3D Reconstruction of Reindeer Antlers Using a Low-Cost Optical Camera system and Gaussian Splatting

Julian Cramb [1], Derek Lichti [2], John Matyas [3], Shabnam Jabari [4]

[1] Dept. of Geomatics Engineering, University of Calgary – 2500 University Drive NW, Calgary, AB - julian.cramb@ucalgary.ca
[2] Dept. of Geomatics Engineering, University of Calgary – 2500 University Drive NW, Calgary, AB - ddlichti@ucalgary.ca
[3] Faculty of Veterinary Medicine, University of Calgary – 280 Hospital Dr NW, Calgary, AB - jmatyas@ucalgary.ca
[4] Dept. of Geodesy and Geomatics Engineering, University of New Brunswick– 3 Bailey Dr, Fredericton, NB - sh.jabari@unb.ca

Keywords: Antler measurement, Gaussian Splatting, Optical Imagery, Reconstruction, Low-Cost

Abstract

Caribou and reindeer antlers provide valuable insights for biodiversity, environmental monitoring, and regenerative medicine. However, traditional methods for measuring antlers rely on manual, time-intensive techniques prone to human error and ethical concerns due to physical animal interaction. This project aims to address these limitations by developing an automated, non-invasive 3D reconstruction pipeline utilizing optical imagery and Gaussian Splatting (GS). The system employs a multi-camera setup to capture high-resolution optical imagery, addressing challenges such as motion artifacts and inconsistent lighting. Prior approaches using time-of-flight cameras suffer from lower quality data, impact from subject movement and interference between cameras. GS-based methods using optical imagery are able to handle these conditions more effectively. Initial development and testing in a laboratory environment replicating the animal pen has shown strong results. With comparison to a reference laser scanner derived point cloud, a cloud-to-cloud mean distance of 3.4-3.7 mm was achieved. The system has been installed at the animal pen and one season of collection was performed; preliminary results showing good quantitative accuracies. The project's objectives include the production of high-quality 3D reconstructions of reindeer antlers, with the capability to handle subject movement. Expected outcomes include an automated 3D modeling pipeline, high-quality antler datasets spanning 2 growth cycles, providing valuable data for interdisciplinary research. These results will progress the understanding of antler growth dynamics and provide a scalable solution for future endeavors. By integrating optical imaging techniques with GS-processing, this research offers a novel approach to generating 3D models of antlers with wide-ranging applications.

1. Introduction

The antlers of Cervidae (deer, elk, etc.) are bony organs that are cast annually and regenerate at an explosive rate (Lincoln, 1992). The study of Cervidae antlers is a vast space with a long history spanning back 2000 years (Kawtikwar et al., 2010). Antlers are a subject of interest for the fields of biomedicine, biomechanics, biochemistry, ecology, and zoology. Indeed, antler regeneration provides valuable insights into the growth and mechanical properties of bony tissue (Currey, 2010), environmental monitoring (Kierdorf and Kierdorf, 2005), and regenerative medicine (C. Li, 2012). Specifically, the study of antlers has the potential to advance regenerative therapies for bone, cartilage, and wound healing in humans. Antlers also exhibit cancer resistance which may provide clues for gene-based treatment of human cancers (Feleke et al., 2021). Reindeer are unique in that both sexes produce antlers, along with calves growing them before puberty (Leader-Williams, 1988). This makes the regeneration of reindeer antlers of particular interest. The physical measurement of antlers is a fundamental aspect in the many fields of research, providing insight into growth characteristics. The most basic method for measuring antler size is with a ruler or tape measure. However, this requires proximity, which can traumatize the animal or cause an unsafe situation. These simple one-dimensional forms of measurement also lack the ability to effectively capture the complex shapes present in antlers, which allows only a rudimentary understanding of their growth properties

In the last decade, there have been advances in the 3D capture of antlers such as the use of optical imagery (Tsuboi et al., 2020) and handheld 3D scanners (Barba et al., 2013), yet both methods still suffer from motion artifacts unless the subject remains still for an extended period or is anesthetized or dead. To partly overcome these limitations, low-cost time-of-flight (ToF) cameras have been reported (Lichti et al., 2016). This work was then further expanded to automate the capture and segmentation of antler points (Cheng et al., 2022). Nevertheless, ToF cameras can also suffer from low data density, quality reductions from subject movement and interference between cameras (Cheng et al., 2022).

The aforementioned limitations of optical imagery, 3D scanners and low-cost ToF cameras have led to the need for a new solution for the measurement of live reindeer antlers. Advances in radiance field-based 3D reconstruction methods have expanded the options for the generation of 3D data. Pairing these methods with a low-cost optical multi-camera system will address the existing limitations. A well synchronized multi-camera system allows for instantaneous capture of the subject. There is no interference between optical cameras. The optical imagery can be streamed at a high rate without major data storage requirements. The point cloud density presented by 3D Gaussian splatting (3DGS) reconstructions is generally higher than the previous ToF camera applications, with the potential to densify these point clouds (Stuart et al., 2025). The goal of this research is to create and implement a low-cost multi-camera system for remote measurement of reindeer antlers via optical imagery, utilising 3DGS for 3D reconstruction. This paper reports on the following aspects, the initial design and implementation of the capture system, preliminary testing of the capture system in a laboratory setting along with the final installed location, and finally initial reconstruction results from live reindeer.

2. Related Work

2.1 Capture Methods

2.1.1 2D Passive Capture

2D imaging methods have been implemented for the classification and measurement of antlers along with estimation of 3D characteristics such as volume and size. Systems such as an aluminum frame which attached two parallel laser pointers and a camera allowed for the remote capture and estimation of ibex horn size (Bergeron, 2007). The lasers must be aligned in the image such that both appear on the subject being measured. As the lasers are separated by a known distance and are parallel, they serve to resolve scale in the image of the subject. The remote measurements were compared to tape measurements, and the error was from 3.9% to 5.2% which corresponds to approximately 2~3mm.

There have also been advances in the 3D estimation of antler size based on 2D imagery. One proposed system uses other anatomical features such as eye-to-eye, ear width, and eye size to scale 2D imagery (Demarais et al., 2013). From this, a predictive regression equation based on at least 100 taxidermy mounted specimens is used to estimate the 3D characteristics (Demarais et al., 2013). The Boone and Crocket score is a commonly used method for classifying antlers based on characteristics such as main beam length, tine length, palm length, point length and circumference. Figure 1 depicts the Boone and Crocket Club measurement diagram for a caribou/reindeer. Together these measurements combine to specify a specific score for that given individual. A thesis project was able to estimate this score from 2D imagery via the use of GIS based georeferencing to resolve scale in generation of the predictive equations (Flinn, 2010). They analyzed images of live deer, relying on ear width as the scaling factor, they were able to achieve a gross antler score with mean error of 4.3% (Flinn, 2010).

