Dear JoVE Editorial Office,

We are pleased to send you this improved scientific manuscript and video resubmissionafter going through peer-review. We feel that its novelty and scientific quality in presenting a novel image processing technique of agricultural significance warrant publication in the top video-graphic methodological journal, JoVE.

We have provided below detailed responses in *italics* and within quotes when indicating changes to the original manuscript with the steps taken with regards to each point raised by the editors and the three separate peer reviewer contributions. We have implemented the changes required of the video resubmission, including completely new recordings of the problematic interviews, for which we did have some scheduling conflicts and thus requested a small extension. Finally, we have uploaded the full version of the revised video in the link provided.

Our hope is that you the Editors of JoVE find that our resubmission of this scientific protocol and video-graphic detailed explanation meet your high standards and those of the reviewers that provided comments and criticisms in the standard scientific peer-review process. We remain dedicated to the scientific process of editorial and peer review and look forward to further communication.

A picture containing object

Description generated with very high confidence

Sincerely,

Shawn C. Kefauver, Ph.D.

**Researcher/Professor**

Universitat de Barcelona. Departament B.E.E.C.A.

Facultat de Biologia, Secció Fisiologia Vegetal

Integrative Crop Ecophysiology Group

https://integrativecropecophysiology.com/

e-mail: sckefauver@ub.edu, sckefauver@gmail.com

web personal/professional: https://sckefauver.com

twitter @sckefauver, tel.: +34 620 738 590

https://ci5.googleusercontent.com/proxy/5buCn-78NhQk6d3Z-byOOL-jgvwyor8i9P9huk6lnP9sbVaZYBv0dJC7by8AzgwZzeb21lBxnzcIjPXpazfNWNUqk5wY5tkrkq8bcnqZ7nzvqw=s0-d-e1-ft#http://www.ub.edu/iub/firma/recursos/AF-firmaelectronica.jpg

The manuscript has been modified to include line numbers and minor formatting changes. The updated manuscript **58695\_R1.docx** is located in your Editorial Manager account. In the revised PDF submission, there is a hyperlink to download the .docx file.**Please download the .docx file and use this updated version for future revisions.** The file is also attached.

*We have downloaded the indicated file from within the Editorial Manager and have worked using this .docx file moving forward with our revisions detailed below.*  
  
You will find Editorial comments and Peer-Review comments listed below. Please read this entire email before making edits to your manuscript.  
NOTE: Please include a line-by-line response to each of the editorial and reviewer comments in the form of a letter along with the resubmission.

*We have provided below detailed responses in italics and within quotes when indicating changes to the original manuscript.*  
  
**Editorial Comments:**  
  
• Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.

*In the process of revision, we have taken an extra look for potential spelling and grammatical errors, albeit after completing all of the revisions as part of the general and specific comments provided below. Furthermore, the first author has requested to include his third official institute affiliation, which has been appended to the start of the manuscript accordingly as well as in the online submission forms.*  
  
• **Protocol Language:** The JoVE protocol should be almost entirely composed of numbered short steps (2-3 related actions each) written in the imperative voice/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 brief “Note” at the end of the step (please limit notes). Please re-write your protocol sections 1 and 2 accordingly. Descriptive sections of the protocol can be moved to Representative Results or Discussion. The JoVE protocol should be a set of instructions rather a report of a study. Any reporting should be moved into the representative results.  
  
1) E.g, Lines 131-135, 148‒160 etc. should be a note.

*The sections starting at Lines 131-135 and Lines 148-160 have been moved to after their respective sections as “Notes.”*

*Moreover, we have moved a large part of the Section 2 Note from lines 149-158 to the Discussion, now at lines 302-310. Lines 186-195 of the paragraph at the start of Section 3 has been moved to the end of the Introduction and Lines 195 to 201 have been moved to the Discussion section and modified according to the recommendations of Reviewer 2 as detailed below.*

2) Steps not in the Imperative voice: entire Section 1, 2.

*The text for all of the steps for the entirety of Sections 1 and 2 have been changed to imperative voice. As this small change often required reorganizing the entire sentence, the authors have not placed it here. In the changing of the voice format for the protocol steps we have tried to minimize the changes to content; however, some changes in voice have resulted in simplification of the steps.*

*In the algorithm adjustments for the CerealScanner, it is not necessary to adjust for image resolution.*

“3.3.1. Adjust algorithm parameter for the camera focal length.

