Robust Manhattan Frame Estimation from a Single RGB-D Image (Supplementary Material)

Bernard Ghanem[1], Ali Thabet[1], Juan Carlos Niebles[2], and Fabian Caba Heilbron[1]
[1]King Abdullah University of Science and Technology (KAUST), Saudi Arabia
[2]Universidad del Norte, Colombia

1. Optimization Details

In this section, we elaborate on the details required to solve the optimization problems proposed in our paper, specifically the MFE (Manhattan Frame Estimation) and RMFE (Robust Manhattan Frame Estimation) problems. These details will help reproduce the results we achieve in the experimental section. For completeness, we reiterate both MFE and RMFE problems in Eqs (1) and (2). Remember that $\mathbf{N}$ is the set of all surface normals and possibly line directions extracted from the RGBD image.

$$(MFE): \quad \min_{\mathbf{R}\in \mathrm{SO}(3),\mathbf{X}} \ \frac{1}{2}\|\mathbf{RN}-\mathbf{X}\|_F^2 + \lambda\|\mathbf{X}\|_{1,1} \tag{1}$$

$$(RMFE): \quad \min_{\mathbf{R}\in \mathrm{SO}(3),\mathbf{X},\mathbf{E}} \ \|\mathbf{E}^T\|_{2,1} + \lambda\|\mathbf{X}\|_{1,1} \quad \text{subject to: } \mathbf{RN}=\mathbf{X}+\mathbf{E} \tag{2}$$

We also reiterate the optimization identities referred to in the paper in Eqs (3), (4), and (5).

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \ \frac{1}{2}\|\mathbf{X}-\mathbf{RN}\|_F^2 + \lambda\|\mathbf{X}\|_{1,1} = \mathcal{S}_\lambda(\mathbf{RN}) \tag{3}$$

$$\mathbf{R}^* = \arg\min_{\mathbf{R}\in \mathrm{SO}(3)} \ \|\mathbf{RN}-\mathbf{X}\|_F^2 = \mathcal{K}(\mathbf{N},\mathbf{X}) \tag{4}$$

$$\mathbf{E}^* = \arg\min_{\mathbf{Y}} \ \frac{1}{2}\|\mathbf{Y}-\mathbf{A}\|_F^2 + \tilde{\lambda}\|\mathbf{Y}\|_{2,1} = \mathcal{Q}_{\tilde{\lambda}}(\mathbf{A}) \tag{5}$$

Here, we have $\mathcal{S}_\lambda(\mathbf{A}_{ij})$ is the soft-thresholding operator on each element of matrix $\mathbf{A}$ and $\mathcal{Q}_{\tilde{\lambda}}(\mathbf{A}_i)$ operates on the i^{th} row of matrix $\mathbf{A}$.

$$\mathcal{S}_\lambda(\mathbf{A}_{ij}) = \mathrm{sign}(\mathbf{A}_{ij})\max(0,|\mathbf{A}_{ij}|-\lambda)$$

$$\mathcal{K}(\mathbf{N},\mathbf{X}) = \mathbf{U}\left[\mathrm{diag}(1,1,\mathrm{sign}(\det(\mathbf{XN}^T)))\right]\mathbf{V}^T$$

$$\mathcal{Q}_{\tilde{\lambda}}(\mathbf{A}_i) = \max\left(0, 1-\frac{\tilde{\lambda}}{\|\mathbf{A}_i\|_2}\right)\mathbf{A}_i$$

Solving the MFE Problem in Eq (1)

As mentioned in the paper, a local minimum of the MFE problem is achieved by alternating optimization, where the rotation matrix $\mathbf{R}$ is updated with the rotated normals $\mathbf{X}$ kept fixed and vice versa. Algorithm 1 shows the details of this optimization. Here, the initial rotation is taken to be identity.

