TeleSculptor

Contents

Introduction

Welcome to TeleSculptor. This software is free and open source software developed by Kitware under multiple SBIR contracts for DARPA and AFRL. The purpose of the software is to calibrate cameras and extract 3D information from video and images, especially aerial video. The software is designed to have different workflows that can either improve the manual modelling process in SketchUp or provide fully automatic 3D reconstruction.

TeleSculptor is a cross-platform desktop application for Windows, Linux, and Mac. This document uses the Windows version of the application to illustrate usage, but everything should work nearly the same on other platforms.

TeleSculptor v1.2.0 marks the third official release of the software. Please refer to the Release Notes for details on all the changes. While TeleSculptor continues to become more broadly useful across imaging scenarios, it is still optimized for the use case of aerial video flying a 360-degree orbit around a target object. TeleSculptor will give best results on a video clip containing one orbit around an object. We will improve robustness on more general imaging scenarios in future releases.

Branding: Origins and Destinations

TeleSculptor has its origins in a software project developed as "MAP-Tk", the Motion-imagery Aerial Photogrammetry Toolkit. The original software was not an end user application but a collection of developer tools and libraries. As the software evolved, it developed a graphical application that we now call TeleSculptor. At the same time, the software libraries of MAP-Tk were reorganized into a new, broader toolkit called KWIVER (Kitware Image and Video Exploitation and Retrieval). So, the original MAP-Tk has dissolved away leaving behind a TeleSculptor application powered by KWIVER. However, some uses of the name MAP-Tk may persist for historical reasons.

Getting Started

- For instructions on how to install Telesculptor, see *Installation*.

- To start a new Telesculptor project, see *Processing Steps*.

TeleSculptor-1.1.
0-Windows-AM
D64.exe
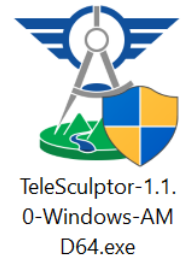
3.1 TeleSculptor Installation

The TeleSculptor software uses a standard installer package on Windows. To install, double click the installer icon and step through the installation steps as shown in the images below. This will require administrative privileges. The name of the 64-bit Windows installer is TeleSculptor-1.2.0-Windows-AMD64.exe.

To run the application, find TeleSculptor in the Start Menu and click the icon. The program will open with an appearance as shown below.

Once the application is open, you can access this documentation from the *Help* menu or by pressing the F1 shortcut key.

3.1.1 SketchUp Plugin Installation

The TeleSculptor application also comes with a plugin for SketchUp that allows SketchUp to read TeleSculptor project files. This is not installed automatically. To install the SketchUp plugin, first locate the plugin in your TeleSculptor installation. If TeleSculptor is installed in the default location, you will find the plugin at:
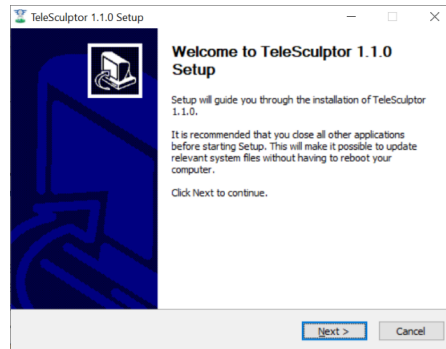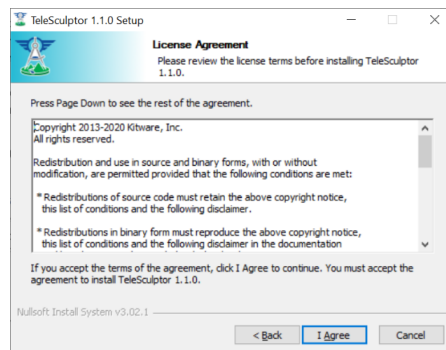
Fig. 1: 1. Click "Next" to begin the setup

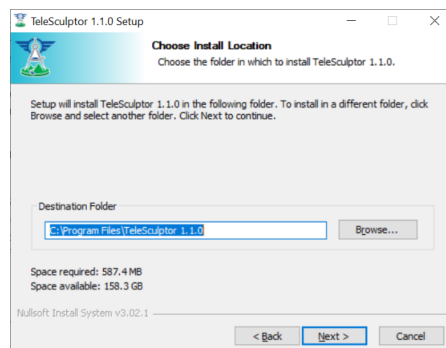Fig. 2: 2. Click "I Agree" to accept the BSD license

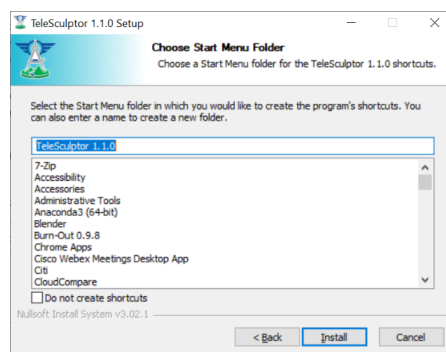Fig. 3: 3. Click "Next" to accept the default destination

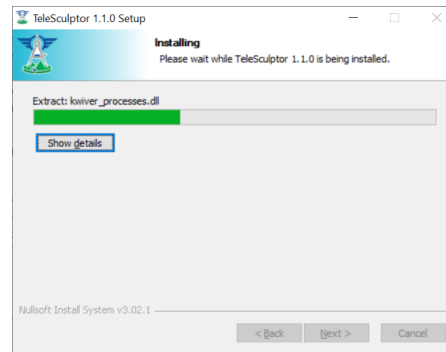Fig. 4: 4. Click "Next" to accept the Start Menu location
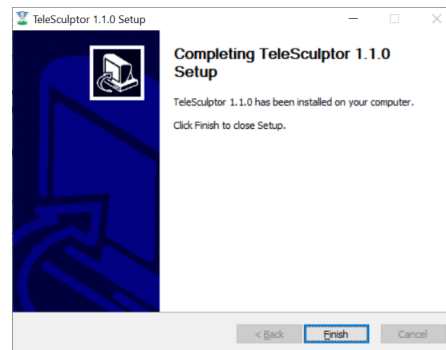
Fig. 5: 5. Wait for files to install

Fig. 6: 6. Click "Finish" to complete

```
C:\Program Files\TeleSculptor 1.2.0\share\telesculptor\1.2.
↪0\plugins\sketchup\kw_telesculptor.rbz
```

To install this plug in SketchUp, first open SketchUp. Next navigate to the *Window → Preferences* menu as shown in the figure below. Within the *System Preferences* dialog, click on Extensions in the menu on the left. Now click the Install Extension button at the bottom of the dialog. Use the file open dialog to locate the kw_telesculptor.rbz file at the location above and click the Open button. Administrative privileges are needed to complete the installation. Once installed, "TeleSculptor Importer" will appear in the Extensions list and the box next to it should be checked to activate the plugin. The plugin may not be fully active until the SketchUp application is closed and re-opened.

3.2 User Interface

3.2.1 Overview

The GUI consists of a primary (world) view, secondary (camera and depth) views, a frame selection panel, and various other ancillary views and panels. The world view is a three dimensional view that shows the reconstruction results and camera poses in the computed world coordinate system. All other views and information panes are dockable, and may be rearranged, resized, detached or closed (hidden) according to user preference. Closed panes may be reopened via the View menu, or using the context pop-up menu of the main menu bar or title bar of any pane.

The secondary panes are:

Camera Selection: Provides controls to view and select the active frame, and to cycle through the same in a manner similar to video playback or a slide show.
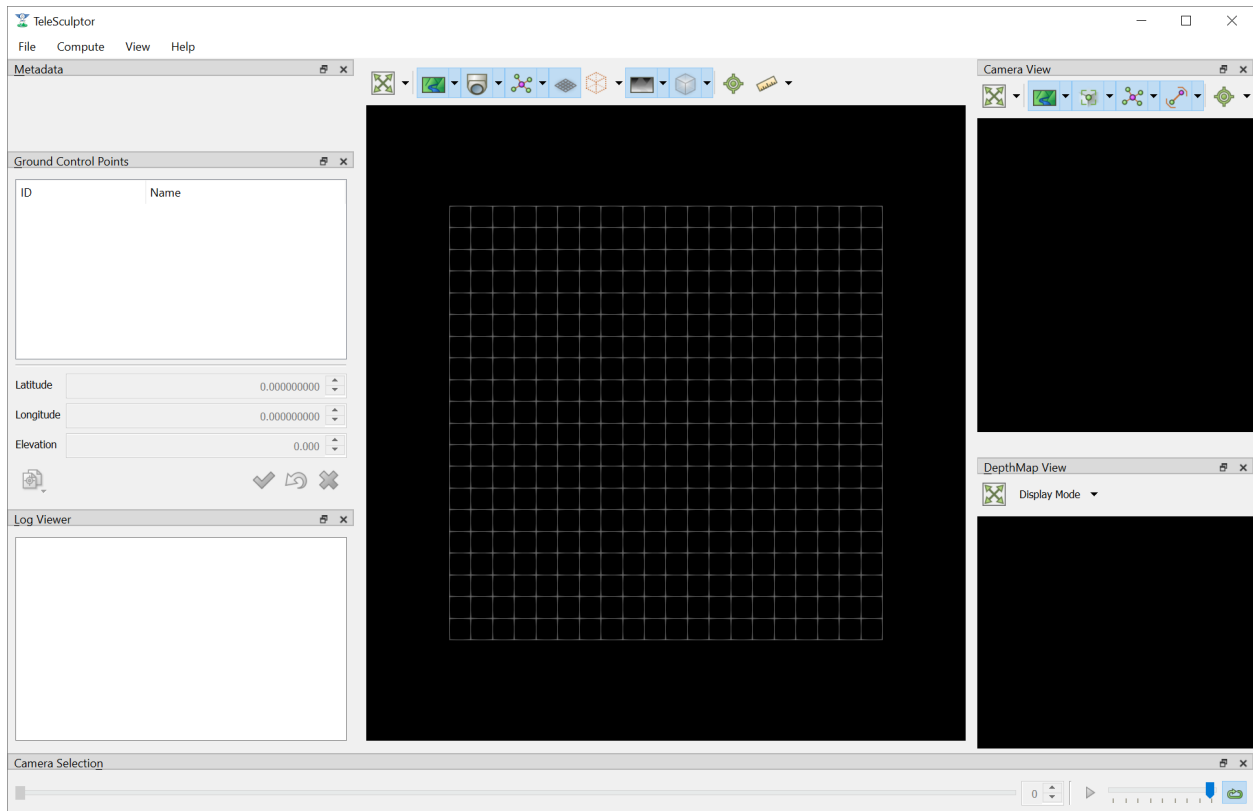
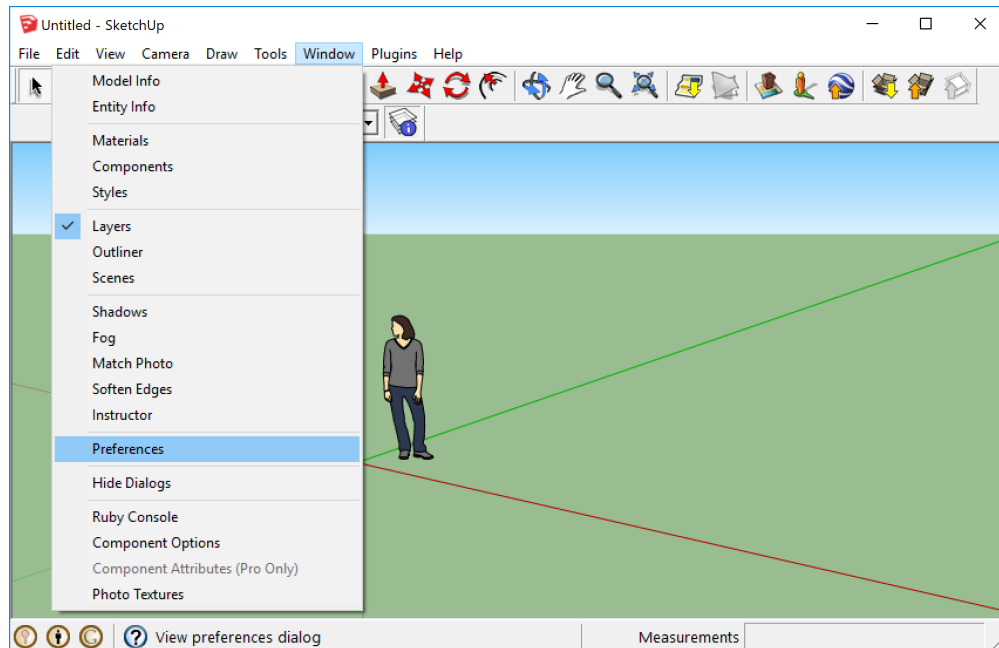Fig. 7: The TeleSculptor Application when first opened.

Fig. 8: Access the SketchUp Preferences to install the TeleSculptor plugin.
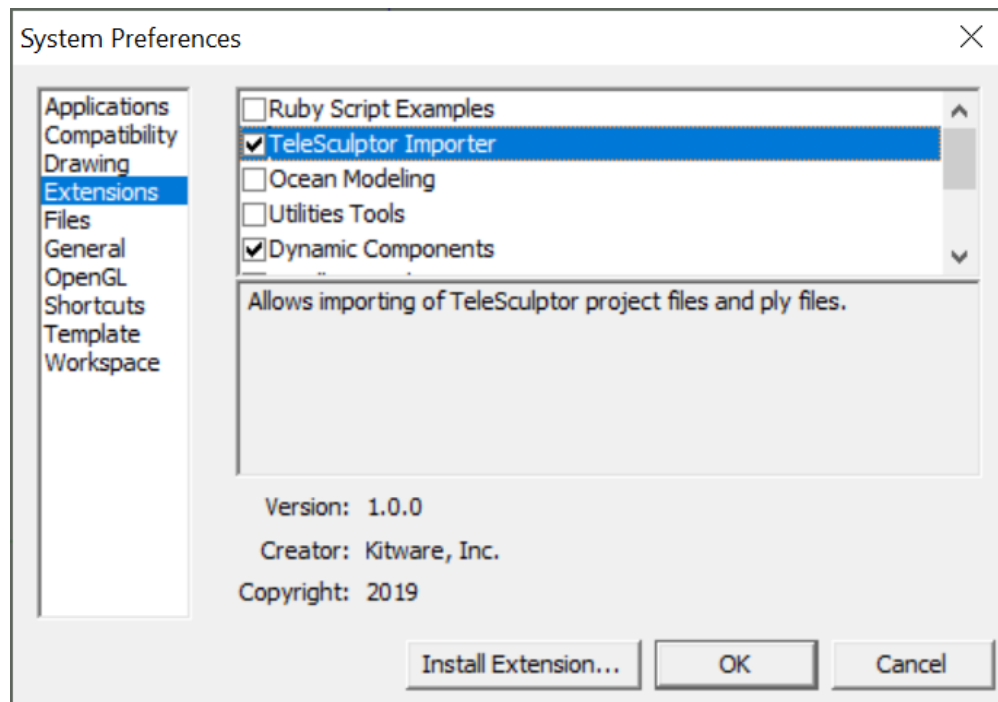
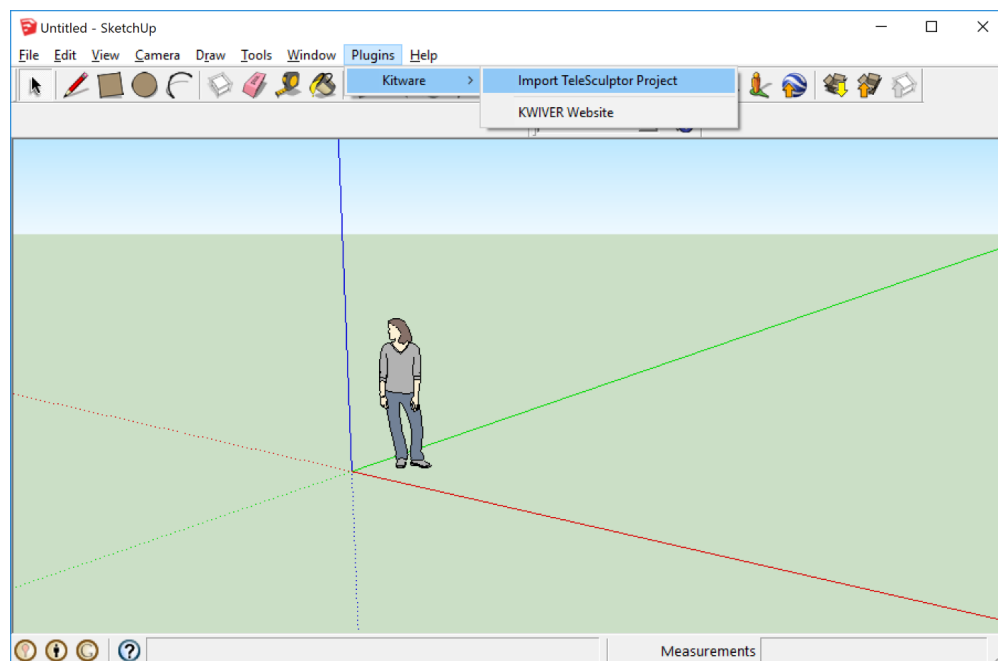Fig. 9: TeleSculptor Importer installed and activated in the SketchUp Extensions list in System Preferences.

Fig. 10: The Import TeleSculptor Project option in the Plugins menu after installing the extension.

Camera View: Shows the imagery from the current frame, along with various two dimensional geometry. If a camera model is available, also shows three dimensional geometry mapped into the camera's point of view, as well as estimation residuals.

Depth Map View: Shows the most recent estimated depth map as a two dimensional image.

Metadata: Shows the metadata values associated with the current frame image.

Ground Control Points: Shows a list of ground control points and/or camera registration points, as well as details (including geodetic location) of the selected point. Also contains controls for editing or deleting points.

Log Viewer: Shows the log output for the application, particularly from running algorithms.

3.2.2 World View

The world view provides a three dimensional view showing all three dimensional products in a common coordinate system. It is especially useful for viewing the spatial relationships between cameras and landmarks, ground control points, dense point clouds, or surface meshes.