3.3.2. Adjust the algorithm parameter for the distance from crop canopy.”

*In the updating of the text of the Protocols to imperative voice, Protocol step 3.5 has also been removed and combined as part of 3.4.3, changing the numbering of the last three validation protocol steps.*

*“3.4.1. In “Options”, enter in Batch Inputs the location of the photos to analyze.*

*3.4.2. In the “Results Files”, select where to save the results file. The results file will include two columns with the image file name and the ear counting results.*

*3.4.3. Finally, click on “Process” and the results file with ear density in sq. meters using a simple ratio using the camera settings and the distance between canopy and camera to convert the image area to an actual canopy area in square meters following Figure 1 will be automatically produced in a few minutes, depending on the computer speed.*

*3.5. Conduct a post-processing validation after data collection by manually counting the number of wheat or barley ears in the image and then converting this to ears/m2 as in Figure 1 for comparison to the algorithm values.*

*3.5.1. Use the simple point placement tool built within FIJI that provides easy support for this process and the FIJI “Analyze Particles” function for producing the counts automatically; this is shown graphically in Fig. 6.”*

• **Protocol Detail:**  
1) Please ensure homogeneity between the written protocol and the protocol presented in the video.  
2) Please ensure that all specific details from the video are mentioned in the manuscript (the manuscript should be able to serve as a standalone article).

*In the edition of the video, we have used the revised version of the manuscript as a guide for making any changes, as well as those suggested by the Editors and two Reviewers. Specifically, as the CerealScanner plug-in has been updated with a new informative web site and some new controls, this part of the video will be re-filmed in order to provide the same exact details as described in the protocol.*  
• **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 (3-6 paragraphs): 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.

*With only minor changes to the text, we have changed slightly the organization of the Discussion section in order to ensure that it loosely follows the above format and have added a short final sixth paragraph on the critical steps within the protocol after the ideas for future applications.*

*“In summary, the critical steps for the implementation of this protocol include first and foremost the proper planning for the time of year and environmental conditions of the crop, being between optimally within growth stages of Zadocks 60-87 and either at solar noon or in diffuse light conditions. Furthermore, the acquisition of digital images should be conducted in a controlled manner accounting for camera angle, distance from the canopy and camera focus for each image. Finally, optimized computer processing options are presented in detail for reproducing the processing code pipeline or contacting the authors for either the original code or the code integrated as a graphical user interface (GUI) in a plug-in for FIJI, the CerealScanner.”*

• **Figures:**  
1) Fig 2/6/7: Please merge panels a and b into 1 page and submit as 1 file.

*We have made the following requested changes and will upload the edited files accordingly.*• **Figure/Table Legends:** Please expand the legends to adequately describe the figures/tables. Each figure or table must have an accompanying legend including a short title, followed by a short description of each panel and/or a general description.

*We have expanded each with an updated and more detailed text as follows:*

*“Figure. 1. Ear counting system using the “phenopole” shown in the field on the left, with a remotely controlled natural color (RGB) large sensor and high-resolution digital camera system with camera tilt and height, indicating the necessary parameters for adjusting the image processing algorithm. The sensor and image resolution are detected automatically by the image properties, while the user should input the specifics for the lens focal length and the distance from the canopy. These are necessary to adjust the algorithm for the estimated number of pixels per ear and also the conversion of the image-based total ear count to ear density (ears/m2) For that reason, it is recommended to use the same camera and lens focal length for all field images.*

*Figure. 2. Durum wheat (a) and barley (b) ear zenithal images for ear counting data set examples with an acceptable stage of growth and senescence from approximately Zadoks 61-87: (a) Durum wheat zenithal image data set example. (b) Barley zenithal image data set example.*

*Figure 3. Image processing pipeline for two row barley ear counting as implemented using specific computer code or using the “CerealScanner” software, both of which operate within FIJI (Fiji is Just ImageJ). Panel 1 shows the original image. Panel 2 shows the results of the applications of the Laplacian filter. Panel 3 shows the application of the Median filter, and Panel 4 shows the results of the final Find Maxima and segmentation for producing the final ear count. Then the calculations are made to convert the image count to ear density as shown in Figure 1. These images are an example taken from the Arazuri field site (NE Spain, 42°48'33.9"N 1°43'37.9"W) in diffuse light conditions.*

*Figure 4. The “CerealScanner” 2.12 Beta central tab on both levels marking the “Ear Counting” function within the “CerealScanner” algorithm collection. The user must select the “…” button to the right of “Batch Inputs” to select the folder where the images files are stored, change the default values of the H Distance (distance from the camera to the top of the crop canopy) and Focal Length, if different from the default values, and then select the “…” button to the right of Results File to choose the name and location of the final results file. The other tabs of the CerealScanner provide algorithms for trait-based phenotyping for Early Vigor and onset of Maturity as part of the CerealScanner code suite. Under the Biomass tab, there are several algorithms for estimations of more general crop vigor and biomass calculations, also for RGB digital images. The example refers to two-row barley as was demonstrated in detail in Figure 3.*

*Figure 5. Adjustments required in the image processing pipeline in order to successfully count both wheat and barley ears using the same algorithm are managed automatically as part of the camera specific adjustments of H Distance (distance between the camera and the crop canopy) and Focal Length and serve to ensure that the number of pixels per ear remains more or less constant between different applications*.

