Tampere University of Applied Sciences

Design of a Photogrammetry Pipeline

Generating 3D Models for Real-time Engines

Manuel Andres Gabarrón Cortés

BACHELOR'S THESIS
March 2021

Media and Arts
Interactive Media

ABSTRACT

Tampereen ammattikorkeakoulu
Tampere University of Applied Sciences
Degree Programme in Media and Arts
Interactive Media

GABARRÓN CORTÉS, MANUEL ANDRES:
Design of a Photogrammetry Pipeline
Generating 3D models for real-time engines

Bachelor's thesis 80 pages
March 2021

The objective of this thesis was to design a reliable pipeline to generate 3D models through photogrammetry that fulfilled the demands and addressed the limitations of current real-time engines. This meant not only to study the current documentation and define the best practices to approach photogrammetry projects, but also to experiment to assess how robust the process was.

As part of this experimentation process, three photogrammetry programs were assessed and compared. Besides the mentioned comparison, a thorough method for optimization of the generated 3D maodels was demonstrated and explained in detail.

The results of the comparison of the photogrammetry programs showed that RealityCapture, one of these programs, consistently provided better results over Agisoft Metashape and Meshroom. The described method for optimization successfully provided a 3D model that fulfilled the requirements of current real-time engines.

Photogrammetry proved to be a fast and reliable method for generating 3D models. The process showed advantages in quality and speed over more traditional methods for creating 3D models.

Key words: photogrammetry, pipeline, optimization, real-time graphics, game art

CONTENTS

ABBREVIATIONS AND TERMS

Digitalization	the process of creating a digital representation of a real-world object
3D model	a digital representation of a geometrical shape in three-dimensional space. Constructed of smaller elements called polygons
Textures	2D images used to increase the detail and realism of 3D models
Materials	a set of 2D images or properties that determine the visual aspect of a 3D model
Point cloud	a visual representation of data points in three-dimensional space
DSLR	digital single-lens reflex camera
Real-time engine	a program that visualizes 3D models and other graphics in real-time
Polycount	the number of polygons that form a 3D model
Polygon	a group of connected vertices that form a plane
Exposure	the amount of light that reaches the sensor of the camera, providing a brighter or darker image
ISO	an electronic signal that brightens a photograph but generates digital noise
Digital noise	color distortions on the pixels of an image caused by an electronic signal
Shutter Speed	a camera setting that controls how long the camera sensor receives light
Aperture	a camera setting that controls how much light a lens allows into the sensor
Motion blur	distortions on the pixels of a photograph caused by the movement of the camera while the sensor was capturing the image
Depth of field	a term that defines the area where elements are in focus in a photograph

Focal length	distance between the optical centre of the lens and the sensor of the camera. Used to describe camera lenses
Prime lenses	lenses without optical zoom
RAW format	lossless file format
Workflow	a series of tasks performed in a specific order
Light box	a device used to diffuse the light that crosses its surface
Occluder	a device used to project shadows
Chromatic aberration	color distortions in an image, caused by imperfections on a camera lens
Metadata	information stored in an image
Open-source software	programs which code can be edited changed by any user
CPU	central processing unit
GPU	graphic processing unit
RAM memory	random-access memory
High poly model	a 3D model with a high number of polygons
Low poly model	a 3D model with a relatively low number of polygons
Topology	term used to describe the structure and arrangement of polygons in a 3D model
Polygroups	Zbrush term used to refer to a group of polygon
UV coordinates	coordinates used by 3D programs to map 2D textures on 3D models
Normal map	a 2D images that simulates surface detail on a 3D model
Albedo map	a 2D images that provides color information to a 3D model
Ambient occlusion map	a 2D image that simulates small shadows based on the geometry of a 3D model
Roughness map	a 2D image that defines how the surface of the 3D model diffuses light.

1 INTRODUCTION

Photogrammetry is a technique that provides accurate measurements between objects by comparing photographs. This technique allows, among other things, the digitalization of real-world objects by generating an accurate representation in the form of a 3D model.

The use of photogrammetry has gained traction in the videogame and film industry as a faster approach to creating 3D models. Traditionally, creating realistic 3D models was a long process that required specialized artists and countless hour of work. By capturing real world objects using photogrammetry, realistic 3D models are produced in a much shorter time. This process can be partially or totally automated.

For instance, in a traditional production pipeline, a 3D artist would spend several hours sculpting rocks and creating the textures and materials to make them look believable. By using photogrammetry, teams can be sent to capture rocks from the real world while 3D artists work on models that are more unique and relevant for the final product. Furthermore, gathering photographs to generate models through photogrammetry is a relatively simple process that does not require extensive training.

The aim of this thesis is to provide practical knowledge on how to approach photogrammetry projects. The thesis focuses on three main parts: capturing the images of the subject, processing them with specialized software and demonstrating techniques to optimize the final 3D model. A discussion of the different methods shown will be presented at the end. Aerial photogrammetry is beyond the scope of this thesis.

2 PHOTOGRAMMETRY IN THE CURRENT DAY

2.1 What is photogrammetry?

“Photogrammetry is the art, science, and technology of obtaining reliable information about physical objects and the environment through processes of recording, measuring, and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena.” (Aber, Marzolff, Ries, Ward Aber 2019, 19).

Photogrammetry is about measuring distances. The goal of specialized programs used in photogrammetry is to compare images to find information about the scene that was photographed. The program finds points that are relatively invariable across different images. Through a process of triangulation, the software places these points in 3D space (FIGURE 1).

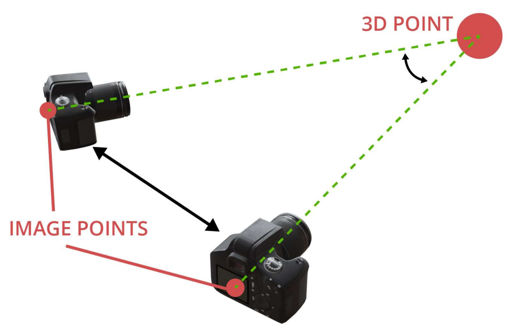
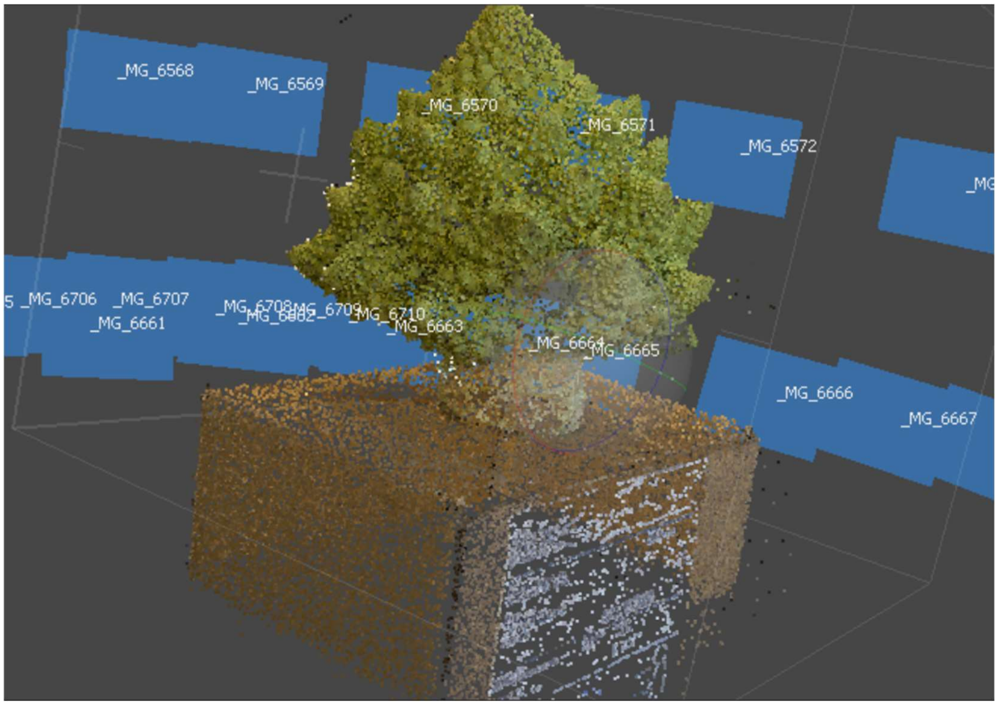


FIGURE 1. The program finds points in 3D space through a process of triangulation.

Once the program has finished analysing all the photographs it will display a point cloud, a visual representation of all the points in 3D space. If the process has been successful, the point cloud will resemble the captured scene from the real world (PICTURE 1).



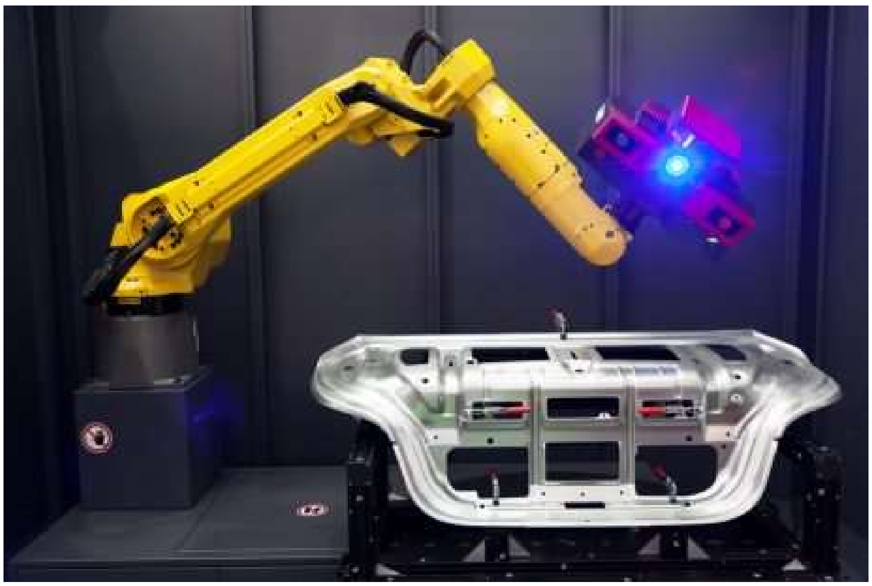
PICTURE 1. Visual representation of a point cloud.

2.2 Uses of photogrammetry

Photogrammetry has many uses across different fields. It has been proven to be an effective and accurate way of preserving monuments and historical artefacts. Generating digital copies of historical items and relics protects their cultural value as the copy will not be affected by time or even the destruction of the original, as has happened during wars (PICTURE 2).

PICTURE 2. Photogrammetry model of The Man of Galera, the second oldest mummy (AD&D 4D 2020)

Different industries have shown interest in the use of photogrammetry due to its versatility and accuracy providing geospatial data. It is used in engineering projects to assess parts and systems, aiding in component manufacturing and quality control. An example of this is the use of photogrammetry to assess mechanical parts manufactured in the production line. A point cloud generated through photogrammetry aids engineers on identifying possible defects in the part (Garnett 2018).



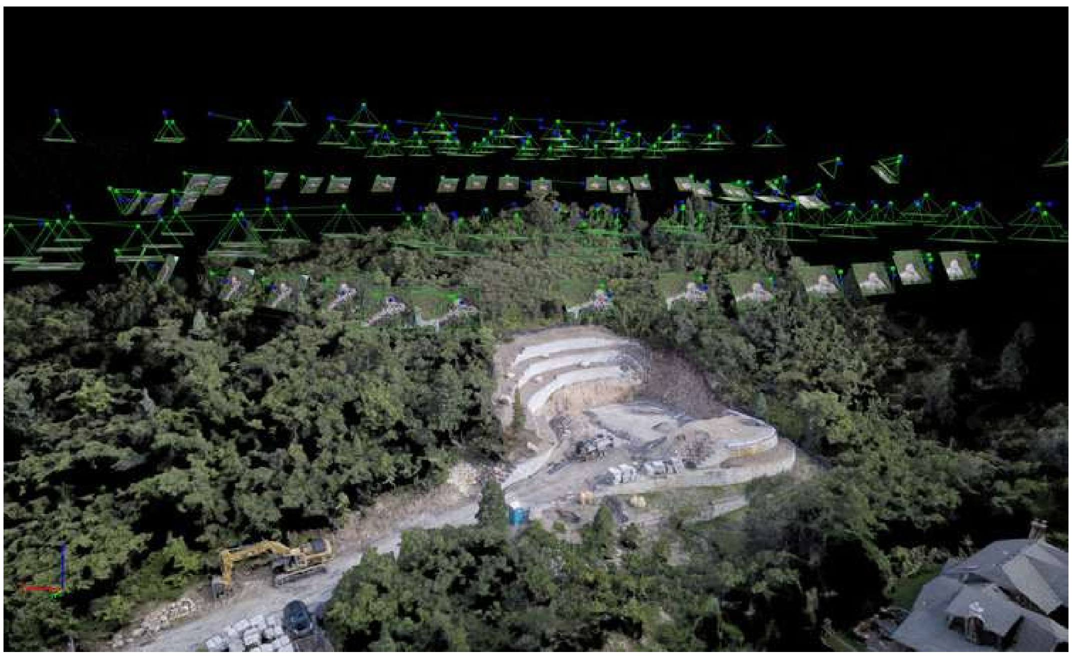
PICTURE 3. Robot taking photographs of a mechanical part (Garnett 2018)

Oil and gas companies use photogrammetry as a way of generating geospatial data of the routes their pipelines use, providing a better understanding of the terrain. Geospatial data is critically important in mining and topographical projects (Lipman 2018). Photogrammetry is not the only method to generate geospatial data. Modern photogrammetry combined with emerging technologies such as the use of drones has proven to be more efficient, faster and relatively inexpensive when compared to more traditional alternatives.

2.3 Types of photogrammetry

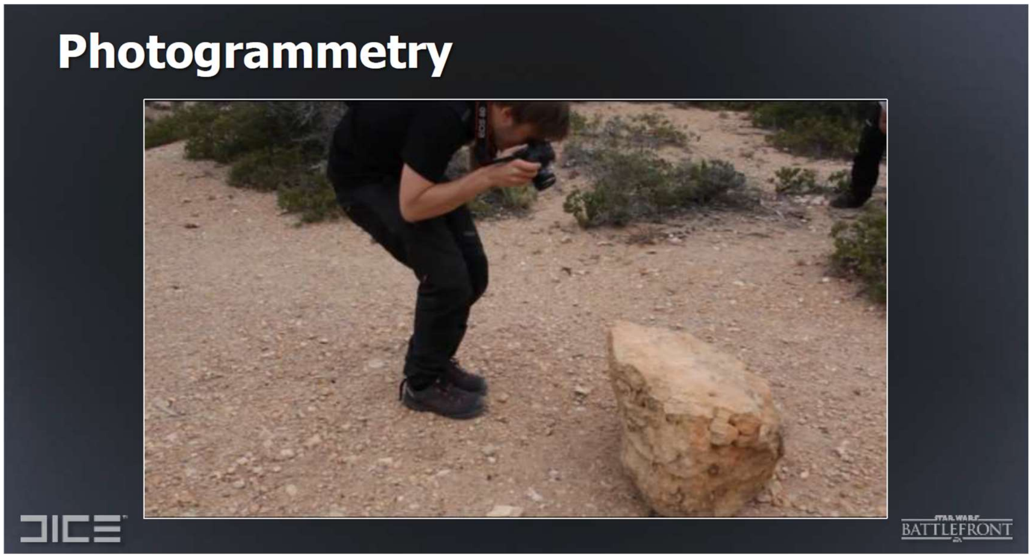
There are two main types of photogrammetry, aerial and terrestrial. Terrestrial photogrammetry is also called close-range photogrammetry.

Modern aerial photogrammetry is commonly accomplished using a drone which carries the camera that takes photographs of the subject. Modern drones use apps and software that allow their users to determine routes and paths that the drone will follow, making the process more efficient compared to manually flying and shooting every photograph. Aerial photogrammetry tends to focus on larger or more inaccessible subjects such as buildings or extensive areas (Drone mapping, Photogrammetry… 2016).



PICTURE 4. Point cloud showing drone positions (Drone mapping, Photogrammetry… 2016).

Terrestrial or close-range photogrammetry can be performed with DSLR cameras or even smart phones. Whether the camera is hand-held or mounted on a tripod depends on the needs of the project (PICTURE 4). This thesis will focus on close-range photogrammetry as the use of drones is out of scope.



PICTURE 4. Screen shot of close-range photogrammetry with a hand-held DSLR (Hamilton, Brown 2016)

2.4 Process of obtaining a 3D model through photogrammetry

This subsection provides an overview of the whole process of obtaining a 3D model through photogrammetry. This includes topics ranging from the acquisition of the photographs to the optimization of the final 3D model, ensuring it is feasible for use in real-time engines such as Unity or Unreal Engine 4.

Acquiring the photographs is one of the most important parts of the process as the quality of the images is directly related to the quality of the final 3D model. To ensure good quality it is important to plan the shooting process, understand the relevant camera settings and consider the lighting conditions of the scene.

Processing the photographs is the next step. This involves not only the use of the specialized software to generate the 3D model, but also editing them to increase the sharpness, and correct light inconsistencies across the photographs.

Optimization is a critical aspect to ensure 3D models have an acceptable polycount. 3D models generated through photogrammetry have many polygons in order to represent the fine detail of the real subject. Nowadays, real-time engines such as Unity or Unreal engine 4 are capable of handling millions of polygons. However, a rock generated through photogrammetry can be composed of millions of polygons. In a scene where other 3D models are rendered at the same time, the number of polygons of the rock can overwhelm the engine, preventing the final game or application from working. Hence, decreasing the polycount of the rock is critical to make its use viable.

3 PHOTOGRAMMETRY PART 1: CAPTURING IMAGES

3.1 Capturing methods

There are two methods that can be employed to capture the subject, photographs and video. Capturing individual photographs generally provides higher resolution. Even when recording 4K video the resolution of the frames is limited to 3840 x 2160 pixels. On the other hand, DSLR cameras can provide resolutions of 5472 x 3648 pixels, but the whole process of shooting the photographs is likely to take more time (Lachambre, Lagarde, Jover 2017, 20).

Whether to choose to capture using video or photographs will depend on the needs of the project. If the subject that is being captured is meant to be used to fill the background, and therefore it will never be shown from a close distance, shooting video might be more beneficial due to its faster approach. However, if the subject is meant to be the protagonist in the final result, choosing to shoot photographs instead of video will likely provide finer detail.

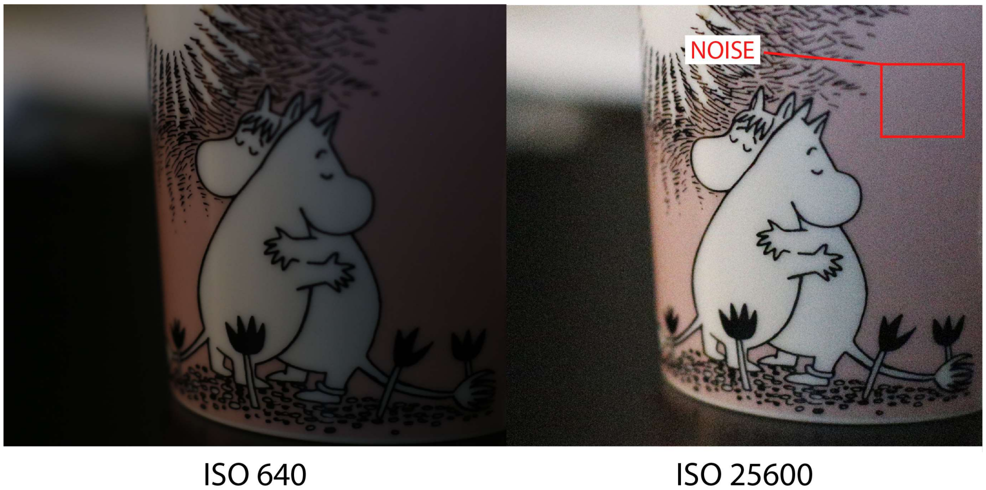
Both methods will be developed in detail in the following subsections.

3.1.1 Capturing photographs

This subsection provides a deeper understanding of the considerations regarding camera settings, image formats and overall good practices for capturing photographs.

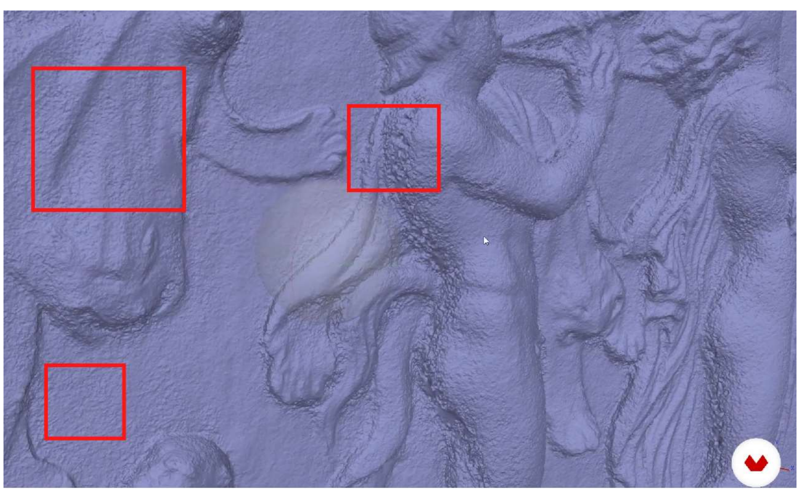
DSLR cameras are ideal for photogrammetry. It is possible to use any other device that can produce photographs, but DSLR cameras offer versatility and control which makes them perfect for the task. Regardless of the device, the user needs to have control over a minimum of three camera settings, ISO, shutter speed and aperture.

