July 24th 2017

Dear JoVE Reviewers,

Thank you for your detailed comments and revisions. The authors of "*Development of new methods for quantifying fish density using underwater stereo-video tools*" have revised the manuscript based on the recommendations of the reviewers. Attached is a copy of the manuscript with track changes, as well as a clean version of the manuscript. We have also attached a folder with supplemental materials. Supplemental screenshots are labeled according to the step numbers in the manuscript. Supplemental R-code and data has been uploaded and referenced on GitHub to help readers understand and perform the bootstrapping procedure. We have included a summary below our actions to reviewer's suggested edits.

In brief, the focus of this manuscript is the description of a MaxN technique for rotating drop camera tools as well as the use of a 95% Z distance metric to define the mean distance of reliable fish measurements. As such, we significantly reduced language in the manuscript that compared our tool directly with other existing survey tools. Similarly, calibration of cameras for 3D photogrametry is well documented in other studies and is not the focus of this present paper. We therefore reduced the description of camera calibration, and instead focus on MaxN and 95% Z Distance.

Changes recommended by the JoVE Scientific Review Editor:

• Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.
We read through the manuscript to check for grammatical and spelling errors.

• Please include at least six keywords/phrases.
Six Keywords and phrases are now included.

• **Protocol Language:** Please ensure that ALL text in the protocol section is written in the imperative tense as if you are telling someone how to do the technique (i.e. "Do this", "Measure that" etc.) Any text that cannot be written in the imperative tense may be added as a "Note", however, notes should be used sparingly and actions should be described in the imperative tense wherever possible.
1) Example NOT in imperative tense: "Some fishes will only be identifiable to higher taxonomic groups"
The protocol was checked for imperative tense at each step.

• **Protocol Detail:** Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. **Please add more details to the following protocol steps.** There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.

Major Edit: The focus of our paper is the description of MaxN for rotating cameras as well as the use of 95% Z-distance values for calculating the mean distance a fish could reliably be measured. We therefore made the choice to limit protocol discussion of camera calibration and instead focus on the intended methods. Calibration and measurements can readily be performed across a variety of software platforms. We made the recommendation of using a 'calibration cube' with references to the table of materials. Again, the choice of calibration method will not affect the MaxN statistic calculated in this paper.

It should be noted that we used specific software for these steps and that complete replication may require referencing both the CAL software manual (80 pages) as well as the EventMeasure manual (110 pages). Neither proprietary manual can adequately be abbreviated in this manuscript. We assume that individuals implementing our protocol will have some basic understanding of their selected software. Nonetheless, we have added specific menu commands for each step. Screenshots of these software steps will be available in the supplemental materials.

1) 1.1: Is it unclear what “cube” is and what it does here (please add a short note after this step to mention this?).
A description of the Calibration Cube was provided as well as a reference to the table of materials.

Please expand the calibration steps. Please mention what button is clicked on in the software to do this, or which menu items need to be selected.
2) 1.1.1: What are the 5 cube orientations?
As per the note above, we have eliminated the description of our specific calibration procedure. The use of 3D photogrametry was not the focus of our paper and has been adequately described in previous research. Menu items and buttons were explicitly referenced.

3) 1.2,1.2.2, 1.2.3, 1.2.4, 1.2.6, 1.2.7, 1.3,1.3.1, 1.3.2, 1.3.3., 1.3.4,1.5 ,1.6, 1.6.1, 1.6.2., 1.6.3., 1.7, and all other software-based steps: Please mention what button is clicked on in the software to do this, or which menu items need to be selected. Please also provide us screenshots to show each software step, these can be added a supplementary files.

4) 2.6: Unclear what is done here.

We have changed the text to try and clarify the step. The accompanying note/example in the protocol is intended to help readers understand the procedure. This step is admittedly tricky to understand, but is important and deserves the long comment and nuanced explanation. For the reviewer: Some fish are only identifiable to the genus - level, and we wanted a count metric that would conservatively avoid over counting fishes. We therefore decided to only count “unknown species’ during rotations where the majority of fishes had in fact been identified to species. This made it more likely that the ‘unknown species’ marked as 2D points had not been previously identified.