*Figure 6. Algorithm validation using manual in-image ear counts for (a) durum wheat and (b) barley. The small dots were created using the FIJI Point Tool and then counted using the Analyze Particles Function with a 0.90-1.00 Circularity constraint after applying a Color Threshold from the Hue Saturation Intensity color space for the color specified by the Point Tool. This method ensures more accurate image-based manual ear counts.*

*Figure 7. Algorithm validation using manual counts in the field and manual in-image ear counts using a circle (a) wheat and (b) barley. (a) Wheat image count validation example using a circle. (b) Barley image count validation example using a circle. The subset counts of the ears within the white circle were counted using the same technique described in Figure 6 with the Point Tool, Color Threshold, and then Analyze Particles Function with Circularity constraints and color selection using Hue.*

*Figure 8. The coefficient of determination between ear density (number of ears/m2) using manual image-based counting and the image algorithm ear counting for durum wheat and two-row barley at different acceptable crop growth stages (Zadoks scale 61-87). Both of the axes show calculations including conversions to ear density rather than image-based result only. The representative results are presented here for two different crops over three different growth stages as well as different light conditions, direct sunlight images for durum wheat at Zadoks 61-65 on the top (R2=0.62, n=72), diffuse light images for barley at Zadoks 71-77 in the middle (R2=0.75, n=30), and diffuse light conditions for durum wheat at Zadoks 81-87 below (R2=0.75, n=24). An example image of each is also shown as inset in the bottom right corner of each.”*  
• **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 Lemnatec FieldScanalyzer (Lemnatec Ltd.),  
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 have removed the specific reference to the Lemnatec FieldScanalyzer and other commercial products from the text as requested. As this was not part of our actual protocol, but part of a literature reference, the commercial product was not added to the table of materials.*

*The specific text on Lines 102-105 have been edited to read as follows:*

*“Another example is an automatic ear counting algorithm developed using a fully automated phenotyping system* *with a rigid motorized gantry was used with good accuracy for counting ear density in a panel composed of five awnless bread wheat (Triticum aestivum* L.*) varieties growing under different nitrogen conditions13.”*  
• **Table of Materials:** Please revise the table of the essential supplies, reagents, and equipment. The table should include the name, company, and catalog number of all relevant materials/software in separate columns in an xls/xlsx file. Please include items such as cameras, software, etc.

*We have updated the table as follows and will provide it separately as an xlsx file.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Name of Material/ Equipment** | **Company** | **Catalog Number** | **Comments/Description** |
| ILCE-QX1 Camera | Sony | WW024382 | Compact large sensor digital camera with 23.2 x 15.4 mm sensor size. |
| E-M10 Camera | Olympus | E-M10 | Compact large sensor digital camera with 17.3 x 13.0 mm sensor size. |
| Multipod | Sony | VCT MP1 | "Phenopole" in the JoVE article |
| Computer | Generic | -- | Data and image analysis |
| ImageJ/FIJI (FIJI is just Image J) | NIH | http://fiji.sc | Plug-in and algorithms for data and image analysis |
| Circle/Metal Ring | Generic | Generic | Metal ring for in-field validation |
| Crab Pliers Clip | Newer | 90087340 | Circle support and extension arm |

• Please define all abbreviations at first use.

*We have completed an extra check of all of the abbreviations in the manuscript after implementing all of the other requested revisions.*  
• 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]."

*All of the figures presented here in this manuscript are original. We have even included new data analyses, so additionally the representative results are also new.*

• **Video-related comments:**  
Branding concerns  
• 0:00-0:09 - The university logos should be removed from the beginning of the video. They can be moved to the end of the video.

*The logos on the start page will be replaced with the textual information regarding the different author affiliations. Logos no only appear at the end of the video.*

• 8:19 - One of the authors (affiliated with Syngenta and wearing a Syngenta shirt) mentions Syngenta in the interview here. Please eliminate/minimize the appearance of commercial bias.

*This video has been re-recorded wearing clothing without any visible logos.*

Audio issues  
• 0:25-1:22, 7:47-8:03 - The audio and video appear to be slightly out of synch during this interview. This should be corrected.

