

Enhancing Rotation Averaging and Global Positioning by an Adaptive Robust Kernel

Chunqi Dai, Sagi Filin

Mapping and Geo-Information Engineering, Technion - Israel Institute of Technology, Haifa, Israel
chunqi.dai@campus.technion.ac.il, filin@technion.ac.il

Keywords: Structure-from-Motion, Motion Averaging, Rotation Averaging, Robust Optimization, Adaptive Kernel, Global Positioning.

Abstract

Motion averaging (MA) offers an efficient and mostly linear means for estimating image sets pose and provides a reliable initialization for large-scale structure-from-motion (SfM) pipelines. Nonetheless, MA, which comprises rotation and translation averaging (RA & TA, respectively), can be severely affected by the presence of outliers that can greatly degrade the performance of the entire optimization process or even lead to the divergence of the SfM solution. While robust loss functions have been applied to mitigate outliers effect, their performance depends heavily on the choice of parameters and requires manual tuning and prior knowledge of the residual distribution. To make MA more robust, we enhance it by incorporating an adaptive robust kernel that automatically adjusts its parameter to the residual distribution. This adaptive behavior balances robustness and sensitivity, removing the need for manual parameter tuning. In addition, to address the ill-posed nature of TA, we adopt a global positioning framework that jointly estimates the camera and 3D point positions. Experimental results show how the adaptive robust kernel consistently outperforms state-of-the-art fixed-parameter functions. It improves accuracy in both RA and global positioning, particularly in scenes with high levels of noise or outliers. These results demonstrate the effectiveness of adaptive robust kernels for improving the reliability and generalization of MA pipelines in challenging reconstruction scenarios.

1. INTRODUCTION

Structure from motion (SfM) is a foundational technique in computer vision and photogrammetry that reconstructs the pose and sparse 3D structure of an unordered image collection. Most modern pipelines adopt a global approach, in which rotations and translations are first solved between image pairs, and then estimated globally through motion averaging (MA). MA is typically decomposed into rotation and translation averaging (RA & TA, respectively, Hartley et al., 2013; Eriksson et al., 2019; Elnashef and Filin, 2022), where RA aims to estimate global camera orientations from a graph of pairwise relative rotations, and TA computes the positions of the camera based on these orientations and relative translation directions. TA per se is an ill-posed problem due to: *i*) the relative translation from two-view geometry lacking scale, *ii*) the sensitivity of relative translation estimates to small baselines, and *iii*) the degeneration of the reconstruction in the case of a collinear motion. Therefore, global positioning has been recently proposed to perform a joint estimation of the camera and point positions (Zhuang et al., 2018; Pan et al., 2024).

Current MA solutions perform well in many scenarios, but outliers which are common due to matching ambiguities, affect the estimated relative motion and significantly degrade their performance. To mitigate outlier influence, robust kernel functions are usually applied to down-weight their effect (Chatterjee and Govindu, 2018; Gao et al., 2021). Many kernels that have been proposed over the years, and selecting one, and its parameters, is crucial for the algorithm performance due to its dependence on the residual distributions. With the lack of prior knowledge about their distribution, it is difficult, in most scenarios, to choose an optimal kernel and set its parameters. The common solution is to resort to a trial-and-error until acceptable solution has been reached. Such approach translates to lengthy and cumbersome processes, which are not always feasible.

In this paper, we propose the introduction of a new robust kernel function to solve the MA problem. This function, first introduced in Barron (2019) and later adapted by Chebrolu et al. (2021), features adaptive properties that allow to relieve the need for parameter tuning and distribution setup. As the paper shows, the introduction of this kernel to both RA and global positioning, contributes to the robustness of the MA solution in the presence of outliers over varied residual distributions. By integrating this solution into the GLOMAP framework we also contribute to a more robust state-of-the-art SfM pipeline. Evaluation on the comprehensive and commonly used ETH3D dataset (Schops et al., 2017), demonstrates how we consistently outperform conventional fixed-parameter robust losses in angular and positional accuracy. As our results demonstrate, the adaptive kernel proves particularly effective in challenging scenes with high outlier rate, offering a robust and generalizable solution for MA in SfM.

2. RELATED WORK

2.1 Rotation Averaging and Global Positioning

Early RA methods, (e.g., Govindu, 2004) studied linearization of the problem based on quaternions and Lie algebra. Later, Martinec and Pajdla (2007) directly dropped the nonlinear constraints for tractability. Some relaxation methods, such as the semi-definite programming (SDP)-based relaxation (Carlone and Calafiore, 2018; Chen et al., 2021) and Lagrangian duality (Eriksson et al., 2019), were adopted to make the problem more tractable. Recent work has focused on improving robustness and accuracy. For example, Chatterjee and Govindu (2013) proposed a two-step solution using the l_1 loss and iterated re-weighted least squares (IRLS). Chatterjee and Govindu (2018) compared the performance of different robust functions, such as the l_1 , Huber, and Geman-McClure, in RA (Huber, 1992;

Geman and McClure, 1985). The results demonstrated that $l_{\frac{1}{2}}$ and Geman-McClure performed best, but while these methods mitigated the outliers impact, they relied on fixed-parameter loss functions, which may not generalize across datasets with different residual distributions. More recently, to further improve the robustness, Zhang et al. (2023) incorporated the uncertainty weighting based on measurement confidence, demonstrating improved resilience to noise.

