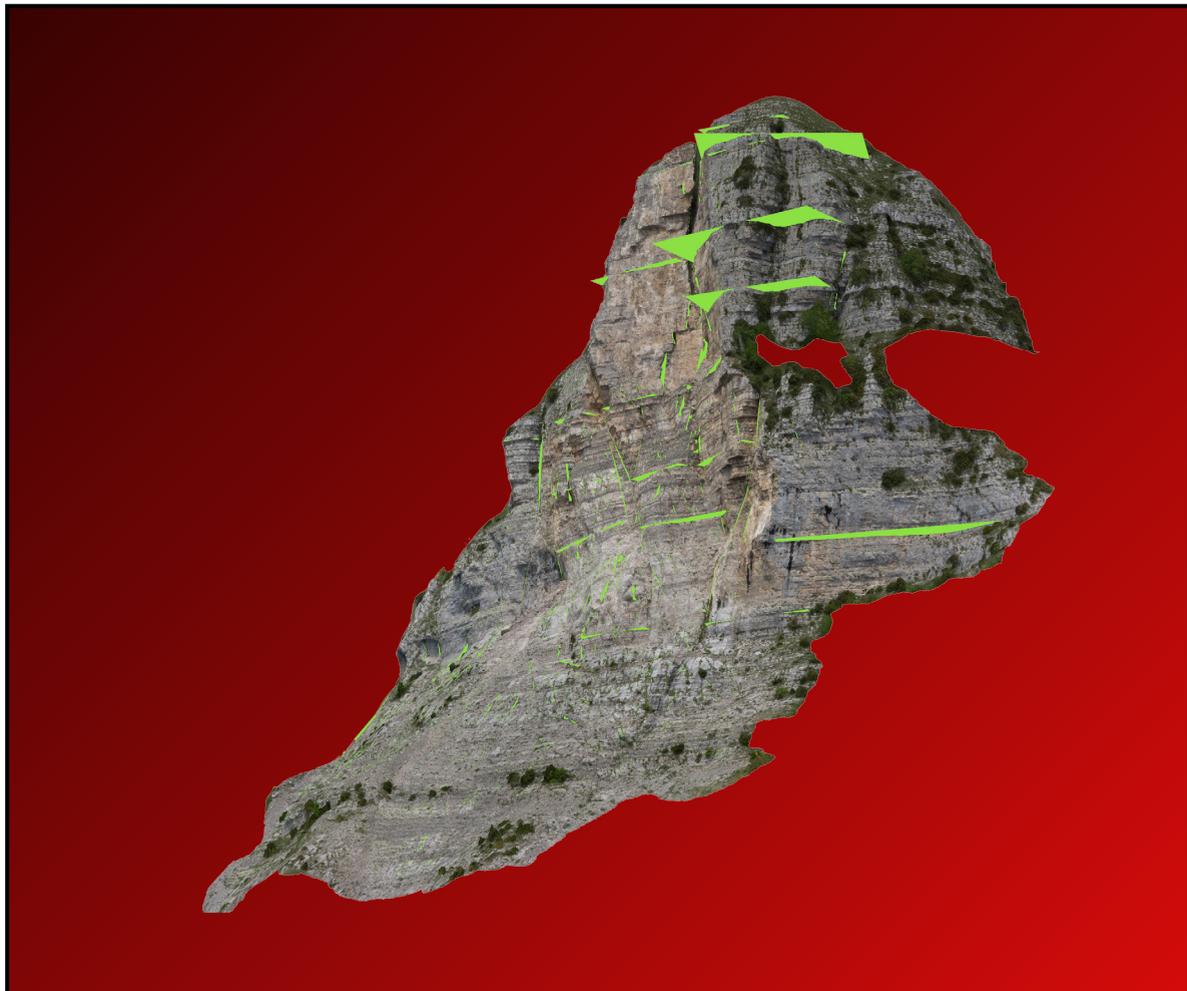


# VIRTUAL OUTCROP MODELS OF GEOLOGICAL STRUCTURES

From the construction of photogrammetric 3D models to their application towards the analysis of geological structures



Stefano Tavani, Amerigo Corradetti, Marco Mercuri, Thomas Seers







**SOCIETÀ GEOLOGICA ITALIANA**

FONDATA NEL 1881 - ENTE MORALE R. D. 17 OTTOBRE 1885



geologica  
internazionale  
**GIGS**

© Società Geologica Italiana, Roma 2024

ISBN 978-88-944844-6-5

<https://doi.org/10.3301/MON.2024.01>

# **Virtual outcrop models of geological structures**

## **From the construction of photogrammetric 3D models to their application towards the analysis of geological structures**

This textbook aims to introduce the reader to the basic principles and methods of virtual outcrop models and their applications within structural geology. We first provide the reader with the basic concepts of virtual outcrop models and an overview of methods commonly employed for digitally scanning outcrop surfaces: lidar and Structure from Motion - Multi-View Stereo (SfM-MVS) photogrammetry. We then focus upon SfM-MVS photogrammetry, illustrating its fundamental concepts. Finally, we illustrate key applications of virtual outcrop models in the field of structural geology. In addition, the book features six training exercises that serve to illustrate the fundamental steps and issues related to the creation and interpretation of models.

### **ABOUT THE AUTHORS**

**Stefano Tavani** is Professor in structural geology at the Department of Earth Sciences, Environment, and Resources at Federico II University of Naples and Associate Professor at the National Research Council of Italy, IGAG. He teaches undergraduate and graduate courses on virtual outcrop models, field mapping, and structural geology. His research expertise is in virtual outcrops, structural geology, tectonics, and geodynamics.

**Amerigo Corradetti** is an Assistant Professor in structural geology at the Department of Mathematics, Informatics and Geosciences of the University of Trieste. His research includes the study of compressional tectonic settings of the Mediterranean and the Middle East region, faults and fractures evolution, distribution and analysis, and the characterization

of reservoir analogues. Amerigo's studies have often taken advantage of digital photogrammetry and virtual outcrop modeling.

**Marco Mercuri** is a post-doctoral researcher in structural geology at the Department of Earth Sciences of the Sapienza University of Rome. His research focuses on the study of brittle deformation structures in carbonate reservoir analogues, with particular focus on fault structure and mechanics. Marco's studies are performed by integrating classical fieldwork and the interrogation of virtual outcrop models.

**Thomas Seers** is an Associate Professor within the Petroleum Engineering Program at Texas A&M University at Qatar (TAMUQ). Dr. Seers teaches graduate and undergraduate courses in geology, petroleum engineering and interdisciplinary engineering. His research focuses upon acquisition, processing and analysis of close-range 3D remote sensing data for geologic applications, including structural geology and reservoir characterization.

# **VIRTUAL OUTCROP MODELS OF GEOLOGICAL STRUCTURES**

From the construction of photogrammetric 3D models  
to their application towards the analysis of geological structures

**STEFANO TAVANI**

*University of Naples Federico II, CNR-IGAG*

**AMERIGO CORRADETTI**

*University of Trieste*

**MARCO MERCURI**

*Sapienza University of Rome*

**THOMAS SEERS**

*Texas A&M University at Qatar*



# Table of contents

Preface	9
Abbreviations	10
<b>1. Introduction to Virtual Outcrop Models</b>	<b>11</b>
1.1 Definition of Virtual Outcrop Model	11
1.2 Overview of acquisition methods	12
1.2.1 <i>Lidar</i>	13
1.2.2 <i>SfM-MVS photogrammetry</i>	15
1.3 Types of VOM	17
1.3.1 <i>Point cloud</i>	17
1.3.2 <i>Mesh</i>	19
1.3.3 <i>Pros and cons</i>	20
1.4 Data products derived from VOMs	21
1.5 Hardware requirements	22
Suggested readings	26
<b>2. Background to Structure from Motion – Multi-View Stereo photogrammetry</b>	<b>29</b>
2.1 How digital cameras work	29
2.2 The collinearity equations	31
2.2.1 <i>Collinearity equation in 2D</i>	31
2.2.2 <i>Collinearity equations in 3D</i>	34
2.3 The Structure from Motion - Multi-View Stereo workflow	36
2.3.1 <i>Feature detection, keypoint correspondence and filtering</i>	37
2.3.2 <i>Structure from Motion</i>	41
2.3.3 <i>Multi-View Stereo</i>	43
2.3.4 <i>Scaling, orienting and georeferencing</i>	44
2.4 Resolution, accuracy, and errors	45
Suggested readings	47
Suggested software tools	49

<b>3</b>	<b>Data acquisition</b>	53
3.1	Know your tools	53
3.1.1	<i>DSLR and mirrorless cameras</i>	53
3.1.2	<i>Smartphones</i>	54
3.1.3	<i>Uncrewed Aerial Vehicles (UAVs)</i>	56
3.1.4	<i>Gimbal stabilizer</i>	57
3.1.5	<i>Global Navigation Satellite System (GNSS) sensors</i>	58
3.1.6	<i>Other tools</i>	59
3.2	Acquisition planning	60
3.2.1	<i>Photo-survey</i>	60
3.2.2	<i>Precautions for registration</i>	61
3.2.3	<i>Multi-camera acquisition</i>	62
3.2.4	<i>Video acquisition</i>	62
<b>4</b>	<b>Applications to geological structures</b>	65
4.1	Detecting and measuring near-planar geological features	66
4.1.1	<i>Background to planar best-fitting</i>	66
4.1.2	<i>Model segmentation for planar best-fitting</i>	69
4.1.3	<i>Best practices</i>	71
4.2	Orthorectifying VOMs	73
	Suggested readings	76
<b>5</b>	<b>SfM-MVS in practice</b>	79
	Exercise 1: Terrestrial DSLR	80
	Exercise 2: UAV model	81
	Exercise 3: Model from Video	82
	Exercise 4: Smartphone + Gimbal stabilizer	83
	Exercise 5: Digitization and orthorectification	84
	Exercise 6: Digitization of traces and the effect of outcrop orientation	85
	Acknowledgments	86

# PREFACE

Over the past two decades, key advancements in geospatial digital data capture have revolutionized outcrop studies in the geosciences. These developments have been fueled by the proliferation of close-range remote sensing techniques capable of accurately reconstructing geologic surfaces (i.e., digital or virtual outcrop models: DOMs/VOMs) amongst geoscience practitioners, such as portable light detection and ranging (lidar), and computer vision-based photogrammetric reconstruction (i.e., Structure from Motion - Multi-View Stereo photogrammetry: SfM-MVS). The diversification and improvement of low-cost sensory platforms capable of capturing these geospatial datasets (e.g., consumer-grade digital cameras, smartphones/tablets, Uncrewed Aerial Vehicles / UAVs) have further catalyzed this digital revolution in field data collection, propelling the use virtual outcrop data into routine field studies. Whilst initially employed as a visualization medium, virtual outcrop models have rapidly evolved into fully interrogatable quantitative geo-data objects from which a diverse suite of rock properties can be extracted. Although early virtual outcrop studies principally relied upon terrestrial lidar for model construction, their rise to prominence within field studies is largely due to the emergence of SfM-MVS photogrammetry. The low cost and portability of tools used in SfM-MVS photogrammetry, such as digital single-lens reflex (DSLR) cameras, mirrorless cameras, camera drones and smartphones/tablets, have had a transformative impact upon the uptake of close-range remote sensing within the geosciences. Such tools provide an egalitarian approach to VOM generation, making virtual outcrop data capture practically available to all. Indeed, low-cost digital cameras enable the construction of models with spatial resolutions down to a few tens of microns, which in most cases compares favorably with models generated by lidar. Consequently, SfM-MVS photogrammetry has rapidly replaced lidar as the de facto close-range remote sensing technique leveraged towards the generation of virtual outcrop models within field studies.

## Abbreviations

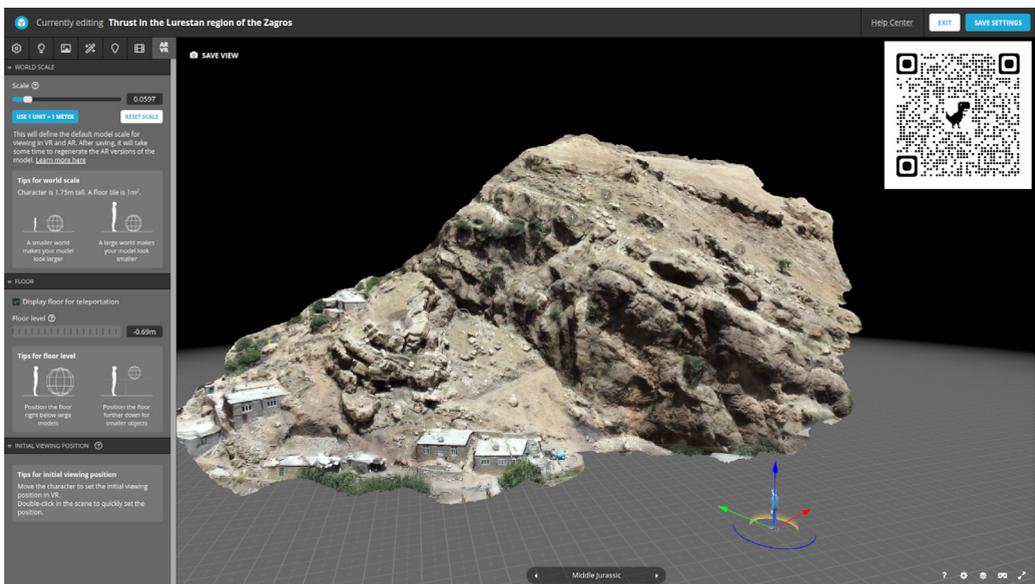
CMYK	Cyan, Magenta, Yellow, Key
CPU	Central Processing Unit
DEM	Digital Elevation Model
dGNSS	differential Global Navigation Satellite System
DOM	Digital Outcrop Model
DSLR	Digital Single-Lens Reflex
GCP	Ground Control Point
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPU	Graphics Processing Unit
GSD	Ground sampling distance
ECC	Error-correction code
EXIF	EXchangeable Image File format
HSV	Hue, Saturation, and Value
lidar	light detection and ranging
MVS	Multi-View Stereo
PCA	Principal Component Analysis
PMVS	Patch-based Multi-View Stereo
radar	radio detection and ranging
RAM	Random Access Memory
RGB	Red–Green–Blue
RGBA	Red–Green–Blue–Alpha
RTK	Real Time Kinematics
RTK-GNSS	Real Time Kinematics Global Navigation Satellite System
SfM	Structure from Motion
SfM-MVS	Structure from Motion–Multi-View Stereo
SIFT	Scale-Invariant Feature Transform
sonar	sound navigation system
TIN	Triangulated Irregular Network
UAV	Uncrewed Aircraft Vehicle
VOM	Virtual Outcrop Model

# Chapter 1

## Introduction to Virtual Outcrop Models

### 1.1 Definition of Virtual Outcrop Model

A virtual outcrop model (VOM) is a 3D surface-based digital representation of a geological outcrop that can be visually navigated and analyzed (Fig. 1).



**Figure 1.** Virtual outcrop model of the Qlaji Thrust, Zagros, NW Iran, as seen in the 3D settings form of the Sketchfab public repository. <https://skfb.ly/6VTL6>

VOMs are generated using close range remote sensing techniques, such as digital photogrammetry or lidar (Light Detection And Ranging), and can incorporate additional attributes and metadata, such as geological features surficial properties, sample locations, and stratigraphic logs, to aid their interpretation. In its raw format, a VOM is simply a reconstruction of the outcrop surface, typically in the form of a colored point cloud, which can be analyzed to extract valuable geological information. Herein, the term 'outcrop' refers to rocks and geological features exposed at the Earth's surface. Therefore, models of geological media that are not physically attached to the original outcrop, such as detached fossils, or hand specimens of rocks or minerals, cannot be considered outcrop models under the above definition.

The surface-based nature of VOMs presents both challenges and opportunities for analysis, setting them apart from volumetric geological datasets, such as 3D seismic and ground penetrating radar (Radio Detection And Ranging). VOMs are typically limited to exposures or structures that span from a few tens of centimeters up to one km, depending on the scalability of the data capture platform used (i.e., terrestrial vs. UAV-based photogrammetry or lidar). Unlike classic topographic geospatial models of the Earth's surface, which are typically presented as raster data (e.g., digital elevation / terrain models), VOMs use unstructured point clouds or triangular irregular networks (TINs) to represent geological surfaces, facilitating the representation of overhanging geometry. This enables VOMs to extract and visualize geological information that is inaccessible from 2.5D raster data.

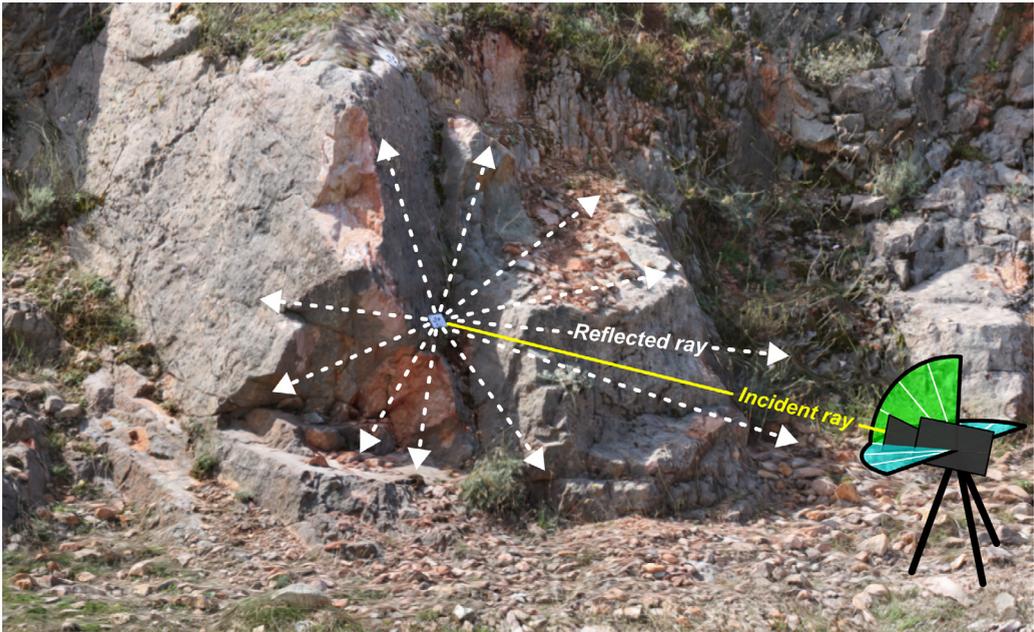
## **1.2 Overview of acquisition methods**

The section provides an overview of the two most commonly used methods generating virtual outcrop models: lidar and SfM-MVS photogrammetry.

