METHODS & TECHNIQUES

A protocol and calibration method for accurate multi-camera field videography

Diane H. Theriault1,*, Nathan W. Fuller2, Brandon E. Jackson3, Evan Bluhm3, Dennis Evangelista3, Zheng Wu1, Margrit Betke1 and Tyson L. Hedrick3,*

ABSTRACT
Stereo videography is a powerful technique for quantifying the kinematics and behavior of animals, but it can be challenging to use in an outdoor field setting. We here present a workflow and associated software for performing calibration of cameras placed in a field setting and estimating the accuracy of the resulting stereoscopic reconstructions. We demonstrate the workflow through example stereoscopic reconstructions of bat and bird flight. We provide software tools for planning experiments and processing the resulting calibrations that other researchers may use to calibrate their own cameras. Our field protocol can be deployed in a single afternoon, requiring only short video clips of light, portable calibration objects.

KEY WORDS: Videography, Stereography, Bundle adjustment, Photogrammetry, Direct linear transformation, Kinematics, Three-dimensional, Tracking

INTRODUCTION
Researchers have studied many diverse topics, including animal behavior and plant seed dispersal, by recording and quantifying the movement of animals and plants using video (e.g. Hayashi et al., 2009; Olaniran et al., 2013; Bhandiwad et al., 2013). For example, in 2012, 70 papers published in *The Journal of Experimental Biology* reported using video to measure kinematics, representing 11% of the papers published in the journal that year. Many of these studies used two or more cameras to measure the three-dimensional (3D) locations of points of interest in the scene. In order to use image observations from multiple cameras to reconstruct 3D world positions via triangulation, the relative position and orientation of the cameras (extrinsic parameters) and their focal lengths and principal points (intrinsic parameters) must be given. The process by which these parameters are estimated is known as ‘camera calibration’ and typically involves matching points on a calibration object across camera views.

In the recent biology literature, the most commonly mentioned method for calibrating cameras is direct linear transformation (DLT) (Abdel-Aziz and Karara, 1971), with some researchers also using a camera calibration toolbox for MATLAB (Bouguet, 1999). When using DLT, it is important to obtain calibration points throughout the volume of interest, otherwise reconstruction accuracy may be reduced (Hedrick, 2008). Previous authors who have used DLT in a field setting have constructed a large physical calibration object at the field site (Clark, 2009; Munk, 2011), limiting the size of the calibration volume. Others have carefully measured the extrinsic parameters by hand (Cavagna et al., 2008), relying on a semi-permanent placement of their cameras in a sheltered location. We propose a different calibration approach that is particularly useful for field settings where the volume of interest may be tens of thousands of cubic meters or where cameras cannot be left in place for multiple sessions. Our approach uses the sparse bundle adjustment (SBA) calibration algorithm (Lourakis and Argyros, 2009), which minimizes the difference between the observed and ideal locations of the calibration points in each camera view. Bundle adjustment has been chosen by biologists for its 3D reconstruction accuracy (Walker et al., 2009), but lacks wide use because of the absence of easily accessible software implementations that are directly applicable to analysis of field data.

The purpose of this paper is to describe a successful three-step workflow and provide software tools (supplementary material Fig. S1) that will allow researchers to perform stereo videography easily and accurately in a field setting. The workflow provides guidance on (1) pre-experiment planning of camera placement to meet observational objectives, (2) an in-field calibration protocol and (3) post-experiment camera calibration. Camera placement planning is supported by our easyCamera software tool in MATLAB; camera calibration is supported by easyWand, which is a graphical user interface to the easySBA routine, or a command-line Python package that calls the SBA routine provided by Lourakis and Argyros (Lourakis and Argyros, 2009). Our software is licensed under GNU public license version 3. Usage instructions are provided with the software download. Sample data and video tutorials are also available online. We here present two cases using the proposed workflow to calibrate cameras and provide accurate estimates of the three-dimensional flight paths of cliff swallows (*Petrochelidon pyrrhonota*) and Brazilian free-tailed bats (*Tadarida brasiliensis*). Although designed for use in the field, the methods outlined here may also be useful in laboratory situations where introducing a calibration object into the volume to be measured is difficult or infeasible.

