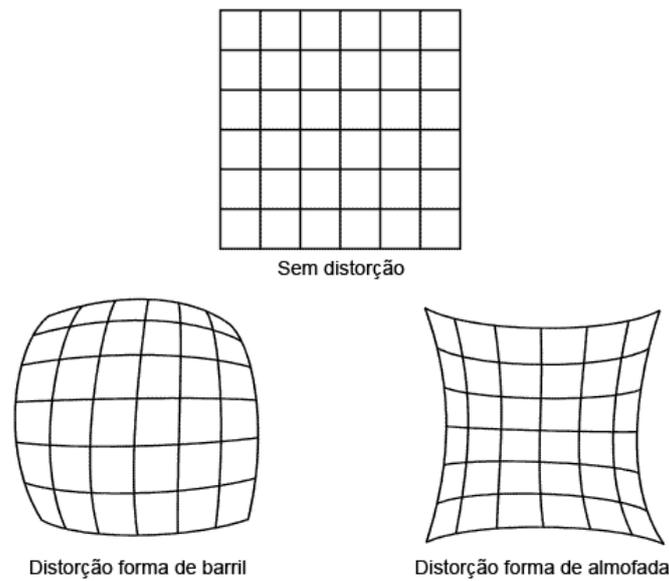


**Figura 12.** Esquema profundidade de campo (Mendonça, 2017)

**Sensibilidade à luz** do sensor ótico: é um parâmetro que pode ser ajustado, e é medido em ASA, DIN ou ISO. Sendo o mais antigo o ASA, e é uma escala aritmética, e o DIN uma escala logarítmica. A escala de medida mais recente combina a aritmética e o logarítmico, e é designada como ISO. Uma sensibilidade maior provoca efeitos secundários, com o aumento de ruído na imagem.

**Velocidade:** representa a velocidade de abertura e fecho do obturador que controla o tempo de exposição da luz ao sensor ótico, quanto maior a velocidade menor tempo de exposição à luz. Este parâmetro é medido em frações de segundos. Como exemplo, uma grande velocidade é útil perante um cenário em que não haja hipótese da utilização de um tripé, suporte ou estabilizador, pois a captura da fotografia é feita em frações de segundos, fazendo com que o risco das fotografias saírem desfocadas seja menor.

**Distorções:** podem ser introduzir nas imagens através das lentes, sendo que há lentes em que estas distorções são praticamente nulas, mas em equipamentos de qualidade inferior estas distorções podem existir. O diafragma impede que os raios de luz oblíquos passem pelo centro da objetiva e, como as superfícies das lentes não são paralelas nos bordos, a luz que forma a imagem inclina-se. Na Figura 13 são representados os dois extremos das distorções mais comuns: a distorção forma de barril e forma de almofada.



**Figura 13.** Distorções mais comuns (Mendonça, 2017)

**Megapixéis:** representam o número de pixéis, em milhões, que a máquina consegue reproduzir nas fotografias captadas. Este valor é diretamente proporcional ao número de sensores na matriz do sensor ótico. Numa câmara digital (DSLR), o tamanho da matriz é maior podendo atingir centímetros de largura e altura, em comparação com os telemóveis em que a matriz atinge apenas milímetros, se não micromilímetros de largura e altura. Isto faz com que estes sensores, mesmo tendo a mesma quantidade de megapixéis, tenham densidades de pixéis diferentes. O megapixel não mede a qualidade da imagem, apenas mede o número de “pontos” onde é captada a informação.

O estudo realizado por Bemis et al. (2014) mostraram que três parâmetros principais afetam a qualidade do modelo 3D:

1. Condições de iluminação – brilho de superfícies refletivas e contraste variável em uma cena podem afetar negativamente a correspondência de pontos.
2. Mudanças no comprimento da sombra e no albedo da superfície como resultado do posicionamento solar também impactam negativamente na correspondência de características. Idealmente, as pesquisas devem ser concluídas em menos de 30 minutos e continuadas nos dias seguintes, se houver condições climáticas semelhantes. Da mesma forma, o operador de câmara deve evitar fazer sombra sobre a área de estudo.

3. Estações limitadas e/ou mal distribuídas resultam em distorções do modelo e/ou áreas de dados perdidos.

Planejar uma pesquisa de campo para SfM-MVS é, portanto, fundamental. Sempre que possível, as imagens para SfM-MVS devem ser:

- Adquirido sob iluminação constante.
- Onde as sombras em movimento e o flash da câmara são minimizados.
- Onde o objeto de interesse é fixo.
- Onde o objeto ou superfície de interesse contém variações de textura e cor.
- Onde o objeto ou superfície alvo é fotografado com imagens sobrepostas tiradas em vários ângulos desse objeto ou superfície.

Além disso, a textura do objeto ou cena é importante e, em particular, a homogeneidade da composição de cores. O SfM-MVS depende dessas diferenças de textura para combinar os recursos, se o objeto ou cena for relativamente desprovido de recursos, o algoritmo de transformação de recursos invariantes de escala (SIFT).

### **1.5. Vantagens e Desvantagens do SfM-MVS**

Em resumo, as vantagens do SfM-MVS em comparação com outros métodos de levantamento digital são (Carrivik et al., 2016):

- SfM-MVS é barato.
- O SfM-MVS é fácil em termos de aquisição de dados de campo e operação de software comercial ou de código aberto.
- SfM-MVS produz dados totalmente 3D, caso contrário, só é possível com TLS.
- Nuvens de pontos derivadas de SfM-MVS e superfícies texturizadas são inerentemente multidimensionais (x, y, z, orientação de ponto, cor, textura) e relativamente facilmente transformadas em ortofotografias.
- Com aplicação cuidadosa, especialmente na aquisição de GCPs, que exigem conhecimento especializado de TS ou dGPS, o SfM-MVS pode fornecer precisão comparável com qualquer outro método de levantamento topográfico (direto ou indireto).

As desvantagens do SfM-MVS em comparação com outros métodos de levantamento digital são as seguintes (Carrivick et al., 2016):

- SfM-MVS apresenta desafios para aquisição e processamento de dados em grandes escalas espaciais.
- A precisão 3D dos dados derivados do SfM-MVS depende de fatores fora do controle e isso levanta questões de repetibilidade.
- O SfM-MVS só produz dados 3D “de volta ao escritório”, portanto, no momento da pesquisa, não se pode saber se uma nuvem de pontos será produzida com sucesso ou quais atributos (extensão, resolução, qualidade de ponto 3D) essa nuvem de pontos terá.
- O SfM-MVS não discrimina localizações de pontos ou tipos de pontos, nem analisa informações de nuvem de pontos automaticamente.
- A visualização eficiente dos dados SfM-MVS é problemática. O software para análise de dados de nuvem de pontos está em sua infância.

## 2. Materiais e Métodos

Um dos objetivos principal desta dissertação é avaliar e comparar a eficiência, precisão e limitações dos softwares selecionados de modo a ajudar na escolha do software mais adequado.

Os materiais utilizados na parte prática apresentados a seguir, foram selecionados de um leque restrito de equipamento disponível para a realização da presente dissertação: Drone (DJI Mavic Air), computadores portáteis, câmara digital (Iphone 11 Pro e Redmi 9), provete cilíndrico, fachada do Bar Académico, VisualSFM, Autodesk Recap, Colmap, Meshroom, Regard3D, Metashape e CloudCompare.

- Drone: DJI Mavic Air, equipada com uma câmara de 12 MP com as seguintes configurações: dimensão das fotos 4056x2280 pixels, resolução 72PPP, abertura f/2.8, tempo de exposição 1/140 seg, ISO-100, distancia focal 4mm e sem flash.
- Câmara Digital: Iphone 11 Pro, equipada com uma câmara de 12 MP com as seguintes configurações: dimensão das fotos 3024x4032 pixels, resolução 72PPP, abertura f/1.8, tempo de exposição 1/50 seg, ISO-250, distancia focal 4mm e sem flash.
- Câmara Digital: telefone Redmi 9 com as seguintes características da câmara: 13 MP, f/2.2, 28mm (wide), 1.0µm, PDAF
- Computador Portátil (Metashape): Placa gráfica: NVIDIA GTX 1650. Processador: Intel® Core™ i7-4790CPU @ 3.60GHz. 32GB de memoria RAM.
- Computador Portátil: Placa gráfica: NVIDIA® GeForce® MX250 | Intel® UHD Graphics 620. Processador: Intel® Core™ i5-8265U Quad Core. 8GB de memoria RAM. Sistema operativo de 64 bits.
- Computador de secretária: Placa gráfica Amd Radeon™ RX Vega 11 Graphics. Processador: AMD Ryzen 5 3400G @ 3.70 GHz. 8GB de memoria RAM. Sistema operativo de 64 bits.
- Modelos utilizados: Provete de granito, fachada do Edifício da Associação de Estudantes da Universidade do Minho em Guimarães e maciço rochoso.

- Softwares: VisualSfM, Colmap, Regard3D, Meshroom, Autodesk Recap, Metashape e CloudCompare.

## **2.1. Software SFM**

### **2.1.1. VisualSFM**

É um software GUI gratuito e de código aberto de reconstrução 3D através de imagens usando a abordagem SFM, desenvolvido pelo Chanchang Wu (Wu, 2011) em colaboração com a Universidade de Washington e a Google. O sistema de reconstrução integra os seus projetos anteriores como SIFT on GPU(SiftGPU), Multicore Bundle Adjustment e Towards Linear-time Incremental Structure from Motion (Wu, 2007, 2011, 2013).

Segundo Wu (2013), VisualSFM deteta os recursos, a sua correspondência e ajuste de pacotes. O software integra na sua interface o “plug-in” PMVS (desenvolvido por Yasutaka Furukawa e Jean Ponce) e CMVS (desenvolvido por Yasutaka Furukawa) para gerar a nuvem densa (Morgan e Brogan, 2016).

Este software permite uma reconstrução totalmente automática a partir da abordagem SFM, iniciado pelo conjunto de imagens tiradas de vários ângulos e posições, terminando com a geração de nuvens de pontos 3D (Olkowicz et al., 2019).

**Fluxo de trabalho:** O fluxo de trabalho começa com o carregamento das imagens no software (etapa 1), importando diretamente várias imagens. VisualSfM determina automaticamente todos os parâmetros usados pela câmara para adquirir as fotos. Na etapa seguinte (etapa 2), o VisualSfM deteta recursos em cada imagem e encontra correspondências. Ele fornece uma variedade de algoritmos diferentes para detecção de recursos, incluindo “Scale Invariant Feature Transform” (SIFT) e SiftGPU (uma implementação GPU do SIFT) (Hafeez et al. 2016).

As correspondências encontradas na etapa anterior são posteriormente convertidas em pontos no espaço 3D (etapa 3) obtidos por meio do ajuste do pacote. Esta etapa é feita usando o atalho Reconstrução Esparsa (fig. 14). Uma nuvem mais densa é obtida usando a ferramenta PMVS/CMVS (etapa 4) através do atalho Reconstrução Densa.

Este comando pede para salvar a saída como um arquivo \*.nvm, que o VisualSFM cria e salva em uma pasta intitulada \*nvm.cmvs e executa o CMVS (Visualização de “cluster” para Multi-view Stereo). Um arquivo \*.ply também é salvo automaticamente no mesmo local.



**Figura 14.** Fluxo de trabalho VisualSfM (adaptado de <http://ccwu.me/vsfm/doc.html>, visitado em 15 março 2022)

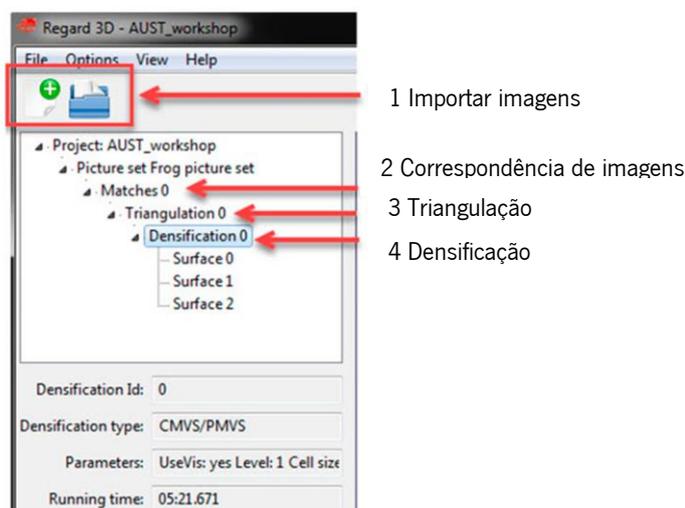
**Requisitos:** O software VisualSfM está disponível para para os sistemas operativos Windows, Linux e MacOS e não necessita de uma instalação ocupando 20Mb no disco. O software precisa do SiftGPU para detecção e correspondência de recursos, PBA para reconstrução esparsa e PMVS/CMVS para reconstrução densa.

### 2.1.2. Regard3D

Segundo Rahaman e Champion (2019), Regard3D é um programa SFM gratuito e de código aberto que suporta várias plataformas tais como o Windows, OS X e Linux. Apresenta uma interface gráfica de usuário (GUI) simples e direta. Dispõe de uma janela de árvores na sua interface onde é possível verificar uma tarefa concluída, as configurações utilizadas para gerá-la e o respetivo tempo de execução. Regard3D utiliza uma serie de imagens de diferentes pontos para desenvolver o modelo 3D. O software tem um tutorial completo no site, de forma a orientar os iniciantes (Setiyadi et al., 2021).

**Fluxo de trabalho:** Os detalhes das tarefas executadas são destacados na visualização em árvore à esquerda (Figura 5). Começa-se por definir um caminho de projeto e um nome para iniciar um projeto. As fotografias devem ser configuradas sendo necessário a introdução da largura do sensor do equipamento utilizado na captação de imagens na base de dados do software, caso não esteja (etapa 1), para que o software calcule as correspondências (etapa 2). Em seguida é realizado o processo de determinação da posição e orientação de cada câmara na cena (passo 3), selecionando os resultados da correspondência e clicando em Triangulação.

Com base nessa nuvem de pontos dispersa, é realizado a Densificação do resultado da triangulação (etapa 4). Na visualização em árvore, é possível destacar os resultados da etapa 4 e escolher 'Criar nuvem de pontos densa'. A nuvem densa (\*.ply, \*.pcd) pode ser exportada no final desta etapa.



**Figura 15.** Fluxo de trabalho Regard3D (adaptado de Rahaman e Champion, 2019)

O Regard3D tem uma GUI simples e direta. Experimentar configurações assim fica mais acessível, pois o usuário só precisa clicar em uma tarefa concluída para ver uma lista dos parâmetros definidos para gerá-la, bem como visualizar o tempo de execução daquela etapa selecionada.

Em seguida segue-se algumas recomendações de modo a obter bons resultados (Regard, 2022):

- Quanto mais imagens, obtém-se melhores resultados.
- As imagens devem ser tiradas numa posição diferente das outras.
- Imagens de maior resolução (pixéis) produz um modelo mais detalhado.
- O objeto deve ser fotografado em todos os ângulos possíveis de forma a evitar furos no modelo resultante.
- O modelo da câmara deve ser armazenado no JPEG e o modelo da câmara deve estar no banco de dados.

As imagens devem satisfazer os seguintes critérios para obter melhores resultados:

- A distância focal (definição de zoom) e o tamanho do sensor de pelo menos algumas das imagens devem ser conhecidos. Isso significa que:
  - Os metadados (EXIF) devem estar presentes, fornecendo o comprimento do foco, o fabricante da câmara e o modelo.
  - A câmara deve ser conhecida pelo Regard3D (está no banco de dados da câmara)

- As imagens devem ter uma boa qualidade: bem iluminadas, em foco e em alta resolução
- As imagens devem fornecer uma visão completa do objeto. Uma boa regra geral é: se algum ponto específico do objeto não for visto por pelo menos 3 imagens, ele não estará no modelo 3D resultante.

**Requisitos:** Está disponível para os sistemas operativos Windows (a partir do Windows 7) e Mac OS X, sendo necessário instalar o software no disco. O ficheiro de instalação possui 40MB e quando instalado o software ocupa 103 MB no disco.

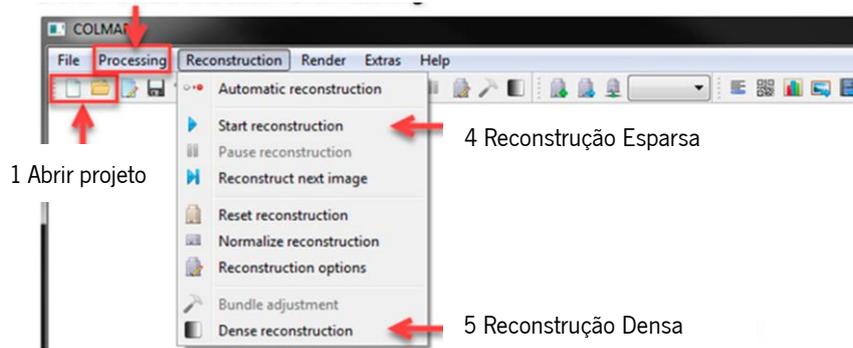
### 2.1.3. Colmap

COLMAP é um software de reconstrução 3D que utiliza o algoritmo SfM para calcular a orientação inicial das fotos e o MVS como base da sua operação. O software foi desenvolvido por Johannes L. Schoenberger (Schoenberger e Frahm, 2016) e está licenciada sob a nova licença BSD. É gratuito e encontra-se disponível para download no site GitHub. O processo de reconstrução pode ser feito através da linha de comando ou pela interface gráfica do usuário. O programa pode gerar uma malha 3D, mas também trabalha com uma nuvem de pontos densa (Schoenberger, 2020).

COLMAP oferece uma reconstrução automática através de um único clique. Este processo automático é muito mais rápido quando se compara ao processo passo-a-passo, mas apresenta um “trade-off” em termos de qualidade do modelo construído. Por outro lado, o processo manual fornece mais flexibilidade nas configurações e uma reconstrução densa mais precisa (Rahaman e Champion, 2019).

**Fluxo de trabalho:** Executa-se o arquivo \*.bat para abrir o programa seguido do menu arquivo para abrir/criar um novo projeto (fig. 16). Deve-se mostrar ao programa a localização das imagens e onde localizar o banco de dados (passo 1). O processo começa com extração de recursos (passo 2) no menu “Processing” e de seguida é feita a correspondência (passo 3). Em seguida, reconstrói-se as posições da câmara e é produzido uma nuvem de pontos dispersa. O comando “Start Reconstruction” está localizado no menu suspenso “Reconstruction” (passo 4). COLMAP produz a visualização 3D enquanto mostra as câmaras sendo adicionadas à cena enquanto simultaneamente forma a nuvem de pontos dispersa. Após concluir esta etapa, a nuvem esparsa pode ser exportada. Um ajuste de pacotes pode ser executado antes da densificação. A reconstrução densa (passo 5) contém três etapas, ou seja, não distorção, estéreo e fusão. Todos os modelos podem ser exportados como formato de arquivo \*.nvm, \*.out, \*.ply e \*.wrl.

