Structure from Motion Introductory Guide

Katherine Shervais (UNAVCO)

This guide is intended as a resource for using Structure from Motion in research applications. It does not detail the algorithms or mathematical background of the methodology, but rather how to use it in practice.

Table of Contents

Introduction

Structure from Motion or SfM is a photogrammetric method for creating three-dimensional models of a feature or topography from overlapping two-dimensional photographs taken from many locations and orientations to reconstruct the photographed scene. This technology has existed in various forms since 1979 (Ullman, 1979), but applications were uncommon until the early 2000's (Snavely et al., 2008). The applications of SfM are wide-ranging, from many sub-fields of geoscience (geomorphology, tectonics, structural geology, geodesy, mining) to archaeology, architecture, and agriculture. In addition to ortho-rectified imagery, SfM produces a dense point cloud dataset that is similar in many ways to that produced by airborne or terrestrial lidar (see figure below).

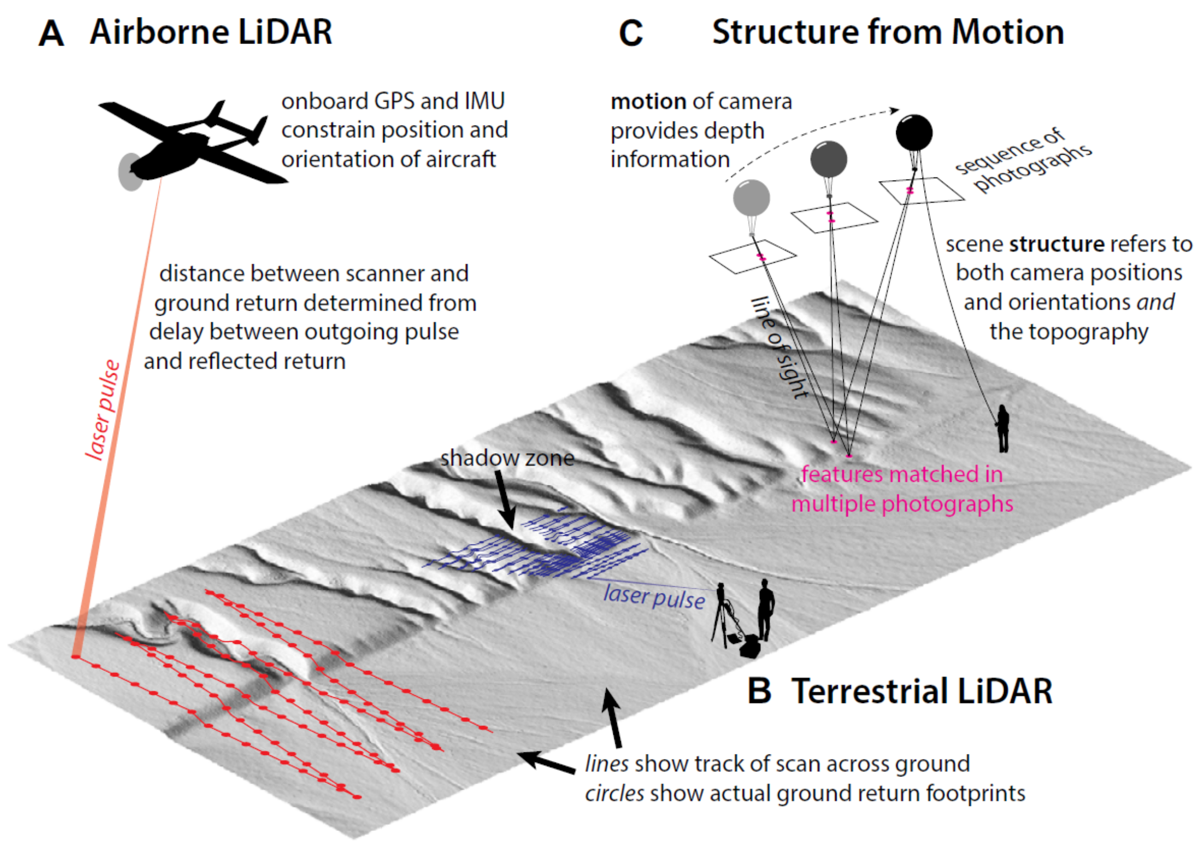

Figure 1: Comparison of SfM to airborne and terrestrial lidar methods. Source: Johnson, et al., (2014).

The advantages of SfM are its relative cost in comparison to lidar, as well as its ease of use. The only required equipment is a camera. A computer and software are needed for data processing. Additionally, an aerial platform like a balloon or drone can also be useful for topographic mapping applications. A major limitation is the processing time for software to align the images to generate a model, which ranges from 10 minutes for a few photographs to days for 100s-1000s of photographs. Because SfM relies on optical imagery, it is not able to generate the "bare earth" topographic products that are typical derivatives of lidar-based technologies – thus, SfM is usually best suited to areas of limited vegetation.

2. Platforms

Depending on the application and survey objective, a camera can be mounted in numerous configurations to capture imagery for SfM processing. This section provides an introduction to various platforms that can be used to acquire imagery (Table 1).