*These two interviews have been re-recorded entirely with a more advanced microphone and video setup.*

Frame size/proportions issues  
• 0:40 - The white background of this figure should be extended to fill the background

*This will be corrected so that the figure is seen larger and has no background framing.*

**Comments from Peer-Reviewers:**   
  
  
**Reviewer #1:**  
Manuscript Summary:  
The manuscript presents a protocol for deriving cereal crop ear counts from zenithal RGB images using a custom software solution that implements image processing, using Laplacian and frequency domain filters. The video is clearly understandable and professionally produced.

*Thank you, we have worked hard to film with the highest possible quality video at each part of the process. As this is a protocol with necessarily two parts, in the field and the laboratory, we were obliged to undertake the filming by ourselves.*  
  
Minor Concerns:  
-Citation 11 on line 90 seems to be wrong. From the text, I expected a review publication summarizing recent agricultural image analysis techniques including the 3d reconstruction based on SfM techniques, but the referenced paper is a publication using the described measurement protocol for ear counting.

*We have checked this error in the listed citations and have changed it to the correct citation with reference to the context of the SfM 3d reconstruction techniques and as they have been used for ear counting.*

-On line 204, the link given to the CerealScanner plugin (htpps://gitlab/sckefauver/CerealScanner) is not working. The installation instructions are a little bit unclear. Is there an option of using the processing code or the CerealScanner plugin instead? Because the webpage is not publically accessible, this can not be verified.

*Section 3.1 has been modified to include a link to an additional new informative website as well as information for how to gain access to the gitlab code and plugin repository. It is not the authors’ intention to make the software available without any control, but rather available upon request. For this purpose, we have created a new informative webpage with a description of the software and contact information for acquiring access to either the original code or the graphical user interface (GIU) plugin version. We have also provided copies of both the code for FIJI and the actual plugin as part of the resubmission if the Reviewers wish to personally test the software functionality. The new text for 3.1 reads as follows:*

*“3.1. Download and install FIJI, Java 8 and the processing code or the University of Barcelona proprietary "CerealScanner” plugin at https://fiji.sc/, https://www.java.com/en/download/ and https://integrativecropecophysiology.com/software-development/cerealscanner/ (information) or htpps://gitlab/sckefauver/CerealScanner (code repository); contact corresponding authors for access permissions. The plugin is installed within FIJI by simply copying into the plugins folder.”*

-The table of materials on line 252 mentions an "Ordenador", this should probably be "Computer" or "Desktop Computer".

*We have corrected this mistake and apologize for the errata.*  
  
**Reviewer #2:**  
Manuscript Summary:  
The manuscript describes an automatic method for counting cereal ears based on digital RGB images.  
  
Major Concerns:  
While the manuscript shows promise, there are several problems that need to be addressed in terms of text clarity, methodology detailing and results analyses. Some major concerns:  
- Line 119: The authors state that "The system may be extended without major problems to bread wheat". Even though morphologies are similar and there is no awn in the latter case, without experimental evidence it is not possible to categorically assert that the system can be easily extended. The authors should state that the system is likely to work with bread wheat, but further experiments are needed to confirm this.

*We have changed the text as listed on line 119 and have changed it to the more conservative language as suggested by Reviewer 2; it now reads as follows:*

*“**The system has been demonstrated on examples for durum wheat and barley but should be extendable in application to bread wheat, which, besides exhibiting ears with similar morphology are frequently awnless, but further experiments would be required in order to confirm.”*

- Lines 142-144: the authors recommend that images be captured close to the middle of the day or under overcast conditions (diffuse lighting). These are vastly different conditions that produce images with quite different characteristics. Under direct sunlight, shadows become a problem, but the image has much more depth and contrasting features. Under diffuse lighting, the image tends to be "flattened", with many details and small features tending to become indiscernible. Did the authors test both types of conditions? Were the results the same?

*Per the Editors guidance, much of this section has been changed. The relevant part now, in sub-part 1.3 specifies, as the Reviewer notes, that the images should be captured either within two hours of solar noon or in diffuse light conditions, as copied in quotes below. As the Reviewer correctly observes, these conditions are indeed quite different in terms of the qualities of images that they produce, especially with relation to observable depth of field, but both light conditions are quite equally effective in reducing the largest source of error in the presented ear counting protocol, the negative impact of shadows. This is a small point that we have added to this protocol step in order to provide some explanation to why we specify these two particular yet seemingly different light conditions. In a similar related publication, which is cited here, we presented also several results from non-optimal conditions with examples of the errors caused by each.* *Here we focused on the presentation of the two main optimal type conditions and have added specifics and examples for each below, including early reproductive stage direct sunlight ears for wheat, post anthesis diffuse sunlight ears for barley and late grain filling growth stage diffuse light for wheat..*

*“1.3. Plan the field excursions to capture the images within two hours of solar noon or alternately on an overcast day in diffuse light conditions in order to avoid the negative effects of ear shadowing on the ear counting algorithm.”*

- Lines 153-156: the information on the number of MPixels is not nearly as relevant as the number of pixels per ear (or grain) necessary for the system to work properly.