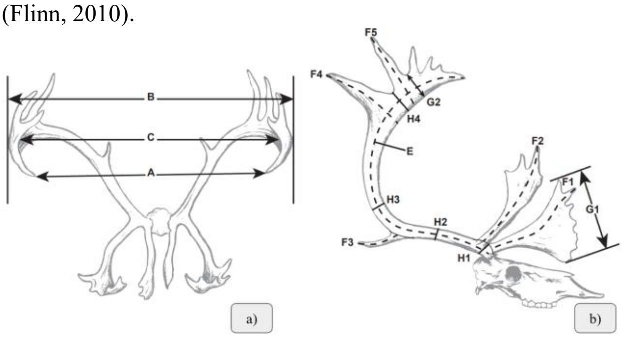

Figure 1. Boone and Crockett Club measurement diagram for a Caribou/Reindeer, (a) top down view (b) side on view (*Download B&C Score Chart PDFs | Boone and Crockett Club*, n.d.)

2D imaging is able to solve several of the limitations found in the traditional forms of measurement, such as the ability to capture the data remotely. This eliminates the ethical concerns of close proximity to live animals. The data volume is also increased significantly, as the user can capture a large number of images in a short period, broadening the scope of analysis. There are still limitations in these systems as they are only providing 3D estimates from single images: they are not directly computing any 3D characteristics. The estimates provided are also more rudimentary, such as gross score, inside spread, main beam length, total tine length, and circumference. These measurements provide general information about the antlers, but volumetric measurements along with any kind of textural analysis is difficult to perform. To utilize multiple stills of the same animal, captured from one camera, total stillness is required. Attempting to match the same point on the animal across frames will fail as in each frame that point will represent a different object space location. Factors including antler orientations, user error, image resolution, and species variation can contribute to errors in the 3D estimations.

2.1.2 3D Active Capture

3D scanners can be placed into one of three general categories, time of flight, phase shift and laser triangulation (Ebrahim, 2013). All three methods are not generally viable for live animal capture due to subject movement.

A study utilized two devices for the 3D capture of a subject, the ARTEC MH and the DAVID scanner. These devices are respectively a handheld flash scanner and a custom capture device consisting of optical imagery paired with a line laser. They used both devices to generate 3D models of a roe deer skull with its antlers intact. The manufacturers of both devices state 3D point accuracies in the sub millimeter range, 0.1 and 0.4mm for the ARTEC and DAVID devices respectively. This study found that both devices were effective for an application where cataloguing of the antlers was the sole purpose. The DAVID device lacked usability and was a more complicated procedure to capture the data. It did however capture textural details where the ARTEC was unable. The ARTECs higher resolution meant that it was able to better capture qualitative details such as the pearling (the small bumps on the antlers). Both devices, however, require absolute stillness of the captured subject, meaning live animal capture would likely require anesthetization or for the subject to be dead.

Further studies have proposed the use of time-of-flight cameras as a non-contact and tractable solution for the measurement of antler growth in live reindeer over time. This system required multiple captures of the same subject from different directions in order to achieve full coverage, the segmentation and registration of the antlers within the point clouds were performed manually. A process that the authors note would need to be automated to make this an enabling technology (Lichti et al., 2016). Another issue is that surface reflectivity affects the range measurement quality (Foix et al., 2011). The relationship between signal amplitude and shot noise (stochastic arrival process of photons) is inversely proportional for range measurements (Lange and Seitz, 2001). The dark felt of reindeer antlers results in low surface reflectivity and therefore higher shot noise which corresponds to an elevated measurement noise. This previous study was expanded by implementing a capture system comprising four low-cost time of flight cameras and one RGB camera along with a more automated pipeline for point cloud processing (Cheng et al., 2022). The system utilized the RGB camera for motion detection such that the animal subject is stationary during the capture of 3D data. The extraction of the antler point cloud is automated based on a fast k-d tree neighbor search, greatly speeding up processing. Multiple captures are required for adequate point cloud density, therefore there needs to be a two-step registration process. First the registration of the point clouds from all four cameras is performed using a target-based method. Then point clouds from different sampling periods are registered via the normal distribution transform to generate a densified cloud. The main limitation of this study is still the low reflectivity of reindeer antler felt causing noisy measurements in addition to the relatively low point density of each camera's capture (Cheng et al., 2022). There also exists methods for the generation of 3D reconstructions from 2D imagery. Thus, eliminating the need for direct 3D capture equipment to address the existing limitations of low point density and requirement for subject stillness. These methods are mainly split into the categories of multi-view-stereo (MVS) and radiance-field based techniques.

2.2 3D Reconstruction Methods

The complex structure of antlers has pushed researchers to the use of 3D imaging techniques as 2D methods are limited to more general antler characteristics (Demarais et al., 2013; Flinn, 2010). These methods generate data that is typically represented as either a point cloud or a mesh. 3D data provides a much more detailed representation compared to 1D or 2D. One the favored methods employed for 3D data generation of antlers is 3D scanning.

2.2.1 Multi-View Stereo

Dense point clouds and meshes can be generated post bundle adjustment via multi-view stereo (MVS) algorithms. The conventional algorithms used can be generally categorized into different scene representations of voxel-based, surface evolution, feature point growing, and depth map based (Stathopoulou and Remondino, 2023). This research has been expanded into methods using machine learning with methods falling into the categories of supervised MVS, unsupervised MVS, learning-aided methods and Neural representations. One study utilised photogrammetric bundle adjustment and MVS for the estimation of antler volume in several species (Tsuboi et al., 2020). This study employed Autodesk's ReCap Photo cloud processing service as the method for generating the 3D models of each specimen. They used repeated measurements and clay models of known volume as the reference for their accuracy assessment. They were able to achieve an average relative error of 8.5%. They found that a majority of the error (70%) originated from scaling error, followed by error in taking photos (19%) and finally rendering (9%). The number of images for each subject varied, requiring a minimum of 30 images with improvements to relative error improving with additional imagery. They estimate the accuracy could be improved to as low as 4.3% using replicated scaling procedures and renderings. The main limitations are the processing time and again, a requirement for stillness of the subject. The inability to effectively capture the number of required images on a live subject limits the use of this method for antler measurement (Tsuboi et al., 2020).