RESULTS AND DISCUSSION
When developing our stereo videography workflow and software, we focused on accuracy and ease. It is important to estimate the level of calibration and reconstruction accuracy when the goal of 3D videography is to quantify the kinematics of airborne animals and facilitate the study of their behavior. Any uncertainty in the estimation of their 3D position affects the uncertainty in derived calculations such as velocity and acceleration, which may be of direct biological interest. We here first describe the results of our experiments that yielded stereoscopic reconstructions of bat and bird flight and then discuss their accuracy and the ease in obtaining them.
Stereoscopic reconstructions of bat field flight

We recorded video of Brazilian free-tailed bats during their evening emergence from Davis Blowout Cave in Blanco County, Texas, USA (Fig. 1), and used the proposed stereo videography workflow to support the estimation of the flight paths of 28 bats flying through a 1400 m³ volume during a 1 s interval with a 7.8 cm root mean squared (RMS) uncertainty in their 3D positions. We used easyWand and easySBA to estimate the extrinsic and intrinsic camera parameters, using, from each view, 226 points manually digitized from the ends of a 1.56 m calibration wand, 2010 points manually digitized from hot packs thrown in the air, and 7135 points on the bats flying through the volume of interest, identified using automated methods (Wu et al., 2009). Manufacturer-provided values served as initial estimates of the intrinsic parameters. The calibrated space was aligned to gravity by calculating the acceleration of a rock tossed through the volume of interest. The standard deviation of the estimated wand length divided by its mean was 0.0056 (0.56%); the respective RMS reprojection errors were 1.16, 1.58 and 1.17 pixels for the three cameras. The swallow observation protocol was approved by the University of North Carolina Institutional Animal Care and Use Committee.

Accuracy of results and ease of experimental setup

As our results demonstrate, the proposed field videography workflow and software yield accurate calibration of multi-camera systems and enable accurate reconstruction of 3D flight paths of bats and birds in field settings. The values of our two measures of calibration inaccuracy were sufficiently small to indicate accurate calibrations without any errors in determining corresponding calibration points.

We balanced our two conflicting observational objectives of recording sufficiently long flight paths in a large volume of interest and recording the animals so that they appeared sufficiently large in all camera views. The resulting camera placement yielded a level of uncertainty in estimated 3D locations of the animals that was less than the length of a bat and half the length of a bird. A different camera placement may have resulted in a different level of accuracy. The Materials and methods section describes how researchers can determine this level in pre-experiment planning using the proposed easyCamera software tool. In post-experiment processing, easyWand can be used to estimate the accuracy of the calibration.

Our recordings were part of extensive multi-day experiments. Camera placement and calibration recordings required 45 min at the beginning of each daily recording bout. This short setup time may be important when the study organism or group is on the move and must be followed, or when site location, daily weather patterns, tides or safety considerations dictate.

We suggest that the methods and implementations provided here will substantially aid biologists seeking to make quantitative measures of animal movements and behavior in field settings, as they have done for our own work on bat and bird flight. While previous studies have gathered similar data, they required heroic attempts at calibration frame construction in the field or carefully controlled field environments. We believe that the ease of setup and...
accuracy of calibration afforded by our methods opens up a wide range of previously unachievable studies and we plan to continue refining the publicly available software implementations to fit a variety of needs.

MATERIALS AND METHODS
The proposed three-step workflow for performing stereo videography provides guidance on field-experiment planning, capture and post-processing. During planning, appropriate camera equipment is chosen and the placement of the cameras is determined so that the captured imagery will satisfy observational requirements. In the field, a protocol is followed for moving calibration objects through the volume of interest. During post-processing, the image locations of the calibration points are digitized in each camera view. Then, easyWand uses corresponding points to estimate the relative position and orientation of the cameras.

Fig. 2. The flight paths of 12 cliff swallows during a 2.3 s interval are shown. Three-dimensional flight trajectories (A) were estimated from video that was captured using three synchronized, high-speed visible light cameras (B). At an observation distance of ~20 m, the birds, which are ~13 cm long, were imaged at an average length of 18 pixels. The baseline distance between the outermost cameras was ~11 m. The RMS reconstruction uncertainty for the 2796 estimated 3D points shown was 5.9 cm, less than half the length of a bird. The 3D trajectories were smoothed using a smoothing spline based on this uncertainty.

Fig. 3. Reconstruction uncertainty due to quantization effects is shown for six hypothetical camera configurations. The cameras were simulated to have a pixel width of 18 μm and a field-of-view angle of 40.5 deg, and be positioned at a fixed height Z and aimed at a common, equidistant fixation point F=(0,0,Z). Horizontal cuts of the 3D view frustums of the cameras at height Z and lines at D_{max}=20 are shown from above. Placing the cameras further apart reduces reconstruction uncertainty (A versus B). If the cameras are placed too far apart (C), however, the view volume is ‘closed’, and there are unobservable regions of space where the cameras will be looking past each other. If the distance between the outermost cameras is held constant, adding additional cameras may not decrease the uncertainty due to image quantization in the common observable region (D versus E). When the image planes of the cameras are parallel (F), the common view volume is smaller and further away from the cameras than in the other configurations. Note that these two-dimensional cuts of the 3D view frustums are at the level and elevation angle of the cameras; cuts at a different level or angle would show slightly greater reconstruction uncertainty but similar trends.

