Robust Geometric Model Estimation Based on Scaled Welsch q-Norm

Jiayuan Li, Qingwu Hu, and Mingyao Ai

***Abstract*— Robust estimation, which aims to recover the geometric transformation from outlier contaminated observations, is essential for many remote sensing and photogrammetry applications. This article presents a novel robust geometric model estimation method based on scaled Welsch q-norm (l_q-norm, $0 < q < 1$). The proposed algorithm integrates a scaled Welsch weight function into the q-norm framework. It, thus, inherits all the advantages of the standard q-norm, i.e., fast and robust. The parameter sensitivity of the standard q-norm is also largely alleviated by integrating such a weight function. These make the proposed algorithm much superior to RANSAC-type methods in real-life applications. We formulate the new cost function as an augmented Lagrangian function (ALF) and divide the ALF into two subproblems [a q-norm penalized least-squares (l_qLS) problem and a weighted least-squares (WLS) problem] by using alternating direction method of multipliers (ADMM) method. For the WLS problem, we introduce a coarse-to-fine strategy into the iterative reweighted least-squares (IRLS) method. We change the weight function by decreasing its scale parameter. This strategy can largely avoid that the solver gets stuck in local minimums. We adapt the proposed cost into classical remote sensing tasks and develop new robust feature matching (RFM), robust exterior orientation (REO), and robust absolute orientation (RAO) algorithms. Both synthetic and real experiments demonstrate that the proposed method significantly outperforms the other compared state-of-the-art methods. Our method is still robust even if the outlier rate is up to 90%.**

***Index Terms*— Image orientation, model fitting, outlier removal, point cloud registration, robust feature matching (RFM).**

I. Introduction

ROBUST estimation is a technique for simultaneous geometric model fitting and outlier detection. Due to the inevitable existence of outliers in real observations, robust estimation has a variety of applications in remote sensing and photogrammetry. For example, robust feature matching (RFM) is a prerequisite for image registration [1]–[3], image mosaicking [4]–[6], image fusion [7], [8], and bundle adjustment [9]; robust exterior orientation (REO) is a prerequisite for digital orthophoto map generation [10] and 3-D reconstruction [11].

Manuscript received November 16, 2019; revised January 4, 2020; accepted February 5, 2020. Date of publication February 21, 2020; date of current version July 22, 2020. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 41901398, in part by the Natural Science Foundation of Hubei Province under Grant 2019CFB167, in part by the China Postdoctoral Science Foundation under Grant 2018M640734, and in part by Wuhan University - Huawei GeoInformatics Innovation Laboratory. *(Corresponding author: Jiayuan Li.)*

The authors are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430072, China (e-mail: ljy_whu_2012@whu.edu.cn).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.2972982

In remote sensing, there are two types of widely used robust estimation methods, i.e., robust estimators and maximum consensus [12].

Robust estimators, such as M-estimators, S-estimators, and MM-estimators, are popular in statistics, whose goal is to give large weights to inliers while small weights to outliers. Thus, the effect of outliers toward the cost function can be largely discounted. Generally, robust estimators are optimized via the iteratively reweighted least-squares (IRLS) [13] algorithm (or called the iteration method with variable weights in the adjustment of measurement), which alternates between weights assignment and weighted least-squares (WLS) optimization until convergence. Robust estimators are efficient and the estimated solutions are optimal. However, they can only deal with cases with no more than 50% of outliers [14], [15].

The maximum consensus is a technique whose goal is to seek a model that maximizes the size of the inlier set. RANSAC [16] is a typical maximum consensus method. RANSAC-type methods can handle more than 50% of outliers. Owing to this property, they are the most popular robust regression methods in computer vision and also in remote sensing. However, RANSAC gives no guarantee of optimality [12]. Because RANSAC and its variants are randomized methods, they cannot absolutely guarantee that the obtained results are satisfactory approximations, let alone optimal estimates. Even if all possible subsets are tried, the optimal solution may not be found because the optimal solution corresponds to all observations rather than subsets [12], [17].

In this article, we propose a new robust geometric model estimation method based on the scaled Welsch q-norm (l_q-norm, $0 < q < 1$). The proposed cost can be regarded as an extension of standard q-norm estimation [18], [19], i.e., we integrate a scaled Welsch function into the q-norm framework to construct a new cost function. It, thus, inherits all the advantages of the standard q-norm estimation. The parameter sensitivity of q-norm estimation is also largely alleviated by integrating a weight function. These make the proposed estimate far superior to RANSAC-type methods in real-life applications. We use augmented Lagrangian function (ALF) [20] and alternating direction method of multipliers (ADMM) [21] to optimize this nonconvex and nonsmooth object function. The ADMM divides the cost into two subproblems, i.e., a q-norm penalized least-squares (l_qLS) problem and a WLS problem. For the WLS problem, we introduce a coarse-to-fine strategy into the IRLS method. We change the weight function by decreasing its scale parameter instead of fixing it. This strategy can largely avoid that the solver gets stuck in

0196-2892 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

local minimums. We also give applications to geometric processing in remote sensing. We develop new RFM, REO, and robust absolute orientation (RAO) algorithms. Extensive experiments demonstrate the power of the proposed method, i.e., it is still robust even if the outlier rate is up to 90%.

II. Related Work

This section reviews recently proposed robust regression techniques (both robust estimators and maximum consensus methods) in remote sensing or computer vision instead of statistics, since the proposed method is designed for remote sensing, photogrammetry, or computer vision.

A. Robust Estimators

As known, the breakdown point of all robust estimators is no better than 0.5. Typical robust estimators cannot handle cases with more than 50% of outliers, such as M-type estimators and cross correntropy.

1) M-Type Estimators [15], [22]: M-estimators can be regarded as generalized maximum-likelihood estimators. Typically, their costs are positive-definite and symmetric functions, such as Huber cost, Welsch cost, and Tukey cost. Generalized M-estimators (GM-estimators) incorporate an appropriate loss weight to bound the influence function of M-estimators. S-estimators are based on a scale estimate. MM-estimators combine the high breakdown of S-estimators and the high efficiency of M-estimators. M-type estimators had been widely used in remote sensing and image processing tasks, such as pixel completion [23] and image denoising [24].

