R6P - Rolling Shutter Absolute Pose Problem

Cenek Albl[1], Zuzana Kukelova[2], Tomas Pajdla[1]
[1]Czech Technical University in Prague, Faculty of Electrical Engineering, Center for Machine Perception [2]Microsoft Research Cambridge

Perspective-n-point problem (PnP) for calibrated cameras is the task of finding a camera orientation and translation from n 2D-to-3D correspondences. It is a key problem in many computer vision applications such as structure from motion, camera localization, object localization and visual odometry.

Standard PnP methods assume a perspective camera model which is a model physically valid for cameras with a global shutter. However, CMOS sensors that are used in vast majority of todays consumer cameras, smartphones etcùse the rolling shutter (RS) mechanism to capture images. The key difference is that with the global shutter, the entire image is exposed to the light at once, whereas when using the RS the individual image rows (or columns) are captured at different times. When a RS camera moves while capturing the image, several types of distortion such as smear, skew or wobble appear.

Recent works have shown that RS is an important effect that should be considered in image rectification [2, 4], structure from motion [1] and multiple view stereo [5]. Those works have shown that existing methods can perform poorly on RS data or even fail completely and that incorporating some sort of RS camera model can solve these issues.

We present the first non-iterative minimal solution to the rolling shutter absolute pose (RnP) problem which improves the precision of camera pose estimates of a standard P3P solution and works on still images as well as video sequences. We investigate several rolling shutter camera models, discuss and verify their feasibility for a minimal solution to the RS abolute pose problem. We describe how to prepare the equations of the model to be solved by a polynomial solver [3]. The selected RS camera model contains a linearized camera orientation and we present a possible method how to keep the data close to the linearization point where the model works well.

Standard PnP for perspective cameras uses the projection function

$$\lambda_i \mathtt{x}_i = \mathtt{R}\mathtt{X}_i + \mathtt{C} \tag{1}$$

where $\mathtt{R}$ and $\mathtt{C}$ is the rotation and translation bringing a 3D point $\mathtt{X}_i$ from world coordinate system to the camera coordinate system with $\mathtt{x}_i = [r_i, c_i, 1]^\top$, and scalar $\lambda_i \in \mathbb{R}$. With RS cameras, when the camera is moving during the image capture, every image row will be captured at different time and hence different positions. $\mathtt{R}$ and $\mathtt{C}$ will therefore be functions of the image row r_i being captured.

$$\lambda_i \mathtt{x}_i = \begin{bmatrix} r_i \\ c_i \\ 1 \end{bmatrix} = \mathtt{R}(r_i)\mathtt{X}_i + \mathtt{C}(r_i) \tag{2}$$

We propose several possibilities of how to describe functions $\mathtt{R}(r_i)$ and $\mathtt{C}(r_i)$. The following model has been chosen since we found it feasible for a polynomial solver. The rotations have been linearized using first order Taylor expansion, resulting in

$$\lambda_i \begin{bmatrix} r_i \\ c_i \\ 1 \end{bmatrix} = (\mathtt{I} + (r_i - r_0)[\mathtt{w}]_x)(\mathtt{I} + [\mathtt{v}]_x)\mathtt{X}_i + \mathtt{C} + (r_i - r_0)\mathtt{t} \tag{3}$$

which are polynomial equations of degree two and with 28 monomials.

The minimal number of 2D-to-3D point correspondences necessary to solve the absolute pose rolling shutter problem is six, because we estimate two rotations (each has 3 unknown parameters), camera center (3 parameters) and unknown translation (3 parameters) and each 2D-3D correspondence gives us two constraints.

This results in quite a complex system of $3 \times 6 = 18$ equations in 18 unknowns. Such a system is not easy to solve for the Gröbner basis method and therefore it has to be simplified. In the paper we show how to reduce the problem to solving a system of 6 equations in 6 unknowns and with 16 monomials.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

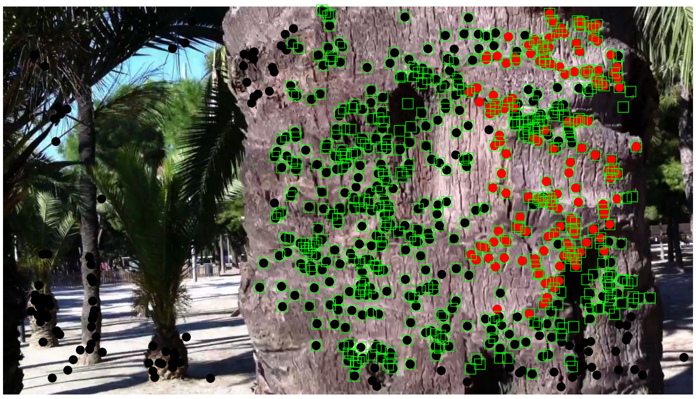

Figure 1: Result of a standard P3P and our R6P algorithm applied on image with rolling shutter distortion. All tentative 2D-3D correspondences are shown in black, inliers found by P3P are red and inliers found by R6P are in green squares. Notice that R6P found many more matches than P3P.

The model has an obvious drawback, that the camera orientation $\mathtt{R}(\mathtt{v})$ can be arbitrary and the linearization is only accurate enough close to $\mathtt{R}(\mathtt{v}) = \mathtt{I}$. If the user is provided with an approximate camera orientation (e.g. from an Inertial Measurement Unit) it is easy to transform the data such that $\mathtt{v}$ is small. In the paper we propose to use the standard P3P algorithm to find a coarse estimate of $\mathtt{R}(\mathtt{v})$ which allows subsequent application of R6P that improves the result.

The experiments show a significant improvement in terms of camera pose accuracy over standard P3P. The P3P has been shown to bring the camera orientation error under 6 degrees when applied on rolling shutter images with camera rotation up to 30 degrees per frame.

This P3P initial orientation is enough for the R6P to reduce the error in orientation under 0.5 degree and relative camera center error under 2%. We demonstrated on real data that R6 is able to correctly classify significantly more inlier matches and is therefore an interesting choice for Structure from Motion applications.

[1] Johan Hedborg, Per-Erik Forssén, Michael Felsberg, and Erik Ringaby. Rolling shutter bundle adjustment. In *CVPR*, pages 1434–1441, 2012.

[2] Chao Jia and Brian L. Evans. Probabilistic 3-d motion estimation for rolling shutter video rectification from visual and inertial measurements. In *MMSP*, pages 203–208. IEEE, 2012.

[3] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Automatic generator of minimal problem solvers. In *Computer Vision - ECCV 2008, 10th European Conference on Computer Vision, Proceedings, Part III*, pages 302–315, 2008. ISBN 978-3-540-88689-1.

[4] Erik Ringaby and Per-Erik Forssén. Efficient video rectification and stabilisation for cell-phones. *International Journal of Computer Vision*, 96(3):335–352, 2012. ISSN 0920-5691.

[5] O. Saurer, K. Koser, J.-Y. Bouguet, and M. Pollefeys. Rolling shutter stereo. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 465–472, Dec 2013.