ICCV2015
International Conference on Computer Vision 2015
December 11-18, Santiago, Chile

Reviews For Paper

Paper ID 408

Title Forget the checkerboard: practical self-calibration using a planar scene

---

Masked Reviewer ID: Assigned_Reviewer_1

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper describes an approach for self-calibration based on planes. It starts the calibration procedure by performing a projective reconstruction, including the estimation of distortion (uncalibrated). Bundle adjustment is used to improve the projective reconstruction. Self-calibration is then performed based on homographies--with the same constraints as in paper [2]. The new contribution is the use of a normalized version of the constraints. The normalized constraints are used in a non-linear optimization procedure that estimates the intrinsic parameters and the plane normal (in the reference frame). Metric reconstruction (including bundle adjustment) is then performed leading to estimates of the intrinsic parameters (including distortion), extrinsic parameters and 3D point positions.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if	The main strength of this paper is a self-calibration plane-based approach (from my point of view it is not a strict self-calibration method) whose accuracy is similar to the state-of-the-art planar calibration (e.g. Bouguet's toolbox)--there is extensive experimentation. Another strength of the paper is the definition and introduction of the normalized version of the constraints of equations 8 and 9. These constraints had already been derived in a previous paper [2]. They are re-derived in this paper but, most importantly, a normalized version (section 5.3) is presented. This normalized version seems to be essential for the accuracy of the results.

you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Do not ask the authors to cite your own work. If you think this is essential, write it in the confidential comments to the AC. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to ICCV and that a theoretical paper need not have experiments. It is not ok to reject a paper because it did not outperform other existing algorithms,	One of the weaknesses of this paper stems from the structure of the approach: to me it is not clear why the procedure starts by a projective reconstruction (including distortion estimation) followed by projective bundle adjustment. From my point of view the authors do not clearly motivate the need for such steps. In addition they do not discuss the effect of these steps on the quality of the results namely, intrinsic parameters, distortion and reprojection errors. The approach also provides the extrinsic parameters and 3D point coordinates. Do the procedures of the first 3 boxes in Figure 3 equally affect all these parameters? Other weaknesses of the approach result from its dependency on the spatial distribution of the camera positions as well as on the number of images required for results with good accuracy. This may be the result of a self-calibration (more dofs) but it is not clear. For the experimental results with synthetic data three configurations (distributions of camera positions) are considered (section 7.1). Most of the results concern scenes A and B.

especially if the theory is novel and interesting. It is also not ok to ask for comparisons with unpublished papers and papers published after the ICCV deadline. Last but not least, remember to be polite and constructive.	
Preliminary Rating. Please rate the paper according to the following choices. Oral: these are papers whose quality is in the top 10% of the papers at ICCV. Examples include a theoretical breakthrough with no experiments; an interesting solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid	Poster

experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers that have major flaws, or have been done before.	
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making. There is no need to summarize the paper.	I think that this paper makes an interesting contribution and also that the method presented may be useful in some applications. However I do have several questions: --How important is the initial projective reconstruction, distortion estimation and projective bundle adjustment? --Figure 7 clearly shows a dependency on the number of images. Apparently 20 or more images are required so that the accuracy is similar to competing approaches. How do you relate the number of images with the number of features? --Can you present or discuss your experimental results for the scenario C, i.e., random positions of the camera? This is important for comparison. --Contrary to what you mention, scenario A also contains translation. How sensitive is this method to the spatial distribution of the camera positions? --Your section on the Experimental results in unclear in several aspects. Namely: --Page 7, 74--87: you mention that fig. 6 shows "what are the best possible results for a planar self-calibration and for a known-target calibration". What exactly do you mean? Best how? Just because you initialize the final metric bundle adjustment with the ground truth? --Page 7, 92-93: you mention that your calibration is very close to the theoretical optimum. How do you prove that? --For your experimental results you use "validation" images. Are they used for anything else other than computing the reprojection error? --Also for the results presented in fig. 7 what are the "best calib" and the "best self"?
Confidence. Write "Very Confident" to stress that you are absolutely sure about your conclusions	

(e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details as confidential comments to PC/AC chairs (point 7.).	Confident

Masked Reviewer ID: Assigned_Reviewer_18

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	This paper describes a calibration method which uses homography between scene planes to estimate intrinsic parameters and then uses bundle adjustment for metric reconstruction to compute the extrinsic parameters.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels of evaluation: a theoretical paper may need no experiments, while a paper	A system has been developed to combine homography based calibration with metric reconstruction. The paper is well written. A possible bias in previous methods has been mentioned and a technique has been

presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	presented to handle it. A different method is presented to derive Strum's equations.
Paper Weaknesses. Please discuss the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Do not ask the authors to cite your own work. If you think this is essential, write it in the confidential comments to the AC. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that	The paper lacks in novel content. Most of the things have already been done in prior work. I think the main contribution was bias removal. The authors could probably explain the bias in detail. Most of the other sections are not new and standard multi view geometry. The whole system is good, but the actual novel contribution has been explained very little.