Some things to consider:

- Many of the platform descriptions include a reference to a picavet; picavets are simply a camera mount designed to keep a camera vertical even if the line the camera is suspended on is not. Picavets are available to purchase, but there are guides on the Internet for those who wish to build their own.
- These platforms are highly customizable. The limiting factor is typically weight, but one has a kite or balloon with enough lift, the system could include a remote control camera mount to change the orientation of the photos during the survey. Other additions include a radio controller to regulate when photographs are taken if your camera does not have a time-lapse, and a GPS tagger on the camera to georeference photographs if the camera does not have an internal GPS. UAVs can now be purchased with GNSS receivers, expediting the process of aligning photographs in the software.
- To conduct an Internet search for the equipment related to one of these platforms, try searching for "<platform> aerial photography" (ex. kite aerial photograph). This should yield platform-specific equipment designed to accommodate a camera.

Table 1. Cost, advantages, and disadvantages of common SfM platforms.

Platform	Cost	Advantages	Disadvantages
Person (hand-held)	$0	Cost, good for detail work (characterize specific, small-scale features), potentially more efficient for outcrop scale work	Limited applications; not useful for areas larger than 100-200 square meters
Pole	~$50-250, depending on pole height. Can purchase kits online, but building your own is often more economical.	Cost, ease of use/ease to set up, good for certain kinds of features (slope underneath an overhang, for example)	Must build mount for camera, limitation for maximum pole height, inefficient in comparison to UAV
Balloon	~$300-5000 – building your own system is inexpensive, but systems on the market can be significantly more expensive because they	Cost (unless you purchase a commercial version), simplicity, camera orientation (can shoot straight down, unlike many pole	Easier with 2 people rather than 1, affected by the wind, requires picavet mount (build or purchase), requires helium (a limited resource)

	can include video systems to view what the camera is viewing. A weather balloon is approx. $100, the picavet mount is approx. $50, kite line ~$20, and helium ~$180 for an appropriately sized tank	setups), height. Balloons are a good option for topographic mapping applications. Tether provided by the line removes legal complications associated with UAVs	
Kite	Cost depends on the weight of the camera mount system. Kites can be used with v. lightweight cameras and cost around $50; kites made for aerial photography can cost $100-400. For both options, you will need to purchase or build a picavet for $50.	Cost, height, camera orientation (can shoot straight down, unlike many pole setups), similar range to a balloon but no helium! Well suited to topographic mapping applications. Tether provided by the line removes legal complications associated with UAVs	Dependent on weather, must build/purchase picavet for camera, kite line can get in the way of photographs, kite must be large/have good “lift”
UAV – motorized glider, multi-rotor copter (quad-, hexa- or octa-)	Cost is highly variable; a motorized glider is around $200-300 + cost of picavet; quadcopters can range from $400-$5000 or more depending on their capabilities.	Height, camera position may be controlled and survey flightlines can be pre-planned and automated, GNSS integration for efficiency	Cost, requires a skilled operator, length of survey depends on the charge of, may require light camera setup. Potentially dangerous if improperly operated. Legal landscape for use of UAVs for anything more than recreation is unclear, and users should consult legal counsel before operation.

Lift

Kite lift is more difficult to determine in a simple chart, as it is highly dependent on the kite and the wind at any give time.

Balloon lift is simpler because it is dependent on the helium. Use this chart (minus the weight of the balloon) to determine the size of balloon needed to lift your camera setup.

Table 1: Balloon size relationship with lift in pounds or grams. Source: University of Hawaii, Manoa, Department of Chemistry (http://www.chem.hawaii.edu/uham/lift.html)

Diameter (feet)	Diameter (meters)	Volume (liters)	Lift (grams)	Lift (pounds)
1	0.3048	14.83	15.2	0.03
2	0.6096	118.62	121.7	0.27
3	0.9144	400.34	410.9	0.91
4	1.2192	948.96	973.9	2.15
5	1.524	1853.45	1902.2	4.19
6	1.8288	3202.76	3287	7.25
7	2.1336	5085.86	5219.7	11.51
8	2.4384	7591.72	7791.5	17.18
9	2.7432	10809.3	11093.7	24.46
10	3.048	14827.58	15217.7	33.55
11	3.3528	19735.5	20254.8	44.65
12	3.6576	25622.05	26296.2	57.97
13	3.9624	32576.18	33433.3	73.71
14	4.2672	40686.87	41757.4	92.06
15	4.572	50043.07	51359.8	113.23
16	4.8768	60733.75	62331.8	137.42
17	5.1816	72847.88	74764.7	164.83
18	5.4864	86474.42	88749.8	195.66
19	5.7912	101702.34	104378	230.12
20	6.096	118620.61	121741	268.4
21	6.4008	137318.18	140931	310.7
22	6.7056	157884.03	162038	357.24
23	7.0104	180407.11	185154	408.2
24	7.3152	204976.41	210369	463.79

Examples of possible platforms

Figure 2. Pole aerial photography (PAP) platform. This example is from an archaeological site in Texas; the person is photographing the talus slope. Note the angle required to photograph given the camera cannot be oriented straight down. Source: https://aswtproject.wordpress.com/category/spring-enc-2014-posts/structure-from-motion/ (accessed July 1, 2015).