Algorithm 1: *MF Estimation (Solving Eq (1))*

Input : $\mathbf{N}$ and λ
Output: $\mathbf{R}$ and $\mathbf{X}$

Initialize $\mathbf{R} = \mathbf{I}_3$
while *not converged* **do**
    Update $\mathbf{X}$ using Eq (3): $\mathbf{X} \leftarrow \mathcal{S}_\lambda(\mathbf{RN})$
    Update $\mathbf{R}$ using Eq (4): $\mathbf{R} \leftarrow \mathcal{K}(\mathbf{N}, \mathbf{X})$
end

Solving the RMFE Problem in Eq (2)

Similar to MFE, we use alternating optimization to achieve a local minimum for Eq (2). Updating $\mathbf{R}$ is done in the same way as in MFE. However, updating the other two variables $\mathbf{X}$ and $\mathbf{E}$ is done using the Inexact Augmented Lagrangian Multiplier (IALM) method. By introducing the augmented lagrange multiplier $\mathbf{\Lambda}$ to incorporate the equality constraint into the cost function, we obtain the augmented Lagrangian function in Eq (6) that we show, in what follows, can be optimized through a sequence of simple closed form update operations. Note that μ is a penalty coefficient.

$$L(\mathbf{X}, \mathbf{E}, \mathbf{\Lambda}) = \|\mathbf{E}^T\|_{2,1} + \lambda\|\mathbf{X}\|_{1,1} + tr\left(\mathbf{\Lambda}^T\left(\mathbf{RN} - \mathbf{X} - \mathbf{E}\right)\right) + \mu\|\mathbf{RN} - \mathbf{X} - \mathbf{E}\|_F^2 \tag{6}$$

In IALM, $\mathbf{X}$ is updated while the other primal variable $\mathbf{E}$ is kept fixed and vice versa. Updating each primal variable is done by minimizing $L(\mathbf{X}, \mathbf{E}, \mathbf{\Lambda})$ w.r.t. to one primal variable at a time. In fact, we update $\mathbf{X}$ in closed form according to Eq (7).

$$\begin{aligned}\mathbf{X}^* &= \arg\min_{\mathbf{X}} \mu\|\mathbf{X} - (\mathbf{RN} - \mathbf{E})\|_F^2 - tr\left(\mathbf{\Lambda}^T\mathbf{X}\right) + \lambda\|\mathbf{X}\|_{1,1} \\ &= \mathcal{S}_{\frac{\lambda}{2\mu}}\left(\mathbf{RN} - \mathbf{E} + \frac{1}{2\mu}\mathbf{\Lambda}\right)\end{aligned} \tag{7}$$

Similarly, the estimate of $\mathbf{E}$ can be updated in closed form according to Eq (8).

$$\begin{aligned}\mathbf{E}^* &= \arg\min_{\mathbf{E}} \mu\|\mathbf{E}^T - (\mathbf{RN} - \mathbf{X})^T\|_F^2 - tr\left(\mathbf{\Lambda}^T\mathbf{E}\right) + \|\mathbf{E}^T\|_{2,1} \\ &= \mathcal{Q}_{\frac{1}{2\mu}}^T\left(\mathbf{RN} - \mathbf{X} + \frac{1}{2\mu}\mathbf{\Lambda}\right)\end{aligned} \tag{8}$$

We summarize the complete algorithm for solving the RMFE problem in Algorithm 2. As mentioned in the paper, we reduce computation by initializing the rotation matrix $\mathbf{R}$ of the RMFE problem with the solution to the equivalent MFE problem denoted as $\mathbf{R}_{MFE}$. The parameter ρ increases the penalty coefficient μ in each iteration. In our experiments, we set $\rho = 1.1$. Convergence is achieved for each *while* loop when the relative change in both variables falls below a threshold $\epsilon = 10^{-4}$. Empirically, we find that this IALM algorithm is insensitive to a large range of ϵ values.