TA usually follows the application of RA, but because of scale ambiguity and linear motion it may become more challenging. Wilson and Snavely (2014) proposed a stabilization by using the direction-based cost and adding camera-to-point constraints, named 1DSfM. An alternative TA solution was introduced by Manam and Govindu (2024), where direction and displacement-based methods were fused. To solve the camera positions more robustly, global positioning was recently proposed to simultaneously estimate the cameras and 3D points positions by measuring the direction from camera center to the points (Pan et al., 2024). As a cost function, the authors built upon the one proposed in Zhuang et al. (2018).

To address the effect of outliers on MA estimation, robust kernel functions have found popularity. Among them, the l_1 , Huber, and Geman-McClure loss functions are the widely used one in MA (Hartley et al., 2013; Sidhartha and Govindu, 2021). Nonetheless, they are unable to perform well in all situations, and choosing the suitable loss and adjusting the parameters by hand is required sometimes (Chebrolu et al., 2021). To handle that, Barron (2019) proposed a generalized loss function that covered several others, e.g., Cauchy, Geman-McClure, Welsch, Leclerc and Charbonnier (Black and Anandan, 1996; Dennis Jr and Welsch, 1978; Leclerc, 1989; Charbonnier et al., 1994). An advantage it has over others is its ability to adjust the parameter automatically and fit the data better than current loss functions. Notwithstanding its shape parameter is only defined for positive values, which means the range of kernels it can adapt to is not general. Hence, Chebrolu et al. (2021) proposed to truncate the kernel, so that the modified robust function can adapt to all ranges. We refer to their kernel as adaptive robust kernel in the following. Though the kernel was verified with the application in 3D registration and bundle adjustment, its performance in MA setups is unclear. Therefore, we integrate it into the linear form of RA and global positioning, to evaluate its performance.

2.2 SfM Framework

Over the years, two main paradigms were adopted to solve the SfM problem: incremental and global approaches. Incremental methods estimate the image set pose and 3D points sequentially, while global ones estimate them simultaneously. Incremental methods are considered more accurate as they are less sensitive to outliers, but are more computationally expensive. Representative incremental methods include the Bundler, VisualSfM, and COLMAP (Snavely et al., 2006; Wu, 2013; Schonberger and Frahm, 2016), where the latter has become the standard tool for 3D reconstruction due to its versatility and robustness in many datasets. Because global methods are more consistent, scalable, and efficient, more attention has been paid to them in recent years. Representatives include the OpenMVG, Theia, GlobalSfMpy, VGGsFM, and GLOMAP (Sweeney et al., 2015; Moulon et al., 2017; Zhang et al., 2023; Wang et al., 2024; Pan et al., 2024), where GlobalSfMpy and GLOMAP were developed based on COLMAP. GLOMAP has demonstrated better performance than OpenMVG, Theia and COLMAP in terms

of accuracy and robustness. Therefore, our robust MA solution has been integrated and evaluated within its framework.

3. Method

Given an image set, SfM estimates their pose and sparse 3D structure from pairwise image matches. We focus on improving the robustness of MA by enhancing both RA and global positioning (an alternative to TA) using an adaptive robust kernel.

3.1 Models of Rotation Averaging and Global Positioning

Rotation averaging Given a set of relative rotations R_{ij} between the cameras, the RA problem is to estimate the global rotations R_i and R_j that minimize the cost function:

$$\arg \min_{R_i, R_j} \sum_{i,j} \rho(d(R_{ij}, R_i R_j^T)), \quad (1)$$

where $\rho(\cdot)$ is the robust kernel function, and $d(\cdot)$ is a distance measure between two rotations. Denoting ω_{ij} , ω_i , and ω_j as the corresponding Lie algebra of R_{ij} , R_i , and R_j , the distance measure can be expressed as:

$$\arg \min_{\Delta\omega_j, \Delta\omega_i} \sum_{i,j} \rho(\|\Delta\omega_{ij} - (\Delta\omega_j - \Delta\omega_i)\|), \quad (2)$$

where Δ relates to updates of the ω values. Further, denoting the correction of the Lie algebra as $\Delta\omega_{rel} = \Delta\omega_{ij}$, $\Delta\omega_{glob} = [\omega_1, \dots, \omega_N]^T$, the cost function becomes:

$$\arg \min_{\Delta\omega_{glob}} \sum_{i,j} \rho(\|\Delta\omega_{rel} - A_{ij} \Delta\omega_{glob}\|), \quad (3)$$

where $A_{ij} = [\dots - I \dots I \dots]$.

Global positioning Given a set of relative directions v_{ik} , global positioning aims to minimize the difference between the actual image rays and the measurements (Fig. 1). The measurements v_{ik} can be obtained by projecting the pixel coordinates p_{ik} to the camera frame, then transforming the direction to the world frame by the global rotations obtained through RA. The camera center position c_i and the 3D point P_k can be estimated by minimizing the cost function:

$$\arg \min_{C, P, d} \sum_{i,k} \rho(\|v_{ik} - d_{ik}(P_k - c_i)\|), \quad (4)$$

where d_{ik} is a slack variable, acting as a normalization factor.