2.2.2 Radiance Field Techniques

A radiance field is a way of describing all of the light present in a scene at any given location and viewing direction. 3D models can be generated from radiance field representations in the form of point clouds or meshes. Radiance field representations of 3D scenes can generally be split into two categories, that being implicit/neural representations and explicit/primitive-based representations. The continuous representation allows for essentially unlimited point sampling when extracting a point cloud. The discrete representation is more limited in point density, but this can be partially overcome by densification algorithms. Both can have their reconstructions outputted as point clouds or meshes depending on the software used.
The first category represents a 3D scene as a continuous vector-valued function, with a 3D location of (x,y,z), a viewing direction of (θ, φ), whose outputs are an emitted colour (r, g, b) and volume density σ (Mildenhall et al., 2020). Neural Radiance Fields (NeRF) use this framework of continuous scene representation to train a model with imagery. The model samples the 5D coordinates along camera rays and feeds that into a multi-layer perceptron (MLP) network, this gives an output of colour and volume density along that ray. These values can then be synthesized into an image with the use of rendering. Because the rendering function is differentiable, the scene can be optimized via comparison of the rendered imagery to the ground truth imagery (Mildenhall et al., 2020). Once the model is fully trained, a novel view can be synthesized by inputting a 5D coordinate and an image is synthesized. Therefore, the NeRF reconstruction is not a literal reconstruction but rather a model that can generate novel views. There exists many iterations on the original methodology for improvements to processing speed, quality and training view requirements (Gao et al., 2025).

The second method represents the scene with discrete primitives in 3D space. The state-of-the-art in explicit radiance field representations is 3D Gaussian Splatting (3DGS). Initial Gaussians are made via structure from motion (SfM) to generate points. The quality of the reconstruction depends on the quality of the SfM point cloud (Kerbl et al., 2023). The Gaussians are 3D ellipsoids with a position, colour, covariance and transparency values. Each point from the sparse cloud is turned into a Gaussian. The training uses stochastic gradient descent. The model will view the Gaussians in a rasterized format from each image's perspective. The difference between the rasterized data and the imagery (ground truth) is used as the loss. Depending on what the loss value is the Gaussians will be modified to better fit the ground truth. It will either split, clone or prune a Gaussian if it is too small, too large, or too transparent. The process is iterated many thousands of times until either an iteration or loss threshold is reached (Kerbl et al., 2023). The scene can then be visualized in 3D space, the Gaussians are an explicit representation of the reconstructed scene. As opposed to the NeRF model where the scene is generated via the trained neural network at each novel view. As previously mentioned, 3DGS can suffer from low point density when outputting point clouds of the reconstructed scenes. Densification algorithms can aid this limitation, this can cause some aberrations but generally improves the textural features of the point cloud. Additionally, there is software that generates a mesh representation of the radiance field. Software such as SuGaR and MILo perform an optimization of the 3DGS algorithm with a secondary mesh estimation (Guédon et al., 2025; Guedon and Lepetit, 2024). SuGaR has modified the algorithm with a regularization term to optimize for flatter Gaussians that are better aligned with the surface. This is achieved by minimizing the smallest axis in the Gaussians and removing the Gaussians in areas of low density. MILo generates a mesh at every iteration of training, using both the imagery and mesh for the loss function. While other software like MeshSplatting use connected triangles instead of points to represent the subject, while optimizing the algorithm to make the triangles opaque at the end of training, directly creating a mesh from the model training (Held et al., 2025).

3. Methodology

3.1 System Design

3.1.1 Measurement Environments

The ultimate measurement environment is a specialized animal holding pen located at the University of Calgary's Wildlife

Research Station (WRS). This pen, pictured in Figure 2 (a), measures 2.5 m x 3.5 m x 2.5 m (L x W x H). To simulate this capture environment, a recreation was made in the test laboratory located at the University of Calgary’s main campus. This recreation, pictured in Figure 2 (b), was a wooden frame of the same dimensions as the holding pen, used for initial testing and design purposes.

Figure 2: Capture areas, (a) holding pen 2.5 m x 3.5 m x 2.5 m (L x W x H) (b) Recreation of the holding pen located in the test laboratory

3.1.2 Capture Subjects

Two separate capture subjects are used in the design and testing of the system. Subject 1 is an amputated, dried, velvet-covered reindeer antler consisting of one main beam with several tines (Figure 3 (a)). It is approximately 0.55 m x 0.10 m x 0.50m (L x W x H) and is mounted on a tripod at a height of approximately 1.5 m. This antler approximates what would be captured in the live scenario, although the test antler is half of a full set.

Figure 3: Test capture subjects, (a) subject 1, single reindeer antler (b) subject 2, 3D printed reindeer mount

Subject 2 used for design and testing was a full-scale 3D printed reindeer mount as seen in Figure 3 (b). The 3D print file was sourced from Thingiverse, the file used was a remix of a full body 3D model from Free3D called “reindeer v3”. This 3D printed subject closely approximates the real capture subjects as it contains both a full set of antlers along with a portion of the upper body and head. The size of this print is approximately 0.85 m x 0.85 m x 0.6 m (L x W x H). For use in testing this subject was held by an individual at a height of approximately 1.5 m. The colour of the print is not of the same colour as the real animals and the fine hair texture of antler felt is not captured. As the print was made of 8 different pieces that needed to be glued together, there are some geometric differences when compared to the triangulation model.

3.1.3 Camera System

To obtain the complete geometric extent of the antlers, full-scene coverage and instantaneous data capture are required. Therefore, a specialized optical multi-camera capture system is proposed for this application. Another requirement of this capture system was the need for it to be low-cost, optical imagery favours this prerequisite as there are a wide range of imaging products available for purchase. Devices used for laser scanning or flash LiDAR are sequential capture devices which suffer from motion artifacts, they are generally more costly and the lower-cost examples of these devices may have higher noise and not provide adequate density (Cheng et al., 2022).

The development of this system was performed in the test laboratory pictured in Figure 2 (b), using capture subject 1. The number of cameras used in the test lab was initially set at 20 total with 5 cameras per side. This imagery network was then manipulated by removing images until a failure presented. These failures could be in the form of failed network orientation in the bundle adjustment step, or in the reconstruction step. Failure in the reconstruction step will appear as either a full failure to render the scene or causing missing segments from the subject. The requirements for the cameras were as follows, independently powered, wireless control, minimization of cost, synchronization capabilities within the range of 50-100ms, and a spatial resolution of approximately 1mm. The high level of synchronization is required for the minimization of movement related error. At 100 ms of delay, at an assumed walking speed of 0.5 m/s would cause a 50 mm error. This highlights the importance of synchronization, as even a small delay will result in a large error. As for spatial resolution, projecting a pixel size of 1.4 µm at a focal length of 4.74 mm (chosen camera specs) into the object space results in a pixel size of 0.15 mm to 1.27 mm for the min and max range respectively. These specifications are from the camera that was ultimately chosen for use. Assuming an antler diameter range of 1-5cm (Reimers, Eigil and Høymork, Amund, 2002), using the same image configuration. The number of samples across the diameter of the antler at the max range would be in the range of 8-39. At the minimum range it would be 66-333 samples. This means that even at the maximum distance, the antler can be resolved at a coarse interval from the imagery. Dropping resolution below this common size may cause a loss of detail as the maximum object space pixel size approaches 2mm. Several cameras will be analyzed to select the one which meets the previously mentioned criteria most effectively.