2 e 3 Extração e correspondência de recursos



**Figura 16.** Fluxo de trabalho Colmap (adaptado de Rahaman e Champion, 2019)

**Requisitos:** O software possui 308 MB e não necessita ser instalado. Está disponível na sua versão 3.7 para mac (no cuda) e Windows (cuda e no cuda). Para Windows é recomendável o binário com Cuda de modo a obter um melhor desempenho do software, mas exige-se uma GPU Nvidia. Windows No-Cuda apresenta todas as funcionalidades, exceto a parte de reconstrução densa.

Entretanto foi desenvolvido uma versão baseada no Colmap Cuda (v 1.5) denominado Colmap-CL desenvolvido com a API OpenCL de modo a obter um desempenho acelerado e que pode ser executado em várias outras plataformas de GPU (AMD, Intel, NVIDIA, etc.) Em hardware semelhante, o COLMAP-CL geralmente processa dados quase tão rápido quanto a versão CUDA, mas às vezes mais rápido e às vezes mais devagar. A versão OpenCL é otimizada de forma diferente do CUDA COLMAP, portanto, o desempenho pode variar dependendo de suas GPUs e parâmetros de entrada específicos. (Colmap, 2022).

#### 2.1.4. Meshroom

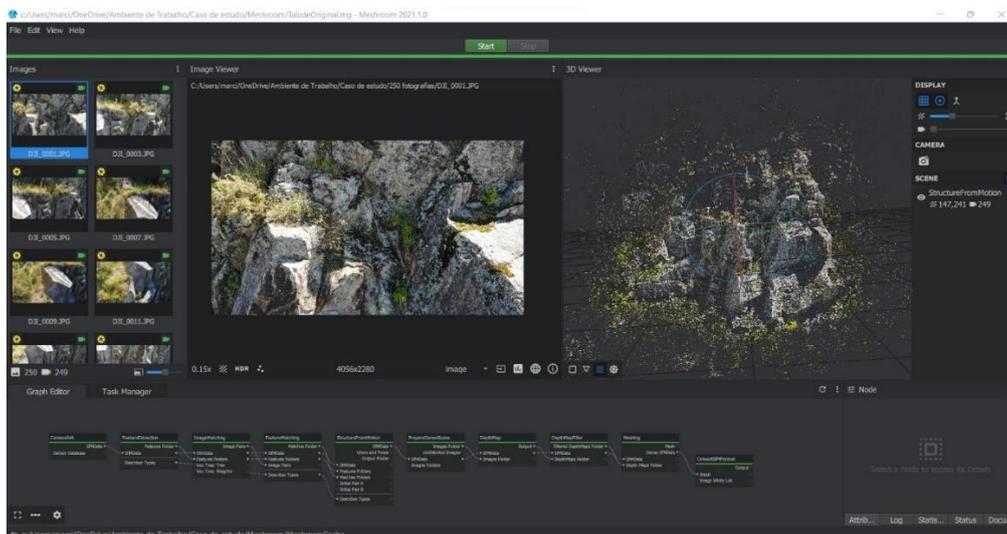
O Meshroom é um software de reconstrução 3D gratuito e de código aberto baseado na estrutura AliceVision1. AliceVision é uma estrutura de visão computacional fotogramétrica que fornece algoritmos de reconstrução 3D e rastreamento de câmara. O AliceVision apresenta uma forte base de software e algoritmos de visão computacional de última geração que podem ser testados, analisados e reutilizados. O projeto é resultado da colaboração entre academia e indústria para fornecer algoritmos de ponta com a robustez e a qualidade necessária para uso em produção. (Meshroom Contributors, 2022)

Meshroom fornece um ambiente nodal para realizar várias tarefas de visão computacional. As tarefas individuais são representadas por nós combinados em gráficos de dependência acíclica direcionados que

são denominados pipelines. Cada nó representa uma ferramenta implementada no AliceVision. (Griwodz et al., 2021)

**Fluxo de trabalho:** O fluxo de trabalho adotado para a criação da nuvem densa no Meshroom baseia-se no estudo realizado por Griwodz et al. (2021). A reconstrução da nuvem densa é composta por uma variedade de nós (fig 17). Cada nós corresponde as etapas de inicialização da câmara, extração de recursos, correspondência de imagens, correspondência de recursos, estrutura de movimento, mapeamento de profundidade (consistindo em preparação, mapeamento e filtragem), geração de malha e em seguida é exportada a nuvem de pontos densa. O processo inicia-se com a importação das imagens para o software, arrastando e soltando-as no painel Imagens (esquerda). Deve-se criar um caminho onde trabalho vai ser guardado.

Através de um único clique botão “Start” (topo) é executado o “pipeline” de reconstrução. Uma barra de progresso aparecerá sob o botão. Quando a barra de progresso chega ao fim, a reconstrução está concluída.



**Figura 17.** Fluxo de trabalho Meshroom

**Requisitos:** Para utilizar totalmente o Meshroom, uma GPU compatível com NVIDIA CUDA é recomendada. Os binários são construídos com CUDA-10 e são compatíveis com a capacidade de computação 3.0 a 7.5. Sem uma GPU NVIDIA compatível, apenas "Draft Meshing" pode ser usado para reconstrução 3D.

### **2.1.5. Autodesk Recap Photo (2021)**

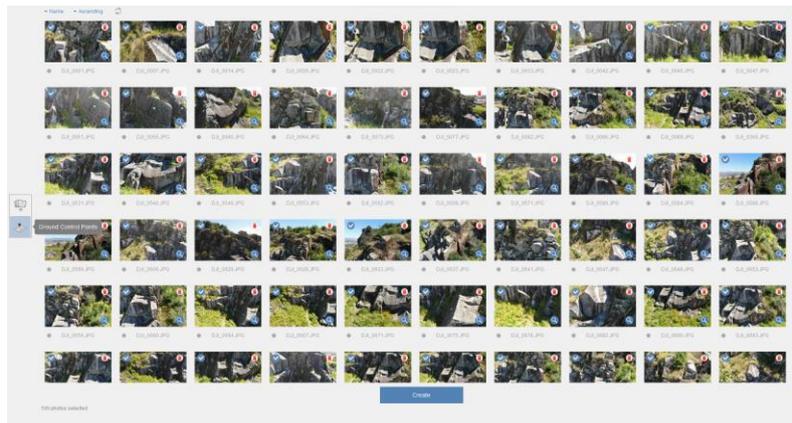
O ReCap Photo é um mecanismo de fotogrametria baseado em nuvem integrado disponível por meio de uma assinatura do ReCap Pro, construído exclusivamente para lidar com fotografias utilizando o poder da computação em nuvem e é exigida uma conexão com a internet.

Recap Photo é capaz de reconstruir modelos 3D aéreos e de curto alcance com recursos avançados para gerar malhas de textura geolocalizadas, nuvens de pontos e fotos ortogonais. Também possui ferramentas poderosas para visualização e edição de malha. Através do aplicativo ReCap Pro, é possível editar a nuvem de pontos RCS e exportá-la para outros formatos (e57, pts, ptg).

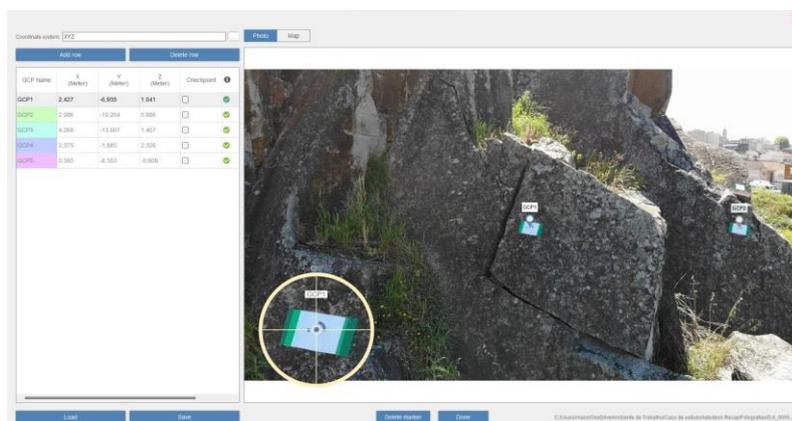
O ReCap Pro pode ser adquirido como uma licença mensal, de um ano ou de três anos. Além disso, o ReCap Photo limita a quantidade de processamento que um único usuário pode fazer, fornecendo um número limitado de “créditos na nuvem”. Se um usuário precisar de mais do que o valor inicial, ele deverá comprar mais do ReCap. (Autodesk Support, 2022)

Embora o objetivo da presente dissertação é testar os softwares gratuitos, considerou-se ser vantajoso incluir o Autodesk Recap (software comercial) pelo acordo da Universidade do Minho com a Autodesk que me permite o uso gratuito do software, mas com limitações. Para a realização desta dissertação é utilizada a versão educacional da Autodesk RecapPro (2021), com um limite de 100 fotos por projeto para o processamento da nuvem.

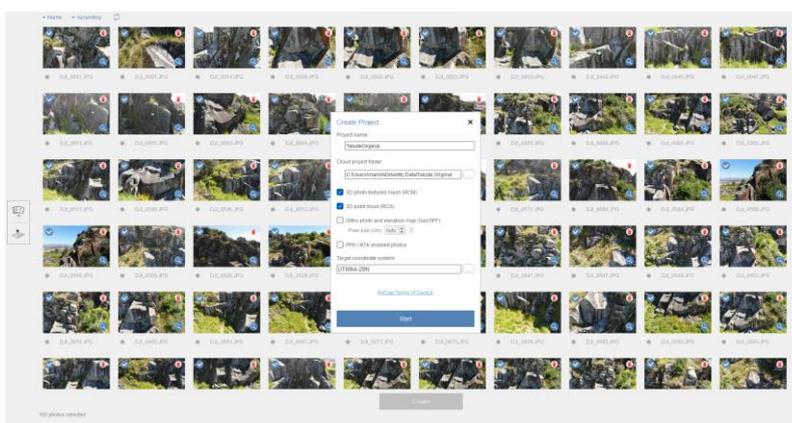
**Fluxo de trabalho:** Como acontece nos outros softwares, o processo inicia-se com a importação das imagens para o software Autodesk Recap Photo (1ª etapa). Em seguida é feita a georreferenciação (2ª etapa), identifica-se todos os pontos de controle (GCP) presentes na fotografia e são atribuídos os números correspondente a cada ponto de controle. Na 3ª etapa é atribuído um nome ao projeto, é definida a pasta do projeto na nuvem e os modelos 3D pretendidos.



**Figura 18.** Fluxo de trabalho Autodesk Recap



**Figura 19.** Processo de georreferenciação do modelo 3D Autodesk Recap



**Figura 20.** Reconstrução do modelo 3D Autodesk Recap

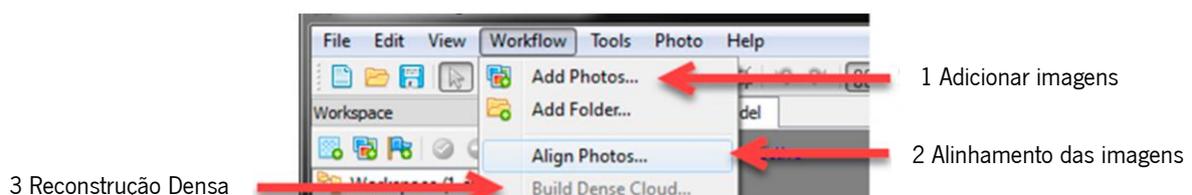
O RECAP não fornece controle do usuário sobre o “pipeline”. Todos os parâmetros que podem ser definidos em imagens de alinhamento, construção de nuvens densas, construção de modelos e geração de estágios de textura são definidos quase automaticamente. Os modelos podem ser criados como malhas texturizadas de foto 3D, nuvem de pontos ou fotos ortomosaicas e mapas de elevação com a

opção de especificar sistemas de coordenadas de destino. As ferramentas de edição no ReCap Photo (v.2021) permitem o controle sobre a qualidade da nuvem de pontos por meio de edição de valores discrepantes, configuração de caixa de limite com base na distância, volume de ponto único e valores de intensidade.

### 2.1.6. Agisoft Metashape

Agisoft Metashape é um software de reconstrução 3D comercial de baixo custo da Agisoft LLC, Rússia. (Agisoft, 2021). O Metashape cria automaticamente modelos 3D texturizados precisos usando fotos digitais (métricas e não métricas) de um objeto ou uma cena, disponível nas versões Standard e Pro. Este programa funciona em sistemas operacionais Windows, Mac OS e Linux em um PC local e, portanto, todos os dados permanecem com o usuário. Em alguns casos, é difícil ou mesmo impossível gerar um modelo 3D de todo o objeto em uma única tentativa. Para superar essa dificuldade, o Metashape oferece opções para dividir o conjunto de fotos em vários "pedaços" separados dentro do projeto. Dessa forma, o processamento/etapas padrão podem ser executados em cada pedaço separadamente e, em seguida, os modelos 3D resultantes podem ser combinados. (Rahaman e Champion, 2019)

**Fluxo de trabalho:** As fotografias são adicionadas ao fluxo de trabalho (1ª etapa), e o alinhamento da imagem é necessário antes da computação (2ª etapa) (fig. 21). No entanto, Metashape recomenda mascarar todos os elementos irrelevantes nas fotos de origem (como o plano de fundo e qualquer primeiro plano acidental) para obter melhores resultados de reconstrução. Na 3ª etapa, Metashape calcula as fotografias e constrói a geometria (cria uma nuvem de pontos) da cena. Nessa etapa pode ser feita edição e limpeza da nuvem de pontos desnecessária. A densidade da nuvem de pontos (gerada a partir da 3ª etapa) pode variar de formato normal, médio a ultra-alto e, em seguida, é feita a exportação do modelo.



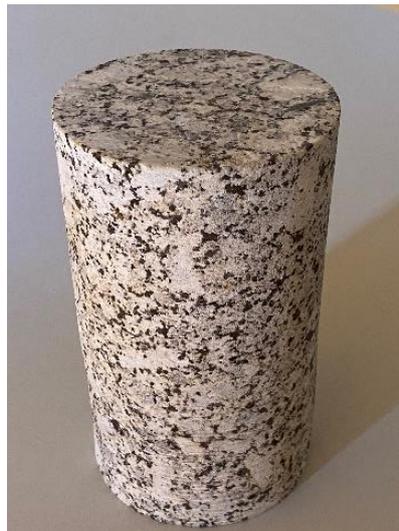
**Figura 21.** Fluxo de trabalho Metashape (adaptado de Rahaman e Champion, 2019)

## 2.2. Caso do Estudo

Um dos primeiros passos foi a exploração dos diversos softwares de fotogrametria, de forma a especificar vários critérios de avaliação e comparação tendo como critério a qualidade do modelo 3D obtido no final do processo, a facilidade de trabalho dos softwares, o tempo de processamento e os requisitos de hardware.

Para a resolução deste trabalho, foram reconstruídos nuvem de pontos 3D em todos os softwares (exceto Agisoft Metashape), de forma a ser realizada análises comparativas em relação a qualidade final dos modelos 3D e o tempo de processamento.

Para iniciar o processo de reconstrução, foi escolhido um provete de granito (fig. 22) de textura lisa, de padrão malhado útil para o processo fotogramétrico, com ângulos curvos e com uma superfície não refletora. A escolha foi feita com base nas características do objeto de modo a possibilitar boa visualização no modelo 3D.



**Figura 22.** Provete de granito

É realizada diversas experiências de modo a obter uma digitalização bem-sucedida. Inicialmente o provete foi posto em cima de uma mesa de base lisa branca e as fotografias foram capturadas a uma distância de 50 cm do objeto com a camara digital do Iphone 11Pro cujo as características das fotografias encontram-se na tabela 1.

Os nomes dos conjuntos têm uma letra seguida de um número que corresponde à ordem em que foram tirados os conjuntos, e a letra representa a primeira letra do nome do modelo usado, neste caso o **P** de provete.

**Tabela 1.** Especificações do sensor Iphone 11Pro

Nome conjunto	Base	Número fotografias (resolução pixéis)	Câmara	Lente	Abertura (f/n)	Tempo de exposição (s)	ISO
P1	Mesa de base lisa branca	50 (3024x4032)	Iphone 11 Pro	4mm	F/1.8	1/50	250

Em seguida, as fotos foram processadas em computador portátil com processador Intel i5, 8GB de memória RAM, placa gráfica NVIDIA GeForce MX250.

Foram feitas novas experiências explorando um outro tipo de máquina digital, lentes e base para auxiliar o processo da fotogrametria. A captação foi feita com o provete em cima de uma base de folha de jornal (fig. 23), utilizado com o intuito de ajudar na os softwares na identificação de recursos comum entre fotografias.



**Figura 23.** Proвете de granito (base com texto)

Na tabela 2, estão os 3 conjuntos de fotografias e as características de captação relacionada com camara digital utilizada. Todas as fotografias foram capturadas com a câmara digital de telefone Redmi 9. Foram capturados 3 conjuntos de 50 fotografias:

- O conjunto P2 foi tirada com distância focal (35mm) de 15mm a uma distância de 50cm do provete;
- O conjunto P3 foi tirada com distância focal (35mm) de 27mm a uma distância de 50cm do provete;
- O conjunto P4 foi tirada com uma distância focal (35mm) de 27mm variando a distância entre 50cm a 1m do provete;

**Tabela 2.** Especificações do sensor do Redmi 9

Nome conjunto	Base	Número fotografias (resolução pixéis)	Câmara	Distancia focal de 35 mm	Abertura (f/n)	Tempo de exposição (s)	ISO
P2	Com texto	50 (3264x2448)	Redmi 9	15mm	F/0	1	0
P3	Com texto	50 (4160x3120)	Redmi 9	27mm	F/0	1	0
P4	Com texto	50 (3120x4160)	Redmi 9	27mm	F/0	1	0
P5	Mesa de base lisa branca	50 (4160x3120)	Redmi 9	27mm	F/0	1	0

O hardware utilizado para gerar a nuvem de pontos é um computador de mesa placa gráfica AMD Radeon™ RX veja 11 Graphics. O software utilizado nesta experiência foi apenas o VisualSFM.

Uma vez que nos testes realizados anteriormente o provete foi colocado por cima de uma folha de jornal, realizou-se uma nova experiência com a mesma camara digital e características de captação utilizados anteriormente para o conjunto P3, alterando apenas a base para a captação de fotografias (fig. 24) para uma mesa de base lisa branca (conjunto P5) a fim de perceber a influência da base na qualidade do modelo final.



**Figura 24.** Provete de granito (base lisa)

Sendo um dos objetivos do presente trabalho a comparação entre os diferentes softwares SfM gratuitos foram realizadas mais pesquisas a fim de explorar as potencialidades de cada um. Essas pesquisas tiveram como enfoque o uso de máscaras e a capacidade para a execução de georreferenciação nos diversos softwares.

Foi realizado diversas experiências explorando as funcionalidades de uso de máscaras, destacando apenas o provete em si para a reconstrução e não a região envolvente de modo a obter melhores resultados no modelo final. A cor branca representa o provete (região de interesse) e a cor preta representa a região envolvente (nenhum interesse) como se pode verificar na Figura 25.



**Figura 25.** Provete de granito e a correspondente máscara

Na tabela 3, está apresentada as condições que foram assumidas na realização das experiências.