### 1.2.1 Lidar

Lidar (originally a portmanteau of light and radar, later adapted to an abbreviation of Light Detection And Ranging) or laser scanning uses the emission of a laser beam with wavelengths ranging between infrared and ultraviolet (typically near infrared for most commercial terrestrial lidar systems), which is diffusely backscattered by a spot on the target object (provided the sensor-object distance is below a threshold value, which depends on the specific properties of the lidar sensor, the angle of incidence between the emitted laser and the target, and the albedo of the mapped surface). The working mechanism of a lidar-based system can be described as two paradigms: phase-based or time-of-flight-based lidar.

Phase-based lidar emits waves of different wavelengths and the distance between the lidar and the backscattering spot is calculated by analyzing the change of phase of the returned signal. In the time of flight-based ranging, by knowing the speed of light and measuring the time interval between the emission and return of the signal, lidar systems determine the distance between the sensor and the backscattering spot. Lidar is thus an active remote sensing technique in which the scanning tool is both the source and receiver of the signal used to determine scene geometry. Its working method is similar to that of sonar (SOund Navigation And Ranging) and radar (RAdio Detection And Ranging), which also represent active range finding techniques. The lidar emits a narrow cone-shaped laser beam, the diameter of which scales with the distance between the target and the sensor. Although differences exist amongst lidar systems, laser beam diameter generally ranges from a few micrometers close to the sensor up to a few millimeters at distances of hundreds of meters. This means that the backscattering spot can be considered a point for most applications. Terrestrial lidar sensors typically rotate around a vertical axis, with tilt adjustment enabling user control over the scan orientation (Fig. 2). The laser beam is deflected by a rotating mirror producing a vertical scan swath, which when coupled with rotation of the sensor allows for hemispherical scanning. During acquisition, the lidar scanner measures the distance of a point on the target object



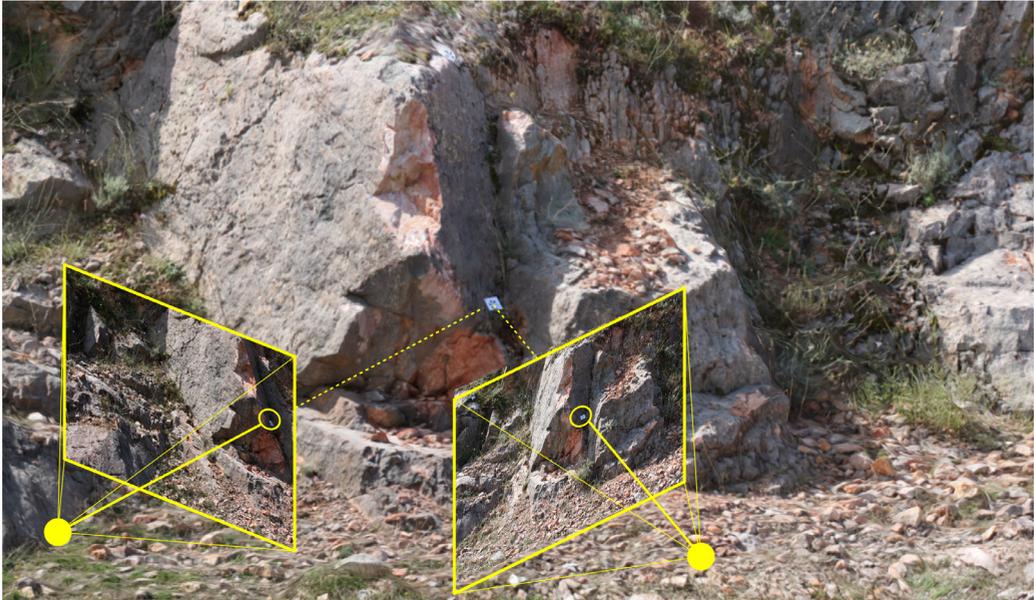
**Figure 2.** Scheme showing the background of lidar acquisition.

and its azimuthal and zenithal coordinates with respect to the detector. This operation, which returns the polar coordinates of a given point, is repeated for different azimuthal and zenithal coordinates within a given angular range (imposed by the particular scan parameters implemented). The azimuthal and zenithal ranges define a scanning window within which the target object is discretized into  $N$  points, of which the coordinates in a reference system whose origin is the lidar sensor are known. These points form a scaled XYZ point cloud representing the surface of the target object. The smaller the azimuthal and zenithal increments (or by analogy, the smaller the distance from the scanner to the target object), the higher the number of points within the scan path and the denser the point cloud. The error associated with each measurement and the threshold lidar-object distance are lidar model-specific and typically provided by the vendor (current survey grade models offer  $\pm 1\text{mm}$  ranging error under common operational conditions). As a general rule, the range increases with increasing target distance and decreasing reflectivity. It is worth noting

that environmental conditions can impact upon the operability of terrestrial lidar. For example, dusty or humid conditions can cause refractions and limit maximum operational range, with wet surfaces enhancing specular reflection of light and causing deleterious scan artifacts. Though having accurate scaling, it is worth noting that entry-level lidar models will output arbitrarily oriented and positioned uncolored point clouds with vertex intensity. Additionally, points have no color information but an intensity value indicating the return strength of the signal. However, most terrestrial lidar models have an integrated camera and Global Navigation Satellite System (GNSS) antenna, which facilitate the production of colored point clouds and colored/textured meshes and enable separate overlapping scans to be co-registered and spatially rectified into the global coordinate frame via the application of a rigid transform (rotation and translation). The intensity of the return signal is, at first approximation, a function of the reflectivity or albedo of the scanned object and can provide information about specific physical characteristics. Note that the angle of incidence and the distance between the lidar unit and the target object modulate the intensity property of a laser scan, which requires normalization to be useful as a quantitative surficial property.

### *1.2.2 SfM-MVS photogrammetry*

Photogrammetry is the science of making measurements from photographs. It is a passive sensor technique in which the sensor (i.e., the camera) has no direct control over the source of the signal (i.e., artificial and natural light sources). Photogrammetry was developed during the 19th century soon after the development of photography, although its basic concepts of perspective, parallax and projective geometry, date back to the 15th century. Stereophotogrammetry, of which Structure from Motion – Multi-View Stereo Photogrammetry (SfM-MVS) represents a specific methodology, is a subclass of photogrammetry. With analogy to human vision, stereophotogrammetry leverages overlapping photographs taken from different positions to reconstruct 3D scenes (Fig. 3).



**Figure 3.** Scheme showing the background to stereoscopic view and SfM-MVS photogrammetry surveying.

As detailed in the proceeding chapters, the position of a point on a photo can be approximately retrieved by a set of two equations (collinearity equations), which relate the X and Y coordinates of a point in a photograph to seven parameters: three for the 3D position of the point with respect to the optical center of the image, three for camera orientation, and one for camera focal length. Stereophotogrammetry is based upon the precept that the position of points and the pose of the camera within an arbitrary 3D space can be determined analytically by solving the collinearity equations for matched pairs of points identified within overlapping images with narrow baseline parallax. Given that the collinearity equations are not able to provide the scale of a scene, the raw output of the SfM-MVS workflow is an arbitrarily scaled, rotated, and translated point cloud, which requires registration (scaling, rotation, and translation). By default, vertex data output by SfM-MVS photogrammetry is typically accompanied by color information. Importantly, whilst lidar requires little post-acquisition processing, SfM-MVS photogrammetric

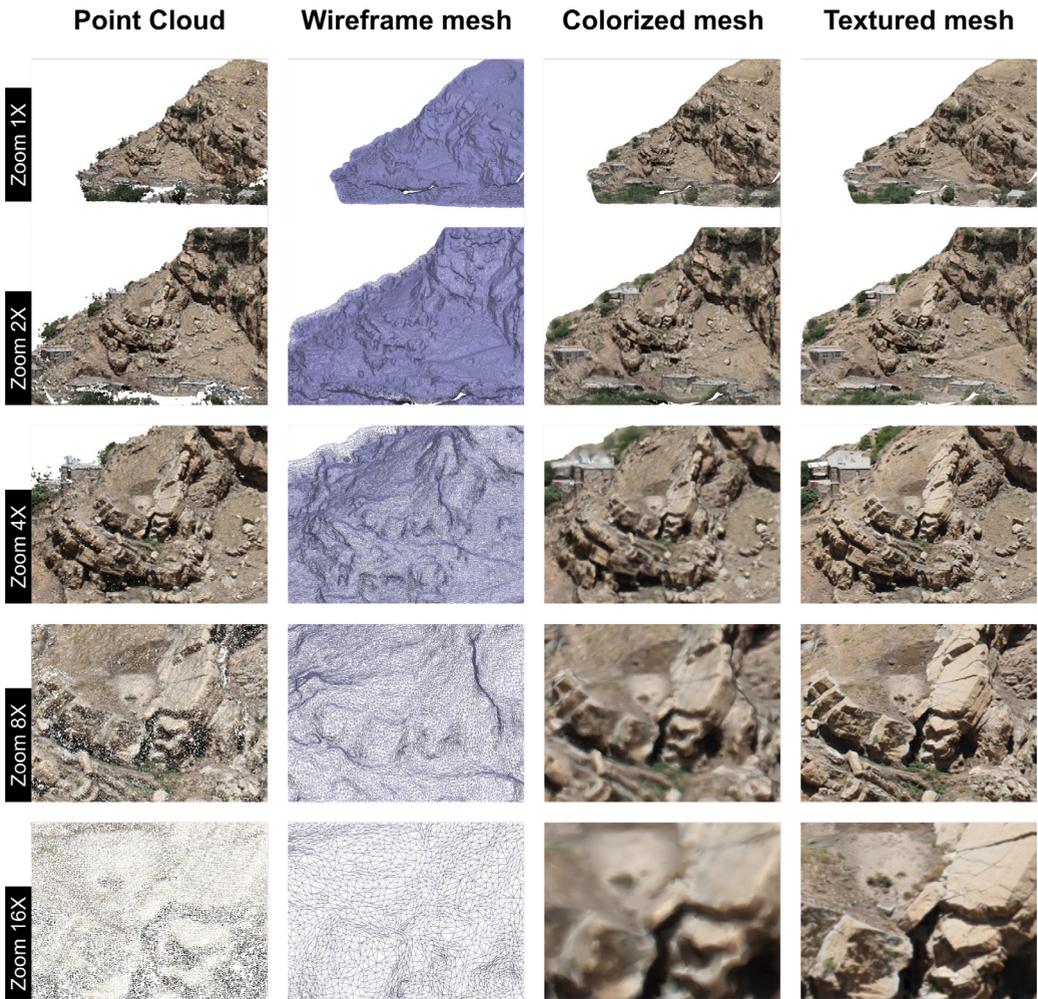
survey data requires significant user intervention during post-acquisition processing. Specifically, finding homologous points in the different photos is a time-consuming step which can impact upon the density and quality of the reconstruction. There are best practices for planning and carrying out an effective photo survey that optimizes the generation of an SfM-MVS photogrammetric model. Chapter 2 illustrates the basic principles of point detection and correspondence determination, with the aim of contextualizing these best practices.

### **1.3 Types of VOM**

#### *1.3.1 Point cloud*

The initial output of both lidar and SfM-MVS photogrammetry is an unstructured point cloud (Fig. 4). This is a list of 3D points whose spatial distribution is disparate and whose neighborhood topology is undefined. An unstructured point cloud can be gridded to produce a structured point cloud, in which two components (typically X and Y) are regularly spaced. Conversion of unstructured point clouds captured using airborne remote sensing platforms to a structured format is a common procedure for Geographic Information System (GIS) applications to enhance computation speed and memory handling. Structured point clouds are, however, rare in the representation of VOMs, as they require downsampling and smoothing of the original point cloud, resulting in considerable loss of information.

Regardless of whether the point cloud is structured or unstructured, its basic form is an array of  $N$  elements ( $N$  points) with three fields, corresponding to X, Y, and Z, (i.e., latitude, longitude, and elevation coordinates). RGB point clouds store color information in three additional fields. In some cases, the color information may include a fourth channel



**Figure 4.** Comparison, for the same outcrop, between the rendering of a point cloud, a wireframe mesh, a colorized mesh, and a textured mesh, at variable zoom values

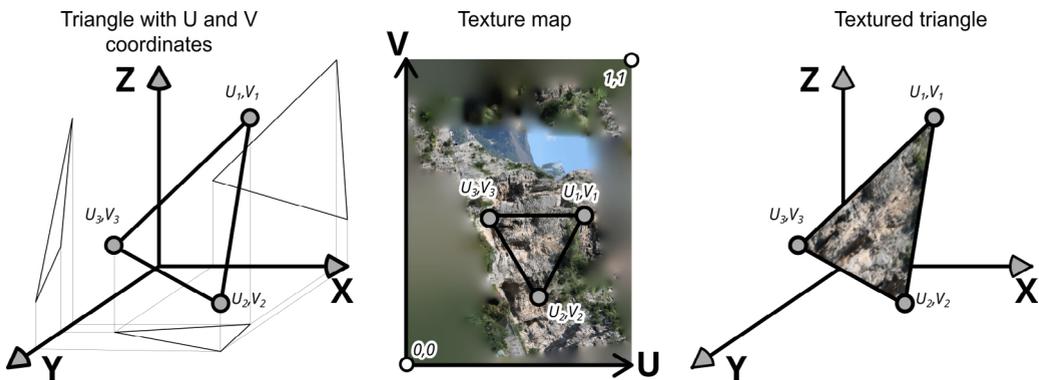
(the Alpha channel) representing transparency/opacity (RGBA format). Rarely, other color models may be used, such as HSV (Hue, Saturation, and Value) or CMYK (Cyan, Magenta, Yellow, and Key). Each point's normal vector may also be included in the vertex list, stored as three property fields representing the X, Y, and Z components of a unit vector. A number of operations require point normals to be computed, including illumination rendering, surface reconstruction, denoising and artifact

removal. Finally, the vertex properties may also encompass bespoke property fields computed directly from the point cloud (e.g., curvature, coplanarity). In summary, a point cloud composed of  $N$  points, whether structured or unstructured, will include an array of  $N$  elements with  $M$  fields, potentially storing color (and opacity), point normals and/or bespoke surficial properties. As detailed in Subchapter 1.5, the higher the number of fields for each point, the higher the hardware requirements.

### *1.3.2 Mesh*

A mesh is a tessellated set of points, forming a subordinate network of polygons connected by row vectors of indices known as a face list. Meshes are commonly used to represent surfaces in various application areas, including computer-generated imagery and 3D gaming. Similarly to point clouds, meshes can be structured (i.e., with regular ordering between nodes) or unstructured. It should be noted, however, that unlike point clouds, the local topology of an unstructured mesh is established within the neighborhood of each vertex. Though quadrilateral meshes are used within some modeling applications, triangles are by far the most efficient and well-established polygon types in computer graphics for 3D geometry representation. Consequently, meshes used to represent VOMs typically use irregular triangles to represent surface geometry (Triangulated Irregular Network / TIN). Since the initial output of both lidar and SfM-MVS photogrammetry is a point cloud, triangular meshes are derived from unstructured point-sets through a process called triangulation. There are many algorithms used in computational geometry to triangulate point clouds. Commonly, triangulation algorithms tend to smooth asperities, discard outlier points, and return a mesh whose number of triangles is defined by the user. Concerning the colors, meshes can be of two types: colorized and textured. In colorized meshes, each vertex of each triangle has a color assigned and each point within the triangle is a blended combination of its three component vertices' colors. Textured meshes have a more sophisticated structure, which can facilitate

improved retention of the optical signature of a VOM (see, for example, the 16X zoom in Figure 4), but also demands more overhead from the Graphics Processing Unit (GPU). A texture (often also called a texture map) is an image draped onto the triangular mesh. Two coordinates, named U and V, are assigned to each vertex of each triangle of the mesh. U and V individuate a point in the texture, and the U and V coordinates of the three vertices of each triangle of the mesh permit mapping textures to their corresponding triangles on the mesh (Fig. 5); this process is called texture mapping. Similarly to point clouds, vertex or face normals, as well as other surficial or rock mass properties can be added to the mesh (e.g., curvature, coplanarity, fracturing intensity, porosity, and equivalent rock mass permeability).



**Figure 5.** Procedure for texturing a triangular mesh.

### 1.3.3 Pros and cons

The main advantage of meshes over point clouds is that they allow for the visualization, picking, and interrogation of the space between vertices, which is not possible with point clouds, particularly when utilizing a limited field of view (see Figure 4). A secondary advantage is that the availability of the topology of each vertex's neighborhood allows faster computation

of surface properties (i.e., via region growing) when compared to point clouds, which typically require computationally demanding kernel based spatial searches to establish the local neighborhood for surface property computation. Textured meshes are especially useful when the user properly balances the level of detail between the resolution of triangles and the texture, resulting in an optimal graphical rendering. However, it should be noted that large textured meshes are computationally demanding and require powerful GPUs.

## **1.4 Data products derived from VOMs**

There are various data products that can be extracted from 3D models. Two important data products for geoscience applications are digital elevation models (DEMs) and orthophotos.

DEMs are structured grids (raster data) with one coordinate representing the distance from a reference plane and the other two defining regularly spaced nodes lying on the reference plane. Typically, the reference plane for DEM is the horizontal plane, with X and Y regularly spaced and Z defining the elevation. However, it is also possible to have other regularly spaced coordinates, such as X and Z (variable Y) or Y and Z (variable X), or even oblique reference planes. 3D point clouds or meshes require resampling to generate DEMs. DEMs are therefore simplified versions of the 3D models, which are more efficient for rendering and analysis within common GIS software. It should be noted, however, that DEMs are considered as 2.5D datasets, being unable to represent overhanging geometries which are common to natural outcrops and manmade excavations.

Orthophotos, on the other hand, are akin to 'screenshots' of the 3D model displayed in an orthographic view. To create an orthophoto from a 3D model, you project the points or textured triangles onto a plane perpendicular to the plane of projection. Typically, the projection plane is the horizontal plane, but for some applications, vertical or oblique

planes can be used. An orthophoto has several advantages over a 3D virtual outcrop model in terms of rendering (for example they allow for delineating polylines using vector graphic software tools). In similitude to DEMs, orthophotos represent 2.5D datasets.

## 1.5 Hardware requirements

Hardware is a crucial aspect when using software tools aimed at creating and interpreting VOMs. These tools must allow for smooth navigation in all directions in 3D (i.e., three axis rotations and translations of the viewport), which is not always possible within common GIS software that often has a preferred zenithal view. To enable this functionality, software tools utilize 3D libraries like OpenGL, Vulkan, or DirectX, which are commonly used in computer game development. As a result, the best PCs and laptops for creating, visualizing, and interpreting VOMs are often those designed for gaming with dedicated graphics cards. Note that whilst arguably offering the best performance, professional grade workstations which have error-correction code (ECC) GPUs and RAM are not generally required for VOM processing and visualization.