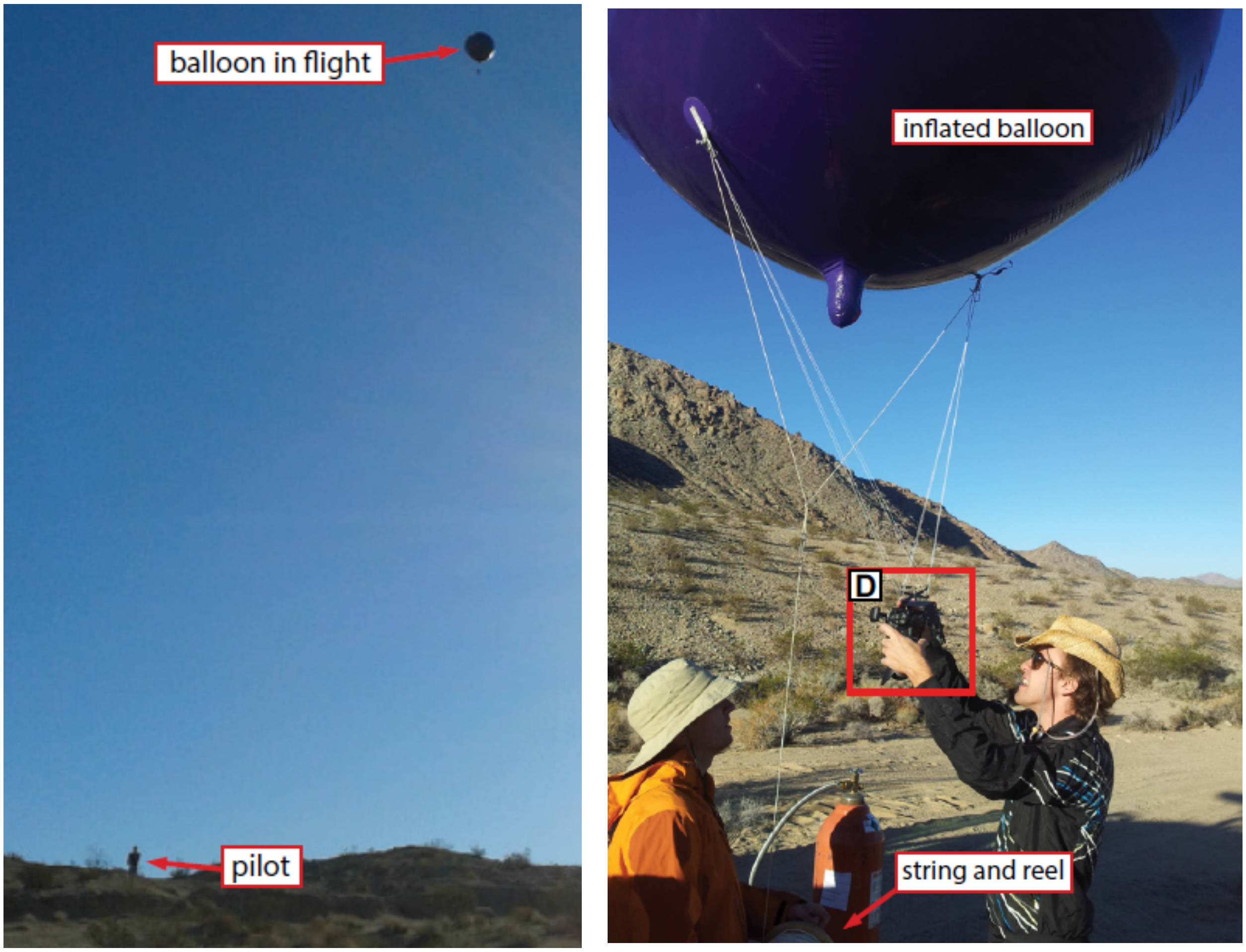

Figure 3. Balloon platform. The photo on the left shows the position of the balloon relative to the pilot while in flight; the photo on the right shows the more detail of the camera setup. Because of the picavet setup, the camera is oriented directly at the ground. See Figure 4 for a schematic of how this works. Source: Johnson et al. (2014).