**Tabela 3.** Condições da experiência realizada

Experiência (nº)	1	2	3	4	5
Nº fotografias/ Nº máscaras	17/0	17/1	17/2	17/3	11/11

Em seguida é utilizado um conjunto de 10 imagens da fachada da sede da Associação de Estudantes da Universidade do Minho, situado em Guimarães na freguesia de Azurém (fig. 26), de modo a explorar mais funcionalidades dos softwares, nomeadamente a georreferenciação dos modelos. As fotografias continham as informações das coordenadas GPS do centro das câmaras.



**Figura 26.** Fachada da sede da Associação de Estudantes da Universidade do Minho (Pólo de Guimarães)

### **2.2.1. Enquadramento Geográfico e Descrição do Talude**

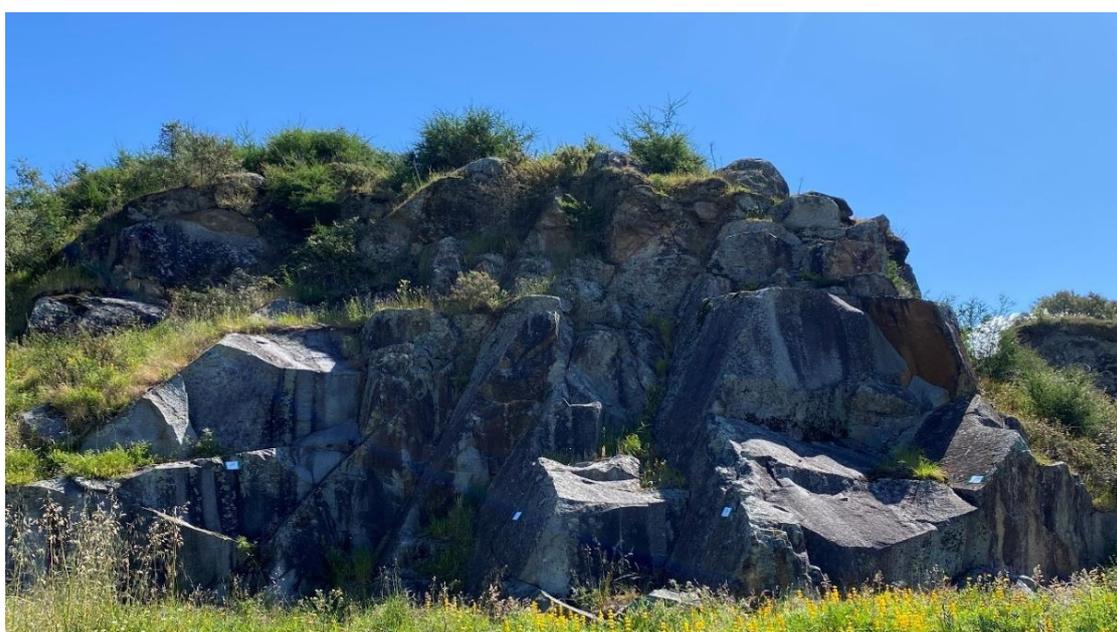
A área de trabalho da presente dissertação localiza-se na freguesia de Dume, junto ao Estádio Municipal de SCB, em Braga (fig. 27), Portugal. A encosta rochosa encontra-se na pedreira de granito de Monte Casto e apresenta um tamanho de aproximadamente 15 m de largura e 10 m de altura.

A escolha da área de estudo (fig. 28) levou em consideração a disponibilidade de informações prévias visto que o maciço a utilizar no estudo foi analisado por Riquelme et al. (2020) e a facilidade de acesso à área. As superfícies planares paralelas expostas representam todos os conjuntos de descontinuidades. A observação das descontinuidades mostra valores de longa persistência e ausência de pontes rochosas. Isso leva a que talvez essas descontinuidades sejam juntas de chapa (Riquelme et al., 2020).



**Figura 27.** Localização do talude (área vermelha marcada) (adaptado do Google Earth)

Segundo Riquelme et al. (2020), dois engenheiros geológicos experientes realizaram uma campanha de trabalho de campo em julho de 2018. Eles usaram métodos tradicionais para identificar os conjuntos de descontinuidade e medir sua persistência. Primeiro, eles identificaram dez conjuntos de descontinuidade. Em segundo lugar, eles mediram as orientações sobre o que eles consideravam planos (bem definidos como vistos) usando uma bússola, e os valores de persistência foram medidos usando uma fita. É importante notar que nos granitos são normalmente encontrados três conjuntos de descontinuidades juntamente com outros conjuntos de descontinuidades secundárias.



**Figura 28.** Imagem adquirida da frente de estudo (Talude)

### 2.2.2. Levantamento Aerofotográfico

O levantamento aerofotográfico do maciço foi realizado com o recurso a um VANT (Veículo Aéreo Não Tripulado) da marca Dji, modelo Mavic Air (fig. 29).



**Figura 29.** VANT utilizado para levantamento aerofotogramétrico

Este VANT está equipado com uma câmara digital de 12 megapixéis com uma distância focal fixa de 24mm e um intervalo de sensibilidade fotográfica (ISO) entre os 100 e 3200 em modo manual (tabela 4).

**Tabela 4.** Especificações do sensor do VANT

Resolução (pixéis)	Distancia focal de 35 mm	Abertura (f/n)	Tempo de exposição (s)	ISO
(4056x2280)	24mm	F/2.8	1	100 - 3200

Para que o levantamento fotográfico seja realizado corretamente é necessário que o sol não incida diretamente sobre a objetiva da câmara, sendo que a situação ideal seria o levantamento ser realizado num dia nublado.

As definições da câmara são reguladas manualmente e é configurada para a captura automática de 3 em 3 segundos de modo que as fotografias adquiridas consecutivamente tenham uma sobreposição

satisfatória para facilitar os softwares na identificação de recursos entre elas (figura 4 fotografias consecutivas).

### **2.2.3. Georreferenciação**

A georreferenciação do levantamento aerofotográfico da pedreira é realizada por várias etapas que têm início mesmo antes do levantamento efetuado pelo VANT. Em primeiro lugar procede-se à fixação de alvos impressos na superfície da rocha, e livre que qualquer obstrução visual, o que se chama de pontos de controlo, ou “Ground Point Control” (GPC), na língua inglesa. Para a realização deste trabalho foram usados cinco alvos de pontos de controlo disponibilizados pelo Metashape (fig 30).



**Figura 30.** Ponto de controlo GCP1

Pontos de controle de solo (GCPs) são coordenadas XYZ locais geralmente obtidas com uma estação total que são usadas como referência ao dimensionar um modelo. Os GCPs também são usados para representar a “verdade básica” ou o realismo do modelo em escala (Jones e Church, 2020).

As coordenadas dos cinco alvos da tabela 5 foram retiradas da nuvem de pontos obtido por Riquelme et al. (2020).

**Tabela 5.** Informação das coordenadas dos cinco GCPs

Coordenadas	X	Y	Z
GCP1	2.427 m	-6.935 m	1.041 m
GCP2	2.086 m	-10.204 m	0.886 m
GCP3	4.268 m	-13.607 m	1.407 m
GCP4	2.375	-1.885 m	2.326 m
GCP5	0.385	-8.353 m	0.808 m

Pode-se ver na figura 31 as localizações dos GCPs. A distância entre o GCP1 e GCP4 medida no terreno é de 5.2 m.



**Figura 31.** Localização dos cinco pontos de controle

#### **2.2.4. Modelação**

A modelação é a parte do trabalho em que se usa as imagens resultantes dos levantamentos aerofotográfico e os dados de georreferenciamento, de modo a obter a nuvem de pontos 3D da área de estudo. São utilizados seis softwares, nomeadamente VisualSfM, Colmap, Meshroom, Regard3D, Autodesk Recap e Metashape para a reconstrução da nuvem de pontos 3D do maciço, seguindo o fluxo de trabalho descrito no subcapítulo 3.1. A georreferenciação só foi feita nos softwares Autodesk Recap e Metashape.

Para reconstrução da nuvem de pontos densa nos softwares VisualSfM, Colmap, Meshroom, Regard3D e Autodesk Recap é utilizada computador portátil com processador Intel i5, 8GB de memória RAM, placa

gráfica NVIDIA GeForce MX250. A modelação no Metashape é utilizada computador portátil com processador Intel i7, 32GB de memória RAM, placa gráfica NVIDIA GTX 1650

Para determinar a precisão das nuvens de pontos 3D é feita uma comparação entre a nuvem de pontos derivadas dos cinco softwares pré-selecionados e a nuvem de pontos 3D de referência produzida a partir do software comercial Metashape.

### **2.2.5. Extração de elementos (CloudCompare)**

CloudCompare (Girardeau-Montaut, 2016) é software código aberto de processamento de nuvem de pontos 3D e de malha triangular disponível online por meio do “Open Source Project” (CloudCompare, 2020). O CloudCompare foi projetado para permitir que um usuário execute comparações entre duas nuvens de pontos densas ou entre uma nuvem de pontos e uma malha triangular. O CloudCompare conta com uma estrutura “octree” específica para essa tarefa. Desde então, ele foi expandido para incluir algoritmos de comparação mais avançados, como registo fino (seleção de ponto), reamostragem, manipulação de campo de cor/normal/escalar, computação estatística (distribuição gaussiana), gerenciamento de sensores, segmentação interativa ou automática e aprimoramento de exibição (CloudCompare, 2020).

O software CloudCompare no presente trabalho é utilizado para tratamentos, determinação do volume e curva de níveis das nuvens de pontos 3D.

Os modelos foram importados para o CloudCompare, como nuvens de pontos 3D. Na 1ª fase é feita a edição das nuvens de pontos 3D retirando os pontos irrelevantes da digitalização. Para realizar esta tarefa é utilizado a ferramenta “Segment” no menu “Edit”. Os pontos selecionados para remoção encontram-se no polígono verde (fig. 32).

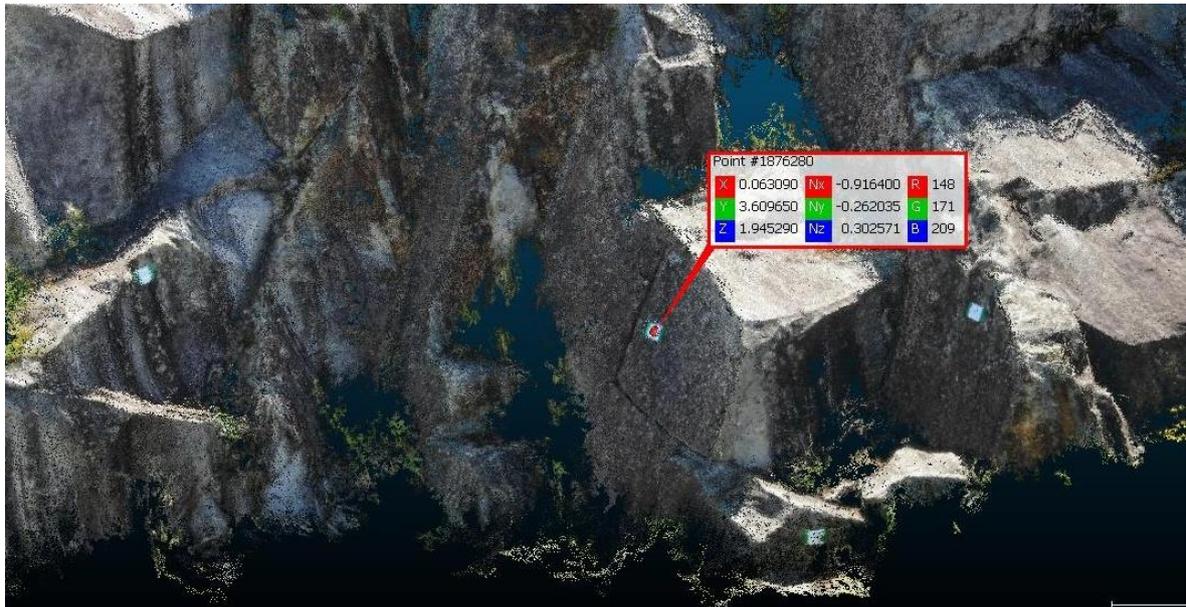


**Figura 32.** Tratamento da nuvem densa 3D



**Figura 33.** Nuvem densa 3D editada

Na 2ª fase é feita a georreferenciação das nuvens de pontos obtidos nos softwares VisualSfM, Colmap, Meshroom e Regard3D. Nas figuras 38 e 39, é possível verificar que a nuvem de pontos 3D (VisualSfM) não se encontra georreferenciada e escala. As coordenadas do ponto abaixo identificado ((fig. 34) não corresponde com as coordenadas GCP1 ( $x = 2.427\text{m}$  ;  $y = -6.935\text{m}$  ;  $z = 1.041\text{m}$ ). A distância real entre o GCP1 e GCP4 é de 5.2 m, o que não se verifica na figura 35.



**Figura 34.** Coordenadas do ponto de controle GCP1

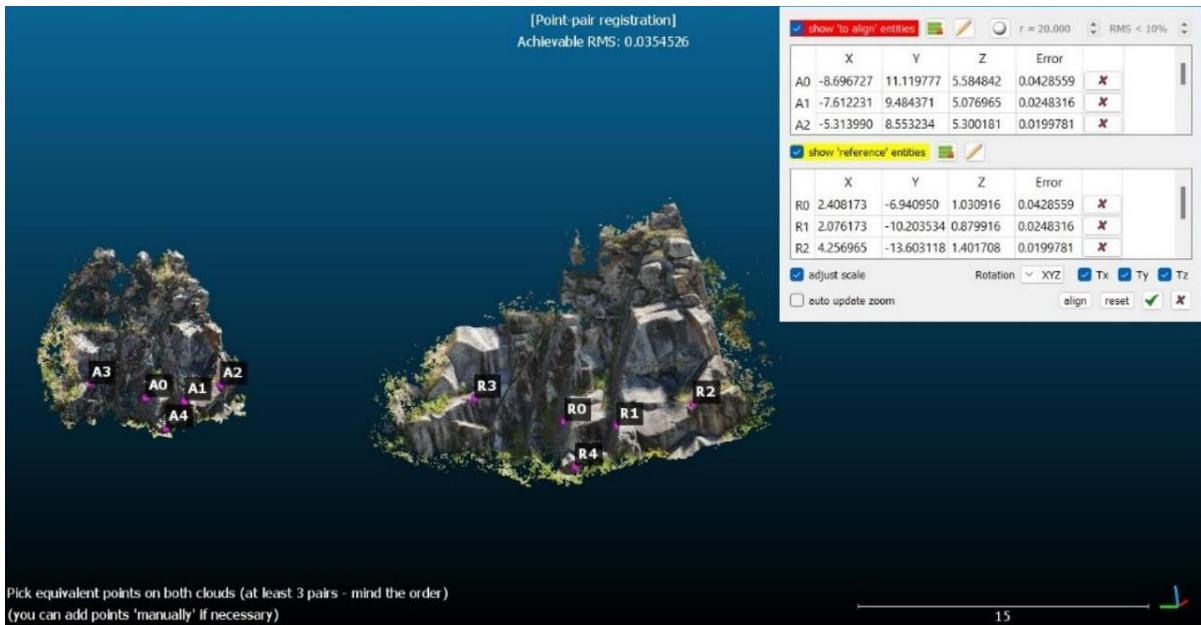


**Figura 35.** Distância entre GCP1 e GCP2

De modo a proceder a correta georreferenciação, a nuvem de pontos obtidos no Metashape que já se encontra devidamente georreferenciada e escalada, é utilizada como modelo base para georreferenciar e escalar as nuvens obtidas nos softwares VisualSfM, Colmap, Regard3D e Meshroom.

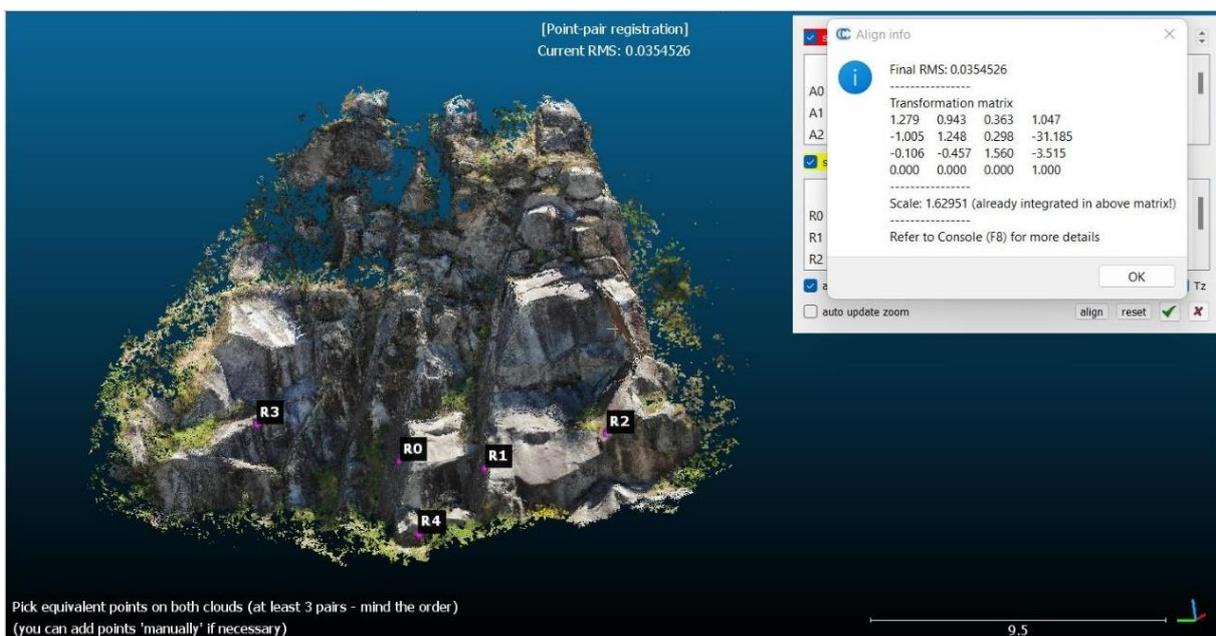
As duas nuvens de pontos 3D são carregadas no software CloudCompare. A nuvem da esquerda obtida com o VisualSfM e da direita Metashape (fig. 36). Esta tarefa é realizada utilizando a ferramenta “Registration > Align (point pairs picking)” no menu “Tools”. São selecionados cinco pontos GCP’s na nuvem de pontos 3D (VisualSfM) e em seguida são identificadas as referências na nuvem de pontos 3D

(Metashape). É selecionado o ajuste de escala e é realizada o alinhamento das duas nuvens de pontos. Na figura 37, as duas nuvens estão sobrepostas, devidamente escalada e georreferenciada com a informação do alinhamento.