Ideal capture conditions are defined as follows, the lighting must be sufficiently bright such that the subject is well contrasted from the background, in the case of this experiment that meant all the overhead fluorescent lights were on. This ensures that the reconstruction software has the best chance of accurately representing the edges of the subject. The subject must also be still for approximately 3-6 frames as the average delay between cameras is around 100ms, at the max frame rate of 56.03 fps that is just under 6 frames. This minimizes the potential effects of motion blur in the imagery and discontinuity (temporal differences) between the cameras. The position of the subject is fully centered in the capture area such that each camera has a full view with no occlusions or portions of the subject outside the FOV. This allows for the maximum amount of subject information to be gathered by each camera. Together, these conditions minimize the potential error sources relating to data capture.

3.1.4 Software

Many forms of 3DGS require that the imagery being used has well estimated camera interior orientation parameters (IOPs) and exterior orientation parameters (EOPs) along with an initial photogrammetric point cloud. Therefore, the cameras must be calibrated, and the imagery network must be oriented. The determination of IOPs is commonly done via a self-calibration bundle adjustment. The network orientation or EOPs are determined through a standard bundle adjustment once the IOPs have been estimated. There are a multitude of options for processing the bundle adjustment from purpose-built software to pre-existing packages such as COLMAP (Schoenebeck et al., 2013; Schönberger and Frahm, 2016). Similarly, there are several renditions of the 3DGS model from different developers including but not limited to NeRFStudio, Kiri Engine, Polycam, Pix4D. Fundamentally, they all use similar methodologies for generating the radiance field rendering. However, the selection of which software to use will depend on ease of use for this specific application. The further processing and analysis of the generated point clouds will be performed in the software CloudCompare.

3.2 Experimental Design

3.2.1 Initial Laboratory Testing

The initial testing was conducted in the laboratory using the holding pen recreation. Subject 1 was used for this stage of development and testing. Using this setup testing was undertaken on the lighting conditions of the room. The lighting was a major concern for the system. At the WRS, there are several sets of overhead fluorescent lights and during anaesthesia events all but one row of lights on the far end of the room are turned off. This one set of lights provides some illumination although it is not direct, therefore the subject is not well contrasted from the background of the pen. As there are limited opportunities to capture the animals, there is a need to test capture data under all potential conditions. There are several potential issues resulting from poor lighting conditions. That being the potential for the bundle adjustment failing to orient the imagery, the decrease in point density and potential for higher error in the reconstruction.

Bundle adjustment requires the extraction and matching of tie points between frames. Automated software often relies on well defined locations with sharp edges or known shapes. The dark scene makes it more difficult for the software to detect and match tie points on the subject (Liu et al., 2025). This can result in the adjustment failing altogether if enough points cannot be identified or matched. As the 3DGS solution also relies on the sparse point network made during the bundle adjustment for initialization of the Gaussians. The reduced number of SfM points generated will result in less initial Gaussian primitives being formed (Zhou et al., 2025). Over-exposed and in this case under-exposed images contain less information, heightened noise, and colour distortion, this poses a challenge for 3DGS to produce a high-quality reconstruction (Zhou et al., 2025). In addition to this the lower signal-to-noise ratio of dark images can result in numerous thin and elongated Gaussians to represent the noise (Z. Li et al., 2024). This results in more noise points being exported into the point cloud.

The laboratory room contains several sets of overhead fluorescent light along with 4 large windows. The testing of the lighting conditions consisted of capturing data under two different states. The first was the bright state where all overhead lights remained on, the second was the dark state, where all lights were turned off and the 4 windows are covered. For both states the capture subject was centred in the capture area. The goal of this test is to best replicate the conditions present at the holding pen where live animal subject will be captured. The initial purpose of this test was to ensure that in the dark state, the bundle adjustment software would not fail due to limited tie points and that the 3DGS reconstruction would still be able to fully reconstruct the subject. The lighting testing was also assessed on the impact to the accuracy of the generated point cloud. To verify the quality of the 3D renderings achieved, a reference point cloud from a FARO TLS will be used as the ground truth. The comparison will be made on the basis of cloud-to-cloud (CTC) distance which can be computed following cloud registration. Registration and accuracy measurement was accomplished using Cloudcompare, the first step is an initial coarse alignment. This involves manual adjustments to the orientation, translation and scale of the generated point cloud, until it approximated the ground truth cloud. Cloudcompares iterative closest point algorithm is then used to refine and finalize the registration of the cloud and the true point cloud. Both the mean CTC distances and the CTC standard deviations will be reported. Heat maps of the CTC distance were generated for visual analysis of reconstruction quality. The primary method for 3D reconstruction of 3DGS was compared to the other leading radiance field reconstruction method of NeRF via Nerfstudio. The same calibrated imagery network of the 3D printed subject will be utilized for each case, each method will be assessed on the metrics of speed, ease of use and final reconstruction accuracy. The metrics used for accuracy assessment will be the same as previously mentioned, that being the CTC distance and heat maps for visual analysis.

3.2.2 Installed System Testing

The camera system was installed at the WRS after initial testing. Wooden boards which were fastened directly to the steel frame of the holding pen, to those boards, the cameras were affixed. To perform the testing of the final installed system, subject 2 was utilized. As previously mentioned, the dark lighting present during the anaesthesia events needed to be tested to ensure no major failures occurred. Using subject 2, the two lighting states were established. The bright state where all the overhead lights remain on, this is the condition for all non-anaesthesia events and would comprise a portion of the live animal captures. The dark state is when all but the far row of overhead lights are turned off, providing minimal lighting. Data of subject 2 centered in the capture area was captured under the bright and dark state. This data was processed using the same pipeline as in the initial laboratory testing. Like the laboratory testing, this experiment will first ensure that no major reconstruction failures occur. Either the bundle adjustment failing or the subject is not fully captured due to a reconstruction failure. The accuracy metrics of CTC distance was computed to assess the impact of the lighting conditions. In this case the triangulation model was used as the ground truth, the Cloudcompare sample points tool was used to generate a surface point cloud of 1 million points from the mesh model. As subject 2 was made of 8 separate pieces glued together there will be some error when compared to the model due to assembly misalignment. The model also did not have a well determined scale and therefore the accuracy assessment was not absolute. For this reason, the CTC metrics will be used in a relative manner to assess the differences caused by the lighting conditions. An example of the lighting conditions in the test laboratory and at the WRS can be seen in Figure 4.