The ISO setting in digital cameras, is a value that gives the user more control over the exposure of the final image. When the ISO value is increased the image becomes brighter, this allows photographers to obtain images with correct exposure in low-light conditions (Mansurov 2019).



PICTURE 5. ISO comparison. Higher levels of ISO provide a brighter image.

However, there is a drawback to higher values of ISO. The ISO is an electronic signal that boosts the brightness of the image but introduces digital noise as a consequence. In photogrammetry the noise present in the photographs shows not only in the final textures of the 3D model (PICTURE 6), but also generates grainy surfaces that the original subject did not have (Mansurov 2019).



PICTURE 6. Noise in the input photographs creates grainy surfaces that is not present in the real sculpture (Chumilla 2019, modified)

The shutter speed of the camera controls how long the camera shutter is open, exposing the sensor to light. Fast shutter speeds allow light into the sensor for a shorter period of time. This translates into darker images, but guarantees that the photograph is properly captured, avoiding motion blur, a phenomenon that occurs when fast moving objects cross the frame while the camera is photographing at a relatively slow shutter speed. Motion blur can also occur when the shutter speed is too slow to shoot with the camera held in hand rather than stabilized on a tripod (Maio 2020).

Shutter speed is usually represented as fractions of a second. For instance, 1/4000 s means a second divided between 4000. This speed can freeze extremely fast objects such as the blades of a helicopter in movement. On the other hand, setting the shutter speed to 1/10 s will cause motion blur on the final image even if the subject is completely still (FIGURE 2). This motion blur is caused by the unavoidable movement that is caused by using the camera hand-held (Maio 2020).

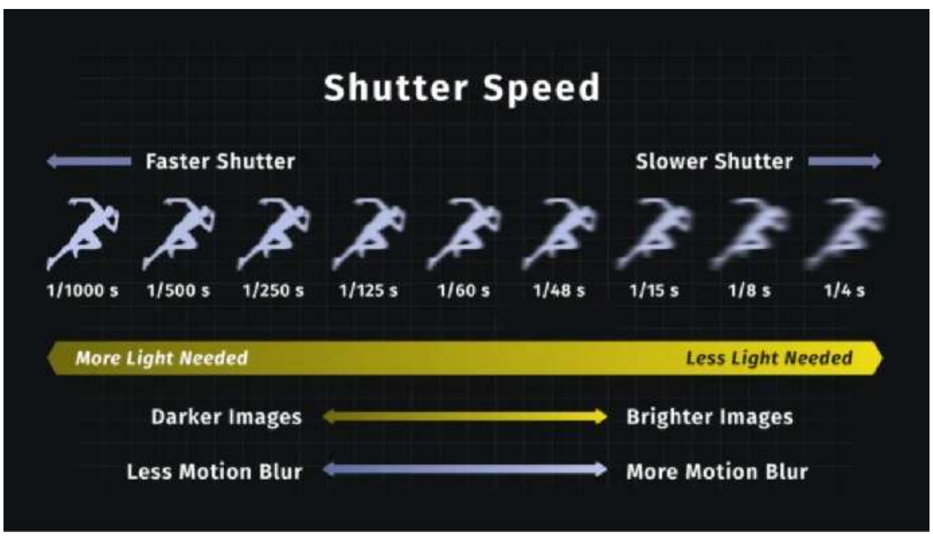


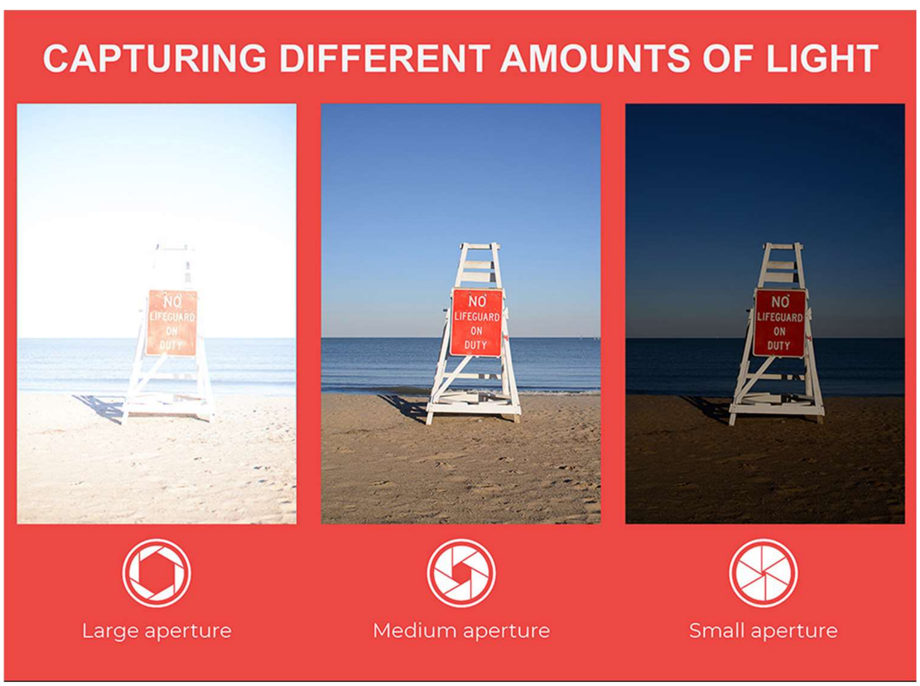
FIGURE 2. Visual representation of different shutter speeds (Maio 2020, modified)

Using fast shutter speeds in low light situations means dark photographs. Examples of low light situations are shooting outdoors at night or in indoor spaces that are not properly lit. It is recommended to use speeds no lower than 1/100 s when shooting hand-held, otherwise there is a risk of obtaining blurry images, even when shooting still subjects.

Blurred images are a problem in photogrammetry. The photogrammetry software struggles to find points that match with the ones in other photographs, and it is unable to determine the position the photo was taken from. As a result, the photograph is discarded and does not contribute to generating the point cloud.

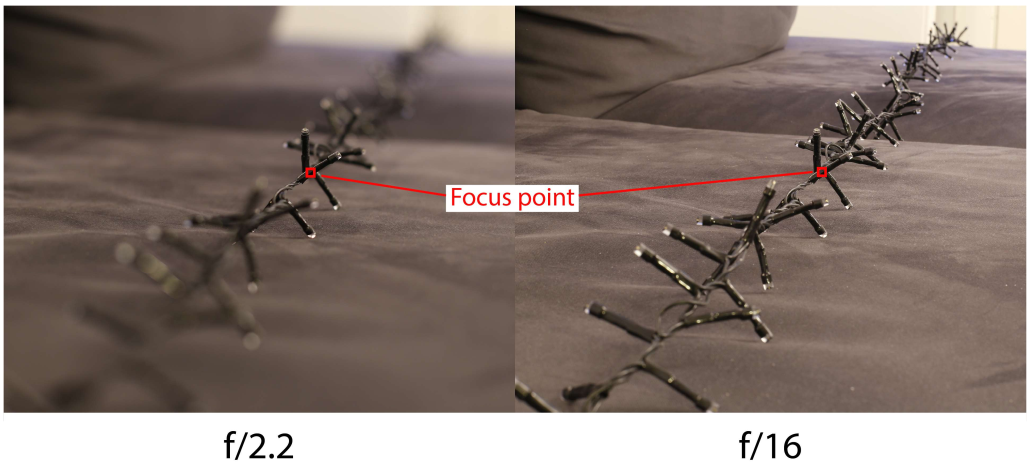
A tripod is the best solution when the low light conditions require to shoot using shutter speeds below 1/100 s. The tripod keeps the camera still while the shutter is open preventing motion blur.

The third setting to consider is the aperture.” Aperture can be defined as the opening in a lens through which light passes to enter the camera.” (Mansurov 2019). The behaviour of the aperture is very similar to the human iris. If the photograph is too bright, lowering the aperture will decrease the amount of light passing through the lens, making the final image darker and thus correctly exposed (PICTURE 7).



PICTURE 7. Different apertures change the brightness of the final image (Mansurov 2020)

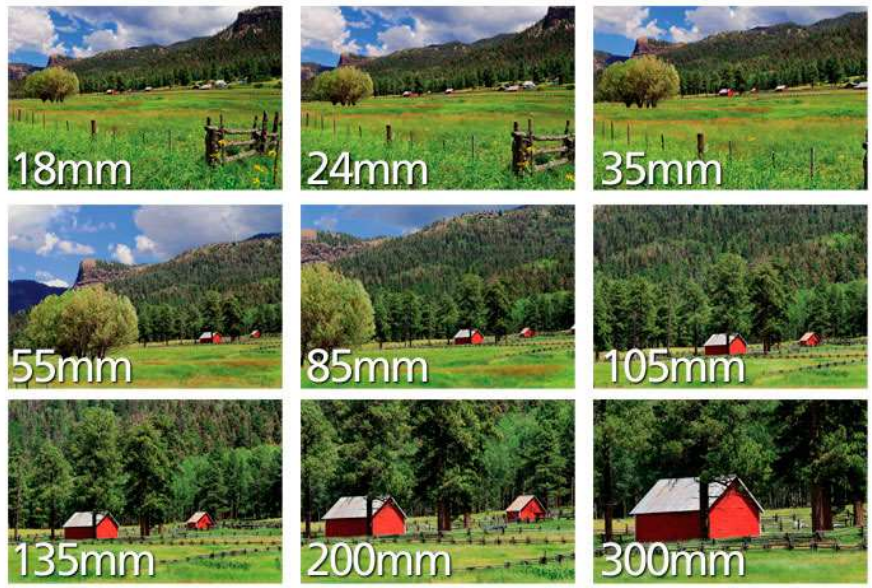
An aperture f/1.4 is considered a large aperture, making the image brighter. On the other hand, f/16 is considered a small aperture, making the image darker. The aperture setting is relevant in photogrammetry because besides changing the brightness of the final image is directly responsible for the depth of field (PICTURE 8). “Depth of field is the distance between the closest and farthest objects in a photo that appears acceptably sharp.” (Gray 2020).



PICTURE 8. Depth of field comparison between different apertures

In photogrammetry it is important to keep the image as sharp as possible. Smaller apertures such as f/16, maximize the sharpness of the elements in the background and the foreground. Photogrammetry software uses all these elements to extract the necessary information to create the point cloud, and subsequently generate the final 3D model. The final textures of the model are also created from the photographs. Images that are quickly out of focus beyond the focus point might result in smudged-looking textures in some areas of the 3D model.

In addition to these three settings, the lens of the camera plays a role. DSLR camera lenses are classified based on their focal length. Lenses with shorter focal lengths can capture a wider field of view (PICTURE 9). The focal length is expressed in millimetres (Berkenfeld, Black, Corrado & Silverman 2021).



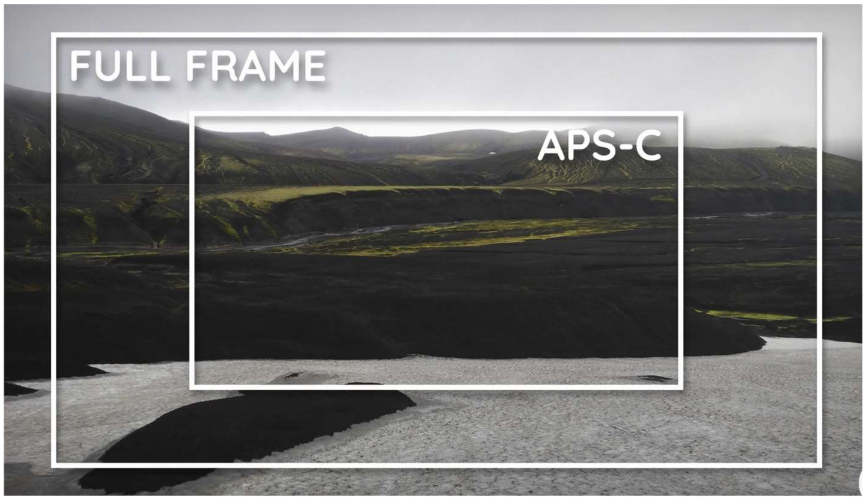
PICTURE 9. Field of view captured by different lenses (Black 2021)

All lenses have some level of imperfection, this translates to distortion of the image around its border. Many photographic programs such as Adobe Lightroom offer ways to correct this distortion. However, lens correction in Adobe lightroom or other photographic programs is not recommended for photogrammetry. Photogrammetry programs take into consideration the lens used to capture the object and apply the necessary corrections that are more suitable for the generation of the 3D model.

Overall, prime lenses, lenses that have no optical zoom and therefore have only one focal length number, provide better quality than zoom lenses. When using a zoom lens for photogrammetry, it is highly recommended to keep the lens fixed on a focal length. Changing the focal length to zoom in and out of the subject could cause problems with the photogrammetry software.

A 50 mm prime lens is a popular choice for photogrammetry. The behaviour of this lens is the closest to the human eye and its distortions are not as acute as the ones in other lenses.

The sensor of the camera plays a role in the choice of the lens. The 50 mm lens described in the previous paragraph corresponds to a camera that is equipped with a full frame sensor, many cameras feature cropped sensors such as APS-C sensors (Jirsa 2016). What is relevant is that a 50 mm lens for a full frame camera and a 50 mm lens for a camera with an APS-C sensor are not equivalent. To compensate for the reduction of field of view that comes with a smaller sensor, a wider lens is needed. In case of the 50 mm lens for full frame cameras, a 35 mm prime lens is considered equivalent (PICTURE 10).



PICTURE 10. APS-C sensors capture a smaller field of view than full frame sensors (Chumilla, 2019)

The format of the photographs is also crucial to guarantee a high-quality result. DSLR cameras, as well as other devices, can save photographs in RAW format. The RAW format is an uncompressed and lossless file that contains all the details of the captured photographs. This allows for greater flexibility during the editing process.

Ultimately, photogrammetry is a robust process that can produce results even when the conditions are not ideal. Not having a 50mm prime lens is not an impediment to produce 3D models, nor is having a camera with an APS-C sensor instead of a full frame sensor.

3.1.2 Capturing video

Capturing video is another possibility to acquire the photographs to input in the photogrammetry software. The differences compared to capturing photographs is that in the case of video the individual frames are the source the software will use to extract information. Some programs such as RealityCapture already accept video as input. However, extracting frames using video editing programs, for instance, Adobe Premiere, is also possible (Lachambre, Lagarde, Jover 2017, 20).

It is inevitable that several of the extracted frames have motion blur, even when moving at a very slow pace around the object. Many cameras record at 1920 x 1080 pixels at 24 or even 60 frames per second. In a minute of footage that would be 1440 images in case of the former and 3600 in case of the later. Recording at 4K produces images of 3840 x 2160 pixels, this is still below the resolution that shooting individual photographs produces.

The advantage of capturing video over individual photographs is the speed. Whether to use one or the other will depend on the project. The benefits of capturing video might outweigh its drawbacks in certain situations.

3.2 Assessing the subject

Photogrammetry has proven to be particularly useful in the game and film industry. Cliffs, rocks and trees are common subjects as photogrammetry is particularly well suited to generate 3D representations of accurate detail. Traditional modelling or sculpting these subjects is difficult and time consuming. Reaching a photorealistic result is easier and faster with photogrammetry.

However, there are subjects that are more easily recreated in 3D using other methods. Understanding and assessing the subject before starting the acquisition of photographs is critical. A coffee mug is a relatively simple object, generating a 3D model through photogrammetry, as opposite to traditional modelling, would be hard to justify.

Furthermore, subjects with surfaces that are reflective or have little to no detail are extremely hard to capture as photogrammetry programs are unable to find common points across the photographs.

3.2.1 Subject surface detail

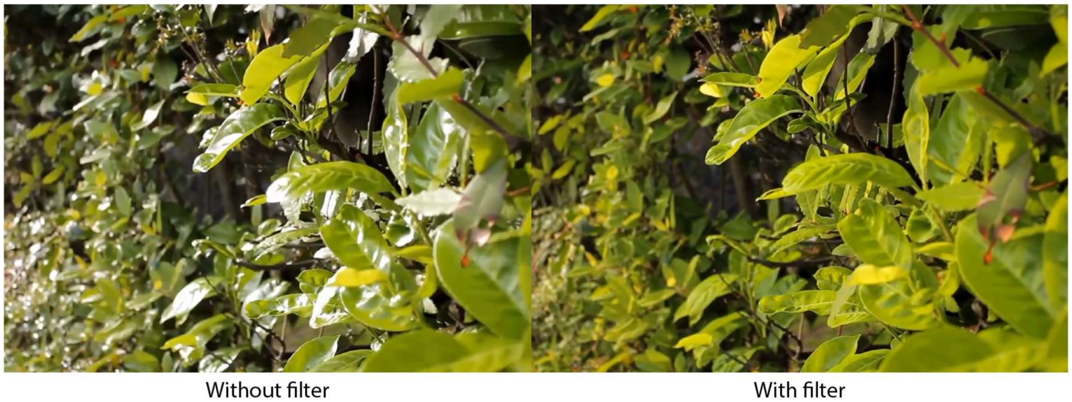
Photogrammetry programs such as RealityCapture, Meshroom or Agisoft Metashape rely on surface detail to find points and subsequently create the point cloud. Subjects that have polished surfaces such as ceramics with little detail cause the programs to struggle. Objects that present shiny or glossy surfaces, such as metals are also troublesome.

Objects made of glass present the most challenging situation. Their polished surface combined with their transparency make them practically impossible to capture without prior treatment. This explains why many 3D models that feature glass parts, such as windows or lights, appear unresolved, often looking like holes in the model (PICTURE 11).

PICTURE 11. The window and the motorbike's front light appear as holes as the program was not able to capture the detail (Bandera 2015, modified)

3.2.2 Troubleshooting

It is possible to overcome the difficulties presented by objects that have difficult surfaces. Polarising filters are camera filters that, when rotated, eliminate highlights on the object, making the surface look matte (PICTURE 12).



PICTURE 12. Polarising filter removes highlights on the subject (Chumilla 2019, modified)

There are occasions in which polarising filters might not be enough. For instance, transparent surfaces will likely need a more extensive treatment before they can be captured. For these cases, it is possible to find coating sprays specifically meant for the purpose. Coating the surface will tint it with a solid, matte colour, making it more suitable for photogrammetry (PICTURE 13).

PICTURE 13. The original object is covered with the 3D coating spray and an additional layer of black paint (Busby 2016, modified)

In the picture above, the artist decided to apply an additional layer of face paint. Although this paint is not specifically made for photogrammetry purposes it is ideal for the task, as it is easy to remove it once the object is photographed. The face paint adds contrast and detail to the surface on top of the white matte coating from the spray (Busby 2016).

However, spraying the subject might not be acceptable in some situations. For examples, projects where the subject is valuable, or it is fragile such as a historical artefact.

3.3 Environment considerations

There are some considerations that should be kept in mind, regardless of whether the subject is captured indoors or outdoors. If the project requires capturing a series of small items, it might be beneficial to capture them in a studio. This would provide more flexibility as weather or outdoor lighting changes would no longer be relevant.

On the other hand, capturing indoors has additional lighting requirements to ensure the object does not cast strong shadows that can be problematic in the texturing phase.

3.3.1 Capturing indoors

To capture indoors it is critical to set up proper lighting. The subject needs to be lit using neutral white lighting, and the light sources need to be positioned correctly to prevent directional shadows on the subject. These shadows would be visible in the final texture of the 3D model, something that is extremely hard to correct. Placing the model in a game engine would show how the light source in the scene does not match the directional shadows on the object, breaking the illusion.

When capturing indoors there are several accessories that help ensuring good lighting conditions. Ring flashes are popular when the subject is relatively small. A polarising filter is a good addition to remove the highlight caused by the flash.

The use of light boxes is also common. These boxes are available in different sizes. The model is placed inside, usually on a turntable platform, the light box diffuses and spreads the light coming from the outside. This is helpful to achieve even light on the model and prevent directional shadows (PICTURE 14).

PICTURE 14. A setup using two lights and a light box (van de Beek 2019)

It is relevant to mention that using a turntable adds complexity to the project. The camera stays still and the subject rotates, as opposite to the model staying still while the camera moves around taking photographs. This will be explained more in detail in the following section of the thesis.

Besides the mentioned accessories, color checkers are widely used. The color checker is a board that displays several colors and shades that photographic programs can use to make sure the colors of the scene look correct. This is necessary as in many situations, sources of light or surfaces in the environment tend to bleed colors into the photographs. For instance, in an environment where there is a yellow wall, and despite the use of white neutral lights and light boxes, there is a possibility that the yellow color of the wall influences the light. This would cause the white neutral light to look slightly yellow. Without correcting this in the editing process the final texture of the 3D model would be affected by the yellow tint. The use of color checkers applies to both indoor and outdoor situations.