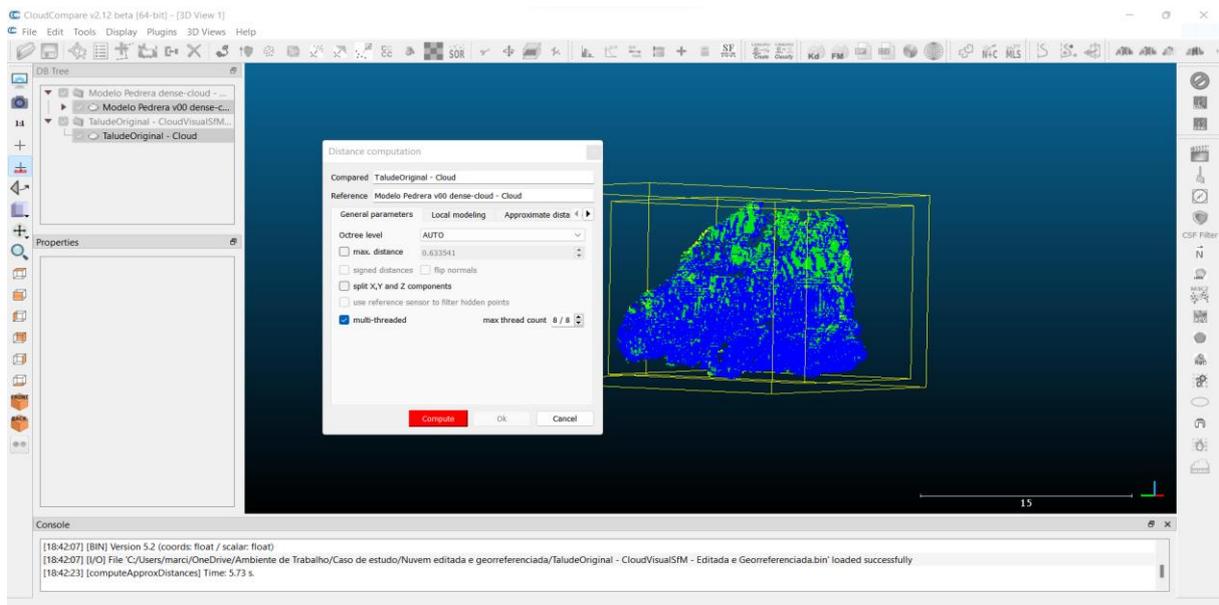
**Figura 36.** Georreferenciação das nuvens densa 3D

Na figura 41, as duas nuvens estão sobrepostas, devidamente escalada e georreferenciada com a informação do alinhamento. Os 2 processos acima descrito são realizadas para as restantes nuvens de pontos 3D (Colmap, Meshroom e Regard3D).



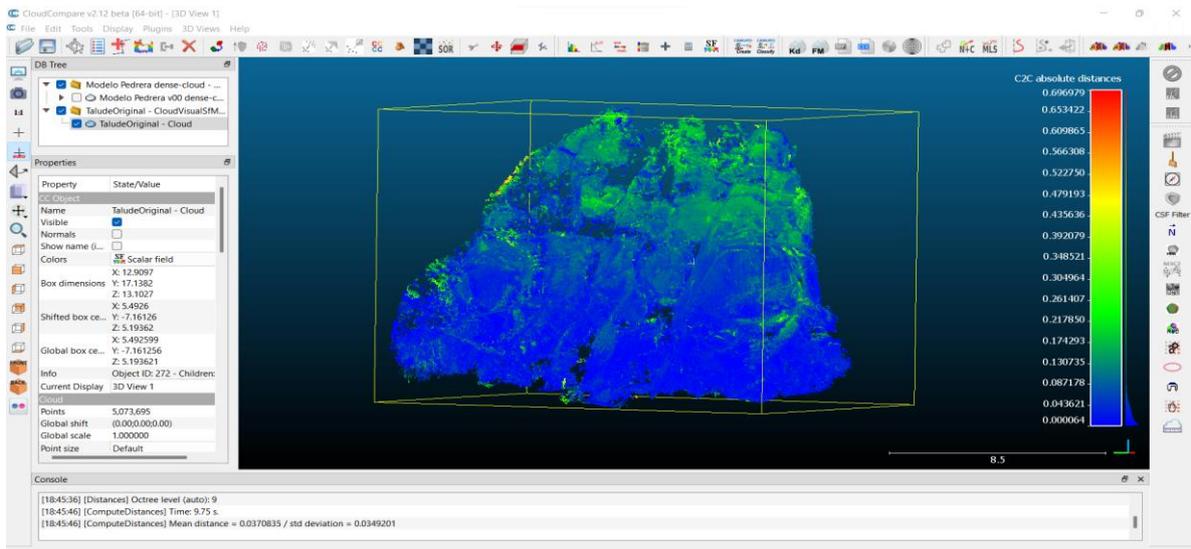
**Figura 37.** Sobreposição das nuvens densa 3D

É então somente após o registo e alinhamento precisos que as distâncias precisas podem ser calculadas entre os dois modelos como distâncias nuvem-nuvem. Para essa tarefa é utilizado a ferramenta “Compute Cloud-Cloud Distance” no menu “Tools - Distances”. Após selecionado a nuvem de pontos 3D de referência (Metashape) e a nuvem pretendida para fazer a comparação, o resultado aparece juntamente com a aba “Distance computation” (fig. 38). Após essa etapa, clicar em “compute e OK”.



**Figura 38.** Cálculo da distância entre nuvens densa 3D

Em seguida, é selecionada a nuvem de pontos já comparada e na aba “properties”, seleciona-se o “color scale” e marca-se a opção “visible”. É selecionado as distâncias absolutas no gráfico “SF display params” (fig. 39).



**Figura 39.** Distância calculada entre nuvens densa 3D

O processo foi feito com todas as nuvens de pontos 3D, e os resultados são apresentados no capítulo seguinte.

### 3. Resultados e Discussão

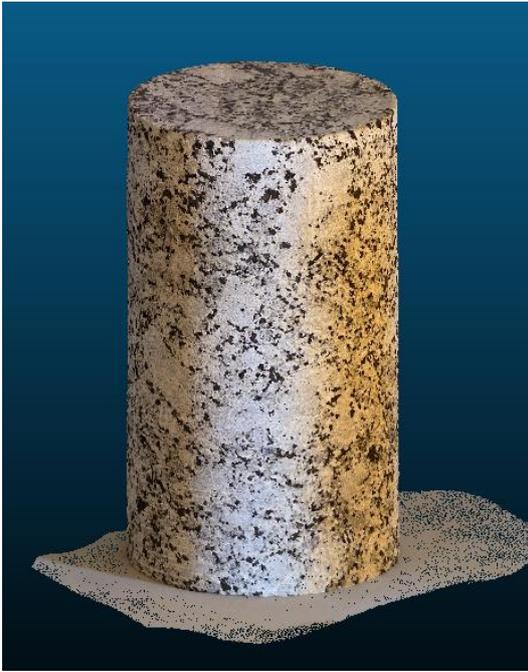
#### 3.1. Provetes de Granito

A tabela 6 apresenta os resultados do tempo de processamento, o número de pontos da nuvem densa obtidos da digitalização do provete por fotogrametria. As fotografias não receberam nenhum tipo de tratamento posteriormente. Nestas tabelas estão incluídos todos os modelos 3D das nuvens de pontos obtidos com os diferentes softwares SfM em análise. A visualização dos resultados nesta seção foi feita pelo CloudCompare. Os softwares VisualSfM, Colmap, Regard3D e Meshroom exportam a nuvem de pontos densa no formato ply. O ReCap web não suportava apenas a exportação de nuvem de pontos, portanto, o arquivo (.rcp) foi baixado e convertido para (.pts), um formato de arquivo de nuvem de pontos pelo ReCap desktop.

**Tabela 6.** Resultados obtidos na modelação do provete de granito

Dense Cloud	Nº de fotografias	Parâmetros definidos	Tempo de processamento	Nº de pontos	Tamanho do ficheiro
VisualSfM	50	Nenhum	22 minutos	1 238 871	35.5 MB
Colmap	50	Qualidade: Média	28 minutos	891 228	23.5 MB
Meshroom	50	Média/Média	3 horas e 20 minutos	486 039	19.3 MB
Regard3D	50	Sensibilidade de ponto-chave e proporção correspondente: Alta	1 hora e 7 minutos	1 293 059	82.5 MB
Autodesk Recap	50	Nenhum	Desconhecido	1 542 548	100 MB

Os softwares VisualSfM, Regard3D, Colmap e Autodesk Recap produzem resultados bastantes satisfatórios. VisualSfM produziu uma quantidade considerável de dados de nuvem de pontos (fig. 40c) e levou consideravelmente menos tempo de processamento. Com exceção do Meshroom, que além de produzir pouca quantidade de número de pontos da nuvem densa, levou mais tempo de processamento e ainda apresenta um modelo muito deformado (fig. 40e). Os outros três softwares (Colmap, Regard3D e Autodesk Recap) produziram uma digitalização bem-sucedida.



a) Autodesk Recap



b) Regard3D



c) VisualSfM



d) Colmap



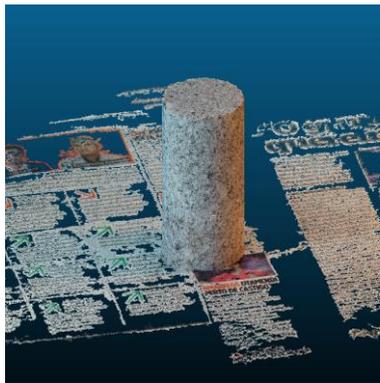
e) Meshroom

**Figura 40.** Nuvem de pontos densa (Provete)

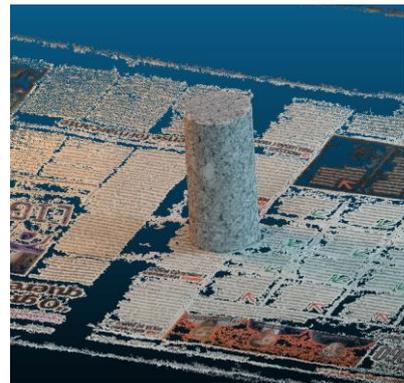
Dos conjuntos das experiências realizadas com o provete em cima da base de folha de jornal, o conjunto P3 produziu a nuvem de pontos com melhor qualidade entre os testes realizados, resultando num modelo 3D com melhor foco e menor caracterização da região envolvente como se pode constar na figura 41.



a) Conjunto P2



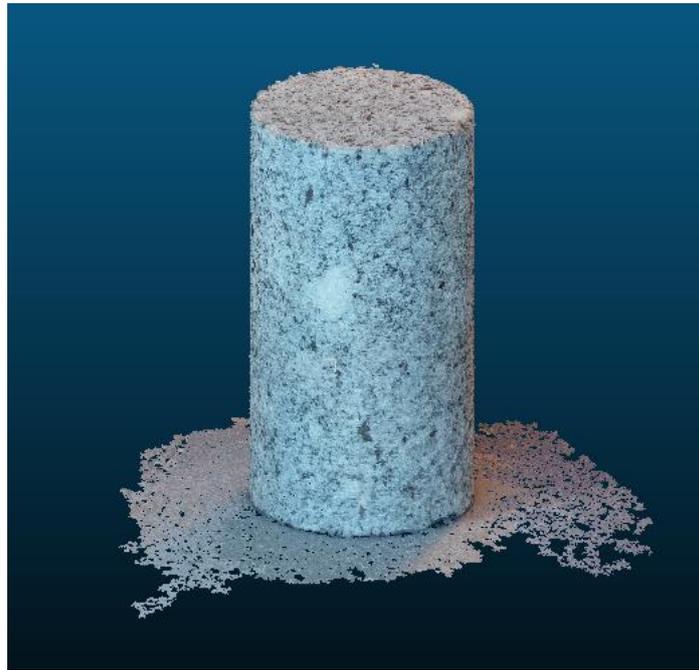
b) Conjunto P3



c) Conjunto P4

**Figura 41.** Nuvem de pontos densa (base de folha de jornal)

A experiência realizada com o conjunto P5 (base lisa) em comparação com o conjunto P3 (base de folha de jornal) conclui-se que não existe vantagem na utilização da base com folha de jornal. A mesa de base lisa branca conduziu ao modelo 3D de nuvem de pontos bastante satisfatório (melhor densidade de pontos), com menor caracterização da região envolvente e com menor tempo de processamento em relação a base de jornal que apresentava maior caracterização da região envolvente, o que não interessa a experiência, tornando assim a digitalização pouco satisfatório. O modelo 3D, com a base de folha de jornal obteve uma nuvem com maior número de pontos (muitos desses pontos pertencentes a região envolvente) o que tornou a digitalização pouco satisfatório e com maior tempo de processamento.



**Figura 42.** Nuvem de pontos densa (base lisa)

Do uso de máscaras pode-se declarar que o Colmap foi o único software permite a utilização de máscaras para ajudar na reconstrução, deste modo destacando apenas o provete em si para a reconstrução e não a região envolvente. Foi realizado diversas experiências no software explorando as funcionalidades de uso de máscaras de modo a obter melhores resultados no modelo final. O software permite utilização de máscaras no formato png. A experiência revelou-se ser pouco vantajoso e pouco contribuiu para a melhoria da qualidade do modelo final. O resultado desta experiência pode ser encontrado no anexo. Os softwares VisualSfM, Regard3D e Meshroom não permitem o uso de máscaras. O Recap não oferece controle do pipeline usado na reconstrução e sendo todo o processo realizado na nuvem, não foi possível averiguar o uso de máscaras.

### 3.2. Fachada da sede da Associação de Estudantes da Universidade do Minho

Quanto a georreferenciação dos modelos 3D, as experiências realizadas permitiram-me averiguar que o Colmap não possui a funcionalidade de georreferenciação por GCP colocados sobre o modelo ou cena, mas pode ser feita a partir da linha de comandos utilizando as coordenadas dos centros das máquinas (GPS). Pode-se consultar o “pipeline” desenvolvido nos anexos.



**Figura 43.** Nuvem de pontos densa (Fachada da sede da Associação de Estudantes da Universidade do Minho - Colmap)

O VisualSFM permite a georreferenciação por GCP colocados sobre a superfície do modelo ou cena, após a obtenção da nuvem dispersa. Ainda apresenta a funcionalidade de georreferenciação baseadas no uso de coordenadas GPS após a obtenção da nuvem dispersa, mas não é possível a sua utilização na criação da nuvem densa. A funcionalidade está indicada como versão beta (ainda em testes).

O Regard3D não permite o uso dos GCP. A partir da última versão do software (lançada em 14 de março de 2019) podem ser usadas as informações GPS das fotografias na fase da triangulação, mas em seguida não permite a geração da nuvem de pontos densa.

O Meshroom e o Recap permitem a georreferenciação tanto por GCP colocados na superfície do objeto ou cena, como também a partir de coordenadas GPS.

### 3.3. Maciço Rochoso

No levantamento aerofotográfico foram adquiridas 500 fotografias com recurso a um VANT descrito no subcapítulo 3.3.2.

Para a criação das nuvens de pontos utilizaram-se apenas 250 fotografias nos softwares VisualSfM, Colmap, Meshroom e Regard3D devido a limitação da capacidade computacional disponível para a obtenção dos modelos. No Autodesk Recap utilizaram-se 100 fotografias devido a limitação do software pelo fato de ser uma licença académica, seguindo o fluxo de trabalho apresentado no subcapítulo. No Metashape utilizaram-se as 500 fotografias obtidas no levantamento aerofotográfico tendo em conta que sistema computacional utilizado apresentava excelentes requisitos para a criação da nuvem de pontos. Tendo obtido as nuvens de pontos densa do maciço para todos os softwares passa-se a fase de tratamento e georreferenciação das mesmas no CloudCompare de forma a obter apenas o talude em estudo como verificado no subcapítulo 3.2.5. Os softwares VisualSfM, Colmap, Regard3D e Meshroom exportam a nuvem de pontos densa no formato ply. O ReCap web não suportava apenas a exportação de nuvem de pontos, portanto, o arquivo (.rcp) foi baixado e convertido para (.pts), um formato de arquivo de nuvem de pontos pelo ReCap desktop. O Metashape exporta a nuvem no formato las.

Na tabela 7 encontram os dados das nuvens de pontos 3D antes e após o tratamento realizado no CloudCompare. Após o tratamento, as nuvens de pontos 3D foram exportadas no formato bin.

**Tabela 7.** Comparação das nuvens de pontos densa antes e após o tratamento

Software SfM	Nº de pontos antes do tratamento	Nº de pontos após o tratamento
VisualSfM	5 328 534	5 073 695
Colmap	2 525 285	2 421 370
Meshroom	1 256 760	1 058 515
Regard3D	4 801 058	4 536 903
Autodesk Recap	6 727 074	6 588 889
Metashape	39 358 364	14 355 786

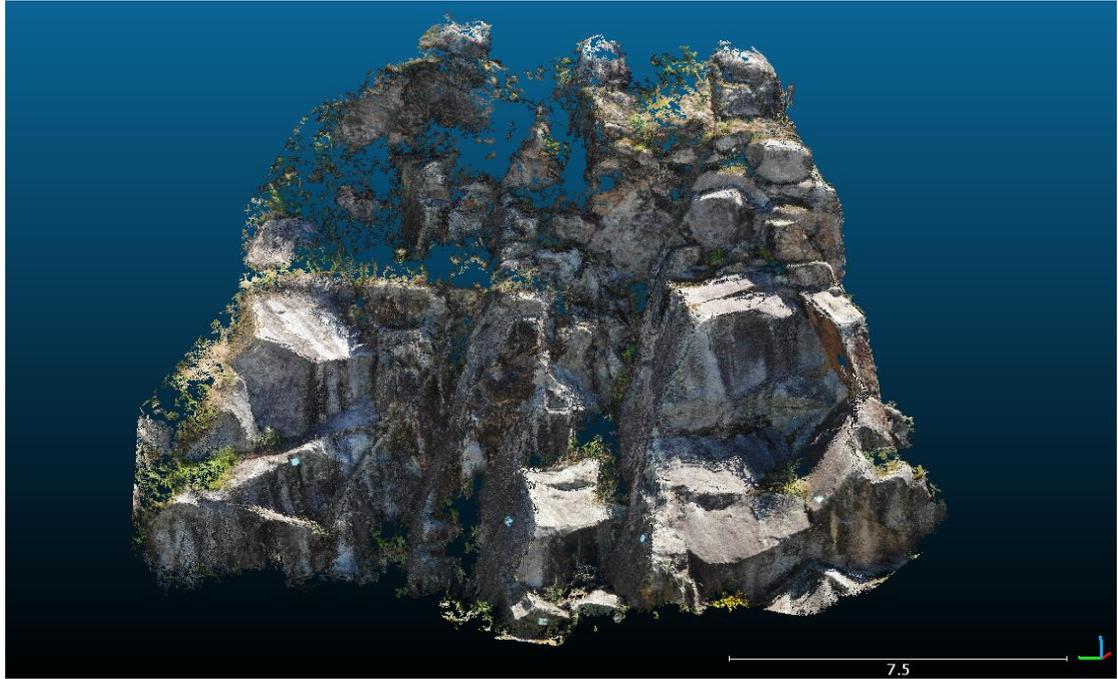
A tabela 8, apresenta o número de fotografias, as definições relevantes, o tempo de processamento, o número de pontos da nuvem densa obtida e o tamanho do ficheiro para cada software estudado. Todas as nuvens de pontos 3D nesta fase estão no formato bin.