We believe that this step will be easier to understand with a visual aid.

• Protocol Numbering: Please adjust the numbering of your protocol section to follow JoVE’s instructions for authors, 1. should be followed by 1.1. and then 1.1.1. if necessary and all steps should be lined up at the left margin with no indentations. There must also be a one-line space between each protocol step.
Care was taken to ensure that all protocol numbering matched the required formatting.

• Protocol Highlight: After you have made all of the recommended changes to your protocol (listed above), please re-evaluate the length of your protocol section. There is a 10-page limit for the protocol text, and a 3- page limit for filmable content. If your protocol is longer than 3 pages, please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps. Please see JoVE's instructions for authors for more clarification. Remember that the non-highlighted protocol steps will remain in the manuscript and therefore will still be available to the reader.

Highlighted text meets the 3 page maximum requirement.

o Underwater step: Can these be filmed in mock? Our videographers cannot film underwater.
As per the notes above, we have reduced the text describing calibration since this was not the focus of the paper. If a segment of calibration were to be included in the video, then this step could easily be filmed in mock. Another possibility is for us to supply video we took ourselves while calibrating underwater. We also thought that some of our GoPro footage of actual fish surveys could be included, both to show the cameras rotating around and also to make the overall video more interesting.

o The highlighting must include all relevant details that are required to perform the step. For example, if step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be included in the highlighting.
o The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.
o Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.
o Notes cannot be filmed and should be excluded from highlighting.
Effort was made to conform to highlight formatting outlined above

o Please bear in mind that software steps without a graphical user interface/calculations (e.g. section 3) cannot be filmed.
We have modified section 3 according to suggestions outlined above. We have included a screenshot in the supplemental material showing this step performed with a GUI.

• Discussion: JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form: 1) modifications and troubleshooting, 2) limitations of the technique, 3)

significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.

We believe that our discussion follows the outlined format well.

• Figures::
1) Fig 3: The text in the inset windows cannot be read.
Figure 3 has been updated to be legible

• Figure/Table Legends:
1) Fig2: Please mention cube dimensions.
Cube dimensions have been added to Fig 2 legend. Note that calibration is no longer a focus of this paper, and so the figure legend describes 'an example calibration'.
2) Fig 4: Please define the red markers.
No red lines were found in Fig 4; however, 'Red' was added to the description of vertical bars in Figure 5

• Commercial Language: JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Examples of commercial sounding language in your manuscript are SeaGIS, Inc, LTD™, Deep Sea Power and Light (DSPL), etc.
1) Please use MS Word's find function (Ctrl+F), to locate and replace all commercial sounding language in your manuscript with generic names that are not company-specific. All commercial products should be sufficiently referenced in the table of materials/reagents. You may use the generic term followed by "(see table of materials)" to draw the readers' attention to specific commercial names.

We removed all commercial sounding language from the text of the manuscript, and instead only reference company names in the materials table.

• Please define all abbreviations at first use.
Abbreviations were defined at first use.
• Please use standard abbreviations and symbols for SI Units such as µL, mL, L, etc., and abbreviations for non-SI units such as h, min, s for time units. Please use a single space between the numerical value and unit.
Unit abbreviations conform to these requirements.

• If your figures and tables are original and not published previously or you have already obtained figure permissions, please ignore this comment. If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."

Our figures and tables are original.

Comments from Peer-Reviewers:

Reviewer #1:

Manuscript Summary:

I have reviewed the paper titled "Development of new methods for quantifying fish densities using underwater stereo-video tools". The manuscript describes a modification to the traditional MaxN approach to enumerating fish abundances for a rotating stereo-video lander. The primary objective of this study was to determine species-specific detectability to estimate densities using a rotating stereo-video camera system. This research is well thought out and clearly presented. I believe that this manuscript represents an important contribution on survey methods for quantifying densities from stereo-video systems. However, there are several clarifications that with regards to the process for determining that an appropriate calibration was created that need to be added to the protocols. Addressing these suggestions will make the manuscript more re-producible. Therefore, I recommend that the manuscript be published after undergoing minor revisions to address the points listed in this review. Comments by section are presented below.

*Title:

I believe that the title accurately describes the scope of the work.