Finally, it is worth mentioning that there are specialized studios that focus on generating 3D characters based on actors and actresses. While it is possible to recreate a human being with a single camera, achieving the high-quality standards that many films and videogames have requires a more demanding setup. Since it is extremely hard for a person to stand completely still for an extended period of time, studios that specialize in this type of photogrammetry have rigs that arrange many cameras focusing on the subject. The cameras are rigged to shoot all at once, generating the necessary images from different angles (PICTURE 15).

PICTURE 15. Screen capture of cameras in a rig that focuses on capturing the actor (Capturing the Apocalypse 2016)

3.3.2 Capturing outdoors

Monitoring the weather becomes a necessity when approaching photogrammetry projects outdoors. Overcast days are considered ideal for photogrammetry. The light filtered through the clouds provides soft, diffuse lights, without high contrast or strong directional shadows.

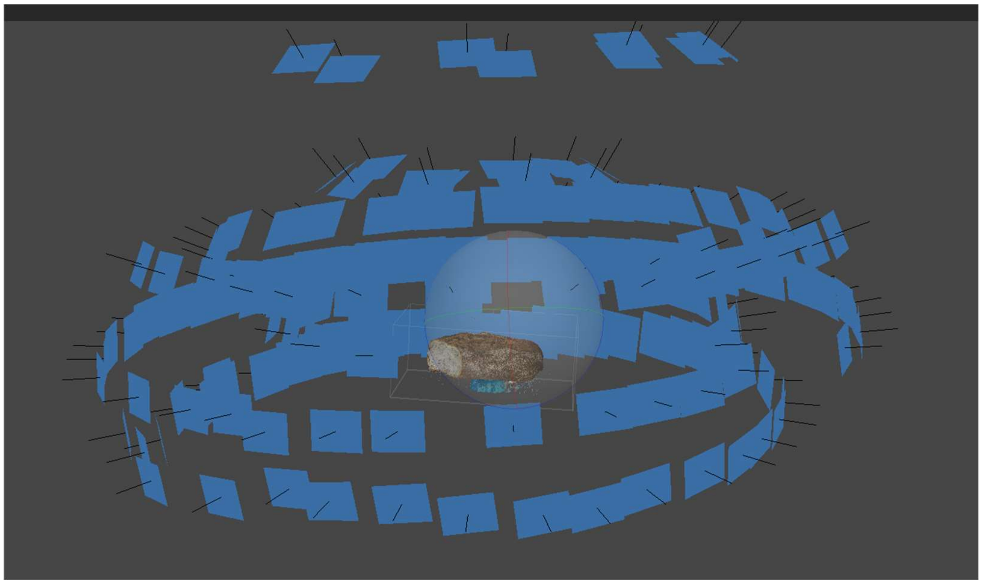
Sunny days are also acceptable as long as it is ensured the subject is in shadow. This can be achieved with the help of an occlude (PICTURE 16).

PICTURE 16. The occluder prevents direct lighting on the subject (Lachambre, Lagarde, Jover 2017)

3.4 How to approach the shooting

Image overlap is the most important concept when shooting photographs for photogrammetry. The number of photos taken depends on the complexity of the model. Overall, 60-70% of overlap between photographs is recommended (ClassyDogFilms 2014).

The photographs can be taken by simply walking around the subject. The goal is to complete a loop of photographs that surround the subject. Once the first loop is completed, a second loop can be started elevating the camera so it can capture the upper parts of the subject. This process should be repeated until capturing the whole subject. If done correctly, the positions of the cameras will resemble a semi-sphere when the photogrammetry software starts to recreate them in 3D space (PICTURE 17).



PICTURE 17. The camera positions resemble a semi-sphere.

Every angle of the subject should be captured in the photographs. The photogrammetry software will not be able to recreate the part of the subject if it does not have photographs of it. This will cause the model to have holes or poorly detailed areas around the parts that were not photographed (PICTURE 18).

PICTURE 18. The 3D model has holes on the areas that were not photographed.

When shooting outdoors it is important to ensure that there is enough space to move around the object in order to photograph it from every angle. There are situations where it will not be possible to photograph the object from all of its sides. The resulting 3D model can still be useful in a game or real-time application as long as the part that has not been generated due to the lack of images is hidden from the viewer (Lachambre, Lagarde & Jover 2017, 25).

To capture flat surfaces, it is recommended that the photographs are taken so the line of sight of the camera is perpendicular to the surface (FIGURE 3). In cases where the surface has many protruding details, such a brick wall where some bricks protrude more and some are missing, it is recommended to tilt the camera to capture those areas that otherwise would not be visible in the photographs (Chumilla, 2019).

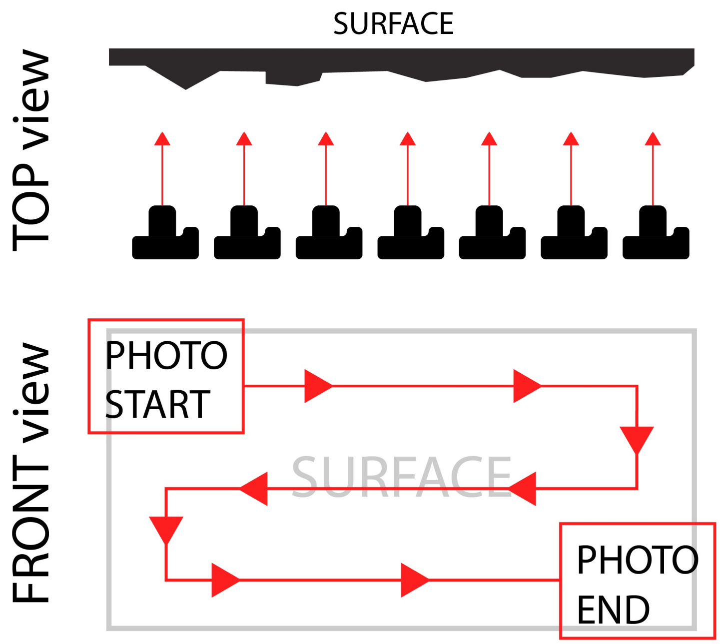


FIGURE 3. Capturing flat surfaces.

The approach is similar when capturing rooms. The back of the photographer should be against the wall while they photograph the opposite wall (FIGURE 4).

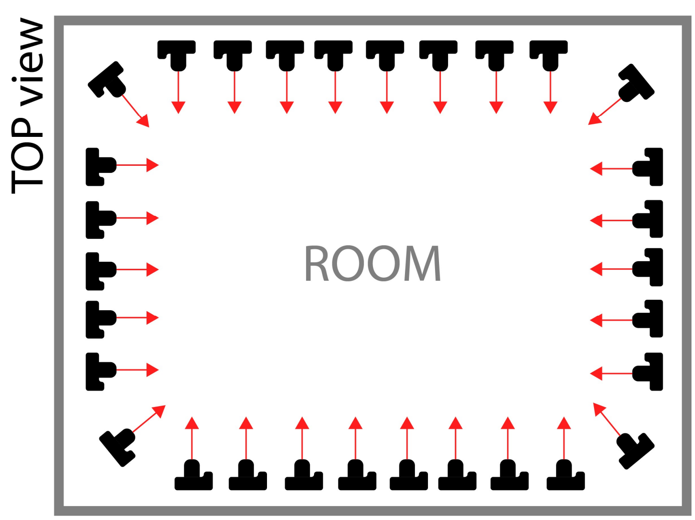


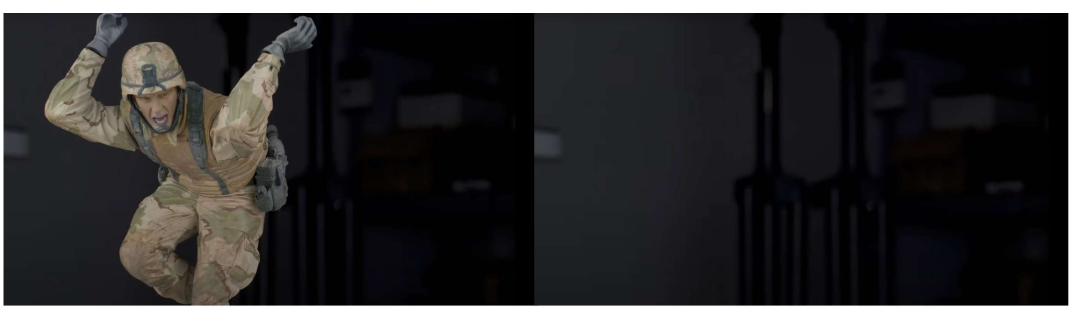
FIGURE 4. Capturing rooms.

When the subject is a small object it can be captured in studio, placing it on a turntable. Using a turntable might confuse the photogrammetry software. As mentioned previously, photogrammetry programs use the background in the photographs to find the position of the cameras in 3D space. When a turntable is used, the camera remains still and the subject rotates. The program does not detect changes in the background and places all the cameras in the same position, as a result, the 3D model is not generated correctly.

There are several solutions to this problem. Some photogrammetry programs support using masks. These masks isolate the subject from the background. When the background is ignored, the program can focus solely on the movement of the subject and use that information to place the cameras in 3D space. If the process is successful, the cameras will form loops as the ones seen in picture 14.

Metashape is one of the programs that supports masking as a way of isolating subject from background. The program offers several tools to adjust and edit the masks. However, manually creating the individual masks for projects that can involve hundreds of images is time consuming and defeats the purpose of using photogrammetry as a fast alternative to generating 3D models.

To solve this issue, Metashape can use photographs where only the background is visible to automatically generate the masks for every photograph taken for the project (Agisoft Metashape User Manual 2019, 82).



PICTURE 19. The photograph of the background is used to automatically create the masks (ClassyDogfilms 2014)

An alternative to this method is to prepare a seamless black or white background for the turntable shots. When the background has no details and is out of focus the photogrammetry software is unable to find points in it. This forces the software to use the rotating subject as the only source to find points and generate the 3D model. To put it simply, the software is fooled to believe that the camera is moving as it cannot find any details on the background.

4 PHOTOGRAMMETRY PART 2: PROCESSING IMAGES

4.1 Image treatment

With the acquisition of the photographs completed, the next part of the process involves doing a preliminary image treatment to ensure the best quality is provided as the input for the photogrammetry software.

This editing is done in photographic software such as Adobe Lightroom or Adobe Camera Raw. All the images are imported but the manual edits are only done in one of the photographs to then apply them to the rest, so the changes are identical across the photographs.

The color checker tool that was mentioned previously is ideal to guarantee that the colors in the photograph are calibrated to match the colors of the subject (PICTURE 20). The purpose is to guarantee that there are no color imbalances that can give a color tint that is not part of the real subject. (Rojas 2016).

PICTURE 20. Color correction done using a color checker (Rojas 2016)

Once the color is corrected, there are other editing changes that can improve the quality of the final result. Increasing the sharpness of the images, as well as adding brightness to the shadows or decreasing the highlights can help to reveal or emphasize small surface details that the photogrammetry program can use to create a more accurate model.

It is also recommended to correct the chromatic aberration of the photographs, that otherwise could show in the final texture. Chromatic aberration is a common issue with camera lenses that manifests producing blurry images or noticeable colors around the edges of an object (Mansurov 2019).

PICTURE 21. Chromatic aberration produces odd colors around objects (Mansurov 2019)

These corrections are ideally done on the RAW format of the photographs as mentioned previously. Examples of RAW formats are CR2 for Canon or NEF for Nikon.

Some artists choose to export two sets of photographs, one set helps the photogrammetry program to reconstruct the subject, and the other to extract the final textures of the 3D model. The idea is that the more heavily edited set of images is beneficial for the photogrammetry program, but it might look slightly unrealistic to be used as textures, hence the second set of images (classyDogFilms 2014).

The exporting format from the photographic software depends on the project's needs. It is recommended to export images in TIFF format, a common lossless format. According to Another World Studios, this format seems to provide slightly better results compared to JPEG or RAW, their test was performed on RealityCapture. However, TIFF images take more space on disk, which might not be suitable to every project or pipeline (anotherworld 2020).

Regardless of the format, the exporting settings should include all the metadata. The metadata is the information contained in every photograph, information about the camera model, the lens, resolution, etc. This information will help the photogrammetry software to consider the lens distortion and other important aspects.

4.2 Photogrammetry software

The workflow photogrammetry programs follow is technically complex. This section of the thesis aims to provide a basic understanding of the different steps photogrammetry programs follow in order to create a 3D model.

Once the user has input the photographs the first process the program will execute is the alignment of the images. The program will analyse the images looking for pixels that remain relatively invariable across the photographs. The program will also use the images to create depth maps (FIGURE 5). "Depth map is a kind of image which is composed of the gray pixels defined by 0 ~ 255 values. The "0" value of gray pixels stand for that "3D" pixels are located at the most distant place in the 3D scene while the "255" value of gray pixels stand for that "3D" pixels are located at the most near place." (I-Art 3D 2019).

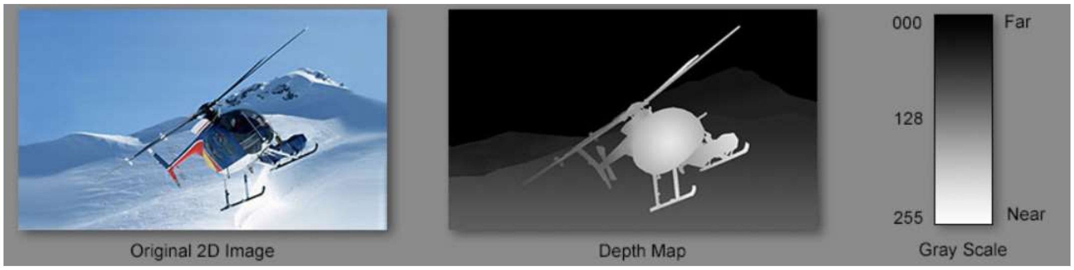
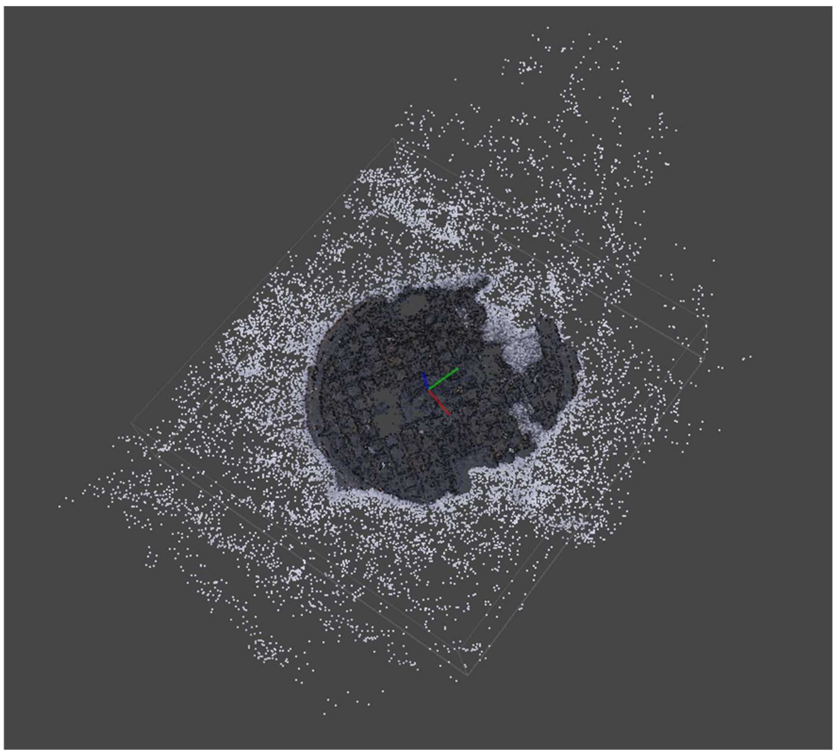


FIGURE 5. Depth map (I-Art 3D 2019)

All this information is used to generate a point cloud (PICTURE 22). “A point cloud is basically a set of data points in a 3D coordinate system, commonly defined by x, y, and z coordinates.” (Kai Chua, How Wong & Yee Yeong 2017, 87).



PICTURE 22 Point cloud visualization, subject is a manhole cover

Once the point cloud is created the program will use it as a base to generate the polygon mesh of the 3D model. The last step of the process is to generate the texture for the 3D model. Many photogrammetry programs allow the user to determine the resolution of the texture. With the textures done, the model can be exported. The whole process is summarized on figure 6.

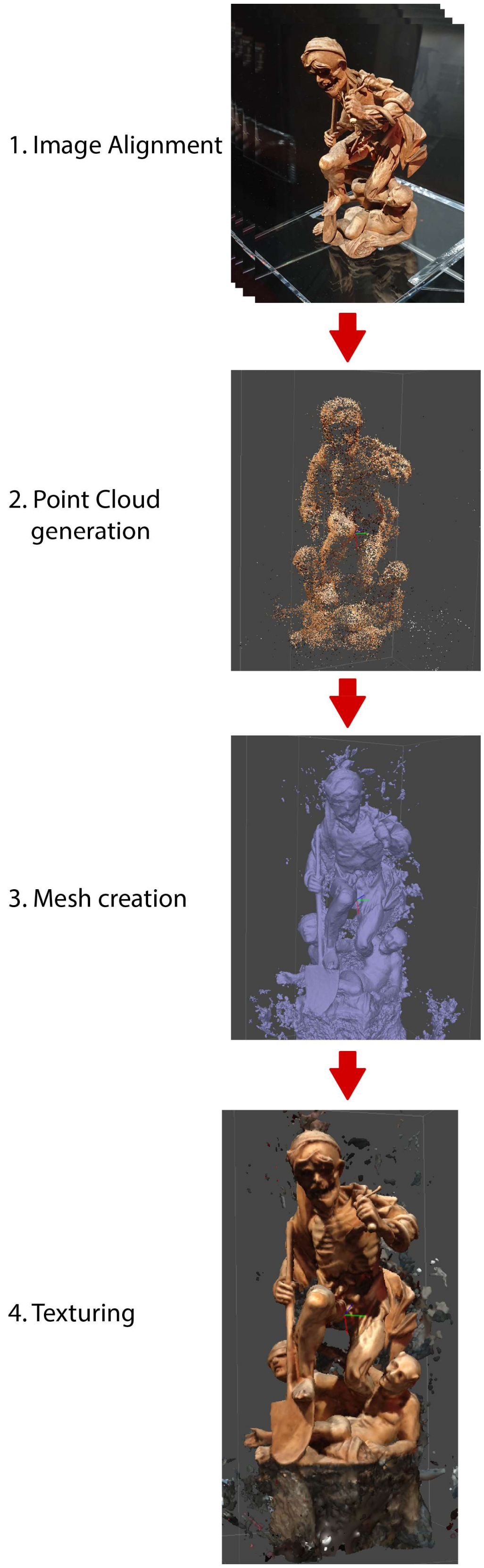


FIGURE 6. Photogrammetry process followed by Agisoft Metashape

4.3 Software comparison

This section offers a brief discussion of the programs that will be used and compared in the following sections of the thesis. Programs approach the photogrammetry process differently and sometimes they use different terminology. It is not the intention of this thesis to attain a deep understanding of the different technical settings used by the different programs to generate 3D models. Mastering the settings in each of the programs would require a level of research and experimentation that is out of the scope of the thesis.

The aim is to compare what a user with a surface-level understanding of each of the programs can achieve with default or close to default settings. The programs that will be compared are Meshroom, Agisoft Metashape and RealityCapture. Meshroom is an open-source program which is currently free. Agisoft Metashape offers several licenses and has a one-time payment model. RealityCapture offers a system called PPI; it stands for pay-per-input. The user buys PPI credits which allow them to process a given number of megapixels or scan points. The program can be used freely, and the process can be done before purchasing PPI credits. The exporting options are unavailable until the user has purchased credits. There is also a subscription-based license offered on Steam, a well-known platform for purchasing digital games and sometimes software.

These three programs were chosen not only for their popularity, but also due to their varying price points. Meshroom represents the most accessible option while Agisoft Metashape and RealityCapture offer more features but at a higher cost.

Three subjects will be scanned for the purpose of the comparison. The chosen subjects are a Romanesco broccoli, a loaf of bread and a boot. Creating photorealistic 3D models of these objects constitutes a complex and time-consuming challenge when using traditional 3D modelling techniques. However, they are ideal subjects for a photogrammetry project. The three sets of photographs will be processed in each program for a total of nine 3D models. Three models produced in Meshroom, three produced in Agisoft Metashape and three produced in RealityCapture.

The aspects that will be taken into consideration are the time the program takes to generate the 3D model, the polycount of the final model and the visual quality. To clarify, the aspects that will be considered when checking the visual quality are whether the model has holes, floating geometry or strange deformations that were not part of the subject (PICTURE 23).

PICTURE 23. Problems that affect the visual quality of the model

The components of a computer that are more relevant for the photogrammetry process are the CPU, GPU and the RAM memory (ClassyDogFilms 2016). The characteristics of the PC that will be used to produce the 3D models are available in the following table.

TABLE 1. PC characteristics.

Operative System	Windows 10
CPU	AMD Ryzen 7 1700 Eight-Core Processor
GPU	Nvidia GeForce GTX 1060 6GB
RAM Memory	16GB

In addition, it is important to mention that the device used to capture the different subjects is a DSLR camera, model Canon 6D. This is a camera equipped with a full frame sensor. The lens used is a Canon EF 50mm f/1.4 USM, a prime lens with no optical zoom. The camera was mounted on a tripod in all three cases. The tripod allowed using slower shutter speeds without producing motion blur.

