MLZ 3D Digitization Protocol

Joshua J. Medina, James M. Maley, Siddharth Sannapareddy, Noah N. Medina,
Cyril M. Gilman, John E. McCormack*

Moore Laboratory of Zoology, Occidental College, Los Angeles, CA 90042

1. setup

1.1. Camera choice:
Though nearly any camera can be used for photogrammetry, a consumer grade DSLR or mirrorless camera will provide the functionality needed for most scans. The camera equipment you choose will largely depend on your budget, your subject matter, and the scope of your project, and can radically change a camera's scanning capabilities.

For scanning bird specimens, we use a Sony a7rii, a full-frame mirrorless camera, largely for its sensor. The sensor quality of your camera sets an upper limit on a resulting scan's geometric fidelity; high-res photos make it easier to resolve high-res scans. Importantly, a high-res sensor can also increase the lower limit of fidelity. This is especially important when digitizing large specimens, where a longer shooting distance might result in a significant loss in resolution.

We found that a 42 megapixel sensor allowed for high-resolution detail without the need for supplemental close-up shots that, given a collections setting, represent a substantial draw on time and manpower. High resolution photos lead to longer processing time, but this also means that fewer photos are required for the reconstruction process.

When considering a camera, it's important to remember that transparency, reflectivity, and fine detail such as hair, can also be difficult or impossible to recreate from photography. While it may be useful to consider supplemental scanners, there are some workarounds. Transparency or reflectivity can be obscured, or diffused with filters in order to avoid problemed scans. Fine, thin, and/or organic detail can be accounted for with the 'brute force' of larger sensors or specialized lenses.

1.2: Lens choice:

The lens used depends on your subject matter and desired scan resolution. Choosing a camera platform will affect the variety of lens available. In order to capture the fine detail of bird skins, we use a 90mm macro lens. Larger specimens will require a lens with a wider field of view and deeper depth of focus. When capturing larger specimens, we use it in tandem with a 15mm wide-angle lens. Since we are using a macro lens, we have to keep a couple of variables in mind to maintain sharpness and a proper representation of the subject. The depth of field will be shallow, so we need the fstop to be high. The magnification should be 1:1 magnification and the focal length should be fixed. We will get to this when discussing camera settings. Future setups may resort to focus-stacking, but current methods multiply the time commitment for image capture.

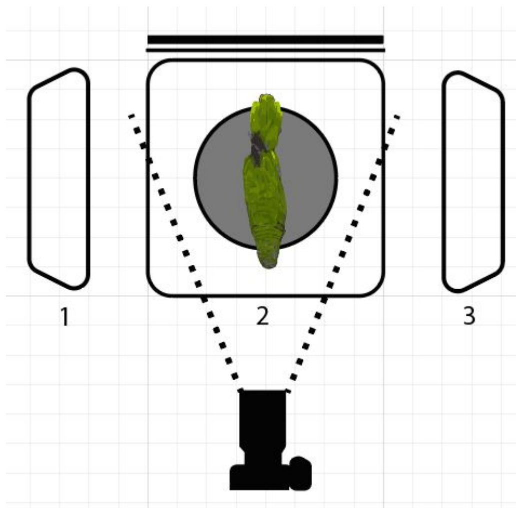
Specular highlights and reflections will confuse the scanning process. A polarizing filter can be used to polarize the light, in order to remove reflection and capture diffuse albedo in the final scan.

1.3: Light Setup:

Lighting can drastically change the outcome of a scan. Excess light and shadow will bake itself into the model's texture. Reflections and bounce-light can make subsequent photos, creating gaps in your image registration.

It is important to set up even, diffuse lighting. Since both a scan's texture and geometric form are constructed from pixel color, it is important that light produces colors accurately, evenly, and predictably. Diffused, even light will help you fight highlights and shadows that get in the way of good texture baking and accurate geometric capture. Hemispherical coverage is ideal.