theoretical results/ideas are essential to ICCV and that a theoretical paper need not have experiments. It is not ok to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not ok to ask for comparisons with unpublished papers and papers published after the ICCV deadline. Last but not least, remember to be polite and constructive.	
Preliminary Rating. Please rate the paper according to the following choices. Oral: these are papers whose quality is in the top 10% of the papers at ICCV. Examples include a theoretical breakthrough with no experiments; an interesting solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have	

one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers that have major flaws, or have been done before.	Weak Reject
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their rebuttals to facilitate your decision making. There is no need to summarize the paper.	Please see the weakness section.
Confidence. Write "Very	

Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details as confidential comments to PC/AC chairs (point 7.).	Confident

Masked Reviewer ID: Assigned_Reviewer_6

Review:

Question	
Paper Summary. Please summarize in your own words what the paper is about.	The paper proposes to use a planar (textured) scene for calibrating (the intrinsics of) images, instead of using a planar scene with with known targets or an arbitrary 3D scene. The method is based on estimating homographies and fusing them to arrive at approximate values for a self calibrating bundle adjustment. The method is described and experimentally evaluated.
Paper Strengths. Please discuss the positive aspects of the paper. Be sure to comment on the paper's novelty, technical correctness, clarity and experimental evaluation. Notice that different papers may need different levels	

of evaluation: a theoretical paper may need no experiments, while a paper presenting a new approach to a known problem may require thorough comparisons to existing methods. Also, please make sure to justify your comments in great detail. For example, if you think the paper is novel, not only say so, but also explain in detail why you think this is the case.	The proposal is an interesting intemediate configuration between full 3D knowledge (testfield) and no knowledge (free set of 3D points), as only 2D coordiantes of the scene points are assumed to be unknown (z=0). As planar objects are omnipresent, this appears as a relevant aspect.
	It took me a while to understand the procedure (not really before Sect. 3, where the used camera model was explained). The description of the procedure is not fully clear. The experiments contain a flaw: Comparing results with 18 and with 35 images of course show, that the accuracy with 18 images is worse. But this is not due to the configuration, but to the pure number of images. Detailed comments. l. 34 "intrinsics" may refer to only the five parameters of the calibration matrix or to all parameters necessary for non-linear distortions (context dependent) it would be good to immediately make this clear. l. 45 Though it is true that planes are omnipresent, most seemingly planar objects are not really planar, when analysed more in detail. The deviations from planarity need to be adressed in their relation to the task of calibration, since deviations from planarity directly lead to errors in calibration. The problem needs (1) adressed, (2) analysed wrt to requirements (what deviations from a plane can be tolerated if a certain accuracy in the image is aimed at?) l. 190 The notion 'homography-based self-calibration' (HBSC) is not precise enough as basis for a discussion, especially when motivating a new procedure: What can (HBSC) do, where are its limits? Why is a projective bundle needed and then a metric bundle? No reasons are given for these statements. l.187: there is no subsection 1.2 ....
Paper Weaknesses. Please discuss	l. 200 Here the notion 'distortion' appears the first time l. 239 better use \cal M(x) for motion, as R not only contains a rotation R