3.2 Adaptive Robust Kernel Function

Traditional robust loss functions such as Huber and Geman-McClure require fixed shape parameters, which are often manually tuned and may not generalize well across scenes. To address this, we adopt an adaptive kernel (Chebrolu et al., 2021), which generalizes many classical robust losses through a learnable shape parameter α . The adaptive robust kernel function is defined as:

$$L(\alpha, c) = - \sum_{i=1}^N \log \frac{e^{-\rho(r_i, \alpha, c)}}{cZ(\alpha)} = \sum_{i=1}^N (\rho(r_i, \alpha, c) + \log cZ(\alpha)), \quad (5)$$

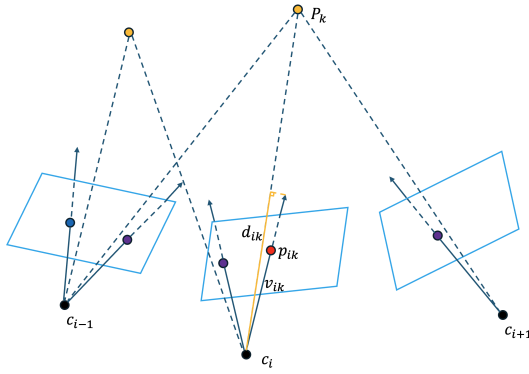


Figure 1. Global positioning. Yellow points are 3D object-space points. Red, blue, and purple points are their projection. Dashed lines represent the actual image rays, while solid lines denote the direction of measurements.

where r is the residual, α and c (> 0) are the shape and scale parameter that control the robustness of the kernel,

$$\rho(r_i, \alpha, c) = \frac{|\alpha - 2|}{\alpha} \left(\left(\frac{(r_i/c)^2}{|\alpha - 2|} + 1 \right)^{\alpha/2} - 1 \right), \quad (6)$$

and

$$Z(\alpha) = \int_{-\tau}^{\tau} e^{-\rho(r, \alpha, 1)} dr. \quad (7)$$

When $\alpha = 2$, the loss becomes quadratic; when $\alpha = 0$, it approximates the Cauchy loss; when $\alpha = -2$, it becomes the Geman-McClure loss; and as $\alpha \rightarrow -\infty$, it approaches the Welsch loss. Fig. 2 plots $\rho(r, \alpha, 1)$ for different α values. Note that smaller α yields smaller $\rho(r, \alpha, c)$, but the inliers may get penalized if the shape parameter is too small. Addressing that, $\log cZ(\alpha)$ acts as the penalty term to solve this matter. Nonetheless, it is tricky to jointly learn both the shape and scale parameters because they contribute to the cost function together. They influence each other and it is unclear whether a balanced state can be reached. So, the scale parameter is fixed according to the measurement noise while the shape parameter is learned by the distribution of residuals. Because the shape parameter is fixed in the optimization process, the adaptive kernel can be further written as:

$$L(\alpha) = \rho(r, \alpha) + \log Z(\alpha), \quad (8)$$

3.3 Robust Motion Averaging with An Adaptive Kernel

With the adaptive robust kernel function, the cost function of RA and global positioning can be expressed as:

$$\min \sum L(\|r(\Delta\omega)\|) = \min \sum (\rho(\|r(\Delta\omega)\|, \alpha) + \log Z(\alpha)), \quad (9)$$

$$\min \sum L(\|r(\Delta P)\|) = \min \sum (\rho(\|r(\Delta P)\|, \alpha) + \log Z(\alpha)), \quad (10)$$

where $r(\Delta\omega) = \Delta\omega_{rel} - A_{ij}\Delta\omega_{global}$, $r(\Delta P) = v_{ik} - d_{ik}(P_k - c_i)$. Because of their same form, we use the RA as an example to demonstrate the optimization. Taking the derivatives of the cost function with respect to the global rotations and

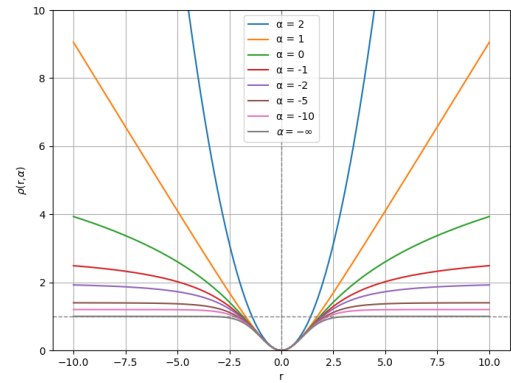


Figure 2. The plot of $\rho(r, \alpha, 1)$ for different α values.