To do this, we recommend setting up two boxes on either side of the scanning stage, and one above. A white backdrop is recommended, and while a green screen may help in masking, it also risks tinted bounce light. It can be useful to calibrate lighting ahead of time to ensure lighting is standard between shooting sessions and setups. Light is additive, so it may be necessary to test them each individually first, then together. (Depending on the shutter speed, kill the room lights so they don't interfere with the exposure.) The recommended exposure will depend on your camera and settings.



An overhead view of the camera setup.
1, 2, and 3 denote diffuse light sources.
Light 2 lights the stage from above.

1.4: Stage Setup:

For low to moderate budgets, we recommend using an automatic turntable to stage your specimens. The turntable should be set at table-height and centered evenly between the hemispherical lights suggested in 1.3. The tripod-mounted camera faces the subject from the open side. It should be at a distance where the subject is fully in frame, a distance which depends on the subject's size, the camera's focal length, and the sensor size of your camera.

If you plan on capturing an object's underside without combining data from multiple scans, it should be placed on a stand. Stands should be both non-invasive enough to minimize impact on the scan's surface, yet sturdy enough to stay rigid when the turntable actuates. The stand should not obstruct the view of the camera during a pre-set rotation. The stand is also a good place to affix scale references such as a ruler or coded markers.

For our bird skins, we use a small stage consisting of a motorized turntable [Comxim] that interfaces with the shutter release of our camera. Taking a picture actuates the turntable, which after turning, signals for another picture. We position coded markers and a scale marker in order to be visible along with the subject. For scale calibration, We place a scale bar right below the bird. For color calibration, we use an X-Rite Colorchecker color chart.

2. image capture

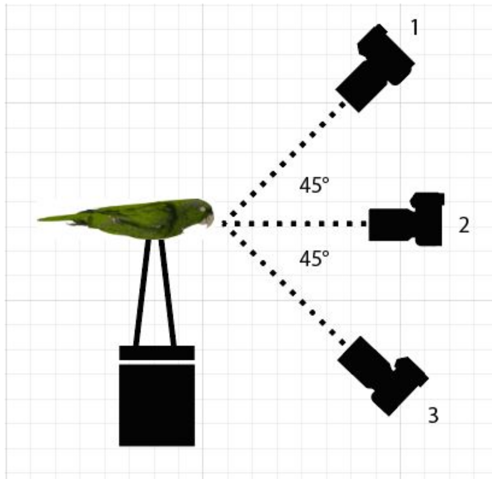
2.1: Camera settings:
Like with lighting, camera settings affect the pixels and thus affect both texture quality and accurate geometric reconstruction. Because image quality directly affects scan resolution, it is important to minimize noise and maximize sharpness in every photo. Blurry photos will both hamper image alignment later on in the process, as well as forming artifacts during texture generation. If a low ISO and a high f-stop significantly decrease the exposure and begins to crush colors, consider using stronger lights or a slower shutter speed.

Considering we use a lens with low depth-of-focus, we use an f22 aperture for maximum depth of field in each shot. We also aim to drop the ISO to 400-800 in order to avoid noticeable noise. We take photos at full resolution, formatted in RAW+JPG. Since we are using an automated setup, we make sure to check the subject's framing and focus before each rotation.

2.2: Documentation:
When recording camera settings, we also take time to document the subject and the shoot. This is done in two steps:
1. Document the name, tag, number, and any other data relevant to the subject and its organization. We take a hand-held photo of the collection tag.
2. Document the ambient color of the setup, for white-balancing and color correction later down the pipe. This can be done by taking a single photo of a ColorChecker or other similar color chart. Make sure the chart is fully in frame and in focus, and using the same camera settings and lighting conditions that you plan to use for the scan.

2.3: Capture:
Photogrammetry software requires consecutive series of photos (at varying angles) to accurately assemble or "solve" the 3d model. In order for the photos to properly register for the scan, a successive series of photos require 50% or higher overlap. When automating the capture, it’s important to keep overlap in mind when determining photo count. When using our turntable, we set it to take a photo every 3-4 degrees until a full rotation has been captured.