*Reviewer 2 is correct that the number of pixels per ear is more relevant to the quality of the results rather than the full resolution of the images taken by the camera. Given our recommendations here for taking images at a distance of 0.8 m, we have estimated that most modern cameras and standard ranges of lenses would meet the minimal requirements as to the number of resulting pixels per ear. Per the Editors guidance we have moved the description of the megapixels of each camera to the Discussion section where it is covered in more detail with the rest of the relevant camera characteristics and recommendations, with some examples of what was used to produce the Representative Results. Moreover, some literature references where a more complete study of the effect of image resolution reduction has on the quality of the results of the presented algorithm have been added. The text is now present on lines 339 to 344 as follows:*

*“The most important relates to the image pixel resolution compared to the ears, followed by a smoothing factor that must be adjusted depending on whether the target crop is wheat or barley. In the case studies presented here, we have used two different compact cameras with large sensors of 20.1 MP and 16.0 MP capturing images with a wide-angle lens of 16-20 mm from a distance of 80 cm from the crop canopy. This has proven more than sufficient to produce detailed canopy barley and wheat information, with simulations demonstrating that the technique maintains high levels of precision down to 8 MP11.”*

- Paragraph 186-200: there is important information missing here. It is not possible to reproduce the work without using FIJI and CerealScanner, and even if such software is employed, it is not clear if the information provided here is enough. For example, what are the kernels used in the Laplacian and median filters? How does this "Find Maxima segmentation technique" works? What about "Analyze particles"?

*This is indeed a pertinent request made by Reviewer 2 and we have added in here the additional information of interest to users. As the location of this text has been moved per the instructions of the Editors, we have expanded this relevant section where it now lays in the Introduction and the Discussion, Lines 186-195 of the paragraph at the start of Section 3 has been moved to the end of the Introduction and Lines 195 to 201 have been moved to the Discussion section. The specific part from Paragraph 186-200 referenced by the reviewer has been expanded with details and incorporated into the last paragraph of the Introduction and now reads as follows:*

*“**The image processing algorithm is composed of three processes that first effectively remove unwanted components of the image in a manner that then allows for the subsequent segmentation and counting of the individual wheat ears in the acquired images. First, a Laplacian frequency filter has been used in order to detect changes in the different spatial directions of the image using the default ImageJ filter settings without window kernel size adjustments (https://imagej.nih.gov/ij/plugins/index.html#filters), which reduces the presence of most of the unwanted background elements in the image, including leaves, soil and some of the awns. Next, a median spatial filter reduces the high frequency noise around the ears due to the presence of the awns. Then, the Find Maxima segmentation technique determines the local peaks after the median spatial filter step, at which stage the pixels related with ears have higher pixel values that soil or leaves. Therefore, Find Maxima is used to segment the high values in the image and those regions are labeled as ears, which identifies ears while also reducing overlapping ear errors. Analyze particles is then used on the binary images to count and/or measure parameters from the regions created by the contrast between the white and black surface created by the Find Maxima step. The result is then processed to create a binary image segmentation by analyzing the nearest neighbor pixel variance around each local maximum to identify the wheat ear shapes in the filtered image. Finally, the ear density is counted using the Analyze Particles, as implemented in FIJI15. Both Find Maxima and Analyze Particles are standalone functions and also available as plugins in FIJI (https://imagej.nih.gov/ij/plugins/index.html).”*

- Figure 8: how many images were used to produce the results depicted here? How many of those were taken under direct sunlight? How many were captured under overcast conditions?

*We have added this information as part of the expansion of the figure caption, per the request of the Editors, and have also included the additional information requested by Review 2.*

*“Figure 8. The coefficient of determination between ear density (number of ears/m2) using manual image-based counting and the image algorithm ear counting for durum wheat and two-row barley at different acceptable crop growth stages (Zadoks scale 61-87). Both of the axes show calculations including conversions to ear density rather than image-based result only. The representative results are presented here for two different crops over three different growth stages as well as different light conditions, direct sunlight images for durum wheat at Zadoks 61-65 on the top (R2 0.62, n=72), direct sunlight images for barley at Zadoks 71-77 in the middle (R2 0.75, n=30), and indirect or diffuse light conditions for durum wheat again at Zadoks 81-87 below (R2 0.75, n=24). An example image of each is also shown as inset in the bottom right corner of each.”*

- Line 303: the authors state that they tested various different techniques before arriving at the protocols presented in the manuscript. What were the techniques that were tested and discarded? This is important information for the scientific community, especially to support the design of future experiments.

