A GENERIC CAMERA CALIBRATION METHOD FOR CONVENTIONAL AND FISH-EYE LENSES BY MINIMIZING OBJECT SPACE ERROR

Panu Srestasathiern , Narut Soontranon

Geo-Informatics and Space Technology Development Agency (GISTDA), Laksi, Bangkok, Thailand - (panu,narut)@gistda.or.th

KEY WORDS: Camera calibration, fish-eye lenses, conventional camera, generic, object space error

ABSTRACT:

With the advancement of sensor manufacturing technology, several kinds of optical cameras such as omnidirectional camera, fish-eye lenses, wide-angle or conventional lenses are more affordable and compact. Therefore, many of mobile platforms i.e. UAV or mobile mapping vehicle are equipped with those optical cameras in order to acquire high information content. To analyse the obtained image, the calibration of camera model is highly required especially in the measurement or mapping applications. In this paper, a generic camera calibration technique for conventional camera and fish-eye lenses is proposed. The proposed method requires only the images of calibration pattern i.e. checker board taken from different poses. The generic backward projection model i.e. polynomial model is employed in this paper. However, other camera models can be used with the proposed technique. The estimation error used in this work is measured in object (as opposed to image) space. Such object space error is called *object space collinearity error*. The advantage of using object space error is that it is more meaningful than the image space error especially in the case of omnidirectional camera or fish-eye lenses.

The calibration process begins with the initializing the interior orientation parameters of the camera. The interior orientation parameters include the affine transformation parameters and the distortion model parameters. The affine transformation maps a point from the image plane to the sensor plane by assuming that the optical axis of the camera is perpendicular to the sensor plane. The distortion model defines the mapping from the 3D camera coordinate system to the sensor plane or vice versa. These parameters can be initialized by the nominal focal length, the maximum field of view and the pixel dimension provided by manufacturer.

Once the initial interior orientation parameters are obtained, the exterior orientation parameters are then initialized. As it is already mentioned, the object space error used in this paper is the object space collinearity error which is formulated as:

$$\mathbf{e} = \left(\mathtt{I} - \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}}\right)(\mathtt{R}\mathbf{p} + \mathbf{T}) \tag{1}$$

where $\mathbf{p}$ is a point in object space, which is a point on the calibration plane, $\mathbf{v}$ the direction of the back-projection ray from the corresponding pixel, $\mathtt{R}$ and $\mathbf{T}$ the rotation and translation of the rigid transformation from object space to camera space. Given the initial interior orientation parameters, the term $(\mathtt{I} - \frac{\mathbf{v}\mathbf{v}^\top}{\mathbf{v}^\top\mathbf{v}})$ is then fixed. The globally convergent rotation and translation estimation can then be obtained.*

The obtained initial interior and exterior orientation parameters are then refined using the non-linear optimization technique. That is, object space collinearity error is minimized with respect to the interior and exterior orientation parameters. Note that sum of error over all points is the objective function of the optimization.

The performance of the proposed camera calibration method was evaluated using both synthetic and real dataset. The synthetic dataset was generate to test the performance of the proposed camera calibration as a function of the image noise level. That is, the result from the experiment with synthetic dataset is reported by the re-projection error versus the image noise level (pixel) varied from 0.5 to 3 pixels. The re-projection error can be computed by:

$$reprojection\ error = \|\mathbf{m} - K^{-1} \circ D^{-1}(\widehat{D} \circ \widehat{K}(\mathbf{m}))\|^2, \tag{2}$$

where $\mathbf{m}$ is an image point. $\widehat{K}$ and $\widehat{D}$ are the estimated affine transformation and back-projection model. respectively. K^{-1} and D^{-1} are the ground truth. For the real dataset, Sony NEX 5 and Ladybug 3 spherical camera was used to collect images. The performance of the proposed method on the real dataset is evaluated in terms of the object space collinearity error. The experimental result shows that the non-linear estimation significantly improvements the object space collinearity error compared with that of initial camera calibration parameters.

*Chien-Ping Lu, Gregory D. Hager, and Eric Mjolsness. 2000. *Fast and Globally Convergent Pose Estimation from Video Images*. IEEE Trans. Pattern Anal. Mach. Intell. 22, 6 (June 2000), 610-622.