Creating a VOM with the SfM-MVS photogrammetry requires uploading tens to hundreds (or in some cases thousands) of photos, which are scanned pixel by pixel in search of hundreds to thousands of keypoints within each photo. Each keypoint of each photo is then compared with those from all photos in the image dataset to detect correspondences. This process, which is only part of the SfM-MVS workflow, can keep the RAM and the GPU busy for hours, even days. The higher the number and resolution of photos, the higher the computational time (assuming hard limits are not placed upon the number of keypoints detected). Having more RAM and a powerful GPU typically results in shorter processing times. There is no straightforward way to predict how long it will take to create a VOM using the SfM-MVS photogrammetry based on the number of photos, their resolution, and the RAM and GPU capacity. The

acquisition quality can also impact processing time, with expert users' image datasets often being processed much faster than those collected by novice users. While most PCs and laptops can be used for SfM-MVS modelling, it is worth noting that the RAM and GPU capacity not only affect processing time but also set a limit on the maximum number of pixels that can be processed. Entry-level PCs will take much longer to process tens of photos compared to cutting-edge hardware. If hundreds to thousands of photos are uploaded, an entry-level PC's CPU and RAM capacity will quickly become depleted, causing it to crash.

To give the reader an idea, a single image with a resolution of 5344x4016 pixels and RGB 8-bit channels will occupy 64 MB of space when uploaded to the RAM. This value is not dependent on the original file size, as even compressed formats like JPG will be fully decompressed in the RAM. Photogrammetric software can save on RAM and GPU usage, so using 1000 such images will not occupy 64 GB of RAM, but only a fraction of that size. However, any PC or laptop with 16 GB of RAM will easily run out of memory with such a dataset. From our experience, a PC with 4 GB of RAM and an integrated graphics card can be used for entry-level models using a few tens of photos for training purposes. High-resolution models using hundreds of photos and resulting in point clouds or meshes with tens of millions of points or faces require at least 16 GB of RAM and, ideally, a dedicated graphics card. Thousands of photos intended for constructing models with hundreds of millions of points require much more powerful hardware, such as 128 GB of RAM and high-end dedicated gaming graphics cards. Concerning storage capacity, regardless of the photogrammetry software used, input photos, point clouds, and textured meshes can easily consume several GB, meaning that the storage capacity of a PC or external hard drive used for routine SfM-MVS model construction should be several TB.

After creating photogrammetric models using SfM-MVS, it is common to export the resulting point clouds or meshes in various file formats. The size of these files can vary greatly, depending on the format used and the number of elements in the point cloud or mesh. Point clouds can be exported in formats such as LAS, PLY, DXF, SHP, and CSV/XYZ. These

formats may include a header and a description of how the data is stored. In its basic binary format, a point cloud includes 64 bits for each of the three coordinates, 8 bits for each of the three- or four-color channels (red, green, blue, and alpha), and 64 bits for each of the three vertex normal components. However, the vertex normal component is not useful for some geological applications and can be excluded from point cloud files to reduce their size. The alpha channel is also often discarded, resulting in only RGB channels being included and reducing the size to 27 bytes per point. Using 32 bits instead of 64 bits for the coordinates produces further compression, which results in 15 bytes per point (despite this can cause issues if the model is referenced using some global coordinate systems due to loss of precision). In text format, each point can take up to 55 to 60 bytes for a XYZ-RGB point cloud. In summary, a medium-sized point cloud with 10 million points can take anywhere between 150 MB and 0.5 GB of storage space, depending on the file format used.

Meshes, on the other hand, can be exported in formats such as GLTF, GLB, PLY, and OBJ. All these formats store the mesh and the texture in separate files, except for GLB, which merges them into a single binary file. The mesh file is typically organized as follows: a short header, a list of properties for each triangle (including the XYZ coordinate), the UV texture coordinate, and the vertex normal of each vertex of each triangle. Each triangle has eight fields for each of the three vertices: X, Y, Z, U, V, and the three vertex normal coordinates. As an example, Figure 6 illustrates the structure of the wavefront file (OBJ). Assuming that the vertex normal is not included in the dataset and using a 32-bit format for XYZ and UV coordinates, each triangle requires  $5 \times 3 \times 32 \text{ bits} = 480 \text{ bits} = 60 \text{ bytes}$ . Therefore, a mesh with 10 million triangles will take 0.6 GB. Since triangles frequently share one or two vertices, space-saving formats write triangles' vertex information (XYZ, UV, and vertex normal) as separate lists. In this case, each vertex will have three fields corresponding to the row in the XYZ, UV, and vertex normal lists in which the data is written. Although storing XYZ and UV coordinates in separate lists may reduce the required storage space, the order of magnitude will be almost the same. Regarding texture files, their format is typically



using highly compressed formats for textures, such as the JPG format, reduces texture space on the disk, but it does not improve rendering speed. Once the texture is uploaded to the graphics processing unit, it will indeed be uncompressed.

It is not uncommon for a PC or laptop that is able to create a certain SfM-MVS photogrammetric model to be unable to render it smoothly. In such cases, it is recommended to downsize or split the point cloud/mesh or use several small textures instead of a single large texture.

In conclusion, working with SfM-MVS photogrammetry and interpretive software tools requires RAM and GPU capacities typical of cutting-edge hardware. One common source of problems is ignoring the hardware limitations during the steps that lead from photo acquisition to model interpretation. This is a particularly relevant issue for practitioners, especially in cases where the model's creator and interpreter do not communicate effectively.

## SUGGESTED READINGS

Bellian, J.A., Kerans, C., and Jennette, D.C., 2005, Digital Outcrop Models: Applications of Terrestrial Scanning Lidar Technology in Stratigraphic Modeling: *Journal of Sedimentary Research*, v. 75, p. 166-176, <https://doi.org/10.2110/jsr.2005.013>

Bemis, S.P., Micklethwaite, S., Turner, D., James, M.R., Akciz, S., Thiele, S.T., and Bangash, H.A., 2014, Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology: *Journal of Structural Geology*, v. 69, p. 163-178, <https://doi.org/10.1016/j.jsg.2014.10.007>

Granshaw, 2020, Photogrammetric terminology: fourth edition: *The Photogrammetric Record*, v. 35(170), p. 143-288, <https://doi.org/10.1111/phor.12314>.

Hodgetts, D., 2013. Laser scanning and digital outcrop geology in the petroleum industry: A review. *Marine and Petroleum Geology*, v. 46, p. 335-354. <https://doi.org/10.1016/j.marpetgeo.2013.02.014>

- Jones, R.R., McCaffrey, K.J.W., Wilson, R.W., and Holdsworth, R.E., 2004, Digital field data acquisition: towards increased quantification of uncertainty during geological mapping: Geological Society, London, Special Publications, v. 239, p. 43-56, <https://doi.org/10.1144/GSL.SP.2004.239.01.04>
- Pearce, M.A., Jones, R.R., Smith, S.A., and McCaffrey, K.J., 2011. Quantification of fold curvature and fracturing using terrestrial laser scanning. AAPG Bulletin, v. 95, p. 771-794. <https://doi.org/10.1306/11051010026>
- Seers, T.D.S., Hodgetts, D.H., Wang, Y.W. and Fadlelmula, M.F., 2017. Direct computation of fracture network equivalent porous medium properties using digital outcrop models. In 79th EAGE Conference and Exhibition 2017 (Vol. 2017, No. 1, pp. 1-5). European Association of Geoscientists & Engineers. <https://doi.org/10.3997/2214-4609.201700647>
- Xu, X., Aiken, C.L.V., Bhattacharya, J.P., Corbeanu, R.M., Nielsen, K.C., McMechan, G.A., and Abdelsalam, M.G., 2000, Creating virtual 3-D outcrop: The Leading Edge, v. 19, p. 197-202, <https://doi.org/10.1190/1.1438576>



## Chapter 2

# Background to Structure from Motion – Multi-View Stereo photogrammetry

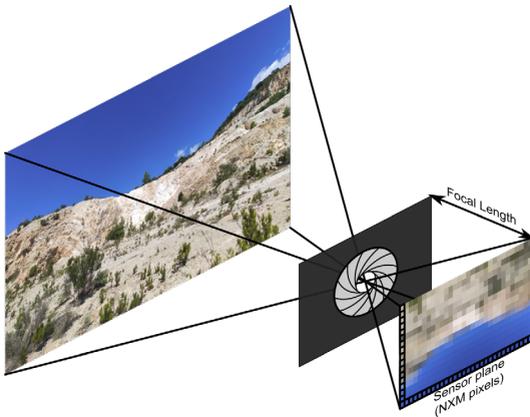
Structure from Motion (SfM) and Multi-View Stereo (MVS) photogrammetry are two key steps in a workflow that involves taking a set of partly overlapping images, detecting common points across the images, and reconstructing the positions and orientations of both the images and detected points (i.e., SfM), ultimately transforming them into a dense point cloud (i.e., MVS). Specifically, after detecting common points in the overlapping images, SfM is responsible for estimating the camera pose (i.e., photos' locations and orientations) and the 3D location of the matched keypoints, whilst MVS algorithms utilize the determined positions and orientations of the images (along with other internal camera parameters, as detailed in the following subsections) to generate a 3D dense point cloud.

To understand the SfM-MVS workflow, it is important to provide a background understanding of how data is captured and how an image pixel can be mathematically transformed into a 3D point.

### 2.1 How digital cameras work

Digital photographs are 2D images composed of  $N \cdot M$  pixels, created by capturing light with a digital sensor. The larger the sensor, the more light is gathered, generally resulting in better image quality. Within a 3D

scene, the location of a point on an image plane is contingent upon the trajectory that light reflected from the point traverses as it reaches the sensor. Light passes through an adjustable aperture diaphragm and a lens positioned at varying distances from the sensor. The mathematics behind the SfM-MVS photogrammetric pipeline is based upon the simplified pinhole camera model, which has no lens and an infinitesimally small aperture for the diaphragm (Fig. 7). In this simplified optical model, there is a direct relationship between points in space and their corresponding projections on the image plane. Correspondences between 3D points in the 3D scene and 2D points on the image plane via rays that pass through the optical center, which coincides with the diaphragm, are defined by a set of paired equations known as the collinearity equations.

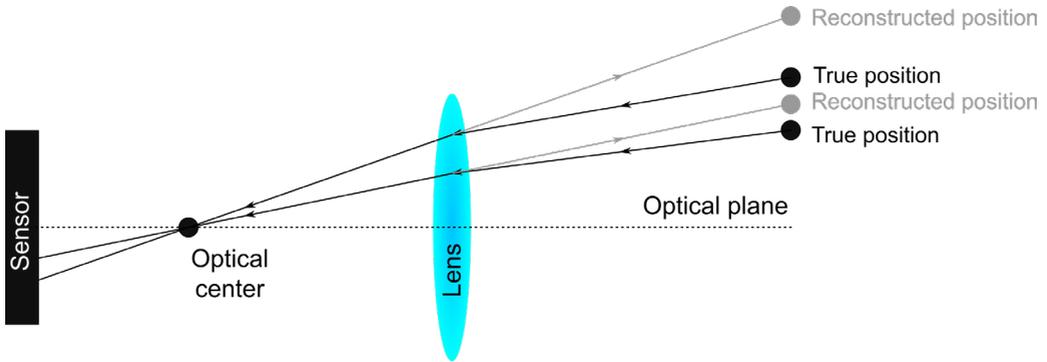


**Figure 7.** Pinhole camera model for a digital camera.

The collinearity equations in the pinhole camera model require seven parameters to determine the XY position of a point on the image plane. These parameters are: (1) the position and orientation of the camera, which are represented by 6 parameters (3 for orientation and 3 for position) known as the extrinsic camera parameters; and (2) the focal length, which is the distance between the

optical center and the image plane. Additionally, in real cameras, since the digital sensor returns data in pixels, whereas the positions of points are in metric reference systems, the pixel size of the sensor (typically expressed in micrometers) must also be known. Within the SfM-MVS workflow, the collinearity equations based on the pinhole camera model are commonly enhanced by incorporating adjustments to account for image distortion caused by the lens (i.e., distortion coefficients). The most common correction is for radial distortion. Radial distortion is caused by the refraction of light within the lens and results in the radial displacement of

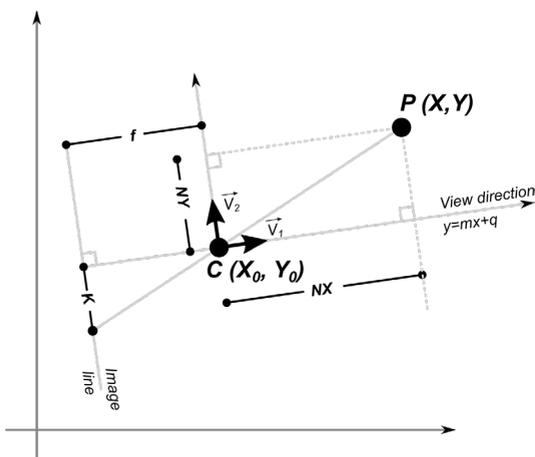
points when the pinhole camera model is assumed (as illustrated in Figure 8). These intrinsic camera parameters, including the distortion factor(s), sensor size, and focal length, are essential for accurate 3D reconstruction.



**Figure 8.** 2D radial distortion produced by the lens when the pinhole camera model is assumed.

## 2.2 The collinearity equations

### 2.2.1 Collinearity equation in 2D



**Figure 9.** Scheme for the derivation of the collinearity equation in 2D.

In this section, we derive the collinearity equation for a 2D space to illustrate the mathematical principles that also apply to the more complex 3D case. In Figure 9, point C ( $X_0, Y_0$ ) represents the optical center of a generic photograph, and the line corresponding to the photo's view direction is  $y = mx + q$ . The focal length, denoted by  $f$ , is the distance

between the image line and the optical center. Consider a point P (X, Y) in space, which is projected onto the image line at a distance K from the center of the image. The projection is the intersection between the image line and the ray connecting C and P. Using the triangle similarity rule, it can be shown that:

$$\frac{K}{f} = \frac{NY}{NX} \quad (1)$$

Here, NX and NY are the X and Y coordinates of P in a new reference system centered at C, with two principal axes parallel to the view direction and the image line, respectively. The two unit vectors  $\vec{V}_1$  and  $\vec{V}_2$  that define this system are:

$$\vec{V}_1 = \left( \frac{1}{\sqrt{1+m^2}}, \frac{m}{\sqrt{1+m^2}} \right) \quad (2)$$

$$\vec{V}_2 = \left( \frac{-m}{\sqrt{1+m^2}}, \frac{1}{\sqrt{1+m^2}} \right) \quad (3)$$

The coordinates NX and NY are expressed as:

$$NX = \overrightarrow{PC} \cdot \vec{V}_1 \quad (4)$$

$$NY = \overrightarrow{PC} \cdot \vec{V}_2 \quad (5)$$

Where PC is the vector from C to P. These equations can be rewritten as:

$$NX = (X - X_0) \cdot V_{1X} + (Y - Y_0) \cdot V_{1Y} \quad (6)$$

$$NY = (X - X_0) \cdot V_{2X} + (Y - Y_0) \cdot V_{2Y} \quad (7)$$

Substituting equations 2 to 7 into equations 1, we obtain the collinearity equation:

$$\frac{K}{f} = \frac{(Y-Y_0) - m \cdot (X-X_0)}{m \cdot (Y-Y_0) + (X-X_0)} \quad (8)$$

Since all points along the line passing through C and P have the same distance K from the image center, knowing K/f and the position and orientation of the photo are not sufficient to retrieve the X and Y coordinates of P (that is all points collinear with P (X,Y) project on the same image

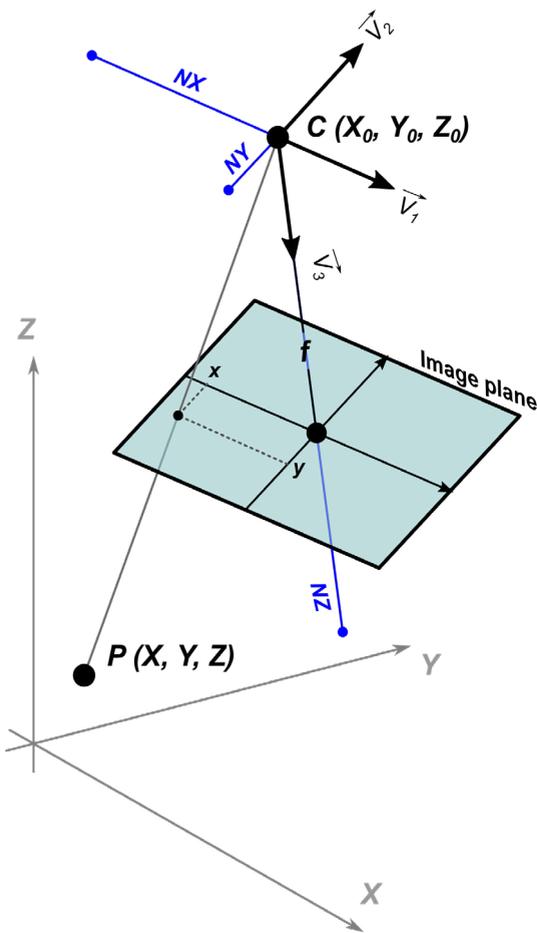
point regardless of their distance from C). To retrieve the position of P, a second photo on which the point P is visible and for which the position, orientation, focal length, and translation of the projected real-world point on the image line must be added. This is indeed the basic requirement for a stereoscopic view (i.e., a stereopair of known position and orientation). Now, let us consider the case where the position and orientation of photos are unknown and need to be retrieved from N points viewed in M photos, of which the focal length (and thus the ratio of K and  $f$ ) is known. In this case, we will have  $N \cdot M$  collinearity equations with a number of unknown parameters equal to  $N \cdot 2$  (i.e.,  $x$  and  $y$  of each point) +  $M \cdot 3$  ( $x$ ,  $y$ , and  $m$  of each photo). To solve the system, the following conditions must be satisfied:

$$N \cdot M \geq 2N + 3M; M \geq \frac{2N}{N-3}; M \geq 2 + \frac{6}{N-3} \quad (9)$$

Equation 9 indicates at least three photos are required to reconstruct a 2D scene (note that somewhat counterintuitively, only two photos are needed in the 3D case: see Section 2.2.2). The collinearity equation has two major drawbacks: (1) The right-hand side of Equation 8 is a ratio between differences, meaning that the reconstructed camera pose and scene cannot be properly scaled, positioned, and rotated in relation to the world coordinate frame. As a result, it is not possible to produce scene reconstructions suitable to metrology using the collinearity equations alone. (2) The collinearity equation only provides a ratio between K and the focal length, and thus knowing the extrinsic camera parameters and the coordinates of each point in the scene does not permit computing the intrinsic camera parameters simultaneously. Either the sensor size or the focal length must be known *a priori* to determine the remaining parameter set. Digital image formats include key intrinsic camera parameters: namely focal length and sensor size. However, establishing other non-linear parameters (i.e., distortion coefficient) and the optical center in image coordinates (i.e., the principal point) require camera calibration. Robust algorithms for self-calibration of images are included in the SfM workflow, accurately determining all intrinsic camera parameters, including distortion parameters (i.e., via a bundle adjustment). These algorithms often work

with pre-defined parameters for the most common sensors/cameras, resulting in a low failure rate for the self-calibration procedure. In rare cases when the self-calibration fails, the software may assign arbitrary values to reconstruct the 3D scene, but with lower accuracy.