*With respect to providing more details on other preliminary tests that did not produce as reliable results as the presented protocol, we have added in the following sentence to this regard at line 336:*

*“Other image filtering attempts based on RGB color or alternative color spaces, such as Hue-Saturation-Intensity or CIE-Lab, were not as effective or consistent as the use of the Laplacian and Median frequency domain filters in removing unwanted image elements, especially the awns.”*

- The authors should discuss the effects of lighting on the accuracy. Did the flattening effect caused by diffuse light influence the results? How did the system deal with images with lesser details?

*Actually, this theme has already been considered as important and is brought up in the third to last paragraph in the Discussion starting on line 367 beginning with the following statement:*

*“Even though the data collection in field conditions requires close attention to such environmental conditions as sunlight intensity and sunlight illumination angles, the robust image analysis algorithm presented here provides some leeway in the image capture window by using spatial techniques that ignore image albedo effects, given that the correct image exposure was used for the particular light conditions at the moment of image capture.”*

*In order to better address this issue, we have added the following text at line 371, following the guidelines set forth by Reviewer 2 to better address the potential effects of lighting.*

*“In previous work, a fuller range of lighting effects were tested, indicating that the only major source of error with regards to light effects is the production of strong shadows in the image when capturing images in direct sunlight either early or late in the day due to the angle of the sun.”*

*With regards to images with ‘lesser details,*’ *we assume that the Reviewer is referring to lower resolution images, either with fewer megapixels or taken from a greater distance. This is discussed in more detail in the second to last paragraph of the discussion, with reference to another previous work in which we simulated lower resolution images and presented its results on the algorithm accuracy, the product of which is the recommended range of spatial resolutions presented in this protocol.*  
  
Minor Concerns:  
There are many other problems, most of them related to writing, that also need to be addressed:  
- The text needs improvement: there are many typos, grammatical errors and overly long and confusing sentences.

*We have rewritten a number of sections of the manuscript per the indications of the Editors towards improved adherence to the style guidelines, and also as part of the editions suggested by the other reviewers. Furthermore, we have endeavored to conduct a detailed revision for all small errors in the text, which are not copied in detail here.*

- Lines 78-81: sentence is too long and difficult to follow.

*This sentence has been simplified and edited for clarity to a much shorter sentence from:*

*“Although previously applied to precision agriculture and more recently to phenotyping activities, remote sensing has traditionally focused on multispectral, hyperspectral, and thermal imaging sensors from aerial platforms for larger studies, in the case of precision agriculture, or ground-based platforms, for plant phenotyping studies at the small plot canopy scale10”*

*And now reads as follows:*

*“Remote sensing has traditionally focused on multispectral, hyperspectral, and thermal imaging sensors from aerial platforms for precision agriculture at the field scale or for plant phenotyping studies at the micro-plot scale10.”*

- Lines 99-102: sentence is too long and difficult to follow.

*This sentence has been edited into two separate sentences for clarity from the original long sentence:*

*“Recent work has advanced in this direction by using a black background structure supported by a tripod in order to acquire suitable crop images by avoiding excessive sunlight and shadow effects, but is cumbersome in field conditions even though demonstrating fairly good results in ear counting12.”*

*And now reads as follows:*

*“Recent work has advanced in this direction by using a black background structure supported by a tripod in order to acquire suitable crop images, demonstrating fairly good results in ear counting12. In this way, these authors avoided excessive sunlight and shadow effects, but such a structure would be cumbersome and a major limitation in application in field conditions.”.*

- Many excerpts are too verbose. For example, the excerpt "More recent work by led Fernandez-Gallego published in 2018…" (line 105) could be replaced with "Recent work…". There are many other instances like this throughout the text.

*We have shortened the text of the manuscript at the indicated location and have edited a number of other locations using more concise language as part of an editorial review of the whole manuscript that is not detailed here as it consisted of a number of small changes.*

- Lines 115-117: convoluted sentence, it should be rewritten for clarity.