2) Cross Correntropy [25]–[28]: Liu *et al.* [25] proposed a novel localized similarity measure called cross correntropy. Compared with the mean-square error (MSE) measure, correntropy dictates the value of the similarity measure by the kernel function while MSE dictates the value by all the samples. Thus, the cost function that maximizes the cross correntropy has no assumption of Gaussian noise and is more robust to outliers than the cost of MSE (l_2-norm cost). Essentially, the maximum correntropy criterion (MCC) can also be regarded as an M-estimate. If we set $\rho(v) = ((1 - \exp(-v^2/2\sigma^2)))/\sqrt{2\pi}\sigma$, where $\rho(\cdot)$ is the cost of M-estimate and v is a residual, the M-estimate cost becomes an MCC cost. MCC has been applied in many computer vision and signal processing problems. For example, He *et al.* [27], [28] introduced the MCC into pattern recognition and principal component analysis [29]; Chen *et al.* [30] applied the MCC to adaptive filtering; Wang and Pan [31] proposed a K-means clustering algorithm based on the MCC measure.

Several recently developed estimators that assume outliers are uniformly distributed can tolerate more than half of the data being outliers, such as generalized projection-based M-estimator (GpbM-estimator), L_2E estimator, and q-norm estimator.

3) GpbM-Estimator [32]: It is an extension of projection-based M-estimator [33], which reformulates M-estimate as a projection-based optimization problem. There are three major steps in GpbM-estimator: it first computes a scale estimate based on the hypothesize-and-verify technique; then, performs a model estimation step and refines the model via local Grassmann Manifolds optimization; finally, an inlier/outlier dichotomy strategy is adapted to distinguish outliers from inliers. The hypothesize and manifolds optimization stages largely decrease its efficiency.

4) L_2E Estimator: Ma *et al.* [34] proposed a robust estimator called L_2E estimation to deal with noise and outliers, the cost function of which is to minimize the l_2 distance between densities. The L_2E estimation is then applied to nonrigid feature matching to show its effectiveness.

5) Q-Norm Estimation [18], [19]: As pointed out in the literature [35]–[37], q-norm ($0 < q < 1$) overcomes some shortcomings of l_1-norm. It is less biased and much sparser than l_1-norm. Therefore, q-norm performs better than l_1-norm in many tasks. For example, Marjanovic and Solo [36] used q-norm for matrix completion. Xie *et al.* [38] developed an image denoising algorithm based on q-norm. Xu *et al.* [37] proposed an iterative half thresholding algorithm based on $l_{1/2}$ regularization for sparse modeling. They demonstrate that the $l_{1/2}$ regularization is superior to l_1 regularization. However, these works did not use the q-norm as a robust cost. In our previous work [18], we introduced the q-norm for robust estimation, which is solved by the ADMM method. Superior to M-estimators and cross correntropy, it can tolerate up to 80% of outliers. The major limitation is the sensitivity to parameters. Different parameters may lead to very different results. For different data sets, the parameters must be carefully tuned, which makes the method is not competitive against RANSAC-type methods. The proposed scaled Welsch q-norm is an extension of q-norm estimation.

B. Maximum Consensus

Maximum consensus methods contain widely used RANSAC-type methods and other exact methods. RANSAC uses a hypothesize-and-verify technique for model fitting, which alternates between randomly minimal subset selection and estimated model verification. The model with the largest inlier set is returned as the solution after a mass of trials. RANSAC has dozens of variants [39]–[43]. For example, Torr and Zisserman [39] proposed a probability-based RANSAC variant called MLESAC, which maximizes the likelihood instead of the number of inliers. Chum *et al.* [40] presented a locally optimized RANSAC (LO-RANSAC). LO-RANSAC improves the sampling stage of the original RANSAC by a local optimization step, which significantly decreases the number of sampling trails. Raguram *et al.* [41] proposed a universal framework for RANSAC-type methods named USAC, which incorporates a number of important practical and computational tricks. Brachmann *et al.* [42] tried to introduce deep learning technique into RANSAC. They proposed a learning-based method called differentiable sample consensus (DSAC). DSAC learns good hypotheses in an end-to-end fashion by a probabilistic selection, where the expected loss with regard to all learnable parameters can be derived. El-Melegy [44] proposed a RANSAC variant that integrates the M-estimate into the traditional RANSAC framework. However, as mentioned earlier,

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

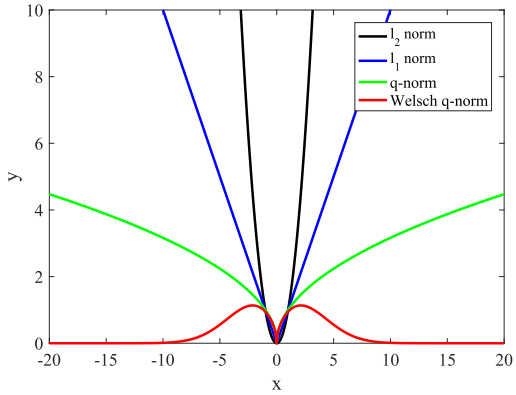

Fig. 1. Cost function curves, the black curve, blue curve, green curve, and red curve, represent the costs of the least-squares estimation (l_2-norm), the l_1-norm estimation, the q-norm estimation ($q = 0.5$), and the proposed scaled Welsch q-norm estimation ($q = 0.5$), respectively.

RANSAC-type methods suffer from a major drawback, i.e., no guarantee of the optimality.

RANSAC-type methods only estimate approximate solutions for a geometric model. Differently, the goal of the exact methods is to find the globally optimal solution. Exact algorithms are usually theoretically justified, which means that optimality guarantees, convergence, and bounds can be provided [12]. The maximum consensus is an intractable problem because of its combinatorial nature. Hence, different search methods should be conducted to find a global solution. The most popular brute force search method is branch-and-bound (BnB) [45], [46]. However, the theoretical computational complexity of these exhaustive search methods is huge, i.e., exponential runtime, which prevents their usages in real applications. More recently, some efforts are made to conduct the search efficiently. For instance, Chin *et al.* [47] proposed a guaranteed outlier removal technique based on mixed-integer linear programming (MILP). Specifically, they used MILP to remove observations that are probably true outliers before global optimization. Therefore, the search space of exact methods can be largely reduced. In another work [48], they transformed the maximum consensus as a tree search problem. They developed a method based on A* search [49] under the framework of LP-type methods [50]. Although many efforts are made, these exact methods are still much slower than approximate methods like RANSAC.