Landmarks are displayed as points, while cameras are displayed as frustums of semi-arbitrary length, with the active camera highlighted (see *Camera Options*). The "camera path", which is a polyline connecting the camera centers, can also be displayed. Ground control points are displayed as "jacks"; three short lines meeting at right angles in 3D.

Additionally, the world view can display a representation of the ground plane (i.e. the local coordinate plane z = 0), and can display the image for the active frame projected to that plane. These projections result in a stabilized view of the video if the scene is largely planar and that plane has been aligned with z = 0. Products from video with geospatial metadata are automatically translated to the z = 0 plane. For data without geospatial metadata, the Estimate Cameras/Landmarks tool attempts to fit a plane to the landmarks as it estimates them to reorient the landmarks to z = 0.

The world view also supports visualization of the depth image for the current frame as either a dense RGB point cloud or a surface mesh. It also can render a three dimensional surface mesh; either a "volume mesh", extracted from the level set of a volumetric array, or an externally provided mesh. The volume is typically the result of the fusion of several depth maps using a truncated signed distance function or similar operator.

Navigation

To rotate the view, press the primary (usually left) mouse button and drag. By default, the view will rotate about the view's "look at" point, which is initially, or after resetting the view, at or near the center of the scene. Holding Ctrl will rotate around the view axis (i.e. will "roll" the view). To pan, hold the middle (wheel) mouse button while dragging. To zoom, hold the secondary (usually right) mouse button and drag, or rotate the mouse wheel. A double-click will set a new "look at" point, centering the view on that point and changing the rotation center.

Tool Bar

Reset View Resets the view extents so that the entire scene is visible. The keyboard shortcut R provides the same effect. Additional actions are available via the action's associated pop-up.

Zoom to Landmarks Resets the view extents so that all landmarks are visible. This action is available via the Reset View button's associated pop-up menu. The keyboard shortcut is Shift+R.

View to World Top/Left/Right/Front/Back Resets the view rotation to a "standard" rotation, such that the view axes are parallel with the world axes. These actions are available via the ⊠ Reset View button's associated pop-up menu.

Perspective Toggles the world view between perspective and parallel projection. Perspective projection more closely models human vision and is often useful for visualizing depth in the scene. Parallel projection can be useful for viewing the scene in profile. This action is available via the ⊠ Reset View button's associated pop-up menu. The keyboard shortcut is P.

Show Camera Frame Image Toggles visibility of the frame image projected onto the ground plane (if an associated camera model is available). The associated pop-up allows the opacity of the same to be adjusted.

Show Cameras Toggles visibility of cameras and related visualizations. The associated pop-up provides additional options; see *Camera Options*.

Show Landmarks Toggles visibility of landmarks. The associated pop-up provides additional options; see *Point Options*.

Show Ground Plane Grid Toggles visibility of the ground plane grid. The ground plane is the z= 0 plane in local 3D coordinates. The grid is centered about x = y = 0, however the grid lines are otherwise strictly aesthetic and do not correspond to any particular values. The size of the grid adjusts to fit the scene.

Show/Edit Region of Interest Toggles visibility of the region of interest selection in the world view. While visible, the ROI may be resized by clicking and dragging on any of the six handles on the faces of the ROI box.

Reset Region of Interest Resets the region of interest to the axis-aligned box containing 80% of the landmark points plus an additional 50% padding. This action is available via the ⬡ Show/Edit Region of Interest button's associated pop-up menu.

Show 3D Depth Map Toggles visibility of the depth map (if available) rendered as a 3D point cloud or mesh; see *3D Depth Map Options*.

Show Surface from Volume Toggles the visibility of the surface mesh extracted from volumetric data; see *Volume Surface Options*.

Edit Ground Control Points Toggles editing of ground control points. See *Editing Ground Control Points* for details. The keyboard shortcut is Ctrl+G.

Enable Measurement Tool Toggles placing or editing of the ruler measurement tool. Initially — when the ruler has not yet been placed, or after it has been removed using Reset Measurement Tool — a ruler can be placed by clicking two points in the view. The depth of the points is calculated based on landmarks, depth map points, or the surface mesh location under the mouse cursor. Turn off visibility of objects to avoid selecting them. The ground plane is selected if no nearby geometry is found. Once placed, the ruler's points may be moved freely. Holding down the z key constrains the ruler to vertical measurements. Holding down the x or y keys constraints the ruler to a horizontal plane. Placement of the ruler may be canceled by pressing the Esc key before placing the second point.

Reset Ruler Removes the currently placed ruler. This action is available via the ⬓ Enable Measurement Tool

---

button's associated pop-up menu.

Camera Options

The ⬤ Show Cameras pop-up provides additional controls that can be used to control the display of the cameras in the world view. These allow changing the color of both the active and inactive cameras as well as the camera path, changing the size of the camera frustums, and toggling visibility of the inactive cameras and camera path separate from the overall camera visibility.

The camera scale controls are relative to a "base size" that is computed from the extents of the scene data. The inactive camera scale is relative to the active camera scale, with the maximum allowed value giving active and inactive camera frustums the same size.

Point Options

The ⚛ Show Landmarks pop-up provides additional controls that can be used to control the display of the landmarks in the world view. (The same controls are also used in the camera view to manipulate the display of feature points and landmarks in that view.) These allow the color of the items to be changed, as well as their size. Feature items (that is, feature points and landmarks) are displayed as dots, with a fixed size-on-screen that is independent of the view.

Several options for color are provided. The simplest is "solid color", which displays all landmarks in the same, user selected color. "True color" displays landmarks in the color estimated to correspond to the actual color of the point in the real world scene, as computed from the input imagery. "Color by data" uses color to visualize other per-point data, such as the number of individual frames that contributed to ("observed") each landmark.

In addition to coloring by data, points may be filtered (selectively displayed) according to their respective values of the currently selected data set. Filtering may exclude points above or below selected lower or upper bounds, or that are not within said bounds.

3D Depth Map Options

The ▭ Show 3D Depth Map pop-up provides additional controls on the display of depth maps in the world view. The options allow the depth map to be rendered either as a 3D point cloud (one point per pixel) or a dense triangular mesh (one vertex per pixel). In either case, the rendered depth data is colored by the RGB color values of the corresponding video frame. A filter option is also available to remove depth points based on thresholds on various attributes. Currently, these attributes are the Weight and Uncertainty. Images of these attributes as well as the depth map itself are also shown in the Depth Map View and the filter options selected here apply to that view as well. See *Depth Map View*.

Volume Surface Options

The ⬡ Show Surface from Volume pop-up provides additional controls on the extraction and coloring of a surface from volumetric data. The "Surface threshold" parameter controls the value of the isosurface at which the surface is extracted from the volume. The "Colorize surface" option, if checked, allows coloring each vertex of the mesh. The "Current frame" mode projects the RGB values from the current frame onto the mesh, while the "All frames" mode combines appearance projected from all frames or a subset of frame sampled at a regular interval. The "Color display" options determine how to color the surface. Options include mean color, median color, surface normal, and number of observations.

Editing Ground Control Points

The ⊙ Edit Ground Control Points action allows the user to enter or leave edit mode for ground control points. When not in edit mode, the scene location of ground control points is fixed and cannot be changed, nor can ground control points be selected in the world or camera views.

In edit mode, clicking on a ground control point in either view selects the point in both views as well as the *Ground Control Points* panel. (Selecting a point in the panel also selects it in both views.) Points may be dragged in either view to change their scene location. Holding the Shift key while moving constrains movement to one of the principle axes.

New points may be added by holding the Ctrl key while clicking. When placing new ground control points in the view, TeleSculptor projects a ray into the scene that corresponds to the location that was clicked and selects a location along this ray based on scene elements in the immediate vicinity. If no nearby landmark points are found, the new point is placed on the ground plane.

Pressing the Del key while in edit mode when one of the views has keyboard focus will delete the currently selected ground control point.

3.2.3 Camera View

The camera view provides a camera space view of detected feature points, computed landmarks and ground control points (all projected to the camera space), camera registration points, and the corresponding input imagery, for the active frame. Additionally, the estimation residuals — the difference between landmarks and feature points which participated in computing their estimated positions — can be displayed as line segments between the feature point location and projected landmark location.

Navigation

To pan the view, hold the middle (wheel) mouse button and drag, or drag while holding Alt and the primary (usually left) mouse button. To zoom, hold the secondary (usually right) mouse button and drag, or rotate the mouse wheel.

Tool Bar

Reset View Resets the view to the frame image extents. Additional actions are available via the action's associated pop-up.

Zoom Extents Resets the view extents so that the entire scene is visible. This action is available via the Reset View button's associated pop-up menu.

Show Camera Frame Image Toggles visibility of the frame image. The associated pop-up allows the opacity of the same to be adjusted.

Show Feature Points Toggles visibility of feature points / trails. The associated pop-up provides additional options; see *Feature Options*.

Show Landmarks Toggles visibility of landmarks. The associated pop-up provides additional options; see *Point Options*.

Show Residuals Toggles visibility of the landmark estimation residuals. The associated pop-up allows the color of the displayed residuals to be changed.

⊕ **Edit Camera Registration Points** Toggles editing of camera registration points. See *Manual Camera Calibration* for details. The associated pop-up allows computing a camera model on the current frame using camera registration points. The keyboard shortcut is Ctrl+R.

Feature Options

In addition to active feature points, which have all the options described in *Point Options*, the position of feature points on adjacent frames may also be displayed by enabling Trails. For image collections where frames adjacent in the frame list also have spatially similar camera poses (especially when using consecutive video frames as input), these may be useful as an additional means of visualizing camera motion.

The trail color and length (number of adjacent frames to be used) may be changed, as well as whether to show trails only for lower-numbered frames ("historic" mode), or for all adjacent frames ("symmetric" mode). In all cases, trails are displayed only for active feature points.

3.2.4 Depth Map View

The Depth Map View provides an image viewer similar to the Camera View but specialized to display depth map images. Depth map images are loaded from VTK image (.vti) files associated with a particular video frame. Often there are only depth maps on a subset of frames. The active (or most recent) depth map is displayed in this view by mapping depth to color. The Depth Map View can also display an image representation of other attributes associated with the depth map, such as the image color. Some attributes like uniqueness and best cost are associated with the algorithms used to generate the depth values. The same depth maps can be rendered in the World View as a point cloud. Furthermore, depth map filtering options in the World View also apply to the image rendering of the depth map in the Depth Map View.

Navigation functions the same as the *Camera View*.

Tool Bar

⛶ **Reset View** Resets the view to the frame image extents.

Display Mode Selects which image mode to display in the in the view: Color, Depth, Best Cost Value, Uniqueness Ratio; see *Color Map Options*. The depth filters apply regardless of which image is shown.

Color Map Options

In addition to selecting the mode under Display Mode, there is also an option to select the color mapping function for each mode except Color. The mapping function describes how the scalar data field (e.g. depth) is mapped to color. Below the color map option are the minimum and maximum values from the data used in the mapping. The Auto check box, which is checked by default, indicates that the values are determined automatically from the range of values in the image data. By unchecking the Auto checkbox, the minimum and maximum values of the range can be adjusted manually for finer control of the visualization.

3.2.5 Camera Selection

The camera selection panel contains a large slider used to select the active frame, and used to control which frame's imagery and feature points are displayed in the camera view. If the camera pose for the frame is known, the active camera is also highlighted in the world view, A spin box next to the slider shows the active frame number, and can

also be used to select the active frame. Note that the frame numbers need not be consecutive. Some video readers are configured to only read every *N*-th frame, where *N* may be 10, for example. This helps cut down on data redundancy in video. The frame sampling rate can be configured by opening the project configuration file (.conf) in a text editor.

The controls to the right of the panel control the application's slideshow mode. Slideshow mode automatically increments through the loaded frames at a fixed rate. This can be used to view the feature points for each frame image in sequence. Setting the delay between cameras sufficiently low can be used to simulate video playback for image sequences taken from a motion imagery source.

The slideshow action controls are also available via the *View* menu. The small slider controls the delay between slides. The slider response is logarithmic, with single steps in one-tenth powers of ten. The slider tool tip includes the current delay in human readable units. Several frame filters are also available in the *View* menu. These filters allow limiting the frames to show to a specific subset, such as key frames or frames with tracking data.

3.2.6 Metadata

The metadata panel displays the collection of video metadata for the current frame, if available. The set of fields is selected from the entire data set; individual frames may be missing some or all fields. The metadata itself is provied by the video reader. For encoded video files, TeleSculptor supports key-length-value (KLV) encoding following the Motion Imagery Standards Board (MISB) 0104 and 0601 standards. Customized video readers can read metadata from other sources, such as supplimentary text files or EXIF data.

3.2.7 Ground Control Points

The ground control points panel displays a list of all ground control points and camera registration points in the current data set, as well as detailed information for the selected point. Points have an automatically assigned ID (which may change between sessions) and an optional user-provided name, which may be assigned or changed by editing that column of the point (by double-clicking or pressing the edit key — usually F2).

Selecting a point in the list will select the same point in the world and camera views. Similarly, selecting a ground control point or camera registration point in either view (when the corresponding edit mode is active) will select the same point in the list.

When a point is selected, changing its geodetic location (as described by the latitude, longitude, and elevation text fields) automatically promotes the point to a "user registered" point. These are points for which the geodetic location has been externally measured and is therefore known to be correct. The geodetic location of points which are not user registered is computed from their scene location and the computed scene to geodetic transformation (if available). User registered points are indicated by a glyph () in the ground control point list.

Note that moving a user registered point in the world or camera views (that is, changing its scene location) does not change its geodetic location.

Tool Bar

Copy Location Copies the geodetic location of the selected point to the clipboard. Several options for coordinate ordering and whether or not to include the elevation are provided.

Apply Similarity Transform Estimates a 3D similarity transformation to best align the ground control point locations with the specified geodetic locations. At least three "user registered" ground control points are required. That is, at least three points must have manually specified latitude, longitude, and altitude. The estimated transform is applied to all data (cameras, landmarks, depth maps, etc.).

Revert Changes Reverts user changes to the active ground control point's geodetic location, such that the point is no longer "user registered". This has no effect on points that are not user registered. Note also that the geodetic location will not change if a scene to geodetic transformation is not available.

Delete Point Deletes the active ground control point.

3.2.8 Match Matrix View

The match matrix view provides a visualization of the feature point associations across frames. Pixels in the image correspond to values in the "match matrix" representing the number of feature points that feature detection has determined correspond to the same real world feature. Several options are provided to adjust the visualization:

- Layout controls the position of "identity" values, i.e. values that compare a frame to itself rather than a distinct frame. The default, "diagonal", simply maps the frame number directly to both the X and Y axes. "Horizontal" skews the image so that the y values are relative to the "identity" values, placing them in a horizontal line at y = 0, with positive y representing "later" frames, and negative y representing "earlier" frames. "Vertical" reverses these axes.

- Orientation controls which screen direction is considered positive Y. The default, "matrix", uses down for positive Y, as in textual value tables (e.g. textual listings of matrices, spreadsheets) or images. "Graph" uses up for positive Y, as in most graphical plots.

- Values controls what values are used for each pixel. The default, "absolute", uses the raw number of feature point correlations (which, for "identity" values is equal to the total number of feature points on that frame). "Relative (combined)" mode uses the percent of common feature points relative to the total number of distinct feature points on each frame being compared. The other two "relative" modes give the percent relative to the total number of feature points for the frame represented by either the X or Y axis.

- Scale controls the scaling function that is applied to the values produced according to the value mode. The choices are "linear", "logarithmic" and "exponential", and should be self explanatory. In absolute value mode, logarithmic scale uses the maximum value as the logarithm base. Otherwise, the base can be adjusted with the "range" control, which applies a pre-scale to the value before computing the logarithm (thereby allowing the shape of the scaling curve to be adjusted). Exponential scale allows the user to select the exponent.

- Color provides the set of colors to which scaled values are mapped. Several presets are available according to user taste. Different presets may help emphasize different aspects of the data.

Moving the mouse over the image will display which frames are being compared and the number or percentage of feature correlations in the status bar. The match matrix view also allows the image to be exported to a file.

3.2.9 Data Files

TeleSculptor supports visualization of various data files (landmarks, cameras, etc.) that are computed in other tools. However the recommended workflow for most users is to simply load a video and derive all other product from it. Video files are loaded using *File → Import → Imagery. . . .*

Before computing any products from video, a "Project" directory is needed to store the results. A project is created with *File → New Project* which asks the user to provide a path to a working directory. Inside this directory, a "Project File" is created (name matching the directory name plus extension .conf) to store project settings. Various other result files are also written to the project directory. To open an existing project, use *File → Open Project. . .* and navigate to an existing .conf file.

Note: When loading cameras or images individually, cameras and images are associated in a first-loaded, first-matched manner. There is no way to load individual camera and image files that allows for cameras without images,

or images without cameras, except at the end of the frame sequence. Similarly, frame identifiers are assigned sequentially based on the order in which files are loaded. In order for feature points to be correctly associated with their corresponding frames, the camera/image files must be loaded so that these automatically assigned identifies match those that were assigned by the feature detection/tracking pipeline.

3.2.10 Menu

File Menu

New Project Select a working directory for a project. A project directory must be set before the tools in the Compute menu can be run. These tool will write files into the project working directory. A configuration file with the same name as the directory is also created in the directory. The project configuration file stores references to the project data such as the source video and computed results like cameras, tracks, or landmarks that will be loaded back in when a project is opened.

Open Project Select an existing project configuration. The project configuration will often include references to various data files which are frequently stored in the same directory as the project configuration.

Import Provides options for importing/loading various types of data into the current project. The user must select the type of data to be loaded, as some data files use the same file extension.

Export Provides options for exporting various data.

Quit Exits the application.

Compute Menu

Run End-to-End Runs the entire processing pipeline from end-to-end. This tools runs Track Features, then Estimate Cameras/Landmarks, then Batch Compute Depth Maps, then Fuse Depth Maps.

Track Features Run feature tracking on the loaded video starting from the current frame. Features and descriptors are detected in each frame and cached into a file in the project directory. Features are then matched between adjacent frames as well as between the current frame as past keyframes. These feature matches form "tracks" through time, and each track has the potential to become a landmark.