the shape parameter, we have:

$$\frac{\partial L}{\partial \Delta\omega} = \sum \frac{\partial \rho}{\partial \Delta\omega} = \sum \frac{\partial \rho}{\partial r} \frac{\partial r}{\partial \Delta\omega} = \frac{1}{2} \sum \left(\frac{\partial \rho}{\partial r} \right) \frac{\partial r^2}{\partial \Delta\omega} \quad (11)$$

where, for simplification, r means $r(\Delta\omega)$. Now, defining $\Phi(\|r\|) = \frac{\partial \rho}{\partial r} / r$ as the weight function, we can write

$$\frac{\partial L}{\partial \Delta\omega} = \frac{1}{2} \sum \Phi(\|r\|) \frac{\partial r^2}{\partial \Delta\omega}. \quad (12)$$

As for the derivative of the shape parameter, we have:

$$\frac{\partial L}{\partial \alpha} = \sum \left(\frac{\partial \rho}{\partial \alpha} + \frac{\partial \log Z}{\partial \alpha} \right). \quad (13)$$

So, the optimal solution is obtained when

$$\frac{\partial L}{\partial \Delta\omega} = 0, \frac{\partial L}{\partial \alpha} = 0. \quad (14)$$

which is equivalent to the following optimization problem, if we treat $\Phi(\|r\|)$ as a constant during each iteration:

$$\min L'(\alpha) = \min_{\Delta\omega, \alpha} \sum \Phi(\|r\|) r^2 + 2 \log Z(\alpha), \quad (15)$$

We now deduce the equivalence between them.

$$\frac{\partial L'}{\partial \Delta\omega} = \sum \Phi(\|r\|) \frac{\partial r^2}{\partial \Delta\omega}, \quad (16)$$

$$\begin{aligned} \frac{\partial L'}{\partial \alpha} &= \sum \left(\frac{\partial (\Phi(\|r\|) r^2)}{\partial \alpha} + 2 \frac{\partial \log Z}{\partial \alpha} \right) \\ &= \sum \left(\frac{\partial (\Phi(\|r\|) r^2)}{\partial r} \frac{\partial r}{\partial \alpha} + 2 \frac{\partial \log Z}{\partial \alpha} \right) \\ &= \sum \left(\Phi(\|r\|) 2r \frac{\partial r}{\partial \alpha} + 2 \frac{\partial \log Z}{\partial \alpha} \right) \\ &= 2 \sum \left(\frac{\partial \rho}{\partial \alpha} + \frac{\partial \log Z}{\partial \alpha} \right). \end{aligned} \quad (17)$$

The derivatives are the same as the original cost function, hence their equivalence. The algorithm for robust rotation averaging is summarized in Algorithm 1.

Algorithm 1 Robust Rotation Averaging with Adaptive Kernel

Input: Relative rotations R_{ij} .

Output: Global rotations R_i , R_j , shape parameter α .

- 1: Initialize the global rotations R_i , R_j and the shape parameter α .
- 2: Compute the residuals r .
- 3: Find the optimal α to minimize Eq. (9) with Ceres.
- 4: Calculate the weight function $\Phi(\|r\|)$.
- 5: Solve the weighted least squares problem.
- 6: Update the global rotations.
- 7: Repeat steps 2-6 until convergence.

4. Experiments

We test our proposed method following its integration into the GLOMAP library. There, the Geman-McClure is used for the RA while the global positioning uses the Huber robust loss function. For fair comparison we implement the Huber and Geman-McClure functions for both RA and global positioning. Thus, the performance of these three kernels can be compared both in RA and global positioning. For RA, the adaptive kernel is implemented within the GLOMAP pipeline using IRLS and Ceres Solver (Agarwal et al., 2012). We use Ceres to calculate the optimal α and then complete the optimization with IRLS. For global positioning, we use the weight function $\Phi(\|r\|)$ to weight the residual in Ceres and add the penalty term through an additional residual block. The shape parameter α is initialized to 1.0 and updated at each IRLS iteration while minimizing the residuals. The integral term is approximated by cubic spline interpolation in the optimization process. Therefore, it does not increase the computational cost by much.

The ETH3D dataset (Schonberger and Frahm, 2016) comprises of 25 high-resolution multi-view indoor and outdoor scenes (including the training and test data) with images, keypoints, LiDAR depth, and millimeter-accurate groundtruth poses. Because GLOMAP fails to solve the scene *exhibition_hall* (Pan et al., 2024), we use the other 24 scenes for validation. COLMAP is used to extract and match the features, while the relative rotations and translations are computed by PoseLib (Larsson and contributors, 2020). We set the integral interval to be $\tau = 10c$, according to the practical implementation of the adaptive robust kernel. For all evaluations, we report the area under the recall curve (AUC) scores calculated from the camera rotation and position error after globally aligning with the ground truth using a robust RANSAC scheme (Schonberger and Frahm, 2016).

4.1 Evaluation of Rotation Averaging

The results for the different robust functions are listed in Table 1. Notably, Geman-McClure loss has been demonstrated to perform well in RA (Chatterjee and Govindu, 2018). It has better accuracy than Huber loss and achieves acceptable accuracy. Nonetheless, the adaptive kernel outperforms it in most scenes. In particular, the adaptive kernel achieves the highest AUC scores in scenes with more outliers or complex structures, such as *botanical_garden*, *statue*, and *terrains*. The new kernel also achieves better performance in simpler scenes, where all kernels perform well (c.f., *lounge*), suggesting strong adaptability to different residual distributions. This confirms that learning the shape parameter from the residuals leads to better robustness without sacrificing inlier sensitivity. Notably, under the strictest threshold (AUC@1°), the adaptive kernel achieves the highest average score (46.72), compared to

Geman-McClure (45.10) and Huber (34.77). However, the performance drops in scenes like *office*, suggesting that the optimization may still fall into local minima, especially in cases with fewer constraints or very ambiguous geometry. Even there, the adaptive kernel can still achieve better performance than the other two kernels.