III. Methodology

A. Motivation

Least-squares regression is widely adopted in remote sensing and photogrammetry, such as exterior orientation, triangulation, and bundle adjustment, because of its efficiency and simplicity. However, it relies on an assumption of Gaussian noise, which is violated by outliers. From Fig. 1, the cost of least squares (the black curve) has no bound and increases quadratically. It is not a robust estimate.

Maximum consensus is a typical technique for robust regression, which finds the model f that maximizes the number of inlier observations [12], [51]

$$\begin{aligned} &\arg\max_{\boldsymbol{\delta}} |I| \\ &\text{s. t. } \| f(\boldsymbol{x}_i, \boldsymbol{\delta}) - \boldsymbol{y}_i \|_2 \le \xi \quad \forall i \in I \end{aligned} \tag{1}$$

where $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ is a multidimensional observation, $\boldsymbol{\delta}$ denotes the parameter set of model f, $H = \{1, 2, \ldots, n\}$ is an index set of the input observations, ξ is an inlier threshold, and $\|\cdot\|_2$ is the l_2-norm. The subset $I \subseteq H$ is an inlier set or called consensus set and $|I|$ represents its size. Essentially, this objective is equivalent to minimize the number of outlier observations, which can be formulated as

$$\begin{aligned} &\arg\min_{\boldsymbol{\delta}} \quad \| \boldsymbol{O} \|_0 \\ &\text{s. t. } \| f(\boldsymbol{x}_i, \boldsymbol{\delta}) - \boldsymbol{y}_i \|_2 \le \xi + o_i \\ &\qquad o_i \ge 0, \quad i = 1, 2, \ldots, n \end{aligned} \tag{2}$$

where $\boldsymbol{O} = [o_1, o_2, \ldots, o_n]^T$ is a nonnegative slack variable vector and $\|\cdot\|_0$ is the l_0-norm. $\|\boldsymbol{O}\|_0$ is the number of nonzero elements in vector $\boldsymbol{O}$. However, this formulation does not make the problem any easier. Optimization of this l_0-norm cost is still very difficult due to the high nonconvexity. A popular choice is to replace the l_0-norm with l_1-norm, since l_1-norm is the closest convex relaxation of l_0-norm. Then, this problem can be drastically simplified

$$\begin{aligned} &\arg\min_{\boldsymbol{\delta}} \quad \| \boldsymbol{O} \|_1 \\ &\text{s. t. } \| f(\boldsymbol{x}_i, \boldsymbol{\delta}) - \boldsymbol{y}_i \|_2 \le \xi + o_i \\ &\qquad o_i \ge 0, \quad i = 1, 2, \ldots, n \end{aligned} \tag{3}$$

where $\|\cdot\|_1$ is the l_1-norm and $\|\boldsymbol{O}\|_1 = |o_1| + |o_2| + \cdots + |o_n|$. The l_1-norm cost has been widely used in robust geometric model estimation, such as triangulation [52], structure from motion [53], [54], and feature matching [55], which is more robust than l_2-norm. It gives equal emphasis to all data while least squares, by squaring the residuals, assign larger weights to outliers. Hence, the cost curve of l_1-norm (the blue curve) only increases linearly. However, l_1-norm still does not largely reduce the influence of outliers. As pointed out in the literature [12], l_1-norm minimization technique should be used only on problems with low outlier rates.

To bridge the gap between the l_0-norm and l_1-norm, q-norm gains a significant interest in image denoising [38], matrix completion [36], [56], and sparse reconstruction [35], [57]. It overcomes some shortcomings of l_1-norm. For instance, q-norm is less biased and much sparser than l_1-norm. Xu et al. [37] demonstrated that $l_{1/2}$ regularization is superior to l_1 regularization. Motivated by these, we introduced the q-norm for robust estimation in our previous work [18]. Different from the l_1 minimization which is based on a maximum consensus framework [see (3)], we use the q-norm cost to construct a robust estimator. From Fig. 1, the cost of q-norm (The green curve, take $q = 0.5$ as an example. Note that the other values between 0 and 1 such as $q = 0.3$, $q = 0.4$, and $q = 0.7$ can also be used.) is much more gradual than l_2-norm and l_1-norm. It can largely discount the effect of outliers and has a high degree of robustness. As demonstrated in [18] and [19], the q-norm estimate can tolerate up to 80%

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

of outliers. However, this cost is unbounded and q-norm estimate is sensitive to parameters. To deal with this problem, we integrate a scaled Welsch weight function into the q-norm framework. Then, the new cost becomes a bounded redescending function. From the curve (the red curve), we can learn that outliers with large residuals do not contribute to the total cost. It almost completely discounts the influence of outliers. Thus, the proposed scaled Welsch q-norm cost has a higher degree of robustness than the standard q-norm estimate.

B. Scaled Welsch Q-Norm Cost

For better illustration of the proposed cost and its optimization framework, we take the RFM problem as an example. Given a set of outlier contaminated matches $C = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^n$ (observations) that is extracted from a remote sensing image pair, where $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are 2-D image coordinates (in pixels). The goal of RFM is to remove false matches inside C, i.e., distinguish inliers from outliers. The geometric relationship between $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ can be modeled by function f. For RFM problem, function f is an affine transformation. If $(\boldsymbol{x}_i, \boldsymbol{y}_i)$ is an inlier, we have

$$\boldsymbol{y}_i = f(\boldsymbol{x}_i, \boldsymbol{\delta}) = \mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} \tag{4}$$

where $\boldsymbol{\delta} = (\mathbf{A}, \boldsymbol{t})$ denotes the parameter set of function f. $\mathbf{A}$ is a 2×2 affine matrix and $\boldsymbol{t}$ is a 2×1 translation column vector. Generally, this problem can be solved by the least-squares regression

$$\arg\min_{\mathbf{A},\boldsymbol{t}} \sum_{i=1}^{n} \| \mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{y}_i \|_2^2. \tag{5}$$

As mentioned above, least squares is sensitive to outliers. The estimated parameters will be skewed to adjust large residuals.