Once the models are generated, they will be exported from the photogrammetry software and imported into the same Blender scene. Blender is a 3D modelling suite, an open-source program that is available for free. Importing the different models into Blender allows for a more equal assessment as the 3D models are visualized by the same program and in the same light conditions.

4.3.1 Case one

For the first case the subject was a Romanesco broccoli. The subject was captured indoors. The light conditions were not ideal and the model texture presents color tints from the sources of light within the room. However, this was not a problem for the reconstruction of the model.

The photographs were captured in RAW format, a total of 160 images. They were edited in Adobe lightroom to improve the sharpness and remove chromatic aberration from the photographs. Once the edits were done the images were exported in jpeg format (PICTURE 24).

PICTURE 24. One of the images for the photogrammetry process

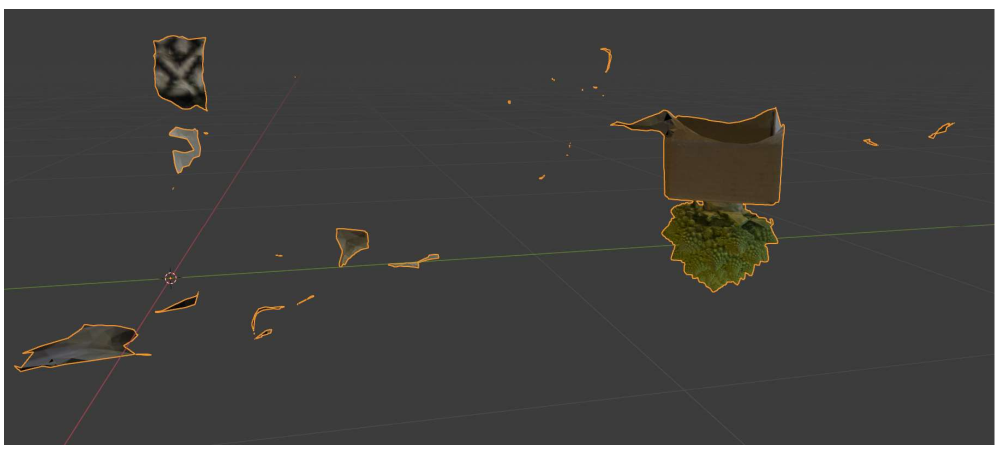
The following table shows the settings of the camera at the time of shooting the photographs.

TABLE 2. Camera settings, case one

ISO	200
Aperture	f/8
Shutter speed	2 seconds
Metadata	Excluded from images

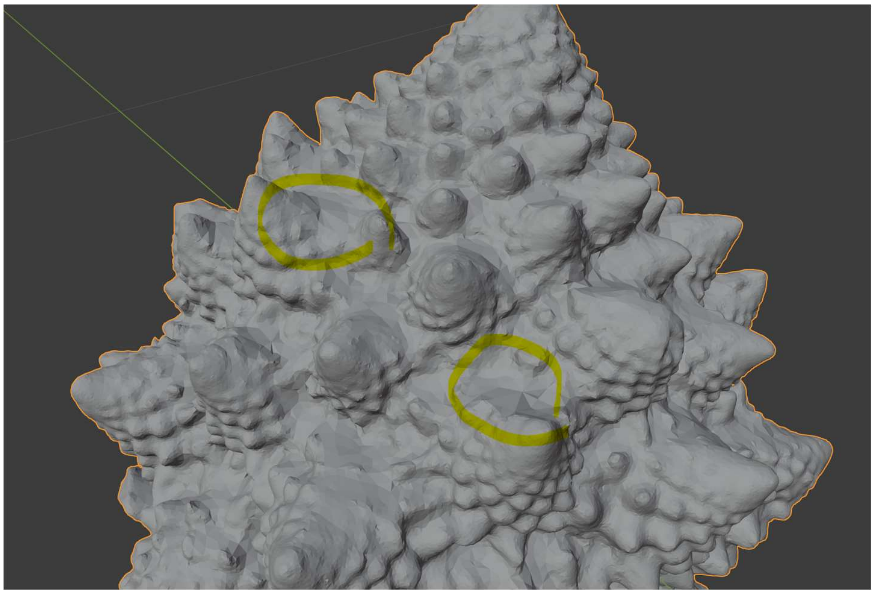
It is important to note that the metadata containing the information about the camera, ISO, focal length, shutter speed and aperture was not included. The idea was to see how the programs would react to the lack of information.

Starting with Meshroom, the program was able to align and use 109 of the 160 images. The process to generate the model took approximately 3 hours and 19 minutes. The model generated consisted of 1.275 million polygons (PICTURE 25).



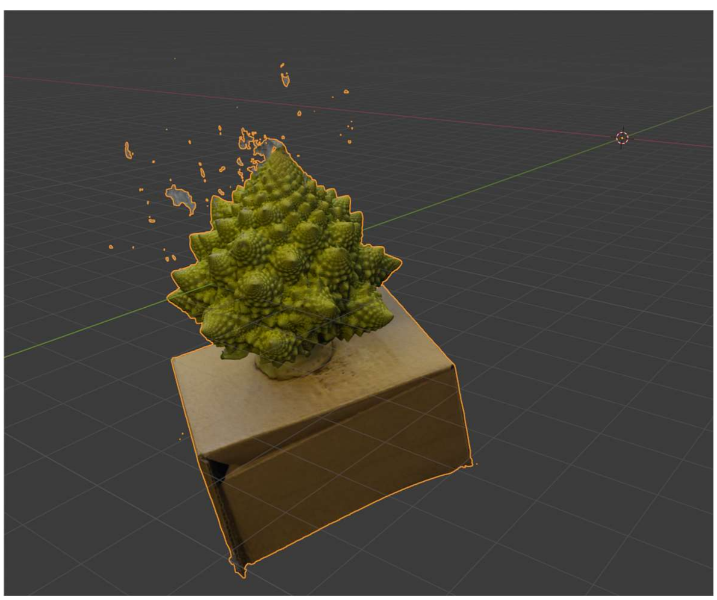
PICTURE 25. First look at the model in Blender, model from Meshroom

The model appears upside down, it is common for 3D models generated through photogrammetry to have the wrong orientation (PICTURE 25). The model presents floating geometry. This floating geometry is generated from details that surrounded the original scene. In addition, there are some areas of lower resolution, the triangle polygons are larger in these areas (PICTURE 26).



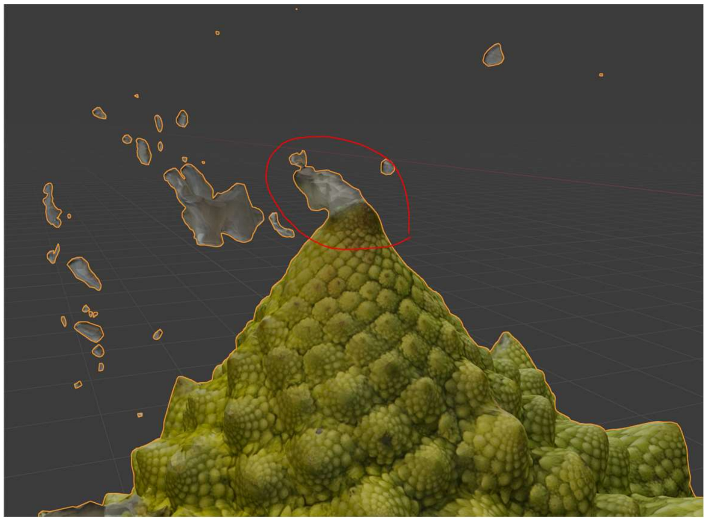
PICTURE 26. The model shows areas of lower resolution. Model from Meshroom

Metashape also aligned 109 of the 160 photographs. The model generated with Agisoft Metashape has a polycount of around 455,000 polygons. Metashape allows the user to choose the desired level of quality in each of the steps of the process. The quality settings were set to “medium” for all the parts of the process from aligning the images to reconstructing the mesh. The texture is a single image, 8192x8192 pixels of resolution. The model was generated in approximately 16 minutes (PICTURE 27).



PICTURE 27. First look at the model in Blender, model from Agisoft Metashape

The model has floating geometry. This is particularly problematic in this case, as part of the floating geometry is merged to the top of the model. Fixing the problem in Blender would require deleting the geometry that should not be part of the model, doing this would leave a hole in the model (PICTURE 28).

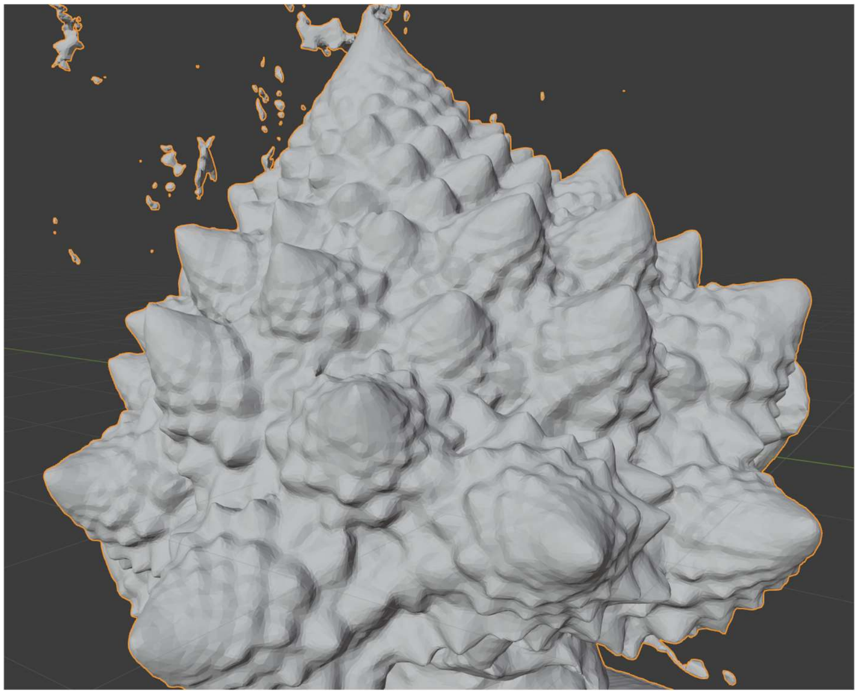


PICTURE 28. Floating geometry merged to the tip of the model.

The hole can be manually fixed in Blender, but the new geometry would not be properly textured. Fixing the problem completely would require some manual work. However, since the area is not large, fixing the issues of the model would still be faster than trying to recreate the subject through traditional modelling.

Furthermore, Agisoft Metashape offers tools to fix this problem during the formation of the point cloud. This solution is the fastest as it happens before the program generates the mesh. The reason this was not done in this case is because Meshroom does not offer tools to edit the point cloud and so the comparison would be even more unequal between the programs.

The consistency of the triangle polygons is better in this model compared to the one generated with Meshroom. There are no big differences between the sizes of the triangles, which translates in even resolution throughout the model (PICTURE 29).

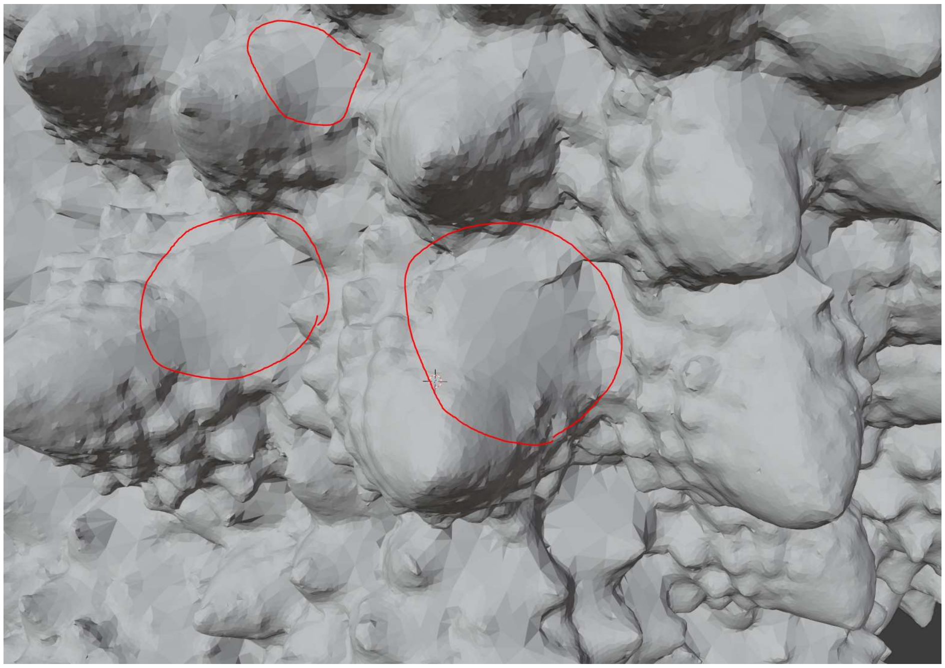


PICTURE 29. There are no noticeable differences in the size of the triangle polygons, the model shows even resolution. Model from Agisoft Metashape

As with the other programs, RealityCapture aligned 109 photographs of the 160. The model has the largest in polycount at around 4.8 million of polygons. RealityCapture took around 12 minutes to generate it. The model looks cleaner than the ones generated in Meshroom and Agisoft Metashape. It barely has any floating geometry (PICTURE 30).

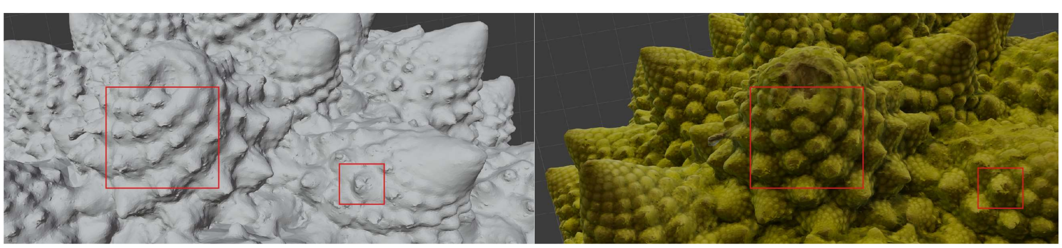
PICTURE 30. First look at the model in Blender. Model from RealityCapture

The resolution of the model is drastically higher than the one in the other models. There are areas of the model where the triangle polygons are noticeably larger than in others. This is not as acute as it was in the model generated in Meshroom (PICTURE 31).



PICTURE 31. Differences in triangle size throughout the model.

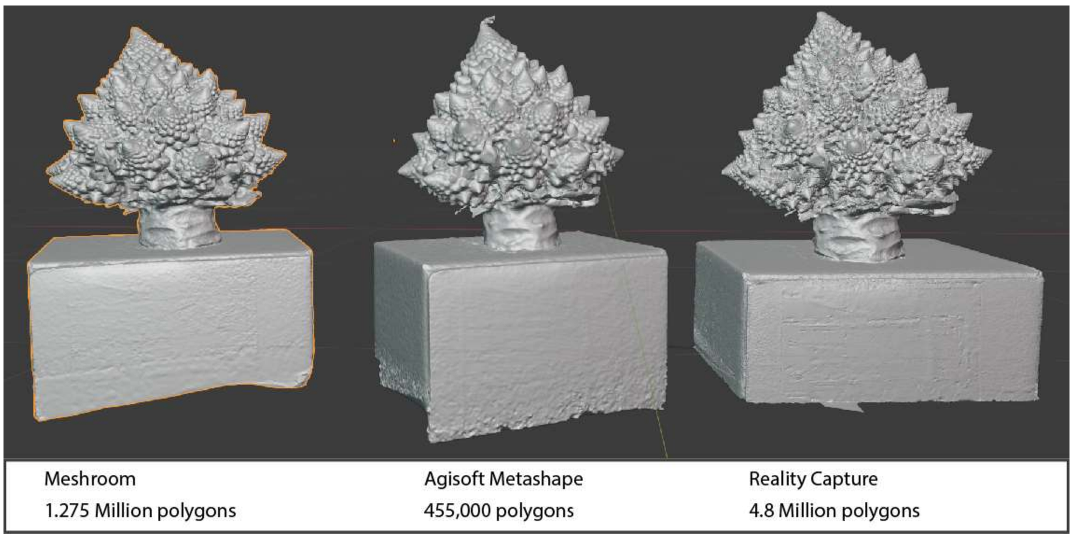
The model also has some areas where the polygons show small bumps or imperfections that do not seem to be part of the real subject. It is unclear whether the other models would show the same surface imperfections if they had a closer polycount to the model from RealityCapture (PICTURE 32).



PICTURE 32. Small surface imperfections in the model

These small imperfections can be easily smoothed either in Blender or with the different tools offered in RealityCapture to clean the model’s geometry. These corrections were not made since Meshroom does not provide any tools for cleaning the model.

Overall, Reality capture provided the best result in this case (PICTURE 33). A high-resolution model in approximately 12 minutes, relatively clean with almost no floating geometry or imperfections.



PICTURE 33. Models side by side. Most of the floating geometry hidden for the screenshot.

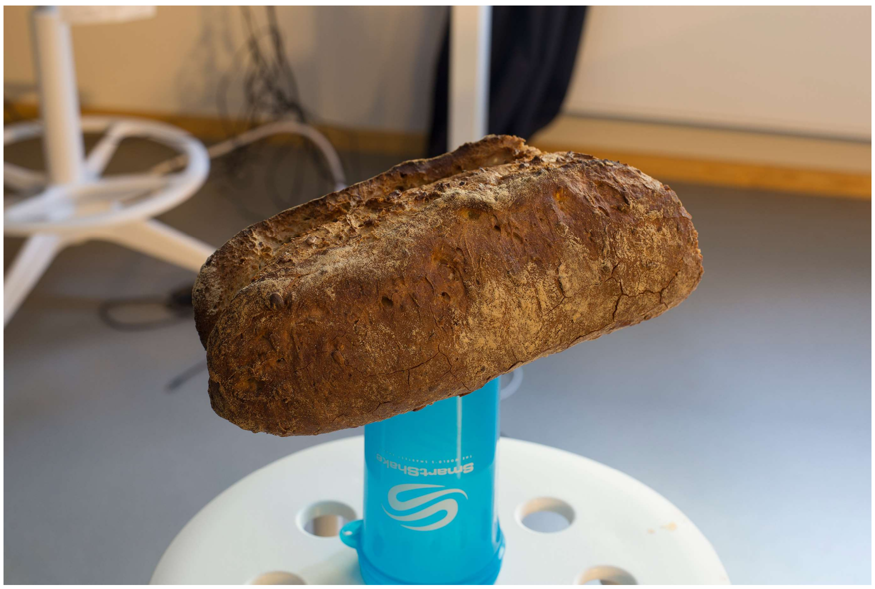
The following table provides an overview of the results of the three programs for this case.

TABLE 3. Case 1 results

SOFTWARE	Meshroom	Agisoft Metashape	RealityCapture
IMAGES USED	109/160	109/160	109/160
PROCESSING TIME	3 hours and 19 minutes	16 minutes	12 minutes
POLYCOUNT	1.275 Million	455,000	4.8 Million

4.3.2 Case two

The second case used a loaf of bread for the subject. The subject was captured indoors. In this case most of the artificial light was turned off, relying mostly on natural light coming from the window. These are not ideal conditions for indoor photogrammetry, but it was not possible to prepare a better environment. To compensate for the lack of artificial lighting, slow shutter speeds were used (PICTURE 34).



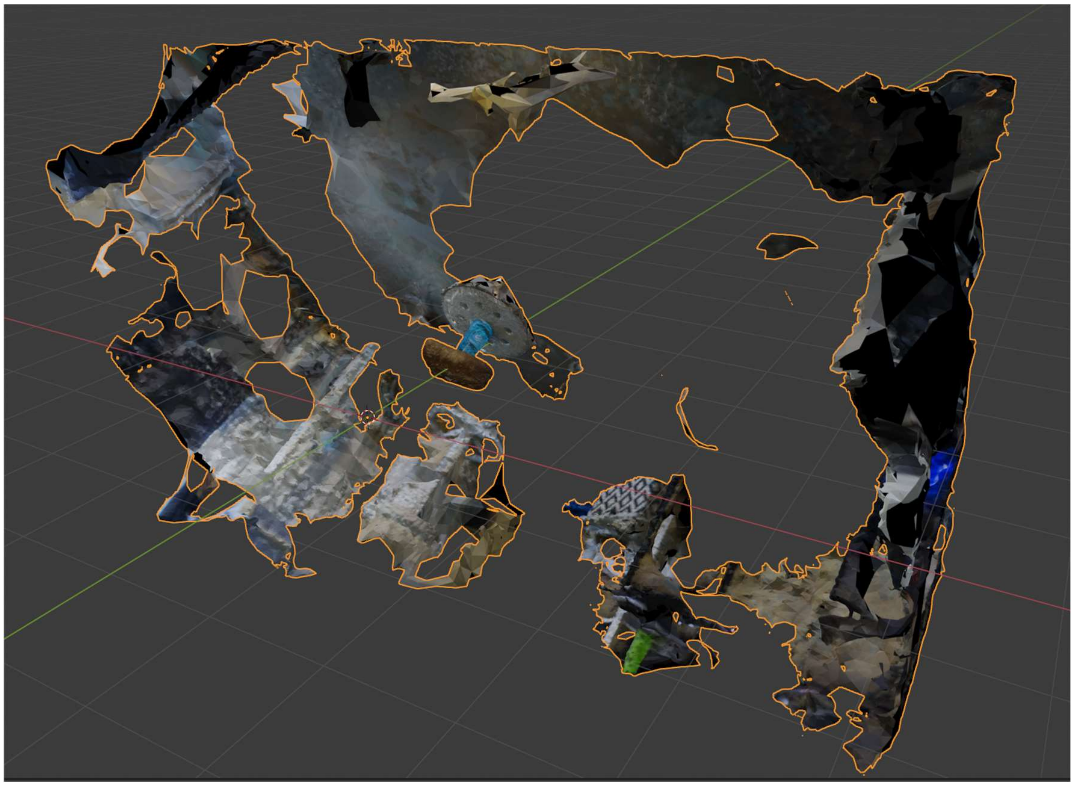
PICTURE 34. One of the images for the second reconstruction

The processing of the images was identical to the first case with the exception of including the camera metadata in the exported images. A total of 147 images were used for this reconstruction. The camera settings during the shooting are shown in table 4.

