Ultra-Precision 6DoF Pose Estimation Using 2-D Interpolated Discrete Fourier Transform

Supplementary Material

A. Detailed Model Formulations

A.1. Full Camera Model Derivation

The standard pinhole camera model is mathematically represented as

$$s\begin{bmatrix}x_i\\y_i\\1\end{bmatrix}=\begin{bmatrix}f&0&0\\0&f&0\\0&0&1\end{bmatrix}\begin{bmatrix}\boldsymbol{R}&\boldsymbol{t}\end{bmatrix}\begin{bmatrix}x_w\\y_w\\z_w\\1\end{bmatrix}\tag{S1}$$

Here $\left[x_w,y_w,z_w,1\right]^T$ and $\left[x_i,y_i,1\right]^T$ denote the homogeneous coordinates of a point in the world coordinate frame and image coordinate frame, respectively. f represents the focal length of the camera and s is a scale factor. $\boldsymbol{R}$ and $\boldsymbol{t}$ are the rotation and translation matrices respectively that define the transformation from the world coordinate frame to the camera coordinate frame. They can be expressed as

$$\begin{aligned}\boldsymbol{R}&=\boldsymbol{R}_z(\alpha)\,\boldsymbol{R}_y(\beta)\,\boldsymbol{R}_x(\gamma)\\ \boldsymbol{t}&=\left[t_x,t_y,t_z\right]^T\end{aligned}\tag{S2}$$

where α, β, and γ are the rotation angles about the z, y, and x axes and t_x, t_y, and t_z are the translations along the x, y, and z axes, respectively.

Assuming the pattern plane coincides with the xy plane of the world coordinate frame (i.e., $z_w=0$), and letting $\boldsymbol{r}_j$ denote the j-th column vector of $\boldsymbol{R}$, Eq. (S1) becomes

$$s\begin{bmatrix}x_i\\y_i\\1\end{bmatrix}=\begin{bmatrix}f&0&0\\0&f&0\\0&0&1\end{bmatrix}\begin{bmatrix}\boldsymbol{r_1}&\boldsymbol{r_2}&\boldsymbol{t}\end{bmatrix}\begin{bmatrix}x_w\\y_w\\1\end{bmatrix}\tag{S3}$$

For $z_w=0$, $s=t_z-x_w\sin\beta+y_w\cos\beta\sin\gamma$.

A.2. Pattern Model Fourier Expansions

Checkerboard and grid patterns are superpositions of sinusoidal components. The spatial representation of a checkerboard pattern in the world coordinate frame is defined as a piecewise function:

$$g(x_w,y_w)=\begin{cases}A, 0\le\frac{2\pi x_w}{T_x}<\pi, 0\le\frac{2\pi y_w}{T_y}<\pi\\ A, -\pi\le\frac{2\pi y_w}{T_y}<0, -\pi\le\frac{2\pi y_w}{T_y}<0\\ 0, \text{otherwise}\end{cases}\tag{S4}$$

The Fourier series expansion of $g(x_w,y_w)$ is given by

$$\begin{aligned}g\left(x_w,y_w\right)=&\frac{A}{2}+\frac{4A}{\pi^2}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{1}{mn}\left[\cos\left(2\pi n\frac{x_w}{T_x}-2\pi m\frac{y_w}{T_y}\right)\right.\\&\left.+\cos\left(2\pi n\frac{x_w}{T_x}+2\pi m\frac{y_w}{T_y}+\pi\right)\right]\end{aligned}\tag{S5}$$

where summations are performed on odd positive integers m and n.

Similarly, the spatial representation of a grid pattern in the world coordinate frame is

$$g(x_w,y_w)=\begin{cases}A, 0\le\frac{2\pi x_w}{T_x}<\pi, 0\le\frac{2\pi y_w}{T_y}<\pi\\ 0, \text{otherwise}\end{cases}\tag{S6}$$

The corresponding Fourier series expansion is

$$\begin{aligned}g\left(x_w,y_w\right)=&\frac{A}{4}+\frac{A}{\pi}\sum_{m=1}^{\infty}\frac{\sin\left(2\pi m\frac{y_w}{T_y}\right)}{m}+\frac{A}{\pi}\sum_{n=1}^{\infty}\frac{\sin\left(2\pi n\frac{x_w}{T_x}\right)}{n}\\&+\frac{2A}{\pi^2}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\cos\left(2\pi n\frac{x_w}{T_x}-2\pi m\frac{y_w}{T_y}\right)-\cos\left(2\pi n\frac{x_w}{T_x}+2\pi m\frac{y_w}{T_y}\right)}{mn}\end{aligned}\tag{S7}$$

where the summations in the final term are again performed on odd positive integers m and n.