We use 96 shots and three different position, or angles: one level with the turntable, one at a 45° angle facing the subject from above the first position, and one at a 45° facing the subject from below. This ensures coverage enough coverage to capture the full subject, while also maintaining the minimal vertical overlap needed for reliable image registration. The process takes 7 minutes per rotation, and the tripod is manually adjusted to the next position between rotations. With the use of three cameras, the shooting time would be cut from 21 minutes to 7 and require no manual adjustment.



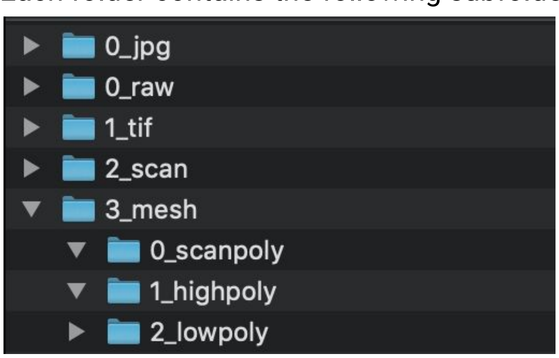
The tripod-mounted specimen is photographed from three different camera positions in order to ensure maximum coverage.

3. image processing

3.1 File Organization
File management is crucial to maintaining an automated/procedural workflow, as mismanaged file paths can break your pipeline. Choose a reliable format, stick with it, and try not to move anything.

We store each scan in the following format:
Date(YYYYMMDD)_Name_Collection#

Each folder contains the following subfolders:



3.2: Photo Processing
In order for the scan to faithfully represent the subject matter, the raw photos should be processed in an image editing software. We recommend Adobe Lightroom for its ability to efficiently manage large batches of photos and its extensive library of plugins and add-ons.

Processing RAW images in Adobe Lightroom:
After the raw photos from image capture have been relocated to the 0_raw folder, they are imported into lightroom through linking their corresponding folder. From here, the following alterations are made:

i. Out-of-focused or obviously troubled photos are removed from the set.
ii. The color checker profile is exported as a new color profile. Lightroom is restarted, and the color profile is loaded into the photo of the color chart.
iii The exposure of the photo will be raised or lowered until the 2nd grey square on the chart is within a 1% margin of 75% grey.
iv. The photo is white-balanced using the 75% grey color chip.
v. Lens profile corrections are applied to the photo, reducing photo distortion and color aberrations according to the make of the camera and lens.
vi. (Optional) Sometimes, a photo set is manipulated to encourage better photo alignment. Settings that affect alignment include lowering highlights and raising shadows in order to flatten the images as much as possible, as well as increasing contrast from 5 to 15%.
vii. The total exposure values are matched across the photoset, using the color checker photo as reference.
viii. The photo settings (minus exposure) are synched across the photoset, using the color checker photo as a reference

The Photos are then exported in one of two formats to the 1_TIF folder. Ensure that the exported images maintain the names they were imported with.

- 8-bit TIFs: the smallest lossless filetype that contains all the information useful for a scan. These photos are lossless, but due to the volume of photos taken, take up significant space.
- High-quality JPEGS. From our testing, high-quality JPGs rendered models of comparable resolution, and take up significantly less space to store. The exported folder should be renamed to 1_HQJPG or equivalent.

3.3: Masking

To optimize photo alignment, the exported photos should be masked. Different subject matter will require different types of masking. Generally, a mask modifies an image in order to obscure the backdrop and isolate the subject. Thought masks can be applied using a number of image editing programs, we use Adobe photoshop for its accessible automation tools.