Estimate Cameras/Landmarks Estimates cameras and landmarks starting with tracks and metadata. This also runs bundle adjustment (refinement) along the way. The goal is to incrementally add cameras and landmarks, while optimizing, to build up a consistent solution.

Save Frames Iterate through a video and save every frame as an image file in a subdirectory of the project directory. This is needed when exporting the data to other tools that do not support video files. This option must be run before importing a project into SketchUp.

Batch Compute Depth Maps Estimates several dense depth maps and corresponding point clouds on several frames spaced throughout the video. This requires valid cameras and computes the results in the active ROI. The algorithm run on each frame is the same as Compute → Advanced → Compute Single Depth Map, but intermediate solutions of each depth map are not rendered.

Fuse Depth Maps Fuse all computed depth maps into a single mesh surface using an integration volume specified by the ROI. Note that this step requires an NVIDIA GPU and may not be able to run if the ROI is too large for the GPU memory.

Cancel Cancel the current running tool. Most, but not, all tools support this cancel action, which causes them to exit early and return a partial result.

Advanced Provides access to additional, lower level tools. See *Advanced Tools* for details.

Options Provides options to alter how the tools operate. See *Compute Options* for details.

View Menu

Play Slideshow Toggles playback of the slideshow.

Loop Slideshow Toggles if the slideshow should restart from the beginning after the last frame. When disabled, the slideshow ends when the last frame becomes active.

Match Matrix Opens a new *Match Matrix View*.

Background Color Changes the background color of the world and camera views.

World Axes Toggles visibility of *X*, *Y*, and *Z* axes in the world view showing numerical values for distances at regular intervals on these axes. The size of these axes is set to span all visible scene objects, including the camera path. If the axes are too large, hiding scene objects like cameras and the ground plane will shrink the coverage to the remaining visible data.

Keyframes Only Limit frame numbers in the camera selection pane to allow only frames that were designated as "keyframes" by the feature tracker. The number of keyframes is typically very small.

Tracked Frames Only Limit frame numbers in the camera selection pane to allow only frames that contain feature tracking results. The feature tracker only processes a fixed number of frames (default: 500) distributed through the video. Enabling this option skips unprocessed frames during playback, which avoids flicker of the display that occurs when unprocessed frames are drawn.

Antialias Views Toggles use of an anti-aliasing filter in the world, camera and depth views. Anti-aliasing is accomplished via a post-processing filter (FXAA) that may produce undesirable artifacts. At this time, anti-aliasing via multi-sampling (MSAA) is not supported.

TeleSculptor User Manual Displays this documentation in the default web browser.

About TeleSculptor Shows copyright and version information about the application.

3.3 User Workflows

This section describes the workflows a user should follow when using the TeleSculptor software. There are several ways to use the software depending on whether the end goal is to produce camera models for import into SketchUp, to create a fully automated 3D models from FMV, or simply to inspect the accuracy of the KLV metadata provided with an FMV clip. The available workflows are described below. The details of each processing step are described in the following Processing Steps section.

3.3.1 Metadata Inspection

This is the simplest capability provided by the TeleSculptor application. When loading a video, the application will read all the KLV metadata (assuming MISB 0601 or 0104 standards) and convert metadata pertaining to the camera location and orientation into 3D visualization of these cameras. This is useful for checking if a video has camera pose metadata and validating that the cameras are pointing at the correct location. This visual check can help find gross errors in metadata that are difficult to interpret by just looking at the raw metadata values.

A metadata panel on the side of the user interface shows all KLV fields associated with the current frame. Values in this panel will update as the video is played or scrubbed. If no metadata is available for a given frame the values are all set to "(not available)". In the example shown below we see a video with MISB 0104 metadata. The platform path is well represented, but this video has only platform orientation and is missing sensor orientation, so the metadata cameras are all rendered pointing up.

To inspect the metadata, follow the instructions for opening a video. There is no need to create a project file if no further processing is anticipated. When a video is opened it is automatically scanned for all metadata to build the 3D representation of the cameras. Scanning a video for metadata may take several seconds or even minutes for very large videos. While TeleSculptor is scanning the video one can still playback the video or seek with the frame selection slider to preview the imagery and metadata values.

3.3.2 SketchUp Enhancement

The goal of this workflow is to estimate accurate camera models for keyframes in the FMV and then import those keyframes and camera models into SketchUp to use the Match Photo interface to build 3D models in SketchUp by drawing directly on the images. The SketchUp workflow is described in a separate document. This document describes how to use TeleSculptor to produce the data that will be loaded into SketchUp. The processing steps are as follows

1. Create a New Project

2. Import a Video

3. Track Features

4. Estimate Cameras/Landmarks

5. Save Frames
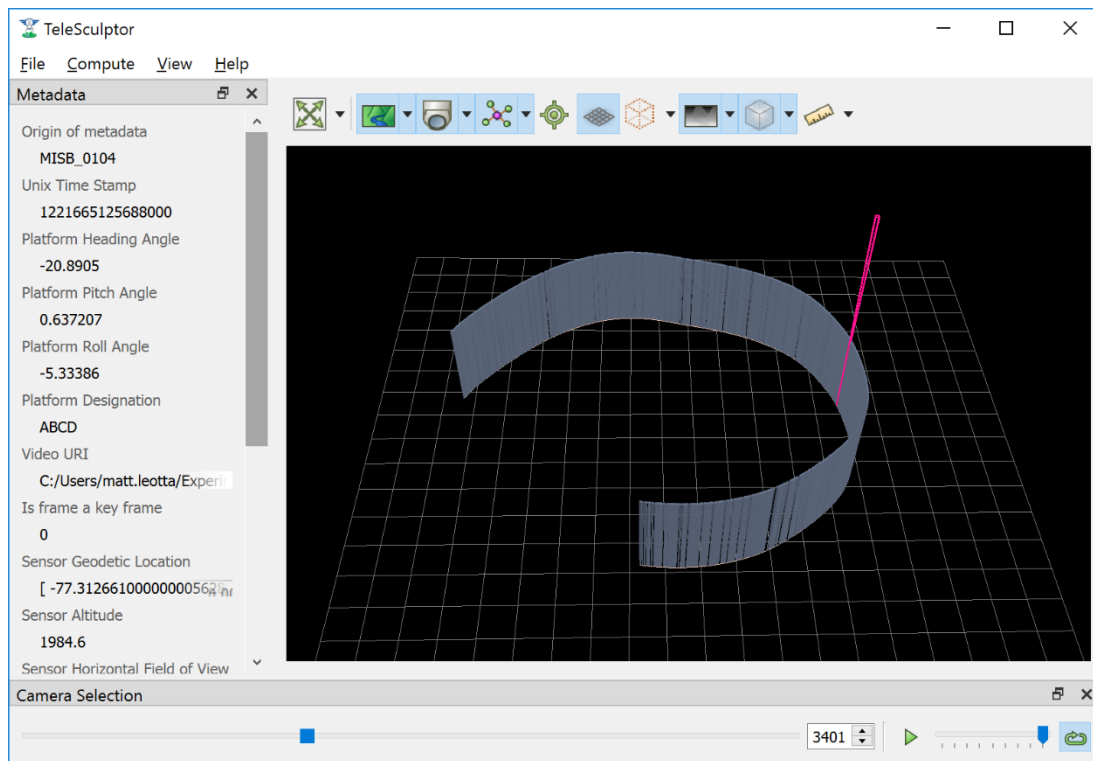
6. Set Ground Control Points

Fig. 11: Inspecting Metadata geometrically and with the metadata viewer pane.

3.3.3 Dense Automated 3D Models

The goal of this workflow is to automatically estimate a dense 3D model of the scene without using SketchUp. The processing pipeline builds on the estimated cameras used in the SketchUp workflow. It estimates dense depth maps from multiple different viewing directions and then fuses the depth maps together into a unified 3D surface mesh. The processing step are as follows

1. Create a New Project

2. Import a Video

3. Track Features

4. Estimate Cameras/Landmarks

5. Batch Compute Depth Maps

6. Fuse Depth Maps

7. Colorize Surface

8. Export Fused Mesh

If running both the SketchUp Enhancement and Dense Automated 3D Models workflows on the same video, there is no need to repeat the common steps (1-4). Simply run steps 5-8 on the existing camera and landmark estimation. Optionally, selecting a smaller 3D region of interest (ROI) before step 6 or 7 will limit the region of computation and compute results more quickly. By selecting "Run End-to-End" in the compute menu, TeleSculptor will automatically run steps 3-6 one after the other.

3.4 Processing Steps

This section describes how to run the key processing steps and what each step does. The previous section describes which of these steps you should run, and in which order, depending on the desired goals. However, processing step are generally run in the order listed below, with some steps only needed for one workflow or another.

3.4.1 Create a New Project

The TeleSculptor application requires a working directory, also called the project folder, in which to save settings and algorithm results when processing a video. To create a new project use the *File → New Project* menu item or keyboard shortcut Ctrl+N. Create a new empty folder at the desired location and press the "Select Folder" button with that new folder highlighted.
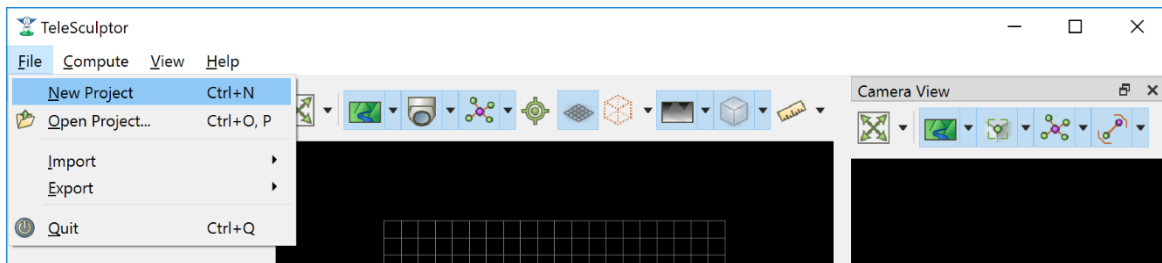

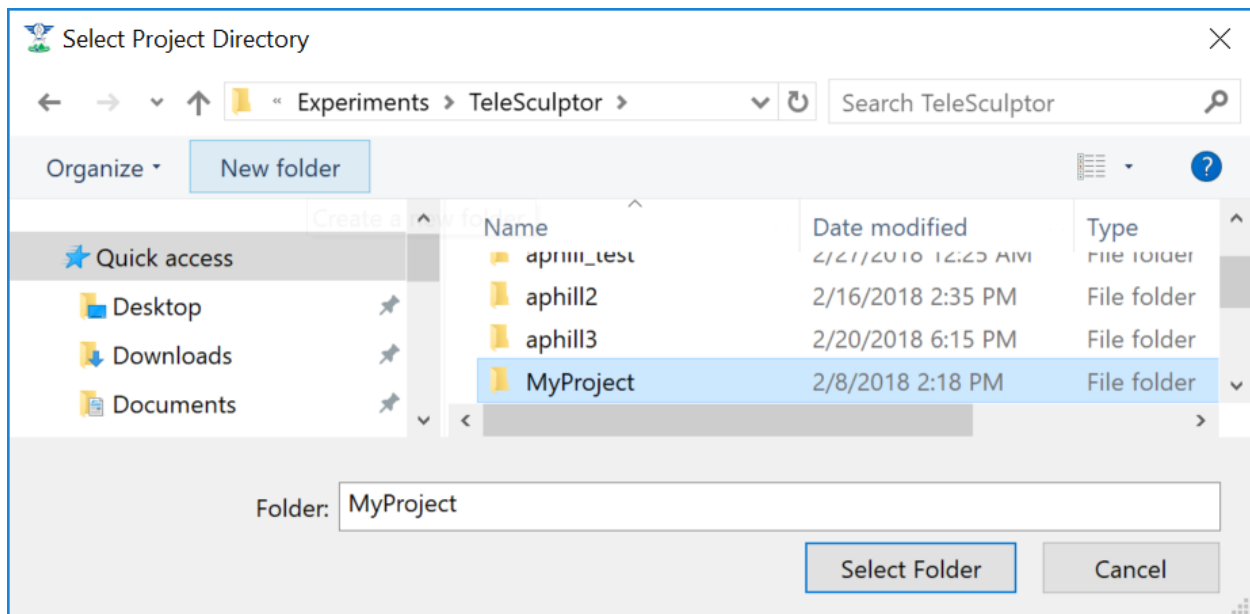Fig. 12: Create a new project from the File->New Project menu item.

Fig. 13: Create a new project folder to store the results of processing.

After creating the new project, the application will create a project file in the project directory with the same name and a ".conf" file extension. In the example shown in the figure above it will create **MyProject/MyProject.conf**. This conf file is known as the project file or the project configuration file. It contains the configuration settings for the project. To reload an existing project after closing the application, use *File->Open* and select this project configuration file.

The user must create or open a project before running any of the tools in the Compute menu. The application can open videos and other files for inspection without a project but cannot process the data without a project. The recommended

workflow is to create a project first and then open a video, however, if the project is created after opening a video the open video will be added to the new project.

Projects can have any name and refer to a video file at any location on the computer. That said, for better organization we recommend giving the project folder the same name video clip that will be processed. If it is possible to move the video clip into the new project folder *before* opening the video the project will be relocatable. That is, if the source video is in the project folder you can move the project folder to another location or another computer and it will still load. If the video is outside the project folder the absolute path will be recorded, but this link could become broken if the video or project is moved later.

3.4.2 Import a Video

To import a video clip, use the *File->Import->Imagery* menu item. In the Open File dialog box browse to select the file to open and then press the Open button. This same menu item can be used to open a list of image paths in a text file. Advanced users can also open intermediate data files, like cameras and landmarks, from the import menu, but these use cases are not covered in this document. Masks are another special type of import. These are also a video or image list but contain black and white images with black indicating which pixels of the image to ignore. Mask images are particularly useful for videos with burned-in metadata as a heads up display (HUD). Kitware has a separate tool call Burn-Out for estimating these mask videos for videos with burned in metadata.
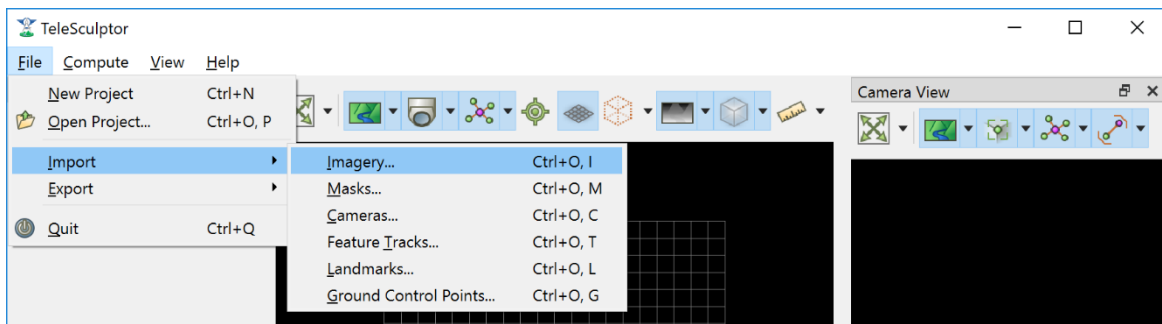

Fig. 14: Import a video clip, image list, or other data.

Once a video is selected to open, the application will scan the entire video to find all metadata. This may take several seconds or even minutes for very large videos. If scanning the video takes too long or it is known that the video does not contain metadata it is okay to cancel the metadata scanning using the "Cancel" option in the Compute menu. If metadata scanning is canceled, then no metadata will be used in any subsequent processing steps.

The first video frame will appear in the Camera View in the upper right. The World View (left) will show a 3D representation of the camera path and camera viewing frustums if relevant metadata is found. When images are shown in the 3D world view, they are always projected using the active camera model onto this ground plane (indicated by the grid). The image button () toggles visibility of images in each view with an opacity slider in the drop-down menu below the button. For images to appear in the world view an active camera model is required. An active camera model is a camera model for the currently selected video frame.

As can be seen in the image above. The aircraft flight path, as given in the metadata, is shown as a curved line in the 3D World View. A pyramid (frustum) is shown representing the orientation and position of the camera at each time step. The active camera, representing the current frame of video, is shown in a different color and is longer than the others. If the video is played with the play button in the lower right, the video frames will play in the Camera View and the active camera will updated in the World View.

The visibility of cameras is controlled by the cameras button () above the world view. In the drop-down menu underneath the cameras button you can individually set the visibility, size, and color of camera path, camera frustums, and active camera.
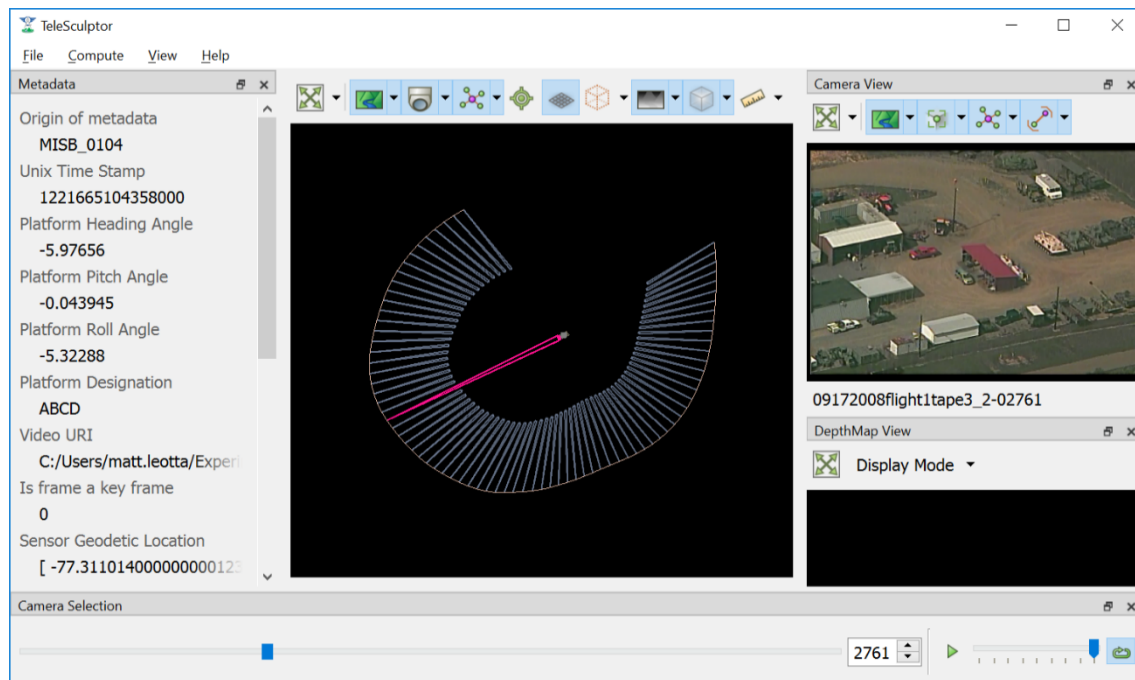
Fig. 15: A video with KLV metadata opened in the application.