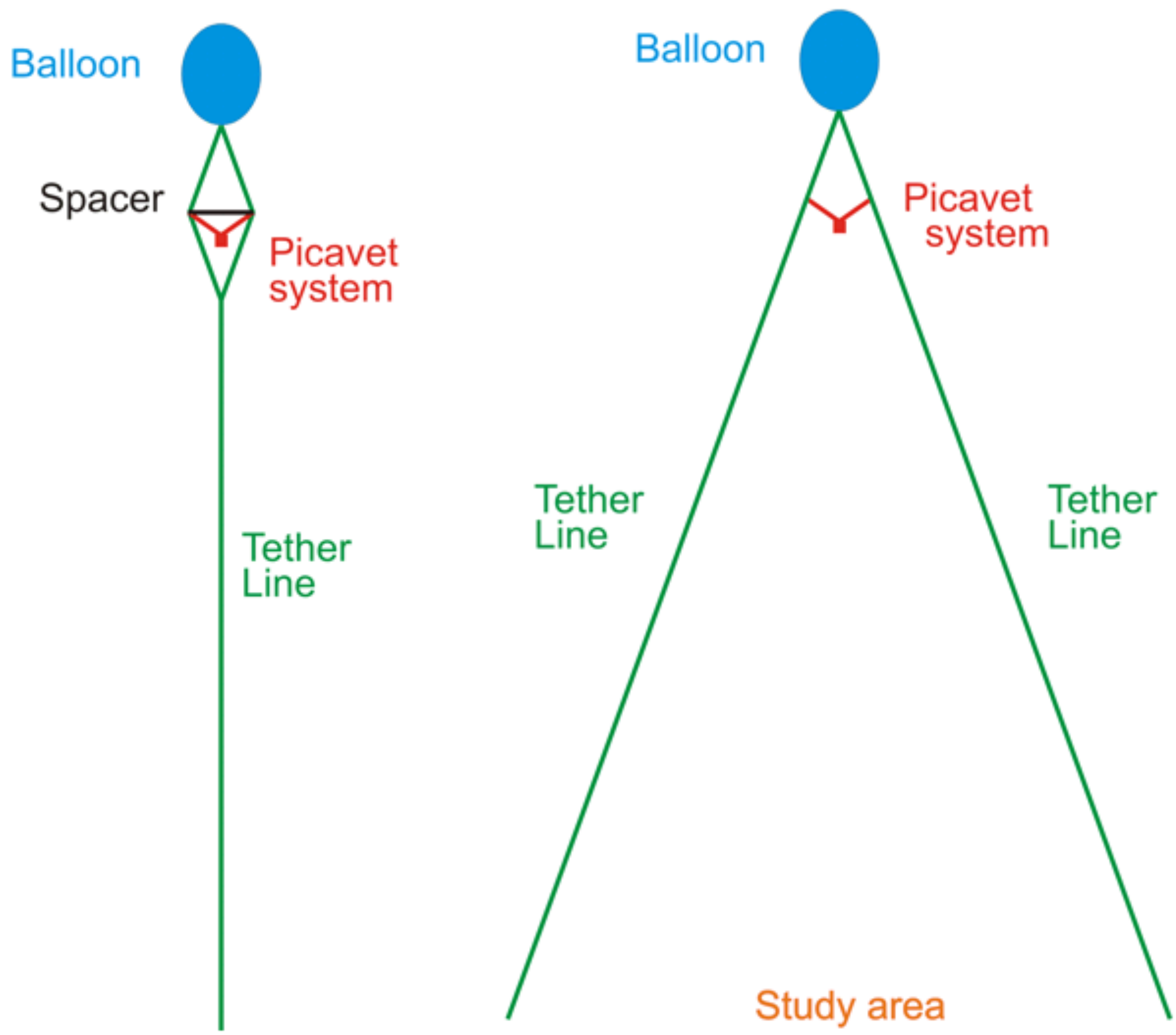

Figure 4. Demonstration of picavet system using a balloon platform. The left schematic demonstrates an ideal setup with one operator, while the right schematic demonstrates an ideal setup with two operators. Source: http://www.paulillsley.com/airphoto/systems/balloons-kites.html (accessed July 1, 2015).