*Keywords:

I would suggest adding in the following: Underwater visual census, groundfish, rotating drop-camera system, and removing stereo video (as it is already in the title).

Underwater visual census, groundfish, rotating drop-camera system were added to the keywords and Stereo Video was removed

*Short Abstract:

The structure and content of the short abstract is suitable.

*Long Abstract:

The long abstract sufficiently describes the background and justification for the application of this modified protocol for future research and is appropriate for this methods article.
*Suggested edits are as follows:
Lines (88-90): Suggest adding in more information about the preliminary studies - what were the analyses used to determine that eight rotations was ideal for this system. Since the author recommends that this metric be determined for a given study site it would help increase reproducibility if more information in conducting these preliminary analyses were added.
A generic description of the determination of the 8 minute soak time was provided. Other reviewers cautioned against inclusion of specific results in the introduction so specific numbers were avoided.

*Protocol:
The steps listed in the procedure overall are clearly explained. All the critical steps are highlighted.

However, the protocol should be clearer in specifying that terms, cited file extensions, and software functions are SeaGis dependent and may not be synonymous with other software applications.
While we wholeheartedly agree that specific software used should be cited within the text, JoVE requires the exclusion of any commercial language from the manuscript. SeaGIS is mentioned in the materials table. The specific steps outlined are only possible in SeaGIS; however the overall approach to the MaxN approach we have outlined is transferable to any stereo software platform.

Also, the syncing steps would be useful to include earlier in the procedure.
Syncing the video is not possible before this point in the procedure.

A possible addition to this protocol maybe a sentence on checking for camera frame drift to ensure that the cameras remained synced.
A sentence was added to our note to recommend checking for camera frame drift during analysis

*Suggested edits are as follows:
Lines (144-145, 149,154): You should clarify if these file extensions are specific to the SeaGIS software and may not be synonymous to all other systems.
See comment above: JoVE will not allow this.

Lines (151-152): Specify the camera specific parameter estimates contained in this file that are needed for calibration.
Example parameters are now included in the text. See supplemental material for a screengrab of all parameters included in the calibration file

Line 154: Specify the cube specific parameter estimates contained in this file for those not using SeaGIS software.

This step in the protocol has been removed; however the cube parameter file contains the precisely calibrated distances and dimensions between calibration cube points.

Line 158: Provide information on how to sync videos. Suggest including information on the process for syncing cameras (i.e. light flash, clapper board, etc.) between 1.1 and 1.1.1 and state how often this process should be conducted during field sampling to identify if there is camera drift.
We have recommended the use of a time UTC time stamp to sync videos and to periodically check that synchronization has been maintained. We also mention other techniques used to calibrate.

Line (160-162): Specify if there is an order needed to when selecting the reference points.
*Overall, in the calibrations section the author needs to specify the acceptable and not acceptable limits for metrics during this calibration. For example, what are the metrics that would make a calibration not meet standards for moving forward (i.e. # of exclusions, relative positioning, etc.).
We have narrowed the scope of our protocol section to focus on the measurements of fishes and the calculation of MaxN and 95% Z-distance values. See previous comments for further details. We did however note that measurement error after calibration should be within 2% of an object's known length. The choice of software and calibration technique can be determined by the researcher and will not affect the strategy employed in the calculations of MaxN and 95% z-distance.

Lines (214-15, 219, 221-222): Specify that these processes are SeaGIS specific and that other software packages may not have these specifications and tabs.
See above comment: JoVE will not allow this

Lines (224-225): This is the first time the syncing process is mentioned see comment above about including this earlier in the protocol.

As addressed above, syncing is now addressed in more detail the first time it is required in the protocol.

*Representative Results:
The results from this study seem reasonable and are useful to readers operating stationary stereo-video drop cameras.

*Suggested edits are as follows:
Lines (379-383): Consider adding a reference to support whether these results are consistent with the known habitat associations of these species.
Reference added citing known habitat associations for both example species

*Discussion:
The discussion details the use and applicability of this approach by highlighting the strengths, limitations, and caveats of both the stationary and rotating lander.