3.4.3 Run End-to-End

Run End-to-End is a new feature in TeleSculptor v1.2.0. Rather than waiting for user input to run each of the primary processing steps, Run End-to-End automatically runs Track Features, Estimate Cameras/Landmarks, Batch Compute Depth Maps, and Fuse Depth Maps. These steps are run sequentially in this order. See details of these steps in their respective sections below.
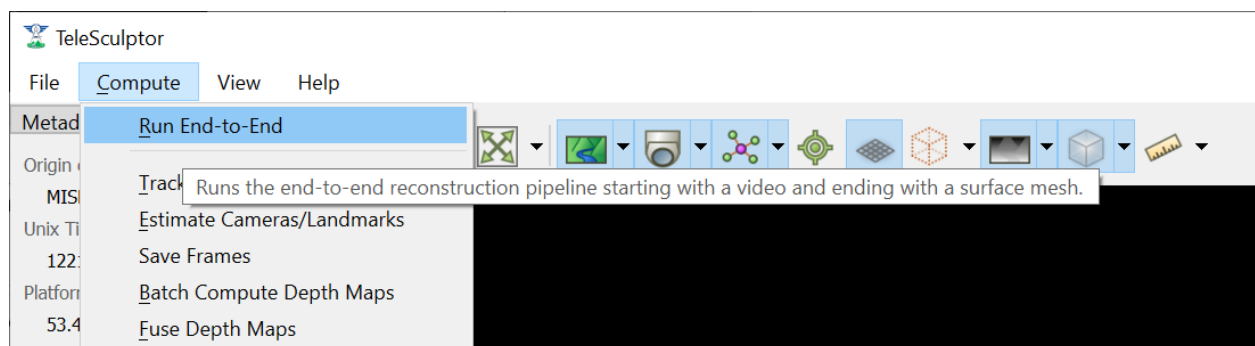

Fig. 16: Run End-to-End Option in the Compute menu.

3.4.4 Track Features

The Track Features tool detects visually distinct points in the image called "features" and tracks the motion of those feature points through the video. This tool is run from the *Compute* menu and is enabled after creating a project and loading a video.

When the tool is running it will draw feature points on the current frame and slowly play through the video tracking the motion of those points as it goes. These tracks are visualized as red trails in the image below. These colors are
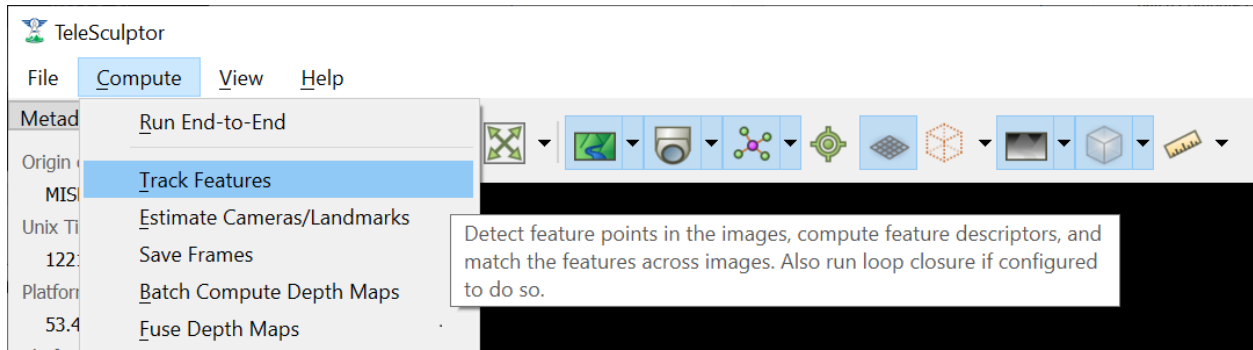
Fig. 17: Run the Track Features algorithm from the Compute menu.

fully customizable. The feature tracks button (🔲) above the Camera View enables toggling the visibility of tracks. The drop-down menu under this button has settings for the display color and size of the feature track points and their motion trails.

The feature tracking tool will start processing on the active frame and will run until the video is complete or until the tool is canceled with the *Cancel* option in the *Compute* menu. We recommend starting with the video on frame 1 and letting the algorithm process until complete for a video clip containing approximately one orbit of the UAV above the scene. However, it is possible to process subsets of the video by scrubbing to a desired start frame before running the tool and then hitting the cancel button after reaching the desired end frame.

When this, or any other, tool is running all other tools will be disabled in the *Compute* menu until the tool completes. Most tools also support exiting early with the *Cancel* button to stop at a partial or suboptimal solution.

Note that to limit redundant computation this tool does not track features on every frame of video for long videos. Instead the algorithm sets a maximum of 500 frames (configurable in the configuration files) and if the video contains more than 500 frames it selects 500 frames evenly distributed throughout the video. Once feature tracking is done, tracks will flicker in and out when playing back the video due to frames with no tracking data. To prevent this flickering select *Tracked Frames Only* from the View menu. With this option enabled, playback is limited to frames which include tracking data.

More technical users who want to understand the quality of feature tracking results may wish to use the *Match Matrix* viewer under the *View* menu (keyboard shortcut M). The match matrix is a symmetric square matrix such that the value in the *ith* row and *jth* column is the number of features in correspondence between frames i and j. The magnitude of these value is colored with a color map and placing the mouse cursor over a pixel prints the actual number of matches in the status bar. Typically, a match matrix has strong response down the diagonal (nearby frames) that drops off as you move away from the diagonal. Flight paths that make a complete orbit should see a second weaker off-diagonal band where the camera returns to see the same view again.

3.4.5 Estimate Cameras/Landmarks

The Estimate Cameras/Landmarks tool in the Compute menu uses structure-from-motion algorithms to estimate the initial pose of cameras and the initial placement of 3D landmarks. It also uses bundle adjustment to jointly optimize both cameras and landmarks. These algorithms use the camera metadata as constraints and initial conditions when available. The algorithm will try to estimate a camera for every frame that was tracked and a landmark for every feature track.

The solution will start with a sparse set of cameras and then incrementally add more. Live updates will show progress in the world view display. During optimization, the landmarks will appear to float above (or below) the ground plane grid because the true elevation is typically not near zero. Once the optimization is complete, a local ground height is estimated, and the ground plane grid is moved to meet the landmarks. This ground elevation offset is recorded as part of the geo-registration local coordinate system.
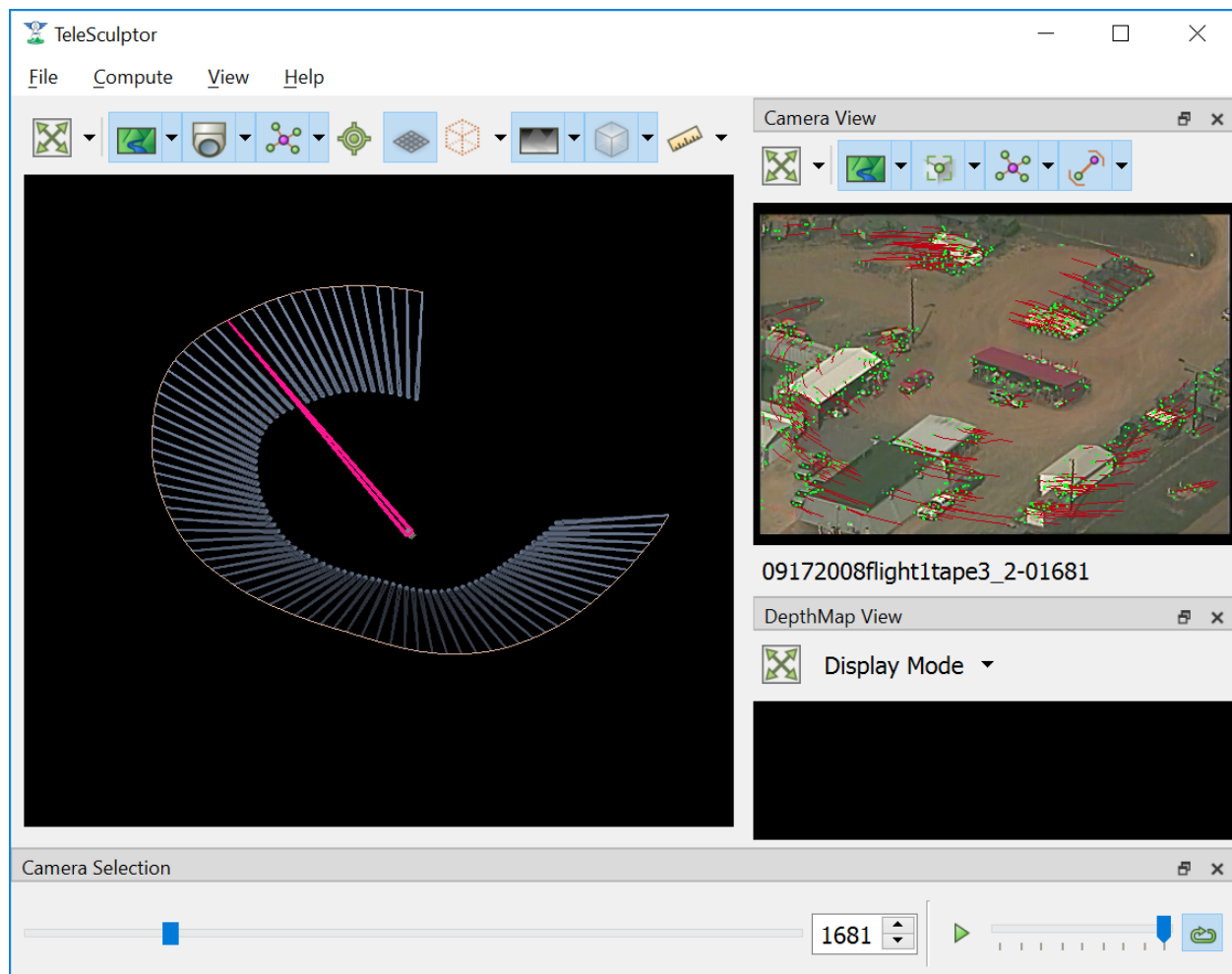
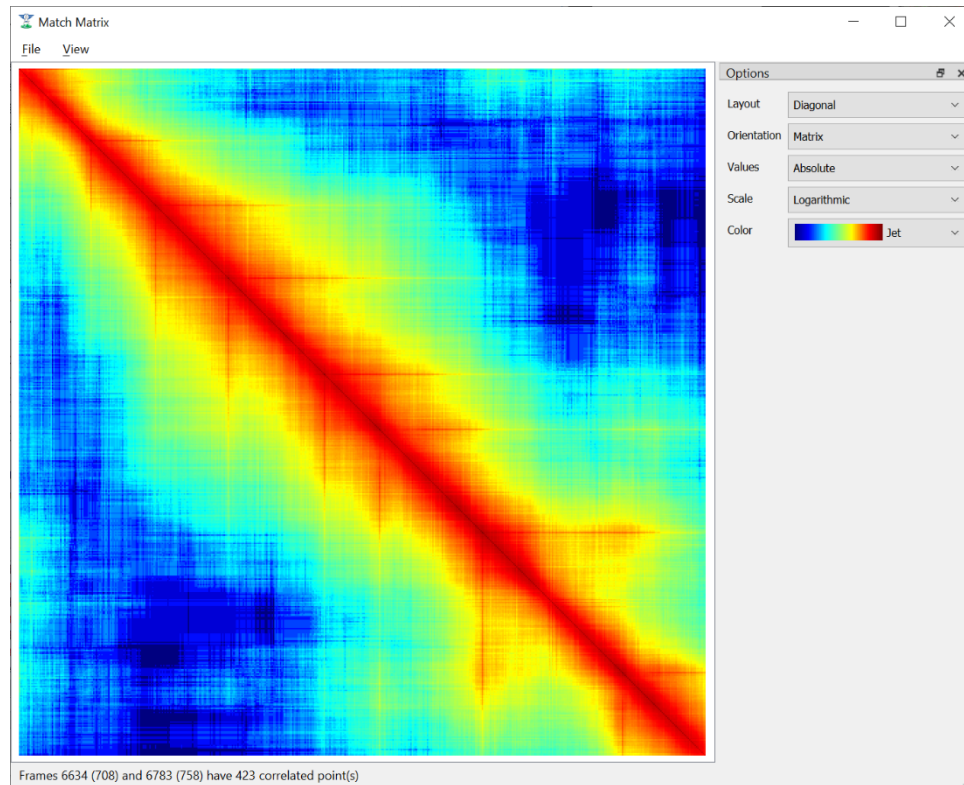Fig. 18: The Track Features algorithm producing tracks on a video.

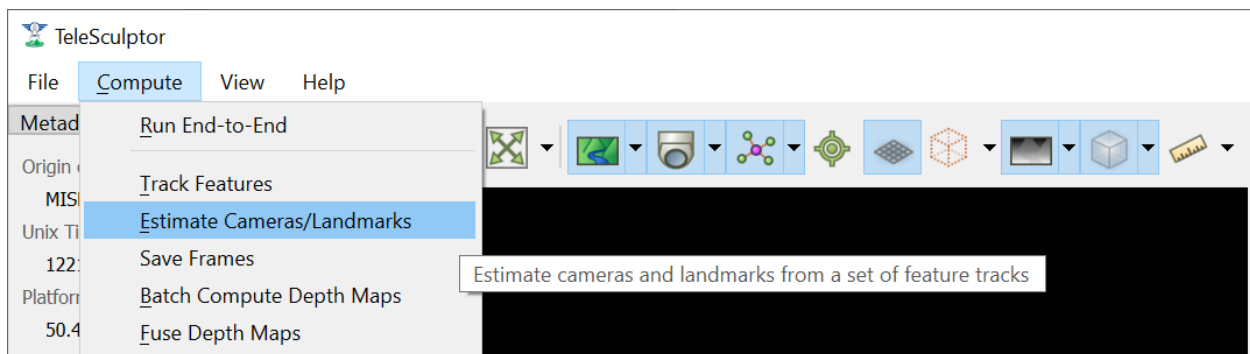Fig. 19: The Match Matrix view of feature tracking results.

Fig. 20: Run the Estimate Cameras/Landmarks algorithm from the Compute menu.

Once landmarks are computed their visibility can be toggled with the landmarks button ( ) in both the world and camera views. The drop-down menu under the landmarks button allows changing the size and color of the landmarks including color by height or by number of observations.

3.4.6 Save Frames

The *Save Frames* tool is quite simple. It simply steps through the video and writes each frame of video to disk as an image file. These image files are stored in a subdirectory of the project directory. Saving frames only requires an open video and an active project. It can be run at any time. Like the feature tracking tool, it plays through the video as it processes the data and can be cancelled to stop early. The primary purpose for saving frames is for using them in the TeleSculptor / SketchUp workflow. SketchUp can only load image images, not video. So, this step produces the image files that are needed when loading a project file into SketchUp.
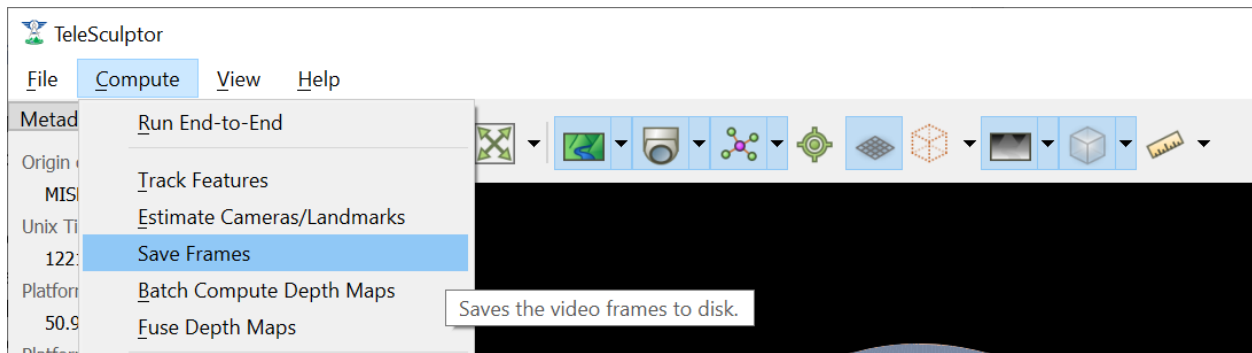

Fig. 21: Save the loaded video frame as image files on disk.

3.4.7 Set Ground Control Points

Ground Control Points (GCPs) are user specified 3D markers that may be manually added to the scene for a variety of purposes. These markers are completely optional features that may be used to provide meaningful guide points for modeling when exporting to SketchUp. GCPs are also used to estimate geo-localization of videos with metadata or to improve geo-location for video with metadata.

To add GCPs press the GCP button ( ) above the 3D World View. To create a new GCP hold the *Ctrl* key on the keyboard and left click in either the 3D World View or the 2D Camera View. A new GCP will appear as a green cross in both views. Initial points are currently dropped into the scene along the view ray under the mouse cursor at the depth of the closest scene structure. If the initial depth of a GCP is not accurate enough it can be moved. To move a point, left click on the point in either view and drag it. Points will always move in a plane parallel to the image plane (or current view plane). It helps to rotate the 3D viewpoint or scrub the video to different camera viewpoints to correct the position along different axes. Holding *Shift* while clicking and dragging limits motion to single coordinate axis. The axis of motion is the direction which has the most motion in the initial mouse movement. Once additional points are added (with *Ctrl* + left clicks) the active point is always shown in green while the rest are shown in white. Left clicking on any point makes it active. Hitting the *Delete* key will delete the current active GCP.