the negative aspects of the paper: lack of novelty or clarity, technical errors, insufficient experimental evaluation, etc. Please justify your comments in great detail. If you think the paper is not novel, explain why and give a reference to prior work. Do not ask the authors to cite your own work. If you think this is essential, write it in the confidential comments to the AC. If you think there is an error in the paper, explain in detail why it is an error. If you think the experimental evaluation is insufficient, remember that theoretical results/ideas are essential to ICCV and that a theoretical paper need not have experiments. It is not ok to reject a paper because it did not outperform other existing algorithms, especially if the theory is novel and interesting. It is also not ok to ask for comparisons with unpublished papers and papers published after the ICCV deadline. Last	but also a translation l. 326 Why does the choice of the reference frame induce a bias? This is unclear. l. 333 (7) the recularization should be weighted with 1/variance of the assumed principal point, e.g. variance = (10 pixel)^2. l. 455 Why diag(f,f,0) and not (f,f,1) ? l. 573 (22) (1) I presume, that the x_k are the (x,y)_k coordiantes of the scene points with z_k=0 - please make this explicit ('position' is not clear enough). (2) If the plane is at z=0, the reference camera can only be approximately at R=[a,b,n] ... (see line 547), Otherwise you fix the relative pose between the reference camera and the plane, which yields suboptimal results. But if the the reference camera is free, then the three parameters of the origin in the plane needs to be specified (and fixed). It is unclear to me how you use 'all available information' and not more? l. 619 If I understand [10] correctly, it assumes varying focal length. Up to know I thought, the focal length would be the same for all cameras here. I am confused. l. 646 I would suspect, that if the angles are smaller the stability of the solution gets worse. So, is there a limit of 35° only, as the experiements were not done for larger angles? l. 652 Fig. 5: use thicker lines to make the colours mor distinct (> 0.3 mm at printing or viewing size) l. 687: add: this is called 'best possible' - though the notion is misleading: These are three cases, which use different information (which cost a different amount). Best possible, could also assume using 50 or more cameras etc ... Less images, other configurations, additional information about coordiantes - all choices change the quality of the result in a predictable manner. In each case, there is a bundle solution as reference: So, if the method is a bundle solution, there is no possibility to make a comparison based on simulations. So, the bundle solution with 18 images for case Fixed-A should coincide with the result shown. But: if the inverse normal equation matrix is taken, being the covariance matrix of all parameters, a conparison could be made between this internal predicion (Cramer Rao bound) and the simulations. It is unclear: what is a fair reference? l. 689 It is known, that rolling the cameras (at leas one camera by 90 degrees) increases the stability of the calibration (referring to scale differences). Was this taken into account?

but not least, remember to be polite and constructive.	see DIGITAL CAMERA CALIBRATION METHODS: CONSIDERATIONS AND COMPARISONS Fabio Remondino, Clive Fraser, In International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,, Vol. Vol. XXXVI, No. part 5. (2006) Key: citeulike:1655218 l. 713 Fig. 6 lower right: self calibration can be done for the 18 images of Fixed-A or the 35 images of Fixed-B. It is unfair to Fixed-A to compare it to selfcalibrating BA with 35 cameras! - Then naturally Fixed-A will turn out to be worse. I miss a solution with a self-calibrating bundle adjustment where all three coordiantes of the 3D points are unknown, thus, the points are not forced to lie on a plane. The result should yield a (hopefully only slightly) smaller reprojection error. The covariance matrix of the intrinsic parameters can directly be derived from the inverse of the normal equation matrix: How do they compare to the empiricl findings in Fig. 6? l. 748: ORB features (give reference) l. 769 Fig, 7: to which value should it converge? If the number of images would be infinity, the distortion error should decrease to 0. Roughly the curves should go with 1/sqrt(I) where I is the number of images (as doubling the number of images, doubles the elements in the normal equation matrix, thus halves the covariance matrix, thus dimninuishes the standard deviations by a factor 1/sqrt(2)) l. 820: what standard deviations do these values have? l. 846 This sounds as, if the angles are large, no good results are obtained - where is this shown?
Preliminary Rating. Please rate the paper according to the following choices. Oral: these are papers whose quality is in the top 10% of the papers at ICCV. Examples include a theoretical breakthrough with no experiments; an interesting solution to a new problem; a novel solution to an existing problem with solid experiments; or an incremental	

paper that leads to dramatic improvements in performance. Oral/Poster: these are very strong papers, which may have one weakness that makes you unsure as to whether they should be oral or poster. Poster: these are strong papers, which have more than one weakness. For example, a well-written paper with solid experiments, but incremental; a paper on a well studied problem with solid theory, but weak experiments; or a novel paper with good experiments, but poorly written. Weak Reject: these are papers that have some promise, but they would be better off by being revised and resubmitted. Strong Reject: these are papers that have major flaws, or have been done before.	Weak Reject
Preliminary Evaluation. Please indicate to the AC, your fellow reviewers, and the authors your current opinion on the paper. Please summarize the key things you would like the authors to include in their	I would like to have anwers to the following questions - What is the 'accuracy of the calibration' mentioned in line 674? - How do the results change, if the 3D points are not forced to lie on a plane?

rebuttals to facilitate your decision making. There is no need to summarize the paper.	
Confidence. Write "Very Confident" to stress that you are absolutely sure about your conclusions (e.g., you are an expert who works in the paper's area), "Confident" to stress that you are mostly sure about your conclusions (e.g., you are not an expert but can distinguish good work from bad work in that area), and "Not Confident" to stress that that you feel some doubt about your conclusions. In the latter case, please provide details as confidential comments to PC/AC chairs (point 7.).	Confident