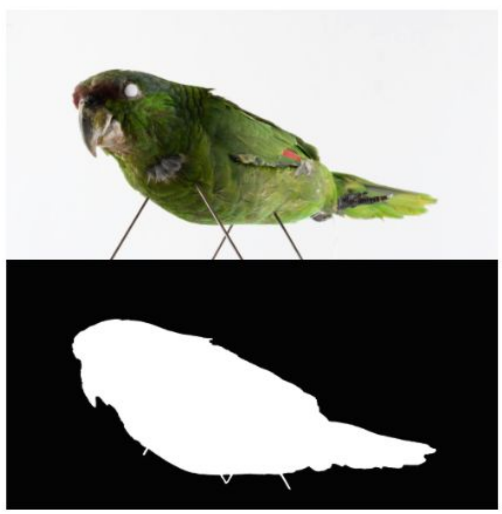
The type of masks we makes depends on the software used for image registration [See section 4.] When using Reality Capture, masking data is stored in the Alpha channel of our exported TIF images. These masks will use transparency to mask out the background of each image. If we are using Metashape, the masked are stored as separate greyscale JPEGs. The white portions will be enabled during image registration, while the black portions will be masked out. Because these masks are generated from the JPEGs captured by the camera, this action can run concurrently to image processing in Lightroom.

When using photoshop, different masking processes or 'actions' can be batched. In Photoshop, an 'action' macro is automatically applied to the photo set, and the photos are exported under 2_mask with the suffix "_masked."

i. If Greyscale masks are required, we import JPEG photos from 0_JPG. The Greyscale macro selects the background and turns it black, inverts the selection, and turns the subject white. The completed greyscale image is then exported to 1_mask, then the batch resets and selects the next photo to mask and export. The exported files will correspond to their counterparts in 1_TIF, over which it will be overlayed inside of Metashape during image registration.

ii. Oversized or oddly-shaped subjects may require the fine-tuning of masking macros. Backup macros are used in these cases, that key the backdrop at different tolerance levels, as well as batches that use other methods of background-detection such as the selection of a color range. Photoshop's "Select Subject" tool has been our most error-free means of quickly producing accurate selections for both Alpha and Greyscale macros.

iii. If Alpha masks are required, we import the TIF photos from 1_TIF. The Alpha Mask macro follows the same steps as the Greyscale macro, only it stores the mask information in the fourth “alpha” channel used for transparency.



Each photo gathered from 0_jpg is given a corresponding Mask, with a black background and a white foreground. When overlaid during image registration, only the white portion will be used to construct the point cloud.

4. image registration and point-cloud generation

Now that the images have been processed, photos are loaded into software that aligns them and generates the initial 3D data that will become your final scan. While there are multiple scan-solving software to choose from, the best one depends on your use case, subject matter, computer hardware, and experience. When the photos are correctly captured and processed, aligning them should be simple and hands-free.

4.1. Software Choice

First, we choose the appropriate software. We use two different image registration packages. Along with registering images and generating point clouds, they both offer meshing and texturing capability. While Reality Capture is the better option for large collections, both programs are recommended according to their particular advantages:

Agisoft Metashape

Positive:

- Forgiving and consistent image alignment, and intuitive image alignment troubleshooting.
- Supports 32bit images, and exports 32 bit textures
- User-friendly and intuitive interface
- Supports automating point-cloud registration.

Negative:

- Significantly longer processing time than reality capture.
- Asset resolution is limited by the PC hardware being used

Reality Capture

Positive:

- Extremely fast processing time
- Less dependent on high-end PC hardware
- Able to generate extremely sharp and detailed results

Negative:

- User interface less friendly
- Less forgiving with
- Point-cloud registration is manual, and less intuitive

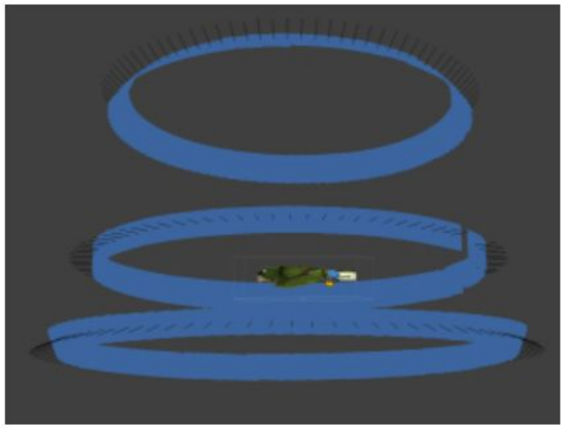
4.2 Image Registration

Regardless of the chosen software, image registration and point cloud generation typically follows the same trajectory.