**Tabela 8.** Resultados obtidos na modelação do talude

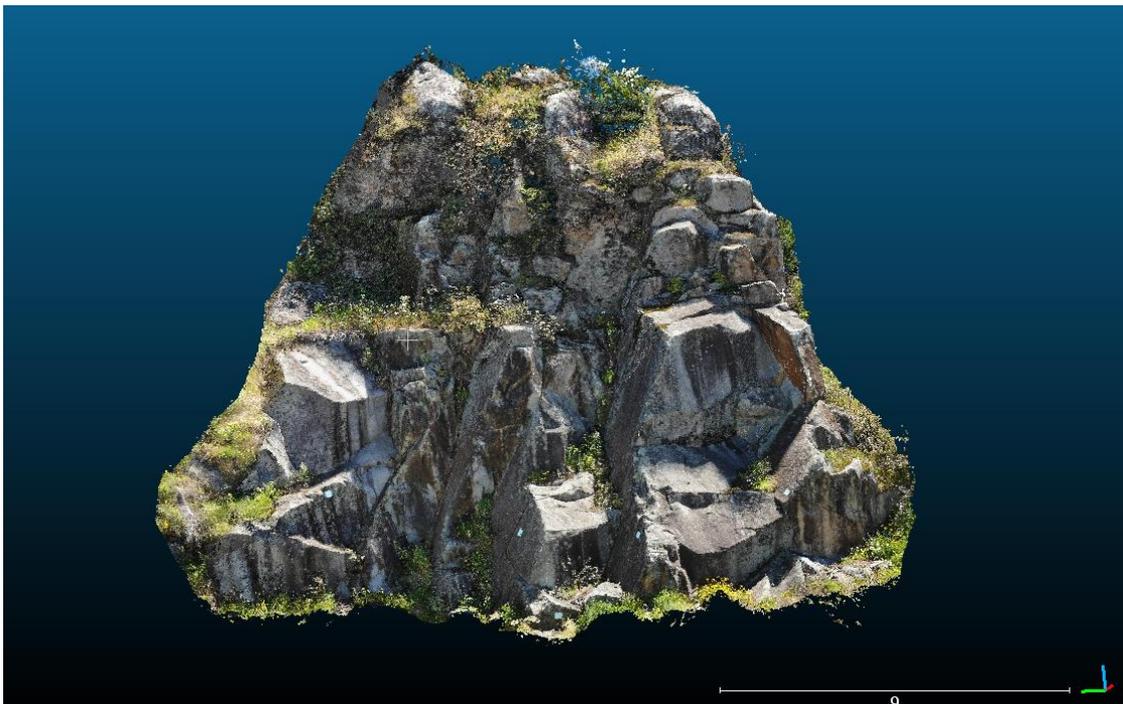
Dense Cloud	Nº de fotografias	Parâmetros definidos	Tempo de processamento	Nº de pontos	Tamanho do ficheiro
VisualSfM	250	Nenhum	6 horas	5 073 695	99.1 MB
Colmap	250	Qualidade: Média	3 horas e 10 minutos	2 421 370	47.3 MB
Meshroom	250	Qualidade: Média	14 horas e 28 minutos	1 058 515	16.5 MB
Regard3D	250	Sensibilidade de ponto-chave e proporção correspondente: Normal	14 horas e 4 minutos	4 536 903	88.61 MB
Autodesk Recap	100	Nenhum	Desconhecido	6 588 889	128.7 MB
Metashape	500	Qualidade: Alta	5 horas	14 355 786	560.7 MB

Baseando na tabela 8 pode-se verificar que o Metashape foi o software que produziu a nuvem densa com mais pontos, seguido pelo Autodesk Recap, tratando-se de dois softwares comerciais, possivelmente devido a uma melhor otimização dos processos SFM nas versões comerciais, pouco aprofundadas nas versões gratuitas. Entre os softwares gratuitos, o VisualSfM foi o que produziu a nuvem densa 3D com mais pontos. Semelhante ao experimento anterior, o Meshroom foi o software que produziu a nuvem densa com menos número de pontos entre todas, levou mais tempo de processamento e gerou um modelo mais ruidoso. O Colmap e Regard3D produziram resultantes consistentes, tendo o Regard3D levado quatro vezes mais tempo de computação em relação ao Colmap.

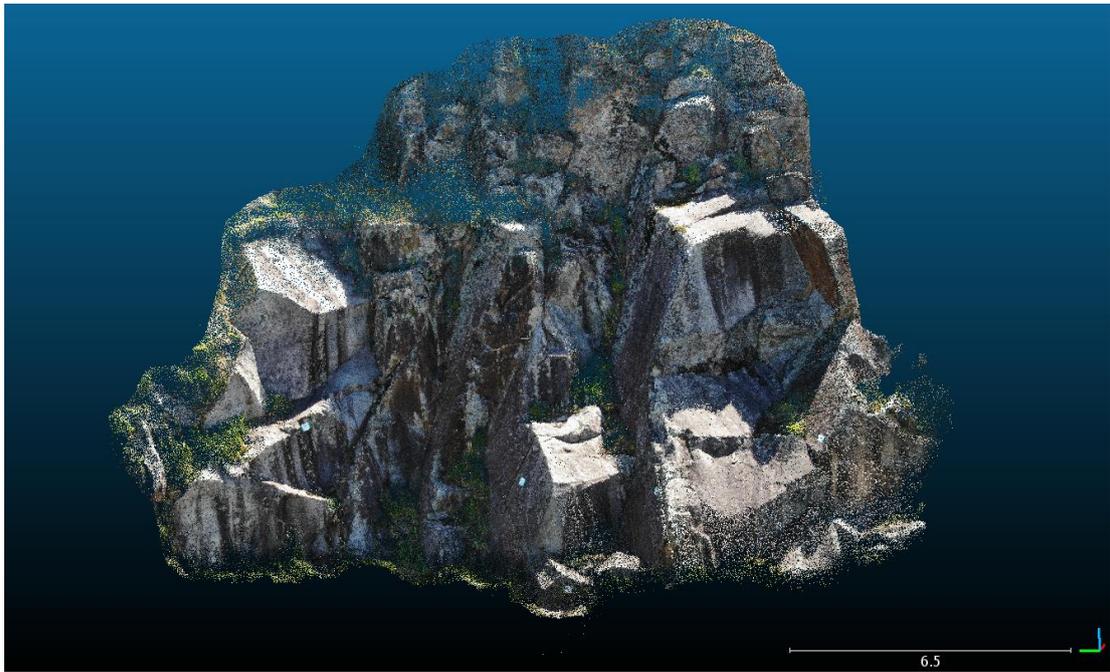
Pode-se ver nas seguintes figuras que as nuvens de pontos são densas o suficiente para apresentar claramente o talude e suas características. Entretanto, os furos são visíveis na nuvem densa do VisualSfM, Colmap, Meshroom, Regard3D e Autodesk Recap. A coloração e as bordas dos elementos na nuvem densa do Metashape são mais nítidas.



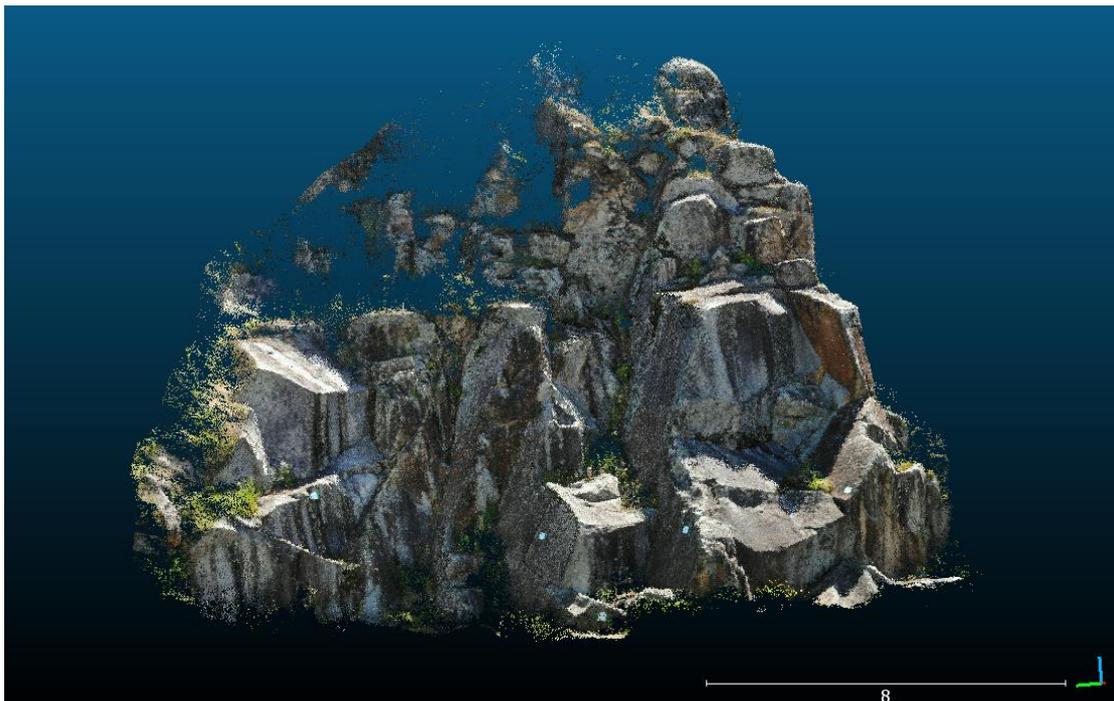
a) VisualSfM



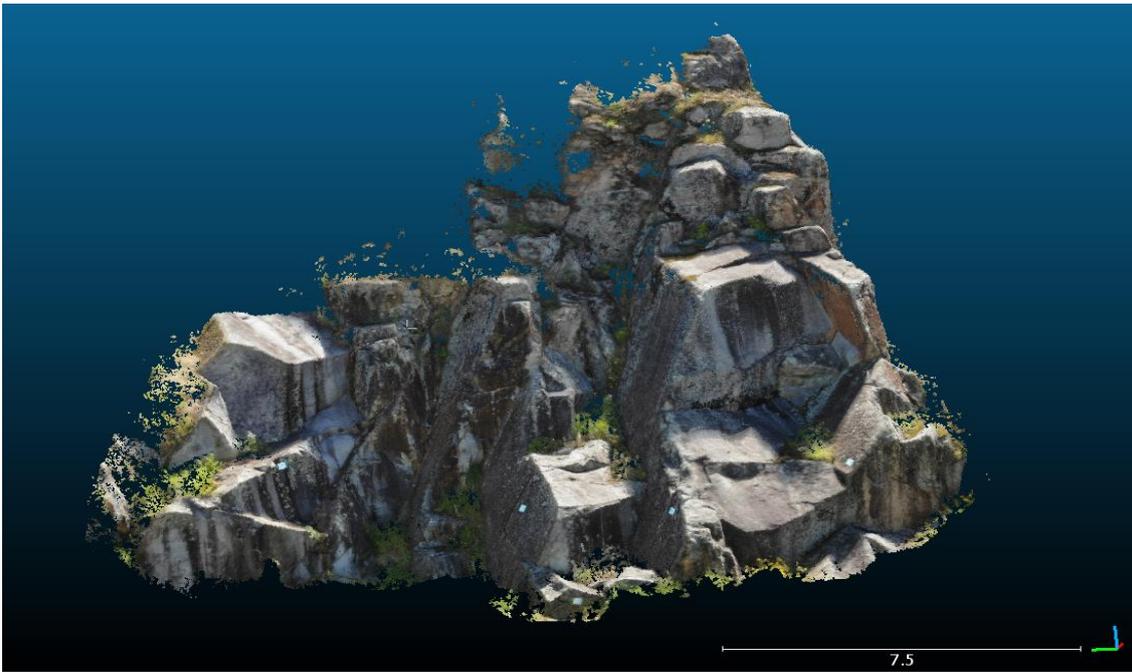
b) Metashape



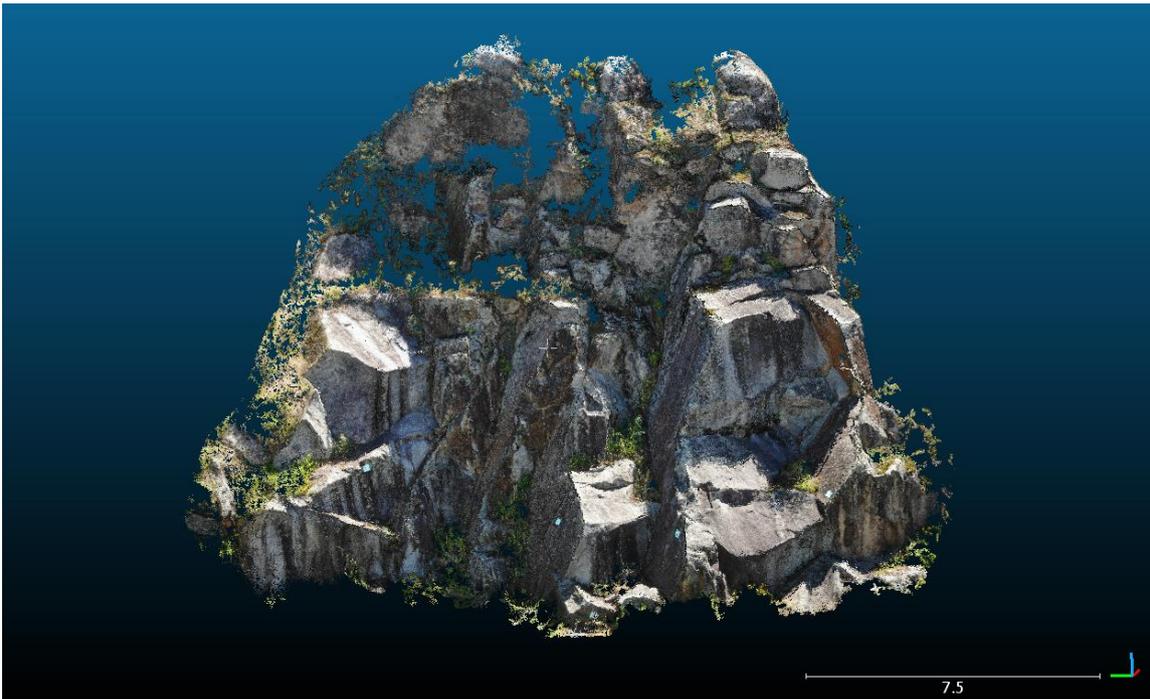
c) Meshroom



d) Colmap



e) Autodesk Recap



f) Regard3D

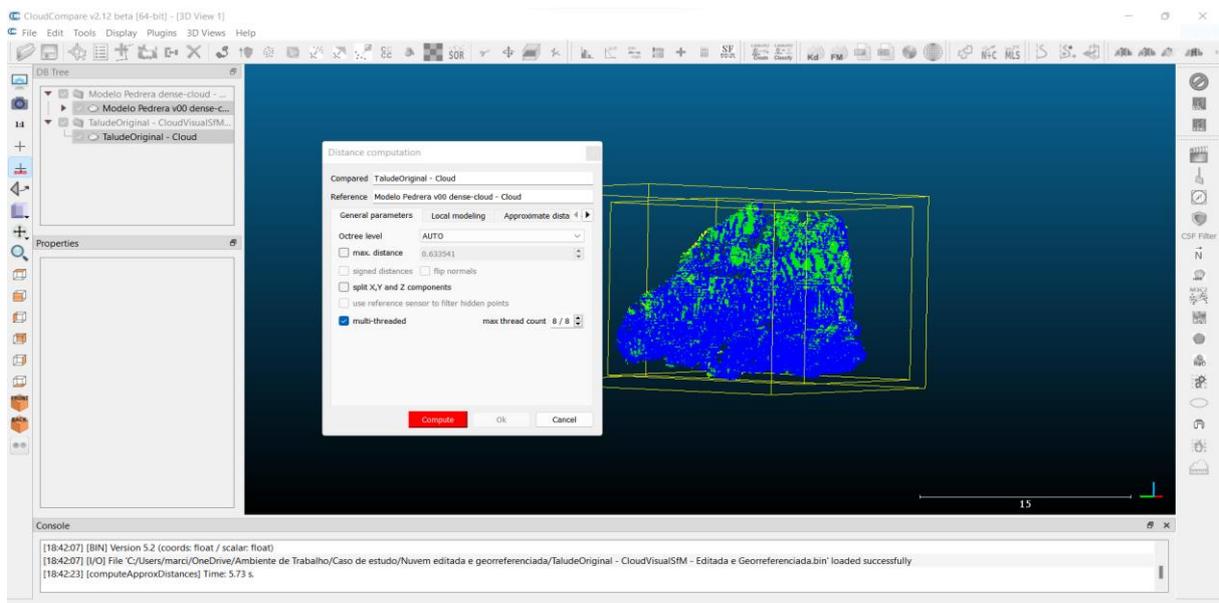
**Figura 44.** Nuvem de pontos densa (Maciço Rochoso)

### 3.3.1. Análise Comparativa (Metashape)

A comparação foi feita entre a nuvem de pontos 3D produzida pelos softwares VisualSfM, Colmap, Regard3D, Meshroom e Autodesk Recap e a nuvem de pontos 3D de referência produzida a partir do software comercial Metashape.

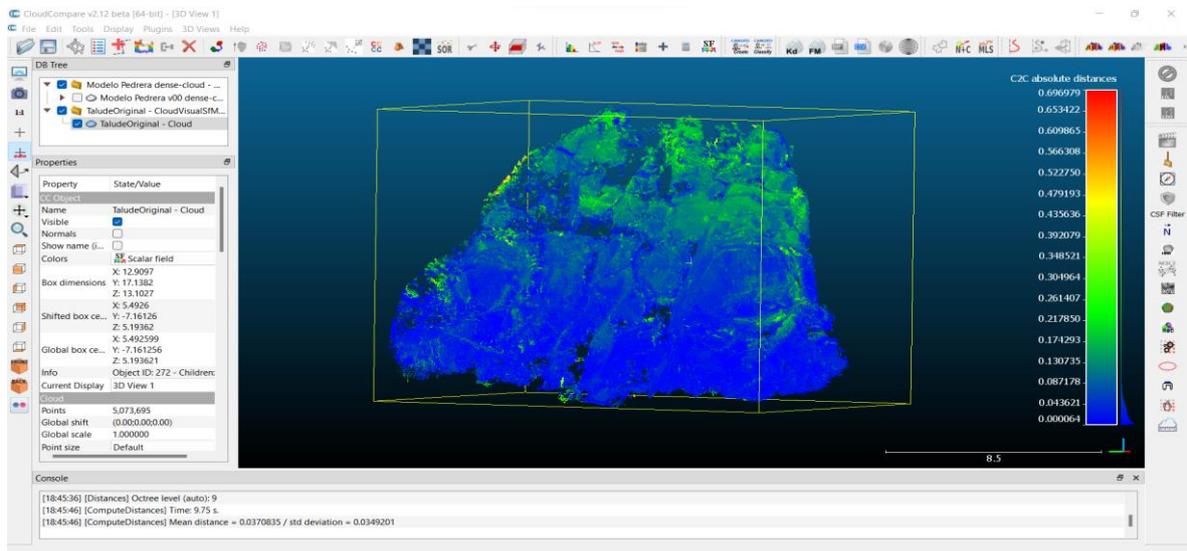
A ideia apresentada por Schöning e Heidemann (2015) foi utilizada para o estudo de comparação. Uma vez que é necessária uma verdade de base para realizar a comparação entre as nuvens de pontos 3D obtidos com softwares gratuitos, e como uma alternativa aos dados LiDAR e tecnologia relacionada, a nuvem de pontos gerada a partir do Metashape foi usada como verdade de base. Isso também significa que os erros na reconstrução do Metashape não foram levados em consideração. Essa precisão deve ser suficiente para propósitos gerais de modelagem e visualização. A reconstrução Metashape será referida como objetos de “verdade do terreno”, enquanto os modelos dos outros softwares gratuitos serão referidos como objetos “reconstruídos”. As comparações foram feitas com o software livre de código aberto CloudCompare.