Fig. 3 features the boxplots of the rotation errors of the six scenes with the most images. The adaptive kernel has the smallest median error in the first, third, and sixth boxplots and smallest first quartile (Q1) in all scenes except the fourth boxplot, which indicates that the adaptive kernel can provide more accurate rotation estimation. Interestingly, for the *living room* dataset (fourth boxplot)—where feature tracks are reliable and outliers are rare—the adaptive kernel’s conservative weighting slightly increases Q1 and Q3, but still achieves the lowest maximum error, indicating a safety margin against rare residual spikes. Overall, Fig. 3 confirms that the adaptive kernel enhances robustness in both structured outdoor and cluttered indoor environments, especially when traditional fixed-parameter losses suffer from degraded performance due to residual heterogeneity.

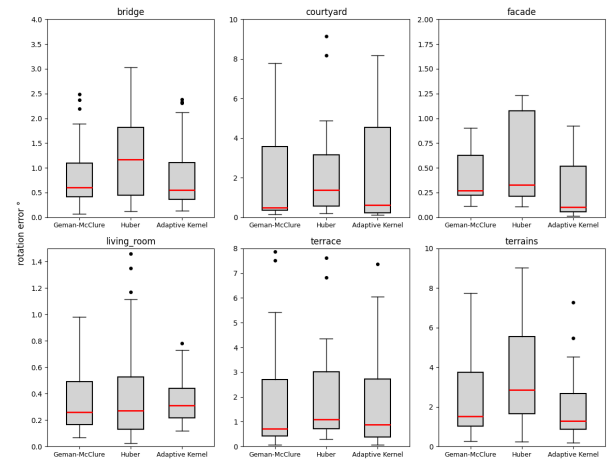


Figure 3. Rotation errors of six scenes with most images.

To demonstrate that the shape parameter is learned from the residuals in MA, we plot the iteration process of the shape parameter in the scene *living room* in Fig. 4 as an example. The shape parameter converges to a stable value after ten iterations, which is consistent with the distribution of residuals. The adaptive kernel can automatically adjust the shape parameter to fit the residuals, which is more efficient than manually tuning the parameters.

4.2 Evaluation of Global Positioning

Table 2 reports the global positioning AUC under three distance thresholds: 0.1m, 0.5m, and 1.0m. The adaptive kernel improves positioning across most scenes and all thresholds. The improvement is particularly significant at the stricter threshold (AUC@0.1m), where the average score increases to 24.53 from 21.27 (Geman-McClure) and 19.91 (Huber). This suggests that the adaptive kernel is more effective at rejecting outliers in global positioning. In scenes such as *office*, *relief_2*, and *terrace_2*, the adaptive kernel yields notable gains. However, the performance is comparable in simpler or well-conditioned scenes like *botanical_garden* or *observatory*, indicating that the

dataset	AUC@1°			AUC@3°			AUC@5°		
	Geman-McClure	Huber	Adapt. Ker.	Geman-McClure	Huber	Adapt. Ker.	Geman-McClure	Huber	Adapt. Ker.
botanical_garden	41.46	39.64	43.41	49.47	48.79	50.09	51.02	50.61	51.39
boulders	48.28	39.19	46.98	70.56	65.74	68.96	77.41	75.45	75.36
bridge	36.47	24.78	38.10	71.01	57.28	71.77	79.76	71.45	80.32
courtyard	36.23	20.69	38.27	58.65	50.01	58.77	66.01	62.58	64.95
delivery_area	50.58	18.57	51.94	69.26	41.97	69.56	75.18	56.72	75.80
door	13.24	14.86	13.24	42.62	44.31	42.24	64.93	66.52	64.70
electro	43.98	28.75	44.31	65.25	57.02	65.28	72.04	66.81	72.14
facade	58.24	51.11	66.07	71.84	68.29	74.96	77.84	74.72	80.12
kicker	40.35	27.03	39.11	62.73	57.12	62.16	68.61	65.24	69.53
lecture_room	42.58	34.90	42.46	60.49	57.86	61.61	69.19	68.04	69.02
living_room	63.87	61.27	63.71	84.47	84.14	84.83	89.50	89.91	89.67
lounge	64.51	46.49	69.90	77.06	70.37	78.86	87.97	82.06	89.23
meadow	14.11	5.78	14.11	18.04	9.90	18.04	18.82	15.42	18.82
observatory	84.07	81.02	85.18	92.22	91.21	92.59	93.85	93.24	94.07
office	49.02	51.57	48.59	68.81	69.03	68.11	73.59	73.73	73.17
old_computer	39.65	20.22	46.26	58.16	39.48	60.44	63.13	51.78	64.55
pipes	38.52	34.61	39.00	66.58	64.84	66.61	76.12	75.35	75.86
playground	73.64	69.88	73.36	91.00	89.50	90.90	94.60	93.70	94.54
relief	48.49	41.25	49.12	74.72	70.21	75.88	82.44	78.58	82.95
relief_2	54.58	26.24	56.97	78.63	58.62	79.51	84.60	70.19	85.12
statue	27.88	23.01	29.57	61.06	55.88	61.00	69.36	66.26	69.33
terrace	32.47	18.29	34.48	58.46	50.48	58.94	68.53	65.26	68.32
terrace_2	69.74	47.16	71.34	79.25	67.09	79.99	83.75	75.48	84.42
terrains	10.42	8.11	15.86	42.42	25.54	48.57	57.32	40.89	63.66
Average	45.10	34.77	46.72	65.53	58.11	66.24	72.73	67.92	73.21