We propose a new robust method which can directly recover the accurate parameter set $\boldsymbol{\delta}$ from $\{\boldsymbol{x}_i\}_1^n$ and $\{\boldsymbol{y}_i\}_1^n$. The new cost function is

$$\begin{aligned}\arg\min_{\boldsymbol{\delta}} \sum_{i=1}^{n} \| w_i[f(\boldsymbol{x}_i, \boldsymbol{\delta}) - \boldsymbol{y}_i] \|_q^q \\ = \arg\min_{\mathbf{A},\boldsymbol{t}} \sum_{i=1}^{n} \| w_i(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{y}_i) \|_q^q\end{aligned} \tag{6}$$

where $0 < q < 1$. w_i is a weight for match $(\boldsymbol{x}_i, \boldsymbol{y}_i)$. Ideally, it gives a large weight (close to 1) to an inlier while a small one (close to 0) to an outlier. We use the scaled Welsch function

$$w_i = e^{-(v_i/u)^2} \tag{7}$$

where u is a scale factor and v_i is a residual.

This cost function is nonsmooth and nonconvex. We use ALF and ADMM to solve this problem. First, auxiliary variables $\mathbf{M} = \{\boldsymbol{m}_i\}_1^n$ are introduced into (6)

$$\begin{aligned}&\arg\min_{\mathbf{A},\boldsymbol{t}} \sum_{i=1}^{n} \| \boldsymbol{m}_i \|_q^q \\ &\text{s.t.} \quad \boldsymbol{\varepsilon}_i = w_i(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{y}_i) - \boldsymbol{m}_i = \mathbf{0}.\end{aligned} \tag{8}$$

This constrained function can be rewritten as an unconstrained one via the Lagrangian function and solved by the dual ascent method. However, the dual ascent method may fail if the cost function is not strictly convex [21]. ALF has no assumptions of convexity and is adopted

$$\begin{aligned}&L_\rho(\mathbf{A}, \boldsymbol{t}, W, \mathbf{M}) \\ &= \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \boldsymbol{\lambda}_i^T \boldsymbol{\varepsilon}_i + \frac{\rho}{2} \| \boldsymbol{\varepsilon}_i \|_2^2 \right) \\ &= \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \frac{\boldsymbol{\lambda}_i}{\rho} + \boldsymbol{\varepsilon}_i \|_2^2 - \frac{1}{2\rho} \| \boldsymbol{\lambda}_i \|_2^2 \right)\end{aligned} \tag{9}$$

where $\{\boldsymbol{\lambda}_i\}_1^n$ are dual variables and $\rho > 0$ is a penalty coefficient.

The optimization of (9) is an intractable problem since it is nonconvex. Fortunately, in the ALF, there are two sets of variables, i.e., parameter set $\boldsymbol{\delta} = (\mathbf{A}, \boldsymbol{t})$ and weight set $W = \{w_i\}_1^n$, and auxiliary variables $\mathbf{M}$. Thus, the optimization process can be performed efficiently by alternating between $(\boldsymbol{\delta}, W)$ estimation and $\mathbf{M}$ estimation via ADMM. ADMM has both superior convergence and decomposability properties. It decomposes (9) into two main parts

$$\begin{aligned}\boldsymbol{RFM}_{\boldsymbol{part1}}: \quad &\arg\min_{\mathbf{M}} L_\rho \\ &= \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \boldsymbol{\beta}_i - \boldsymbol{m}_i \|_2^2 \right)\end{aligned} \tag{10}$$

$$\begin{aligned}\boldsymbol{RFM}_{\boldsymbol{part2}}: \quad &\arg\min_{\mathbf{A},\boldsymbol{t},W} L_\rho \\ &= \arg\min_{\mathbf{A},\boldsymbol{t},W} \sum_{i=1}^{n} \frac{\rho}{2} \| w_i(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{\gamma}_i) \|_2^2\end{aligned} \tag{11}$$

$\{\boldsymbol{\beta}_i\}_1^n$ and $\{\boldsymbol{\gamma}_i\}_1^n$ are used for notation compactness

$$\begin{cases}\boldsymbol{\beta}_i = \dfrac{\boldsymbol{\lambda}_i}{\rho} + w_i^{k-1}(\mathbf{A}^{k-1}\boldsymbol{x}_i + \boldsymbol{t}^{k-1} - \boldsymbol{y}_i) \\ \boldsymbol{\gamma}_i = \boldsymbol{y}_i + \dfrac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}}.\end{cases} \tag{12}$$

For $part1$, $(\mathbf{A}, \boldsymbol{t}, W)$ is fixed and $\mathbf{M}$ is the variable. The equation is separable and can be efficiently solved via coordinate-wise optimal minimization methods [36]. (See Appendix VI for the derivation of (10∼11) and the details of $part1$ optimization.)

For $part2$, $\mathbf{M}$ is fixed and $(\mathbf{A}, \boldsymbol{t}, W)$ is the variable. Equation (11) is a weighted sum of l_2-norm cost which is generally solved by IRLS. Different from the classic IRLS which fixes the weight function, we change the weight function through the scale parameter u in each iteration. See in Fig. 2, large values of u make the width of the cost curve larger. So, more observations are allowed to take part in the optimization stage. In contrast, the width of the cost curve becomes narrower as u decreases, which makes the parameter estimation more precise. For example, when $u = 36$, matches with residuals smaller than 80 pixels all contribute to the cost of (11). However, when $u = 6$, matches with $v_i > 10$ almost have no impact on the cost. In our optimization stage, we first give a very large value to u ($u = D$, where D is the largest distance between two matches in $\{\boldsymbol{x}_i\}_1^n$). During the iteration process, we decrease the value of u until the optimization is

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

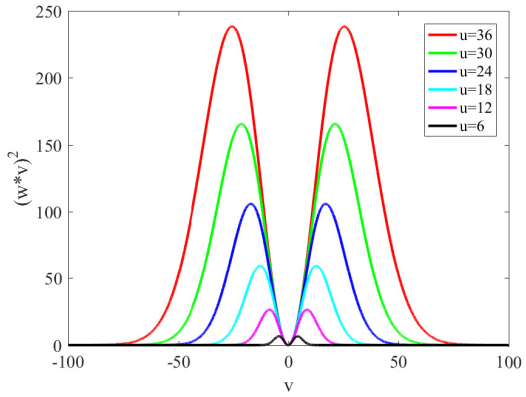

Fig. 2. Illustration of the proposed coarse-to-fine IRLS strategy. Larger values of scale u have smoother weight function curves, so more observations are allowed to take part in the optimization stage. As the scale u decreases, the weight function becomes sharper and the estimation is more accurate.

converged or u reaches 3ξ, where ξ is a threshold for the identification of inlier matches. For RFM, the affine model is linear, which can be solved without the initial guess. Benefit from this coarse-to-fine strategy, the effect of local minimums can be largely alleviated. Algorithm 1 summarizes the optimization process of $part1$ and $part2$.