Figure 4: Frames from lighting conditions testing, (a) lab bright, (b) lab dark, (c) WRS bright, (d) WRS dark

3.2.3 Densification Testing

Both during the initial laboratory and installed system testing, a densification test was performed to determine its effects on accuracy and qualitative attributes. The CTC metrics are compared before and after densification. The subject 1 clouds can be analysed on true accuracy whereas the subject 2 clouds are just analysed on a relative basis. The clouds are also checked for additional noise and any artifacts caused by the densification, this is accomplished manually by visual inspection.

3.3 Final System Implementation

3.3.1 Camera System

The number of capture points was chosen based on the criteria of minimizing the total number of cameras needed while maintaining the ability to fully generate 3D reconstructions of the subject. From initial lab testing with a smartphone camera in a reconstruction of the WRS pen, it was decided that 16 cameras with 4 per side would be used. The cameras used for this system were custom made using Raspberry Pi components, consisting of the camera module 3, a Pi board 4 with 1 GB of RAM, along with a 32GB SD card. Each camera is powered individually, for this power banks were used, each with a 20,000-milliamp hour capacity. 3D printing was used to create the case with the designs taken from Thingiverse. Each camera case consisted of 8 separate components fixed together with machine screws and nuts. The final implemented design can be seen in Figure 5. This figure shows four of the cameras mounted to the laboratory test frame. The boards communicate through secure shell protocol (SSH), with the ability to control any aspect of each individual camera via the master. The board clocks were synchronized via Chrony (Lichvar and others, 2026).

The raw captured data consists of 1920 x 1080 pixel optical imagery acquired at 30-56.03 fps depending on the experiment. As described in the previous section, the videos are stored individually on each camera as .MP4 files. Through SSH protocol, the videos from each camera are downloaded onto the master Pi board. They are then removed from the master board to a portable drive and transported to a different machine for further processing. Image convergence was set to have each camera oriented towards the center of the capture area at a height of approximately 1-1.5 m. Being pointed centrally ensures that there is sufficient overlap between images, treating the center of the pen as the optimal capture location. The height was chosen based on the observed height of the reindeer present at the WRS. With the goal of capturing from their upper neck to the top of the antlers.

Figure 5: Camera modules installed on wooden test frame

3.3.2 Software Pipeline

As the Pi boards are synchronized and start from a predetermined time based on their internal clocks, selecting the same frame number from each video file results in 16 images of the same moment in time. There is of course some synchronization error present in this method which means the user should select a moment where the subject is briefly still for approximately 3-6 frames depending on the frame rate. These 16 images are then processed through a photogrammetric bundle adjustment using COLMAP. This provides estimates for the EOPs and IOPs of the capture network along with a sparse point cloud used for Gaussian initialization. This bundle adjustment must be performed for each unique image set as the 3DGS performance heavily depends on the SfM points. Due to the use of the Nerfstudios built-in version of COLMAP, the precision metrics are not provided in the outputs. Importantly, the SfM point cloud lacks true scale due to the lack of predetermined EOP estimates and is given an arbitrary scale. This leads to the 3D reconstructions also lacking true scale. This will need to be addressed to make measurements of the live subjects. This will be addressed in future by performing a full system calibration at the WRS.

To generate the 3D rendering with the 3DGS algorithm, Nerfstudios Splatfacto model was used in conjunction with the adjusted network of imagery (Kerbl et al., 2023). Nerfstudio was selected as the reconstruction software as it provided many different reconstruction models including 3DGS and NeRF, it also contains a built-in viewer and point cloud exporting feature. The only way to export the 3DGS scene to a point cloud from Nerfstudio is by having each Gaussians origin represented by a point.

The densification is accomplished with the software package called 3DGS-to-PC (Stuart et al., 2025). This software takes the splat output PLY file along with the transforms JSON file, it will generate a cloud with the specified number of points. The authors do not report on an optimal density increase to avoid under or over sampling of the scene. The densification is accomplished by the following. For every pixel, the software calculates the contribution that each Gaussian has towards calculating the pixel colour. The pixel to which a Gaussian contributes most to is assigned the colour of that pixel. Only the zeroth order spherical harmonic for colour is used in calculations to reduce rendering time. The Gaussians that contribute zero to the scene are then removed to eliminate redundant Gaussians. Finally, they sample each Gaussian as a multivariate standard distribution, while removing erroneously generated points until the user selected number of points is reached (Stuart et al., 2025).

The point clouds in this paper will be referred to as either raw or densified, with raw clouds being those outputted directly from

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1011-2026 | © Author(s) 2026. CC BY 4.0 License.

the Nerfstudio software, and densified referring to those outputted from the 3DGS-to-PC software. Both have been cleaned for outlier and noise points, any deviation from this will be noted. The raw and densified cloud require some initial cropping to extract the antler points. This is done manually via the cross-section tool. The densified clouds often suffer from some additional noise points along the surface of the antler. These noise points are reduced using a tool which computes the average Euclidian distance of each point to its k-nearest neighbours. Points that exceed the global threshold, defined by the mean distance plus a user specified multiple of the standard deviation, are pruned. The clouds directly outputted from Nerfstudio do not suffer from the same level of noise points, often only containing a small number of them. If any outliers need removal, the segment tool is used.

4. Results

4.1 Design And Testing

4.1.1 Radiance Field Methodology Selection

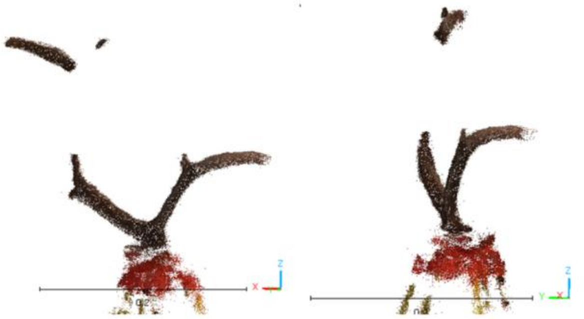

Figure 6: Example of failed reconstruction of the test subject, made during initial testing of the system in the test lab

The Nerfstudios Nerfacto model was used initially for testing as the reconstruction method as it could provide a higher density point cloud output. It also performed similarly to the Splatfacto model in terms of time required to train a model. With both methods taking around 20 minutes on a machine with an 11th gen Intel Core i7 and an NVIDIA GeForce RTX 3060. However, problems were encountered when the number of images was reduced. Below 16 images, the NeRF model tended to fail either at reconstructing the scene altogether or containing major gaps in the subject. An example of this failure can be seen in Figure 6. This failure pushed us to use 3DGS as a method as it tended to be much more reliable with a sparser image set. Especially in the case of dark lighting, the 3DGS proved to be much more effective at reconstructing the subject than NeRF was. As one of the main goals of this project was a low-cost system, reducing the number of cameras needed was an effective way to accomplish that goal.