*Suggested edits are as follows:
Line 590: This citation is not cited in the paper.
Citations were updated to excluded references not used

Reviewer #2:
Manuscript Summary:
Nice work. I think this is an interesting paper that warrants publication. I checked "major revision" only because I think you need to think hard about the way in which the work is contextualized. Papers that compare approaches are fraught with peril, because there are so many potential explanations for why one approach might be better under one set of circumstances, but not others. For that reason, I recommend that rather than juxtaposing this approach (lander with stereo video) as the solution to problems inherent in other tools/approaches (ROVs/subs using lasers), I suggest that you describe the universe of challenges in collecting data from deep water, perhaps describing how some other tools address some of those challenges, and then introduce your approach as a means to addressing selected challenges without directly juxtaposing it to other tools. I think that could make a really useful contribution. I'm happy to talk further about this if it would be helpful. Good luck!

Lines 54-55 Should read "…these known habitat associations…"
Line 56 " …that are better suited to sampling low-relief sedimentary environments" This provides a better juxtaposition to the prior sentence.

Lines 59 -65 Since you assert above that relief and substrate is the primary issue, I think it will be important to better flesh out that particular issue with respect to these traditional gears, particularly trawls. For instance, it is not simply that trawls may underestimate counts, rather trawl surveys will intentionally avoid the types of high relief habitat you are discussing due to the way the gear is deployed and fished. Explain that. This is an important point that sets the stage for your approaches' utility. Next I am not convinced by your critiques of acoustic surveys and hook and line methods. While both may have the limitations you describe, they can be dealt with by someone who understands how to use the tools. They are not necessarily problems unique to those tools. If your goal is to cast landers in a better light than these other tools, I would make a different case..

Lines 68-71 Similar to my comment above, while it is true that most ROV surveys to-date have relied on paired sizing lasers rather than stereo video, it is not true that this limitation is inherent to ROVs or subs (which this statement implies). Indeed prior to your tool, the same can be said of landers (e.g., ODFW's). If you want to flesh out differences between transect-based visual tools here to set the stage for a discussion of your lander, that could be useful. For instance, for a given day at sea, an ROV/Sub can survey x amount of seafloor, distributed over a total area of y, given the logistics of getting the vehicle in and out of the water and down to the seafloor. How do lander stats compare? Smaller area surveyed, but over a wider area? This would be an interesting way to compare while avoiding the economic comparison I mention below.

Also, I recommend the use of human-occupied submersibles (HOVs) rather than manned submersibles.

Lines 72-73 Consider removing value judgments about the "expense" of particular tools. Expense is relative. The widely variable types of ROVs and associated vessels make this type of statement difficult to support without a more dedicated evaluation of cost associated with the use of any particular tool.

Lines 75-76 same comment as above. This critique is too broad. Does "relative ease" refer to smaller vessels and fewer surface support staff? Either explain that, or don't make that claim. I can put a small ROV on a small boat, using student support, and the cost would be minimal. Of course, conditions would be a huge issue, and perhaps your lander would be preferable for certain sea states or certain water depths, but you don't talk about that. Also, what about the cost of stereo video processing software? ROV/sub video can be reviewed using a free VLC media player. What about the time required to calibrate the camera systems with a cube? I've got a small ROV that I can turn on and deploy in two minutes, no time intensive calibration needed.

Line 82 As you can probably surmise by now, I'm not sure that "more efficiently" actually applies.

Major Edit: In reference to the above comments from Reviewer #2, the entire introduction has been refocused on the MaxN and 95% characteristics. We agree with the reviewer's concerns that 'tool comparison' studies are difficult and often wrought with problems. We therefore

significantly reduced language that compared our tool with existing tools. We agree with the reviewers overall comment that those comparisons were overly broad and not necessary for the primary aims of the manuscript. Our philosophy is that this paper best explains a technique to researchers who have *a priori* made the decision to use a rotating camera system.

Line 102 This paragraph describes issues associated with species ID. You need to explain that along with species morphology, color, and behavior, the environmental conditions (light level, suspended particulate) are critical determinants of our ability to ID species. Like an ROV or Sub, the lander will be similarly impacted by these conditions.
A sentence addressing these concerns has been added.