Experiment planning
Camera systems
When selecting a multi-camera system, scientists should consider whether the frame rate, spatial resolution, field-of-view and synchronization ability of the cameras are appropriate for the size and speed of the study organisms. Camera synchronization, in particular, is a requirement for successful stereoscopy. In our field work, we used hardware synchronization to ensure accurate temporal alignment of frames across cameras. Multi-camera systems without precise frame synchronization could be calibrated using these methods if wand and animal pixel motion per frame is small, a large number of calibration points are used and some motionless background points are visible in all cameras.

For capturing video of Brazilian free-tailed bats, we used three thermal infrared cameras (FLIR SC8000, FLIR Systems, Inc., Wilsonville, OR, USA) with variable-focus 25 mm lenses and a pixel width of 18 μm,
providing a 40.5 deg field of view. The 14 bit grayscale–1 video has a frame size of 1024×1024 pixels and frame rate of 131.5 Hz. For capturing video of cliff swallows, we used three high-speed cameras (N5r, Integrated Design Tools, Inc., Tallahassee, FL, USA) with 20 mm lenses (AF NIKKOR 20 mm f/2.8D, Nikon Inc., Melville, NY, USA), recording 10 bit grayscale video with a frame size of 2336×1728 pixels at 100 Hz.

**Camera placement**

When designing camera placement, scientists should consider their observational objectives, the amount of 3D reconstruction uncertainty they may tolerate, and potential additional requirements introduced by the manual or automatic post-experiment video analysis.

Two observational objectives that commonly conflict are the size of the observation volumes and the uncertainty due to quantization within the pixels. Ensuring that the angle between the optical axes is not larger than the field-of-view angle of the cameras leads to ‘open’ intersection volumes that extend infinitely far away from the cameras (all examples in Fig. 3 except C), which is desirable because it facilitates recording even if the animals appear in an unexpected location. The level of uncertainty increases with the distance from the cameras because the volume of the intersection of the pixel frustums also increases with this distance.

In addition to the reconstruction uncertainty created by quantization, we also consider the reconstruction uncertainty arising from the difficulty in identifying the location of an animal in an image (Fig. 4). The location of an animal is often thought of as a single point, e.g. at the center of its body. Localization accuracy of this ill-defined point depends on the resolution of the animal in the image (Fig. 4B). To estimate uncertainty in localization, we included a stochastic element to the simulation procedure described above by adding noise to the two-dimensional projections before quantization (supplementary material Fig. S2). Over 100 trials, the RMS distance between the original and reconstructed scene points gives an estimate of the reconstruction uncertainty at the original point (Fig. 4D). The camera placements we selected for our bat field experiments (Fig. 1A) were similar to the configuration shown in Fig. 3B. With our simulation, we were able to determine, prior to any field work, that the levels of uncertainty due to quantization and localization issues would be acceptable for use (supplementary material Figs S2–4).

Reconstruction error occurs when image locations corresponding to different animals are mistakenly used to reconstruct 3D positions. These ‘data association’ errors are commonly made by automated tracking methods, especially if the animals appear similar and small in the images. Camera selection and placement can reduce the potential occurrence of data association errors by imposing appropriate geometric constraints on the triangulation (Fig. 5). We recommend use of three or more cameras and a non-collinear camera placement that ensures that the image planes are not parallel (avoiding the configuration in Fig. 3F).

**Protocol for field experiment**

The protocol we recommend for field work includes two phases. In the first phase, prior to any camera setup and recordings, a preliminary plan for the location of the recording space and placement of the cameras is made. The easyCamera software is then used to estimate the uncertainty in localizing the study organism in this space, and adjustments to the plan can be made by experimenting with other hypothetical camera configurations. In the second phase, the actual camera setup in the field can be done easily because no field measurements of camera pose or distances to the animals of study are needed. The only measurements required are references for scene scale and orientation. In our experiments, the known length of a calibration wand (supplementary material Figs S5, S6) moved through the scene provided a scale reference, and gravitational acceleration, estimated from the ballistic trajectories of thrown objects, provided a reference for scene orientation.