4.1.2 Lighting Conditions

The lighting conditions test previously described was performed twice with the finalized system design during the initial laboratory testing. This resulted in two point clouds from the bright state and two from the dark state. The CTC accuracy metrics were averaged across both tests. The accuracy statistics of this testing can be found in Table 1.

It is noted that the accuracy of the cloud derived from dark imagery was higher than that of the bright state. Reflected in a lower mean distance and standard deviation for the dark cloud. Although the sample size is small for this test, all the results for mean distance were within 0.5mm for the four clouds analyzed. Similarly, the standard deviations were within 0.4mm. This indicates that for the test laboratory setup, there is not a significant impact from the reduced lighting. The lighting test for subject 2 was performed 2 times with the installed system. Again, resulting in two point clouds from the bright state and two from the dark state. The average relative CTC statistics can be found in Table 2. The results from the WRS installed system testing indicate that the lighting may have a larger impact on accuracy than in the laboratory. The mean distance has a similar change although in this case the dark state provided a lower accuracy. Additionally, the change in standard deviation between the two states is much larger than in the laboratory test. As can be seen in Figure 4 (d), the dark state of the WRS appears to have less lighting than that of the laboratory in Figure 4 (c). This may have contributed to the enhanced error found in the dark state. There may also be some contributions due to the subject and therefore the point cloud being larger for subject 2. With it being a more complex object the 3DGS algorithm may struggle more. If converted to mm based on the subject 1 statistics, 26.64% would result in a increase to standard deviation of 0.8 mm. This is still relatively low in terms of scale.

Lighting State	CTC Mean Distance (mm)	CTC Standard Deviation (mm)
Bright	3.7	3.0
Dark	3.4	2.9
Bright to Dark	-8.9%	-2.85%

Table 1: Lighting conditions test statistics for subject 1 raw cloud generated during initial laboratory testing of the final system in the test lab

Of note is the fact that the raw point clouds generated from the dark state imagery are much less dense than the bright state derived ones. The bright derived clouds varied from approximately 4000 to 6000 points, whereas the dark derived ones were from 1000 to 1500 points. This is potentially due to the previously mentioned limitations of 3DGS at its reliance on good initial SfM points and imagery with good a signal to noise ratio. Overall, the lighting tests do not show any significant impact to accuracy from the dark state at the laboratory or the WRS. The dark state at the WRS was more impactful to the standard deviation than that of the laboratory, potentially pointing to more noise points present in that cloud. The decrease to point density in dark lighting may be overcome through densification of the point clouds, this will be analysed in the next section for impact to accuracy.

Lighting State	CTC Mean Distance (% change from bright)	CTC Standard Deviation (% change from bright)
Bright	0.00	0.00
Dark	8.04	26.64

Table 2: Lighting conditions test statistics for subject 1 raw cloud generated during testing of the final system installed at the WRS

4.1.3 Densification

As previously mentioned, the lower point density problem can be aided with additional processing. An example of point clouds made during initial testing of subject 1 using the final system can be seen in Figure 7. The red cloud was generated directly from Nerfstudios Splatfacto model and the points represent the center of each Gaussian (Tancik et al., 2023). The green point cloud was generated using 3DGS-to-PC to densify the raw cloud (Stuart et al., 2025). The Blue cloud is the TLS scan of subject 1. Both the raw and densified point clouds were processed in Cloudcompare to remove noise and outlier points.

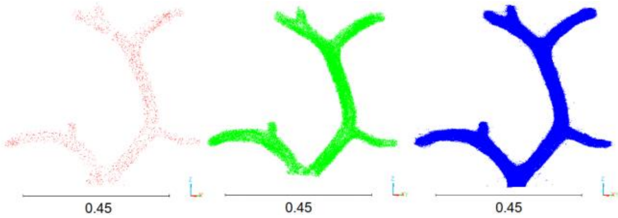

Figure 7: Raw (red), densified (green), and reference TLS point cloud (blue) of the test subject 1 made during initial testing of the final system in the test lab

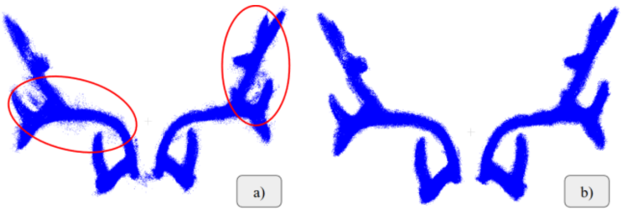

Figure 8: Densified point cloud of subject 2 antlers (a) wo/noise removal, (b) w/noise removal, made during testing of the final system installed at the WRS, arbitrary scale

From a qualitative perspective, it can be seen that the geometry and edges of the antler are well preserved in the densified point cloud. Although the density of the directly exported cloud is still high enough to visually detect all the main features of the antler. There are no obvious major artifacts created from densification of the raw point cloud. The volume density of each point cloud was computed using a spherical neighbourhood search of radius 3cm. An average density of 56, 2459 and 5403 points per sphere was computed for the raw, densified and TLS point clouds seen in Figure 7. Demonstrating a significant increase to the point density. As mentioned, the densified clouds suffer from additional noise points surrounding them. An example this noise can be seen in Figure 8. This figure shows a point cloud of subject 2 from the WRS installed system. First the densified point cloud which has only been cropped is shown, below, the same cloud after being processed to remove noise and outlier points is seen, examples of noise are denoted by the red circles.

Cloud	CTC Mean Distance (mm)	CTC Standard Deviation (mm)	Number of points
Raw	3.60	2.95	1,470
Densified	3.65	2.89	71,000
Change	1.39%	-2.16	N/A

Table 3: Subject 1 densification testing accuracy statistics, distances are computed from the reference TLS cloud, change is computed from the raw to densified clouds, made during initial testing of the final system in the test lab

The effects of densification were tested 3 times during initial testing of the final system. Bright state imagery was used for all 3 cases. The average accuracy statistics of the subject 1 tests can be found in Table 3, along with average number of points. Both the raw and densified point clouds were processed to remove noise and outlier points.