B. Pose Estimation Model Derivation

B.1. Quasi-Orthographic Projection Model Derivation

Starting from Eq. (S3) and expanding s, we obtain

$$x_i=f\frac{x_w r_{11}+y_w r_{12}+t_x}{-x_w\sin\beta+y_w\cos\beta\sin\gamma+t_z}\tag{S8}$$

$$y_i=f\frac{x_w r_{21}+y_w r_{22}+t_y}{-x_w\sin\beta+y_w\cos\beta\sin\gamma+t_z}\tag{S9}$$

where

$$\begin{aligned}r_{11}&=\cos\alpha\cos\beta\\r_{12}&=\cos\alpha\sin\beta\sin\gamma-\sin\alpha\cos\gamma\\r_{21}&=\sin\alpha\cos\beta\\r_{22}&=\sin\alpha\sin\beta\sin\gamma+\cos\alpha\cos\gamma\end{aligned}\tag{S10}$$

Under the quasi-orthographic assumption, $t_z\gg x_w\sin\beta$ and $t_z\gg y_w\cos\beta\sin\gamma$. Thus, the denominator is $\approx t_z$. This leads to

$$\begin{aligned}\frac{x_i}{f}&=\frac{x_w\cos\alpha\cos\beta+y_w\left(\cos\alpha\sin\beta\sin\gamma-\sin\alpha\cos\gamma\right)+t_x}{t_z}\\ \frac{y_i}{f}&=\frac{x_w\sin\alpha\cos\beta+y_w\left(\sin\alpha\sin\beta\sin\gamma+\cos\alpha\cos\gamma\right)+t_y}{t_z}\end{aligned}\tag{S11}$$

which is Eq. (3) in the main paper.

B.2. Relationships between Pattern Parameters and 6DoF Pose

Substituting Eq. (S11) into Eq. (2) yields the pattern image expression in Eq. (4). The full expressions for $\omega_{x_1}, \omega_{y_1}, \omega_{x_2}, \omega_{y_2}, \varphi_1$, and φ_2 are

$$\omega_{x_1} = 2\pi \frac{t_z}{f}\left(\frac{\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma}{T_x\cos\beta\cos\gamma} - \frac{\sin\alpha}{T_y\cos\gamma}\right) \tag{S12a}$$

$$\omega_{y_1} = 2\pi \frac{t_z}{f}\left(\frac{\sin\alpha\cos\gamma - \cos\alpha\sin\beta\sin\gamma}{T_x\cos\beta\cos\gamma} + \frac{\cos\alpha}{T_y\cos\gamma}\right) \tag{S12b}$$

$$\omega_{x_2} = 2\pi \frac{t_z}{f}\left(\frac{\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma}{T_x\cos\beta\cos\gamma} + \frac{\sin\alpha}{T_y\cos\gamma}\right) \tag{S12c}$$

$$\omega_{y_2} = 2\pi \frac{t_z}{f}\left(\frac{\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma}{T_x\cos\beta\cos\gamma} + \frac{\cos\alpha}{T_y\cos\gamma}\right) \tag{S12d}$$

$$\varphi_1 = 2\pi\left(\frac{-t_x\left(\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma\right)}{T_x\cos\beta\cos\gamma} + \frac{t_y\left(\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma\right)}{T_x\cos\beta\cos\gamma} + \frac{t_x\sin\alpha - t_y\cos\alpha}{T_y\cos\gamma}\right) \tag{S12e}$$

$$\varphi_2 = 2\pi\left(\frac{-t_x\left(\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma\right)}{T_x\cos\beta\cos\gamma} + \frac{t_y\left(\cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma\right)}{T_x\cos\beta\cos\gamma} - \frac{t_x\sin\alpha - t_y\cos\alpha}{T_y\cos\gamma}\right) \tag{S12f}$$

B.3. Solution of Parameter η

The constraint equation is

$$\left(\omega_{x_1} + \omega_{x_2}\right)^2 + \left(\omega_{y_2} - \omega_{y_1}\right)^2 = \frac{\eta^2\left(1 - \cos^2\beta\sin^2\gamma\right)}{T_x^2\cos^2\beta\cos^2\gamma} \tag{S13}$$

Substituting expressions for $\cos\beta$ and $\cos\gamma$ (from Eq. (6)) into Eq. (S13) leads to the quadratic equation in η^2

$$\rho_1\eta^4 - \rho_2\eta^2 + 1 = 0 \tag{S14}$$

where

$$\begin{aligned}
\rho_1 &= \frac{1}{T_x^2T_y^2\left(\omega_{y_1} + \omega_{y_2}\right)^2\left(\omega_{x_1} + \omega_{x_2} - \left(\omega_{y_2} - \omega_{y_1}\right)\tan\alpha\right)^2} \\
\rho_2 &= \frac{\left(\omega_{x_1} + \omega_{x_2}\right)^2 + \left(\omega_{y_2} - \omega_{y_1}\right)^2}{T_y^2\left(\omega_{y_1} + \omega_{y_2}\right)^2\left(\omega_{x_1} + \omega_{x_2} - \left(\omega_{y_2} - \omega_{y_1}\right)\tan\alpha\right)^2} \\
&+ \frac{1}{T_x^2\left(\left(\omega_{x_1} + \omega_{x_2}\right)\cos\alpha - \left(\omega_{y_2} - \omega_{y_1}\right)\sin\alpha\right)^2}
\end{aligned} \tag{S15}$$