Lines 127 onward. The following protocol is obviously based on SeaGIS' technology. I would precede this discussion with a short paragraph indicating that a variety of stereo video tools exist out there in the world, possibly name a few, and then explain that for the purposes of this paper you are going to limit the discussion to SeaGIS.
See above comments: JoVE will not allow commercial language in Manuscript text. SeaGIS is referenced in the materials table.

Lines 448 onward. I think that there are three related, but distinct, dimensions to this paper: landers vs other platforms, stereo video vs traditional sizing lasers, and MaxN versus other estimates. The problem is that the three dimensions are not uniformly addressed throughout the paper. For instance, your introduction (subject to my comments above) really sets the stage for a comparison between landers and other visual tools/traditional tools, without really talking about MaxN at all. But the first paragraph of your discussion, asserts the importance of MaxN as assumedly your most important finding.
The introduction has been modified to de-emphasize comparisons with other tools and instead focus on MaxN and 95% Z distance calculations

Reviewer #3:
Manuscript Summary:
The manuscript describes the use of a rotating stereo video lander for estimating fish densities and discusses statistical modifications of established methods for estimations of MaxN and maximum distance estimations for fish detection.

-The challenges of estimating fish densities for mobile species that are distributed quite variably in time and space are many, and the development of this rotating stereo video lander system is a great step forward. But this manuscript does not do it justice in it present form. The modification of MaxN and estimations of maximum detectable distances seem appropriate and were

necessary developments for incorporating the results of this tool to other scientists and managers. Clearly, this instrument covers a larger area than static video systems as shown by the comparisons of MaxN between rotating video and non-rotating video subsampled from the video. Each tool produces estimates of MaxN that are unique to each and are only useful in a relative context. For the rotating system (which by the way should be named for ease of use in manuscripts) to be accepted and useful it must be used solely in all areas of interest, otherwise estimates derived from its use are not scalable to other static video systems. While static systems underestimate fish densities relative to the rotating system, the rotating system also underestimates fish densities but less than static systems as the author's claim. I think it is important to develop a method to intercalibrate the rotating system with static systems (especially if the rotating system becomes more prevalent) to prevent historical baseline estimates developed using static systems from becoming incomparable. This should be attempted using multiple methods besides video as mentioned in the Discussion.
We fully agree that to properly do 'tool comparison', all tools need to be inter-calibrated. Because that was not the focus of this paper, we have reduced language of tool comparison as per the response to Reviewer #2's comments.

Finally, I am concerned about pooling distance estimates among different locales given the highly variable nature of water clarity and camera light penetration among settings. In my mind, this was not dealt with adequately in the manuscript or the study methods, yet it is presented as an important modification element.

We have added text to a note to try and clarify this issue. We agree that each particular survey will have differences in water clarity and camera light penetration. The 95% Z distance represents the average water clarity conditions across the entire study. We believe that it is statistically valid to pool results this way; however, this procedure would need to be recalculated if a different study were performed (say a different survey year).

-The Introduction is well written and introduces the topics of importance, however it includes results starting at line 85 which are not appropriate in the introduction. These should be limited to the results section.
We have added a few more sentences to clarify the introduction section, but it was unclear what 'results' were included. We have eliminated specific quantitative 'result' language from the introduction.

-The Protocol section is vague in describing general concepts and how these were applied (e.g., bootstrapping details are vague and inadequate).

We have added additional details to many protocol steps as was requested by other reviewers. It is unclear what level of detail is needed for a widely used (and software generic) statistical approach such as bootstrapping. In subsection 3.1.1.1, we describe taking '1000 random draws of the data with replacement' for each tested sample size and calculating a statistic (95% Z distance). We have also included code online at: https://github.com/rfields2017/JoVE-Bootstrap-Function

-Line 394: "Welches" should be "Welch's" - the use of Welch's t-test assumes independence of samples and the way that the "pseudo-stationary" estimates were derived by subsampling video sections of the rotating system means that they are not independent. I understand why this was done given the difficulty and logistics of the work and think this is a minor flaw given the magnitude of the differences.

We agree with the reviewer that assumptions for Welch's t-test were not fully met, and also agree that this is a "minor flaw given the magnitude of the differences". We therefore left the result unaltered.