Algorithm 2: *Robust MF (Solving Eq (2))*

Input : $\mathbf{N}$, λ, ρ
Output: $\mathbf{R}$, $\mathbf{X}$, and $\mathbf{E}$

Initialize $\mathbf{R} = \mathbf{R}_{MFE}$ and $\mathbf{E} = \mathbf{0}$
while *not converged* **do**
  //IALM: Update $\mathbf{X}$ and $\mathbf{E}$ keeping $\mathbf{R}$ fixed
  while *not converged* **do**
    Update $\mathbf{X}$ using Eq (7): $\mathbf{X} \leftarrow \mathcal{S}_{\frac{\lambda}{2\mu}}\left(\mathbf{RN} - \mathbf{E} + \frac{1}{2\mu}\mathbf{\Lambda}\right)$
    Update $\mathbf{E}$ using Eq (8): $\mathbf{E} \leftarrow \mathcal{Q}^{T}_{\frac{1}{2\mu}}\left(\mathbf{RN} - \mathbf{X} + \frac{1}{2\mu}\mathbf{\Lambda}\right)$
    Update the Langrange multiplier: $\mathbf{\Lambda} \leftarrow \mathbf{\Lambda} + \mu\left(\mathbf{RN} - \mathbf{X} - \mathbf{E}\right)$
    Update the penalty coefficient: $\mu \leftarrow \rho\mu$
  end
  Update $\mathbf{R}$ using Eq (4): $\mathbf{R} \leftarrow \mathcal{K}(\mathbf{N}, \mathbf{X} + \mathbf{E})$
end

2. Repeatability

Given an aligned scene using ground truth values, we look at the repeatability of both our method and ES. We rotate every scene at discrete angles around each of the 3 axis independently, and look at the accuracy of each method's alignment under these new conditions. More specifically, for every ground truth aligned scene S_i, we rotate it from 0 to $75°$, in increments of $5°$, around each axis independently. We also randomize the sign of each rotation angle. We then evaluate every method on the new scene S_i'.

We use this analysis to check how well each technique works under different initial conditions. The results are presented in Figure 1. These results show how our method can reliablly estimate a scene's MF for initial rotations of up to $45°$ on each axis. In contrast, ES breaks at lower bounds, reaching $20°$ in some occasions. The plot in Figure 1 shows that our method is completely robust to perturbations up to a rotation angle of $45°$; beyond that point our RMFE algorithm will find it less costly to flip the axis by $90°$ as most of the normals are more aligned towards this orthogonal direction. ES also flips the axis beyond $45°$ rotations but it also induces large errors ($20° - 40°$) at smaller perturbations. These repeatability experiments show that our method is more reliable at estimating MFs even in the presence of difficult scene orientations.

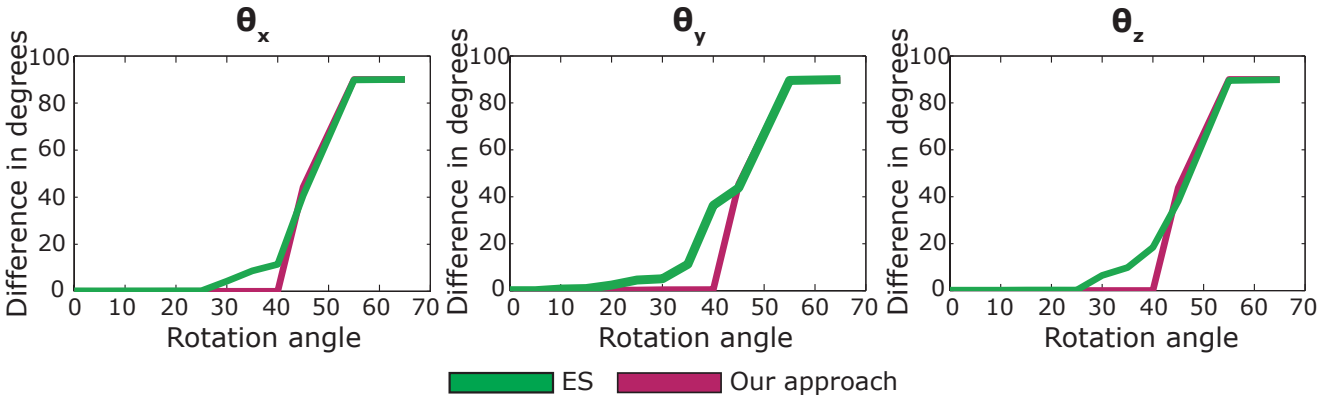

Figure 1: We align each scene using ground truth and then rotate around each axis independently by discrete amounts. We compute the performance of both our algorithms and ES, and plot the angle differences for θ_x, θ_y and θ_z at each rotation. We notice how our method confidently estimates the MF for a larger range of rotations. The break in our case happens at $45°$; we expect this as at that range the method will find it easier to flip the axis (and therefore we see the $90°$ error after the break. The same break occurs in ES but a much earlier stage, reaching sometimes as low as $20°$.