The Ground Control Points pane provides a way to select and manage the added GCPs. The pane lists all added points and allows the user to optionally assign a name to each. The GCP pane also shows the geodetic coordinates of the active point, and these points can be copied to the clipboard in different formats using the copy location button ( ). If the value of the geodetic coordinates is changed that GCP becomes a constraint and is marked with an icon ( ) in the GCP list. Constrained points will keep fixed geodetic coordinates when the GCP is moved in the world space. A constraint can be removed by pressing the reset button ( ). Once at least three GCP constraints are added
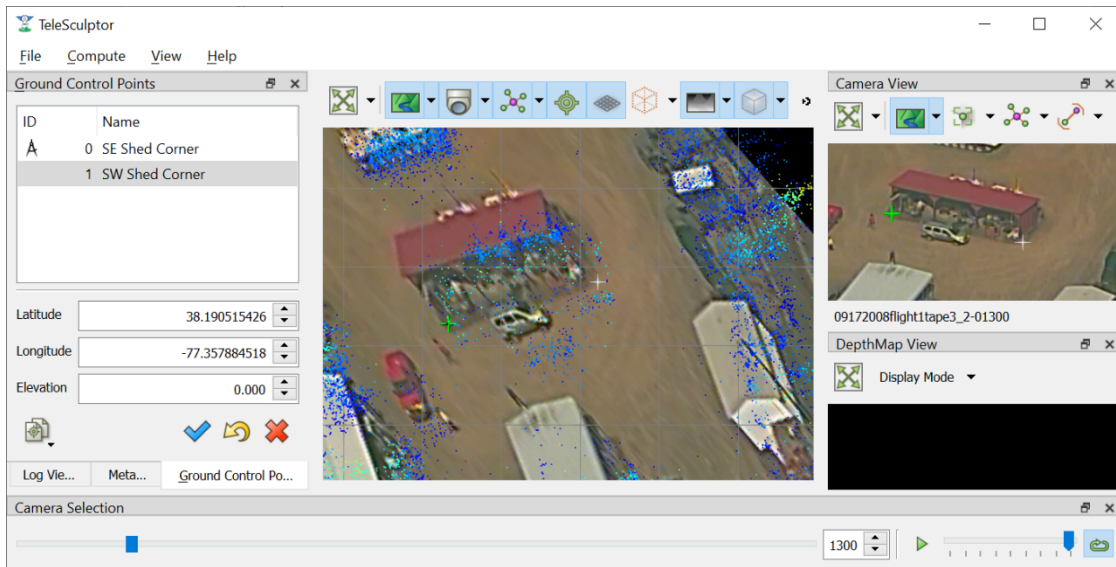
Fig. 22: Setting ground control points. The active point is shown in green.

with geodetic coordinates, the apply button (✔) can be used to estimate and apply a transformation to geolocalize the data. While three GCPs are the minimum, five or more are recommended. The transformation will be fit to all GCP constraints. After applying the GCPs, all cameras, landmarks, depth maps, and GCPs are transformed to the new geographic coordinates. Currently the mesh fusion is not transformed because the integration volume is axis-aligned. Instead the fusion results are cleared and need to be recomputed.

It is often helpful to compute a depth map (or even fused 3D model) before setting GCPs to provide additional spatial reference in the 3D view. It is possible set GCPs entirely from the 2D camera views by switching between video frames and correcting the position in each. However, this is more tedious. When the 3D position is correct the GCP should stick to the same object location as the video is played back.

GCPs are currently not saved automatically. To save the GCP state use the *File->Export->Ground Control Points* menu option and create a PLY file to write. This file path is cached in the project configuration, so GCPs are automatically loaded when the project is opened again. They are also automatically loaded when importing the project configuration into SketchUp.
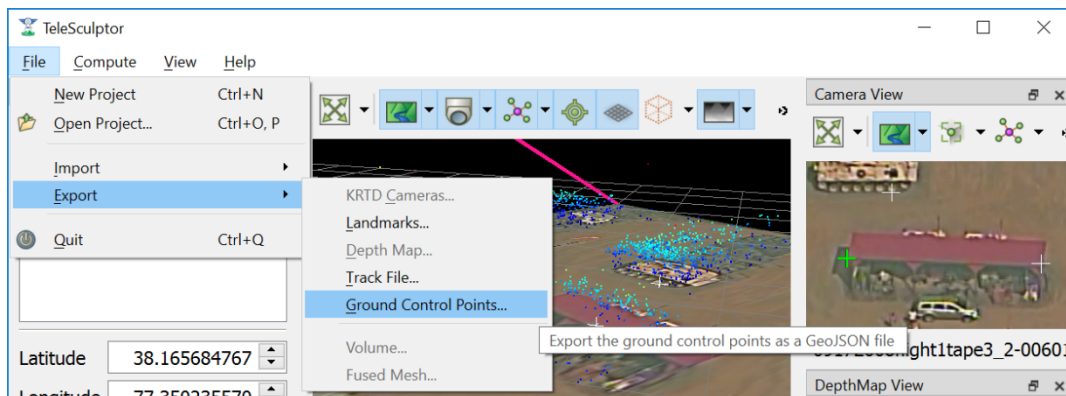

Fig. 23: Export the ground control points.

3.4.8 Set 3D Region of Interest

Before running dense 3D modeling operations, it is beneficial to set a 3D region of interest (ROI) around the portion of the scene that is of interest. This step is optional. By default, a ROI is chosen to enclose most of the 3D landmark point that were computed in the triangulate landmarks step. Some outlier points are rejected when fitting the ROI and the estimated ROI is padded to account for missing data. The default ROI is generally sufficient for further processing, but may be larger than necessary. The advantage of picking a smaller ROI is a significant reduction in compute time and resources. Furthermore, the quality and resolution of the result often improves when focusing on smaller subset of a large scene because we can focus more compute resources on that location.

To see and manipulate the 3D ROI click the 3D ROI button (⬚) above the 3D World View. A 3D axis-aligned box is shown which contains the set of 3D landmarks. Inside the box are axis lines along the center of the box in each of the three coordinate directions. At the ends of these lines are spheres which act as manipulation handles. Left click on any of these handles and drag to reposition the corresponding face of the box in 3D. Left click the sphere at the center of the box and drag to translate the entire box. A middle click (or *Ctrl* + left click or *Shift* + left click) and drag anywhere in the box has the same translation effect as using the center handle. A right click and drag will scale the box uniformly about its origin. Note that the ground plane grid will adjust size relative to the ROI size.
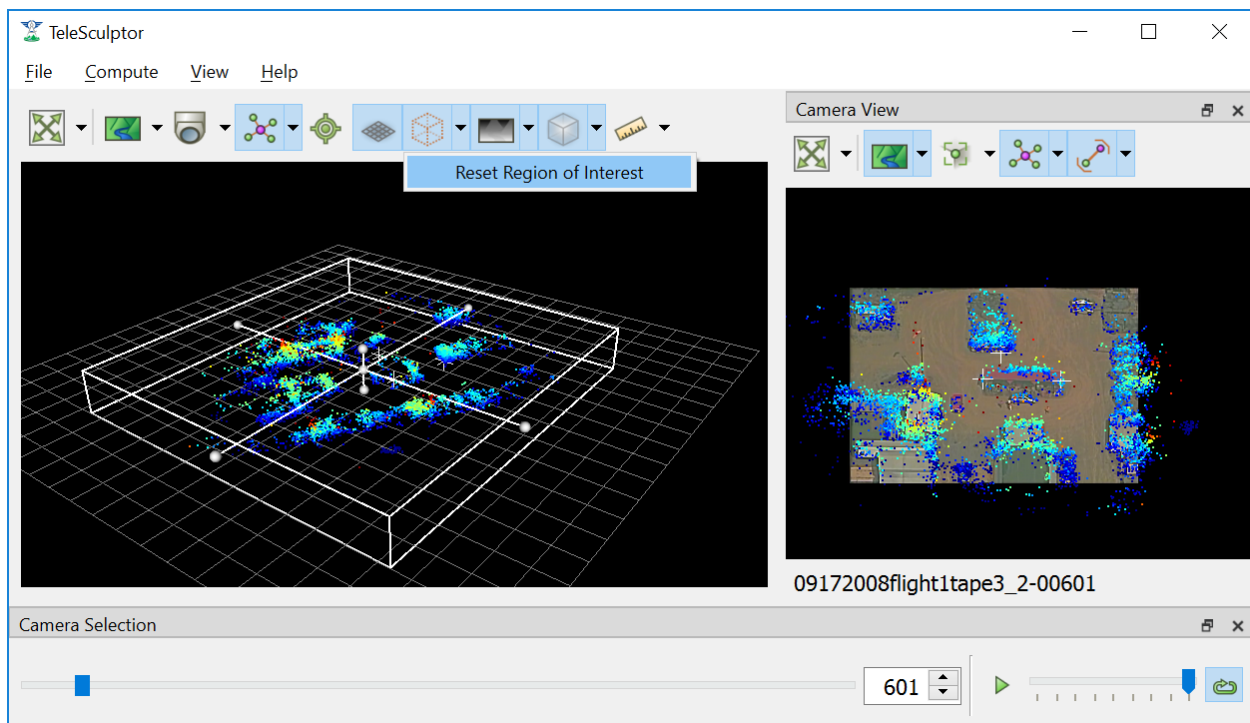

Fig. 24: The default 3D bounding box is fit to the sparse landmarks and is often bigger than needed.

A good practice is to set the bottom of the ROI box just below the ground and the top just above the tallest part of the structure. Likewise, set the sides to be just a bit outside the object of interest. It may be difficult to determine the bounds accurately from the sparse landmarks. A good strategy is to start with a slightly larger guess, then use the *Compute Single Depth Map* advanced tool to compute a single depth map. The depth map gives more detail which helps pick a tighter box. Then use *Batch Compute Depth Maps* to compute the additional depth maps with the revised box.

To reset the ROI to the initial estimated bounding box, use the *Reset Region of Interest* option in the drop-down menu under the ROI button.
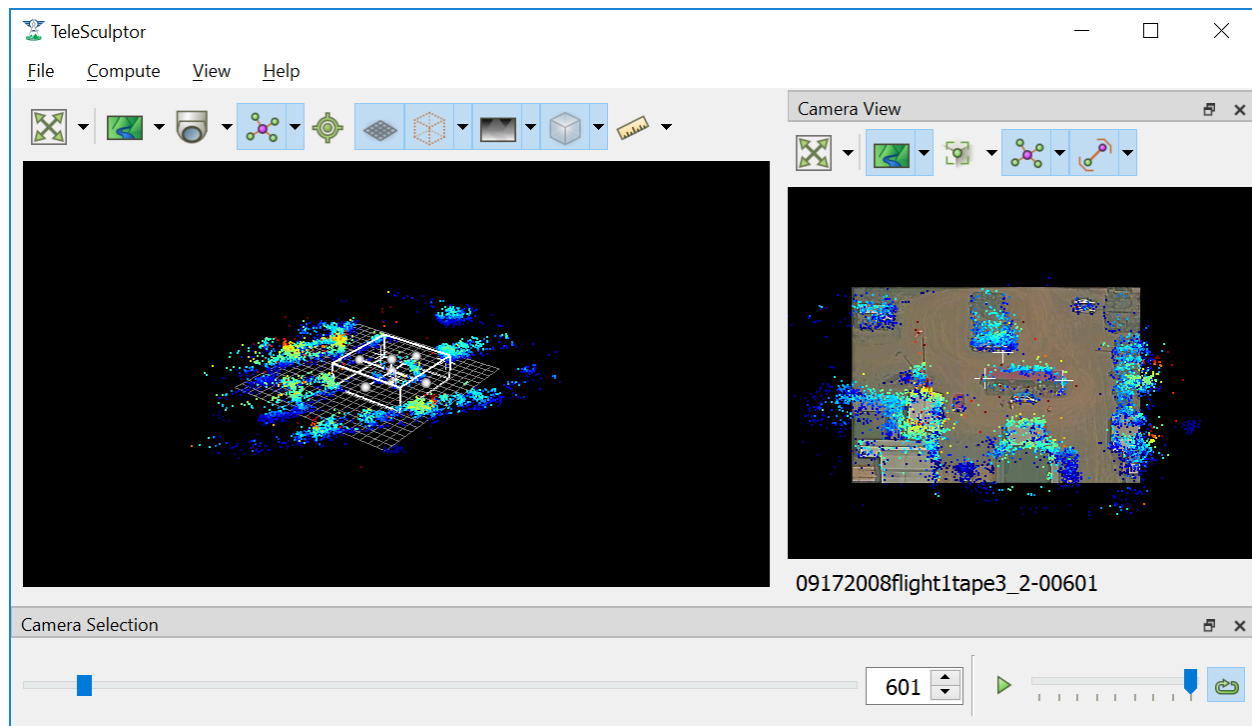
Fig. 25: Setting a tighter bounding box around one structure of interest.

3.4.9 Batch Compute Depth Maps

Computation of dense depth maps is part of the fully automated 3D reconstruction pipeline. Several depth maps are needed to compute a final 3D result. Running the *Batch Compute Depth Map* tool from the *Compute* menu will estimate depth maps (2.5D models) for twenty different frames sampled evenly through the video. To compute a depth map only on the current frame, see the *Compute Single Depth Map* option in the advanced menu.

This algorithm requires very accurate camera and landmark data resulting from the previous *Estimate Cameras/Landmarks* step above. Furthermore, cameras models on multiple frames in nearby positions are required. By default, the algorithm uses the ten frames before and ten frames after each selected depth frame for reference.
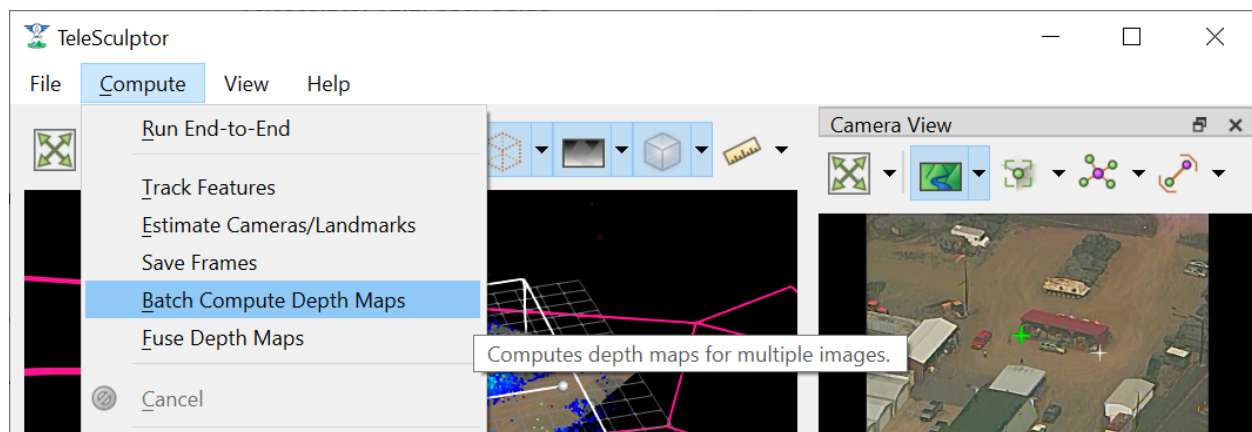

Fig. 26: Compute dense depth maps on key frames.

The results of depth map estimation are shown in two ways. In the world view the depth maps are shown as a dense

colored point cloud in which every image pixel is back projected into 3D at the estimated depth. Use the Depth Map button ( ) to toggle depth point cloud visibility. The second way depth maps are visualized is as a depth image in the Depth Map View. Here each pixel is color coded by depth and the color mapping is configurable.
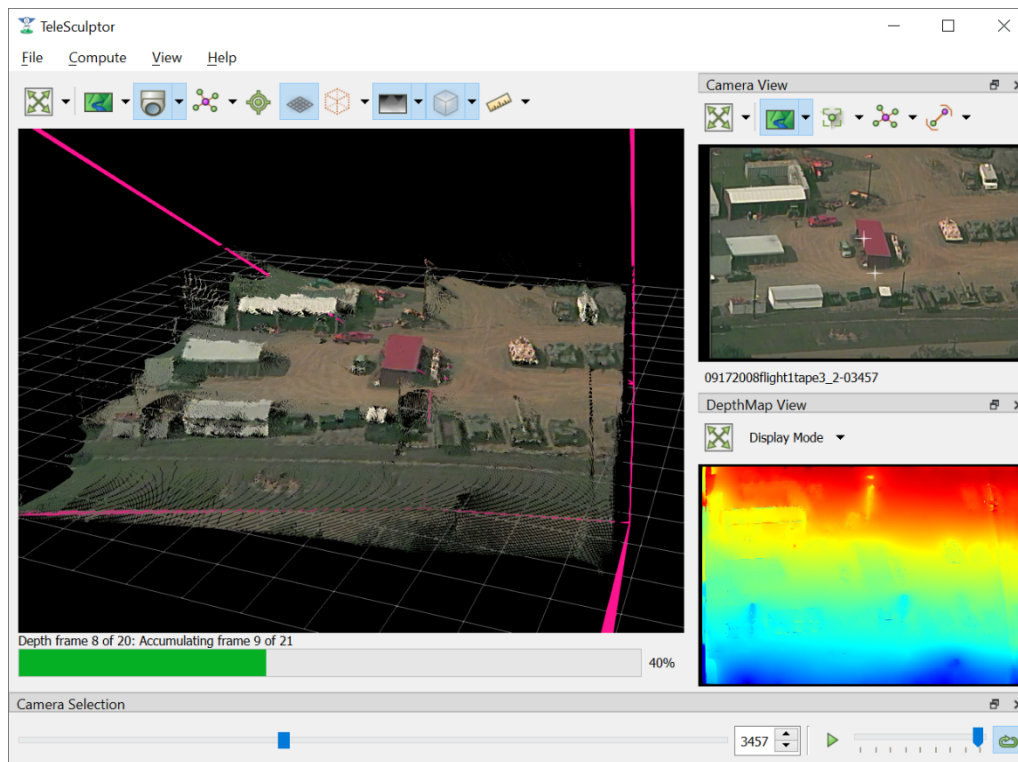

Fig. 27: Results of depth map computation.