### 2.2.2 Collinearity equations in 3D



**Figure 10.** 3D Pinhole camera models and scheme for deriving the collinearity equations in 3D. For the sake of simplicity, the image plane is placed ahead of the optical center.

The image that was represented by a line in 2D is a plane in 3D. A point in a 3D space is projected at the x and y coordinates of the image plane (Fig. 10). Additionally, image orientation, which is a function of a single parameter in 2D, in 3D requires three mutually perpendicular axes to be defined: one for the view direction (perpendicular to the image plane), one for the long axis of the photo, and one for the short axis of the photo. In the scheme of Figure 10,  $X_0, Y_0,$  and  $Z_0$  represent the coordinates of the optical center (C); x and y represent the coordinates in the image plane of a point of interest P  $(X, Y, Z)$ . The three orthogonal unit vectors defining the image plane are  $\vec{v}_1, \vec{v}_2,$  and  $\vec{v}_3$ . The coordinates of P in the reference system centered in C and having  $\vec{v}_1, \vec{v}_2,$  and  $\vec{v}_3$  as bases are NX, NY, and NZ.

Analogously to the 2D case, in 3D we have:

$$\frac{x}{f} = \frac{NX}{NZ} \quad (10)$$

$$\frac{y}{f} = \frac{NY}{NZ} \quad (11)$$

where  $f$  is the focal length in the pinhole camera model of Figure 10, and where:

$$NX = (X - X_0) \cdot V_{1X} + (Y - Y_0) \cdot V_{1Y} + (Z - Z_0) \cdot V_{1Z} \quad (12)$$

$$NY = (X - X_0) \cdot V_{2X} + (Y - Y_0) \cdot V_{2Y} + (Z - Z_0) \cdot V_{2Z} \quad (13)$$

$$NZ = (X - X_0) \cdot V_{3X} + (Y - Y_0) \cdot V_{3Y} + (Z - Z_0) \cdot V_{3Z} \quad (14)$$

Substituting equations 12 to 14 into equations 10 and 11, we obtain the two collinearity equations in 3D:

$$x = f \frac{V_{1X}(X-X_0) + V_{1Y}(Y-Y_0) + V_{1Z}(Z-Z_0)}{V_{3X}(X-X_0) + V_{3Y}(Y-Y_0) + V_{3Z}(Z-Z_0)} \quad (15)$$

$$y = f \frac{V_{2X}(X-X_0) + V_{2Y}(Y-Y_0) + V_{2Z}(Z-Z_0)}{V_{3X}(X-X_0) + V_{3Y}(Y-Y_0) + V_{3Z}(Z-Z_0)} \quad (16)$$

It is worth noting that if there are  $N$  points visible in  $M$  photos and the focal length is known, then in 3D, we will have  $2 \cdot N \cdot M$  collinearity equations with  $N \cdot 3$  (i.e.,  $x$ ,  $y$ , and  $z$  of each point) +  $M \cdot 6$  ( $x$ ,  $y$ , and  $z$ , + three parameters defining the orientation of the photo) unknown parameters. To solve the system, the following condition must be satisfied:

$$2 \cdot N \cdot M \geq 3N + 6M; M \geq 1.5 + \frac{1.5}{N-3} \quad (17)$$

As  $N$  increases, the right-hand side of the equation tends to zero and  $M$  approaches 1.5. Therefore, as mentioned earlier, at least two overlapping images are required to reconstruct a 3D scene. Finally, similar to the 2D example, the reconstructed 3D scene is arbitrarily scaled, positioned, and oriented.

## 2.3 The Structure from Motion - Multi-View Stereo workflow

The Structure from Motion - Multi-View Stereo (SfM-MVS) workflow results in the reconstruction of a 3D scene in the form of a point cloud from a set of images with initially unknown extrinsic and intrinsic calibration parameters. This workflow can be applied using various software tools, both open-source and commercial. Though Structure from Motion - Multi-View Stereo photogrammetry is coined from two critical projective geometry problems (i.e., SfM and MVS), the solutions to which it is heavily reliant upon, the SfM-MVS workflow comprises six main steps:

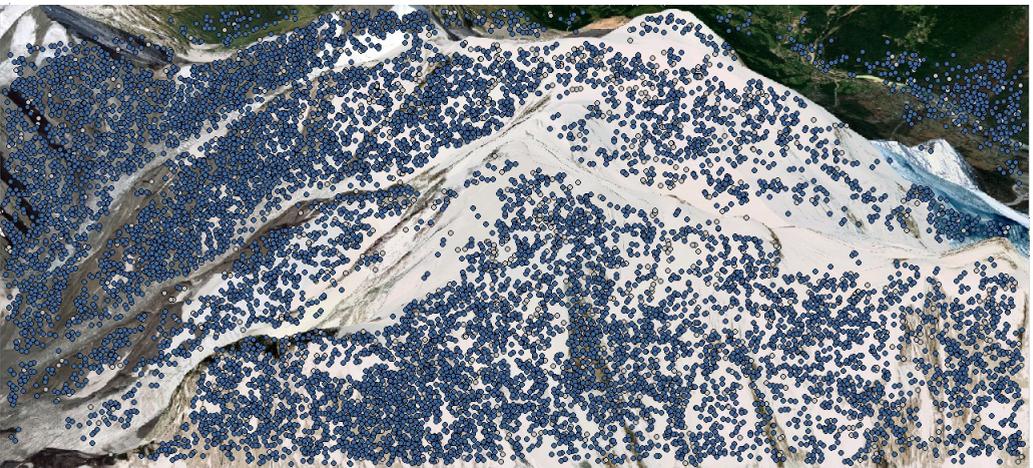
- 1) Feature detection
- 2) Keypoint correspondence
- 3) Keypoint filtering
- 4) Structure from Motion
- 5) Multi-view stereo
- 6) Model registration

The first step involves detecting features in each image, such as points, lines, or edges. The second step involves identifying corresponding features between images. In the third step, the detected features are filtered to remove any false matches. In the fourth step, the SfM technique is used to estimate the camera positions and the sparse 3D points. This results in a sparse point cloud that represents the scene. In the fifth step, the MVS algorithm is used to densify the sparse point cloud, generating a dense point cloud. This step typically involves estimating the depth of each pixel in each image, refining the camera positions and orientations, and optimizing the final point cloud. Finally, in the last step, the models are registered using a geometric transformation (typically a similarity transform: translation, rotation, scale) to spatially rectify the point cloud and/or camera positions into a global or local coordinate system.

### *2.3.1 Feature detection, keypoint correspondence and filtering*

Recognizing homologous points in overlapping photos (i.e., feature point detection) is a prerequisite for SfM estimation. This process involves identifying common points, or sets of pixels, on overlapping images, known as keypoints. Keypoints allow for matching different images to reconstruct scene geometry, providing the nascent structure for the sparse reconstruction of the scene. While several techniques for keypoint detection exist, the most widely used is the Scale-Invariant Feature Transform (SIFT). SIFT offers the advantage of detecting features and describing keypoints even under challenging conditions, such as large variations in viewpoint (camera rotation and translation) and scale, partial occlusion, and variable illumination. The density of keypoints identified in an image varies widely and depends on the texture, sharpness, and resolution of the image. Typically, high numbers of keypoints are extracted from heterogeneous textures, sharp images, and high-resolution images. Accordingly, since the individuation of keypoints relies on color patterns, regions exhibiting uniform colors will generate a restricted set of keypoints. Figure 11, showing the Monte Bianco peak as seen in Google Earth, exemplifies this issue, illustrating that areas covered by snow yield fewer keypoints compared to areas in which the bedrock is exposed.

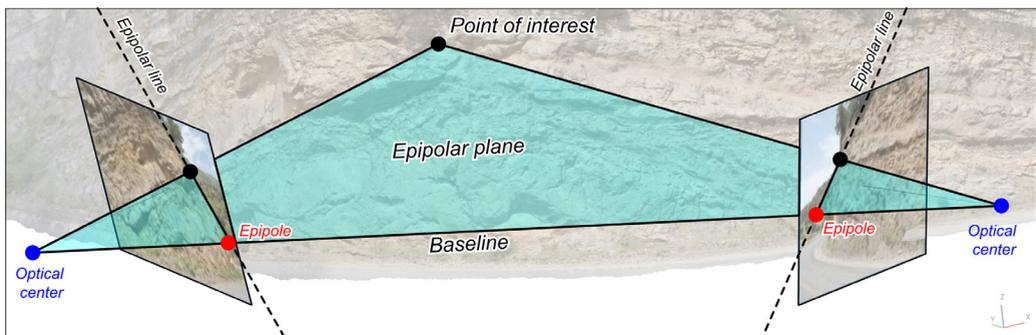
The keypoints that have been detected in each image through feature detection need to be compared with those in other images to identify potential matches. This step is called keypoint correspondence, which involves searching for and evaluating each potential keypoint correspondence within an image dataset and discarding keypoints that cannot be matched. Although SIFT can work with different scales and variable illumination, minimizing variations in scale between overlapping images and capturing image data under constant illumination conditions (i.e., by restricting the acquisition window during photogrammetric surveys) are best practices that generally enhance the performance of the establishing keypoint correspondences.



**Figure 11.** View of the Monte Bianco peak in Google Earth, with detected keypoints shown as dots.

Some erroneous correspondence can still be present even after the keypoint correspondence step. To refine the matches between different images, an additional step called keypoint filtering is applied. In this step, a fundamental matrix is calculated for each pair of images, which constrains the locations of correctly identified keypoints in both images and permits the verification of the geometrical consistency of the matches for the reconstruction of a 3D scene. To understand this process, it should be noted that, in addition to a detected point shared by two camera views,

the optical center of each camera capturing two views of the same scene center maps onto a unique location on its paired camera's image plane (epipoles or epipolar points) (Fig. 12). The line which connects the paired camera centers and intersects their epipolar points is known as the *baseline*. Both optical centers and the detected 3D point form an *epipolar plane* which intersects both the epipolar points and projected real world point on the stereo pair's respective image planes. Finally, the line that connects the epipolar point to a projected real-world point on the image plane is known as the *epipolar line*. The above-described geometric system (i.e., epipolar geometry) provides the epipolar constraint over the matching of the location of a point in one image to its equivalent location in a second image, reducing the search space from the entire image to a single line.



**Figure 12.** Epipolar geometry of a stereo view.

Correspondences between matched keypoints are used to estimate the position of points and camera pose, with false correspondences being progressively removed (Fig. 13). Several algorithms using different parameters for the construction of the fundamental matrix (e.g., RANSAC, the eight-point algorithm, Levenberg-Marquardt, and others) are commonly used to elicit outlier points.



**Figure 13.** Correspondence between keypoints detected in two sets of images, with the blue and red lines connecting points that have and have not passed the correspondence filtering step, respectively. (A) Example of image pair with both successful and discarded connections. (B) Example of keypoints correspondences that have not passed the filtering step. Keypoints in the two photos look exactly the same, being part of two target objects having identical appearance. Nevertheless, the two targets are placed in different positions, making impossible to reconstruct the photo positions using these correspondences.

At the end of the keypoint filtering step, a link between images sharing a set of matching keypoints, called a track, can be defined. Additionally, most software tools allow the user to manually add some keypoints of known correspondence, which are generally referred to as markers (Ground Control Points or GCPs when their position in real-world coordinates is known).

### 2.3.2 Structure from Motion

Structure from motion is the SfM-MVS workflow component that uses the keypoints defined in the previous steps for the simultaneous recovery of: (1) the tridimensional structure of the scene, including 3D object points camera positions and orientations (i.e., pose estimation or extrinsic calibration parameters) (Fig. 14); and commonly (2) intrinsic camera calibration parameters.



**Figure 14.** Scene after the SfM step, with camera and keypoints position reconstructed.

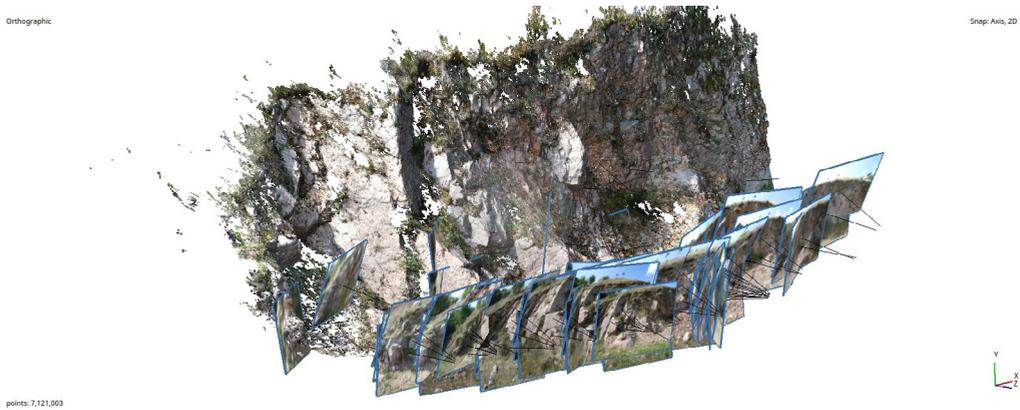
As described in Chapter 2.2, the extrinsic parameters are defined by a matrix that relates the 3D scene coordinates to the camera coordinate system. The intrinsic camera parameters include the distortion coefficient(s), sensor size, and focal length and are defined by a 3x3 matrix called the camera calibration matrix or intrinsic matrix that, amongst other factors, depends on the camera type (e.g., full frame camera, fisheye camera, 360° camera). Note that some intrinsic parameters (e.g., focal length and sensor size) typically accompany the image file (e.g., EXIF: EXchangeable Image File).

Bundle adjustment is the method used to estimate the three-dimensional scene and requisite camera parameters while minimizing the model fitting error. The term ‘bundle’ refers to the bundles of rays connecting

camera centers to 3D object points in the scene, and 'adjustment' refers to the minimization of the reprojection error. To initiate scene reconstruction, a pair of images, known as the 'initial pair', is selected. For a robust reconstruction, the initial pair should be characterized by a large number of matched points, as well as a significant shift in perspective producing significant displacement (i.e., parallax) between object points viewed in the matched views. The parameters are then refined by progressively adding more camera views. As more cameras are added, the camera poses and the 3D scene are progressively optimized. The process ends when no remaining cameras contain the sufficient number of reconstructed 3D points to be reliably added to the model. The Structure from Motion process produces a sparse point cloud as its output, which may not be sufficient to satisfactorily reconstruct geological surfaces for most geoscience applications, due to its low point density. To address this, a further image processing step (Multi-View Stereo reconstruction) is typically needed to densify the point cloud.

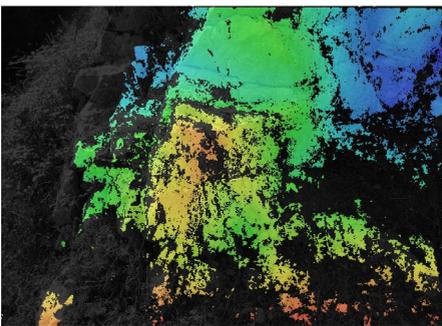
In some cases, there may be overlapping areas with poor image coverage, leading to the SfM algorithm returning multiple independent models. To address this issue, user-defined markers can be manually imposed to force the connection of these models and return a single scene. However, this approach can result in some distortion of the model. Therefore, it is generally recommended to consider placing markers as the last resort, with the repetition of image acquisition often being the preferred solution when SfM fails to produce a cohesive model of a single scene.

### 2.3.3 Multi-View Stereo



**Figure 15.** Reconstructed scene after the MVS step.

Multi-View Stereo reconstruction is used to increase the point cloud density obtained from Structure from Motion estimation, resulting in a dense point cloud that in most practical applications, is at least two orders of magnitude denser than the sparse point cloud (Fig. 15).



**Figure 16.** Example of depth map of an image.

Multi-View Stereo algorithms can be classified into four categories: voxel-based, surface evolution-based, depth map merging, and patch-based methods. Voxel-based and surface evolution-based methods reconstruct the 3D scene using voxel (i.e., a 3D pixel) and polygonal meshes, respectively. Depth map merging algorithms create a depth map for each image (Fig. 16) and combine them into a single 3D model. In essence, each image is converted into a DEM-like

domain where pixels assume distance information in the reference system of the image from the viewpoint of the image to the scene in the image. Patch-based methods grow small patches called surfels to reconstruct the scene. One widely used patch-based algorithm for Multi-View Stereo is patch-based MVS (PMVS), which has three main steps. In the first step, the algorithm detects specific shapes like corners and blobs in the images and matches them across different images. The matches are then used in the second step to identify neighboring pixels that can be used to expand the reconstructed scene, with matches being filtered in the final step.

#### *2.3.4 Scaling, orienting and georeferencing*

After the reconstruction of the scene using the SfM-MVS workflow, the objects in the scene (points and cameras) are positioned in an arbitrary coordinate system. This means that the scene is likely not properly scaled or oriented with respect to the world frame (i.e., unregistered), which may be necessary for many geological applications. To register the scene, a similarity transformation involving rigid rotation, rigid translation, and isotropic scaling needs to be applied. In some cases (i.e., when geographical context is not necessary), only scaling and rotation may be required to register the scene into a local reference system, resulting in a partially registered scene/model. Depending on the software used or specific user requirements, registration can be undertaken at any stage of the SfM-MVS workflow.

To fully or partly register the scene, the simplest way is to use the built-in functionality of the SfM-MVS software tools used to reconstruct the scene, to derive the 3D similarity transform matrix (a 4x4 matrix representing the uniform scale, translation and rotation of the registered scene to the target scene). This requires measuring the position of at least three matched non-collinear points in both the arbitrary and real-world/local coordinate systems. These points can be either those belonging to the scanned outcrop or camera positions. Accordingly, there are two

ways to register models: (i) using Ground Control Points (GCPs) or (ii) using cameras positional information.