Algorithm 1 RFM via Scaled Welsch q-Norm

Input: initial match set: $C = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^n$
Output: affine parameters $\boldsymbol{\delta} = (\mathbf{A}, \boldsymbol{t})$ and inlier set I
Initialize q, ρ, $\{\boldsymbol{\lambda}_i\}_1^n = \mathbf{0}$, $\{w_i\}_1^n = \mathbf{1}$, and $u = D$;
while $\boldsymbol{\delta}$ *unconverges* **do**
  (1) part1 optimization:
  Compute $\{\boldsymbol{\beta}_i\}_1^n$ according to Eq. (12);
  Optimize $\boldsymbol{m}_i$ via shrinkage algorithm;
  (2) part2 optimization:
  Compute $\{\boldsymbol{\gamma}_i\}_1^n$ according to Eq. (12);
  while W *unconverges or* $u > 3\xi$ **do**
    Optimize $\boldsymbol{\delta} = (\mathbf{A}, \boldsymbol{t})$ via WLS;
    Update $u = u/\eta$, $\eta > 1$ is a step-size;
    Update $\{w_i\}_1^n$ based on Eq. (7);
  Update dual variable set $\{\boldsymbol{\lambda}_i\}_1^n$, $\boldsymbol{\lambda}_i = \boldsymbol{\lambda}_i + \rho\boldsymbol{\varepsilon}_i$;
Apply affine model f on C to find inlier set I.

IV. Applications

The proposed cost is general. It is suitable for not only RFM but also other geometric model estimation problems, such as exterior orientation [also called Perspective-n-Point (PnP)] and absolute orientation. Exterior orientation and absolute orientation play a very important role in remote sensing, which are the foundations of bundle adjustment, coordinate transform, point cloud registration, and so on. In the above, we take RFM as an example; in this section, we apply the proposed cost to exterior orientation and absolute orientation, developing new REO and RAO algorithms.

A. Robust Exterior Orientation (REO)

Suppose we are given n 3-D–2-D contaminated correspondences $\{(\boldsymbol{Q}_i, \boldsymbol{p}_i)\}_1^n$, where $\{\boldsymbol{Q}_i\}_1^n$ are coordinates of noncollinear 3-D points in an object reference system and $\{\boldsymbol{p}_i\}_1^n$ are 2-D image projections. Our goal is to recover the pose of the image. Assume that the camera internal parameters $\mathbf{K}$ are known. Then, the camera projection model which maps 3-D object points into 2-D image points is as follows [58]:

$$d_i \begin{bmatrix} \boldsymbol{p}_i \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{R}, \boldsymbol{T}] \begin{bmatrix} \boldsymbol{Q}_i \\ 1 \end{bmatrix} \tag{13}$$

where $\mathbf{R}$ and $\boldsymbol{T}$ are rotation matrix and translation vector, respectively. d_i is the depth of image point $\boldsymbol{p}_i$. After eliminating d_i, the geometric model f of camera exterior orientation problem is

$$\boldsymbol{p}_i = f(\boldsymbol{Q}_i, \boldsymbol{\delta}) = \frac{\mathbf{P}_{1:2}[\ \boldsymbol{Q}_i\ 1\ ]^T}{\mathbf{P}_3[\ \boldsymbol{Q}_i\ 1\ ]^T} \tag{14}$$

where $\mathbf{P} = \mathbf{K}[\mathbf{R}, \boldsymbol{T}]$ is a 3×4 camera matrix; $\mathbf{P}_j (j = 1, 2, 3)$ represents the jth row of $\mathbf{P}$; and $\boldsymbol{\delta} = (\mathbf{R}, \boldsymbol{T})$.

The scaled Welsch q-norm cost of REO is

$$\begin{aligned} &\arg\min_{\boldsymbol{\delta}} \sum_{i=1}^n \| \ w_i[f(\boldsymbol{Q}_i, \boldsymbol{\delta}) - \boldsymbol{p}_i] \ \|_q^q \\ &\quad = \arg\min_{\mathbf{R}, \boldsymbol{T}} \sum_{i=1}^n \left\| w_i \left(\frac{\mathbf{P}_{1:2}[\ \boldsymbol{Q}_i\ 1\ ]^T}{\mathbf{P}_3[\ \boldsymbol{Q}_i\ 1\ ]^T} - \boldsymbol{p}_i \right) \right\|_q^q. \end{aligned} \tag{15}$$

Similarly, ADMM also obtains two main subproblems (see Appendix B for details). In $part2$, the unit quaternions are used to format the rotation matrix and the Gauss–Newton algorithm is adapted to optimize the nonlinear least squares (line 9 of Algorithm 1). The Gauss–Newton method first linearizes the nonlinear function via Taylor series expansion and solves the cost iteratively. Hence, the initial values of parameters $\boldsymbol{\delta} = (\mathbf{R}, \boldsymbol{T})$ are required to guarantee convergence.

B. Robust Absolute Orientation (RAO)

Recovering the geometric relationship between two different Cartesian coordinate systems from n given 3-D–3-D correspondences $\{(\boldsymbol{X}_i, \boldsymbol{Y}_i)\}_1^n$ is known as the absolute orientation problem in photogrammetry and remote sensing. It is usually applied in registering a freenet adjustment result to a geodetic coordinate system and point cloud registration. Mathematically, this problem is defined by a seven-parameter rigid transformation

$$\boldsymbol{Y}_i = f(\boldsymbol{X}_i, \boldsymbol{\delta}) = s\mathbf{R}\boldsymbol{X}_i + \boldsymbol{T} \tag{16}$$

where $\mathbf{R}$ and $\boldsymbol{T}$ are rotation matrix and translation vector, respectively; s is a scale factor; and $\boldsymbol{\delta} = (s, \mathbf{R}, \boldsymbol{T})$. The correspondences $\{(\boldsymbol{X}_i, \boldsymbol{Y}_i)\}_1^n$ contain outliers.