Table 1. Comparison of RA errors under 1°, 3°, and 5° thresholds using different robust kernels.

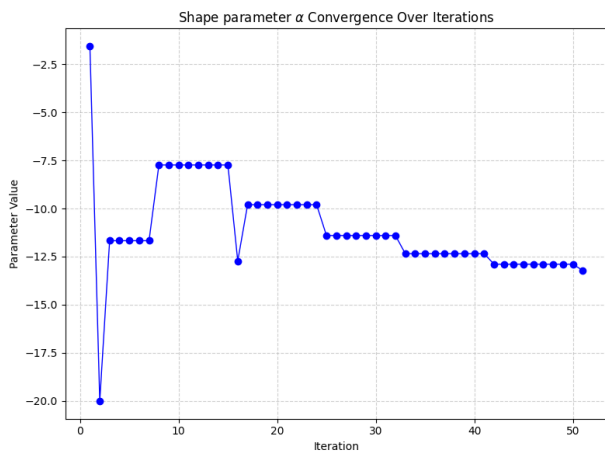


Figure 4. Iteration process of shape parameter α in the scene *living_room*.

benefit is most prominent in challenging cases. Fig. 5 shows the position error distributions across six scenes with most images. The adaptive kernel nearly consistently achieves lower or more concentrated errors compared to the fixed robust losses. In the first, second, and third boxplots, the adaptive kernel has both lower Q1 and medians, indicating improved accuracy and consistency. In the fourth and fifth boxplots, it has narrow interquartile. It can be summarized that the robust kernel can produce smaller medians in difficult scene and narrow interquartile in easy scenes, which suggests the kernel’s effectiveness in rejecting poor constraints that typically affect global positioning in reconstructions. To demonstrate that the shape parameter is also learned from the residuals in global positioning, we plot the iteration process of the shape parameter in the scene *old_computer* in Fig. 6. In terms of a practical implementation, we set the lower bound of the shape parameter $\alpha_{min} = -20$ as it does not have much difference with $\alpha = -\infty$ for large residuals. So the parameter doesn’t converge after 5 iterations and just reaches the lower bound. The shape parameter converges to a stable value after 17 iterations, which is consistent with

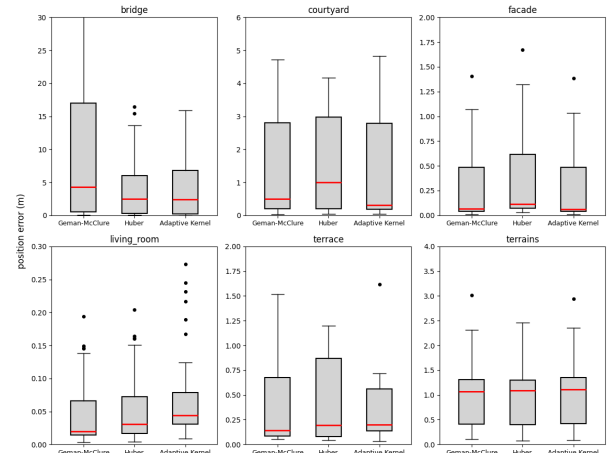


Figure 5. Position errors of six scenes with most images.

the distribution of residuals. The adaptive kernel can automatically adjust the shape parameter to fit the residuals, which is more efficient than manually tuning the parameters. Note that the IRLS is used in RA and Ceres is used in global positioning, so the iteration process is obviously different from Fig. 4.

5. CONCLUSIONS

This paper introduced enhanced RA and global positioning forms by incorporating an adaptive robust kernel into the MA framework. The adaptive kernel adjusts its shape parameter based on the residual distribution, enabling it to down-weight outliers more effectively than fixed-parameter loss functions.