Para essa tarefa é utilizado a ferramenta “Compute Cloud-Cloud Distance” no menu “Tools - Distances”. Após selecionado a nuvem de pontos 3D de referência (Metashape) e a nuvem pretendida para fazer a comparação, o resultado aparece juntamente com a aba “Distance computation” (fig. 45). Após essa etapa, clicar em “compute e OK”.



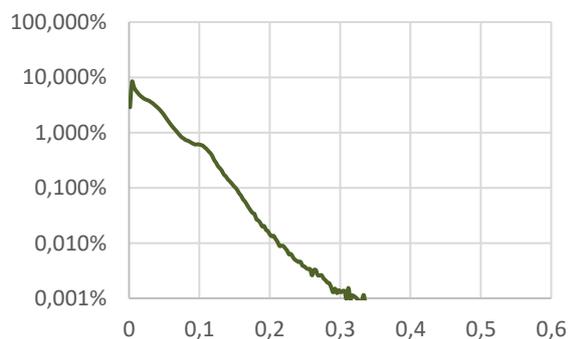
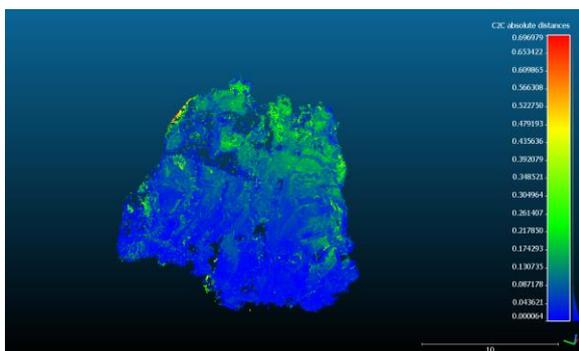
**Figura 45.** Cálculo da distância entre nuvens densa 3D

Em seguida, é selecionada a nuvem de pontos já comparada e na aba “properties”, seleciona-se o “color scale” e marca-se a opção “visible”. É selecionado as distâncias absolutas no gráfico “SF display params” (fig, 46).

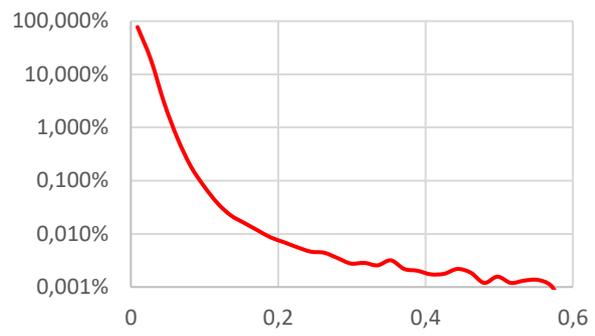
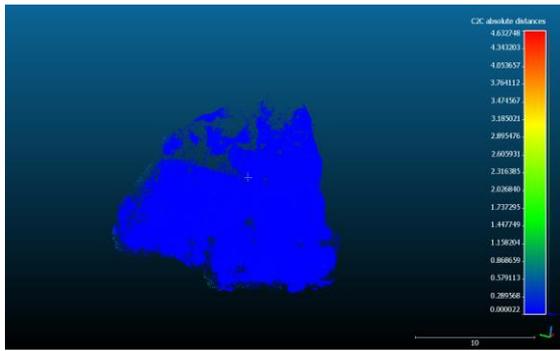


**Figura 46.** Distância calculada entre nuvens densa 3D

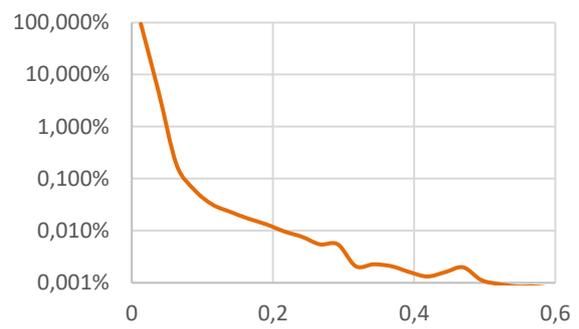
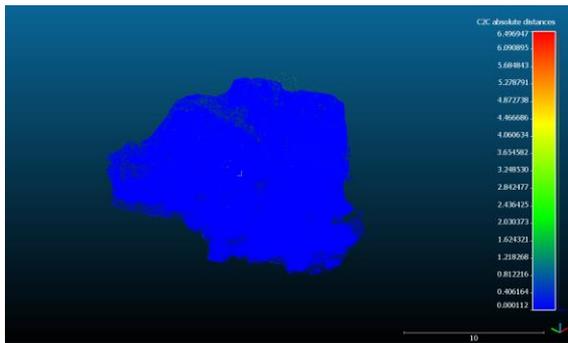
É apresentada a seguir, a tendência na distribuição de distâncias e ilustrada através dos histogramas. Colmap, Meshroom e Regard3D registaram pontos com diferença de distância absoluta maiores, aproximadamente de 0,6 m. Apesar de todas as diferenças, os resultados gerados pelos softwares VisualSfM, Colmap, Meshroom, Regard3D e Autodesk Recap são bastante comparáveis ao do Metashape.



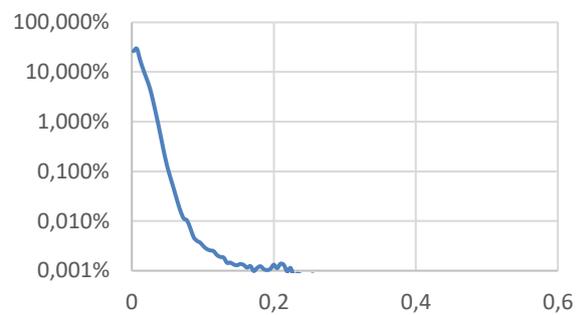
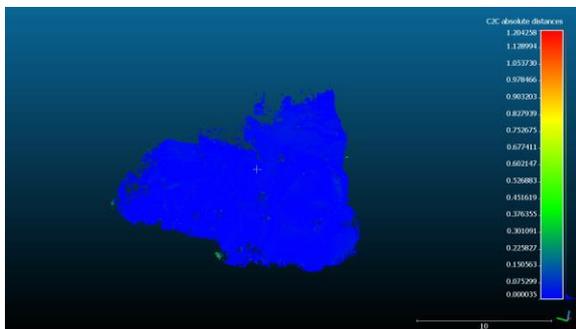
a) VisualSfM



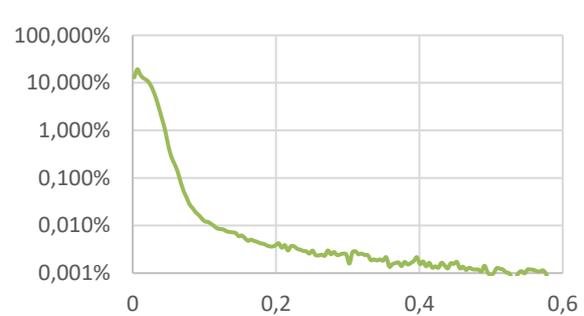
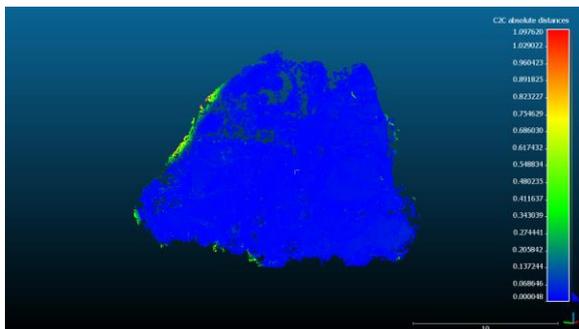
b) Colmap



c) Meshroom

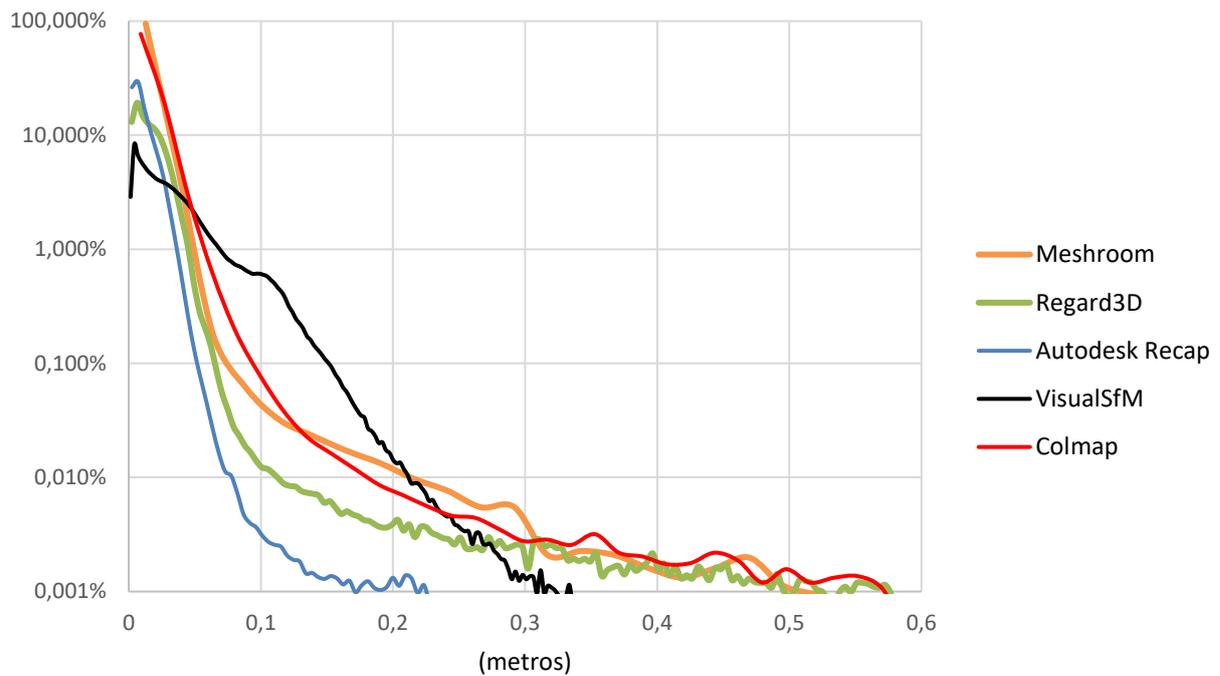


d) Autodesk Recap



e) Regard3D

**Figura 47.** Mapa de calor de distâncias (esquerda) e histograma (direita)



**Figura 48.** Histograma Geral - Distância Absoluta

Com base neste estudo comparativo, baseando no histograma e mapa de calor de distâncias, Autodesk Recap, produziu o melhor resultado, ou seja, poucos desvios da verdade do terreno (o modelo Metashape). Em relação aos softwares gratuitos, o Regard3D produziu uma nuvem densa 3D com melhor qualidade embora foi necessário muito tempo de processamento. Por outro lado, foi considerado mais equilibrado em documentação, GUI e opções de ajustes. Além disso, é totalmente gratuito e pode produzir superfície e textura. Sendo um software livre, o COLMAP oferece uma interface gráfica agradável e muitas opções de ajustes. É utilizada a reconstrução automática para a reconstrução do modelo 3D e produziu resultados impressionantes. No entanto, o Meshroom produziu as nuvens com mais ruídos. De referir que os parâmetros da reconstrução da nuvem densa foram configurados atendendo a capacidade computacional disponível tornando possível a conclusão da reconstrução dos modelos. Nenhum dos softwares gratuitos oferece qualquer edição de nuvem de pontos ou malha, portanto, o usuário precisa usar um aplicativo externo como o CloudCompare para limpeza e edição 3D adicionais. Contudo, tal não consitui qualquer inconveniente na utilização dos softwares gratuitos.

### **3.4. Experiências do usuário**

Existem algumas diferenças entre os softwares em relação à quantidade de controle ao longo de certas etapas do processo, e mais obviamente nas interfaces do usuário. O VisualSfM e o Regard3D exigem várias etapas antes de criar o modelo final, considerando que apenas o Meshroom tem uma solução de etapa única para todo o processo. O Colmap pode utilizar a reconstrução automática onde o usuário obtém o produto final, ou se preferível uma reconstrução manual que apresenta várias etapas. Todos os pacotes de software usados podem criar 3D modelos em diferentes níveis de qualidade, e no presente estudo foram explorados diferentes níveis de qualidade de acordo com a capacidade computacional disponível.

Em geral, o VisualSfM é uma excelente ferramenta para pegar algo estático, converter e depois importar nuvens de pontos para um ambiente 3D, mas requer processamento adicional e habilidade humana para criar um ambiente digital perfeito. Ferramentas adicionais (PMVS/CMVS) também são necessárias para executar o VisualSfM, elas precisam ser baixadas de seus respectivos "sites" e copiadas para as pastas locais. Isso pode ser complicado para usuários novos e não técnicos, no entanto, o restante do processo de instalação é principalmente automático.

O VisualSfM não oferece nenhuma opção de edição integrada ou limpeza de ruído, portanto, é necessário um software externo. A última atualização do software foi em 2012. É possível orientar e corrigir a escala no software. A distância focal é retirado no EXIF das fotos, caso não houver essa informação é calculado por 1.2 multiplicado pela máxima dimensão da foto. A máxima resolução das fotografias permitida é 4096 píxeis. Fazer a análise por vídeos ("sequence match") não é muito eficiente, pelo que recomendo "Compute Missing Match" (analisa as correspondências entre todas as fotografias). Relativamente ao uso de coordenadas GPS, a funcionalidade está indicada como versão beta (ainda em testes). Conseguir-se utilizar na nuvem dispersa, mas não é possível a sua utilização na criação da nuvem densa. O software permite orientação da nuvem baseando no GCP após gerar a nuvem dispersa.

O COLMAP oferece uma reconstrução simples em uma etapa, que pode ser conveniente como uma solução rápida para usuários iniciantes. O processo manual ou passo a passo, por outro lado, oferece melhor qualidade de reconstrução com muitas opções de ajuste. No entanto, a etapa "estéreo" do processo de reconstrução densa pode travar o programa se a memória da placa de vídeo for inadequada. Portanto, recomenda-se a redução do uso de memória durante um processo de reconstrução densa. Além disso, a etapa 'estéreo' leva muito tempo durante a computação. Apesar disso, o COLMAP oferece uma infinidade de variáveis para ajustar, podendo ser possível minimizar o tempo de computação e obter resultados razoáveis.

O Regard3D oferece um processo de instalação simples em uma etapa, e o arquivo do instalador pode ser baixado de seu site oficial. O Regard3D é totalmente gratuito para uso com um fluxo de trabalho simples que oferece um menu vertical para fluir várias etapas. Opções adicionais são oferecidas no lado direito da GUI. Embora não ofereça edição nuvem de pontos, oferece reconstrução e texturização de superfície. O software não permite a utilização de máscaras. A partir da última versão do software (lançada em 14 de março 2019) podem ser usadas as informações GPS das fotografias na fase da triangulação. Utilizando as informações de GPS para inicializar a triangulação, o processo não revelou ser eficiente na calibração das câmaras e em seguida não foi possível obter a nuvem densa. A funcionalidade não apresenta vantagens.

O Recap apresenta uma interface de usuário direta. Os comandos e funções eram limitados, portanto, causando menos confusão na seleção de ferramentas. É um aplicativo baseado em nuvem que diminuía a importância do forte poder de computação e por outro lado, requer uma conexão de internet estável. Os usuários podem procurar ajuda no fórum a qualquer momento. Alguns componentes do projeto que os usuários podiam controlar eram materiais de entrada (ou qualidade da fotografia) e pontos de controle. O software oferece a opção de edição e tratamento da nuvem.

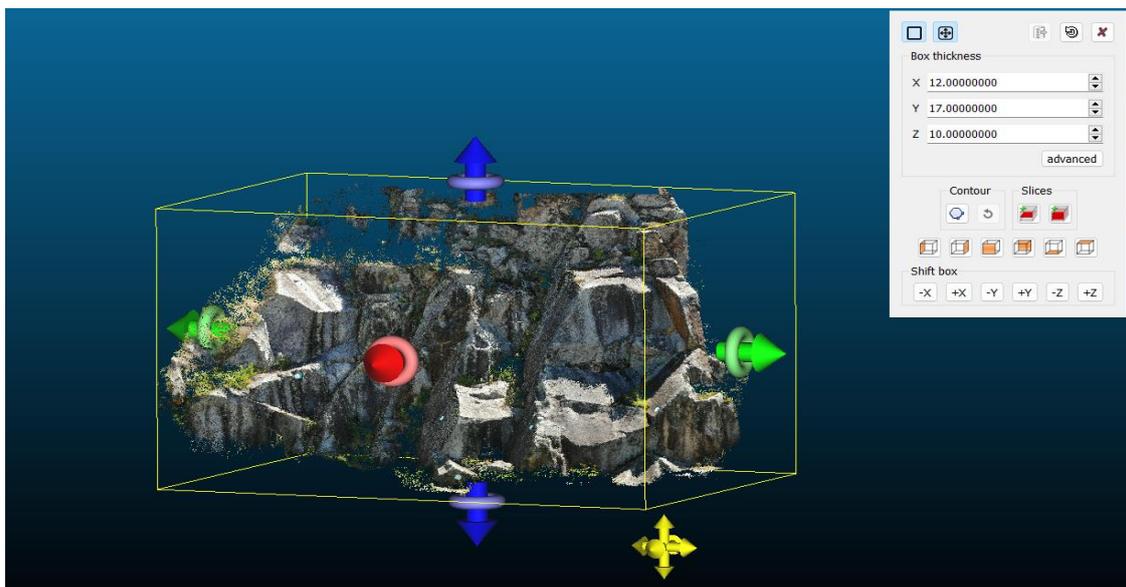
Metashape é um dos pacotes comerciais mais populares, portanto, decidiu-se a sua utilização para produzir os dados de verdade. Possui um “pipeline” robusto, mas direto, que pode produzir resultados precisos. O Metashape tem a capacidade de dividir o conjunto de fotos, remover manualmente nuvens de pontos extras e também possui alguns recursos extras interessantes, como a capacidade de fechar buracos automaticamente.

## 4. Aplicação Prática

No presente capítulo vão ser apresentados algumas aplicações práticas com o intuito de apresentar possíveis trabalhos futuros com as nuvens de pontos obtidas.