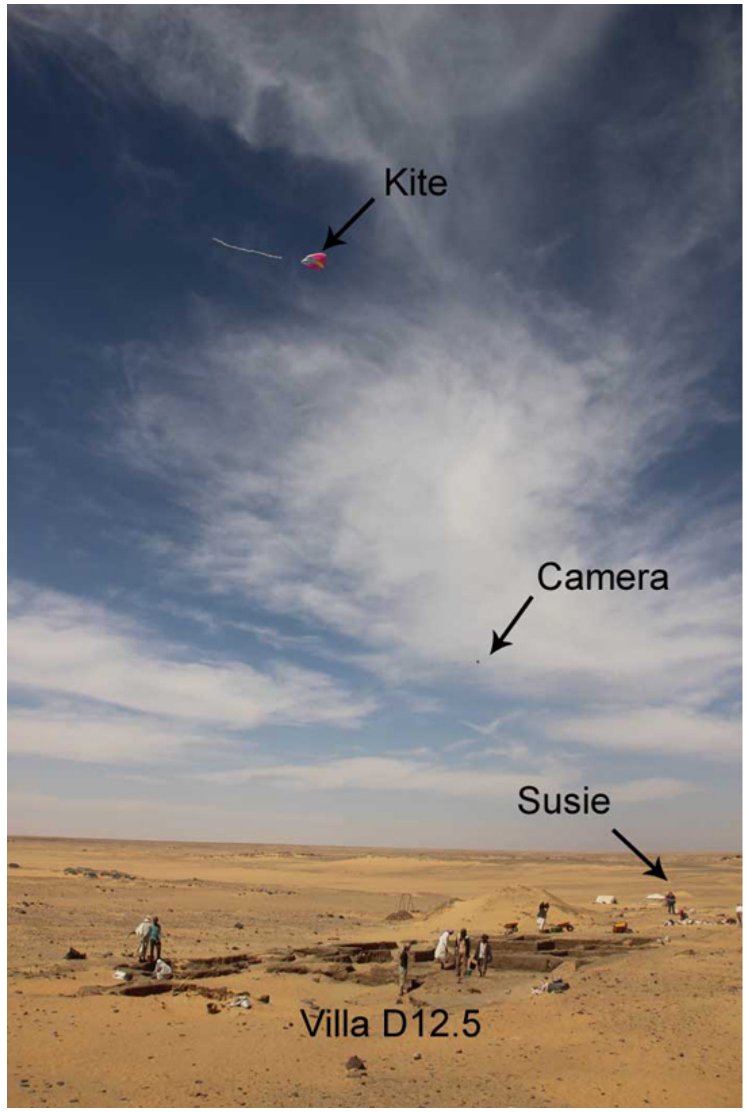

Figure 5. Above is a picavet setup for a kite – note this is basically the same as the setup for the balloon. However, this system is remote controlled with the camera mounted on a gimbal so the pilot can change the orientation of the camera while the kite is in flight. To the right is the location of the pilot (Susie) relative to the camera and the kite while in flight photographing Villa D12.5, the study area of interest. Source: http://blog.britishmuseum.org/category/archaeology/amara-west/ (accessed July 1, 2015)

Figure 6. Examples of pre-made and built UAV camera mounts. The top image is of a typical camera mount available to purchase in most hobby shops for UAVs. The example below is a camera attached to a quadcopter using bungee cords and a piece of foam padding to reduce the vibration transmitted to the camera. Sources: Turner et al. (2012), Remote Sensing, and http://www.paulillsley.com/airphoto/systems/Phantom/ (accessed July 1, 2015)

3. Camera choices

Rather than recommend a specific camera, compiled here is a list of general guidelines to follow when picking a camera. This list was synthesized from Johnson et al. (2014), http://www.paulillsley.com/airphoto/systems/Phantom/ (accessed July 1, 2015), Raugust and Olsen (2013), and personal communication with Kendra Johnson (Colorado School of Mines).

1. Consider the weight your platform is capable of carrying. A large balloon may be able to carry a heavier camera setup than a kite, for example.
2. Opinions about the usefulness of DSLR vs. point-and-shoot cameras vary. Most recommend using a DSLR (digital single-lens reflex) or a point-and-shoot that has faster ISO levels than average.
3. GPS: built in GPS tagger saves weight and simplifies the photo stitching process in some modeling software. It also produces more accurate georeferenced point clouds. If the camera you select does not have GPS, it is best to buy a GPS tagger in addition to help the data processing.
4. Shooting: the camera will need to take photographs at certain intervals. This can be accomplished one of three ways but cameras with time-lapse (also referred to as an intervalometer) or the ability to modify the camera to take time-lapse photographs are recommended.
    a. Time-lapse: the camera has the capability to take photographs at specific time intervals
    b. Continuous shooting: similar to time-lapse but the camera button needs to be pushed (this is simple – use a rubber band!)
    c. Remote controlled shooting: this option is the most complex, as systems may rely on radio so you will need to limit the height of the camera.
5. Resolution: some recommend staying at or above 12 megapixels, but the need for this has been debated.
6. Picture format
    a. RAW image files are most useful, so selecting a camera that is capable of this is a good idea.
    b. Do not use JPEGs because they introduce unnecessary noise.
7. If you would like the option of using a First Person View setup to view a live video feed of the camera sees, make sure the camera has a live video out option.
8. Manual exposure and focus: this ensures the images have a similar exposure. Manual focus helps the camera record images even if the autofocus is not working perfectly.
9. Avoid ultra-wide lenses such as those found on the GoPro!
10. Turn off "shake reduction"
11. VIDEO: do not use a video camera! Not all SfM algorithms work well with images recorded from video because of differences in how the shutter functions.

4. Software

SfM software typically falls into one of two categories: 1) Commercially available software, for which the workflow is more streamlined but the software is a "black box" and 2) Open source software, for which the workflow is more complex (several programs may need to be used in sequence, not some may not have graphical user interface [GUI]).

Commercial software

The primary commercial software used for research in geoscience and archaeology is Agisoft Photoscan. This software is available in Professional ($3499, $549 academic as of 07/01/15) and Standard ($179, $59 academic as of 07/01/15) editions. The Professional Edition is best for geologic SfM purposes, as it allows the use of ground control points, measurements, DEM (rather than just point cloud) export, and georeferenced orthomosaic export. Agisoft is a "black-box" but uses the SIFT algorithm used in the open source software (Verhoeven, 2011).

Pix4Dmapper is available for purchase or to rent. This program is $350 for a month-long rental, $3500 for a year, and $8700 for one time purchase (2 computer license). More analysis and data processing can be done in this program rather than exporting the point cloud or DEM to another program such Cloud Compare or ArcGIS.

Another commercial program is PhotoModeler Scanner, which is the PhotoModeler program optimized for UAV mapping. This program is $2495 and does not seem to have an education license. This program has been used less frequently for research applications, so has less peer-reviewed documentation about its capabilities.