The proposed method was integrated into the GLOMAP pipeline and evaluated on the ETH3D dataset. Quantitative results demonstrate consistent improvements over standard robust kernels such as Huber and Geman-McClure, particularly under stricter angular and positional error thresholds. The adaptive kernel achieves higher average AUC scores in both rotation av-

dataset	AUC@0.1m			AUC@0.5m			AUC@1.0m		
	Geman-McClure	Huber	Adapt. Ker.	Geman-McClure	Huber	Adapt. Ker.	Geman-McClure	Huber	Adapt. Ker.
botanical_garden	32.56	32.65	32.80	48.39	48.38	48.51	50.86	50.86	50.92
boulders	5.93	6.67	6.34	47.17	44.96	48.49	62.05	60.94	62.71
bridge	1.68	5.76	4.35	12.38	19.82	21.43	20.41	26.42	28.39
courtyard	3.64	3.39	5.15	29.82	28.01	35.64	43.73	38.59	50.05
delivery_area	14.06	2.18	12.87	51.03	26.03	54.30	60.74	48.19	62.38
door	0.00	0.00	0.00	18.31	24.62	17.85	37.64	41.32	37.35
electro	12.12	0.00	8.14	42.03	33.35	40.91	54.56	49.86	53.68
facade	33.32	15.24	33.56	62.50	54.12	62.62	70.09	65.33	70.14
kicker	20.52	14.87	21.91	53.07	51.05	54.91	62.49	62.87	65.07
lecture_room	11.94	6.34	11.67	51.79	44.21	50.27	64.84	64.15	64.16
living_room	61.40	56.20	47.83	86.41	84.99	82.56	90.13	89.42	88.20
lounge	21.25	36.09	23.09	68.55	72.03	68.74	75.94	77.68	76.04
meadow	4.12	0.00	0.00	16.43	8.99	14.71	21.55	17.55	20.69
observatory	66.08	66.45	66.92	90.25	90.33	90.42	93.27	93.31	93.36
office	0.00	17.41	24.03	36.73	61.57	63.64	59.61	72.86	74.13
old_computer	6.15	3.12	6.83	45.54	45.18	46.75	55.18	55.00	55.78
pipes	34.04	35.73	37.64	76.56	77.58	78.99	88.28	88.79	89.49
playground	46.79	49.33	39.79	87.33	87.90	83.62	93.67	93.95	92.09
relief	23.29	10.24	20.69	67.47	59.55	66.25	78.37	74.07	77.87
relief_2	49.04	40.68	51.77	69.04	66.62	69.45	71.61	70.40	71.82
statue	49.47	56.74	49.39	75.82	76.80	75.79	78.82	79.31	78.81
terrace	11.61	12.96	11.21	53.84	50.66	52.46	65.29	62.32	64.85
terrace_2	0.00	3.89	71.21	34.70	39.48	75.78	60.97	64.15	76.35
terrains	1.55	1.94	1.58	15.03	15.55	14.59	24.95	25.33	25.27
average	21.27	19.91	24.53	51.67	50.49	54.95	61.88	61.36	63.73

Table 2. Comparison of global positioning errors under 0.1m, 0.5m, and 1.0m thresholds using different robust kernels.

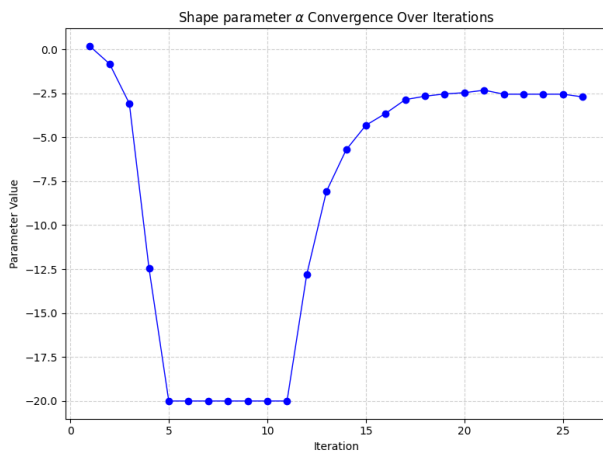


Figure 6. Iteration process of shape parameter α in the scene *old_computer*.

eraging and global positioning tasks, indicating improved robustness and accuracy.

These findings confirm the potential of adaptive robust kernels in challenging multi-view reconstruction scenarios. Notably, in this paper, the scale parameter is fixed according to the measurement noise. However, the noise level may not always be known in practice. So, our aim in future work is to explore how to make the scale parameter adaptive. Our plan is also to investigate the potential of adaptive kernels in other works, such as deep learning-based SfM, to further enhance robustness and accuracy in large-scale 3D reconstruction.

References

- Agarwal, S., Mierle, K. et al., 2012. Ceres solver: Tutorial & reference. *Google Inc*, 2(72), 8.
- Barron, J. T., 2019. A general and adaptive robust loss function. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 4331–4339.

Black, M. J., Anandan, P., 1996. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer vision and image understanding*, 63(1), 75–104.

Carlone, L., Calafiore, G. C., 2018. Convex Relaxations for Pose Graph Optimization With Outliers. *IEEE Robotics and Automation Letters*, 3(2), 1160–1167.

Charbonnier, P., Blanc-Feraud, L., Aubert, G., Barlaud, M., 1994. Two deterministic half-quadratic regularization algorithms for computed imaging. *Proceedings of 1st International Conference on Image Processing*, 2, IEEE Comput. Soc. Press, Austin, TX, USA, 168–172.