3.4.10 Fuse Depth Maps

After computing depth maps in batch, or manually computing multiple depth maps, the next step is to fuse them into a consistent 3D surface. Running *Fuse Depth Maps* from the *Compute* menu will build an integration volume and project all depth maps into it for fusion. This integration step requires a modern CUDA capable Nvidia GPU (Requires at least Nvidia driver version 396.26). The size of the integration volume and the ROI covered is determined by the same ROI box used in the depth map computation. This processing step runs in only a few seconds and may cause lag in the display during this time due to consumption of GPU resources. Once the data is fused into a volume, a mesh surface is extracted from the volume.

To toggle the view of the fused mesh, press the Volume Display button ( ) above the 3D World View. The surface mesh can be fine-tuned if desired by adjusting the surface threshold in the drop-down menu under the Volume Display button. Setting the threshold slightly positive (e.g. 0.5) often helps to remove unwanted outlier surfaces that tend to appear in areas with only a few views.

3.4.11 Colorize Mesh

The fused mesh is provided initially in a solid grey color. To add color, use the drop down menu under the Volume Display button ( ). Use the left drop down menu to select between *No Color* and *Image Color*. If, and only if, a mesh was loaded from a file, a third option for *Original Color* is available. *Original Color* displays the color loaded from the mesh file.
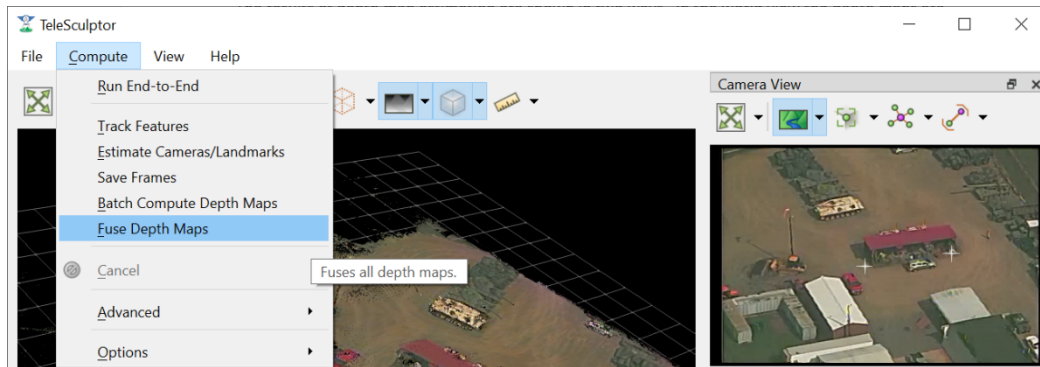
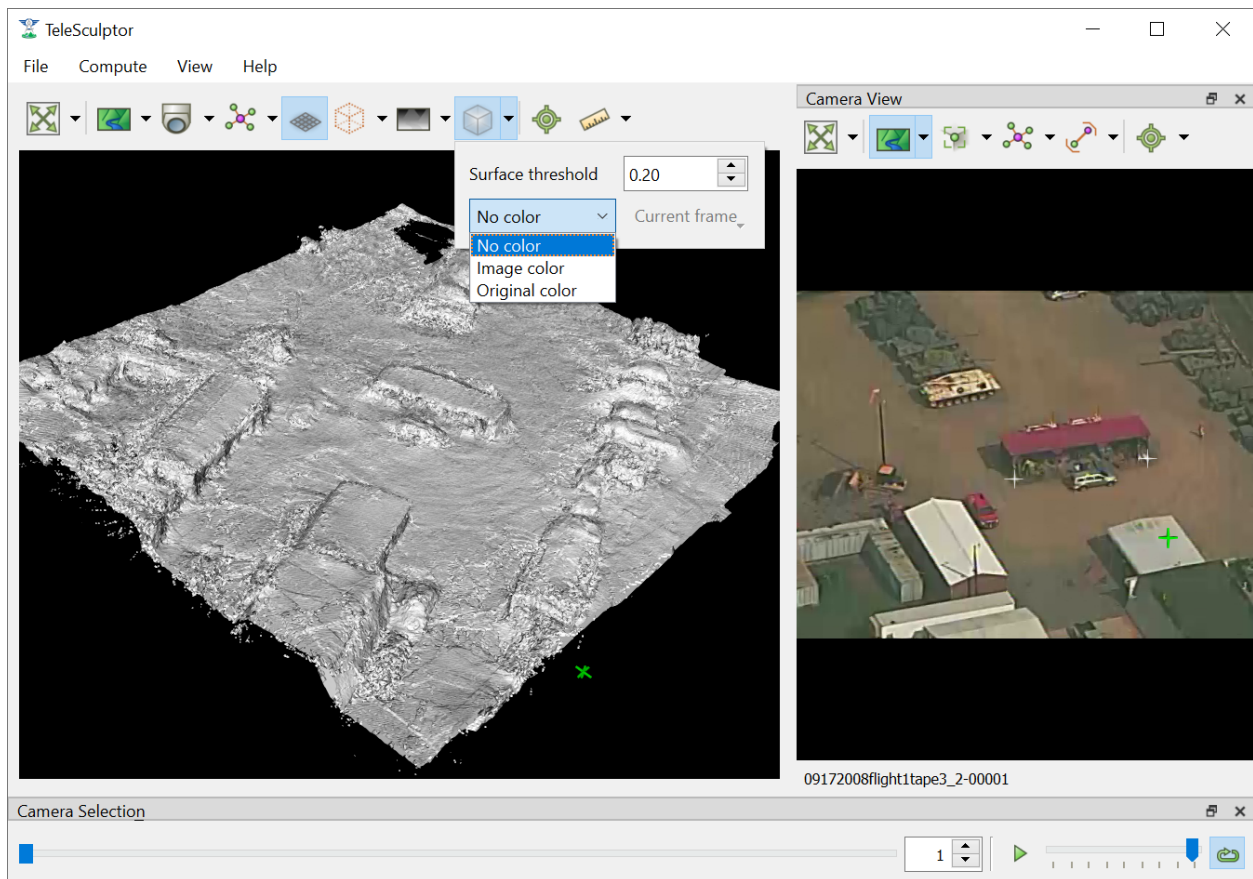Fig. 28: Fuse the depth maps into a consistent mesh.

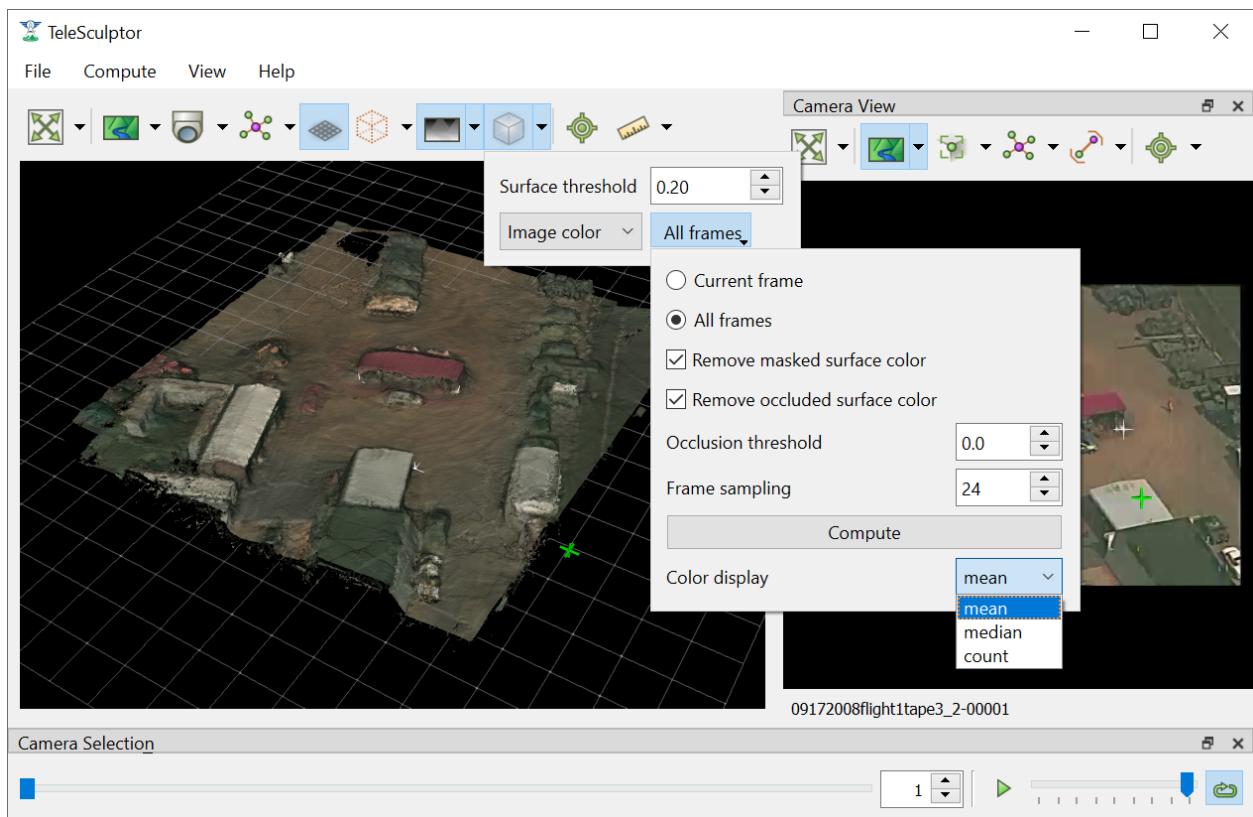Fig. 29: View the fused mesh, adjust surface threshold if desired.

Fig. 30: The mesh colorization menu options.

If *Image Color* is selected then the mesh vertices are colored by projecting color from one or more images onto the mesh. Additional settings on how the mesh is colored are in the right drop down menu. The *Current frame* option always projects the current frame onto the mesh and the color updates when you play back the video. The *All frames* option estimates a static mesh coloring by projecting multiple images onto the surface and combining them. The *Frame sampling* combo box allows configuration of how frequently to sample frames for coloring. Smaller sampling uses more frames for better color but more computation time. Press the *Compute* button to compute mesh color. When complete, the *Color display* option can be changed without needing to recompute color. The recommended color display option is *median*, however, *mean* is often quite similar. The *count* option colors by the number of depth map views that observed each part of the surface.
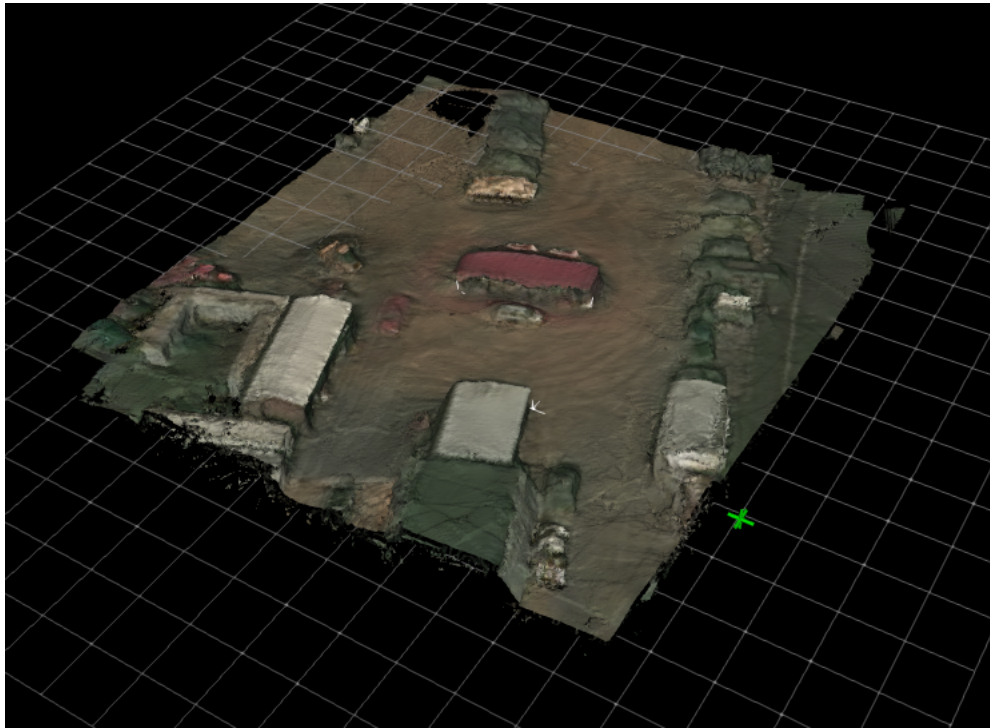

Fig. 31: A fused mesh colored by the mean of multiple frames.

Additional check boxes are available to enable removal of color on some surfaces that should not be colored. The *Remove masked surface color* option uses a mask video loaded into TeleSculptor and does not apply color at points under this mask. The *Remove occluded surface color* option enables visibility checking such that occluded surfaces will not receive color. The *Occlusion threshold* varies the sensitivity of the occlusion depth test. Negative values cause larger occluded regions while positive values cause smaller occluded regions.

3.4.12 Export Data

To export the finale colorized mesh for use in other software, use the *File->Export->Fused Mesh* menu item. This will provide a file dialog to save the model as a mesh in standard PLY, OBJ, LAS, or VTP file formats. The LAS file format will only save the dense mesh vertices as a point cloud and does include geo-graphic coordinates. The other formats save the surface mesh but only in local coordinates. Note that all formats (except OBJ) will also save RGB color on the mesh vertices. This color matches whatever display options are currently set.

To export the active 2.5D depth map for use in other software, use the *File->Export->Depth Map* menu item. This will provide a file dialog to save the model as an RGB colored point cloud in the standard PLY or LAS file formats.

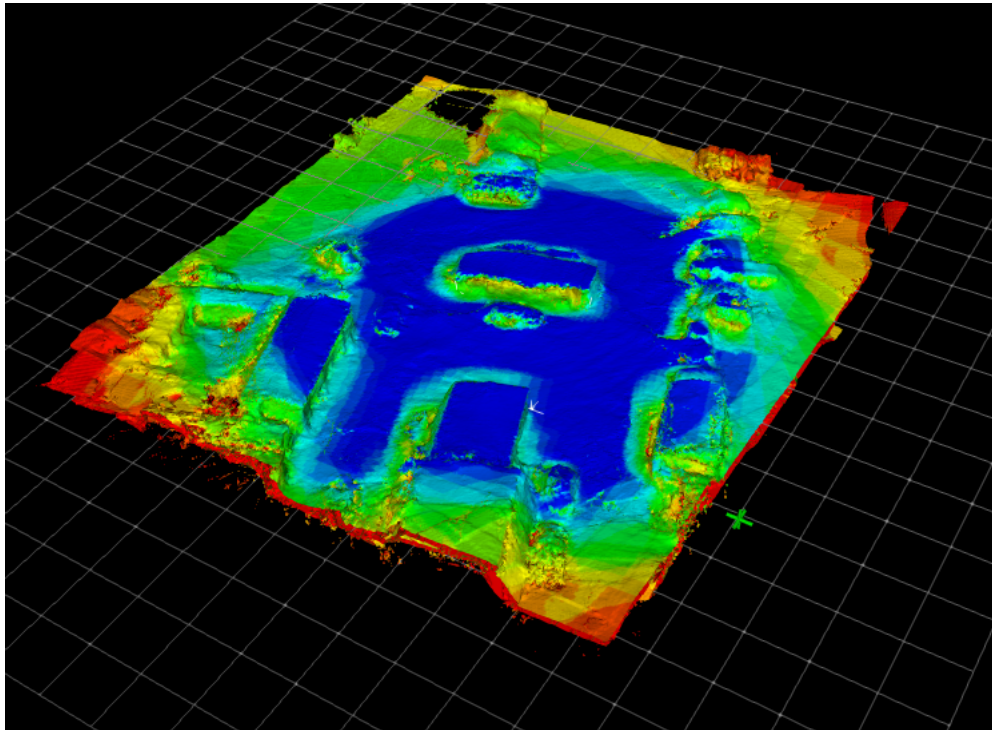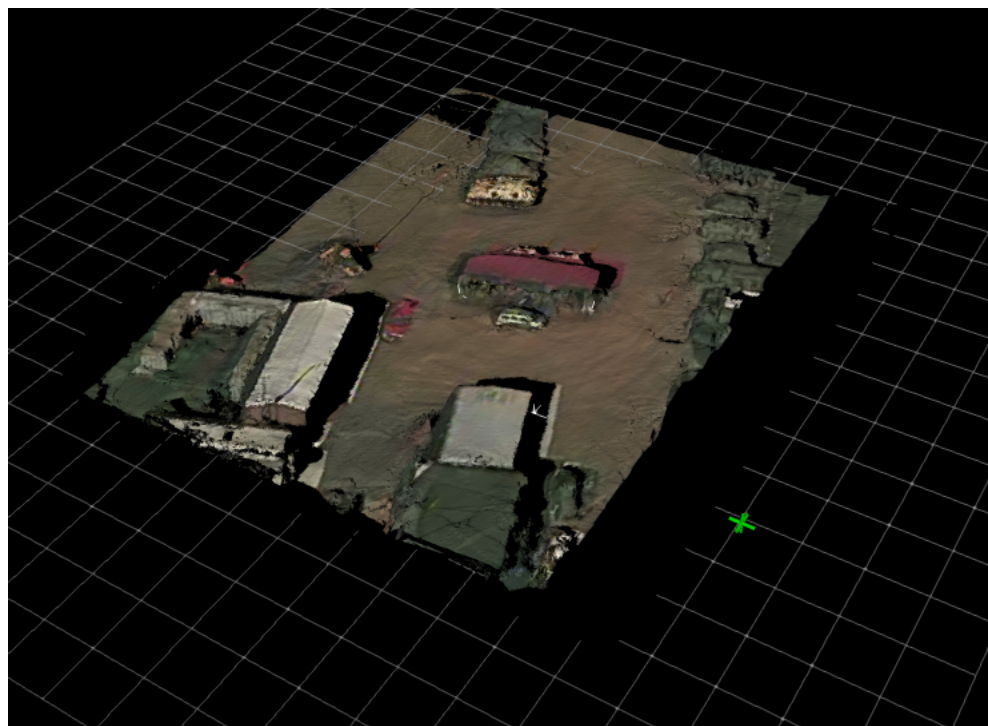Fig. 32: A fused mesh colored by the number of views that see each part of the surface.