Solving this quadratic equation gives

$$\eta = \sqrt{\frac{\rho_2 \pm \sqrt{\rho_2^2 - 4\rho_1}}{2\rho_1}} \tag{S16}$$

C. Pattern Parameter Estimation Algorithm (2D-IpDFT) Derivation

C.1. Window Function Details

The 2D maximum sidelobe decay window function $w(u, v)$ is separable: $w(u, v) = w_1(u)w_1(v)$, where $w_1(\cdot)$ is a 1D window. For a window of length L, $w_1(\cdot)$ has the form

$$w_1(x) = \sum_{h=0}^{H-1}(-1)^h a_h \cos\frac{2\pi xh}{L} \tag{S17}$$

where the coefficients a_h are

$$a_0 = \frac{C_{2H-2}^{H-1}}{2^{2H-2}},\ a_h = \frac{C_{2H-2}^{H-h-1}}{2^{2H-3}},\ h = 1, \dots, H-1 \tag{S18}$$

The DFT of $w\left(u, v\right)$ can be expressed as

$$W\left(k, l\right) = W\left(k\right)W\left(l\right) = \frac{D\left(k\right)}{P\left(k\right)}\frac{D\left(l\right)}{P\left(l\right)} \tag{S19}$$

where

$$\begin{aligned}
D\left(\cdot\right) &= \frac{M\left(2H-2\right)!}{\pi 2^{2H-2}}\sin\left(\pi\cdot\right)e^{-j\pi\cdot} \\
P\left(\cdot\right) &= \cdot\prod_{h=1}^{H-1}\left(h^2 - \cdot^2\right)
\end{aligned} \tag{S20}$$

Note that the function $D\left(\cdot\right)$ exhibits the property

$$D\left(\cdot - 1\right) = D\left(\cdot\right) = D\left(\cdot + 1\right) \tag{S21}$$

C.2. Full DFT of Windowed Image Signal

The DFT of the windowed image signal $g(u, v)$ is

$$\begin{aligned}
G\left(k, l\right) = &\frac{1}{2}\left[W\left(k - \lambda_1, l - \mu_1\right)e^{j\varphi_1} + W\left(k + \lambda_1, l + \mu_1\right)e^{-j\varphi_1}\right. \\
&\left. + W\left(k - \lambda_2, l + \mu_2\right)e^{j\varphi_2} + W\left(k + \lambda_2, l - \mu_2\right)e^{-j\varphi_2}\right]
\end{aligned} \tag{S22}$$

C.3. Linear System for Frequency Estimation

Rearranging Eq. (S22) and evaluating at DFT frequencies $k-1, k, k+1$ yields a system of linear equations $\boldsymbol{Ad} = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{bmatrix} \frac{1}{P(k-1-\lambda_1)} & \frac{1}{P(k-1+\lambda_1)} & G\left(k-1, l\right) \\ \frac{1}{P(k-\lambda_1)} & \frac{1}{P(k+\lambda_1)} & G\left(k, l\right) \\ \frac{1}{P(k+1-\lambda_1)} & \frac{1}{P(k+1+\lambda_1)} & G\left(k+1, l\right) \end{bmatrix} \tag{S23}$$

$$\boldsymbol{d} = \begin{bmatrix} \frac{D(k-\lambda_1)D(l-\mu_1)e^{j\varphi_1}}{P(l-\mu_1)} \\ \frac{D(k+\lambda_1)D(l+\mu_1)e^{-j\varphi_1}}{P(l+\mu_1)} \\ -2 \end{bmatrix} \tag{S24}$$

$$\boldsymbol{b} = \begin{bmatrix} -W\left(k-1-\lambda_2, l+\mu_2\right)e^{j\varphi_2} - W\left(k-1+\lambda_2, l-\mu_2\right)e^{-j\varphi_2} \\ -W\left(k-\lambda_2, l+\mu_2\right)e^{j\varphi_2} - W\left(k+\lambda_2, l-\mu_2\right)e^{-j\varphi_2} \\ -W\left(k+1-\lambda_2, l+\mu_2\right)e^{j\varphi_2} - W\left(k+1+\lambda_2, l-\mu_2\right)e^{-j\varphi_2} \end{bmatrix} \tag{S25}$$

The condition $\det(\boldsymbol{A}) = 0$ for a non-trivial solution of $\boldsymbol{d}$ (when $\boldsymbol{b}$ is considered part of the augmented matrix for homogeneous system or by Cramer's rule if solving for components of $\boldsymbol{d}$) leads to the frequency estimation formula Eq. (9).