Free web-based programs exist, like Microsoft Photosynth. Photosynth builds a point cloud, but the point cloud is not linked to physical coordinates.

Open source software

MANY open source software programs exist for SfM or portions of the SfM workflow. These programs range from GUIs to programs run from the command line to programs run through MATLAB (note that MATLAB requires its own, fairly expensive license - $2150, $500 for academic, $49-99 for student edition). An incomplete list is in this Wikipedia article: https://en.wikipedia.org/wiki/Structure_from_motion, which is one of the few sites to aggregate this kind of list.

Papers by Westoby et al. (2012), James and Robson (2012), Fonstad et al. (2013), and Green et al. (2014), go through different workflows using existing open source software. Green (2012) is a master's thesis that outlines the workflow for one method of using open source software and will be a good resource for anyone who prefers the open source option.

Figure 7 has a comparison between the Agisoft Photoscan Pro workflow developed by Johnson et al. (2014) and the workflows using open source developed by Westoby et al. (2012), James and Robson et al. (2012) and Fonstad et al. (2013).

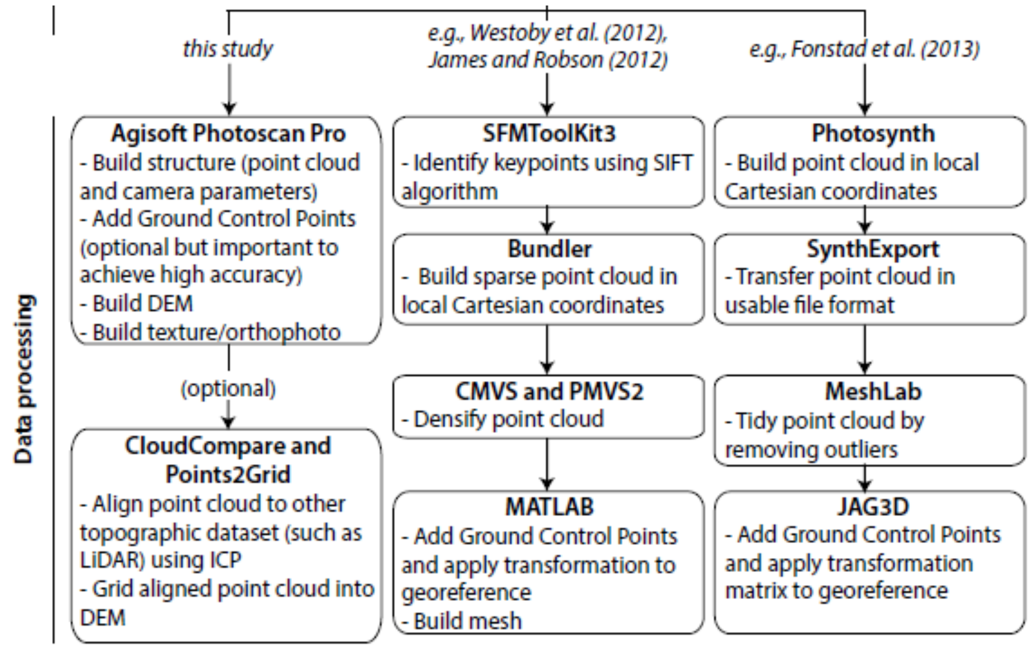

Figure 7: Comparison of Agisoft Photoscan Pro workflow to workflows using open source software. The end product is the same, but the open source software requires significantly more steps and a greater time commitment. Source: Johnson et al. (2014)

Table 2: Comparison of differing software for each step in the SfM process from Green et al., 2014. Though this covers similar software to above, it includes a few that are not listed in the Johnson figure. Note that all steps for Photoscan are in one software, but the others require knowledge of multiple programs.

The key steps in the SfM process for different toolchains.

Software	Low density pointcloud	High density pointcloud	Georeferenced mesh and texture
Bundler & PMVS2	Using one of a number of available scripts (eg. ArcheOS) to point at a set of photographs, the bundler pipeline returns a text file containing camera locations, parameters and a low density collection of points.	A second script uses the bundler output to remove distortions from the photographs and launch the PMVS2 process. The result is one or more .ply files containing a high density pointcloud	Further processing must be done in other software such as CloudCompare (to georeference points) and Meshlab (to create and texture meshes)
VisualSFM & PMVS2	VisualSFM has a GUI that allows a set of photographs to be loaded, and takes the user through a series of steps to match points and find camera positions. The cameras and low density points are displayed in a 3D viewer.	PMVS2 can be launched from within the VisualSFM GUI. The result is a series of .ply pointcloud files very similar to those produced by Bundler and PMVS2. These can be viewed in the GUI or opened in an external editor such as Meshlab.	Further processing must be done in other software such as CloudCompare (to georeference points) and Meshlab (to create and texture meshes)
Photosynth	A set of photographs is uploaded to the Photosynth website. The result is an online 3D viewer which displays a low density collection of points.	No further processing is currently possible	
123D Catch			A series of photographs are uploaded to the 123D Catch website and the result is a fairly low resolution textured 3D model that can be downloaded. Georeferencing can be taken from points on the images.
Photoscan	A set of photographs is loaded into the software and when processed the camera positions and low density points are displayed in a 3D viewer	The next stage in the process produces a high density pointcloud which can also be displayed in the viewer.	Further processing within Photoscan will produce a model and a high resolution texture. It also has the facility to georeference points.