I. After the software has been chosen, we import the TIF photos for registration. When using reality capture, the masked TIFS can be imported straight from 1_mask. When using metashape, the TIFS can be loaded from 1_TIF and the masks can then be loaded separately from 1_mask.

ii. Photos are aligned, automatically generating a point cloud. In Metashape, this process takes anywhere from 1 to 4 hours using our recommended settings. In Reality Capture, it takes between 10 and 30 minutes. When using default or medium settings, a decent point cloud will have in excess of 200,000 points and should recognizably resemble the subject.

iii. When correctly aligned, the camera positions should read as three perfect rings per rotation set: each ring representing a 360 degree rotation of 96 individual photos. At this stage, remove any photos that fall outside of these rings or that don't contribute to the visible point cloud.



In Metashape, simulated camera positions appear as blue rectangles. Three distinct orbits represent the three camera angles and the rotation of the turntable. The presence of the complete rings suggests good photo alignment.

4.3 Point Cloud Editing

Before generating a mesh from the point cloud, it is important to double-check image registration to make sure that the photos have properly aligned. After a point cloud is generated, you should have enough visual feedback to note the incomplete data of the scan, usually towards the direction of the stand the subject was resting on. Editing point cloud typically follows a series of steps.

i. Look for misaligned cameras. They will fall outside of the uniform circular rings visualized in the software's viewport, and often correspond with errors in the point cloud. Reset misaligned cameras, then try re-registering them.

ii. Look for missing data in the point cloud. If the subject's silhouette is visibly noisy or isn't visible at all, some of the cameras may have failed alignment. Try re-registering photos using different alignment settings, such as image overlap and alignment quality.

iii. If automatic photo registration is failing, you can manually adjust photo alignment align by placing markers. Some software, like Agisoft Metashape, allow you to place markers directly onto the 3d model in the viewport. Otherwise, you can place markers directly onto the photographs. Place them on visual landmarks visible from multiple photos. Often three or more photos are required for proper alignment.

iv. Lastly, the scale is defined by placing markers on a defined unit of measurement present within the photo. A scalebar appropriate to the subject's size is often placed on the stand for this purpose. Two markers should set the boundaries of the largest unit of measurement you want to use to calibrate the scale of your model. Once placed, a digital scalebar can be registered between the points. After the scale accuracy has been set and checked, the model must be updated with the new scale reference in mind.

v. We finish out work in the Image Registration software by generating a high-poly mesh from the finished point cloud. The geometry is generated at the highest resolution possible; a point cloud of 200,000+ should generate a mesh with a polycount exceeding 500,000 triangles. The new mesh is exported under 0_scanpoly withing 3_mesh as an OBJ file with the "_scanpoly" suffix.

5. mesh processing - manual

Image registration and Mesh Processing largely occur without supervision. However sometimes it is useful to check the mesh for errors and, if needed, to manually address issues with the geometry before sending it further down the pipe. Manual cleanup of high-poly meshes should be preformed in a 3D sculpting package. We use Pixologic ZBrush because of its agility with very high-polycount meshes and its "tessemation" (tesselation + decemation) toolset. In ZBrush, we take the following steps:

i) Small excess pieces are removed.
- In the Polygroup menu of the Tool palette, select 'Autogroups'.
- Select the body of the main scan geometry. The extraneous pieces will disappear.
- In Geometry menu of the Tool palette, under Modify Topology, select 'Delete Hidden.'