TABLE 4. Camera settings during case two

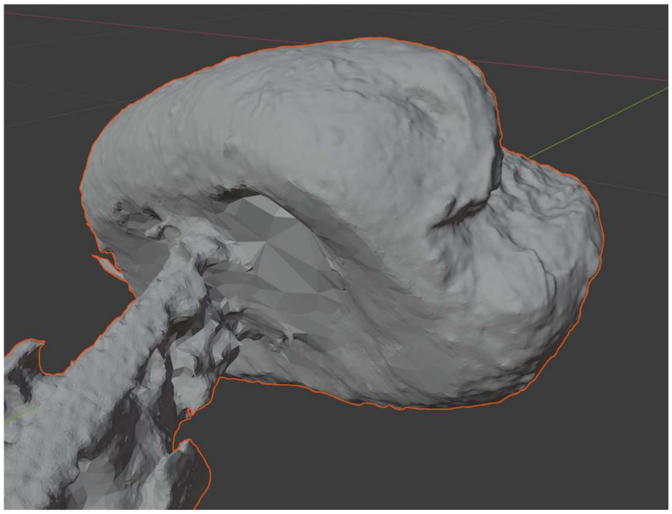
ISO	200
Aperture	f/8
Shutter speed	10 seconds
Metadata	Included in the images

A first look at the model generated by Meshroom shows that the program captured a lot of the detail from the background of the photographs. Meshroom did not discard any of the 147 photographs. The process took around 4 hours and 5 minutes to be completed. The model generated is made of approximately 1.24 million of polygons. However, it is important to consider that part of this number is used to recreate elements of the background and not the subject (PICTURE 35).



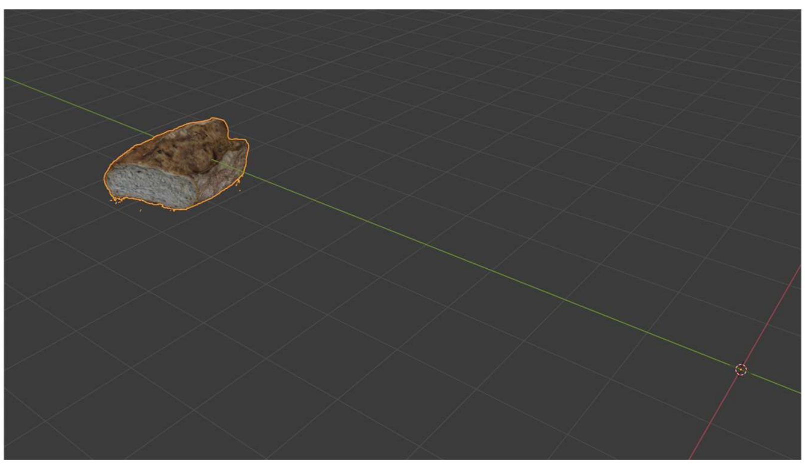
PICTURE 35. First look at the model generated by Meshroom

The model of the subject itself looks clean. There are no strange deformations or floating geometry merged with the subject. Meshroom was even able to close the bottom of the subject despite having little information of the area in the photographs. The lower resolution in this area is expected due to the small number of photographs showing this part (PICTURE 36).



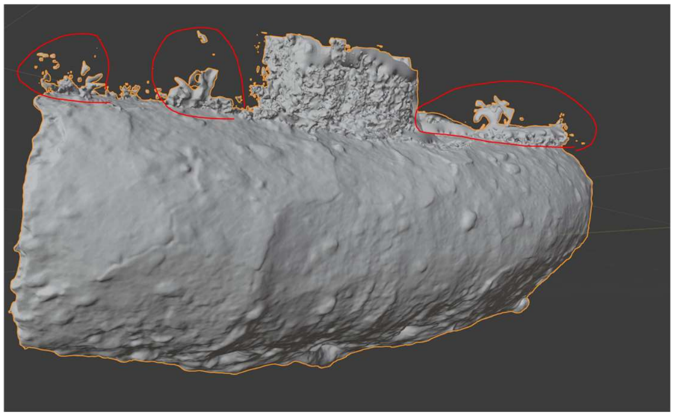
PICTURE 36. Bottom side of the loaf of bread

Agisoft Metashape used all the images for the reconstruction. The process took approximately 42 minutes. In this case, the quality settings for all the process were set to “high”. This includes the alignment of the images, the generation of the point cloud and the mesh reconstruction. The resulting model consists of approximately 1 million polygons (PICTURE 37).



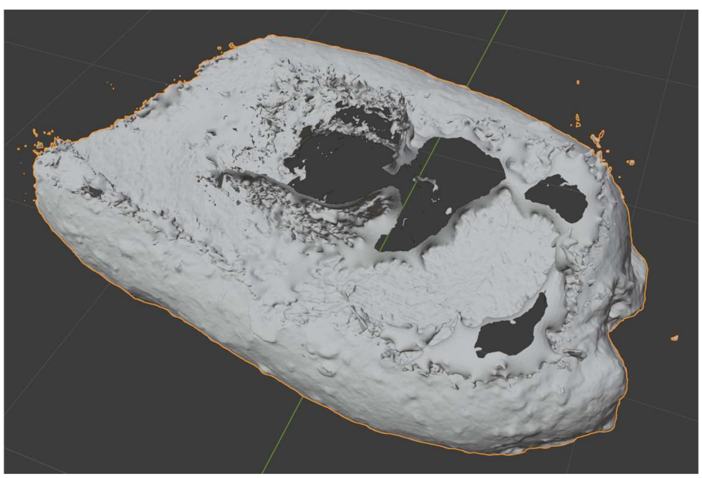
PICTURE 37. First look at the model generated by Agisoft Metashape

A first look at the model shows that Agisoft Metashape did not reconstruct any of the background in the photographs. However, the model presents floating geometry and deformations around the bottom part that would require manual work if the intention were to show the model from all its sides (PICTURE 38). As in case one, Agisoft Metashape offers tools to edit the point cloud that would make fixing the model drastically easier. The reason none of these tools are used in either of the cases is because Meshroom does not support any of these editing tools and the comparison between the programs should be as equal as possible. This applies to RealityCapture as well.



PICTURE 38. The bottom side of the loaf of bread presents deformations

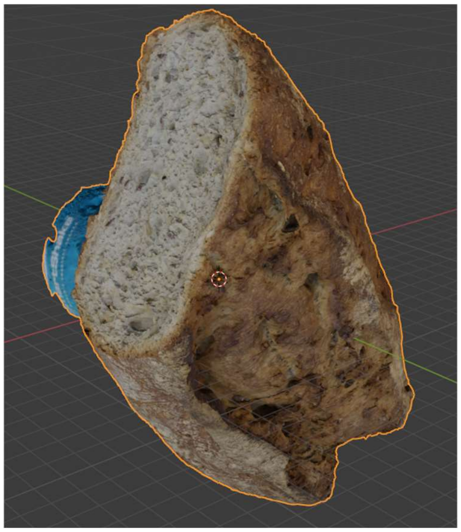
Furthermore, Agisoft Metashape was not able to solve part of the bottom side of the loaf of bread. The result is a model with holes (PICTURE 39). Models with holes are common but problematic for some techniques of optimization. However, this will be explained further in section 5 of the thesis.



PICTURE 39. Holes in the bottom of the subject. Model by Agisoft Metashape

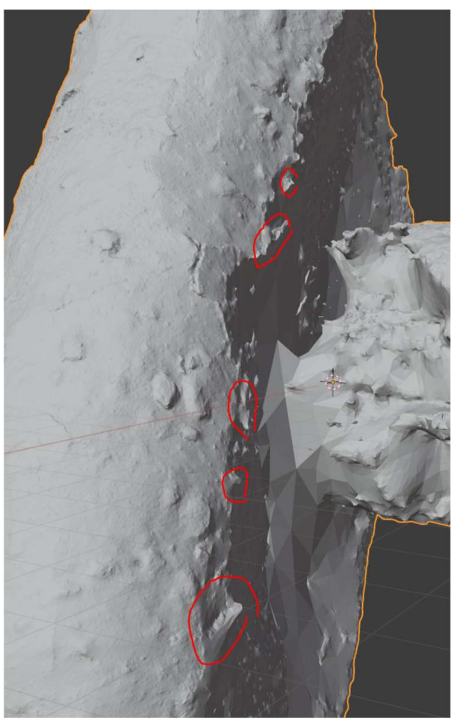
As it was the case in the other programs, RealityCapture used all the photographs. The whole process took 54 minutes approximately. The quality level was set to “high” for this case. The resulting model is composed of approximately 14.23 million polygons.

The model looks clean and has a lot of detail compared to the other models. RealityCapture managed to solve the bottom side of the loaf of bread despite the lack of information of this part of the model in the photographs (PICTURE 40).



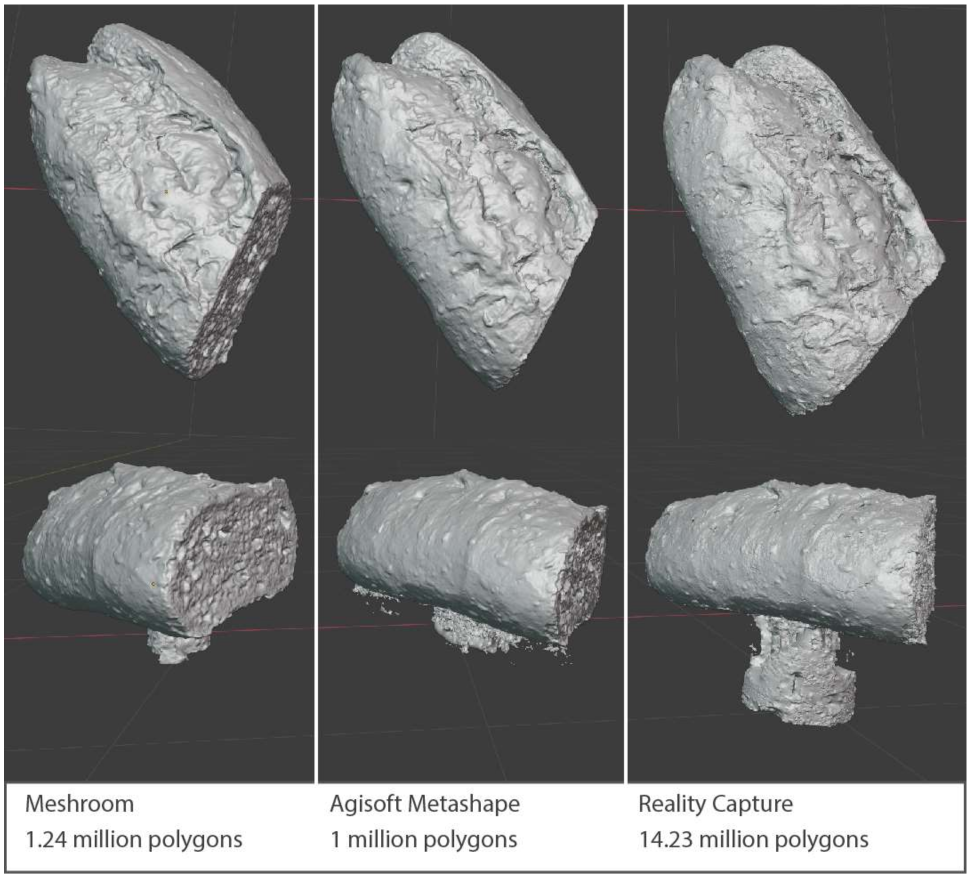
PICTURE 40. First look at the model created in RealityCapture

The model shows some small deformations around its bottom side, but they are not as severe as the ones in the model created in Agisoft Metashape (PICTURE 41).



PICTURE 41. Small deformations in the model generated in RealityCapture

Overall, considering processing time, resolution and minimum mesh problems, RealityCapture produced the highest quality model (PICTURE 42).



PICTURE 42. Models side to side, floating geometry hidden for the screenshot

The following table provides an overview of the results of the three programs for this case.

TABLE 5. Summary of case two

SOFTWARE	Meshroom	Agisoft Metashape	RealityCapture
IMAGES USED	147/147	147/147	147/147
PROCESSING TIME	4 hours and 5 minutes	42 minutes	54 minutes
POLYCOUNT	1.24 Million	1 Million	14.23 Million

4.3.3 Case three

The third's case subject was a boot. The subject was captured indoors, but unlike case one and case two the setup was different. The subject was captured on a turntable with a grey seamless background (PICTURE 43). There were some complications as a high angle position of the camera revealed part of the floor, a surface the programs could use to find detail. Masks were created but since Meshroom does not currently support the use of masks they were not used in the process of any of the programs. It is clear that none of these cases fulfils the ideal conditions described in previous parts of this thesis. However, there is value in assessing how the photogrammetry programs behave when working in difficult conditions to understand and find their limitations.

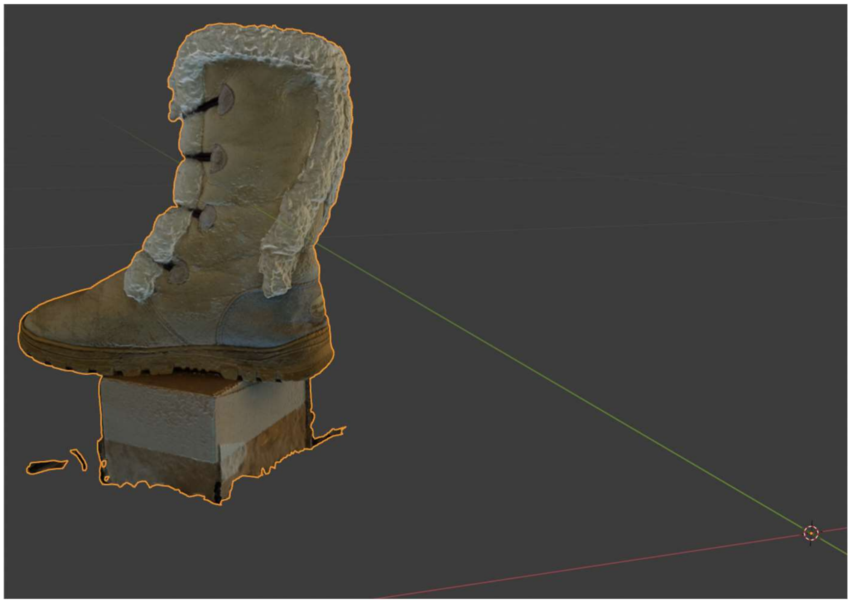
PICTURE 43. One of the photographs used in case three

For this case, a total of 157 photographs were taken. The camera remained stationary, and it only changed position when a higher or lower angle was required. The camera settings during the shooting are available in the following table.

TABLE 6. Camera settings during case three

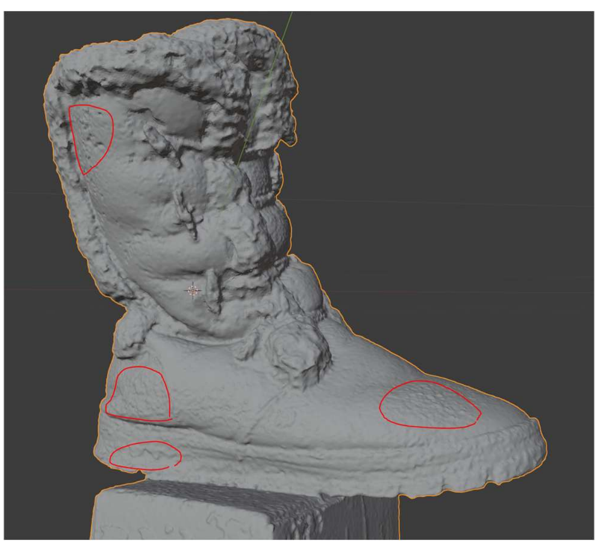
ISO	400
Aperture	f/11
Shutter speed	1.3 seconds
Metadata	Included in the images

Meshroom took around 4 hours to generate the model. It used 156 of the 157 photographs provided. The resulting model was around 1.57 million polygons. The model has almost no floating geometry and no parts of the background were reconstructed (PICTURE 44). This is likely due to the setup during the shooting.



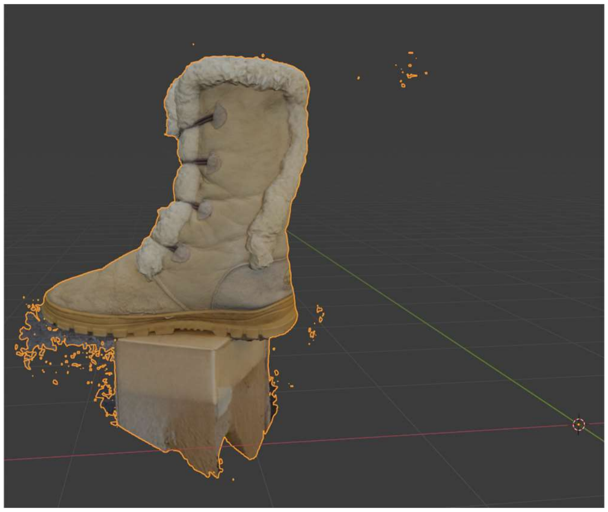
PICTURE 44. First look at the model generated in Meshroom

However, the model seems to have many small surface imperfections (PICTURE 45). For example, the model contains small bumps and indentations that were not part of the real surface of the boot. These imperfections could be fixed by smoothing the surface where needed, using digital sculpting, for instance.



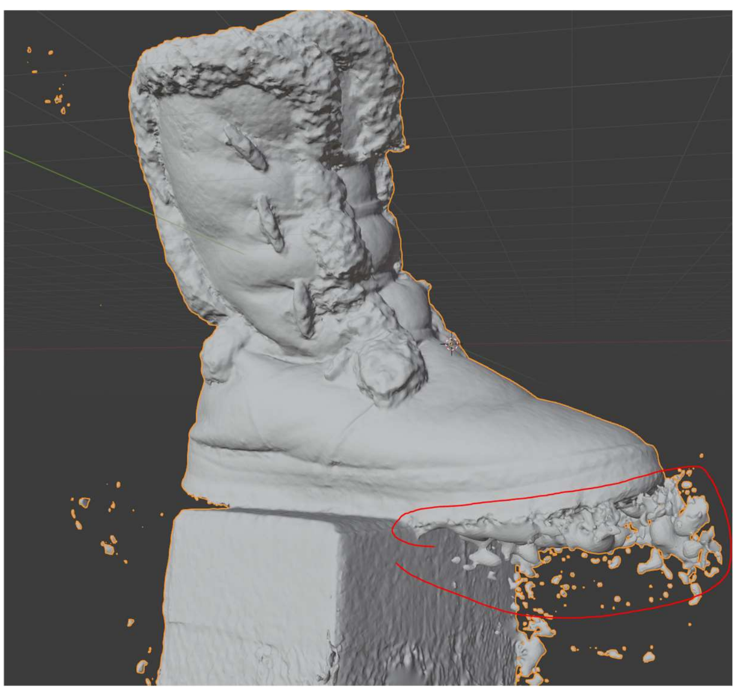
PICTURE 45. Surface imperfections

Agisoft Metashape generated the model in around 27 minutes. The program used all the images. The quality settings were set to “medium”; the model is made of 576,244 polygons. The quality of the surface is better than in the model from Meshroom, more even and consistent throughout the model (PICTURE 46).



PICTURE 46. First look at the model generated in Agisoft Metashape

However, the model shows floating geometry, part of it is merged with the sole of the boot (PICTURE 47). This issue would require a substantial amount of manual work now that the mesh has been generated.