**Fig. 4. Reconstruction uncertainty due to quantization and resolution issues is shown.** In a video frame obtained for a flight study (A), the automatically detected locations of the animals may not be at their centers (colored dots in B). When estimating reconstruction uncertainty (C,D), we include this effect by corrupting the image projections of simulated world points, generated throughout the whole space, with Gaussian noise where the standard deviation is one-sixth of the calculated apparent size of an animal at that location (circles in B). When estimating the reconstruction uncertainty, including image location ambiguity (D) increases the estimated uncertainty more than threefold over image quantization alone (C) (note the change in color scale).
Our calibration method generally produces more accurate results the more sets of corresponding image points it uses (supplementary material Table S1, column 2). Thus, we used recordings of the animals, digitized automatically using a preliminary wand-only calibration, to augment the number of calibration points and volume encompassed by them. This augmentation is a feature of our SBA-based calibration pipeline not possible with calibration-frame-based DLT methods. See supplementary material Table S1 for an exploration of the effects of using animal points in the calibration. We typically recorded our study videos of bats and birds after obtaining videos of calibration objects, but this order can be reversed.

Post-experiment camera calibration

Our easyWand calibration software bundles a modular pipeline of algorithms that can be used to estimate the relative positions and orientations of the cameras and their intrinsic parameters. The first, most time-consuming step of the calibration procedure is to manually or automatically digitize the image locations of objects recorded in all views. In our post-experiment analysis, we identified thousands of sets of matching image points.

Using the focal lengths and principal points obtained directly from the camera, we compute preliminary estimates of the camera pose and 3D positions of the calibration objects. Our software then applies the SBA algorithm (Lourakis and Argyros, 2009) to obtain refined estimates for all positions of the calibration points. A collinear camera placement (Fig. 3F) would result in parallel epipolar lines, making the matching of corresponding image points difficult.

Acknowledgments

We wish to thank Brian Borucki, Ashley Banks, Ann Froshauer and Kimmi Swift for assisting with the bat recordings, and Nick Deluga for assisting with the swallow recordings. We also thank Dewayne Davis and David Bamberger for property access, and the Department of Texas Parks and Wildlife for permitting assistance. Thanks also to two referees for providing useful feedback.

Competing interests

The authors declare no competing financial interests.

Author contributions


Funding

This work was partially funded by the Office of Naval Research [N0001410100952 to M.B. and T.H.], the National Science Foundation [0910908 and 0855065 to M.B. and 1253276 to T.H.] and the Air Force Office of Scientific Research [FA9550-07-1-0540 to M.B.].

Supplementary material

Supplementary material available online at http://jeb.biologists.org/lookup/suppl/doi:10.1242/jeb.100529/-/DC1

References


Figure S1:
Software packages for easyCamera, easyWand, and easySBA and documentation can be downloaded from the OpenBU repository at http://hdl.handle.net/2144/8456. The Python SBA source code is also available at https://bitbucket.org/devangel77b/python-sba and the Python PIP stable release at https://pypi.python.org/pypi/sba/1.6.0
Figure S2: Reconstruction uncertainty due to quantization and localization ambiguity effects is shown for six hypothetical camera configurations. The cameras were simulated to have a pixel width of 18 µm and field-of-view angle 40.5° and be positioned at a fixed height Z and aimed at a common, equidistant fixation point \( F = (0,0,Z) \). Horizontal cuts of the 3D view cones of the cameras at height Z and lines at \( D_{\text{max}} = 20 \) m are shown from above. Image localization ambiguity was modeled by adding Gaussian noise to the image points before image quantization. Noise added to the image points corresponding to the 3D points had a standard deviation that was one-sixth of the calculated apparent size of a 10 cm animal, given the distance between each camera and corresponding 3D point. When the 3D points are close to the camera, the image localization ambiguity contributes more to the uncertainty, whereas when the 3D points are further away, the image quantization effects contribute more.
Figure S3: 3D flight trajectories of 28 Brazilian Free-tailed Bats during a 1-s interval are shown in the context of the spatially-varying reconstruction uncertainty arising due to both image quantization and image localization ambiguity from an oblique view (A) and from the top (B). The tracks are shown from the point of view of the cameras (C) and from the side (D). The observation distance between cameras and bats was approximately 10 m (B, D), chosen so that the nose-to-tail span of a bat in an image was at least 10 pixels. The baseline distance between the outermost cameras was approximately 6 m, chosen so that the expected uncertainty in reconstructed 3D positions at the observation distance due to image quantization and image localization ambiguity was less than 10 cm, the length of a bat. The RMS reconstruction uncertainty for the 1,656 estimated 3D positions shown was 7.8 cm.
Figure S4: The flight paths of 12 Cliff Swallows during a 2.3-s interval are shown in the context of the spatially-varying reconstruction uncertainty arising due to both image quantization and image localization ambiguity from an oblique view (A) and from the top (B). The tracks are shown from the point of view of the cameras (C) and from the side (D). At an observation distance of approximately 20 m (B,D), the birds, which are approximately 13 cm long, were imaged at an average length of 18 pixels. The baseline distance between the outermost cameras was approximately 11 m. The RMS reconstruction uncertainty for the 2,796 estimated 3D points shown was 5.9 cm, less than half the length of a bird.
**Figure S5: Bat field recording wand and site.** The 1.56-\(m\) wand is shown attached to an approximately 3.5-\(m\) pole (A), which was used to wave the wand through the camera scene (B). The wand was equipped with hot and cold packs to provide thermal infrared contrast between points on the wand and the background sky and vegetation. The bat field site is shown with three thermal infrared cameras arrayed for recording and the calibration wand in the foreground (C).