*This sentence has been edited from its original version,*

*“This work proposes a simple system for the automatic quantification of ear density, using the examples of durum wheat and barley crops grown in field conditions, based on VHR RGB (natural color) images acquired from commercially available digital cameras.”*

*And now reads as follows:*

*“This work proposes a simple system for the automatic quantification of ear density in field conditions using images acquired from commercially available digital cameras.”*

- Lines 117-119: This sentence needs to be better explained.

*This sentence has been edited from its original version,*

*“The system takes maximal advantage of natural light in field conditions and therefore needs to take into account various environmental factors but remains in effect simple to implement.”*

*And now reads as follows:*

*“This system takes advantage of natural light in field conditions and therefore requires consideration of some related environmental factors, such as time of day and cloud cover, but remains in effect simple to implement.”*

- Lines 121-124: confusing. It is not clear how these two sentences relate, nor the reason the authors mention manual counting using images or in the field, since the proposed scheme is supposed to be automatic. This whole paragraph has many problems and should be rewritten for the sake of clarity.

*These sentences have been edited from their original version,*

*“In our protocol, zenithal images were taken by holding an RGB camera by hand above the crop. In real field conditions the easiest way and the cheapest way to acquire information is walking across the plots in the field and counting the ears manually in the image or in the field for validation.”*

*And now reads as follows:*

*“In our data capture protocol, zenithal images are taken by simply holding by hand or using a monopod for positioning the digital camera above the crop. Validation data can be acquired by counting the ears manually for sub-plots in the field or during postprocessing by counting ears in the image itself.”*

- Line 133: please give examples of factors that affect image quality (shadowing is certainly not the only relevant factor).

*All of Protocol 1 and 2 have been completely re-written at the behest of the Editors. The relevant text is now part of the Note for Protocol 1:*

*“Protocol 1 Note: In considering the objectives of this protocol, it is important to first consider that the growth stage of the crop is suitable for applying ear counts. Capturing images outside of the recommended growth stage will either result in sub-optimal or meaningless results (if ears are not present or fully emerged). Image quality also has a considerable impact on processing results, including resolution and sensor size, and some environmental conditions, such as time of day and cloud cover, need to be carefully considered before proceeding with image capture.”*

- Lines 148-149: this needs a better explanation.

*This section has been entirely removed and most of Protocol 2 has been completely re-written. The lines at 148-149 have been redone and reduced to a small Note for Protocol 2.*

*“Protocol 2 Note. Three major considerations in selecting a camera therefore include: (1) camera specifications, in this case the sensor physical size; (2) focal length of the image lens, and (3) distance between the canopy and the camera: closer distances or greater zoom lenses will capture less area while images captured from a greater distance or with lower zoom will capture more crop area. See Figure 1 for the details on the relevant camera specifications.”*

- Figure 1: the legend should be more descriptive, so the figure can be understood standing alone.

*All of the Figure legends have been re-written and expanded.*

*“Figure. 1. Ear counting system using the “phenopole” shown in the field on the left, with a remotely controlled natural color (RGB) large sensor and high-resolution digital camera system with camera tilt and height, indicating the necessary parameters for adjusting the image processing algorithm. The sensor and image resolution are detected automatically by the image properties, while the user should input the specifics for the lens focal length and the distance from the canopy. These are necessary to adjust the algorithm for the estimated number of pixels per ear and also the conversion of the image-based total ear count to ear density (ears/m2) For that reason, it is recommended to use the same camera and lens focal length for all field images.”*

- Line 162: the way the sentence is written, it looks like the images are captured in motion, which is probably not the case.

*All of Protocol 2 has been redone completely. This sentence was previously:*

*“2.1. Images captured are taken walking between the plots to obtain zenithal images. Using the “phenopole” acquisition system shown in Fig. 1, it is possible to capture images quickly and yet in a controlled and consistent manner.”*

*It now reads:*

*“2. Image capture in field conditions with natural light*

*2.1. Prepare a “phenopole” as shown in Fig. 1 or a similar acquisition system (even handheld) to capture images at each plot or target location quickly and yet in a controlled and consistent manner.*

*2.2. Position the camera on a suitable monopod or “selfie” pole such that it may be maintained level either using level bubbles or an in-camera stabilization system to obtain zenithal images.”*

- Figure 2: too many examples. One sample from each row in the images would be more than enough. The legend is highly redundant, as the information before and after ":" is exactly the same.

*We have reduced the number of example images here to show only one row of example images for wheat and barley. All of the Figure legends have been completely re-written.*

*“Figure. 2. Durum wheat (a) and barley (b) ear zenithal images for ear counting data set examples with an acceptable stage of growth and senescence from approximately Zadoks 61-87.”*

- Lines 174-175: confusing.