5. Ground Control Points (Targets)

Ground control points (GCP's) will need to be recorded in the field to link the generated model to the global coordinate system. Ground control points can be recorded on specific points of distinctive features or of targets that are photographed in the field. These points can then be georeferenced after the model is generated. Survey the targets in using a global position system (GPS).

Examples of ground control points:

1. Duct tape: on a rock outcrop, duct tape with a written X
2. Frisbees with X: the center of an X on a Frisbee
3. Reflective lidar targets

6. Field Workflow (and field prep)

Based on personal experience and Johnson et al. (2014).

Before leaving for the field

1. Decide on a platform that works best for the data you would like to collect. In selecting a platform, be sure to consider your need for power (to charge batteries) or refill helium (if using a balloon).
2. Select a camera – remember to use the section above to guide your choice, and ensure that the camera has GPS tagging or can integrate GPS tagging.
3. Select what your targets will be – make sure they will be visible given the terrain you are surveying and make sense with what you're interested in – if you are mapping the terrain, a flat target (Frisbee, etc) makes more sense but if you're photographing an outcrop, a target that attaches to a tripod or a target that attaches to the outcrop makes more sense.
4. Have an SD card for your camera that will hold a large number of photos and a way to back up the data.
5. Field supply list: platform, camera, camera mount, extra charging supplies, and GPS (to survey in the targets)

At the field site

1. Make a plan: what is the furthest extent you would like dense photographs? Where should the targets go to not obstruct features of interest? Is everything of interest visible? If you are mapping an outcrop, you may want to clean it beforehand. If using a UAV, how much flight time do you have and how should that influence your survey design? What time of day will give you the best lighting to photograph this area? Does that place a limitation on the time you have available – and as a result, the area to survey?
2. Setup and survey in the targets as ground control points.
3. Take photographs. Remember the important parts of taking SfM photos: overlap and changing position. Do not stand and take photographs in a circle around yourself, for example. Move the camera locations for best results. If using an aerial platform, set a time (5-10 s) that makes sense with the speed you've moving the platform and the number of photographs you would like to take/the spacing of those photographs.

4. At the end of the field day (especially if surveying the same location again, or if you have extra days to fix problems) generate an initial model to ensure the photographs you took captured the feature you are interested in (see next section).

7. Agisoft Photoscan Pro

One consideration with Agisoft is the hardware used. Below are the recommendations from Agisoft as of July 2015.

Basic configuration

up to 32 GB RAM

CPU: Quad-core Intel Core i7 CPU, Socket LGA 1155 (Sandy Bridge, Ivy Bridge or Haswell)

Motherboard: Any LGA 1155 model with 4 DDR3 slots and at least 1 PCI Express x16 slot

RAM: DDR3-1600, 4 x 4 GB (16 GB total) or 4 x 8 GB (32 GB total)

GPU: Nvidia GeForce GTX 780 or GeForce GTX 980 (optional)

Advanced configuration

up to 64 GB RAM

CPU: Six-core Intel Core i7 CPU, Socket LGA 2011 (Sandy Bridge-E)

Motherboard: Any LGA 2011 model with 8 DDR3 slots and at least 1 PCI Express x16 slot

RAM: DDR3-1600, 8 x 4 GB (32 GB total) or 8 x 8 GB (64 GB total)

GPU: Nvidia GeForce GTX 780 Ti, GeForce GTX 980 or GeForce GTX TITAN X

Extreme configuration

more than 64 GB RAM

For processing of extremely large data sets a dual socket **Intel Xeon Workstation** can be used.

For resources on memory:
http://www.agisoft.com/pdf/tips_and_tricks/PhotoScan_Memory_Requirements.pdf

For a comparison of differing CPU/GPU's: http://www.anandtech.com/show/7648/gigabyte-brix-pro/3

Based on personal experience

Though many SfM software packages exist, this is my preferred workflow for working in Agisoft Photoscan Pro. If you would rather use open source software, see: James and Robson, (2012), Westoby et al. (2012), Fonstad et al. (2013), and Green et al. (2014). I would also suggest looking into the blog http://archaeologysfm.blogspot.co.uk/, in which Susie Green (of Green et al., 2014) includes additional details, as well as a link to her master's thesis on the topic.

Field Workflow for Agisoft:

1. Take a cursory glance through your photographs to ensure that none are outright blurry. This will help with the texture overlay later and get rid of errors in the generated point cloud.
2. Load photos, as well as associated GPS information (known as "camera position"). The details of how to do this are in the Agisoft document here: http://www.agisoft.com/index.php?id=31
3. Go to the Workflow menu and select "Batch Process." This way you can set the model up to run without having to constantly check whether a step has finished.