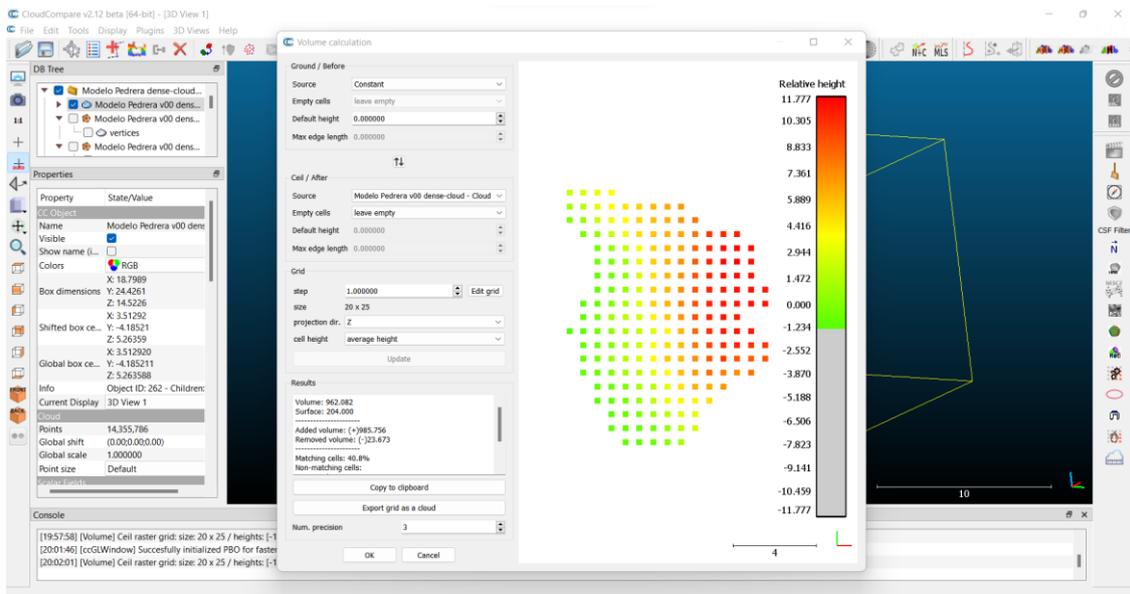
### 4.1. Volume

Primeiramente é criada uma caixa delimitadora (fig. 49) com as mesmas medidas para todos os modelos 3D de modo a proceder à correta determinação do volume para assim realizar a comparação entre as nuvens de pontos 3D.



**Figura 49.** Determinação da caixa delimitadora

O volume da nuvem de pontos 3D é calculada recorrendo a ferramenta “Compute 2.5D Volume” (fig. 50) no menu “Tools – Volume”.



**Figura 50.** Determinação do volume da nuvem densa 3D

Na tabela 9 apresenta-se o volume calculado das nuvens de pontos dos diferentes softwares.

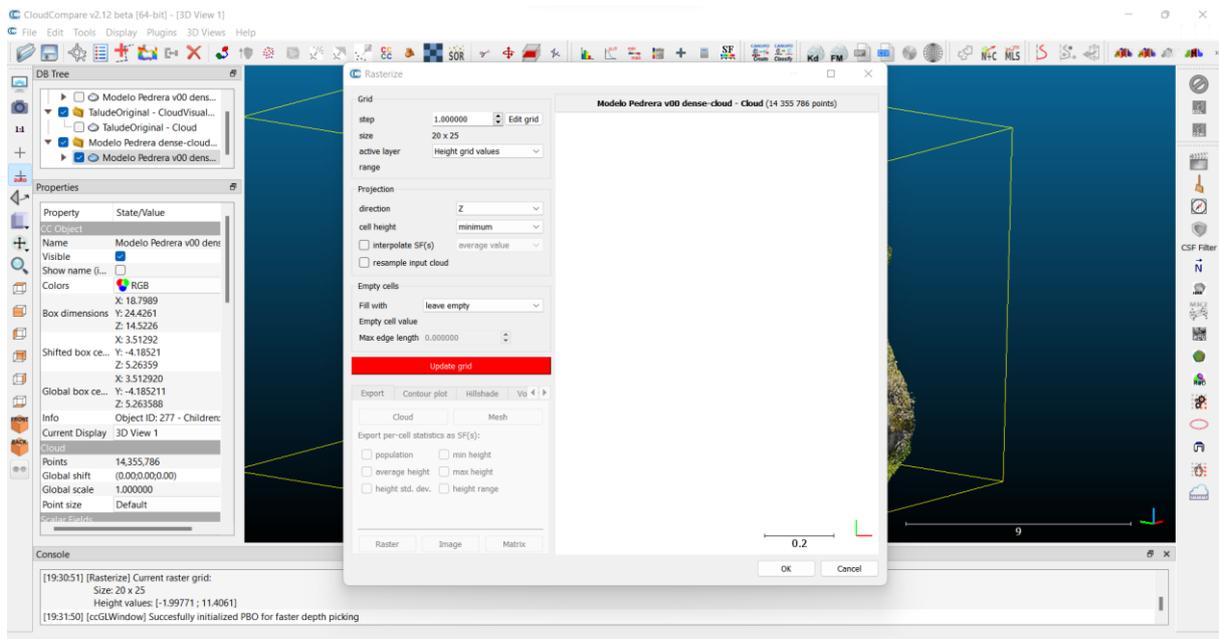
**Tabela 9.** Volume das nuvens de pontos

	Colmap	Meshroom	Autodesk Recap	VisualSfM	Regard3D
Volume (m <sup>3</sup> )	294.82	276.50	164.23	267.89	330.85

A nuvem de pontos do software Regard3D produziu um maior volume como se pode verificar na tabela acima. O Autodesk Recap por seu lado, foi o software que produziu uma nuvem com menor volume, sendo que os restantes softwares produziram volumes similares.

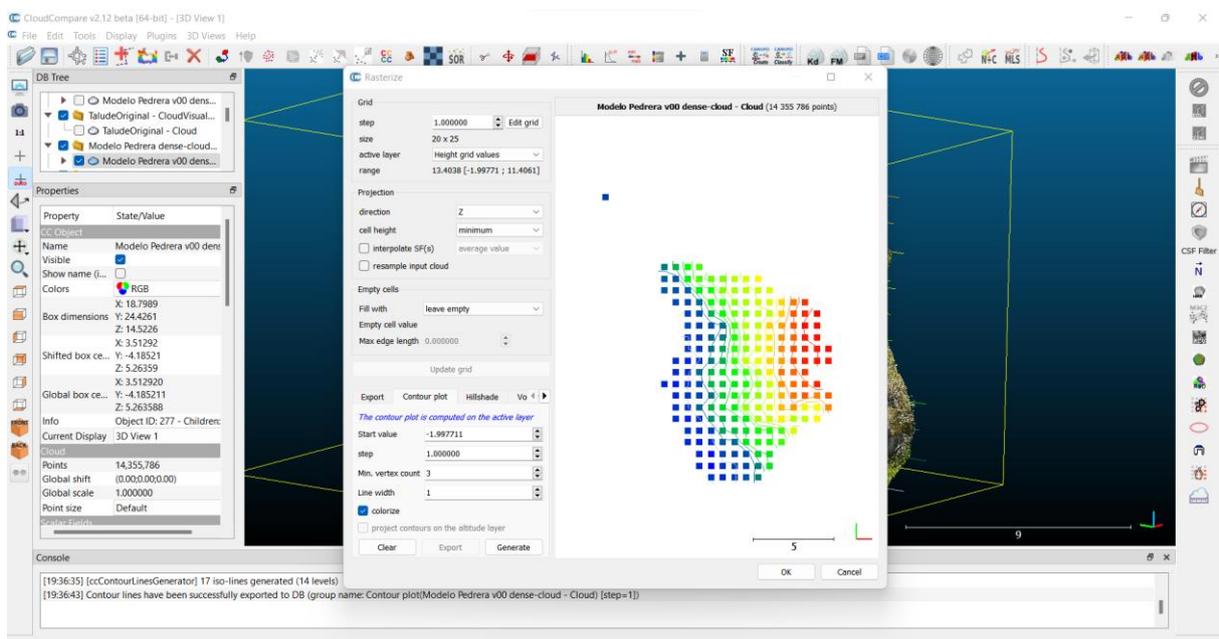
#### 4.2. Curva de níveis

A nuvem de pontos 3D é carregada no software CloudCompare e para a obtenção da curva de nível é utilizada a ferramenta na barra superior “Convert a cloud to 2D raster”. Na aba de Rasterize, é preenchido a resolução do modelo de acordo com a escala real em step (valor em metros) e em seguida “Update grid” (fig. 51).

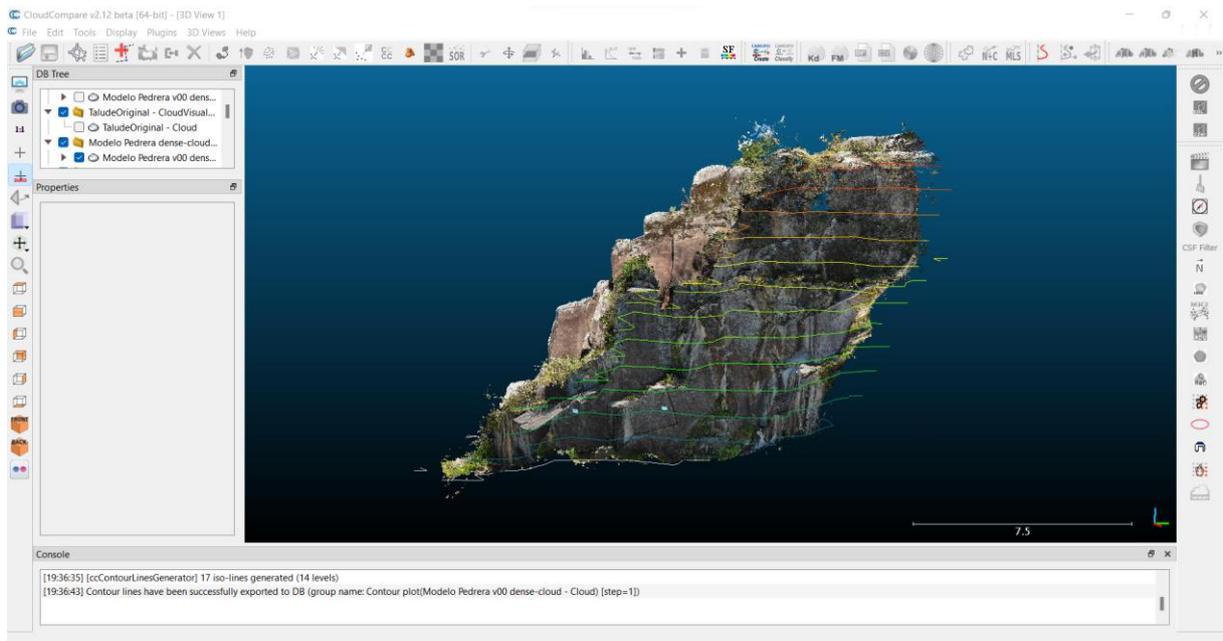


**Figura 51.** Determinação da curva de nível - passo 1

Após esta etapa e ainda na aba de “rasterize”, é definido em metros a distância de intervalo entre as curvas de nível em step, é selecionado a opção colorize, após “generate” seguido de “export” e “OK” (fig. 52). Deste modo é obtido a curva de nível da nuvem de pontos 3D (fig. 53).

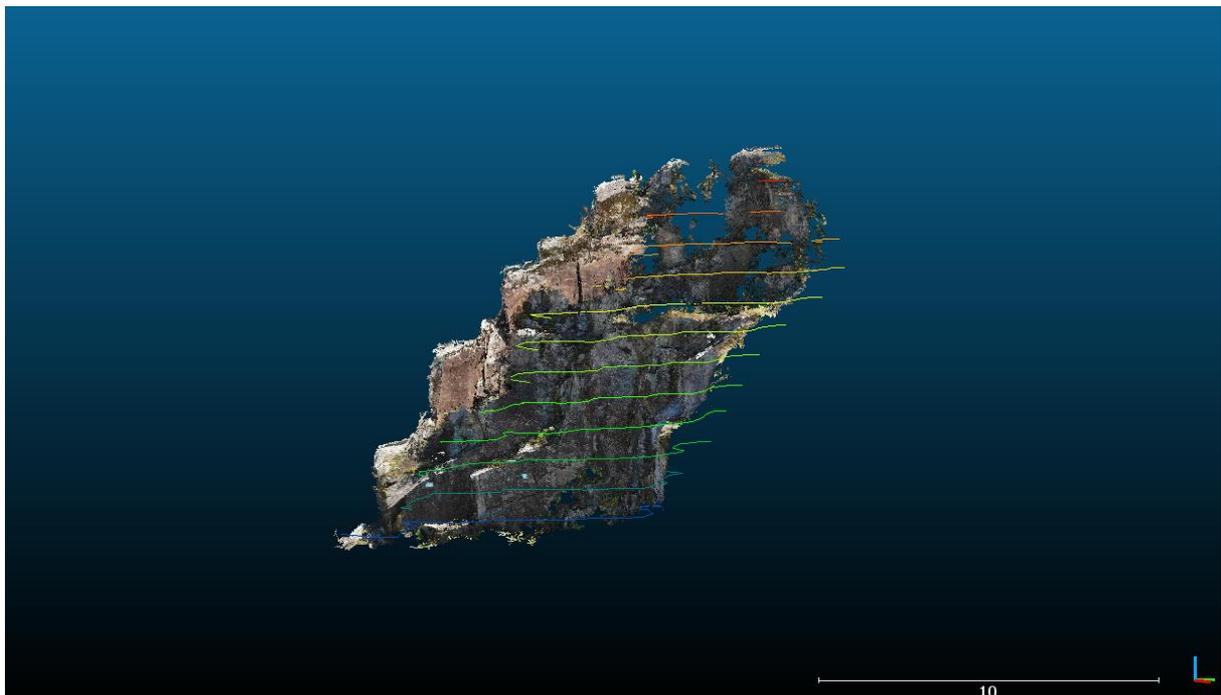


**Figura 52.** Determinação da curva de nível - passo 2

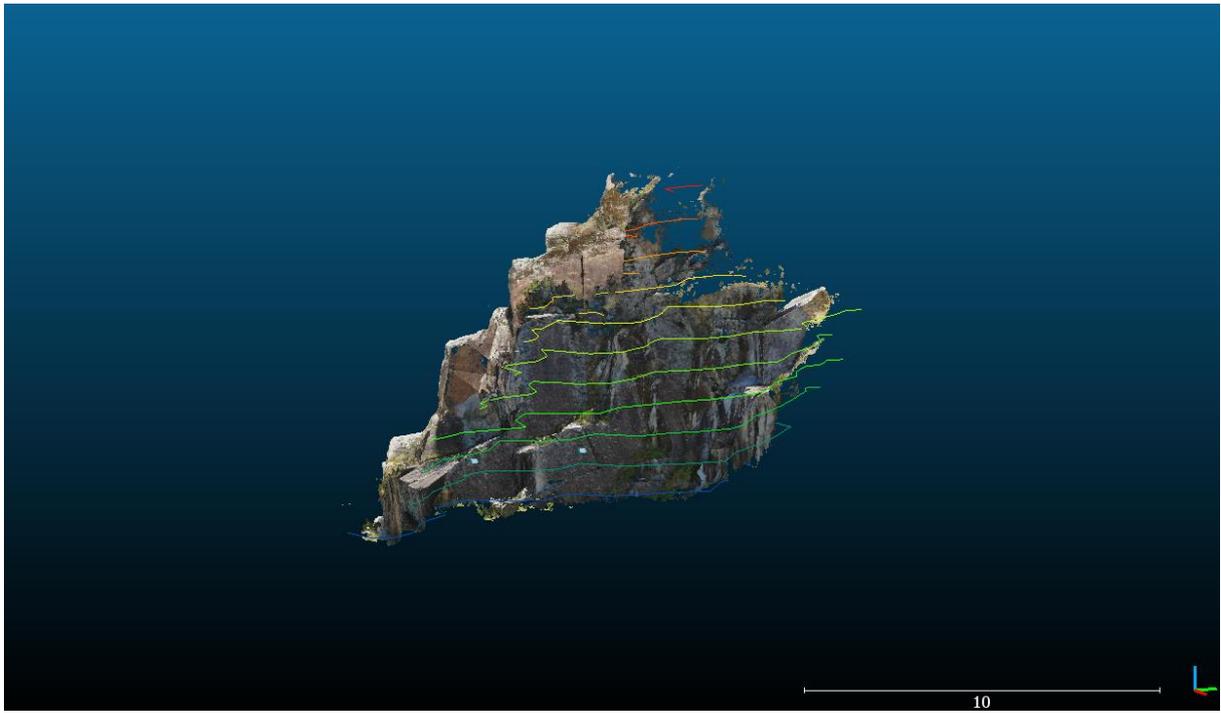


**Figura 53.** Determinação da curva de nível - passo 3

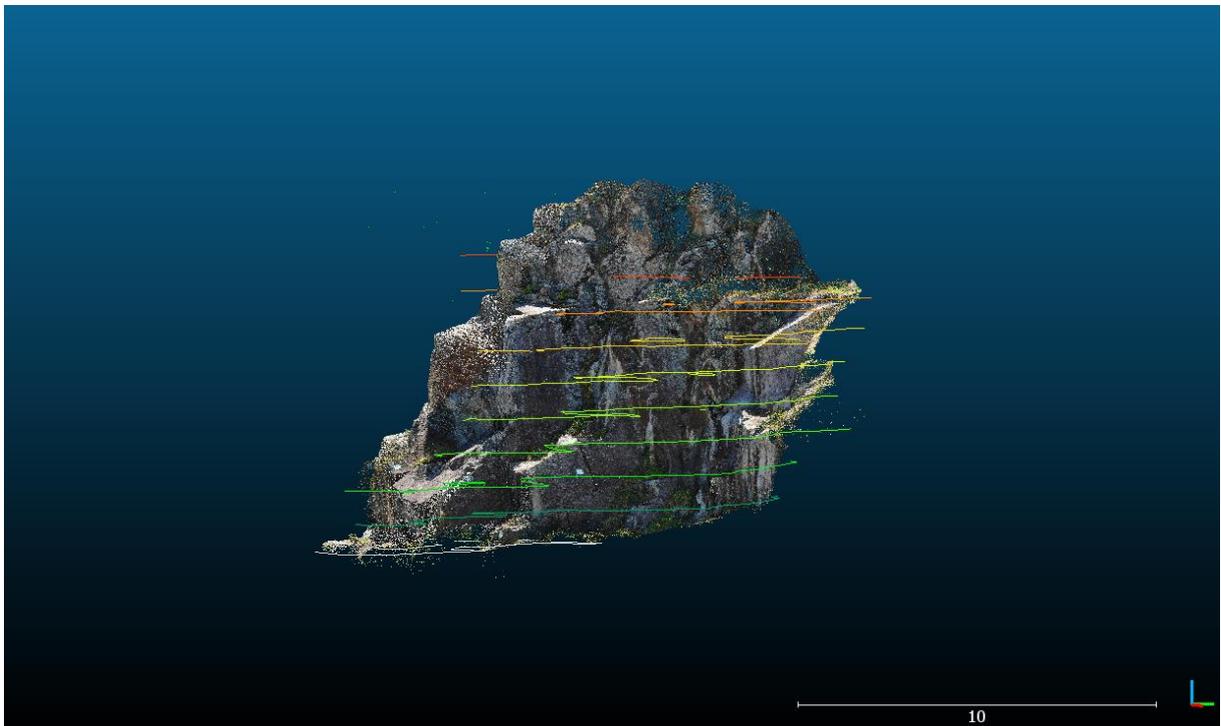
O processo foi feito com todas as nuvens de pontos 3D, e os resultados são apresentados em seguida.