*All of Protocol section 2 has been re-written entirely. Section 2.4:*

*“2.4. Note the image number prior to image capture in order to match the images correctly with the field plots. An image of the general field area to start and on image of the ground/field between blocks is recommended for control and counting.”*

*Now reads:*

*“2.4. Take note of the image number prior to image capture in order to match the images correctly with the field plots. Record one image of the general field area at start and one image of the ground/field between blocks for post-processing controls.”*

- Lines 186-188: sentence is too long.

*This whole section has been deleted, with some relevant parts incorporated into either the Introduction or Discussion sections accordingly.*

- Figure 5: legend should be more descriptive, as is it is a bit ambiguous.

*All of the Figure legends have been re-written and expanded.*

*“Figure 5. Adjustments required in the image processing pipeline in order to successfully count both wheat and barley ears using the same algorithm are managed automatically as part of the camera specific adjustments of H Distance (distance between the camera and the crop canopy) and Focal Length and serve to ensure that the number of pixels per ear remains more or less constant between different applications.”*

- Figures 6 and 7 should be better explained. Are the points marked manually or automatically? The way text and legends are written, it is very difficult to understand the process.

*All of the Figure legends have been re-written and expanded.*

*The captions for Figures 6 and 7 were previously:*

*“Figure 5. Adjustments required in the image processing pipeline in order to successfully count both wheat and barley ears using approximately the same algorithm.*

*Figure 6. Algorithm validation using manual in-image ear counts for (a) durum wheat and (b) barley.”*

*They now read:*

*“Figure 6. Algorithm validation using manual in-image ear counts for (a) durum wheat and (b) barley. The small dots were created using the FIJI Point Tool and then counted using the Analyze Particles Function, with a 0.90-1.00 Circularity constraint, after applying a Color Threshold from the Hue Saturation Intensity color space for the color specified by the Point Tool. This method ensures more accurate image-based manual ear counts.*

*Figure 7. Algorithm validation using manual counts in the field and manual in-image ear counts using a circle. (a) Wheat image count validation example using a circle. (b) Barley image count validation example using a circle. The subset counts of the ears within the white circle were counted using the same technique described in Figure 6 with the Point Tool, Color Threshold, and then the Analyze Particles Function with Circularity constraints and color selection using Hue.”*

- Figure 3 (legend): what do you mean by "Fiji is Just ImageJ"?

*All of the Figure legends have been re-written and expanded. The Figure 3 legend now reads:*

*“Figure 3. Image processing pipeline for two row barley ear counting as implemented using specific computer code or using the “CerealScanner” software, both of which operate within FIJI (Fiji is Just ImageJ). Panel 1 shows the original image. Panel 2 shows the results of the applications of the Laplacian filter. Panel 3 shows the application of the Median filter, and Panel 4 shows the results of the final Find Maxima and segmentation for producing the final ear count. Then the calculations are made to convert the image count to ear density as shown in Figure 1. These images are an example taken from the Arazuri field site in diffuse light conditions.”*

*With regards to this specific text, the acronym* ***FIJI*** *is* ***F****IJI* ***i****s* ***j****ust* ***I****mageJ.*   
  
  
**Reviewer #3:**  
Manuscript Summary:  
This manuscript describes a protocol for collecting top down images of wheat and barley fields, as well as a software approach to identifying and counting wheat heads within the resulting images. The methodology is clearly written and the technique is likely to be of interest to a significant number of breeders and quantitative geneticists working in these two species. I have only a couple of minor points about the manuscript as currently written.  
  
Minor Concerns:  
Lines 119-121: "The system may be extended without major problems to bread wheat, which besides exhibiting ears with similar morphology than bread wheat, the ears frequently are awnless, which would make ear counting easier."

*We have changed the text as listed on line 119 and have changed it to the more conservative language as suggested by Reviewer 2 and Reviewer 3; it now reads as follows:*

*“The system has been demonstrated on examples for durum wheat and barley but should be extendable in application to bread wheat, which, besides exhibiting ears with similar morphology are frequently awnless, but further experiments would be required in order to confirm.”*

Line 136: Zadoks is a sufficiently narrowly used scale that it would make sense to describe it a bit more here as some of the target audience is presumably researchers working in other crops who may want to adapt this technique to their target species.

*We have re-written all of the Protocol 1 section and have added additional text and also language in the video to indicate that the optimal image data acquisition should be as follows:*

*“1.1. Make sure that the crop growth stage is approximately in the range of Zadoks 60-87, between grain filling and near crop maturity, with ears that are still green even if the leaves are senescent14. Some yellowing of the leaves is acceptable, but not necessary.*