The scaled Welsch q-norm cost of RAO is

$$\begin{aligned} &\arg\min_{\boldsymbol{\delta}} \sum_{i=1}^n \| \ w_i[f(\boldsymbol{X}_i, \boldsymbol{\delta}) - \boldsymbol{Y}_i] \ \|_q^q \\ &\quad = \arg\min_{s, \mathbf{R}, \boldsymbol{T}} \sum_{i=1}^n \| \ w_i(s\mathbf{R}\boldsymbol{X}_i + \boldsymbol{T} - \boldsymbol{Y}_i) \ \|_q^q. \end{aligned} \tag{17}$$

The details about the ADMM subproblems can be found in Appendix C. The orthonormal matrix [59] method is used to

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

TABLE I
DETAILED SETTINGS OF THE COMPARED ALGORITHMS IN SIMULATIONS

Method	Parameters	Implementation
Welsch M-estimator	Tuning constant: 2.985; maximum iterations of IRLS: 100.	MATLAB code; single thread https://www.mathworks.com/help/stats/robustfit.html
Tukey S-estimator	Number of sampling trails: 500; tuning constant: 4.685; maximum iterations of IRLS: 100.	MATLAB code; single thread http://rosa.unipr.it/FSDA/fsdadownload.html
GpbM-estimator	Number of sampling trails: 1000; kernel type: Gaussian; maximum iterations of scale estimate: 2000.	MATLAB&C++ code; single thread http://rci.rutgers.edu/ meer/RIUL/research/code/GPBM/index.html
Q-norm estimation	Parameter q: 0.2; penalty coefficient: ρ: 3×10^{-6}; step size α: 1.45; maximum iterations of IRLS: 50; maximum iterations of ADMM: 100.	MATLAB code; single thread http://www.escience.cn/people/lijiayuan/index.html
RANSAC	Transformation model: affine; Confidence: 0.99; inlier threshold ξ: 3 pixels; maximum iterations: 10^4.	C++ code; single thread https://www.peterkovesi.com/matlabfns/index.html#robust
LO-RANSAC	Transformation model: affine; Confidence: 0.99; inlier threshold ξ: 3 pixels; maximum iterations: 10^4; local iterations: 10.	MATLAB code; single thread https://www.peterkovesi.com/matlabfns/index.html#robust
FLO-RANSAC	Transformation model: homography; Confidence: 0.99; inlier threshold ξ: 3 pixels; maximum iterations: 10^4; local subset size: 21; local iterations: 50.	MATLAB code; single thread https://www.peterkovesi.com/matlabfns/index.html#robust
Our scaled Welsch q-norm	Parameter q: 0.2; penalty coefficient: ρ: 3×10^{-6}; step size α: 1.45; maximum iterations of IRLS: 50; maximum iterations of ADMM: 100.	MATLAB code; single thread

solve the seven-parameter rigid transformation, which provides a close-form solution and does not require parameter initial values.

V. EXPERIMENTS AND EVALUATIONS

In this section, parameter sensitivity analysis is first carried out. Then, both simulated and real experiments are performed to qualitatively and quantitatively evaluate the proposed scaled Welsch q-norm cost. For a comprehensive assessment, both the proposed RFM and REO algorithms are evaluated (RAO is not included since it is similar to the RFM).

A. Simulation Experimental Evaluations

This section compares the proposed method with several robust estimators and RANSAC-type methods through simulation experiments, including Welsch M-estimator, Tukey S-estimator, GpbM-estimator [32], q-norm estimation [18], RANSAC [16], LO-RANSAC [40], and FLO-RANSAC [60]. The experimental details of these methods are summarized in Table I, including parameter settings and implementations. All the experiments are performed on a laptop with an Intel Core i7-8550U at 1.8-GHz CPU, 8-Gb RAM. Five indicators are adapted to quantitatively evaluate the results, including root-mean-square error (RMSE), precision, recall, F-score, and success rate. Precision describes the correct proportion of the detected inliers under the groundtruth transformation, i.e., the proportion of observations with residuals less than ξ. Recall is the ratio of the detected inlier number and the groundtruth inlier number. F-score combines the two indicators of precision and recall. The calculation formulas of F-score and RMSE are as follows:

$$\begin{cases} F\text{-score} = \dfrac{2\text{Pr} * \text{Re}}{\text{Pr} + \text{Re}} \\ \text{RMSE} = \sqrt{\dfrac{1}{n}\sum\limits_{1}^{n} v_i^2} \end{cases} \tag{18}$$

where Pr and Re denote precision and recall, respectively. To alleviate the randomness of simulation experiments, 1000 independent tests (one group) are repeated for each experiment under the same configurations. Thus, success rate is the ratio of successful estimation times in 1000 tests. If the RMSE of a test is less than ξ, the estimation is successful. We use the average result of one group test for evaluation. Namely, subsequently reported RMSE, precision, recall, F-score, and success rate accuracies are all the average values of 1000 independent tests.

1) Parameter Sensitivity Analysis: Parameter sensitivity analysis is performed by simulating image feature matching process. First, 100 2-D feature points $\{\boldsymbol{x}_i\}_1^{100}$ are randomly generated. The horizontal and vertical coordinates of $\{\boldsymbol{x}_i\}_1^{100}$ are distributed within the interval of $[-500,\ 500]$ pixels. Then, a groundtruth global affine transformation $\mathbf{T}^{gt}$ is randomly generated

$$\mathbf{T}^{gt} = \begin{bmatrix} 1 & \tan\kappa & t_x \\ \tan\varphi & 1+\tan\varphi\tan\kappa & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x\cos\theta & s_x\sin\theta & 0 \\ -s_y\sin\theta & s_y\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{19}$$

where $\boldsymbol{t} = [t_x\ t_y]^T$ is a translation; φ and κ are shear angles; θ is a rotation angle; s_x and s_y are anisotropic scales. Specifically, we randomly generate a rotation angle within $[-\pi/2,\ \pi/2]$, two shear angles φ and κ within $[-\pi/6,\ \pi/6]$, and anisotropic scales within $[0.5,\ 1.5]$. We use an average value of $\{\boldsymbol{x}_i\}_1^{100}$ as the translation $\boldsymbol{t}$. The groundtruth correspondences $\{\boldsymbol{y}_i^{\text{true}}\}_1^{100}$ of $\{\boldsymbol{x}_i\}_1^{100}$ can be obtained by

$$\boldsymbol{y}_i^{\text{true}} = \mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} \tag{20}$$

where $\mathbf{A} = \mathbf{T}^{gt}_{(1:2,1:2)}$ is a 2×2 submatrix of $\mathbf{T}^{gt}$ in the top left.