4. Hit the "add" button and select Align Photos. Keep accuracy at high – this helps with later steps. As you will have camera locations from GPS, selecting high does not take significantly longer. Choose "reference" for the pair preselection option as you have camera locations. Hit OK.
5. Now add the next step: build mesh. Choosing to this before you make the dense cloud makes the processing faster. Surface type should be height field if a planar surface and arbitrary if something like a building. Source data from the sparse cloud. Choose the medium or high option for polygon count. Hit OK.
6. Add the next step: select the high or medium option, then hit OK.
7. The last step is adding texture. Select arbitrary if working on something like a building; adaptive orthophoto if there is a flat part and a vertical part; orthophoto if it is flat, spherical if it is spherical. Blend a mosaic for slightly higher quality texture. Blending should be modified based on your results, but mosaic works best for a quick model. Hit OK.
8. Now run the batch process: using these parameters, it should only take a few hours (if you have less than 200 photos) to check that you photographed the area well.

High Resolution Workflow for Agisoft:

This basically follows the same steps. However, you should always select high when it is an option, and you should make the dense point cloud before the mesh and then source the mesh from the dense cloud.

8. Photo acquisition considerations

Based on personal experience and Raugust and Olsen (2013).

1. When in the field, consider taking photographs at multiple distances. If using a balloon or UAV, collect a flight path at a single height and then either increase or decrease the height depending on whether you would like more surrounding context or more detail
2. Always photograph a larger region than you anticipate needing. The edges of the area you are surveying will have a lower photo density, so ensure that these are not areas that apply to your research.
3. OVERLAP IS KEY. Overlap your images as much as possible. Different software programs have different guidelines, but it is essential all portions of the area of interest are covered by multiple photographs. Less than 70% overlap will affect the interpreted scene, while more than 90% overlap may significantly increase processing time.
4. When considering overlap, keep in mind the goals of your project. Do you want decimeter resolution? Centimeter resolution? If you are interested in lower resolution models, high overlap percentages are unnecessary.
5. Lighting the photos well is important. If it is too dark, features (especially texture) do not stand out, but if is too bright, these features are washed out. Ideally, if working on smaller scale features, take photographs when the feature is in shadow or lower lighting but the sun is still out.
6. Georeferencing targets rather than geologic features is best. If necessary, use parts of the outcrop as ground control points, but some features may be less visible in the model and georeferencing is essential for scale.
7. Surface texture is essential. If there is little variation on the surface, SfM is difficult. Problematic materials: glass, mirror, very smooth dirt, painted surfaces, surfaces that do not vary in color or texture (ie a box), many other manmade materials, snow. PRACTICE photographing a similar feature beforehand to ensure that the texture will be represented.
8. Break areas into blocks if there are a large number of photographs (greater than 250) and processing is extremely slow. This is a way to speed up processing time, and with georeferenced data, all blocks can be tied together later.
9. To work with data while in the field, make lower resolution models that require less processing time and computer power to check that the model covers all of your area of interest. In addition, if using Agisoft Photoscan Pro, generating the mesh prior to generating the dense point cloud (after the photos are aligned) will be a quicker way to stitch the photographs together. This method is less reliable – sometimes models generated are just black blobs rather than textured accurate representations – but is useful in a field setting when you're prepping for the next day of work.

9. References

Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L., and Carbonneau, P.E., 2013, Topographic structure from motion: a new development in photogrammetric measurement: Earth Surface Processes and Landforms, v. 38, p. 421-430, doi: 10.1002/esp.3366

Green, S., 2012, Structure from motion as a tool for arachaeology, MSc thesis, University College London, 128 p.

Green, S., Bevan, A., and Shapland, M., 2014, A comparative assessment of structure from motion methods for archaeological research: Journal for Archaeological Science, v. 46, p. 173-181, doi: 10.1016/j.jas.2014.02.030

James, M.R., and Robson, S., 2012, Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application: Journal of Geophysical Research, v. 117, F03017, doi: 10.1029/2011JF002289

Johnson, K., Nissen, E., Saripalli, S., Arrowsmith, J.R., McGarey, P., Scharer, K., Williams, P., and Blisniuk, K., 2014, Rapid mapping of ultrafine fault zone topography with structure from motion: Geosphere, v. 10, no. 5, p. 969-986, doi: 10.11130/GES01017.1

Raugust, J.D., and Olsen, M.J., 2013, Emerging technology: structure from motion: LiDAR Magazine, v. 3, no. 6, 5 p.

Snavely, N., Seitz, S.N., and Szeliski, R., 2008, Modeling the world from internet photo collections: International Journal of Computer Vision, v. 80, p. 189-210, doi: 10.1007/ss112263-007-0107-3

Ullman, S., 1979, The interpretation of structure from motion: Proceedings of the Royal Society of London B, v. 203, p. 405-426, doi: 10.1098/rspb.1979.0006

Verhoeven, G., 2011, Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan: Arcchaeological Prospection, v. 18, p. 67-73, doi: 10.1002/arp.399

Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., and Reynolds, J.M., 2012, Structure from motion photogrammetry: a low-cost, effective tool for geoscience applications: Geomorphology, v. 179, p. 300-314, doi:10.1016/j.geomorph.2012.08.021