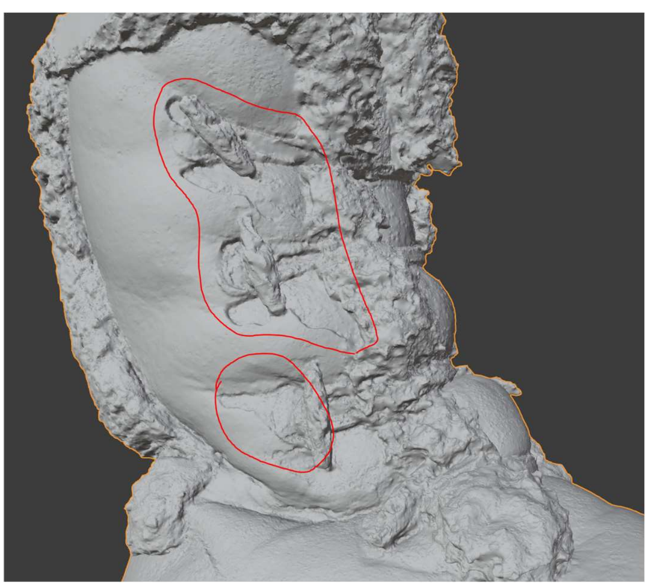


PICTURE 47. Floating geometry merged with the model

RealityCapture generated the model in approximately 19 minutes. All images were used. The model has a polycount of approximately 6.25 million polygons. Overall, the model looks clean, there is no floating geometry compared with the model generated in Agisoft Metashape (PICTURE 48).

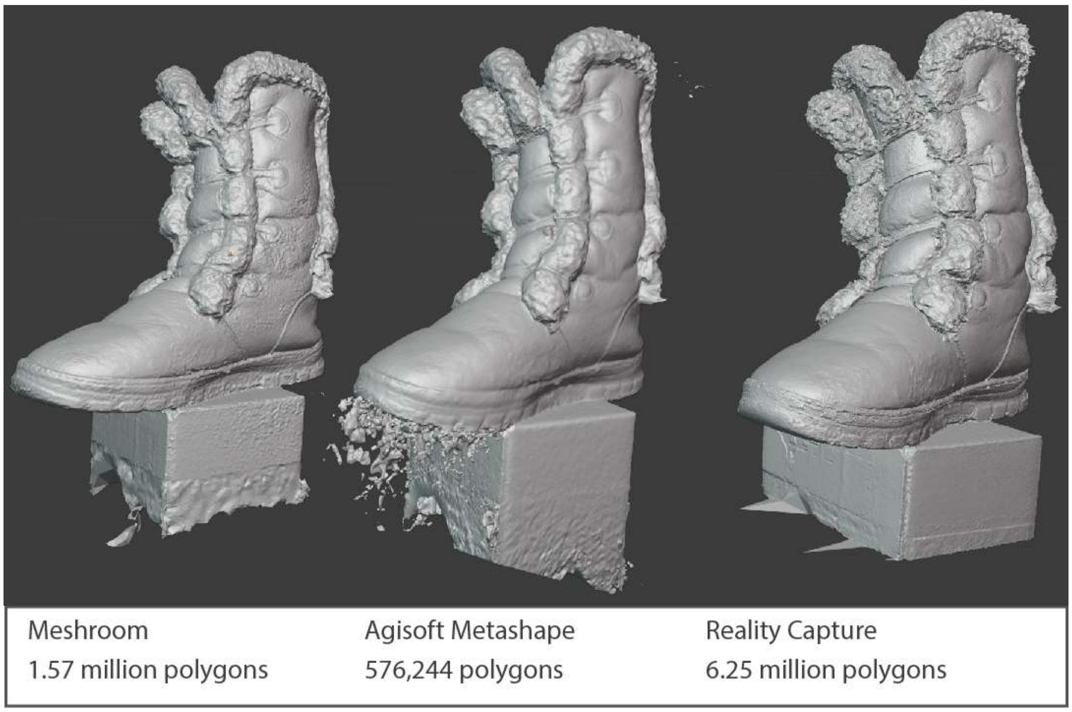
PICTURE 48. First look at the model from RealityCapture

The surface of the model looks more even and consistent unlike in the model generated in Meshroom. However, there is an area of the model that shows surface deformations (PICTURE 49). This particular area looks cleaner in the model from Agisoft Metashape.



PICTURE 49. Surface deformations, model generated in RealityCapture

Looking at the model from Agisoft Metashape and the model from RealityCapture, it seems that Agisoft Metashape was able to better reconstruct the area shown in picture 48. However, keeping in mind the problems described about the model from Agisoft Metashape, the processing times and the resolution, it seems that the best result was provided by RealityCapture (PICTURE 50).



PICTURE 50. Models side by side, case 3

The following table provides an overview of the results of the three programs for this case.

TABLE 7. Results summary for case three

SOFTWARE	Meshroom	Agisoft Metashape	RealityCapture
IMAGES USED	156/157	157/157	157/157
PROCESSING TIME	4 hours	27 minutes	19 minutes
POLYCOUNT	1.57 Million	576,244	6.25 Million

5 PHOTOGRAMMETRY PART 3: CLEAN-UP AND OPTIMIZATION

5.1 Goals

At this stage of the workflow the photogrammetry software has generated a 3D model. However, the model needs to go through several steps before it can be properly used in a real-time engine such as Unity or Unreal Engine.

Real-time engines have limitations regarding the number of polygons that they can visualize. Scenes that require showing several 3D models might require that those models are optimized to ensure good performance of the game or application (Geometry Best Practices… 2021). For instance, the loaf of bread produced by RealityCapture could be used as part of a scene of a bakery, which would include many other 3D models. The problem is that the model of the loaf of bread is currently 14.23 million polygons. This leaves little to no room for any other 3D models as the model of the loaf of bread uses all the polygon budget for the scene. To solve the issue the 3D model must be optimized.

The first part of the optimization will be decreasing the number of polygons in the model to make it feasible for real-time visualization. This includes cleaning the current model, which will be referred to as high poly model, and generating a lower poly version. To preserve the resolution of the high poly model despite the decrease in polygons, a normal map can be baked and applied to the low poly model. A normal map is a 2D image, commonly used for texturing 3D models. This type of texture affects how the 3D model is lit making the low poly model look like it has the surface details of the high poly model (Understanding Normal Maps 2015).

The second part of the optimization involves working with the textures of the 3D model. Even when the acquisition of the photographs for the photogrammetry process has been done flawlessly, small shadows remain in the model. These shadows are noticeable in areas where the light cannot reach, such as small crevices in the subject's surfaces. In computer graphics, this phenomenon is called ambient occlusion (Pluralsight 2014). The effect is artificially recreated in real-time engines, but if the shadows already exist in the texture of the model the final result may look slightly unrealistic. To solve this, the texture of the model must undergo a process called de-lighting or un-lighting (Lachambre, Lagarde & Jover 2017, 84).

In addition, 3D models that aim for a realistic result tend to use the PBR texturing approach. PBR stands for physically based rendering (McDermott 2021). This method of texturing 3D models ensures a consistent look regardless of the real-time engine or 3D software that visualizes them. PBR texturing uses different 2D textures to define the characteristics of the model.

Some of the common PBR texture maps are an albedo map, a metal map, a normal map and a roughness map. An albedo map indicates the color of the model. A metal map helps the artist define which parts of the 3D model are metallic. The normal map helps to transfer surface detail from a high poly to a low poly model. The roughness map indicates what parts of the model are more reflective and which parts have more of a matte surface (McDermott 2021).

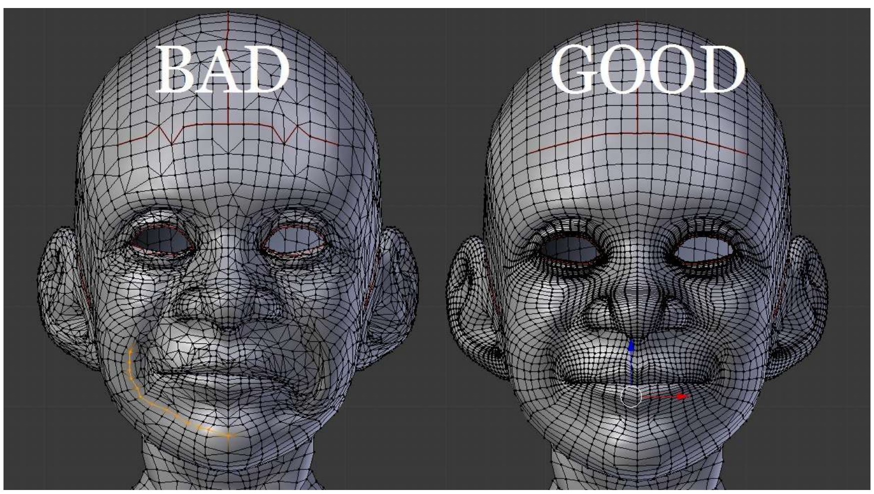
FIGURE 7. PBR texture maps (McDermott 2021, modified)

In case of the 3D model of the loaf of bread there is no need for a metal map since no part of the model is metallic. RealityCapture provided an albedo texture, but it contains some lighting information that needs to be removed. The normal map can be baked from the high poly model, provided by RealityCapture, to a low poly version that will be created in the following sections of the thesis. Finally, creating a roughness map will increase the realism of the 3D model, this will be demonstrated in section of the thesis which covers creating PBR textures.

5.2 Creating a low poly model

There are different options to approach the creation of the low poly model. RealityCapture and Agisoft Metashape offer several tools to clean the mesh and simplify the polycount of the model. However, these simplification tools reduce the number of triangles of the mesh rather than creating a new mesh with cleaner topology.

“Topology is the organization, flow and structure of vertices/edges/faces of a 3D model.” (Danan 2016). This thesis has talked mostly about triangles; however, triangles can create shading problems and strange deformations when animated. Quads are the ideal topology when editing the mesh is required. They also offer the best results when the mesh is deformed during animations. Thus, exploring an alternative method to generate a quad-base mesh is relevant (PICTURE 51).

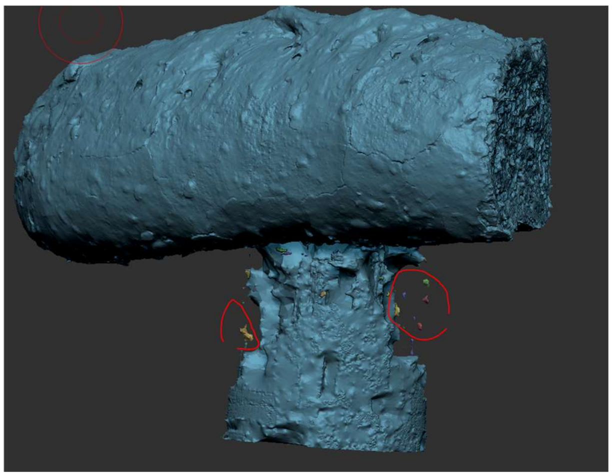


PICTURE 51. Good topology speeds up the editing of the mesh (Danan 2016)

The following workflow to generate the low poly model is based on Alex Alvarez’ Workflow (Gnomon 2016). Zbrush will be the main software used to generate the final mesh. Zbrush can handle high numbers of polygons without slowing down, and it has tools to automatically generate a new mesh with acceptable topology-Furthermore, it offers tools to easily remove floating geometry and close holes in the model.

The first step of the workflow is to import the model into Zbrush. An initial clean-up is required to eliminate the floating geometry and the pedestal that was used to hold the loaf of bread.

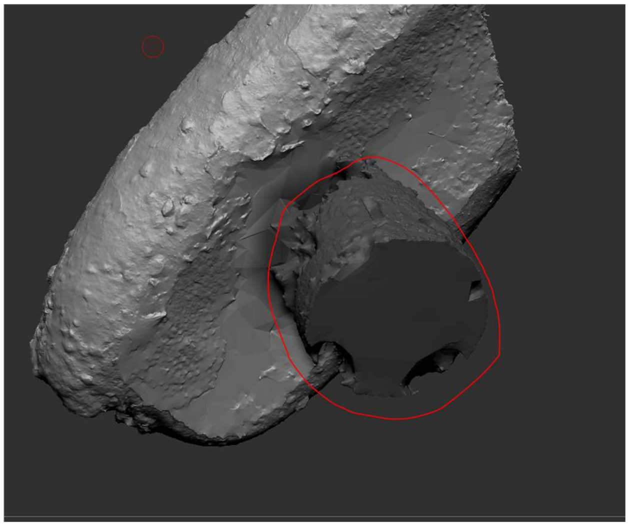
To remove the floating geometry, Zbrush can assign groups of polygons based on whether they are connected (PICTURE 52). This function is available in the menu “polygroups” on the right side of the screen. In the menu there is a button named “assign polygroups.” Zbrush also assigns automatic colors to polygroups so the user can see where a polygroup ends and another begins. In this case the floating geometry has a different color than the rest of the model.



PICTURE 52. Polygroups are shown in different colors

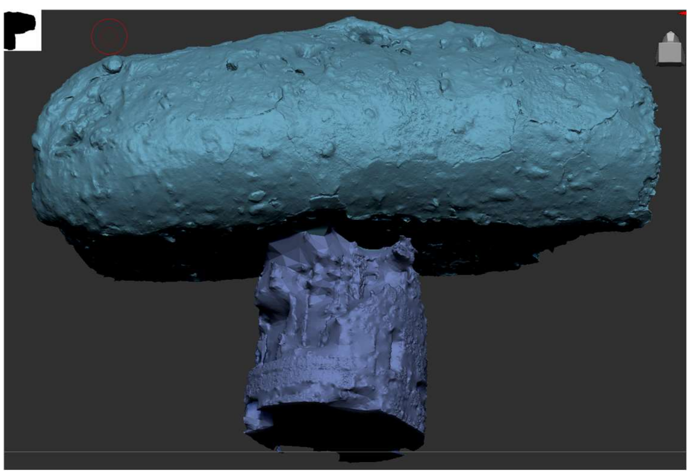
By pressing “Shift + click” on the bread model, Zbrush isolates its polygroup, hiding the rest of the polygroups, which in this case are the floating geometry. The floating geometry can be deleted by using the command “delete hidden” that can be found in the menu “geometry”, and subsequently in the sub-menu “modify topology”.

The next step is to delete the polygons that make the pedestal underneath the bread model. Zbrush has a masking tool, by pressing and holding the Ctrl key, the user can paint over the polygons that need to be masked. Zbrush also offers a lasso tool for this (PICTURE 53).



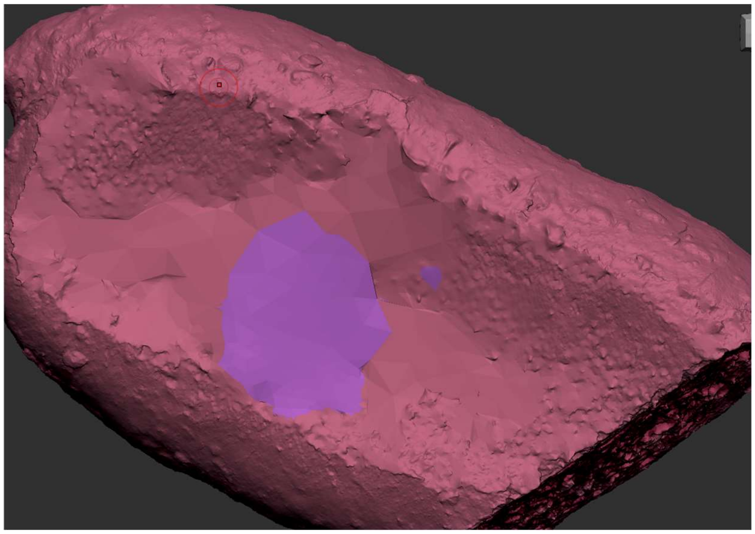
PICTURE 53. The pedestal underneath the bread is now masked

There is a function in the menu “polygroups” that creates polygroups based on masked parts in the model, this function is called “group masked”. Now there is a rough separation between the pedestal and the model of the bread (PICTURE 54).



PICTURE 54. The pedestal and the bread model are in separated polygroups

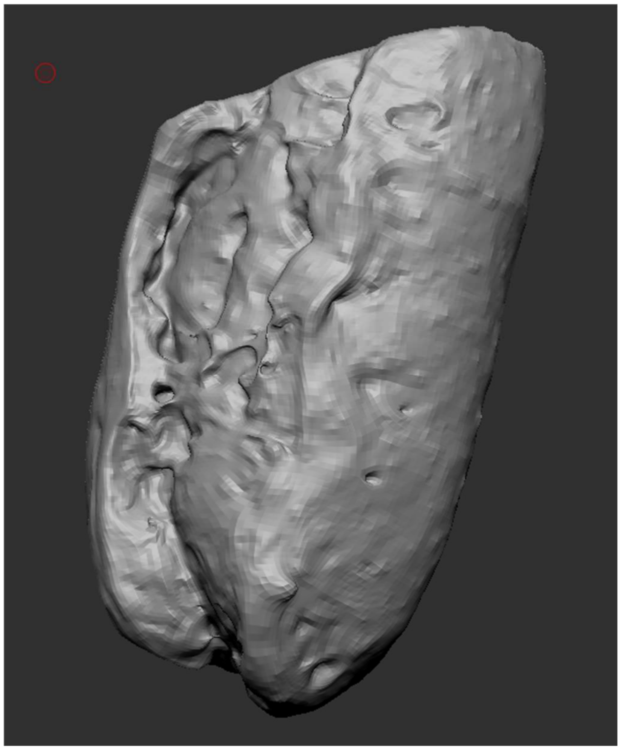
By pressing “Shift + click” over the bread model as was done in the previous step, Zbrush will hide the pedestal polygroup. Now it is possible to delete the hidden polygroup as it was explained in the aforementioned steps. Deleting the pedestal polygons will leave a hole in the model. Fortunately, Zbrush offers a function for closing holes automatically, it can be found in the sub-menu “modify topology” which can be found in the menu “geometry” (PICTURE 55).



PICTURE 55. Holes closed with new geometry

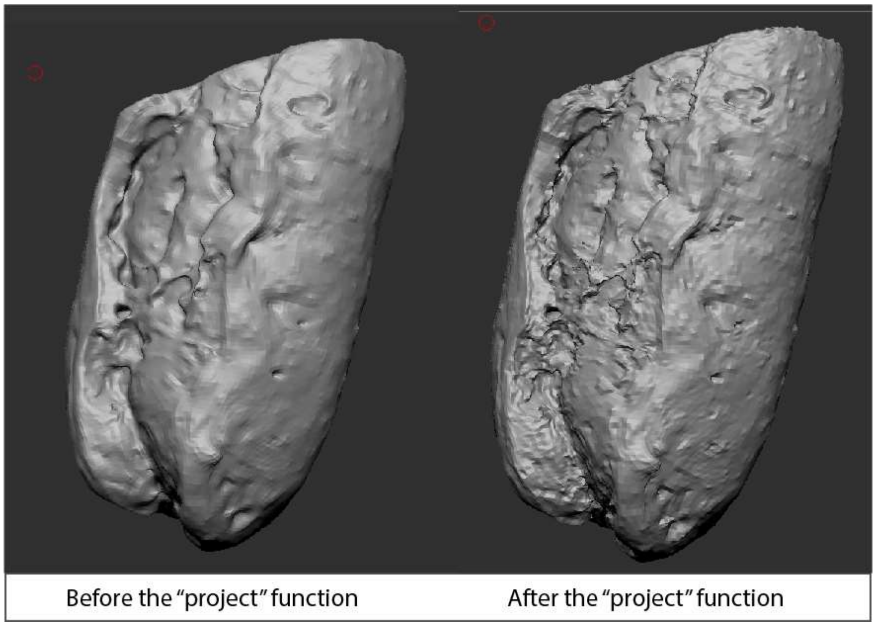
Now that the model has been cleaned from undesired geometry, it is time to create the low poly model. In a Zbrush scene models are called “tools.” When the model of the bread was imported a new “tool” was created. Different objects or models can exists separated within a tool. Each separated mesh is called “sub-tool”. In the current scene there is only the bread model in the sub-tool menu. To create the low poly model the bread model needs to be duplicated from the sub-tool menu. The copy of the model will be used as a guide to create the low poly model.

“Zremesher” is a function that creates quad-based topology. Quads are essential in digital sculpting, especially when working with low poly models as they provide better results and react better to deformation. The Zremesher operation is often used to decrease the number of polygons of a model. The model’s geometry is replaced by quad-based geometry automatically generated by Zbrush. In case of the bread model, the intention is to replace the triangle-based geometry for uniform quad-based geometry, as well as greatly decreasing the number of polygons (PICTURE 56). The Zremesher option can be found in the “geometry” menu.



PICTURE 56. Low poly model, quad polygons generated with Zremesher

The Zremesher operations tends to smooth the model in a way that some surface details might be lost in the process. While some small details will no longer be visible, as the lower polycount cannot represent them, surface detail can be recovered by using the “project” operation. The project operation can be found on the sub-tool menu. To perform this function, it is necessary to have both the high poly model and the low poly model visible. With the low poly model selected in the sub-tool menu the “project all” button in the project menu can be pressed. This function will change the position of the vertices of the low poly model to match closely the position of the vertices of the high poly model. To put it simply, Zbrush will deform the low poly model’s surface to approximate the high poly model’s surface (PICTURE 57).