To make the simulation realistic, random noise within $[-2,\ 2]$ pixels are added to $\{\boldsymbol{y}_i^{\text{true}}\}_1^{100}$, obtaining $\{\bar{\boldsymbol{y}}_i\}_1^{100}$. We then randomly select 50 feature points from $\{\bar{\boldsymbol{y}}_i\}_1^{100}$ to add

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

TABLE II
DETAILS OF PARAMETER SETTINGS

Experiment	Variable	Fixed parameters
Parameter q study	$q = [0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9]$	$\rho = 3 \times 10^{-4}$, $\alpha = 1.65$
Parameter α study	$\alpha = [0.85,\ 1.05,\ 1.25,\ 1.45,\ 1.65,\ 1.85,\ 2.05,\ 2.25,\ 2.45]$	q= 0.2, $\rho = 3 \times 10^{-4}$
Parameter ρ study	$\rho = 6 \times 10^{-6} \times 5^{\chi}$, $\chi = [1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9]$	q= 0.2, $\alpha = 1.45$

TABLE III
PARAMETER q STUDY

Method	Metric	q, $\rho = 3 \times 10^{-4}$, $\alpha = 1.65$								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Q-norm estimation	F-score/%	97.3	96.9	94.4	93.8	91.9	79.7	59.5	32.8	5.1
	Success rate/%	96.4	96.2	92.2	91.4	88.0	73.2	50.4	24.4	2.6
Scaled Welsch q-norm	F-score/%	100	100	100	99.9	100	99.9	100	100	100
	Success rate/%	100	100	100	100	100	100	100	100	100

TABLE IV
PARAMETER α STUDY

Method	Metric	α, q= 0.2, $\rho = 3 \times 10^{-4}$								
		0.85	1.05	1.25	1.45	1.65	1.85	2.05	2.25	2.45
Q-norm estimation	F-score/%	0.1	0.1	18.9	99.6	96.5	92.3	79.5	62.2	41.4
	Success rate/%	0.0	0.0	0.0	99.6	95.4	89.4	72.8	51.8	29.8
Scaled Welsch q-norm	F-score/%	79.2	92.6	99.9	100	100	100	100	100	100
	Success rate/%	49.6	88.9	100	100	100	100	100	100	100

TABLE V
PARAMETER ρ STUDY

Method	Metric	φ, $\rho = 6 \times 10^{-6} \times 5^{\varphi}$, q= 0.2, $\alpha = 1.45$								
		1	2	3	4	5	6	7	8	9
Q-norm estimation	F-score/%	99.6	99.3	99.2	99.6	99.6	99.2	99.8	67.6	2.3
	Success rate/%	99.4	99.0	99.0	100	100	99.4	100	66.4	2.0
Scaled Welsch q-norm	F-score/%	100	100	100	100	100	100	100	96.4	99.0
	Success rate/%	100	100	100	100	100	100	100	93.2	99.1

errors and obtain the final observed correspondences $\{\boldsymbol{y}_i\}_1^{100}$. The errors are distributed in $[-500,\ 500]$ pixels, i.e., the entire image region. The 50 matches with errors are outliers and the outlier rate is 50%.

Next, we use $\{\boldsymbol{x}_i, \boldsymbol{y}_i\}_1^{100}$ as the input and perform an RFM algorithm to estimate an affine transformation $\mathbf{T}^e$. The residual of each feature correspondence is calculated based on the transformation $\mathbf{T}^e$. Feature matches whose residuals are larger than $\xi = 3$ pixels are removed as outliers, obtaining the filtered feature matches $I = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^{K}$, where K is the number of remaining matches.

There are three main parameters in both standard q-norm estimation and the proposed scaled Welsch q-norm cost: q, α, and ρ. The parameter q controls the shape of the cost function of q-norm, the value of which ranges from 0 to 1. Parameter ρ is a penalty coefficient. ρ is not a fixed value, which varies with the number of iterations. Thus, ρ herein represents the initial penalty. α is an iteration step size. After an iteration, we have $\rho^{k+1} = \alpha\rho^k$. Based on the above simulation process, three independent experiments are designed to study these parameters, where each experiment has only one parameter as a variable, and others are constant. The experimental details can be found in Table II. The results of both methods are reported in Tables III–V.

From the results of the standard q-norm estimation, we can learn the following:

1) Larger values of q result in worse performance. However, if q is too small, this method will be sensitive to noise.
2) The standard q-norm estimation totally fails when the values of α are small, such as α= 0.85, α= 1.05, and α= 1.25. Too large values of α also seriously influence the performance, i.e., larger values of α result in lower accuracies, such as α= 2.25 and α= 2.45.
3) The initial penalty coefficient ρ has a relatively small impact on the performance.