Chatterjee, A., Govindu, V. M., 2013. Efficient and Robust Large-Scale Rotation Averaging. *2013 IEEE International Conference on Computer Vision*, IEEE, Sydney, Australia, 521–528.

Chatterjee, A., Govindu, V. M., 2018. Robust Relative Rotation Averaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 958–972.

Chebroly, N., Labe, T., Vysotska, O., Behley, J., Stachniss, C., 2021. Adaptive Robust Kernels for Non-Linear Least Squares Problems. *IEEE Robotics and Automation Letters*, 6(2), 2240–2247.

Chen, Y., Zhao, J., Kneip, L., 2021. Hybrid Rotation Averaging: A Fast and Robust Rotation Averaging Approach. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Nashville, TN, USA, 10353–10362.

Dennis Jr, J. E., Welsch, R. E., 1978. Techniques for non-linear least squares and robust regression. *Communications in Statistics-simulation and Computation*, 7(4), 345–359.

Elnashef, B., Filin, S., 2022. Rotation Only Bundle-Adjustment for the Generalized Camera Model and its Application for Large-Scale Underwater Image-Sets. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, V-2-2022, 325–332.

- Eriksson, A., Olsson, C., Kahl, F., Chin, T.-J., 2019. Rotation Averaging with the Chordal Distance: Global Minimizers and Strong Duality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–1.
- Gao, X., Zhu, L., Xie, Z., Liu, H., Shen, S., 2021. Incremental rotation averaging. *International Journal of Computer Vision*, 129, 1202–1216.
- Geman, S., McClure, D. E., 1985. Bayesian image analysis: An application to single photon emission tomography. *Proceedings of the American Statistical Association*.
- Govindu, V., 2004. Lie-algebraic averaging for globally consistent motion estimation. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, 1, IEEE, Washington, DC, USA, 684–691.
- Hartley, R., Trunf, J., Dai, Y., Li, H., 2013. Rotation Averaging. *International Journal of Computer Vision*, 103(3), 267–305.
- Huber, P. J., 1992. Robust estimation of a location parameter. *Breakthroughs in statistics: Methodology and distribution*, Springer, 492–518.
- Larsson, V., contributors, 2020. PoseLib - Minimal Solvers for Camera Pose Estimation.
- Leclerc, Y. G., 1989. Constructing simple stable descriptions for image partitioning. *International journal of computer vision*, 3(1), 73–102.
- Manam, L., Govindu, V. M., 2024. Fusing Directions and Displacements in Translation Averaging. *2024 International Conference on 3D Vision (3DV)*, IEEE, Davos, Switzerland, 75–84.
- Martinec, D., Pajdla, T., 2007. Robust Rotation and Translation Estimation in Multiview Reconstruction. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, Minneapolis, MN, USA, 1–8.
- Moulon, P., Monasse, P., Perrot, R., Marlet, R., 2017. Open-MVG: Open Multiple View Geometry. B. Kerautret, M. Colom, P. Monasse (eds), *Reproducible Research in Pattern Recognition*, 10214, Springer International Publishing, Cham, 60–74.
- Pan, L., Baráth, D., Pollefeys, M., Schönberger, J. L., 2024. Global structure-from-motion revisited. *European Conference on Computer Vision*, Springer, 58–77.
- Schönberger, J. L., Frahm, J.-M., 2016. Structure-from-Motion Revisited. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Las Vegas, NV, USA, 4104–4113.
- Schops, T., Schönberger, J. L., Galliani, S., Sattler, T., Schindler, K., Pollefeys, M., Geiger, A., 2017. A Multi-view Stereo Benchmark with High-Resolution Images and Multi-camera Videos. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Honolulu, HI, 2538–2547.
- Sidhartha, C., Govindu, V. M., 2021. It Is All In The Weights: Robust Rotation Averaging Revisited. *2021 International Conference on 3D Vision (3DV)*, IEEE, London, United Kingdom, 1134–1143.
- Snively, N., Seitz, S. M., Szeliski, R., 2006. Photo tourism: exploring photo collections in 3d. *ACM siggraph 2006 papers*, 835–846.
- Sweeney, C., Hollerer, T., Turk, M., 2015. Theia: A fast and scalable structure-from-motion library. *Proceedings of the 23rd ACM international conference on Multimedia*, 693–696.
- Wang, J., Karaev, N., Rupprecht, C., Novotny, D., 2024. VGSfM: Visual Geometry Grounded Deep Structure from Motion. *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Seattle, WA, USA, 21686–21697.
- Wilson, K., Snively, N., 2014. Robust Global Translations with 1DSfM. D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (eds), *Computer Vision – ECCV 2014*, 8691, Springer International Publishing, Cham, 61–75.
- Wu, C., 2013. Towards Linear-Time Incremental Structure from Motion. *2013 International Conference on 3D Vision*, IEEE, Seattle, WA, USA, 127–134.
- Zhang, G., Larsson, V., Barath, D., 2023. Revisiting Rotation Averaging: Uncertainties and Robust Losses. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Vancouver, BC, Canada, 17215–17224.
- Zhuang, B., Cheong, L.-F., Lee, G. H., 2018. Baseline Desensitizing in Translation Averaging. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Salt Lake City, UT, 4539–4547.