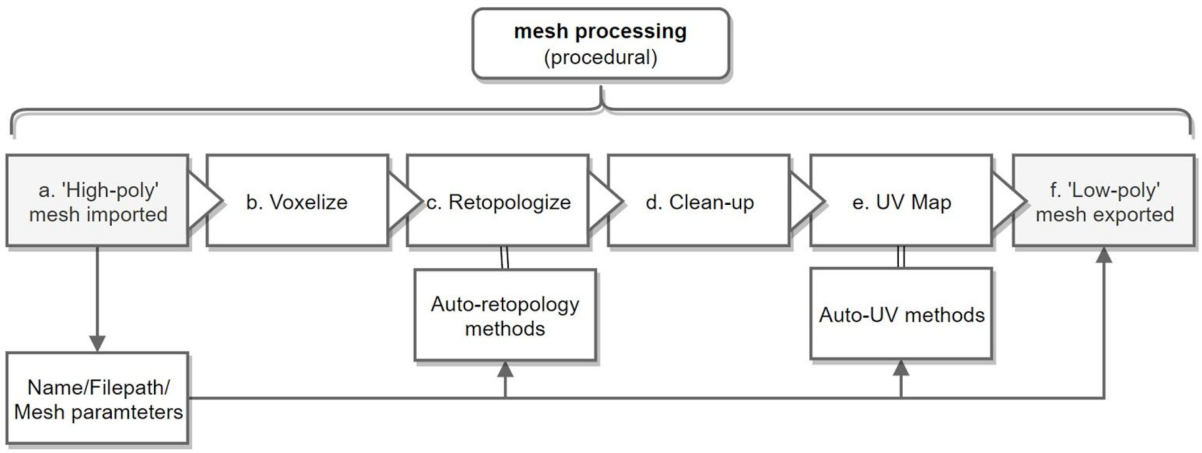
ii) The Stand and other extraneous geometry is removed
- The Select Lasso tool is used to remove visible portions of the stand.
- Holes are filled by accessing the Geometry menu, and under Modify Topology, selecting 'Close Holes.'
- The Sculptris brushes are used to any leftover noise or stand geometry.

iii) The Export parameters are changed to match the desired output
- Triangles is selected to export a triangular mesh.
- 'no groups' is selected to avoid exporting polygroup information from Zbrush.

The result is exported under 1_highpoly withing 3_mesh as an OBJ file with the "_highpoly" suffix.

6. mesh processing - procedural

To ensure the most flexible possible results, we run the "high-poly" mesh through a series of steps that control and refine its resolution settings, fix any topology errors, and smoothly export the remaining files. Each step represents a node, or series of nodes in SideFX Houdini's procedural node graph. Our node graph works like this:



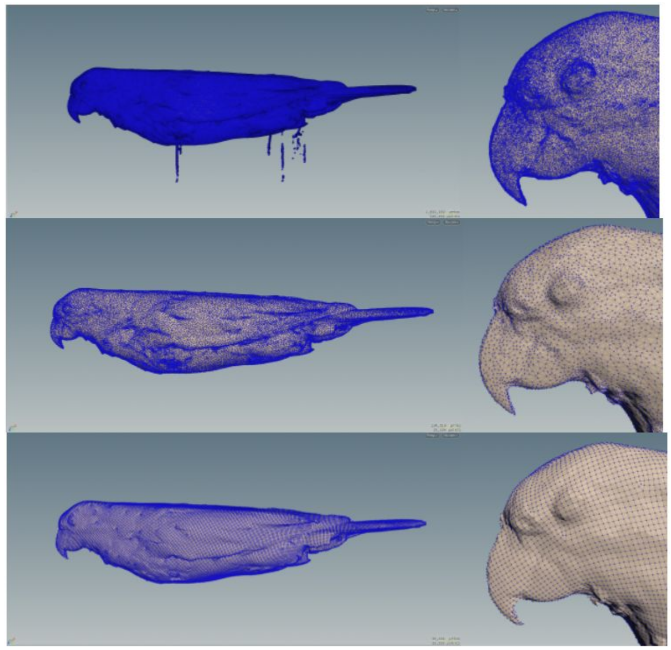
A. 'High-poly' mesh imported
The mesh is imported from the 1_scanpoly or 2_highpoly subfolder within the 3_mesh folder. Parameters related to retopology and UV mapping can be set during import.