Both methods, have their own advantages and disadvantages. Using GCPs tends to be more accurate within practical applications, as the coordinates of the markers can be obtained with high-precision GNSS antennas or from existing terrain models. However, GCPs typically need to be manually inserted into the scene prior to surveying, which can be challenging in difficult terrain or weather conditions. Using camera location, on the other hand, does not require physical access to the scene. However, the accuracy of the camera locations obtained through GNSS antennas equipped in smartphones or UAVs is usually significantly poorer than measurements conducted with RTK or differential antennas, typically used for geolocating GCPs. Therefore, the choice between these two methods depends on the specific needs and characteristics of the project, such as the size of the scene, accessibility, and required accuracy.

In some cases, solutions that utilize both orientation data (such as planar/linear features in the scene or cameras' view directions) and positional data can be adopted. These solutions can be less straightforward, as they may require some computations outside of the SfM-MVS software. Generally, their use is limited to specific situations, which are discussed in Chapter 3 of this book.

## **2.4 Resolution, accuracy, and errors**

In its general meaning, resolution is the slightest change that can be detected by an instrument. For instance, in the case of a thermometer, the resolution is the smallest increment that is detectable from it, for example, 0.5 degrees Celsius. In the field of photogrammetry, the term most often refers to the spatial resolution achievable from a camera device, which is the smallest discernible detail in an image. Resolution is therefore the combined performance of the entire camera system. Is resolution equivalent

to the number of pixels on a camera sensor? The short answer is no! The resolution of a camera (or spatial resolution) is of course dependent on the sensor size and the number of pixels that define the pixel size, but it is also dependent on how the light reaches the sensor. Therefore, additional factors, including lens sharpness/distortion and exposure parameters such as the ISO number, shutter speed, and diaphragm aperture (which is the reciprocal of the f-stop number), contribute to defining how sharp/detailed/focused the image will be. Finally, the camera focal length and the aperture (f-stop) control the depth of field (DoF), which is the range of distances that will appear in sharp focus in the final image. For instance, narrow apertures contribute to keeping a deep DoF with an extended ability to capture details that are in sharp focus over a wide range of distances. The spatial resolution is not to be confused with the ground sampling distance (GSD), the latter being the pixel size expressed by the size of the pictured object in the real-world unit. If we know the width of the camera sensor ( $S_w$ , expressed in mm), the focal length of the camera ( $f$ , expressed in mm), and the image width ( $I_w$ , expressed in pixels), we can estimate the GSD as a function of the distance to the outcrop ( $D$ ) as follows:

$$GSD = \frac{S_w \cdot 100 \cdot D}{f \cdot I_w} \quad (18)$$

During the SfM-MVS pipeline, other things being equal, images with finer resolution will result in denser point clouds. By association with the resolution of laser scanner devices, the resolution of an SfM-MVS-derived point cloud is often associated with the total number of points by scanned area or by point-to-point average distance. Nevertheless, by analogy with image resolution, 3D points constituting noise degrade the spatial resolution of the point cloud since those points do not contribute to resolving details. At the same time, a sparse point cloud would necessarily have a lower resolution than a denser point cloud. This combination of factors (i.e., how dense and noisy a point cloud is) implies that, given a certain dataset, there exists an achievable optimal resolution through the SfM-MVS workflow. Similarly, the spacing of vertices quantifies the resolution of a mesh. Finally, we can describe the texture resolution by its pixel count.

When dealing with 3D geometries reconstructed through the SfM-MVS workflow, we always come across two non-exchangeable terms: precision and accuracy. The general definition of accuracy is how close a certain measurement is to a true value. In SfM-MVS-derived products, accuracy is generally defined as the closeness of the internal geometry and/or the absolute position of points in a model in relation to the world frame. In essence, to define the accuracy of a model, it is necessary to know the absolute or relative coordinates of points in the world frame. The best way to evaluate the accuracy of an SfM-MVS-derived product is to compare it against a benchmark. A benchmark should be obtained by a different survey method of known accuracy, for example, by a laser scanner or a total station. The closer the SfM-MVS-derived model is to the benchmark, the higher the accuracy. The deviation between the model and the benchmark is the error associated with the SfM-MVS reconstruction. Frequently, during SfM-MVS surveys, two sets of GCPs are collected: one is used for registering the models, and the other is used as a benchmark and will serve to compute the error between the estimated and measured positions. Unlike accuracy, precision refers to how close different measurements are to each other, even if there is a large error. In essence, if a SfM-MVS-derived model is uniformly distant from a benchmark, it can be considered precise but not accurate.

## SUGGESTED READINGS

- Carrivick, J.L., Smith, M.W., and Quincey, D.J., 2016, Structure from Motion in the Geosciences: Chichester, UK, John Wiley & Sons, Ltd, <https://doi.org/10.1002/9781118895818>
- Furukawa, Y. and Ponce, J., 2009. Accurate camera calibration from multi-view stereo and bundle adjustment. International Journal of Computer Vision, v. 84, p 257-268, <https://doi.org/10.1109/CVPR.2008.4587681>

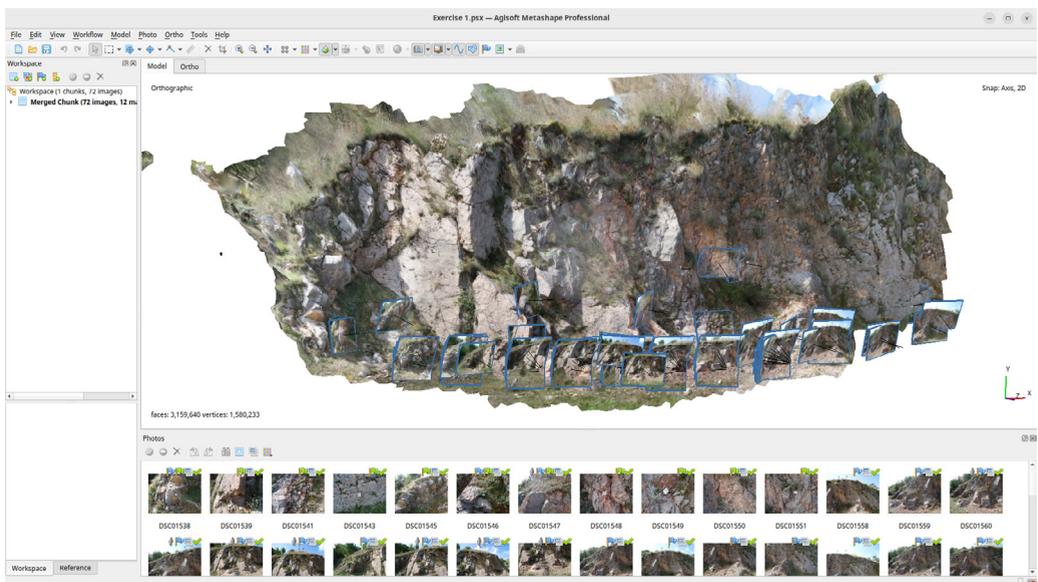
- James, M.R., and Robson, S., 2012, Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application: *Journal of Geophysical Research*, v. 117, p. F03017, <https://doi.org/10.1029/2011JF002289>
- Lowe, D., 1999. Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision*, v. 2, p. 1150-1157. Ieee.
- Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W., 2000, Bundle adjustment—a modern synthesis. In *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings* (p. 298-372). Springer Berlin Heidelberg.
- Ullman, S., 1979, The interpretation of structure from motion: *Proceedings of the Royal Society of London. Series B Biological Sciences*, v. 203, p. 405-426, <https://doi.org/10.1098/rspb.1979.0006>
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M., 2012, 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications: *Geomorphology*, v. 179, p. 300-314, <https://doi.org/10.1016/j.geomorph.2012.08.021>

## SUGGESTED SOFTWARE TOOLS

Amongst standalone software tools for Structure from Motion – Multi-View Stereo Photogrammetry available in 2024, we suggest the following:

**Agisoft Metashape** (previously named Agisoft Photoscan) is the benchmark software for SfM-MVS, running on different Operating Systems (Win, Mac & Linux). The standard and pro versions are amongst the more affordable options in this category of software, with educational licenses (having however limited functionalities) available at reduced prices.

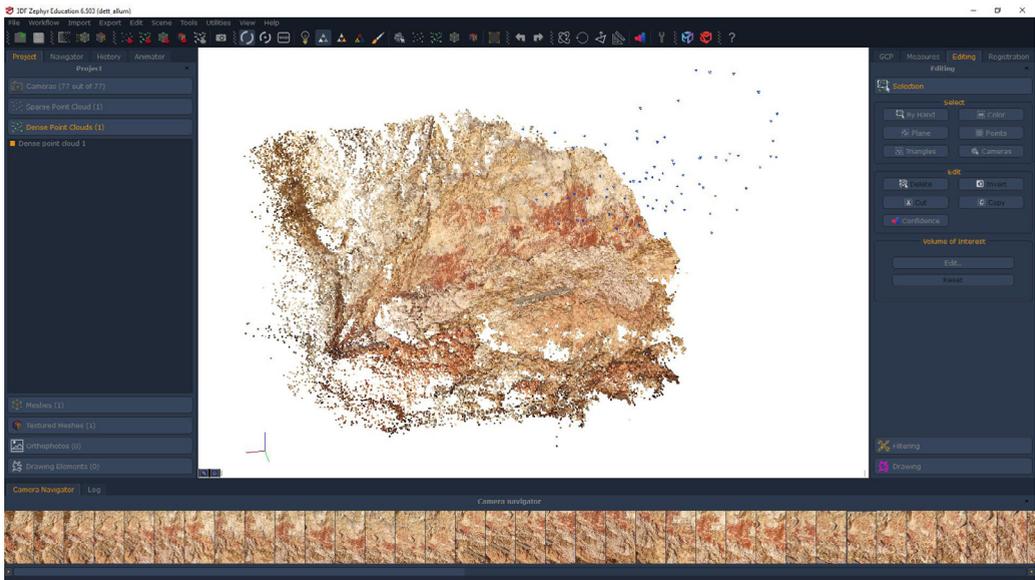
<https://www.agisoft.com/>



*Figure 17. Metashape graphical user interface.*

**3DF Zephyr** is a viable alternative that operates exclusively on the Windows OS. Despite its pro license being more expensive than Metashape, it provides temporary educational licenses for free, making it particularly beneficial for classes with large numbers of students.

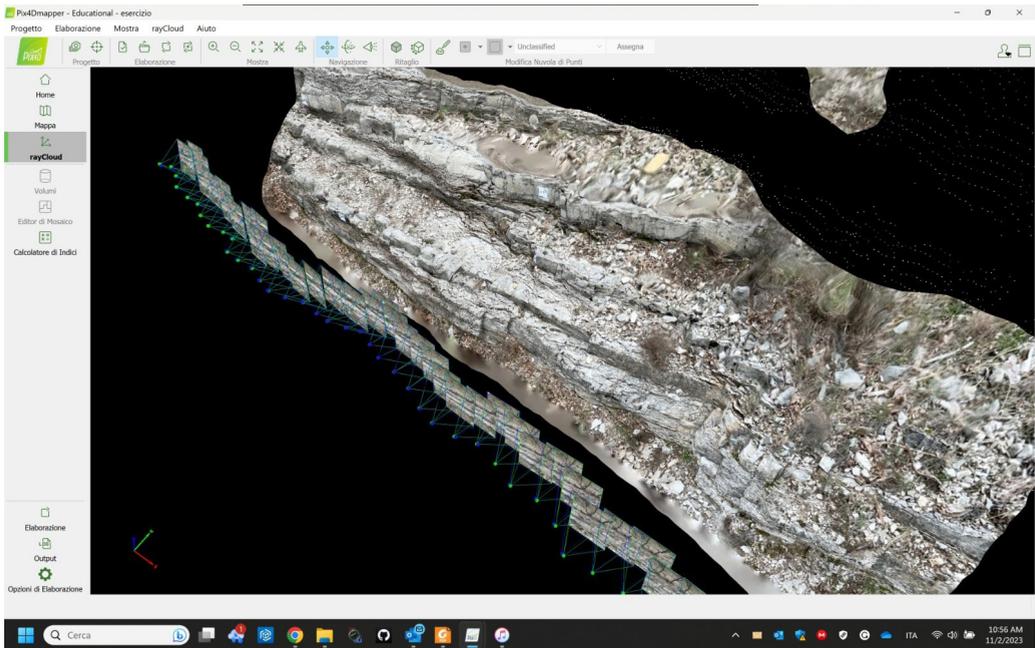
<https://www.3dflow.net/>



*Figure 18. 3DF Zephyr graphical user interface.*

**Pix4D.** This is another viable alternative to Metashape, available only for Windows OS, whose principal advantage is its integration with apps for data capture by means of UAVs and smartphones.

<https://www.pix4d.com/>



*Figure 19. Pix4D graphical user interface.*

**Meshroom.** Developed by Alicevision, it is one of the few free available SfM-MVS software tools. It is compatible with Windows and Linux operating systems, although it carries certain hardware prerequisites. Specifically, the software requires an NVIDIA CUDA-enabled GPU to complete the MVS part of the workflow.

<https://alicevision.org/#meshroom>

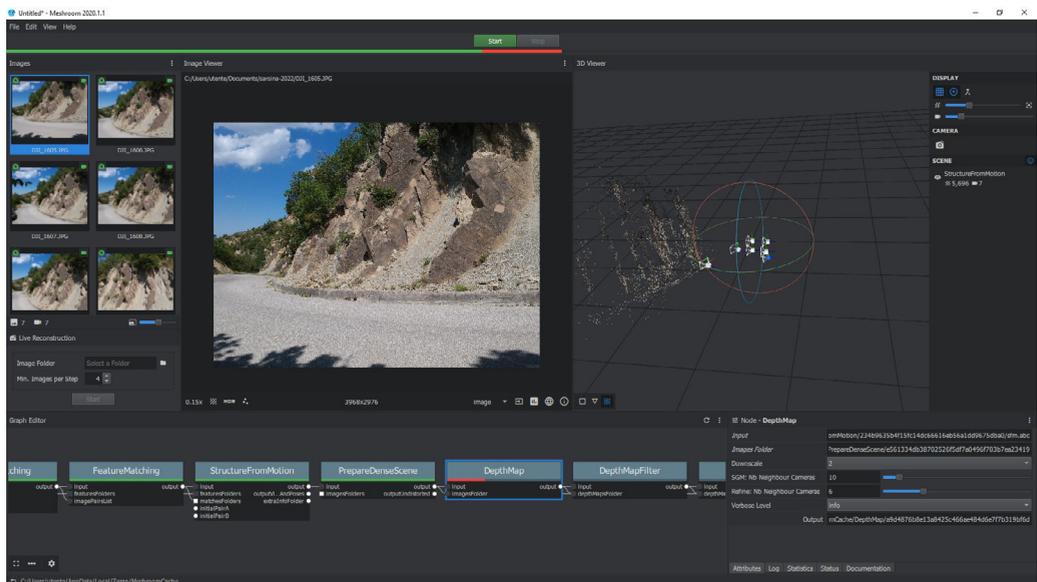


Figure 20. Meshroom graphical user interface.

## Chapter 3

# Data acquisition

Any tool equipped with a camera can be employed for SfM-MVS photogrammetry. However, handheld (DSLR and mirrorless) cameras, smartphones, and UAVs are by far the most commonly deployed image acquisition tools. The selected camera system forms part of a multi-tool setup, which must be considered as a whole when planning and carrying out an image acquisition survey. These setups have advantages and disadvantages that are strictly related to the scene that must be surveyed and the expected results. The best practices associated with each one of these tools are laid out in this chapter, starting with the sensor specifics of each tool.

### 3.1 Know your tools

#### 3.1.1 DSLR and mirrorless cameras



**Figure 21.** Digital Single Lens Reflex camera.

In Digital Single Lens Reflex (DSLR) cameras (Fig. 21), light passes through the lens and hits a mirror, which reflects the light up into an optical viewfinder. This allows the user to see exactly what the camera is pointing at. When taking a photo, the mirror briefly moves out of the way (producing a distinctive sound),

and the light goes directly to the camera sensor. On the other hand, mirrorless cameras do not have mirrors. Instead, light goes straight to the sensor, and the image is displayed on a digital monitor on the back of the camera (Fig. 22).



**Figure 22.** *Mirrorless camera.*

Both DSLR and mirrorless cameras can be used for photogrammetry, and their sensors come in a variety of sizes (with larger sensors generally producing better image quality). These cameras can offer resolutions of up to 50 megapixels. DSLR cameras typically have a larger range of lenses (which allows for a wider range of optical zoom) compared to mirrorless cameras, but they can be heavier and bulkier. Some DSLR and mirrorless cameras are equipped with GNSS

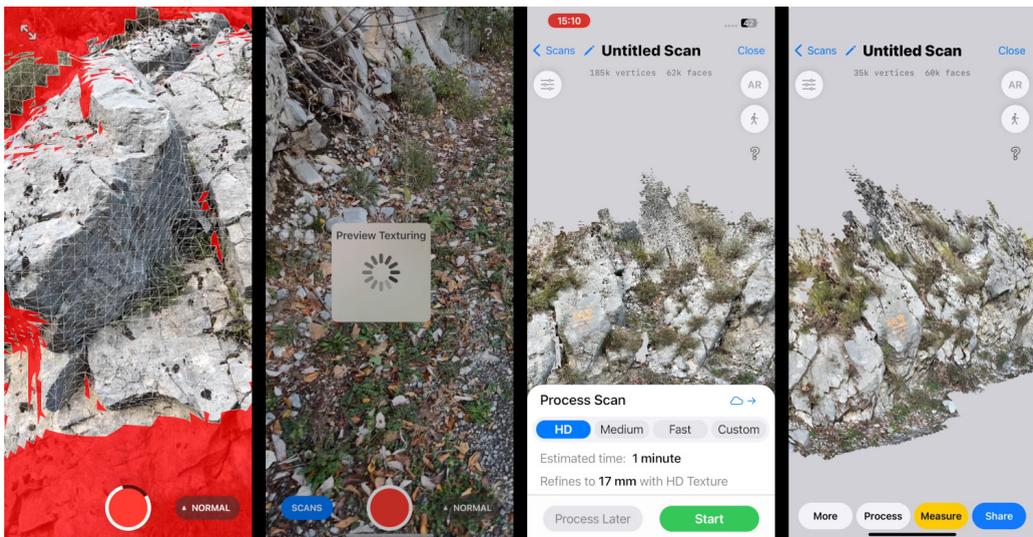
antennas, although accelerometers, gyroscopes, and magnetometers are less common. This means that geotagging (i.e., adding position and orientation data to photos) is not always automatic when using these cameras.