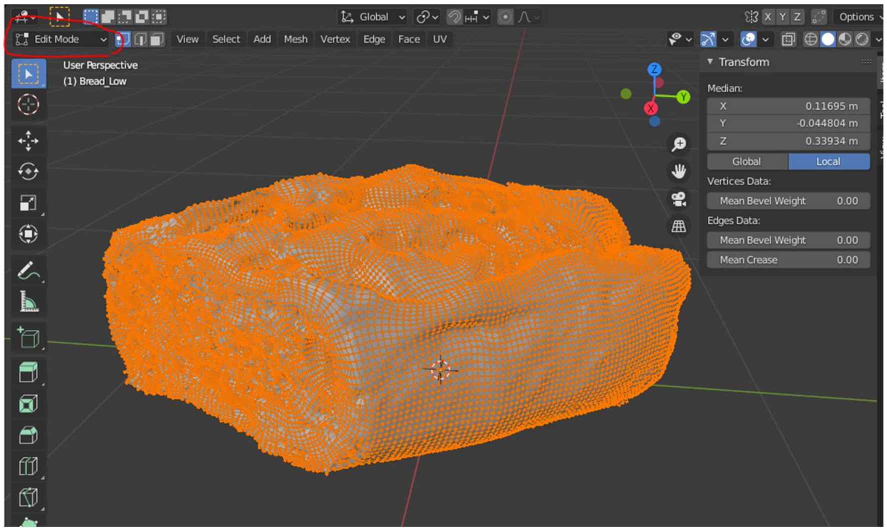
Picture 57. The low poly model before and after the "project" function

The low poly model is approximately 30,000 quad-polygons. If this number were to be converted to triangle polygons the result would be around 60,000 polygons. This number might be high depending on the purpose, but It is always possible to generate a lower poly model using the Zremesher.

At this point the low poly model has been generated, both the high poly and the low poly model should be exported from Zbrush to ensure they overlap as closely as possible, this is important for the creation of the normal map.

Before the normal map can be generated, the low poly model needs UV coordinates. "UVs are two-dimensional texture coordinates that correspond with the vertex information for your geometry." (Pluralsight 2014). UV coordinates are necessary to apply 2D textures on a 3D model. Many 3D programs offer the possibility of creating UV coordinates either manually or automatically. In this case Blender will be used for this task. Since the intention of this thesis is to limit manual work as much as possible the UV coordinates will be generated using the "smart UV project" function in Blender.

Once the model is imported it is necessary to switch from "object mode" to "edit mode". This can be done by selecting "edit mode" from the drop-down menu on the 3D viewer of the program (PICTURE 58).

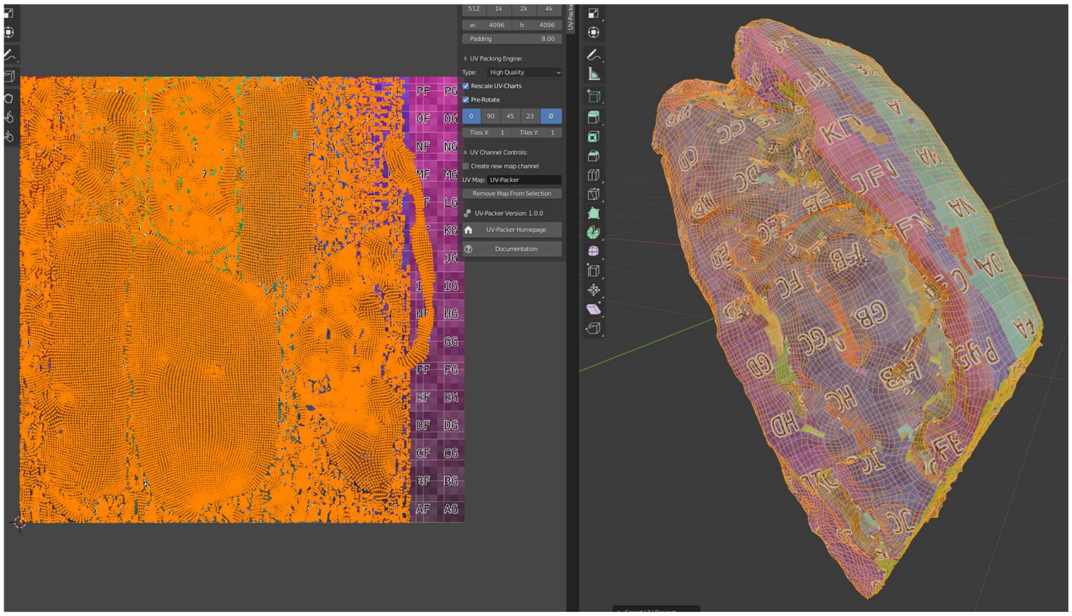


PICTURE 58. Edit mode

Now that “edit mode” is enabled it is possible to select all polygons by pressing the “A” key on the keyboard and select “smart UV project” from the UV menu, this menu is visible in the upper part of picture 55.

Blender’s algorithm for creating UV coordinates could be more efficient using the UV space. The UV space is the space the 2D projection of the 3D model occupies. Using as much of this space is important because empty space in the UV coordinates is space that the 3D model is not using for texturing. Simply put, a 3D model that uses the UV space efficiently will make better use of the 2D textures than a model that has inefficient UV coordinates.

While Blenders’ default algorithm for creating UV coordinates is inefficient, addons can override it which improve results. This is why Blender was chosen to create the UV coordinates of the low poly model. An addon named UV-Packer optimizes the use of the UV space (PICTURE 59).

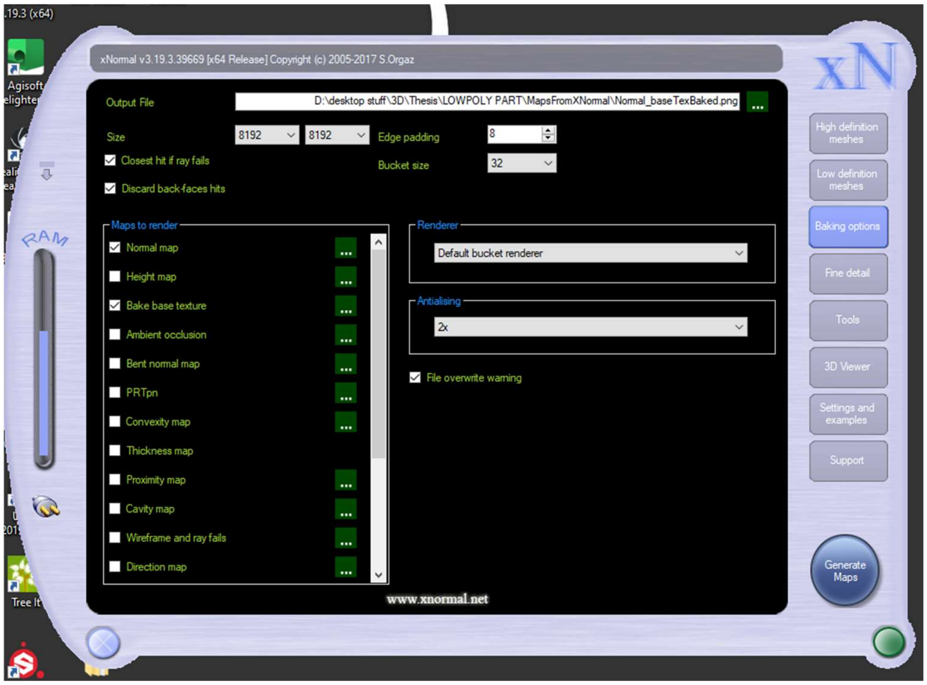


PICTURE 59. UV coordinates created and packed with UV-Packer

The low poly model is now ready for creating the normal map. This will be done using a different program named xNormal. xNormal was chosen as it is easy to use and provides good results consistently.

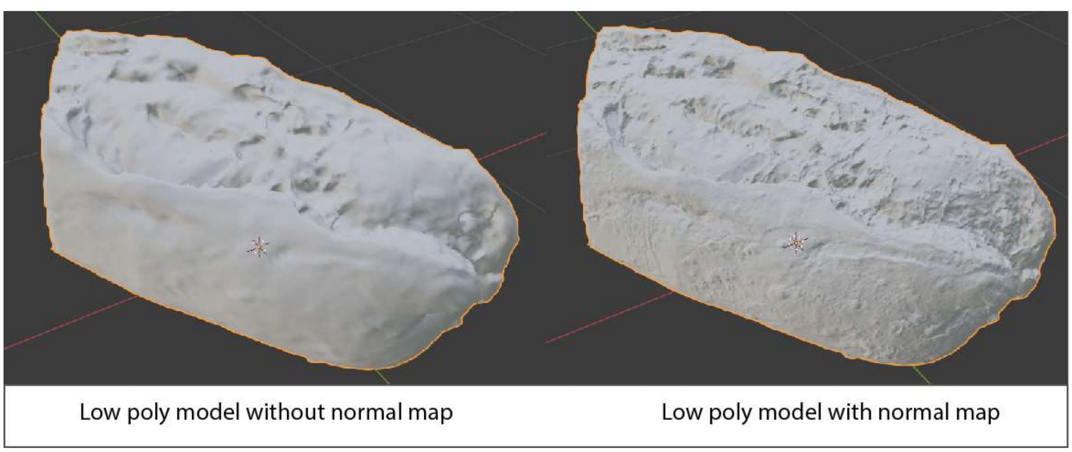
Once the low poly model is exported from Blender, the high poly and low poly model can be imported into xNormal. This program provides a straightforward workflow.

The high poly model is imported in the first tab on the right and the low poly model is imported in the second tab. The tab named “Baking options” allows the user to specify the output resolution of the textures and the 2D maps that will be generated (PICTURE 60). In this case, xNormal will generate the normal map to recover surface detail from the high poly model, and an albedo texture to transfer the color information of the high poly model to the low poly model.



PICTURE 60. The baking tab from xNormal

This would complete the process. A low poly model with a normal map to imitate the surface detail of its high poly version and the color information that was extracted from the photographs during the photogrammetry process. Picture 59 shows the difference between the low poly model without a normal map and with a normal map applied.



PICTURE 61. Surface detail comparison

5.3 Refining and creating PBR textures

Addressing the textures of the low poly model is the last part of the workflow. As it was mentioned in previous sections of the thesis, the texture generated from the photographs by the photogrammetry software, the albedo texture, still contains small shadows that should be removed. Some of these shadows are directional while others are located in crevices in the surface of the model where the light cannot reach.

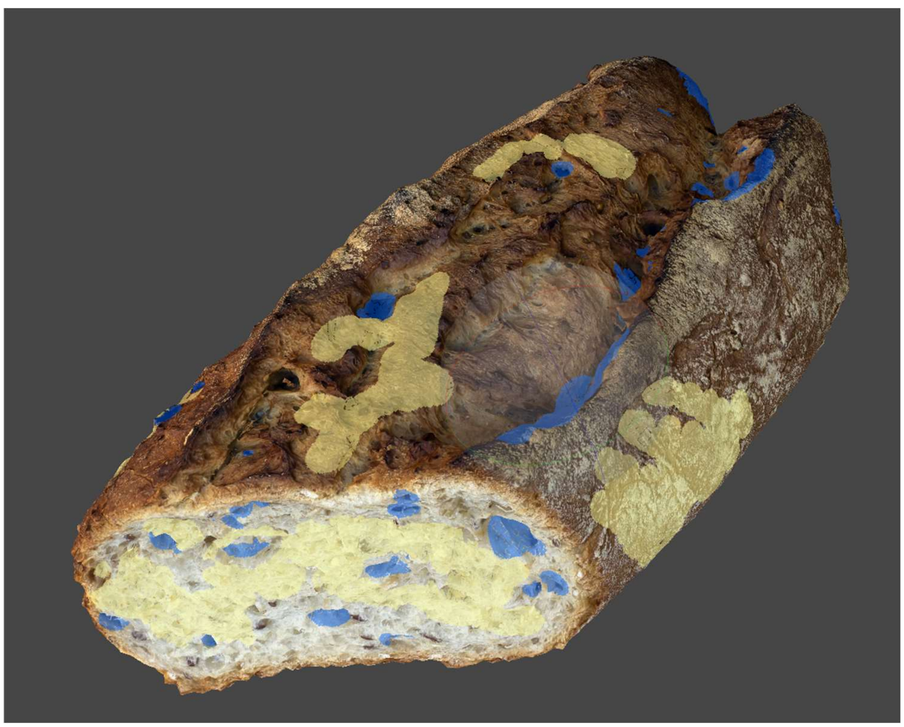
The better the light conditions are at the time of shooting the photographs, the easier refining and correcting the shadows will be. The loaf of bread did not have the best lighting conditions during shooting. However, improving the texture is still possible.

There are different approaches to removing or minimizing the shadows in the texture of the final model. One of these approaches suggests creating two sets of photographs, as it was mentioned in previous sections of the thesis. One of the sets is edited in the photographic software to aid the photogrammetry program during the reconstruction of the 3D model, while the other set is used to generate the texture of the 3D model. The second set focuses on flattening the images by darkening its highlights and adding brightness to its shadows (PICTURE 62). Thus, the contrast of the images is minimal, minimizing shadows and highlights (Kuzmin 2018).

PICTURE 62. The same photograph edited for different purposes

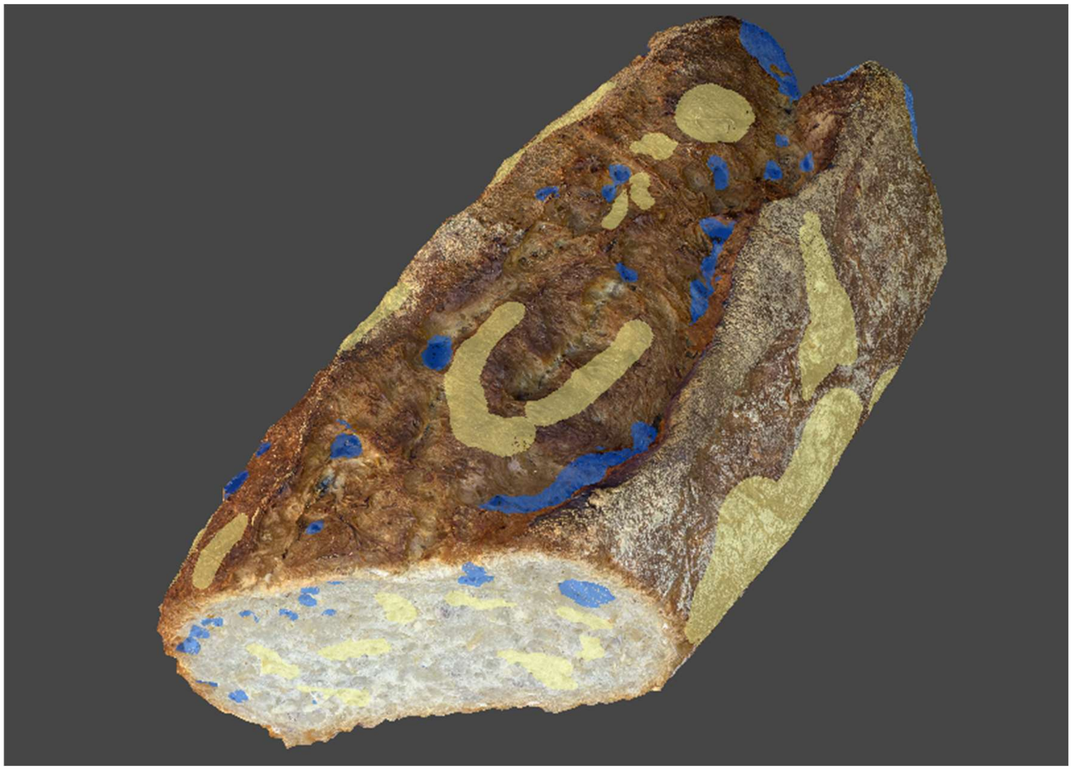
There are situations in which this approach does not work or does not provide the best results. Fortunately, Agisoft offers a free standalone program meant to help delighting the texture of a 3D model with little manual input required.

Agisoft Texture De-lighter uses heuristics in its algorithm to attempt to remove cast shadows and ambient occlusion shadows from the texture of the model. The program requires the user to provide input to work. The user is asked to roughly paint brush strokes over shadowed areas in the model, as well as paint areas that are clearly in light (PICTURE 63). This is done by using two different brushes, a blue brush for shadowed areas and a yellow brush for light areas. While the process might sound tedious, the program does not require the user to be precise and only some samples of these areas are needed, making the process fast and painless (Agisoft Texture De-lighter… 2020).



PICTURE 63. The blue brush roughly marks areas in shadow while the yellow brush marks lit areas.

Once the relevant areas have been marked the button “remove shadows” will start the process. Agisoft Texture De-lighter offers two main operations, “remove shadows” focuses on removing cast shadows caused by the lighting conditions in which the photographs were taken. In this case, the program was not able to remove all the shadows from the texture of the bread model, however it helped minimize the most noticeable ones (PICTURE 64).

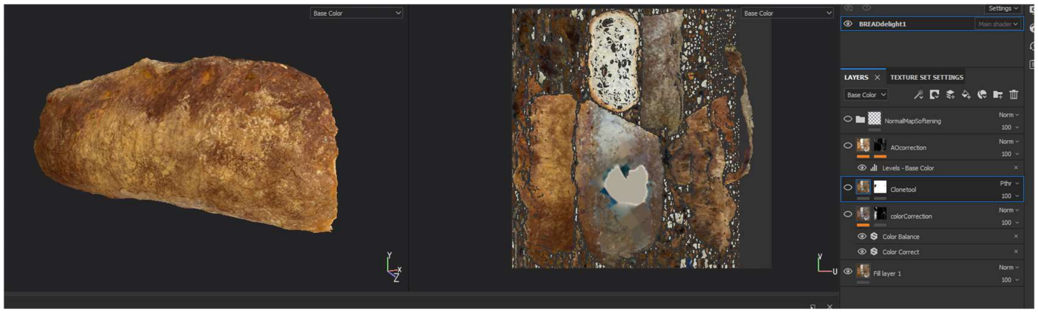


PICTURE 64. The texture has less contrast as many of the shadows have been removed.

The second operation focuses on removing shadows caused by ambient occlusion. However, in this case the operation failed to provide acceptable results. The texture showed color inconsistencies that could not be fixed despite the instructions provided by their tutorial (Agisoft Texture De-lighter… 2020).

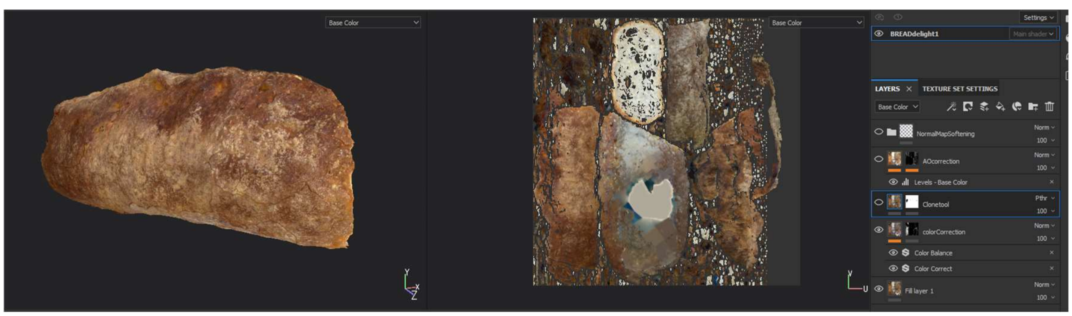
The results of this second operation were discarded and only the results from the first operation were exported as a new albedo texture. In order to improve the texture further, the low poly model, the albedo texture and the normal map texture were imported to Substance painter, a software specialized in texturing 3D models.

Substance painter offers several tools for creating and editing texture sets for 3D models. In this case it was used to correct the color of the albedo texture, as well as removing part of the shadows caused by ambient occlusion. The albedo texture that was generated through the photogrammetry process presented tints of yellow that were not originally part of the color of the loaf of bread. This tint is likely due to the suboptimal lighting conditions in which the photographs were taken (PICTURE 65).



PICTURE 65. The albedo texture looks slightly too yellow

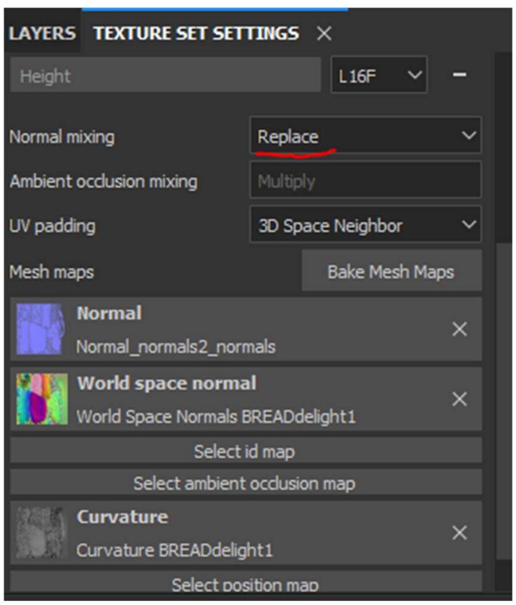
Fortunately, Substance Painter offers several filters to correct colors. The program follows a layer-based workflow that allows a non-destructive approach, which means all changes can be easily edited or reverted. To correct the color, the albedo texture was duplicated and edited in another layer using the filters “color correction” and “color balance” (PICTURE 66). A black and white mask was used to indicate substance painter which areas needed to be more heavily corrected than others.