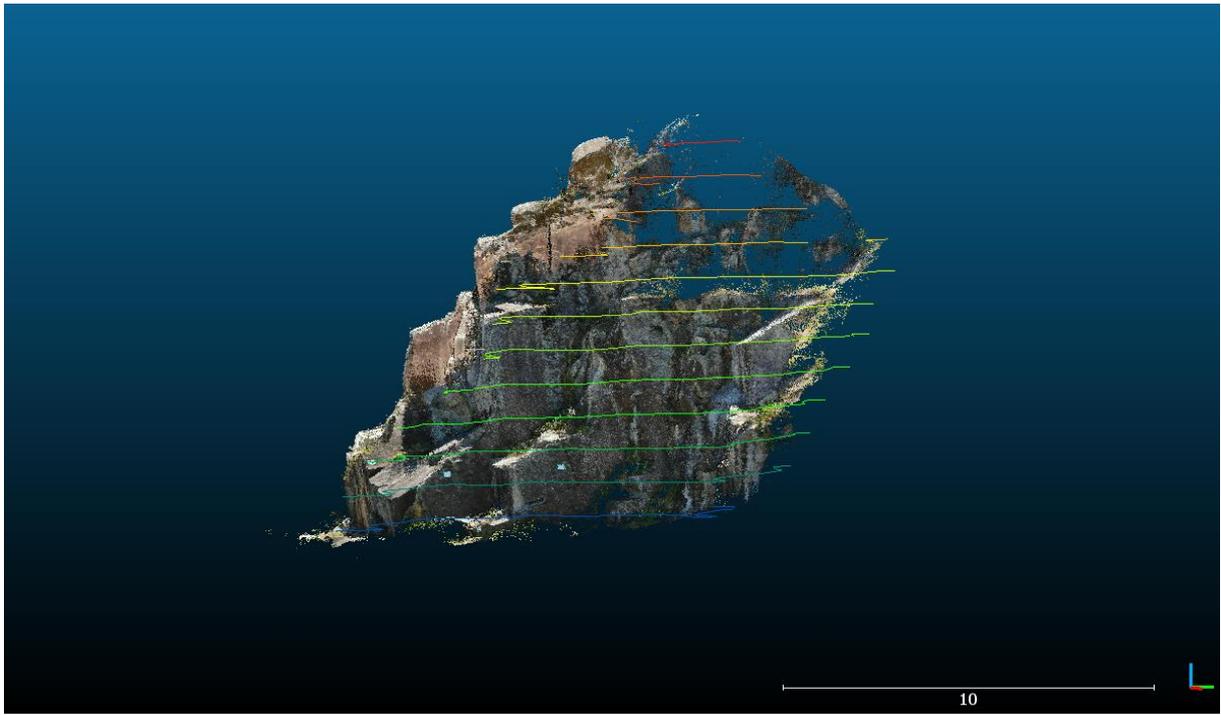
a) VisualSfM



b) Autodesk Recap



c) Meshroom



d) Colmap



e) Regard3D

**Figura 54.** Curva de nível

## 5. Conclusão

Apontar para um programa de modelagem baseado em imagem 3D específico como a melhor solução gratuita e de código aberto é uma tarefa difícil. Os quatro aplicativos gratuitos selecionados (VisualSfM, Regard3D, Colmap e Meshroom) oferecem várias opções ou manipulação de conjuntos de dados, opções de ajustes, tempo de computação, usabilidade, GUI e curva de aprendizado. Os softwares foram testados primeiramente para realizar um estudo piloto e em seguida utilizados para o nosso caso de estudo. Os fluxos de trabalho desses softwares foram estudados primeiro e, posteriormente, seus resultados de reconstrução foram avaliados em relação a objetos de verdade com base na medição de distância e no tempo de computação.

Deparamos, durante o estudo, que cada um desses programas possui fluxos de trabalho básicos e são muito semelhantes. O Meshroom revelou ser menos sofisticada para iniciantes tanto na GUI quanto nos resultados de saída. O Colmap e Regard3D oferecem uma GUI sofisticada e limpa, que pode atender a uma ampla gama de necessidades dos usuários. No entanto, tanto o COLMAP, Regard3D e Meshroom travaram várias vezes durante o estudo. Por outro lado, VisualSfM produziu resultados relativamente bons. O software GUI é simples e de fácil uso, podendo gerar o modelo 3D gerado ao fim de quatro cliques. Considerando a automação do processo, o tempo de processamento, a GUI, a densidade da nuvem de pontos e, não menos importante, a precisão, portanto, classificamos o software de estudo começando com o melhor como Regard3D, COLMAP, VisualSfM e Meshroom.

Esta dissertação se propôs a investigar soluções de reconstrução 3D baseando nos softwares gratuitos, levando em consideração os potenciais benefícios com base na nuvem de pontos densa gerada tornando possível realizar vários tipos de análise e medições que atualmente carecem de recursos tecnológicos de ponta e habilidades relacionadas.

O Regard3D, um software livre e de código aberto é, portanto, sugerido como a solução mais conveniente, pois é fácil de instalar, não requer conhecimento de programação para usar, pode produzir um bom resultado significativo e oferece uma curva de aprendizado suave. O site oficial também fornece documentação e tutoriais detalhados.

A fotogrametria provou-se vantajosa de várias formas em relação aos levantamentos manuais “in situ”. No caso do tempo de execução, como referido o levantamento fotográfico e dos pontos de controle demorou cerca de duas horas, sendo o restante trabalho de fotogrametria feito no escritório de forma relativamente rápida. Acreditamos que esta dissertação ajudará os usuários não especialistas a entender e selecionar o software mais adequado para a produção de modelos 3D baseadas em imagens a baixo custo.

Os trabalhos futuros propostos fixam na exploração das potencialidades do software Regard3D e a obtenção de uma nuvem de pontos com máxima qualidade. Torna-se de especial interesse a aplicação

das nuvens de pontos no âmbito da Engenharia Civil no controlo das fases de obra, monitoramento das estruturas, levantamentos, terraplanagens ou ainda análise de descontinuidades em maciços rochosos, permitindo a visualização de dados precisos economizando assim tempo e custos. São vários os benefícios da utilização da nuvem de pontos, mas é necessário investimento.

## 6. Bibliografia

Abel-Aziz, Y.I. & Karara, H.M. (1971) *Direct Linear Transformation from Comparator Coordinates Into Object Space Coordinates*, pp. 1–18. American Society of Photogrammetry, Urbana, IL.

Ahmadabadian, A.H., Robson, S., Boehm, J., Shortis, M., Wenzel, K. & Fritsch, D. (2013) A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. *ISPRS Journal of Photogrammetry and Remote Sensing*, 78: 157–167.

Baumberg, A. (2000) Reliable feature matching across widely separated views. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1, June 2000, pp. 774–781. Hilton Head Island, SC.

Bay, H., Ess, A., Tuytelaars, T. & van Gool, L. (2008) Speeded-up robust features (SURF). *Computer Vision Image Understand*, 110: 346–359, Similarity Matching in Computer Vision and Multimedia.

Bemis, S.P., Micklethwaite, S., Turner, D. et al. (2014) Ground-based and UAV-based photogrammetry: a multi-scale, high-resolution mapping tool for structural geology and paleoseismology. *Journal of Structural Geology*, 69: 163–178.

Bustos, B., Deussen, O., Hiller, S. & Keim, D.A. (2006) A graphics hardware accelerated algorithm for nearest neighbour search. *International Conference of Computational Science*, 4: 196–199.

Calonder, M., Lepetit, V., Strecha, C. & Fua, P. (2010) BRIEF: binary robust independent elementary features. In: K. Daniilidis, P. Maragos & N. Paragios (eds), *Computer Vision—ECCV 2010, Lecture Notes in Computer Science*, 6314, pp. 778–792. Springer, Berlin Heidelberg.

Carrivick, J., Smith, M., Quincey, D., (2016) *New Analytical Methods in Earth and Environmental Science – Structure from Motion in the Geosciences*

Colmap (2022) <https://colmap.github.io/>, acessado em 13 abril 2022

Duong, C., (2017) A Comparative Case Study between Agisoft Photoscan and Autodesk Recap Photo-to-3D Web Services

Fischler, M.A. & Bolles, R.C. (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24 (6), 381–395.

Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L. & Carbonneau, P.E. (2013) Topographic Structure from Motion: a new development in photogrammetric measurement. *Earth Surface Processes and Landforms*, 38: 421–430.

Furukawa, Y. & Ponce, J. (2009) Carved visual hulls for image-based modelling. *International Journal of Computer Vision*, 81: 53–67.

Furukawa, Y., Curless, B., Seitz, S.M. & Szeliski, R. (2010) Towards internet-scale Multi-View Stereo. In: *IEEE Conference, Computer Vision and Pattern Recognition (CVPR)*, June 13–18, 2010, pp. 1434-1441. San Francisco, CA.

Granshaw, S. (1980) Bundle adjustment methods in engineering photogrammetry. *The Photogrammetric Record*, 10: 181–207.

Granshaw, S. (2016), Photogrammetric Terminology: Third Edition. *Photogram Rec*, 31: 210-252.

Griwodz, C., Gasparini, S., Calvet, L., Gurdjos, P., Castan, F., Maujean, B., De Lillo, G., e Lanthony. Y., 2021. AliceVision Meshroom: An open-source 3D reconstruction pipeline. In Proceedings of the 12th ACM Multimedia Systems Conference (MMSys '21). Association for Computing Machinery, New York, NY, USA, 241–247.

Hartley, R. & Zisserman, A. (2003) *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge.

Hartley, R.I. & Sturm, P. (1997) Triangulation. *Computer Vision and Understanding*, 68: 146 157.

Hartley, R.I. (1997) In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19: 580–593.

Hafeez, J., Hamacher, A., Son, J., Pardeshi, S., Lee, H., (2016) Workflow Evaluation for Optimized Image-Based 3D Model Reconstruction. In Proceedings of the International Conference on Electronics, Electrical Engineering, Computer Science (EEECS): Innovation and Convergence, Qingdao, China, pp. 62–65.

James, M.R. & Robson, S. (2012) Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. *Journal of Geophysical Research, Earth Surface*, 117, F03017.

Jones, C., Church, E., (2020) Photogrammetry is for everyone: Structure-from-motion software user experiences in archaeology, *Journal of Archaeological Science: Reports*, Volume 30, 102261, ISSN 2352-409X

Lhuillier, M. & Quan, L. (2005) A quasi-dense approach to surface reconstruction from uncalibrated images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27: 418–433.

Li, J., Li, E., Chen, Y., Xu, L. & Zhang, Y. (2010) Bundled depth-map merging for Multi-View Stereo. In: *IEEE Conference, Computer Vision and Pattern Recognition (CVPR)*, June 13–18, 2010, pp. 2769–2776. San Francisco, CA.

Linder, W., (2009) *Digital Photogrammetry 3rd Edition A Practical Course*

Longuet-Higgins, H.C. (1981) A computer algorithm for reconstructing a scene from two projections. *Nature*, 293: 133–135.

Lourakis, M.I.A. & Argyros, A.A. (2009) SBA: a software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software*, 36 (1article 2). doi: 10.1145/1486525.1486527.

Lowe, D.G. (1999) Object recognition from local scale-invariant features. In: *International Conference on Computer Vision*, September 20–27, 1999, pp. 1150–1157. Corfu, Greece.

Lowe, D.G. (2001) Local feature view clustering for 3D object recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2001, 7 pp. Kauai, Hawaii.

Lowe, D.G. (2004) Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60: 91–110.

Matas, J., Chum, O., Urban, M. & Pajdla, T. (2004) Robust wide baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22 (10), 761–767.

Meshroom Release v2021.0.1 (2022) Meshroom Contributors

Micheletti, N., Chandler, J.H. & Lane, S.N. (2014) Investigating the geomorphological potential of freely available and accessible Structure-from-Motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, 40 (4), 473–486.

Micusik, B. & Pajdla, T. (2006) Structure from Motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (7), 1135–1149.

Mikolajczyk, K. & Schmid, C. (2005) A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27: 1615–1630.

Mikolajczyk, K., Tuytelaars, T., Schmid, C. et al. (2005) A comparison of affine region detectors. *International Journal of Computer Vision*, 65 (1/2), 43–72.

Morel, J. & Yu, G. (2009) ASIFT: a new framework for fully affine invariant image comparison. *SIAM Journal on Imaging Sciences*, 2: 438–469.

Morgan, J., e Brogan, D. (2016). *How to VisualSFM*, Department of Civil & Environmental Engineering Colorado State University, Fort Collins, Colorado

Nistér, D. (2004) An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26: 756–777.

Nocedal, J. & Wright, S.J. (2006) *Numerical Optimization*. Springer, New York.

Olkowicz, M., Dabrowski, M., e Pluymakers, A. (2019), Focus stacking photogrammetry for micro-scale roughness reconstruction: a methodological study. *Photogram Rec*, 34: 11-35.

Osmana, M., e Tahar, K. (2016) "3D accident reconstruction using low-cost imaging technique," *Advances in Engineering Software 100*, vol. 100, p. 231–237.

Pons, J.-P., Keriven, R. & Faugeras, O. (2007) Multi-View Stereo reconstruction and scene flow estimation with a global image-based matching score. *International Journal of Computer Vision*, 72: 179–193.

Rahaman, H., e Champion, E. (2019). "To 3D or Not 3D: Choosing a Photogrammetry Workflow for Cultural Heritage Groups" *Heritage* 2, 3: 1835-1851

Regard3D (2022) <http://www.regard3d.org/index.php/documentation/introduction>, acessado em 15 abril 2022

Reljić, I., e Dunder, I. (2019) Application of Photogrammetry in 3D Scanning of Physical Objects TEM Journal 1:94-101

Rippin, D.M., Pomfret, A. & King, N. (2015) High resolution mapping of supraglacial drainage pathways reveals link between micro-channel drainage density, surface roughness and surface reflectance. *Earth Surface Processes and Landforms*. doi: 10.1002/esp.3719.

Riquelme, A., Araújo, N., Cano, M., Pastor Navarro, J., Tomás, R., Miranda, T. (2020), Identification of Persistent Discontinuities on a Granitic Rock Mass Through 3D Datasets and Traditional Fieldwork: A Comparative Analysis, em: Proceedings of the 3rd International Conference (ICITG), Guimarães, Portugal.

Rosnell, T. & Honkavaara, E. (2012) Point cloud generation from aerial image data acquired by quadcopter type micro unmanned aerial vehicle and a digital still camera. *Sensors*, 12: 453–480.

Schoenberger, J. L. e Frahm, J. (2016) Structure-from-Motion Revisited: Conference on Computer Vision and Pattern Recognition (CVPR)

Schöning, J., e Heidemann, G. (2015) Evaluation of multi-view 3D reconstruction software. In Proceedings of the International Conference on Computer Analysis of Images and Patterns, Valletta, Malta. 450–461

Seitz, S.M. & Dyer, C. (1999) Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35 (2), 151–173.

Seitz, S.M., Curless, B., Diebel, J., Scharstein, D. & Szeliski, R. (2006) A comparison and evaluation of Multi-View Stereo reconstruction algorithms. In: *IEEE Computer Society Conference, Computer Vision and Pattern Recognition*, Vol. 1, June 17–22, 2006, pp. 519–528.

Setiyadi, S., Mukhtar, H., e Cahyadi, W. (2021) "A Comparative Study of Affordable Photogrammetry Software for Reconstructing 3D Model of a Human Foot," *IEEE 7th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2021, pp. 286-291

Shen, S. (2012) Depth-map merging for Multi-View Stereo with high resolution images. In: *Pattern Recognition (ICPR), 21st International Conference*, November 11–15, 2012, pp. 788–791. Tsukuba.

Shortis, M.R., Bellman, C.J., Robson, S., Johnston, G. J. & Johnson, G.W. (2006) Stability of zoom and fixed lenses used with digital SLR cameras. In: *Proceedings of the ISPRS Commission V Symposium of Image Engineering and Vision Metrology*, September 25–27, 2006, pp. 285–290. Dresden.

Snavely N. (2008) Scene reconstruction and visualization from internet photo collections, PhD thesis, University of Washington.

Snavely, N., Seitz, S.N. & Szeliski, R. (2008) Modeling the world from internet photo collections. *International Journal of Computer Vision*, 80: 189–210.

Strecha, C., Bronstein, A., Bronstein, M. & Fua, P. (2012) LDAHash: improved matching with smaller descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34: 66–78.

Sunglok, C., Taemin, K. & Wonpil, Y. (2009) Performance evaluation of RANSAC family. In: *Proceedings of the British Machine Vision Conference (BMVC'09)*, September 7–10, 2009. London.

Szeliski, R. (2011) *Computer Vision: Algorithms and Applications*. Springer, London.

Tannús, J., Favareto, I., Lamounier, E., Cardoso, A. (2019) Comparação entre Técnicas de Fotogrametria e Escaneamento de Luz Estruturada para Reconstrução de Objetos em 3D. In: WORKSHOP DE INICIAÇÃO CIENTÍFICA - SIMPÓSIO DE REALIDADE VIRTUAL E AUMENTADA (SVR), 21 p. 21-22.

Thoeni, K., Giacomini, A., Murtagh, R. & Kniest, E. (2014) A comparison of Multiview 3D reconstruction of a rock wall using several cameras and a laser scanner. In: *Proceedings of ISPRS Technical Commission V Symposium*, June 23–25, 2014, pp. 23–25. Riva del Garda, Italy.

Tomasi, C. & Kanade, T. (1992) Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9: 137–154.

Triggs, B., McLauchlan, P.F., Hartley, R.I. & Fitzgibbon, A.W. (2000) Bundle adjustment – a modern synthesis. In: B. Triggs, A. Zisserman & R. Szeliski (eds), *Vision Algorithms '99, LNCS 1883*, pp. 298–372. Springer-Verlag, Berlin Heidelberg.

Turner, D., Lucieer, A. & Wallace, L. (2014) Direct georeferencing of ultrahighresolution UAV imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52, 2738–2745.

Ullman, S. (1979) The interpretation of Structure from Motion. *Proceedings of the Royal Society B*, 203: 405–426.

Wu, C., Agarwal, S., Curless, B., e Seitz, M. (2011) "Multicore bundle adjustment," *CVPR 2011*, pp. 3057-3064

Wu,C. (2013) "Towards Linear-Time Incremental Structure from Motion," *2013 International Conference on 3D Vision - 3DV 2013*, pp. 127-134

## 7. Anexos



a) Experiência nº 1



b) Experiência nº 2



c) Experiência nº 3



d) Experiência nº 4



e) Experiência nº 5

**Figura 55.** Nuvem de pontos densa

```

colmap feature_extractor^
--image_path workspace/images^
--database_path workspace/database.db^
|rem

colmap exhaustive_matcher^
--database_path workspace/database.db^
|rem

mkdir workspace\sparse

colmap mapper^
--database_path workspace/database.db^
--image_path workspace/images^
--output_path workspace/sparse^
|rem

mkdir workspace\dense

colmap image_undistorter^
--image_path workspace/images^
--input_path workspace/sparse/0^
--output_path workspace/dense^
--output_type COLMAP^
--max_image_size 3600^
|rem

colmap patch_match_stereo^
--workspace_path workspace/dense^
--workspace_format COLMAP^
--PatchMatchStereo.geom_consistency true^
|rem

colmap stereo_fusion^
--workspace_path workspace/dense^
--workspace_format COLMAP^
--input_type geometric^
--output_path workspace/dense/fused.ply^
|rem

```

**Figura 56.** "Pipeline" coordenadas GPS Colmap