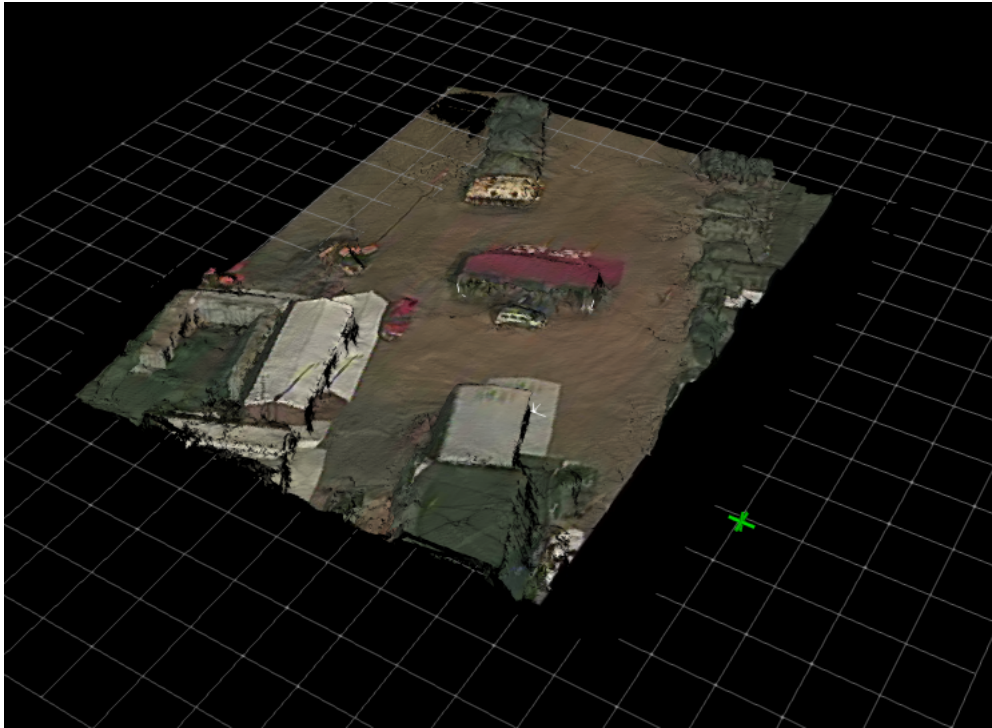Fig. 33: A fused mesh colored by a single frame using occlusion detection.

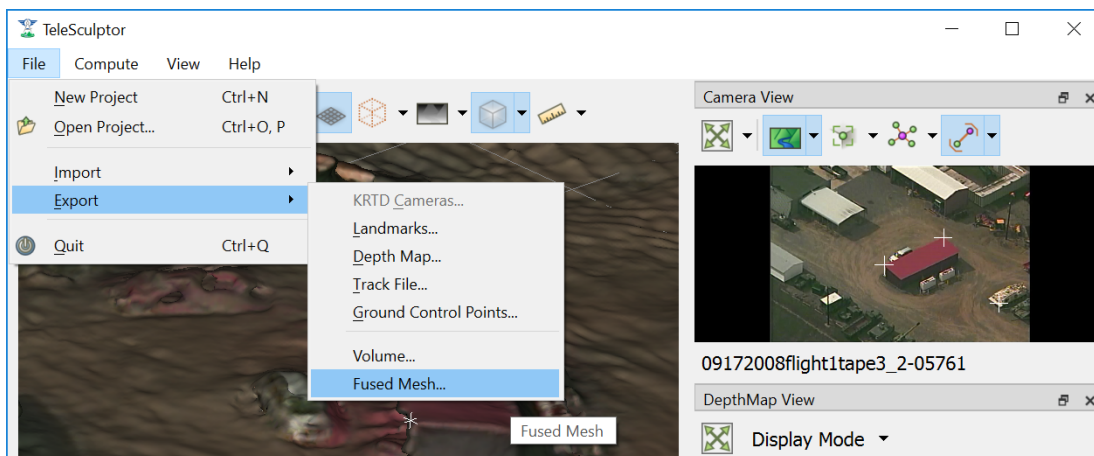Fig. 34: A fused mesh colored by a single frame without occlusion detection.

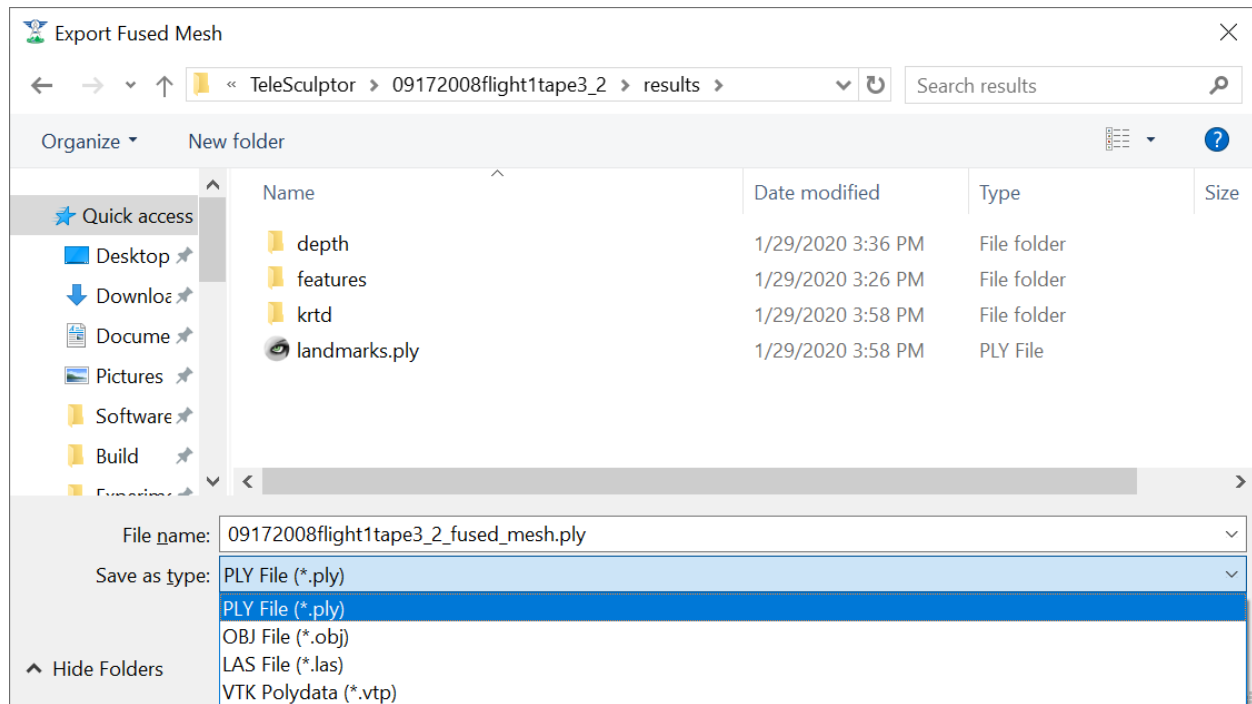Fig. 35: Save a colorized mesh as a PLY, LAS, or VTP file.

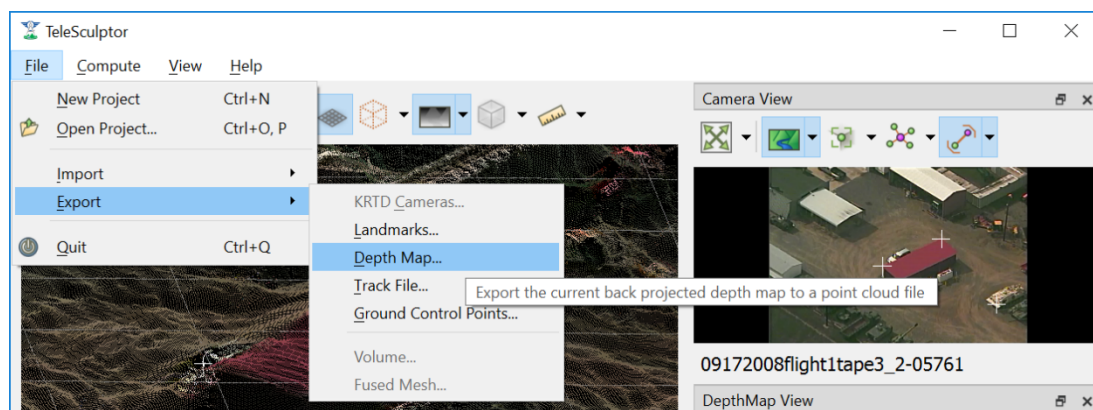Fig. 36: Change the Save as type to export as LAS.

Fig. 37: Save a computed depth map as a point cloud in PLY or LAS formats.

3.4.13 Measurement Tool

The measurement tool ( ) allows the user to measure straight line distance in world coordinates. Placing the end points of the ruler uses a similar interface to placing GCPs, and each end point can be adjusted independently just like GCPs. The number displayed next to the green line in both world and camera views represents the distance in world space. If geo-spatial metadata is provided the measurements are in units of meters. Without metadata (as in the example below) the measurements are unitless. The ruler can be drawn and adjusted in either the world or camera views. Often it is easier to get more accurate alignment in the image space. As with GCPs, if the ruler sticks to the correct location when playing back the video then the 3D coordinates are correct.
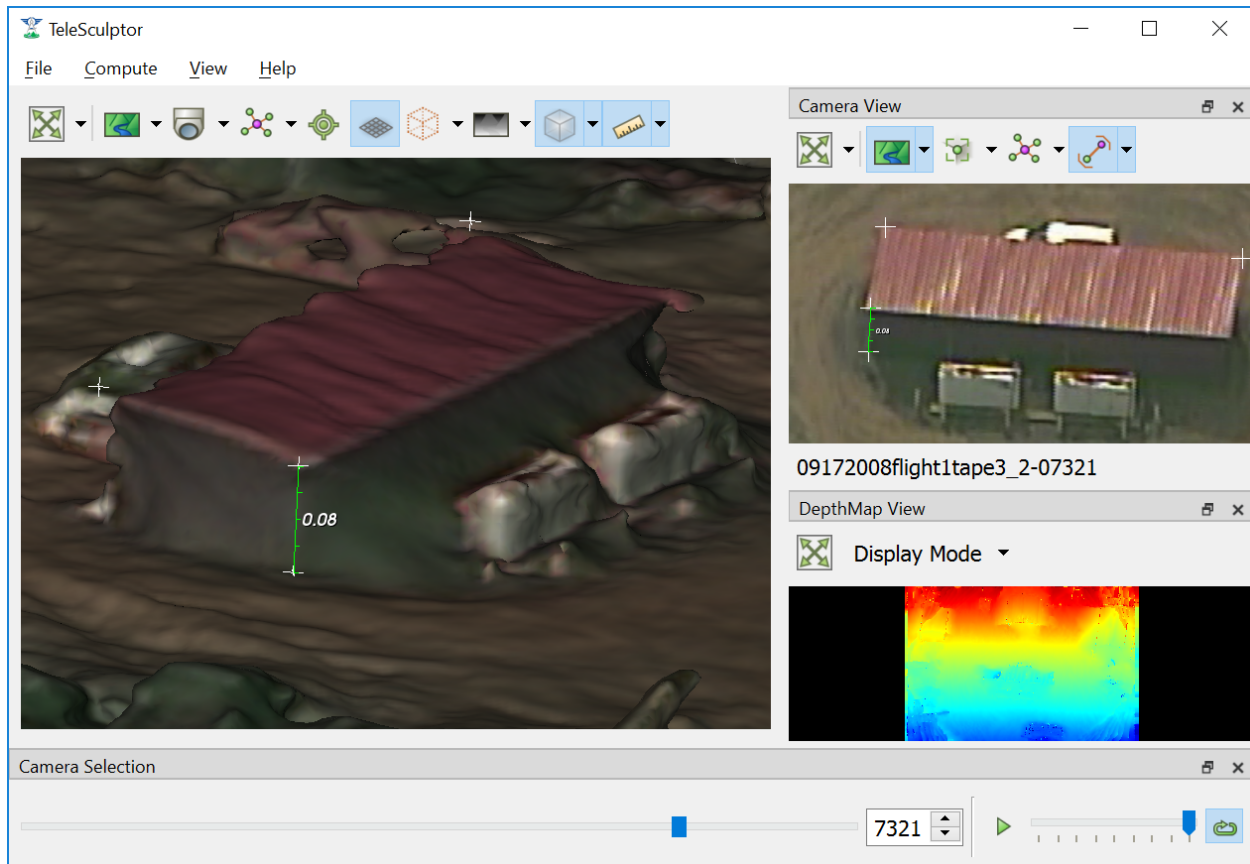
Fig. 38: Measuring a building height with the measurement tool.

When measuring it is sometimes convenient to constraint the measurements to the horizontal or vertical directions. After the initial ruler is placed in the scene, click and drag one end point. If the Z key is held on the keyboard the moving point will be constrained to lie on a vertical axis through the point at the other end of the ruler. If either the X or Y keys is held, the moving point will be constrained to lie in a horizontal (X-Y) plane that passes through the other ruler point. Each of these constraints is indicated by an indicator as shown below.
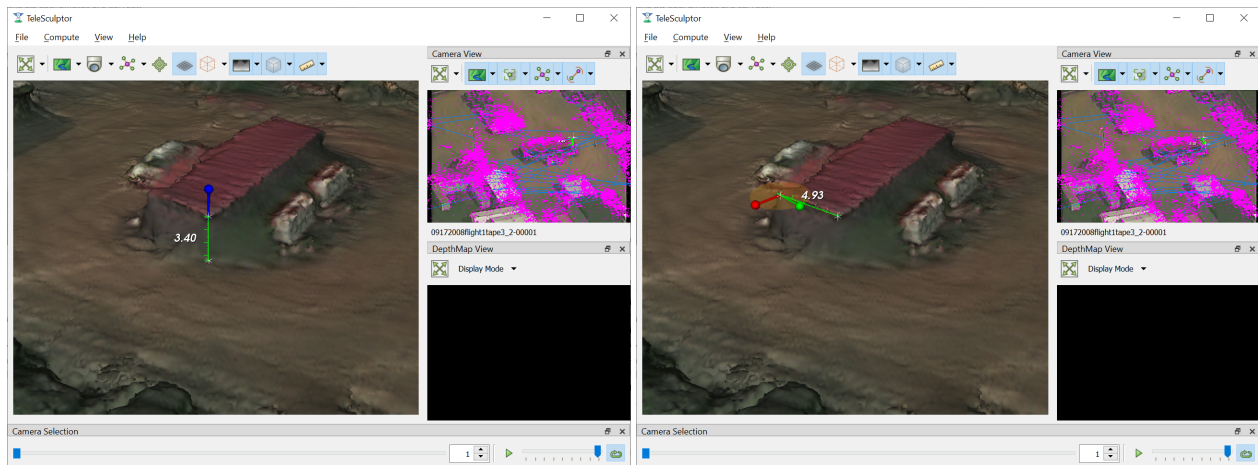
Vertical Constraint (hold Z) Horizontal Constraint (hold X or Y)

3.5 Advanced Tools
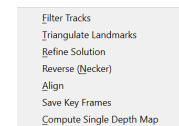
The following tools are under the Compute->Advanced menu. Most users should not need these tools, but they may come in handy on challenging data sets where the normal compute steps do not work ideally.

3.5.1 Filter Tracks

This tool filters the set of tracks to find a reduced set of tracks that spans the same frame range. It tries to keep the longest, most stable tracks and throws out many short tracks. The goal is to make bundle adjustment more efficient by limiting the solution to the most important tracks. Since we now limit how many features are tracked to begin with, this tool does not usually provide a benefit.

3.5.2 Triangulate Landmarks

Triangulate Landmarks attempts to create a 3D landmark position for each feature track by back-projecting a ray through the feature point locations in each image and intersecting these rays in 3D space. The triangulation is fast and should finish almost instantly, however it requires accurate camera models to work. The *Triangulate Landmarks* tool requires both feature tracks and cameras to work.

If the metadata for camera poses is very accurate one can bypass the Estimate Cameras/Landmarks step and directly triangulate landmarks. The triangulated positions will likely still be very noisy, but this can be improved by bundle adjustment using the *Refine Solution* tool below.

3.5.3 Refine Solution

Refining the solution optimizes the calibration and pose of the cameras as well as the locations of the 3D landmarks in a process known as bundle adjustment. This tool requires tracks, landmarks, and cameras to run. It adjusts the parameters such that the bundle of rays used to triangulate each point meets more precisely at a single location.

Bundle adjustment is already run as part of *Estimate Cameras/Landmarks*, but this option allows it to be run directly. Direct refinement is useful for features that are directly triangulated from metadata cameras.

Running this algorithm can take some time. While it is running, the solution incrementally improves, and updated results are displayed as the solution evolves. Bundle adjustment tends to make very large corrections very quickly and then spend lots of time fine tuning the final solution to get it just right. The cancel option allows the user to exit the optimization early and keep the current state of progress. Cancelling is useful when the solution is taking too long to complete but appears to have found a reasonable solution. It helps get a solution more quickly but beware that a suboptimal solution will impact the quality of later 3D reconstruction stages. It is better to wait for completion when time is available.