Picture 66. The yellow tint is no longer present

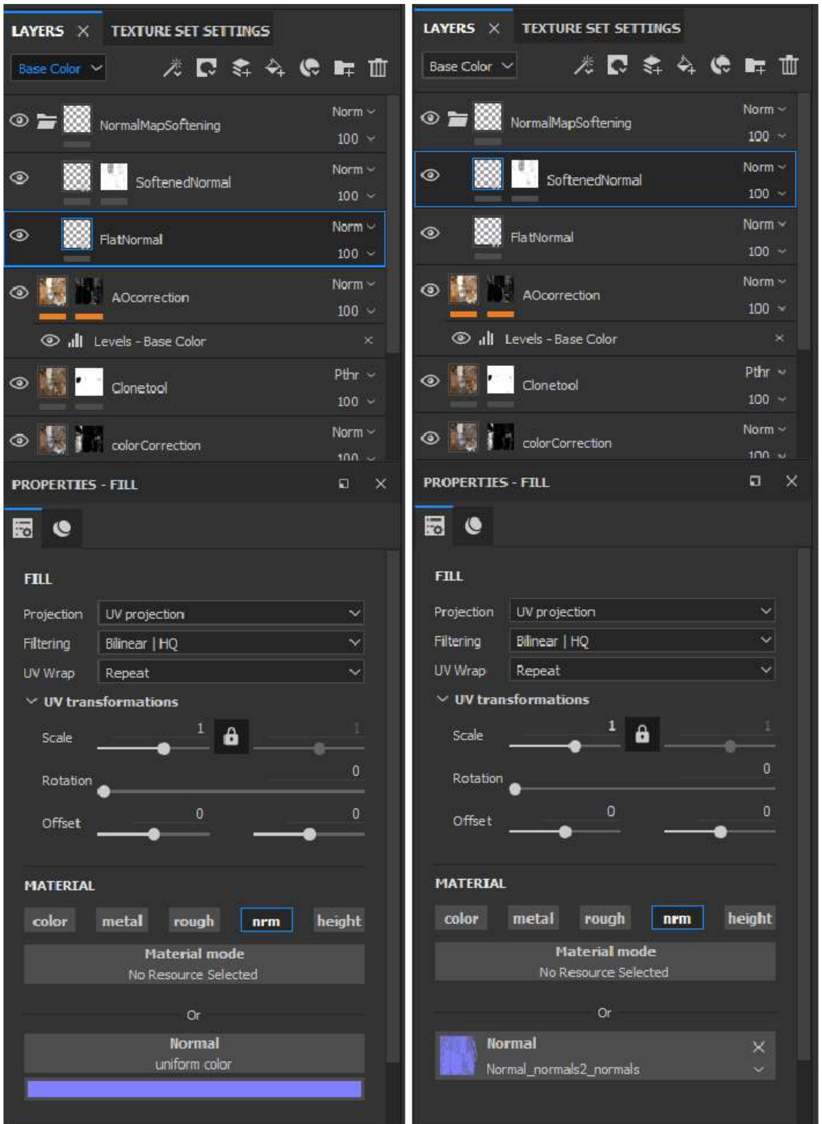
Substance Painter offers the possibility of baking texture maps based on the geometry of the model. An ambient occlusion map is a 2D texture that re-creates the shadows originated in small crevices and corners where light has difficulty to reach in the real world. It is a black and white image that, when inverted, can be used as a mask to indicate in which parts of the models there will be shadows caused by this effect. To remove these shadows, a brighter duplicate of the albedo texture was created and was applied based on the inverted ambient occlusion mask.

Additionally, the normal map of the 3D model was softened around the sliced part of the loaf of bread. The strong normal map effect in this area gave the model an unrealistic look with hard edges. Substance painter allows the user to import texture images and apply them to the 3D model. The normal map was generated in xNormal as shown in the previous section of the thesis. This map was imported and applied to the 3D model and the option “replace” was selected in the drop-down menu named “normal mixing” (PICTURE 67). This option is what allows the user to edit the baked normal map (Normal Map Painting 2021).



Picture 67. Texture set settings in Substance Painter

Two layers were created to soften the normal map effect. One layer with the normal map uniform color, corresponding to a flat surface and a second layer with the baked normal map loaded as an image (PICTURE 68). The second layer made use of a black and white mask to partially hide the baked normal map in the desired area.

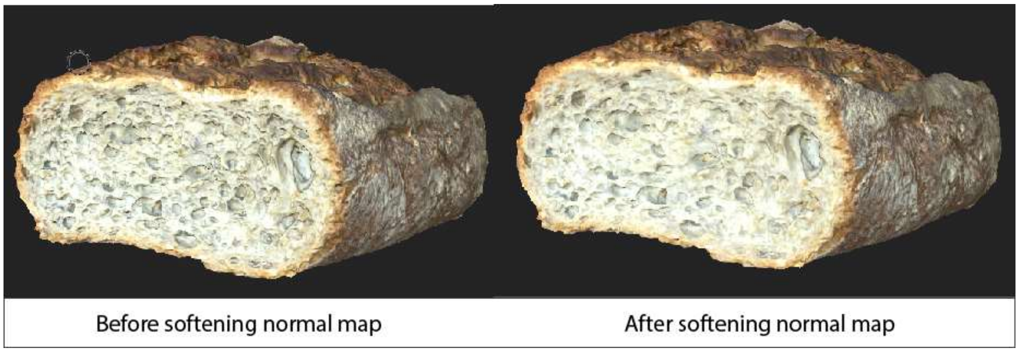


Flat normal layer

Layer with partially hidden baked normal map

PICTURE 68. Layer setup

The resulting combination conserved the strength of the normal map effect on most of the model except for the area indicated in the mask of the second layer, which was softened (PICTURE 69). Editing a normal map is possible in other image editing programs. However, the advantage of Substance Painter is that it allows to see results of the changes immediately.

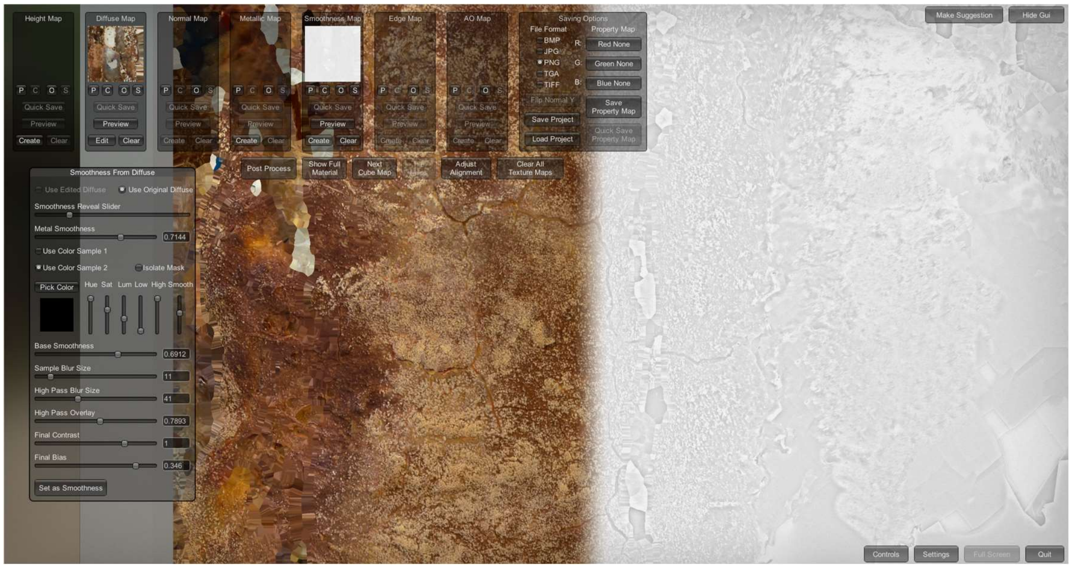


PICTURE 69. Normal map changes

Three texture maps were exported from Substance painter, the corrected albedo, the improved normal map and a baked ambient occlusion map. The most relevant texture map left at this point was the roughness map. The roughness map is a black and white image that indicates how rough the surface of the 3D model is. Rougher surfaces will diffuse the light they receive more than smoother surfaces (The PBR Guide… 2018).

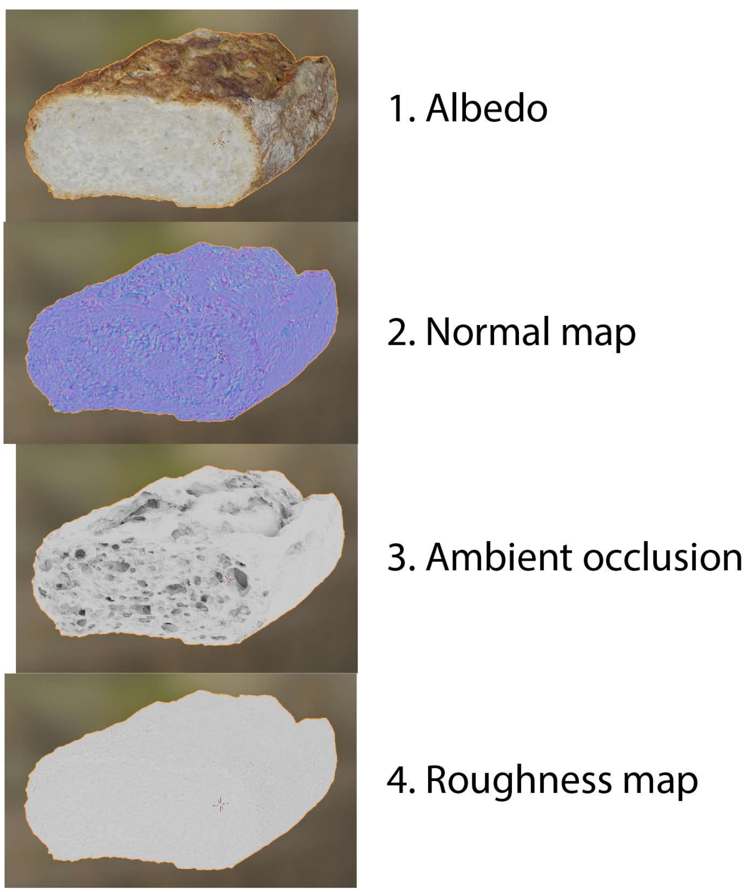
The powdery surface of the flour coating the bread is rougher than the uniform crust of the bread that lays underneath, this is the principle that was kept in mind to generate the roughness map. In order to limit the manual work, the roughness map was generated in a program named Materialize, created by Bounding Box Software. This free software is able to generate texture maps from a photograph or in this case, an albedo map.

Materialize does not generate a roughness map directly, instead it generates a smoothness map. The smoothness map is the inverted version of a roughness map. The program allows the user to edit the result of the generation process through several sliders. These sliders allow the user to invert the smoothness map to obtain the roughness map (PICTURE 70).



PICTURE 70. Generating the roughness map of the 3D model

With the roughness map exported, the process of generating the most important PBR textures was completed. The albedo texture, normal map, ambient occlusion map, and roughness map are shown in picture 71.



PICTURE 71. The complete texture set applied to the 3D model

6 DISCUSSION

Photogrammetry is an established technique that makes generating photorealistic 3D models easier and faster. Its applications go beyond the entertainment industry and its workflow and requirements make it relatively unexpensive, scalable and potentially automatable. Companies such as Quixel Megascans and game studios such as DICE have demonstrated its effectiveness in the creation of real-time models for videogames. The available documentation allows interested parties to learn and understand the basis of the technique with ease. Real-time engines such as Unity and Unreal Engine are among the publishers of said documentation, ensuring it is reliable and clear.

The information available regarding the acquisition of the photographs explains how and why it is important to maximize the quality of the images, as they are closely related to the quality of the final model.

The aim of this thesis was to find a reliable pipeline to generate 3D models through photogrammetry. The comparisons performed using different photogrammetry software give a clear picture of the most popular program options at the time of writing. The results of the nine processed models showed what can be expected of the programs.

The optimization and texture creation part of the thesis successfully demonstrated how it is possible to generate a lower polygon version of the generated model. A 3D model that fulfils the requirements and respects the limitations of current real-time engines. Furthermore, it showed how it is possible to accomplish this task with relatively low amounts of manual work.

If a similar project were to be done in the future, more attention would be put in the lighting conditions during the acquisition of the photographs. As it was demonstrated during the final texturing process, correcting undesired color variations caused by light sources during the acquisition of the images is possible. However, less manual work would have been necessary if the lighting conditions during the shooting of the subject had been optimal.

Figure 8 summarizes the pipeline demonstrated throughout this thesis, including the main tools used to perform the different tasks.

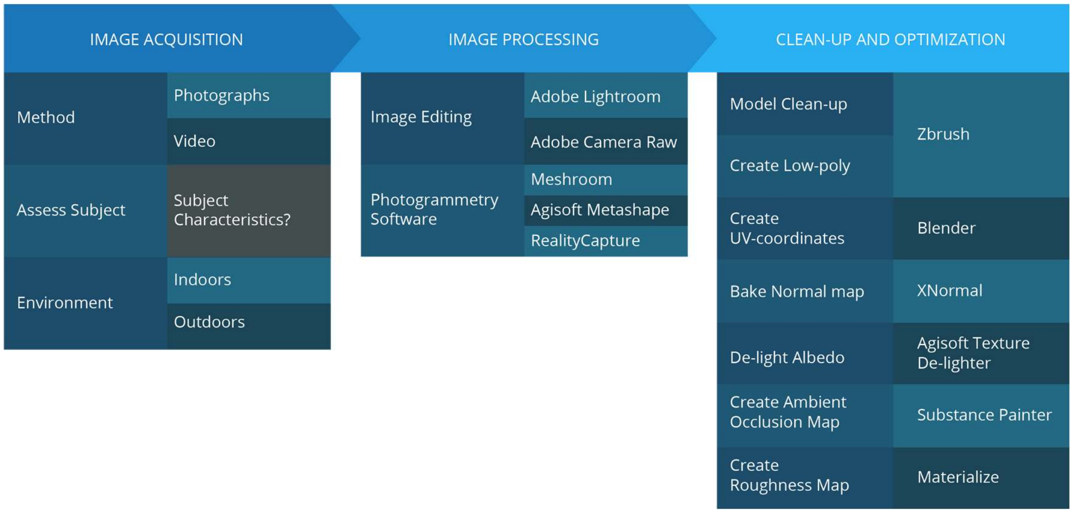


FIGURE 8. Photogrammetry pipeline

REFERENCES

Ad&D 4D. The Grandfather of Europe. 3D Model. Accessed on 17.03.2021. https://sketchfab.com/3d-models/the-grandfather-of-europe-granada-spain-decf3d333d9d4345a35ac5d524d71e1f

Agisoft Metashape User Manual Version 1.5 2019. Read 15.03.2021. https://www.agisoft.com/pdf/metashape-pro_1_5_en.pdf

Agisoft Texture De-Lighter: General Workflow. 2020. Tutorial. Read 22.03.2021.

Another World Studios. Photogrammetry: Raw vs Tiff16 vs JPG. Blog. Read 11.03.2021. http://www.anotherworldstudios.com/photogrammetry-data-comparison-nef-vs-tiff16-vs-jpg/

Bandera, M. 2015. Japanese Barber Shop Photogrammetry Raw Scan. 3D Model. Accessed on 05.03.2021. https://sketchfab.com/3d-models/japanese-barber-shop-photogrammetry-raw-scan-87fcb67998cf4c6fb6012c0c547f2bc4

Berkenfeld, D., Black, D., Corrado, M., Silverman, L. 2021. Understanding Focal Length. Blog. Read on 19.03.2021. https://www.nikonusa.com/en/learn-and-explore/a/tips-and-techniques/understanding-focal-length.html#

Busby, J. 2016. 3D Scanning Reflective Objects with Photogrammetry. Blog. Read on 11.03.2021. https://80.lv/articles/full-photogrammetry-guide-for-3d-artists/

Capturing the Apocolypse. Screen capture of cameras in a rig that focuses on capturing the actor. Ten24 Media n.d. Watched on 14.03.2021. https://ten24.info/apocalypse/

Chua, K.C., Wong, H.C., Yeong, Y.W. 2017. Standards, Quality Control, and Measurements Sciences in 3D printing and Additive Manufacturing. London:Elsevier.

Chumilla, D. 2019. Introducción a la fotogrametría 3D. Online Course. Requires access right. https://www.domestika.org/es/courses/666-introduccion-a-la-fotogrametria-3d/course

Classy Dog Films. 2014. Guide to PhotoScan Part 2: Shooting with Turntables. Youtube video. Published 24.11.2014. Referred on 11.03.2021. https://www.youtube.com/watch?v=9_F-b2hxP_o&t=2s

Danan. 2016. Why do We Need Topology in 3D Modeling. Blog. Read on 22. 03. 2021. http://thilakanathanstudios.com/2016/09/why-do-we-need-topology-in-3d-modeling/

Garnett, L. 2018. Photogrammetry in Mechanical Engineering. Blog. Read on 04.03.2021. https://www.engineeringclicks.com/photogrammetry-mechanical-engineering/

Geometry Best Practices for Artists. n.d. Blog. Read on 21.03.2021. https://developer.arm.com/solutions/graphics-and-gaming/developer-guides/game-artist-guides/geometry-best-practices/single-page

Gnomon. 2016. Creating Natural 3D Environments with Alex Alvarez. Published 14.12.2016. Referred on 23.03.2021. https://www.youtube.com/watch?v=T4uej9tppsU&list=PLzsm4rXzmwno17kKONMYd4r5WuTgIVzyq&index=3

Gray, E. 2020. Understanding Depth of Field: A Beginner's Guide. Blog. Read on 01.03.2021. https://photographylife.com/what-is-depth-of-field

Hamilton, A., Brown, K. 2016. Photogrammetry and Star Wars Battlefront. Screen shot of close-range photogrammetry with a hand-held DSLR. Watched on 02.03.2021. https://www.gdcvault.com/play/1023272/Photogrammetry-and-Star-Wars-Battlefront

I-Art 3D. n.d. What is "Depth Map"? Blog. Read on 25.03.2021. https://www.i-art3d.com/Eng/About_Depth.htm

Jirsa, Pye. 2016. Photography 101: Crop vs. Full Frame Cameras. Blog. Read on 20.03.2021. https://www.slrlounge.com/workshop/crop-vs-full-frame-cameras/

Kuzmin, V. 2018. Full Photogrammetry Guide for 3D artists. Blog. Read on 11.03.2021. https://80.lv/articles/full-photogrammetry-guide-for-3d-artists/

Lachambre, S., Lagarde, S., Jover, C. 2017. Photogrammetry Workflow. Manual. Read on 09.03.2021. https://unity3d.com/files/solutions/photogrammetry/Unity-Photogrammetry-Workflow_2017-07_v2.pdf

Lipman, A. 2018. 5 Industrial Photogrammetry Applications. Blog. Read on 18.03.2021. https://industrytoday.com/5-industrial-photogrammetry-applications/

Maio, A. 2020. What is Shutter Speed and its Role in the Exposure Triangle? Blog. Read on 13.03.2021. https://www.studiobinder.com/blog/what-is-shutter-speed/

Mansurov, N. 2020. Understanding Aperture in Photography. Blog. Read on 13.03.2021. https://photographylife.com/what-is-aperture-in-photography

Mansurov, N. 2019. What is Chromatic Aberration? Blog. Read on 14.03.2021. https://photographylife.com/what-is-chromatic-aberration#:~:text=Chromatic%20aberration%2C%20also%20known%20as,positions%20in%20the%20focal%20plane

Mansurov, N. 2019. What is ISO? The Complete Guide for Beginners.Blog. Read on 11.03.2021. https://photographylife.com/what-is-iso-in-photography#:~:text=In%20very%20basic%20terms%2C%20ISO,aperture%20and%20shutter%20speed%20settings

Marzolff, I., Ries, J., Aber, S., Aber J. 2019. Small-format photography: principles, techniques and geoscience applications .Amsterdam: Elsevier.

McDermott, W. 2018. The PBR Guide Part 2. Manual. Read on 21.03.2021. https://academy.substance3d.com/courses/the-pbr-guide-part-2

Normal Map Painting. n.d. Blog. Read on 22.03.2021. https://docs.substance3d.com/spdoc/normal-map-painting-109608964.html

Pendelton Design Management. 2016. Drone Mapping, Photogrammetry and Landscape Design. Read on 04.03.2021. http://www.pendletondm.com/blog/2016/10/11/drone-mapping-photogrammetry-and-landscape-design

Pluralsight. 2014. Understanding Ambient Occlusion. Blog. Read on 22.03.2021. https://www.pluralsight.com/blog/film-games/understanding-ambient-occlusion#:~:text=Ambient%20occlusion%20shading%20is%20actually,the%20area%20will%20stay%20brighter

Pluralsight. 2014. Understanding UVs: Love Them or Hate Them, They're Essential to Know. Blog. Read on 22.03.2021. https://www.pluralsight.com/blog/film-games/understanding-uvs-love-them-or-hate-them-theyre-essential-to-know

Rojas, J. 2015. Learn How to Use a Color Checker in Two Minutes. Blog. Read on 11.03.2021. https://fstoppers.com/originals/learn-how-use-color-checker-two-minutes-144866

Understanding Normal Maps. Blog. 2015. Read on 22.03.2021. https://blog.teamtreehouse.com/understanding-normal-maps

Vand de Beek, N. 2019. Photogrammetry. Blog. Read on 18.03.2021. https://nickyvandebeek.com/projects/photogrammetry/