### 3.1.2 Smartphones

Smartphones are equipped with more sensors than cameras, including GNSS receiver, magnetometer, and gyroscope/accelerometer, and some advanced models even feature lidar. These sensors need to be small and light, and smartphone image sensors are generally smaller than those in cameras. Smartphone images typically have a resolution of 12 megapixels. Techniques for sensor manipulation can lead to

images of up to 108 megapixels, but for photogrammetric purposes, the raw resolution of the sensor is what matters. Due to space and weight constraints, smartphones have basic lens systems and limited optical zoom capabilities, with most zoom achieved digitally, resulting in a loss of quality. Some cutting-edge smartphones can achieve up to 5X optical zoom, but this is not comparable to the zoom on DSLR cameras. Thus, smartphones are mainly suited for close-range surveys.

Smartphones are also equipped with an IMU (Inertial Measurement Unit), consisting of inertial sensors such as a magnetometer, gyroscope, and accelerometer. When combined with a GNSS receiver, the IMU provides position and attitude information, theoretically allowing retrieval of all extrinsic camera parameters. However, this data must be handled with caution, as the positioning error of smartphones can be in the order of a few meters, making them unsuitable for registering scenes smaller than tens of meters. The error associated with orientation data is highly variable and depends largely on the smartphone model.



**Figure 23.** Screenshots illustrating the various steps of lidar 3D acquisition with an iPhone 13 Pro, using the 3d Scanner App.

Among the various smartphones, the iPhone Pro deserves a special mention. The iPhone 12 Pro, released in 2020, was the first smartphone equipped with a lidar. Based on the simultaneous emission of 576 beams, its maximum operational range was <5 m under optimal conditions. As of the present year, 2024, only models within the Apple Pro series are equipped with lidar, and there have been no significant improvements to either the operational distance or the number of beams. Available scanning apps for iPhone employing lidar can be grouped into two types: (1) standalone lidar-based apps (Fig. 23), and (2) hybrid photogrammetry-lidar apps. The latter type of application relies on cloud processing of images, usually requiring a subscription, to generate an SfM-MVS model that is scaled and oriented with the assistance of the lidar data. Apps of the first type enable users to perform similar functions without the need for a subscription, but with a less straightforward procedure: The raw data employed to construct the model's texture encompasses images; These images can be uploaded into any SfM-MVS software to produce high-resolution models; The key lies in aligning these high-resolution models with the corresponding lidar model. An indisputable advantage of this approach is that after scanning a section of the outcrop within less than 5 meters using the lidar, additional images can be captured to illuminate more remote portions of the outcrop. These images can later be integrated with the ones acquired during the lidar scanning process to build a large model, registered by aligning it to the lidar model.

### *3.1.3 Uncrewed Aerial Vehicles (UAVs)*

Small or entry-level UAV (Fig. 24) cameras share the same limitations as smartphones, with small image sensors and positioning errors of a few meters. However, larger and higher-quality UAVs can be equipped with cutting-edge cameras and RTK GNSS antennas, enabling centimeter-level positioning accuracy. The advantage of UAVs over smartphones and cameras lies in their ability to acquire aerial imagery in difficult-to-access areas or on vertical cliffs. However, UAVs have a limited battery life of



**Figure 24.** Uncrewed aerial vehicle.

a few tens of minutes, and only lightweight models can really form part of the basic equipment during geological surveying, while larger UAVs require customized survey planning.



**Figure 25.** Gimbal stabilizer for smartphones. This tool ensures that one of the three axes of the image is always oriented horizontally.

### 3.1.4 Gimbal stabilizer

UAVs come equipped with a gimbal stabilizer that ensures the camera's long side (i.e., one of the mutually perpendicular axes that define the image orientation in 3D) is always oriented horizontally, which facilitates model registration and helps identify any distortion in the scene reconstruction. Similarly, inexpensive gimbal stabilizers can be paired with smartphones (Fig. 25) to keep either the long or short side of

the photo horizontal. The benefit of having a horizontal axis in the images is that such information can be used both to independently evaluate the registration of a model and to register it.

### *3.1.5 Global Navigation Satellite System (GNSS) sensors*

Since the 1960s, satellite navigation and positioning systems have been used for military purposes. These systems are called GNSS, or Global Navigation and Satellite System, as they can transmit positioning and timing data signals from space. GNSS receivers installed on devices process this data, received from one or multiple constellations, to determine a location. Currently, there are four main operational constellations, including Global Positioning System (GPS), which has been fully operational since 1995. However, the term GPS is not equivalent to GNSS, as GPS only refers to the US constellation. The GLONASS constellation, provided by Russia, has been fully operational since 2011, while the European Galileo and Chinese BeiDou constellations are the most recent in operation. Other regional GNSS systems include the Indian NavIC and Japanese QZSS. Satellites in each constellation emit signals at different frequency bands, and the more constellations and bands a GNSS receiver can read, the more accurate its positioning will be. Dual-band multi-constellation GNSS receivers installed on consumer smartphones can achieve real-time positional accuracy on the order of a few meters, which can be improved further by post-processing carrier phase fix. GNSS provides altitudes referred to as ellipsoid heights, but these are not the typical altitudes used on maps. To convert ellipsoid heights to orthometric heights, which are commonly used on maps, the geoid-ellipsoid separation is used. Differential GNSS (dGNSS), which uses a reference base station to provide correctional data to a GNSS receiver, can dramatically reduce systematic errors in positioning. Real-time kinematics (RTK-GNSS) refers to the dGNSS technique when space observations are provided by a reference base station in real-time to the rover. Accurate positioning can also be achieved using a single dual-

frequency GNSS receiver through the precise point positioning (PPP) technique in kinematic or static mode, using precise satellite orbits and clock data.

### 3.1.6 Other tools

When it comes to small outcrops, smartphone GNSS receivers are generally not accurate enough for model registration. While dGNSS



**Figure 26.** Laser distance meter. These handheld tools can calculate distances of up to 100 meters with an accuracy of a few millimeters or less.

antennas can provide more precise positioning, they are not typically included in a geologist's standard equipment. However, for small outcrops that only require scaling and orientation, a compass and measuring tools such as a meter tape, a laser distance meter (Fig. 26), or any object of known dimension that is placed in the scene can be sufficient for partial registration. This process becomes even simpler when combined with a smartphone gimbal stabilizer.

## 3.2 Acquisition planning

### 3.2.1 Photo-survey

The golden rule of image acquisition is that, in order to help the keypoint detection and correspondence steps, illumination conditions should be as stable as possible. Outdoor image acquisition should thus be carried out in a restricted time window, ideally with sunshine (keeping your back to the sun) or with a homogeneous cloud cover. In the case of long-lasting acquisitions (several hours), it may be better to split the acquisition into more days but during the same hours. Although the Scale-Invariant Feature Transform object recognition system can work even with large variations in scale, another good practice is to keep the GSD (i.e., the distance on the ground represented by one pixel in an image) within the same order of magnitude. Concerning the GSD and the optimal zoom of images, it should be noted that the resolution of a dense point cloud is generally one order of magnitude higher than the GSD. To obtain a dense point cloud with a resolution of 10 cm, the input photos should have a GSD of less than 1 cm per pixel, keeping in mind that zooming is allowed only with an optical zoom. In practice, using DSLR and mirrorless cameras equipped with a good optical zoom permits capturing images from a certain distance to the outcrop. Conversely, to improve the GSD using UAV or smartphones that do not implement optical zooming, the acquisition must be carried out closer to the outcrop.

Since the entire pipeline of SfM-MVS photogrammetry is based on the recognition of points seen in overlapping images, it is obvious that: (1) if a portion of the outcrop is not seen, it will not be reconstructed; and (2) images must overlap. An overlap between adjacent images of 60% or higher is recommended, and each point should be seen in at least four or five images (the more the better). Photos should be taken both perpendicular and oblique to the target surface. This is because the use of a monotone view direction increases the risk of large-scale folding of the scene, known as the doming effect. For example, given a flat vertical

surface, an ideal terrestrial (i.e., handheld cameras or smartphones) acquisition should include photos taken using view directions both perpendicular to the surface and up to  $60^\circ$  with it. The acquisition of a vertical cliff by using a UAV will require images with both a horizontal view direction and a  $30^\circ$  plunging view direction. Similarly, the surveying of an almost flat landscape by an UAV will be done using a zenithal view and a  $60^\circ$  plunging view. Another golden rule is “one shooting position, one photo”, i.e., the user should avoid shooting photos from the same position but with different view directions and/or zoom values.

### *3.2.2 Precautions for registration*

Registration of the model is the process of aligning the 3D model created from the photos with a coordinate system. It is part of the software processing, but the setup for registration must be designed during photo acquisition. As mentioned in Section 2.3.4, the registration of a model can be done using: (1) GCPs or a mix of GCPs and objects for which the orientation has been measured; (2) the cameras' pose information. In the first case, it is obvious that GCPs must be included in the scene before acquisition, and, like any other point, they must be seen in at least four or five images. Similarly, to use any geometrical feature of known orientation to register the model (e.g., a bedding surface), such an object must become part of the model. GCPs should be less collinear and as distant as possible. It is not rare that when the target scene is a narrow strip, the scanned area must be expanded to avoid collinearity of GCPs. Similarly, when camera position information is used to register the model, collinearity of images must be avoided too. To achieve this purpose, frequently two sets of images are acquired: the first set is aimed at providing a good 3D reconstruction, and the second is instead used to guarantee the occurrence of non-collinear images for the registration step. It is worth mentioning that this second set of images must be correctly positioned in the scene, and thus they must have an overlap of more than 60% with the other photos.

### *3.2.3 Multi-camera acquisition*

Given that camera pose information is available for smartphone and UAV photos, the rough registration of models made using smartphone or UAV photos is, in many applications that do not require topographic accuracy, quite straightforward and does not need any GCPs. On the other hand, the quality of the lenses and sensors of handheld cameras ensures high-quality models. It is possible to combine images acquired with different tools (i.e., smartphones, handheld cameras, and UAVs) in the same reconstruction to take advantage of each tool. However, integration of images acquired with tools having different image sensors (even with two different cameras) may lead to problems in scene reconstruction. In some circumstance, when multiple acquisition devices are available, it may be convenient to use one of the models to register another. In this case, the model with more accurate constraints can be used to register a poorly constrained but much denser model. In detail, for terrestrial acquisitions carried out with smartphones and handheld cameras without GNSS, a good practice is to: (1) Build and register the smartphone model; (2) Build the handheld camera model; and (3) recognize homologous points in both models and pick their coordinates in the smartphone model to register the handheld camera model.

### *3.2.4 Video acquisition*

Using video records for SfM-MVS reconstruction is possible via the extraction of frames. In many software tools, such an operation is automated, while in others, we must use additional software to extract the frames, save them as images, and then import them into the SfM-MVS software. Video acquisition has some attractive advantages but also many drawbacks. The two main advantages are that the acquisition time is strongly reduced and that high overlap among photos is guaranteed. However, there are several problems with video acquisition:

(1) Frames have no camera pose information.

- (2) When using smartphones or UAVs, the resolution and quality of photos are generally higher than the resolution and quality of video frames.
- (3) Blurry frames greatly complicate the SfM reconstruction and can be avoided only if the user ensures a smooth video acquisition, carried out at a low movement velocity and without sudden movements.
- (4) Given the slow velocity required to avoid blurry frames and considering that commonly the frame rate of videos ranges from 20 to 60 frames per second, only a small fraction of frames should be extracted (e.g., from 1 to 4 frames per second) to avoid memory overload, which in any case will not produce a better model. Extracting only a few frames per second also reduces the risk of uploading images with an overlap of nearly 100%, which may induce problems during reconstruction.
- (5) To avoid frames with an overlap of nearly 100%, the user should never stop moving during the acquisition.

All of these problems make video acquisition frequently challenging, and we recommend it only for expert users.



## Chapter 4

### **Applications to geological structures**

After VOMs are created, these can be interpreted, which in the majority of cases consists of, or starts with, digitally collecting the attitudes of geological surfaces like bedding, fractures, faults, etc. Collecting an attitude in a digital 3D space involves finding a function, or a set of functions, that approximates the surface. The level of approximation that we consider acceptable depends on several factors, including the aims and objectives of the geological interpretation as well as the surface's exposure conditions. In the ideal case, the geological surface of interest is fully exposed and the topography of the VOM, or part of it, coincides with the surface of interest. In this case, sophisticated analyses of the surface can be carried out, including, for example, the evaluation of the surface's roughness. This condition of exposure is, however, quite rare. More commonly, only small and unevenly distributed patches of a surface are exposed. Even more commonly, geological surfaces intersect the outcrop's surface at a certain angle, so that only curves corresponding to the intersection between the geological surface and the outcrop surface are exposed. In the latter two cases, if the surface of interest or parts of it display a near-planar character, detecting and measuring the geological surface (or part of it) implies finding the best-fitting plane and the boundary of the area that can be approximated by the best-fit plane, not necessarily in this order.

## 4.1 Detecting and measuring near-planar geological features

### 4.1.1 Background to planar best-fitting

Given a set of 3D points, two methods can be used to derive the best-fit plane: (1) planar regression; and (2) moment of inertia, also called Principal Component Analysis (PCA). Planar regression computes the best-fit plane of a set of points by minimizing the error between the plane and the original points, i.e., by minimizing the difference between the expected and real positions of the points. This method returns the orientation of the best-fit plane and the scalar value of the error. The moment of inertia method, instead, measures the concentration of points about three orthogonal principal axes, the direction of the minimum concentration of points being the normal vector of the best-fit plane. Mathematically, PCA consists in taking the covariance matrix of the considered 3D points and finding its eigenvalues and eigenvectors. The eigenvector associated with the minimum eigenvalue (i.e., the direction of the minimum concentration of points) is the normal vector of the best-fit plane.

Specifically, given a set of  $n$  3D points  $(x_1, y_1, z_1), (x_2, y_2, z_2), \dots, (x_n, y_n, z_n)$ , the covariance matrix  $T$  is:

$$T = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n (x_i - Ax)^2 & \sum_{i=1}^n (x_i - Ax)(y_i - Ay) & \sum_{i=1}^n (x_i - Ax)(z_i - Az) \\ \sum_{i=1}^n (x_i - Ax)(y_i - Ay) & \sum_{i=1}^n (y_i - Ay)^2 & \sum_{i=1}^n (y_i - Ay)(z_i - Az) \\ \sum_{i=1}^n (x_i - Ax)(z_i - Az) & \sum_{i=1}^n (y_i - Ay)(z_i - Az) & \sum_{i=1}^n (z_i - Az)^2 \end{bmatrix} \quad (19)$$

in which  $Ax$ ,  $Ay$ , and  $Az$  are the coordinates of the center of mass of the point set.  $\lambda_1 > \lambda_2 > \lambda_3$  are the eigenvalues of  $T$  divided by the trace of  $T$  (i.e., the sum of the diagonal elements of  $T$ , so that  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ ) and define how much the point set concentrates around the three corresponding eigenvectors ( $\xi_1, \xi_2, \xi_3$ ).

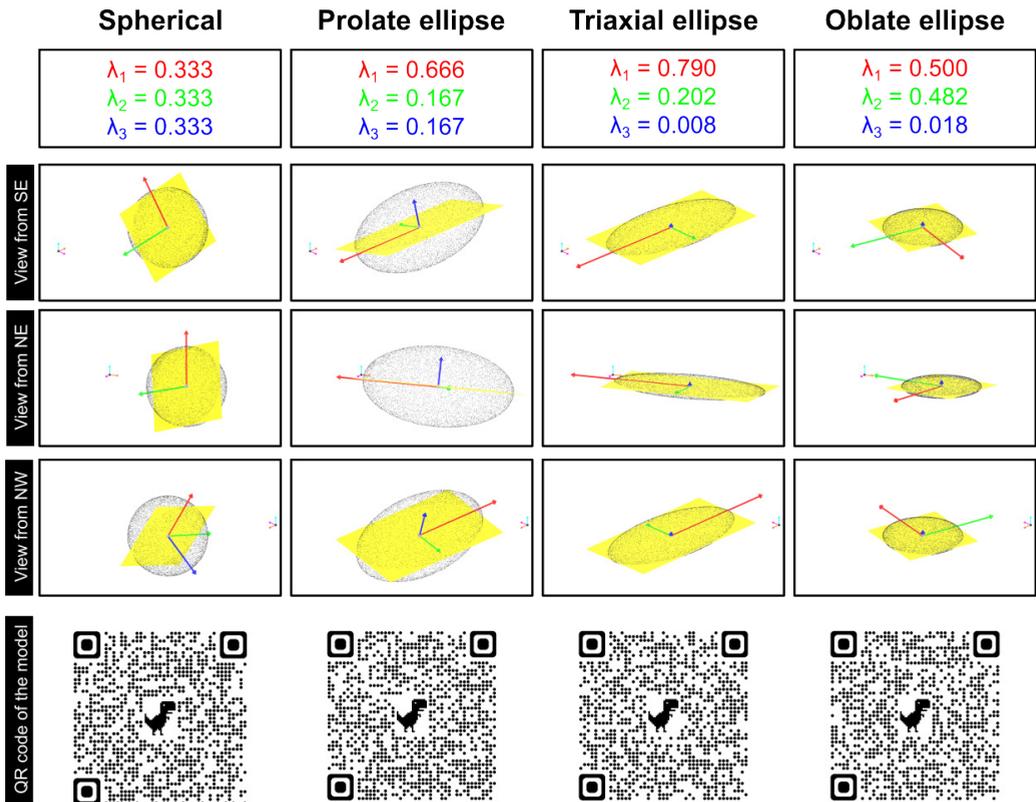
In the majority of cases, planar regression and moment of inertia return similar best-fit planes for the same point distribution. Despite this and despite being less intuitive, the moment of inertia method is preferred over the planar regression method as the latter provides only the total

error, which is a measure of the degree of fit. The vectorial output of the moment of inertia approach, on the other hand, offers estimators of the shape of the point cloud, allowing the user to determine how far it varies from the ideal coplanar and collinear conditions. Specifically, in the moment of inertia analysis the coplanarity and collinearity indexes are defined as:

$$\text{Coplanarity Index} = \frac{\lambda_1 + \lambda_2 + \lambda_3}{\lambda_3} = \frac{1}{\lambda_3} \quad (20)$$

