Supplemental Material with Toward Coherent Object Detection And Scene Layout Understanding in CVPR2010

Sid Y. Bao

July 21, 2010

yingze@umich.edu

Version 1.0

Proposition: Eq. 1 admits one or at most two non-trivial solution of $\{f, n_1, n_2, n_3\}$ if at least three non-aligned observations (u_i, v_i) (i.e. non-collinear in the image) are available. If the observations are collinear, then Eq. 1 has infinite solutions.

$$\begin{bmatrix} u_1 & v_1 & f \\ u_2 & v_2 & f \\ u_3 & v_3 & f \\ & \vdots & \\ u_n & v_n & f \end{bmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} -\cos\phi_1 \sqrt{u_1^2 + v_1^2 + f^2} \\ -\cos\phi_2 \sqrt{u_2^2 + v_2^2 + f^2} \\ -\cos\phi_3 \sqrt{u_3^2 + v_3^2 + f^2} \\ \vdots \\ -\cos\phi_n \sqrt{u_n^2 + v_n^2 + f^2} \end{pmatrix} \tag{1}$$

Proof: Eq.1 must have solution since it is deducted from real meaning. Under the given condition, assuming three objects located at $\{u_i, v_i;\ i = 1...3\}$ are not aligned in a line in a image plane. Then we have

$$\begin{bmatrix} u_1 & v_1 & f \\ u_2 & v_2 & f \\ u_3 & v_3 & f \end{bmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} -\cos\phi_1 \sqrt{u_1^2 + v_1^2 + f^2} \\ -\cos\phi_2 \sqrt{u_2^2 + v_2^2 + f^2} \\ -\cos\phi_3 \sqrt{u_3^2 + v_3^2 + f^2} \end{pmatrix} \tag{2}$$

If objects are not aligned in a line in a image plane, then the rank of the left matrix of the left-hand part of Eq. 2 is equal to 3. We can represent $\{n_1, n_2, n_3\}$ each as a function of f. Utilizing the additional constraint, $\|n\| = 1$, we obtain an Eq. of f where only f^2, f^4, and known values appear. Therefore, at most two real positive f can be the solutions. Given f, $\{n_1, n_2, n_3\}$ is uniquely determined. Consequently, at most two solutions of Eq.1 exists. Furthermore, these solutions must be the solutions of Eq.1, when we have more than three measures $\{u, v\}$, otherwise it contradicts the premise that Eq.1 must have solutions.

On the other hand, If (u_i, v_i) are aligned in a line of the image plane, then more than one solutions of Eq.1 exist.

If (u_i, v_i) are aligned in a line in the image plane, then the rank of the left matrix of the left-hand part of Eq. 1 is equal to 2. Arbitrarily let $(u_1, v_1) \neq 0$, then using Gaussian elimination, Eq. 1 will be equivalent to the following form

$$\begin{bmatrix} \alpha & \beta & f \\ \gamma & \epsilon & 0 \\ 0 & 0 & 0 \\ & \vdots & \end{bmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} = \begin{pmatrix} \zeta \\ \eta \\ 0 \\ \vdots \end{pmatrix} \tag{3}$$

If $\widehat{f}, \widehat{n1}, \widehat{n2}, \widehat{n3}$ are the true values, then $\widehat{f}, \widehat{n1} + km_1, \widehat{n2} + km_2, \widehat{n3} + km_3$ is also a solution of Eq. 3, where (m_1, m_2, m_3) is non-trivial solution of

$$\begin{bmatrix} \alpha & \beta & f \\ \gamma & \epsilon & 0 \end{bmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = 0$$

Eq. 1 guarantees that as long as at least three objects do not lie on same line in the image, is possible to estimate camera focal length and supporting planes normal from object locations and zenith pose measurements in the image. Notice that this Eq. does not assume all objects are placed on one unique plane and it does not require there is no in-plane rotation of camera. Such estimates can be used to measure the probability to observe $n > 2$ objects in the image (at certain locations and poses in the image) given a geometrical configuration (defined by f $\{f, n_1, n_2, n_3\}$); or conversely, the probability of having a certain geometrical configuration given at least three object detections (and relevant measurements).