B. Voxelize and C. Retopologize: First, the high-poly mesh is voxelized at a resolution exceeding that of the input polycount. This ensures that the topology used for decemation is evenly distributed across the models surface. The polycount is then reduced through decimating the mesh according to an angle-based "quality tolerance" that conforms to the shape of the original high-definition model. Though this will lead to some (minor) variation in final polycount, this ensures that the decemation is controlled according to an error threshold instead of a pre-specified percentage or target polygon count. The quality tolerance should be adjusted with respect to your intended geometric resolution.

D: Mesh Cleaning: On occasion, the retopology node can introduce minor geometric errors or expose holes left in the previous scan. This can lead to a number of topology defects, many of which are invisible to the eye but impact UV Mapping and rendering down the line. Houdini comes packaged with nodes that automatically resolve overlapping polygons, non-manifold geometry, holes, and other such errors. We have strung together a divide, clean, delete small parts, and a pollyfill node to deal with such errors.

E: UV Mapping: To match the new topology, the mesh is given UV coordinates. When choosing a UV algorithm, the aim is to balance maximally efficient texture-usage with even texile density across the model. Houdini's autoUV node contains multiple methods for automatically generating UV coordinates. We use the 'UV Unwrap' method with 8 simultaneous planar projections. For quicker results, Houdini's AutoUV SOP has automatic seam-drawing parameters that can be tuned to match the shape of your model.

F: ‘Low-Poly’ Export: The final node of the Houdini procedure exports the mesh in the 3_lowpoly subfolder of the 3_mesh folder in the main directory provided by the import node. This final 3d model, with a "_lowpoly.obj" extension, will be the most optimized version of the model to texture and upload.

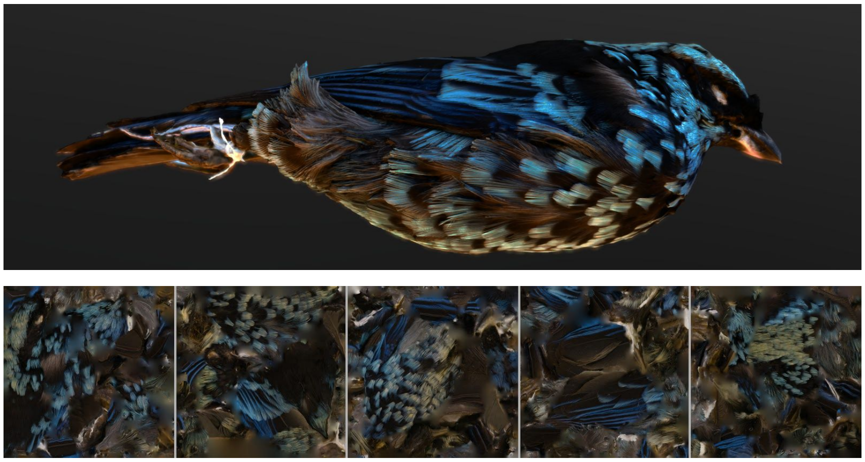


Three different levels of decemation with vertices highlighted in blue and close-ups shown to the right. Iterations like these can be generated and compared in quick succession within Houdini. The Last iteration uses quad-based topology instead of triangles.

7. texture generation

After processing, the retopologized low-poly model is imported back into the Image Registration software, to replace the previous high-poly model. From here, the previously-aligned photos can be used to generate a texture map using the fresh UV coordinates. The resolution of your texture cannot exceed the resolution of your camera's sensor, which becomes especially apparent rendering textures. We usually render 4k textures for our models.

After generating the texture, export it as a full-sized .jpg into the same 3_lowpoly folder from which you saved the low poly model with proper UVs.



Texture information is stored in image files called texture maps, which flatten the 2D surface of the 3D model onto an image plane and assign color values to the corresponding coordinates.

8. finalization

The final mode and texture file is taken from the “3_lowpoly” folder loaded into a 3D viewer to finalize geometry and positioning. The scale calibration should be double-checked by comparing measurements on the 3D model to its real-world counterpart. Since these models are optimized for real-time rendering, they can be shared online and loaded into 3D viewers and other scenes. Additional texture maps and post-production can be applied to better characterize shiny, translucent, or even transparent surfaces that were lost to the digitization process.