The densification test of subject 1 indicates that there is not a significant impact to the accuracy. The change is smaller than the impact due to the lighting conditions, which is already relatively small. The densification test was performed using one point cloud of subject 2 at two different levels of densification. Again, bright state imagery was used to generate the point cloud. The relative CTC statistics can be seen in Table 4. Again, the subject 2 densification testing indicates that there is not a significant change from the additional points generated. There is still the problem of a smaller sample size for the tests undertaken. However, as previously stated the variance in accuracy amongst all the point clouds was very low. To get a better understanding of where the error tends or originate in the point clouds the heat maps are analysed. A visualisation of the CTC mean distance can be seen in Figure 9. This figure represents the CTC distance error from a densified, 3DGS derived point cloud to the TLS scan ground truth point cloud.

Cloud	CTC Mean Distance (% change from raw)	CTC Standard Deviation (% change from raw)	Number of points (#)
Raw	0.00	0.00	5,164
Densified 1	-1.35	-0.66	46,539
Densified 2	-0.97	1.95	307,592

Table 4: Subject 2 densification testing accuracy statistics, are distances computed from the reference TLS cloud, made during testing of the final system installed at the WRS

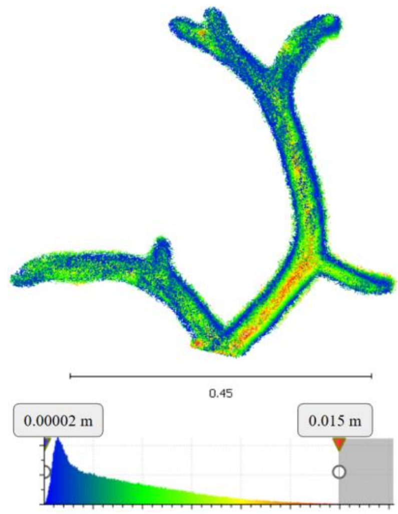

Figure 9: Subject 1 densified point cloud with error shown in meters, made during initial testing of the final system in the test lab

There are some areas along the edges of the point cloud where the points extend past the surface. However, from visual inspection it was noted that a larger portion of the points with higher error were those within the surface of the antler. Worth noting is the fundamental issue that the point clouds derived via the method of using the centre of the Gaussian results in points that appear inside the subject. This can be seen in Figure 9 as the points that are orange to red in the lower section of the main beam. 3D scanning methods, and other methods such as MVS generate surface points or meshes. This means that the CTC or cloud-to-mesh distances to surface models are impacted as all points that lie inside of the subject surface will add error. This is one of the reasons that generating meshes from 3DGS is not a straightforward objective, requiring algorithm modifications or additional steps (Guédon et al., 2025; Guedon and Lepetit, 2024; Held et al., 2025). These mesh models would be able to resolve the issue of points lying below the surface, allowing for a more realistic measure of accuracy for these point clouds. This also indicates that the CTC statistics computed in the lighting and densification tests may have inflated error values due to these internal points.

4.2 Live Animal Capture

The system can also generate qualitatively accurate reconstructions of live reindeer antlers under various lighting conditions. An example of one 3DGS rendering is seen in Figure 10 (a) with the corresponding non-densified point cloud seen in Figure 10 (b). These figures show reindeer 1F from the herd, captured after a non anaesthetic event in the bright state.

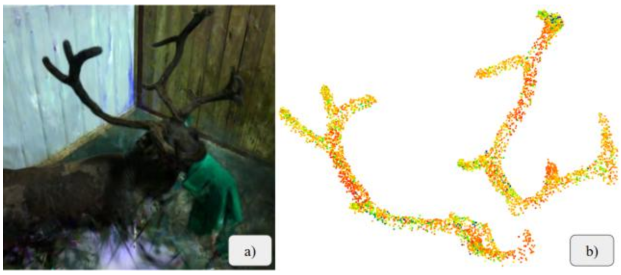

Figure 10: Reindeer 1F, (a) view of 3DGS reconstruction (b) raw antler point cloud, captured using final system installed at the WRS

5. Conclusion

This study investigated the use of a low-cost optical multi-camera system and Gaussian splatting for automated, non-invasive 3D reconstructions of reindeer antlers. The proposed pipeline addresses key limitations of previous methods, such as the need for prolonged subject stillness, ethical concerns of close contact, and low point density. This is accomplished by enabling instantaneous capture from 16 synchronized cameras and high-fidelity reconstruction via 3D Gaussian splatting.

Testing was performed on the lighting conditions of the lab and installed capture, both bright and dark states were analysed. The test demonstrated that the various lighting conditions present in the capture areas do not have a large impact to accuracy for this system. However, point density does suffer in the darker conditions of the installed system. This was overcome through the use of densification algorithms which were also found to not impact the accuracy in a meaningful way. Accuracy measures of CTC mean distance varied from 3.4 to 3.7 mm across tests, indicating a good level of precision and repeatability. In real-world deployment at the Wildlife Research Station, the system produced qualitatively accurate reconstructions of live reindeer antlers. While solving most limitations from previous studies, the methods used in this paper present new challenges. The stillness of the subject is still required for an optimal capture, although only brief stillness is required in this case. The point cloud contains points that lie within the surface of the subject, causing issues for surface modelling and accuracy computations. The system suffers in low light conditions, decreasing point density and increasing noise. There also still exists the issue no true scale from the SFM point cloud, resulting in an arbitrary scaling in the 3D reconstruction.

Future work for this project includes a rigorous full system calibration of EOPs and IOPs, EOPs will resolve the arbitrary scaling issue and IOPs will improve the final accuracy. A more rigorous testing of the camera system, testing of variance in capture rate, subject positioning, and subject motion. A comparison of other reconstruction methods such as 3DGS mesh methods and MVS. Collecting another growth period of data and performing a year-to-year change analysis. By providing a non-invasive, accurate, and scalable tool for antler measurement, this research contributes to advancing wildlife monitoring, regenerative medicine studies, and ethical biometric data collection. The system’s modular design and open-source components offer an adaptable framework for similar applications in ecology, veterinary science, and beyond.

Acknowledgments

The authors are indebted to Drs. Rob McCorkell, Greg Muench, Jeff Biernaskie whose research made animals available for study, and to the staff at the Wildlife Facility for their care of the animals. These studies have been approved by the institutional animal care committee Protocol #aAC23-0011. Funding for this work was provided by the Natural Sciences and Engineering Research Council (NSERC) of Canada (RGPIN-2024-03793) and the UCalgary Research Excellence Chairs program.

References

A G Stuart, L., Morton, A., Stavness, I., and Pound, M. P., 2025. 3DGS-to-PC: 3D Gaussian Splatting to Dense Point Clouds. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, 3730–3739.