$$\text{Collinearity Index} = \frac{\lambda_1}{\lambda_2} \quad (21)$$

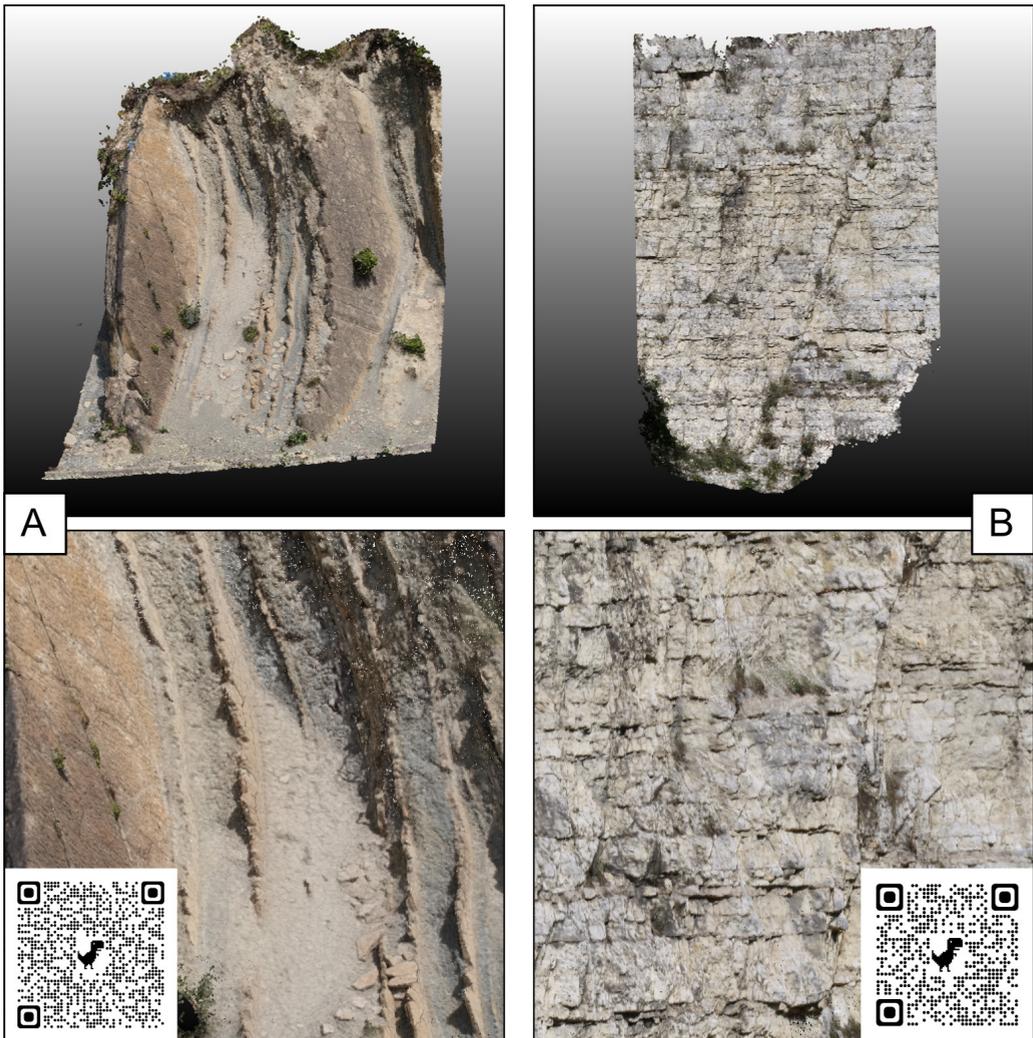
with the coplanarity index ranging from 3 to  $\infty$  and the collinearity index ranging between 1 and  $\infty$ . A good planar fit is one applied to a distribution with high coplanarity (the points are roughly distributed on a planar surface) and low collinearity (the points do not align on a straight line), which means  $\lambda_1 \approx \lambda_2 \gg \lambda_3$ . Examples of the moment of inertia outcome for ellipsoidal distributions are provided in Figure 27. In particular, low values of coplanarity characterize spherical and ellipsoid prolate distributions, for which a planar best fit should not be applied. In both cases, since the best fit plane is the plane normal to the  $\xi_3$  eigenvector associated with the minimum eigenvalue  $\lambda_3$ , the similar values of  $\lambda_2$  and  $\lambda_3$  indicate that the fit plane is not a reliable outcome. In a triaxial ellipsoid distribution, the coplanarity increases, making the application of a planar best fit more reliable. Highly flattened oblate distributions, representing the best case for planar fitting, are characterized by high coplanarity and low collinearity.



**Figure 27.** Moment of inertia analysis applied to four elliptical point clouds.  
 Spherical: <https://skfb.ly/oKznr>; Prolate: <https://skfb.ly/oKzoV>; Triaxial: <https://skfb.ly/oKzpq>; Oblate: <https://skfb.ly/oKzpG>

#### 4.1.2 Model segmentation for planar best-fitting

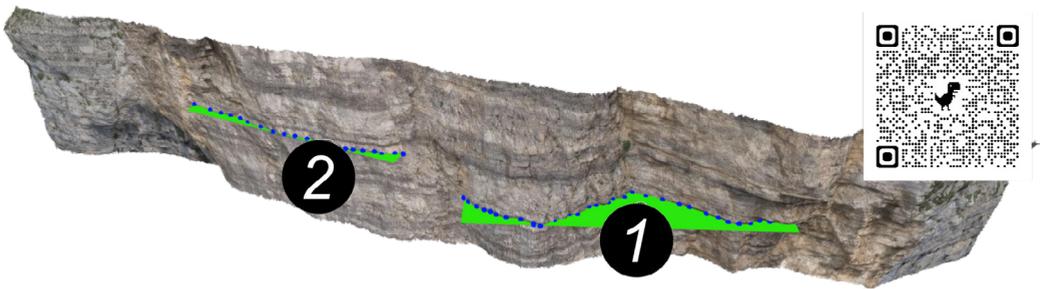
Point cloud and/or mesh segmentation and classification refer to the various methods that group parts of the model based on certain similarities. In our case, these terms refer to the individuation, within the model, of clusters belonging to near planar surfaces. There are two groups of methods for detecting near-planar geological surfaces and deriving them



**Figure 28.** Examples of outcrops displaying: (A) conformable patches <https://skfb.ly/oKAw6>; and (B) Geological traces <https://skfb.ly/oKAzq>

from a 3D model: conform patch analysis and geological trace analysis. Conform patch analysis involves direct model geometry analysis, which is the outcrop's topography, while geological trace analysis is based on tracing features along their intersection with the model geometry (requiring the creation of a subset of points, i.e., the picked points). The choice between the two ways is not a matter of preference, as it is determined by the exposure conditions. To introduce the subject, we are going to analyze the two 3D models in Figure 28. The first is a model of an outcrop characterized by well-exposed, steeply dipping strata. In this case, large parts of the exposure coincide with near-planar geological surfaces, i.e., several patches of the point cloud align with geological surfaces, and we call them conform patches. The orientations of those geological features may be derived through the local best-fitting of the points that form part of the considered geological surface. There are various procedures that can be used for segmenting the model, with variable levels of automation, ranging from manual to fully automated. In the case of manual interpretation, the standard procedure is that of manually selecting an area that the user considers part of a surface, applying a best fit, and eventually evaluating the result (such as the coplanarity) to confirm or discard it. Automatic segmentation and classification involve the a priori definition of the threshold of certain parameters (e.g., minimum number of points, coplanarity, distance between points, and many others). These thresholds are then used to automatically select points that form part of near-planar patches. The second model in Figure 28 is an outcrop characterized by sub-horizontal bedding, near vertical joints, and steeply dipping faults. None of these geological surfaces is imposed over the topography, and they intersect the outcrop at a high angle. If data are extracted based on the local best-fitting of the point cloud distribution, the orientation of the outcrop rather than the geological surface orientation is obtained. In such an outcrop, orientation data needs to be obtained from the analysis of the trace of the intersection between the discontinuity and the topography. In essence, we need to sample this trace along segments to derive the orientation of the plane fitting the sampled trace. At present, this kind of procedure is, in the vast majority of cases, carried out manually, and

includes the picking of points, the derivation of the best fit plane for the picked points, and, eventually, the evaluation of the quality of the best fit. An example of this manual picking is illustrated in Figure 29, showing two best-fit planes derived from point sets with different collinearity indexes. The point set with high collinearity (labelled 2 in Figure 29) results from the selection of a bedding surface intersecting the outcrop along an almost straight segment, while the point set with low collinearity (labelled 1 in Figure 29) results from the selection of a bedding trace that follows a more curved, and therefore more reliable path.



**Figure 29.** Example of picked bedding surfaces. The surface on the right intersects the outcrop along a curvilinear path, whereas that on the left intersects the outcrop along a nearly straight segment. <https://skfb.ly/oMSoU>

#### 4.1.3 Best practices

For both conform patch and geological trace analyses, the major drawbacks of manual interpretation are: (1) the non-systematic character of results; and (2) the long time that may be required for data extraction. The first point is particularly critical and largely relates to the fact that the quality of results strongly depends on the geological skills of the interpreter, particularly in challenging VOMs. This, however, may be a major advantage of manual interpretation (when carried out by geologists) over automatic interpretation. An expert can bring experience in recognizing geological structures that extends beyond the mere evaluation of geometrical

parameters, to include the recognition of geological coherence of the results. On the other hand, once thresholds are set, automatic data extraction is fast and repeatable. There is, however, no guarantee that the outcome of the automated data extraction will yield surfaces that are indeed geological structures. A critical evaluation of any automated outcome is always required, in order to iteratively tune the (geometric) thresholds until a certain geological coherence of results is attained.

Both conform patch and geological trace analyses have specific problems. In the first case, for manual interpretation, the geological expertise of the interpreter is what guarantees the selected patch is actually a geological surface. Despite the fact that best-fitting procedures return statistical parameters such as coplanarity, the quality of results in these cases is mostly evaluated by visual inspection. Generally, when the interpreter is an experienced geologist, the results of manual interpretation of conform patches are reliable, although the procedure may be time-consuming. Automatic extraction is quick in ideal outcrops where the topography is entirely made up of conform patches. In real cases, in which vegetation and debris occur or in which there are oblique cuts, tuning thresholds has the twofold purpose of including geological surfaces and excluding debris and vegetation. By increasing the portion of the model that does not correspond to geological surfaces, filtering out areas that do not correspond to any geological surfaces may become time-consuming. As a rule of thumb, in conform patch analysis, when less than 30% of the exposure coincides with near-planar geological surfaces, the manual interpretation can become more straightforward compared to the automatic one.

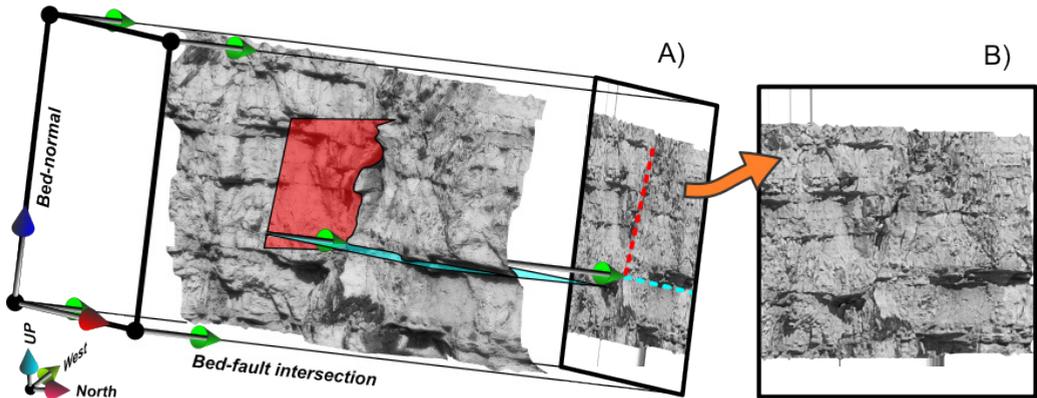
For geological trace analyses, to the best of our knowledge, there is not yet publicly available software for automatic tracing that can be successfully applied to real outcrops. Manual interpretation consists of picking points from a geological trace (to this end, meshes are preferred over point clouds) and applying planar best-fitting to the picked points. In conjunction with visual inspection, the coplanarity and collinearity of the best-fit plane can be used to filter out those planes associated with traces that are too collinear or not coplanar enough. In this sense, major

problems arise when dealing with nearly flat exposures. In these cases, the intrinsic but variable roughness of any geological surface, combined with the fact that traces are collinear along the outcrop direction, makes extracting any best-fit plane with decent statistical coherence arduous.

The last issue that must be considered, for both manual and automatic data extraction and for both conform patch and geological trace analyses, is the resolution of the point cloud or of the mesh. This resolution impacts the minimum size of the near-planar feature that can be extracted. It is quite intuitive, indeed, that given a resolution of 1 cm, any trace shorter than 1 cm or any patch with a radius smaller than 1 cm would return unreliable best-fit planes. Our experience is that the extraction of data should be limited to traces and patches that are two orders of magnitude larger than the resolution of the mesh or of the point cloud.

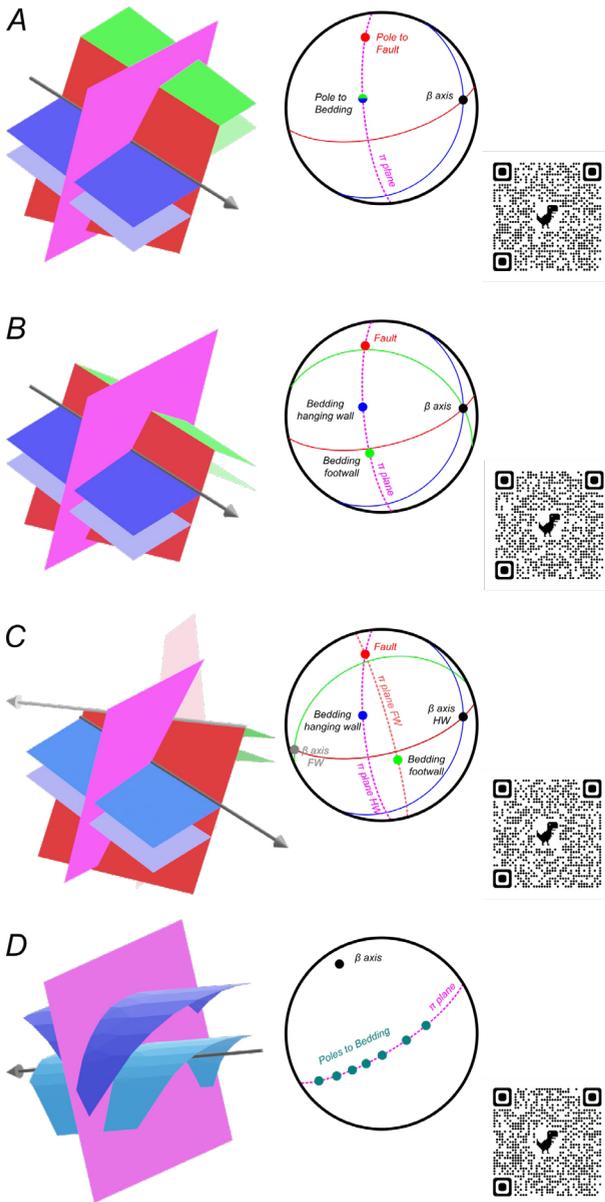
## **4.2 Orthorectifying VOMs**

One of the most important applications of virtual outcrop representations for structural geology and geosciences is the ability to view a photo-realistic virtual outcrop in orthographic projection mode, using any view direction. This capability enables the user to virtually orthorectify the outcrop with respect to any direction of geological interest (e.g., fold axis, fault slip-normal direction, fracture-bedding intersection) to obtain true measurements of geological features, such as bed thickness and fault displacement for optimally oriented and undistorted photo-realistic cross-sections. The procedure is rather intuitive and straightforward when carried out in a qualitative way, as it just involves rotating the model to attain a visually satisfying orientation, which is to say projecting all the points of the VOM onto a plane perpendicular to the view direction. The accurate quantification of the proper view direction (and of its perpendicular plane) instead requires data extraction (following the procedures described in section 4.1) and analysis (Fig. 30).



**Figure 30.** Orthorectification of a Virtual Outcrop Model (VOM) along a direction of interest. This procedure essentially consists of two steps. (A) Geological features are digitized in the model, to compute the viewing direction of interest, which in this image is the intersection between bedding (blue) and the fault (red). (B) Desired viewpoint shown as orthographic projection.

In any case, the definition of a geologically correct view direction for the orthorectification may not always be possible. The procedure for determining the view direction for orthorectification resembles that for the derivation of projection direction for the construction of geological cross sections, but there is a major limitation that must be considered. The possibility of orthorectifying a VOM along a geologically meaningful view direction is dependent on whether such a direction exists and if it is valid for the entire geological structure contained in the VOM under consideration. Let us consider the case of a multilayer affected by a fault oriented obliquely to the strata, with the strata in the hanging wall and footwall being parallel to each other. The stereoplot in Figure 31a corresponds to such a situation. A plane that contains the poles of the fault and of strata does exist (we call it  $\pi$ ), and the direction perpendicular to  $\pi$  (we call it  $\beta$ ) coincides with the intersection between the plane of fault and strata. If we want to build a geological cross section for this structure, we will use  $\beta$  for projecting data onto the section plane.  $\beta$  is also the proper



**Figure 31.** Examples that demonstrate the determination of the proper view direction ( $\beta$  axis) and the corresponding  $\pi$  plane from stereonet. (A) <https://skfb.ly/oNFH8>; (B) <https://skfb.ly/oNFH9>; (C) <https://skfb.ly/oNFHB>; (D) <https://skfb.ly/oNFHE>

view direction for building an orthorectified model. Now we consider the case in which hanging wall and footwall strata are oblique, but a  $\pi$  plane containing the poles of the fault and of the strata persists (Fig. 31b). Also in this case,  $\beta$  coincides with the intersection between the plane of fault and strata. In this second case, if we want to build a geological cross section or orthorectify a VOM, we will still use  $\beta$ . In the third case (Fig. 31c), there are two distinct  $\pi$  and two  $\beta$ , one for the hanging wall-fault intersection, and one for the footwall-fault intersection. This configuration is not at all a problem when constructing a cross section, as we will project data from the hanging wall and footwall using the two distinct  $\beta$  directions. Instead, this is not possible for the orthorectification of the model, as all the points of the model must be projected at once, using a single direction. A similar situation occurs in folds that deviate

from a cylindrical shape. In cylindrical folds, indeed, all the bedding poles are contained along a single  $\pi$  plane, and thus a single  $\beta$  direction exists that can be used for orthorectification of a VOM containing a cylindrical fold (Fig. 31d). In any other case of fold (e.g., conical fold), since a single  $\beta$  axis does not exist, models cannot be orthorectified.