**Figure S6: Bird field recording wand** The 1.0-\(m\) calibration wand (A) used in the swallow recordings was attached to a stout fishing pole, also shown (A), to facilitate tossing the wand through the camera viewing volume and retrieving it afterward. The calibration wand itself was painted white and black to enhance contrast and visibility against different backgrounds. A video frame from a wand calibration toss is shown with the wand visible against the sky and the fishing pole appearing in the lower left corner (B).
<table>
<thead>
<tr>
<th># of bird points</th>
<th>Calibration RMSE</th>
<th>Bird RMSE</th>
<th>Wand score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wand only (0)</td>
<td>0.2653</td>
<td>4.5150</td>
<td>0.6475</td>
</tr>
<tr>
<td>5</td>
<td>0.3800 (0.0533)</td>
<td>2.2002 (0.5969)</td>
<td>0.7007 (0.2645)</td>
</tr>
<tr>
<td>10</td>
<td>0.4580 (0.1011)</td>
<td>1.8825 (0.3664)</td>
<td>0.6926 (0.2771)</td>
</tr>
<tr>
<td>50</td>
<td>0.8403 (0.2044)</td>
<td>1.4389 (0.1207)</td>
<td>0.6146 (0.1627)</td>
</tr>
<tr>
<td>500</td>
<td>1.1477 (0.0808)</td>
<td>1.1021 (0.0404)</td>
<td>0.5418 (0.0466)</td>
</tr>
<tr>
<td>1000</td>
<td>1.1818 (0.0718)</td>
<td>1.0738 (0.0277)</td>
<td>0.5238 (0.0332)</td>
</tr>
<tr>
<td>All (4946)</td>
<td>1.1985</td>
<td>1.0480</td>
<td>0.5235</td>
</tr>
</tbody>
</table>

**Table S1. Effect of including animal points in the calibration set.** Here we examine the effect of including animal data in the calibration procedure by creating a swallow scene calibration using 58 points on the wands only, the wand and a range of randomly selected bird points from 5 to 1000, and the full set of bird points used in the calibration (4946). In cases where a random sample of points was used, 100 samples were drawn for each case and the results are reported as: median (inter-quartile range). The results are assessed by the root mean square reprojection error (RMSE) in pixels for all calibration points, the RMSE for a second set of 4303 bird points not used in calibration and the ratio of the standard deviation of wand length divided by its mean, multiplied by 100, listed as the wand score. Note that the bird points were less precisely digitized than the wand points. This causes the calibration RMSE to rise as additional bird points are added since the RMSE reflects the distance in the image between the reprojected points and the observations; if the observations are less precisely known then the RMSE rises. Incorporation of bird points also causes a negligible rise in the wand score when only a few bird points are used; in all cases the wand score indicates a high quality calibration. On the whole, incorporating bird points in the calibration greatly increases the reconstruction fidelity for other bird points, not used in the calibration, as shown by the middle column. This is because the bird points used in the calibration provide image correspondences in regions not accessible to a ground-launched wand. Generally, the more points that can be used, the smaller the errors become (middle column, last row). Even incorporating 5 bird points in the calibration decreased the RMSE of bird points not used in the calibration by more than one half (middle column, row 2). In all cases we allowed optimization of the focal length and principal point; focal length estimates were always 10% greater than the manufacturer-provided values.