Barba, S., Fiorillo, F., and Feo, E., 2013. 3D-ANTLERS: Virtual reconstruction and three-dimensional measurement. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 15–20. https://doi.org/10.5194/ISPRSARCHIVES-XL-5-W1-15-2013

Bergeron, P., 2007. Parallel Lasers for Remote Measurements of Morphological Traits. The Journal of Wildlife Management, 71(1), 289–292. https://doi.org/10.2193/2006-290

Boone and Crockett Club, n.d. Download B&C Score Chart PDFs. https://www.boone-crockett.org/download-bc-score-charts (accessed 4 February 2026).

Cheng, S., Lichti, D. D., and Matyas, J., 2022. Automated generation of high-quality 3D point clouds of antlers using low-cost range cameras. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B2-2022, 531–538. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-531-2022

Currey, J. D., 2010. Mechanical properties and adaptations of some less familiar bony tissues. Journal of the Mechanical Behavior of Biomedical Materials, 3(5), 357–372. https://doi.org/10.1016/j.jmbbm.2010.03.002

Demarais, S., Strickland, B. K., Flinn, J., and Webb, S. L., 2013. Method and system for estimating antler, horn, and pronghorn size of an animal. United States Patent No. US8483446B2.
Ebrahim, M. A.-B., 2013. 3D Laser Scanners’ Techniques Overview. International Journal of Science and Research (IJSR), 4(10).

Feleke, M., Bennett, S., Chen, J., Hu, X., Williams, D., and Xu, J., 2021. New physiological insights into the phenomena of deer antler: A unique model for skeletal tissue regeneration. Journal of Orthopaedic Translation, 27, 57–66. https://doi.org/10.1016/j.jot.2020.10.012

Flinn, J. J., 2010. Accuracy of estimating age and antler size of photographed deer. M.S. thesis, Mississippi State University.
Foix, S., Alenya, G., and Torras, C., 2011. Lock-in Time-of-Flight (ToF) Cameras: A Survey. IEEE Sensors Journal, 11(9), 1917–1926. https://doi.org/10.1109/JSEN.2010.2101060

Gao, K., Gao, Y., He, H., Lu, D., Xu, L., and Li, J., 2025. NeRF: Neural Radiance Field in 3D Vision: A Comprehensive Review (Updated Post-Gaussian Splatting). https://doi.org/10.48550/arxiv.2210.00379

Guédon, A., Gomez, D., Maruani, N., Gong, B., Drettakis, G., and Ovsjanikov, M., 2025. MILo: Mesh-In-the-Loop Gaussian Splatting for Detailed and Efficient Surface Reconstruction. ACM Transactions on Graphics. https://doi.org/10.1145/3763339

Guedon, A. and Lepetit, V., 2024. SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 5354–5363. https://doi.org/10.1109/CVPR52734.2024.00512

Held, J., Son, S., Vandeghen, R., Rebain, D., Gadelha, M., Zhou, Y., Cioppa, A., Lin, M. C., Van Droogenbroeck, M., and Tagliasacchi, A., 2025. MeshSplatting: Differentiable Rendering with Opaque Meshes. https://doi.org/10.48550/arxiv.2512.06818

Kawtikwar, P. S., Bhagwat, D. A., and Sakarkar, D. M., 2010. Deer antlers- Traditional use and future perspectives. Journal of Herbal Medicine and Toxicology, 9(2).

Kerbl, B., Kopanas, G., Leimkuehler, T., and Drettakis, G., 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics, 42(4), 1–14. https://doi.org/10.1145/3592433

Kierdorf, U. and Kierdorf, H., 2005. Antlers as biomonitors of environmental pollution by lead and fluoride: A review. European Journal of Wildlife Research, 51(3), 137–150. https://doi.org/10.1007/s10344-005-0093-0

Lange, R. and Seitz, P., 2001. Solid-state time-of-flight range camera. IEEE Journal of Quantum Electronics, 37(3), 390–397. https://doi.org/10.1109/3.910448

Leader-Williams, N., 1988. Reindeer on South Georgia: The ecology of an introduced population. Cambridge University Press.
Li, C., 2012. Deer antler regeneration: A stem cell-based epimorphic process. Birth Defects Research. Part C. Embryo Today, 96(1), 51–62. https://doi.org/10.1002/bdrc.21000

Li, Z., Wang, Y., Kot, A., and Wen, B., 2024. From Chaos to Clarity: 3DGS in the Dark. https://doi.org/10.48550/arxiv.2406.08300

Lichti, D. D., Steward, J., Chow, J. C. K., and Matyas, J., 2016. The Practical Application Of 3D Vision in the Field: Measuring Reindeer (Rangifer Tarandus) Antler Growth Velocities. The Photogrammetric Record, 31(156), 394–406. https://doi.org/10.1111/phor.12167

Lichvar, M. et al., 2026. chrony: A versatile implementation of the Network Time Protocol. https://chrony-project.org/
Lincoln, G. A., 1992. Biology of antlers. Journal of Zoology, 226(3), 517–528. https://doi.org/10.1111/j.1469-7998.1992.tb07495.x

Liu, C., Wang, Y., Luo, W., and Peng, Y., 2025. LDFE-SLAM: Light-Aware Deep Front-End for Robust Visual SLAM Under Challenging Illumination. Machines, 14(1). https://doi.org/10.3390/machines14010044

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R., 2020. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. https://doi.org/10.48550/arXiv.2003.08934
Reimers, E. and Høymork, A., 2002. Antler development in reindeer in relation to age and sex. Rangifer, 22(1), 75-. https://doi.org/10.7557/2.22.1.692

Schoenebeck, C. W., Peterson, B. C., and Obermiller, J. A., 2013. Accuracy of antler metrics in predicting age of white-tailed deer and mule deer. Proceedings of the Nebraska Academy of Sciences, 23.

Schönberger, J. L. and Frahm, J.-M., 2016. Structure-from-Motion Revisited. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Stathopoulou, E. K. and Remondino, F., 2023. A survey on conventional and learning-based methods for multi-view stereo. The Photogrammetric Record, 38(183), 374–407. https://doi.org/10.1111/phor.12456

Tancik, M. et al., 2023. Nerfstudio: A Modular Framework for Neural Radiance Field Development. ACM SIGGRAPH 2023 Conference Proceedings.

Tsuboi, M., Kopperud, B. T., Syrowatka, C., Grabowski, M., Voje, K. L., Pélabon, C., and Hansen, T. F., 2020. Measuring Complex Morphological Traits with 3D Photogrammetry: A Case Study with Deer Antlers. Evolutionary Biology, 47(2), 175–186. https://doi.org/10.1007/s11692-020-09496-9

Zhou, H., Dong, W., and Chen, J., 2025. LITA-GS: Illumination-Agnostic Novel View Synthesis via Reference-Free 3D Gaussian Splatting and Physical Priors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 21580–21589. https://doi.org/10.1109/CVPR52734.2025.02010