## SUGGESTED READINGS

### Mathematics background

Allmendinger, R.W., Cardozo, N., and Fisher, D.M., 2011, Structural Geology Algorithms: Cambridge, Cambridge University Press, <https://doi.org/10.1017/CBO9780511920202>

Woodcock, N.H., 1977, Specification of fabric shapes using an eigenvalue method: Geological Society of America Bulletin, v. 88, p. 1231, [https://doi.org/10.1130/0016-7606\(1977\)88<1231:SOFSUA>2.0.CO;2](https://doi.org/10.1130/0016-7606(1977)88<1231:SOFSUA>2.0.CO;2)

### Data extraction

Battulwar, R., Zare-Naghadehi, M., Emami, E., and Sattarvand, J., 2021, A state-of-the-art review of automated extraction of rock mass discontinuity characteristics using three-dimensional surface models: Journal of Rock Mechanics and Geotechnical Engineering, v. 13, p. 920–936, <https://doi.org/10.1016/j.jrmge.2021.01.008>

Cawood, A.J., Bond, C.E., Howell, J.A., Butler, R.W.H., and Totake, Y., 2017, LiDAR, UAV or compass-clinometer? Accuracy, coverage and the effects on structural models: Journal of Structural Geology, v. 98, p. 67–82, <https://doi.org/10.1016/j.jsg.2017.04.004>

Fernández, O., 2005, Obtaining a best fitting plane through 3D georeferenced data: Journal of Structural Geology, v. 27, p. 855–858, <https://doi.org/10.1016/j.jsg.2004.12.004>

- Seers, T.D., and Hodgetts, D., 2016, Extraction of three-dimensional fracture trace maps from calibrated image sequences: *Geosphere*, v. 12, p. 1323–1340, <https://doi.org/10.1130/GES01276.1>
- Seers, T.D., Sheharyar, A., Tavani, S., and Corradetti, A., 2022, Virtual outcrop geology comes of age: The application of consumer-grade virtual reality hardware and software to digital outcrop data analysis: *Computers & Geosciences*, v. 159, p. 105006, <https://doi.org/10.1016/j.cageo.2021.105006>
- Thiele, S. T., Grose, L., Samsu, A., Micklethwaite, S., Vollgger, S. A., & Cruden, A. R. (2017). Rapid, semi-automatic fracture and contact mapping for point clouds, images and geophysical data. *Solid Earth*, 8(6), 1241-1253. <https://doi.org/10.5194/se-8-1241-2017>
- Thiele, S.T., Grose, L., Cui, T., and Cruden, A.R., 2019, Extraction of high-resolution structural orientations from digital data : A Bayesian approach: *Journal of Structural Geology*, <https://doi.org/10.1016/J.JSG.2019.03.001>.
- Umili, G., Ferrero, A., and Einstein, H.H., 2013. A new method for automatic discontinuity traces sampling on rock mass 3D model. *Computers & Geosciences*, v. 51, p.182-192. <https://doi.org/10.1016/j.cageo.2012.07.026>
- Wang, X., Zou, L., Shen, X., Ren, Y., & Qin, Y., 2017. A region-growing approach for automatic outcrop fracture extraction from a three-dimensional point cloud. *Computers & geosciences*, v. 99, p. 100-106. <https://doi.org/10.1016/j.cageo.2016.11.002>

### **Applications in structural geology**

- Bistacchi, A., Griffith, W.A., Smith, S.A., Di Toro, G., Jones, R., and Nielsen, S., 2011. Fault roughness at seismogenic depths from LIDAR and photogrammetric analysis. *Pure Appl. Geophys.*, v. 168, p. 2345-2363, <https://doi.org/10.1007/s00024-011-0301-7>
- Ceccato, A., Viola, G., Antonellini, M., Tartaglia, G., & Ryan, E. J., 2021. Constraints upon fault zone properties by combined structural analysis of virtual outcrop models and discrete fracture network modelling. *Journal of Structural Geology*, 152, 104444. <https://doi.org/10.1016/j.jsg.2021.104444>
- Camanni, G., Vinci, F., Tavani, S., Ferrandino, V., Mazzoli, S., Corradetti, A., Parente, M., and Iannace, A., 2021, Fracture density variations within a reservoir-

scale normal fault zone: a case study from shallow-water carbonates of southern Italy: *Journal of Structural Geology*, v. 151, p. 104432, <https://doi.org/10.1016/j.jsg.2021.104432>

Cawood, A.J., Corradetti, A., Granado, P., and Tavani, S., 2022, Detailed structural analysis of digital outcrops: A learning example from the Kermanshah-Qulqula radiolarite basin, Zagros Belt, Iran: *Journal of Structural Geology*, v. 154, p. 104489, <https://doi.org/10.1016/j.jsg.2021.104489>

Corradetti, A., Tavani, S., Parente, M., Iannace, A., Vinci, F., Pirmez, C., Torrieri, S., Giorgioni, M., Pignalosa, A., and Mazzoli, S., 2018, Distribution and arrest of vertical through-going joints in a seismic-scale carbonate platform exposure (Sorrento peninsula, Italy): insights from integrating field survey and digital outcrop model: *Journal of Structural Geology*, v. 108, p. 121–136, <https://doi.org/10.1016/j.jsg.2017.09.009>

Fleming, Z. and Pavlis, T., 2018. An orientation based correction method for SfM-MVS point clouds - Implications for field geology. *Journal of Structural Geology*, v. 113, p. 76-89, <https://doi.org/10.1016/j.jsg.2018.05.014>

Menegoni, N., Giordan, D., Perotti, C., and Dwayne, D., 2019, Detection and geometric characterization of rock mass discontinuities using a 3D high-resolution digital outcrop model generated from RPAS imagery - Ormea rock slope, Italy: *Engineering Geology*, v. 252, p. 145–163, <https://doi.org/10.1016/j.enggeo.2019.02.028>

Tavani, S., Corradetti, A., and Billi, A., 2016, High precision analysis of an embryonic extensional fault-related fold using 3D orthorectified virtual outcrops: The viewpoint importance in structural geology: *Journal of Structural Geology*, v. 86, p. 200–210, <https://doi.org/10.1016/j.jsg.2016.03.009>.

## SUGGESTED SOFTWARE TOOLS

Among the standalone software tools allowing the user to carry out basic structural analysis, we suggest the following:

**CloudCompare.** <https://www.danielgm.net/cc/>

**Lime.** <https://virtualoutcrop.com/lime>

**VRGS.** <https://www.vrgeoscience.com/>

**OpenPlot.** <http://www.openplot.altervista.org/>

## Chapter 5

### **SfM-MVS in practice**

In this chapter, we illustrate the acquisition and processing of six case studies, each presenting an increasing level of complexity. The SfM-MVS exercises are conducted using version 2.02 of Agisoft Metashape. For specific tasks, we utilize CloudCompare and OpenPlot software tools.

All the necessary materials, including the input images or video, the Metashape project, and the output textured meshes, have been provided in the Zenodo external repository (<https://zenodo.org/>).

Below, we provide a description of the materials and objectives of each exercise. The step-by-step processes for each exercise are detailed in 5 to 10-minute YouTube videos, accessible through the QR codes and links associated with each exercise.

It is recommended to view the videos at least twice. The first time should be uninterrupted, and the second time with pauses to faithfully replicate all the operations.

# Exercise 1: Terrestrial DSLR

This is “my first model” exercise, aimed at building a model of a 10-meter-wide outcrop using terrestrial acquisition and registering it using GCPs.

## Objectives

Illustrating the software's interface and explaining the basic steps and parameter settings for: (i) photo alignment; (ii) dense cloud building; (iii) meshing; (iv) mesh texturing, (v) model registration via GCPs

## Material

The dataset consists of 72 images of unknown position and orientation taken from ground level using two reflex cameras and a text file labeled marker.txt containing the coordinates, in a local reference frame, of ground control points included in the scene that are used to register the model. 28 photos were taken with a Sony DSC-W530 (14 MP = 4320 x 3240 pixels), and 42 photos were taken with a Canon EOS 450D (12 MP = 4272 x 2842 pixels).



**Figure 32.** View of the 3D model used in exercise 1. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/orbVnukfULk?feature=shared>

Material: <https://doi.org/10.5281/zenodo.10148701>

## Exercise 2: UAV model

This exercise is aimed at building a model of a 50-meter-wide outcrop using aerial acquisition and registering it using the camera position information. The outcrop consists of a Z-fold involving Jurassic limestones near San Potito village, Italy.

### Objectives

Explaining the basic steps and parameter settings for: (i) model's registration via photos' position; (ii) orthomosaic building; (iii) fast testing of the registration quality

### Material

The dataset consists of 41 images of known position and unknown orientation taken using a DJI Mini 2 UAV.



**Figure 33.** View of the 3D model used in exercise 2. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/8kOUEbxznt0?feature=shared>

Material: <https://zenodo.org/doi/10.5281/zenodo.10148744>

## Exercise 3: Model from Video

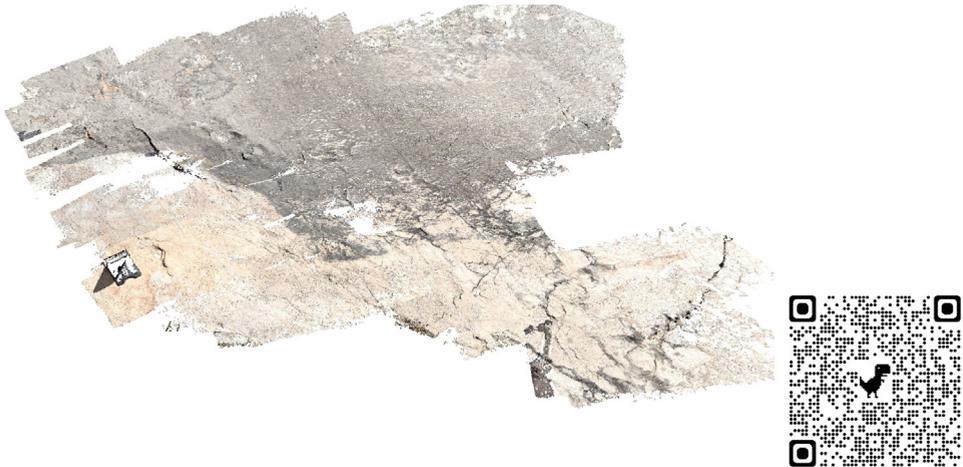
This is the “my first fail” model exercise. It is aimed at showing the problems associated with the building of a model from frames extracted from a video. The outcrop consists of a fault ribbon located near the Castelluccio Village, Italy.

### Objectives

Explaining the basic steps and parameter settings for building a model from video frames. Illustrate the main problems with this procedure. Understanding when it is time to give up

### Material

A 5-minute video is provided, with a resolution of 3840 x 2160 pixels and a video bit rate of 42025 kbps. The video was acquired using a Xiaomi 9T Pro smartphone in association with a gimbal stabilizer.



**Figure 34.** View of the 3D model used in exercise 3. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/ZlpJ3x2vWzw?feature=shared>

Material: <https://zenodo.org/doi/10.5281/zenodo.10148779>

## Exercise 4: Smartphone + Gimbal stabilizer

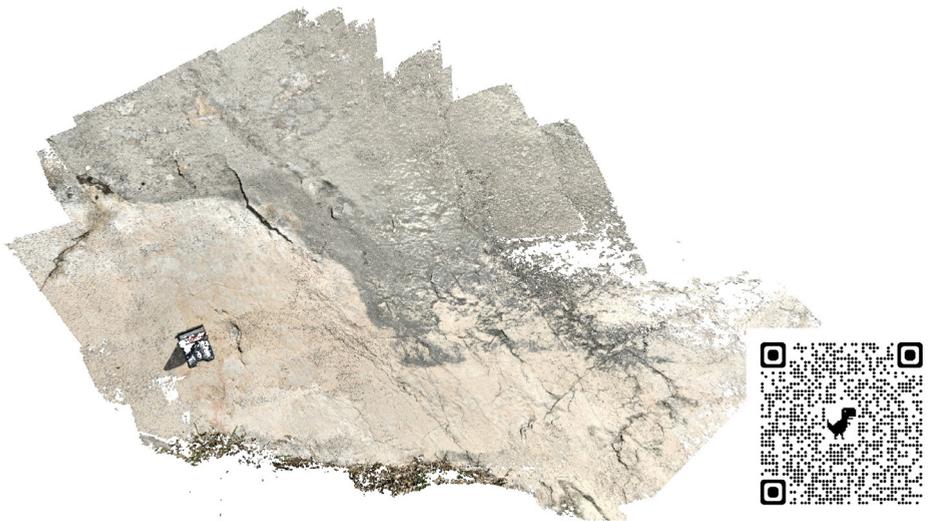
This exercise is aimed at building a model of a 2-meter-wide outcrop consisting of a fault ribbon located near Castelluccio Village, Italy (the same outcrop as Exercise 3), using images acquired by using a smartphone in conjunction with a gimbal stabilizer.

### Objectives

Explaining the steps for rotating and scaling a model using the image orientation and the dimension and orientation of an object placed in the scene.

### Material

The dataset consists of 142 images, with a resolution of 4000 x 2250 pixels, of unknown position and partly unknown orientation, taken using a Xiaomi 9T Pro smartphone in association with a DJI OSMO Mobile 2 gimbal stabilizer.



**Figure 35.** View of the 3D model used in exercise 4. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/TUbddjlO1s8?feature=shared>

Material: <https://zenodo.org/doi/10.5281/zenodo.10148912>

## Exercise 5: Digitization and orthorectification

This exercise is aimed at orthorectifying a model of a 100 meter-wide outcrop consisting of a syncline-anticline pair located near Passo Rolle, Italy.

### Objectives

Explaining the steps for digitizing bedding surfaces using different approaches (manual and automatic), evaluating the cylindricity of the structure, defining the view direction, and eventually orthorectifying the model.

### Material

The dataset consists of 266 images of known position and with a resolution of 5472 x 3648 pixels, taken using a DJI Air 2S UAV.



**Figure 36.** View of the 3D model used in exercise 5. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/RzSq26k3iYA?feature=shared>

Material: <https://zenodo.org/doi/10.5281/zenodo.10148914>

## Exercise 6: Digitization of traces and the effect of outcrop orientation

This exercise is aimed at digitizing bedding surfaces along a 10 m-wide outcrop, exposing homogeneously dipping limestone beds along a near vertical cliff close to the Vallepietra Village, Italy

### Objectives

Illustrating the procedures and issues related to the digitization of traces for planar best fitting, with particular attention to the effect of the outcrop's orientation on the orientation of traces.

### Material

The dataset consists of 61 images of known positions with a resolution of 4000 x 2225 pixels, taken using a DJI Mini 2 UAV.



**Figure 37.** View of the 3D model used in exercise 6. The QR code allows access to the YouTube video of this exercise.

Video: <https://youtu.be/wmaXX2kFXo?feature=shared>

Material: <https://zenodo.org/doi/10.5281/zenodo.10148924>

## Acknowledgments

**Stefano Tavani.** *First and foremost, I would like to express my gratitude to Josep Anton Muñoz, Pau Arbues, and Oscar Fernandez, from the University of Barcelona, for introducing me to the world of 3D digital geology in the early 2000s. I also want to extend my thanks to Francesco Vinci and Pablo Granado, who contributed to testing and improving software and workflows presented in this book. Antonio Piglanlosa, from Stage S.r.l. is acknowledged for providing various high resolution 3D models since 2014, including the model featured on the front page of this book. Lastly, I wish to thank the president of the Geological Society of Italy, Sandro Conticelli, for fostering and encouraging the publication of this book.*

**Amerigo Corradetti.** *I am indebted to numerous individuals who have played a crucial role in guiding me into the world of virtual outcrop geology. I will not repeat those expressions of gratitude already expressed by Stefano and that I share. Nevertheless, while the list of individuals deserving acknowledgment is extensive, I will limit myself to expressing gratitude to the coauthors of this work and to Ken McCaffrey at Durham University.*

**Marco Mercuri.** *I would like to thank everyone who supported me through my exploration of virtual outcrop models. Prof. Carminati and Collettini from Sapienza University, my PhD supervisors, deserve my deepest gratitude for their invaluable support. Special appreciation to Ken McCaffrey of Durham University for guiding my understanding of virtual outcrop interpretations, and to Paolo Mazzanti of Sapienza University for introducing me to diverse construction techniques.*

**Thomas Seers.** *Over the past decade and a half, I have had the privilege of witnessing numerous incredible hardware, software and methodological developments that have contributed to the burgeoning field of virtual outcrop geology, making it one of the most exciting subdomains within the Earth Sciences. I would like to acknowledge Dr David Hodgetts of VRGeosciences Limited, who was instrumental to my entry into this field and represents a true pioneer of virtual outcrop data analysis.*



Finito di stampare nel mese di gennaio 2024  
Pixartprinting S.p.A. stabilimento di Lavis (TN)



# Virtual outcrop models of geological structures

## From the construction of photogrammetric 3D models to their application towards the analysis of geological structures

This textbook aims to introduce the reader to the basic principles and methods of virtual outcrop models and their applications within structural geology. We first provide the reader with the basic concepts of virtual outcrop models and an overview of methods commonly employed for digitally scanning outcrop surfaces: lidar and Structure from Motion - Multi-View Stereo (SfM-MVS) photogrammetry. We then focus upon SfM-MVS photogrammetry, illustrating its fundamental concepts. Finally, we illustrate key applications of virtual outcrop models in the field of structural geology. In addition, the book features six training exercises that serve to illustrate the fundamental steps and issues related to the creation and interpretation of models.

**STEFANO TAVANI** is Professor in structural geology at the Department of Earth Sciences, Environment, and Resources at Federico II University of Naples and Associate Professor at the National Research Council of Italy, IGAG. He teaches undergraduate and graduate courses on virtual outcrop models, field mapping, and structural geology. His research expertise is in virtual outcrops, structural geology, tectonics, and geodynamics.

**AMERIGO CORRADETTI** is an Assistant Professor in structural geology at the Department of Mathematics, Informatics and Geosciences of the University of Trieste. His research includes the study of compressional tectonic settings of the Mediterranean and the Middle East region, faults and fractures evolution, distribution and analysis, and the characterization of reservoir analogues. Amerigo's studies have often taken advantage of digital photogrammetry and virtual outcrop modeling.

**MARCO MERCURI** is a post-doctoral researcher in structural geology at the Department of Earth Sciences of the Sapienza University of Rome. His research focuses on the study of brittle deformation structures in carbonate reservoir analogues, with particular focus on fault structure and mechanics. Marco's studies are performed by integrating classical fieldwork and the interrogation of virtual outcrop models.

**THOMAS SEERS** is an Associate Professor within the Petroleum Engineering Program at Texas A&M University at Qatar (TAMUQ). Dr. Seers teaches graduate and undergraduate courses in geology, petroleum engineering and interdisciplinary engineering. His research focuses upon acquisition, processing and analysis of close-range 3D remote sensing data for geologic applications, including structural geology and reservoir characterization.