3.5.4 Reverse (Necker)

There is a special type of failure mode in camera calibration that only happens with very long focal length cameras. This failure mode happens because of a depth reversal ambiguity that occurs when perspective distortion is lost and the projection is nearly orthographic. The solution is bistable, just like to famous Necker Cube optical illusion. Under this "Necker Reversal" the one can invert the height of landmarks and flip cameras upside down and mirror them across the orbit to produce nearly identical geometric projection. The Reverse (Necker) tool flips the data into this alternate configuration to rectify this invalid solution. Note that Necker Reversal is not a problem when geospatial metadata is available or when one can make assumptions about the orientation of the cameras (e.g. up in the world is up in the image). The Initialize Cameras/Landmark tries to automatically detect and correct for this ambiguity, so this the manual correction is rarely needed.
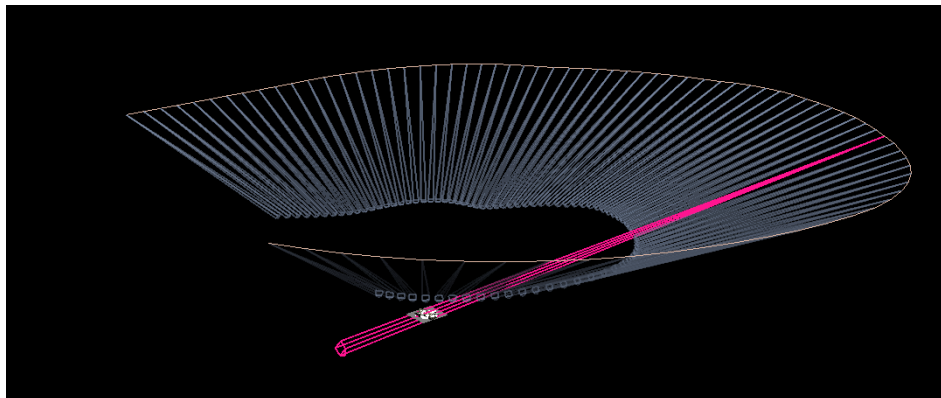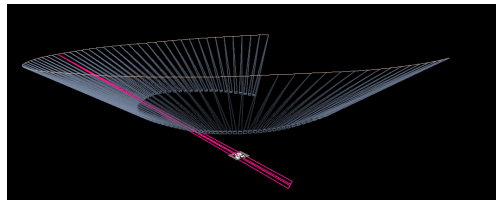
Fig. 39: Effect of Necker Reversal on camera orbits. Bottom cameras are flipped and upside down.

3.5.5 Align

The *Align* tool is for videos that do not have metadata. Without metadata the orientation, position, and scale (7 degrees of freedom) of the solution are completely undetermined. The solution floats freely in space. The *Align* tool attempts to align the data to a canonical coordinate system with the following properties: The centroid of the landmarks is aligned to the origin. The direction of minimal landmark variance is aligned with the Z axis with cameras on the positive Z side. The scale is set such that the variance is landmark positions from the origin is one. The origin is

shifted along Z such that 90% of landmarks are above Z=0 (ground plane estimation). This algorithm is also now run automatically as part of Initialize Cameras/Landmarks, so this manual tool is rarely needed.

3.5.6 Save Key Frames

The *Save Key Frames* tool is the same as the Save Frames tool except that it only saves frames that are marked as key frames by the *Track Features* tool. Saving only key frames makes more sense than saving all frames for use in SketchUp. However, the selection of key frames is not currently reliable. Sometimes only one keyframe is selected. One could try this option first and then save all frames if not enough keyframes are available. To preview the key frames select *Keyframes Only* from the *View* menu and play back the video.

3.5.7 Compute Single Depth Map

The *Compute Single Depth Map* tool is the same as the *Batch Compute Depth Map* tool except that it only computes on depth map on the current frame. The tool provides live visualization of the intermediate results to visualize how the solution evolves over compute iterations. The initial estimated point cloud will be quite noisy and then will continue to improve with live updates shown in both views as the algorithm progresses. Much like the refine solution tool, the solution improves quickly at first and then spends a longer time fine tuning the solution. The cancel option will also end the optimization early for this tool.

3.6 Compute Options

A few basic switches are available under the Compute->Options menu to control the behavior of some of the algorithms. Each of these options corresponds to a boolean configuration value in the project file. Checking these options will save that option in the project file. In the future, we will provide a visual interface to configure many more of the TeleSculptor algorithm options. For now these basic options are presented in the menu:

3.6.1 Ignore Metadata

When checked, this option causes the Initialize Cameras/Landmarks algorithm to ignore any metadata that was loaded with the video. This option is useful because it is often the case that the metadata is incorrect and negatively impacts the algorithm rather than helping it.

3.6.2 Variable Lens

By default, TeleSculptor assumes that all frames in a video are collected with the same lens and zoom setting. The intrinsic camera parameters are shared across the entire sequence. This assumption gives the best results as long as the assumption holds true. When the assumption does not hold, and the lens zooms or is changed in the middle of the sequence, checking "Variable Lens" will instruct TeleSculptor to estimate unique camera intrinsic parameters for each video frame.

3.6.3 Fix Geo-Origin

TeleSculptor automatically selects a local origin near the centroid of the data and records the geographic location of this origin point. When the data is updated by running algorithms that origin point is recalculated and may change. In some cases, there are benefits to specifying the geographic origin to use and keeping it fixed, for example, forcing two data sets to share a common local origin for easier comparison. Checking "Fix Geo-Origin" instructs TeleSculptor to keep the current geographic origin and not recalculate a new one.

3.6.4 Use GPU

Some TeleSculptor algorithms (currently, depth map fusion) are able to use CUDA to improve performance. However, this requires an NVIDIA GPU with sufficient GPU RAM. Disabling this option instructs TeleSculptor to rely solely on the CPU. This is slower, but should give the same results and may be used if a suitable GPU is not available.

3.7 Advanced Configuration Options

Most users should not need these advanced features, but some may be interested.

3.7.1 Changing Frame Sampling Rate

By default, all video frames are loaded when opening a video (though not all are processed in feature tracking). If the number of frames is too large to manage it is possible to read only every Nth frame by changing a setting in the project configuration. First, close the application. Next, open the project .conf file in a text editor like Notepad. Look for the following line:

```
video_reader:filter:output_nth_frame=1
```

Increase the number from 1 to 10 to sample every tenth frame, for example. Save the .conf file in the text editor and open that file again in the TeleSculptor application.

3.7.2 Modifying Algorithm Parameters

TeleSculptor is a highly configurable application, though most of the configuration options are not yet exposed to the user interface. Each tool in the compute menu calls an algorithm from the KWIVER toolkit and each algorithm is configurable at run time. Algorithms can even be swapped out for other algorithms at run time. One can contribute a new algorithm without recompiling TeleSculptor by dropping in a new .dll and updating the configuration files. All this configurability is managed with configuration files. The project file is one example of configuration file, but there are also many default configuration files loaded by TeleSculptor at run time. The default configuration files for a standard install path are found in these two locations:

```
C:\Program Files\TeleSculptor 1.2.0\share\telesculptor\1.2.0\config
```

```
C:\Program Files\TeleSculptor 1.2.0\share\kwiver\1.6.0\config
```

TeleSculptor specific configurations are found in the first directory and these include configurations for KWIVER algorithms found in the second directory. It is recommended that you not modify these values, but instead copy some of these files into your project directory and modify the copies. TeleSculptor will load configuration files from the project directory first. Each of the tools in the *Compute* menu loads a configuration file when it is run. These have names starting with a gui_ prefix. For example:

Compute Menu Name	Configuration File Entry Point
Track Features	gui_track_features.conf
Estimate Camera/Landmarks	gui_initialize.conf
Save Frames	gui_frame_image_writer.conf
Batch Compute Depth Maps	gui_compute_depth.conf
Fuse Depth Maps	gui_integrate_depth_maps.conf

Most of these configuration files also reference other configuration files with an `include` statement. Configuration values for tools can also be added to the project file but copying the GUI configuration files into the project directory adds more flexibility because these files are reloaded each time the tool is run, which allows changing parameters between runs without loading the project.

As an example, consider changing the maximum number of frames to use in the feature tracker. First, copy gui_track_features.conf into your project directory. Open this file and look for the following section:

```
# Parameters for the feature tracker
block feature_tracker
  include core_feature_tracker.conf

  # The maximum number of frames for the GUI to process.
  # The tracker will choose frames distributed over the video
  max_frames = 500
endblock
```

Modify the line with `max_frame = 500` to use a different value, such as `1000`. Note that you could also make this change in the project file by appending the following line:

```
feature_tracker:max_frames = 1000
```

Note that the `max_frames` parameter is in the `feature_tracker` scope and scope must be specified either using `block`/`endblock` notation or with a prefix before a colon.

3.7.3 Printing all KLV metadata

The TeleSculptor application loads KLV metadata and display it in a viewer, but there is no way to export this data in batch. However, the installer does provide the kwiver command line tool that has an applet that will print out all metadata in a video. This applet is called dump_klv. The default installation path is:

```
C:\Program Files\TeleSculptor 1.2.0\bin\kwiver.exe
```

To run dump_klv, open up a command prompt (search for "cmd.exe" in the Start Menu). Then run:

```
"C:\Program Files\TeleSculptor 1.2.0\bin\kwiver.exe" dump_klv video_file.mpeg
```

and replace video_file.mpeg with the path to the video file to process. This will print out all the metadata. To redirect the output to a file use:

```
"C:\Program Files\TeleSculptor 1.2.0\bin\kwiver.exe" dump_klv video_file.mpeg
↪> metadata.txt
```

3.8 Manual Camera Calibration

This section describes the functions for manual camera calibration capabilities in TeleSculptor (TS) and steps to achieve, possibly correct, and export a camera calibration model.

3.8.1 Preparation

TeleSculptor can work with ffmpeg-readable videos (e.g. mpg, mp4, avi) or with image sequences. The latter can be specified by listing paths to individual images in a plain text file (e.g. images.txt) using one line per path/to/image/file. Standard image formats (e.g. bmp, png, jpg) are supported.

TeleSculptor has a capability to import PLY mesh or point cloud files via File → Import → Mesh. Point clouds are imported as faceless meshes. The imported mesh is used as a 3D target for calibration. RGB[A] color space is currently supported. The alpha channel can be used to set mesh transparency levels.

We recommend geo-registering the mesh before importing it into TS to have the output calibrated camera parameters in mesh metric units and also automatically geo-registered with the 3D model.

3.8.2 Calibration

Start the TeleSculptor GUI and proceed with the following steps in a new or an existing project.

New project

Create a new project using File → New Project (Ctrl+N) and select a folder for the new project, creating one if necessary. The project .ini file name will have the same name by default.

Select a video or image sequence text file using File → Import → Imagery... (Ctrl+O, I). Opening a large videos will initiate a potentially long search for metadata, which may be canceled via the menu action Compute → Cancel.

Select a PLY file with a mesh or a point cloud using File → Import → Mesh... (Ctrl+O, M). The loaded 3D geometry should appear in the 3D world view.

Locate the desired video frame or image using the scrubber or spin box in the Camera Selection pane. Rotate the model in the world view as needed so the desired point correspondences are visible.

Select ⊕ Edit Ground Control Points (in the world view) to activate editing mode for ground control points in the world view. Place points using Ctrl+Click. Telesculptor attempts to place points "on" the loaded mesh, but it may be helpful to rotating the world view to verify that points are placed in the right 3D location. While editing, points may be dragged as needed. Newly placed points will appear in the Ground Control Points list, and may be given names if desired.

To create 3D-to-2D point correspondences, select ⊕ Edit Registration Points (in the camera view). Ensure that the desired 3D point is selected in the Ground Control Points list (the selected point is highlighted in a different color). Place a corresponding point using Ctrl+Left Click, positioned to match the location in the camera image to which the 3D ground control point would be projected. The Ground Control Points list will show a glyph () to indicate that the point has a corresponding camera registration point associated with the active video frame or image. Once created, points may be dragged as needed to fine tune their locations. Select another 3D ground control point and repeat until at least six correspondences have been created for the active video frame or image. The relevant edit mode actions may be used at any time to switch between editing points in the camera or world views.

Once six or more correspondences have been defined, the camera model calibration process may be invoked from the Compute Camera action located in the drop-down associated with ⊕ Edit Registration Points. Check the Log Viewer for the re-projection error or for any error messages (e.g. insufficient calibration points). Upon successful calibration, the world view will show the projected camera image (if 🖼 Show Camera Frame Image is enabled), and the camera view will show the ground control points projected onto the camera image.
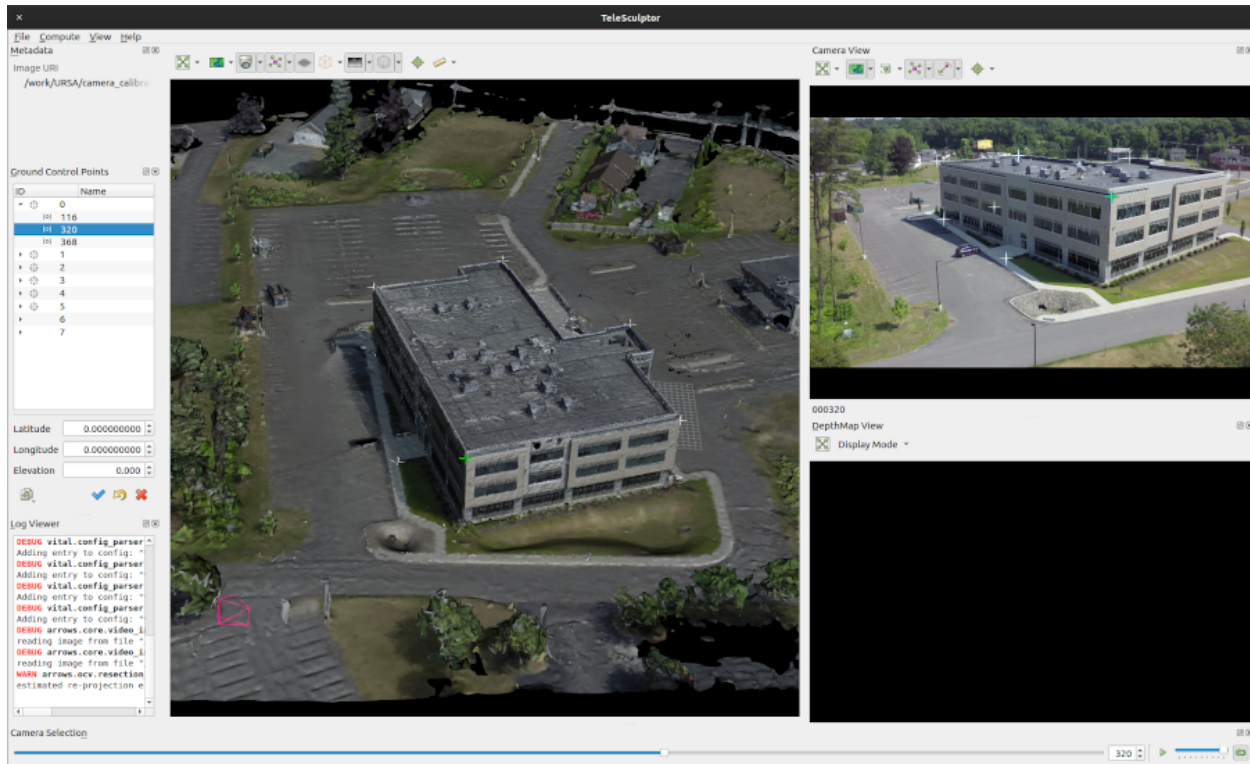
Repeat this process on additional frames. Point locations may be adjusted as needed, as described above, to refine the camera model calibration. Camera calibrations may be refined at any time; it is not necessary to "finalize" one camera model before calibrating additional camera models.

When satisfied, File → Export → KRTD Cameras. . . may be used to save the parameters of all calibrated cameras to KRTD text files. Each file will capture the intrinsic matrix K, the rotation matrix, the translation vector and the lens Distortion parameters.

File → Export → Ground Control Points. . . will export all of the ground control points and their corresponding camera registration points as a GeoJSON-like file. (Only the geodetic locations, if available, conform to GeoJSON; the world and camera locations are stored as extensions.) This will output all points, whether or not calibration was performed successfully.

Existing project

Open an existing project using File → Open Project. . . (Ctrl+O, P). The project's corresponding video or image sequence, and the 3D mesh or point cloud, will be loaded automatically. Follow the calibration procedure described for new projects, above.

3.9 Glossary

Frame: In order to work, photogrammetry requires multiple images of a scene that have been captured from distinct angles. For simplicity, TeleSculptor refers to these distinct images as frames, whether or not they are derived from a single, continuous video. Although a frame corresponds to a camera pose, the term "frame" is used to indicate instances where the pose may not yet be known.

Camera Model: In TeleSculptor, a camera model (often shortened to "camera") describes the intrinsic and extrinsic parameters of a camera at a given frame. Intrinsic parameters include focal length and lens distortion. Extrinsic

parameters are the camera pose — orientation and position. A frame may have at most one camera pose, which is unique to that frame. While extrinsic parameters are never shared across frames, intrinsic parameters may or may not be shared depending on the algorithm that produced them.

In the GUI, cameras are represented as rectangular pyramids. The peak of the pyramid is the camera center of projection. The rectangle base is proportional to the image aspect ratio, and a triangle attached to the base indicates which direction is "up" in the image. The pyramid visualizes the field of view of the camera.

Feature: A feature describes a salient point in an image. Features have enough visual texture that they can be reliably localized in images across time. Features are also known as interest points or corner points.

Track: A (feature) track is a collection of correlated features; that is, detected feature points estimated to correspond to the same landmark.

Landmark: A landmark is an estimated 3D world point that, when projected into the images, gave rise to an observed feature track.

Residual: A residual, in general, is the difference between an observed value and an estimated value[1]. In TeleSculptor, the observed value is typically a detected feature point, and the estimated value is the projection of its corresponding landmark into the image.

Ground Control Point (GCP): A ground control point is a 3D point that is manually placed in the scene by a user. Ground control points are often attached to an identifiable location from a reference image or map. The user can assign the true geodetic location (latitude, longitude, elevation) to a ground control point, and that ground control point can then serve as a constraint when geo-registering the model.

Camera Registration Point (CRP): A camera registration point is a 2D point that is manually placed by the user in one or more frame images. Camera registration points are associated with a ground control point and can tie that ground control point to multiple images, even when no camera model has been estimated. Whereas a ground control point is the manually placed version of a landmark, a camera registration point is the manually placed version of a feature track. A collection of camera registration points can be used to manually estimate a camera model when automated feature matching is not possible.

---

[1] https://en.wikipedia.org/wiki/Errors_and_residuals_in_statistics