In general, the standard q-norm estimation has very strong dependence on parameters. Different parameters may lead to completely different results, which greatly affects its practicability. In contrast, the results show that the proposed scaled Welsch q-norm cost is not sensitive to parameters. Different parameters have little effect on the proposed method. Only too small α values and too large ρ values will decrease the performance. Obviously, the proposed cost has much better

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

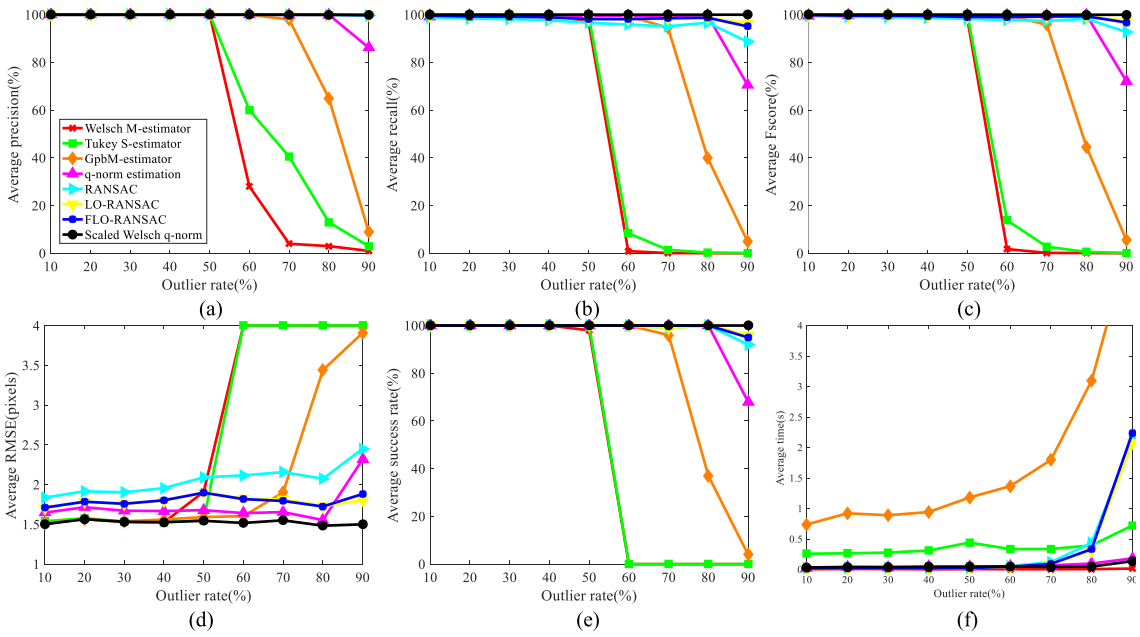

Fig. 3. RFM results of the eight compared methods on simulated data, including (a) precision, (b) recall, (c) F-score, (d) RMSE, (e) success rate, and (f) running time.

practicability than the original q-norm estimation. We fix $q = 0.2$, $\alpha = 1.45$, and $\rho = 3 \times 10^{-6}$ for the following experiments.

2) RFM Simulation: The simulation process is similar to the one in the parameter sensitivity analysis section. The only difference is that the parameter sensitivity section sets the outlier rate to 50% and fixes it, while this section treats the outlier rate (denoted by r_{out}) as a variable. We fix the number of inlier feature matches to 50, and sequentially increase the outlier rate r_{out} from 10% to 90%. For each value of the outlier rate r_{out}, we generate $n = 50/(1 - r_{\text{out}})$ feature matches. Then, we randomly select $n - 50$ feature matches to add errors.

Fig. 3(a)–(c) reflects the precision, recall, and F-score accuracies of each compared method, respectively. As can be seen, when the outlier rate is less than 50%, Welsch M-estimator and Tukey S-estimator work very well. Their performances are comparable with the proposed scaled Welsch q-norm cost. Unfortunately, once the outlier rate is higher than 50%, these methods are completely ineffective. Therefore, their curves in Fig. 3 show a "cliff-like drop." GpbM-estimator performs better than these two traditional robust estimators, whose performance decreases after the outlier rate reaches 70%. RANSAC-type methods have very high precision accuracies and better resistance to outlier rates. However, their recall and F-score accuracies are usually not as high as the q-norm estimation. The standard q-norm estimation and the proposed scaled Welsch q-norm cost achieve the second best and the best performances in terms of F-score, respectively. The proposed method is more robust than the standard q-norm estimation. When the outlier rate is within 80%, the accuracy of the scaled Welsch q-norm cost is consistent with the standard q-norm estimation. Once the outlier rate exceeds 80%, the proposed method is obviously better than the standard q-norm estimation. Our cost is still robust even if the outlier rate is up to 90%. Fig. 3(d) plots the RMSE results. Note that if a test fails, we assign its RMSE to two times the noise level (the noise level is two pixels in this experiment). As shown, the RMSEs of RANSAC, LO-RANSAC, and FLO-RANSAC are the largest, because RANSAC-type methods have no guarantee of the optimality of the estimated solutions and are sensitive to noise. Welsch M-estimator, Tukey S-estimator, and GpbM-estimator get similar RMSEs with the proposed scaled Welsch q-norm cost if the outlier rate is less than 50%. Our RMSE is better than the one of the standard q-norm estimation, which means that our model estimation accuracy is higher than the original q-norm estimation. Fig. 3(e) reports the success rate. Again, the proposed method ranks the best, whose success rate is still almost 100% under 90% of outliers. In terms of running time, Welsch M-estimator is the fastest while GpbM-estimator is the slowest. Standard q-norm estimation is comparable to the proposed method. RANSAC-type methods become much slower as the outlier rate increases. When the outlier rate is 90%, our method is 15+ times faster than RANSAC-type methods.

3) REO Simulation: We evaluate the proposed REO algorithm by simulating the process of camera perspective projection. Specifically, suppose that we are given a well-calibrated pinhole camera with a 1500-pixels focal length, a zero principal point offset, and a 2000 $\times$ 2000 pixels image size. We first randomly generate n 3-D image space points $\{\boldsymbol{Q}_i^c\}_1^n$ inside a square box of $[-8, 8] \times [-8, 8] \times [8, 16]$. We use the mean of $\{\boldsymbol{Q}_i^c\}_1^n$ as a translation vector $\boldsymbol{T}$, and randomly generate a 3×1 rotation angle vector to construct a 3×3 Rodrigues rotation matrix $\mathbf{R}_{3\times3}$. The translation vector and rotation matrix $(\mathbf{R}_{3\times3}, \boldsymbol{T})$ are regarded as the groundtruth camera exterior parameters. Based on these parameters, we can obtain the corresponding object space points $\{\boldsymbol{Q}_i\}_1^n$ of $\{\boldsymbol{Q}_i^c\}_1^n$

$$\boldsymbol{Q}_i = \mathbf{R}_{3\times3}^{-1}\left(\boldsymbol{Q}_i^c - \boldsymbol{T}\right). \tag{21}$$

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

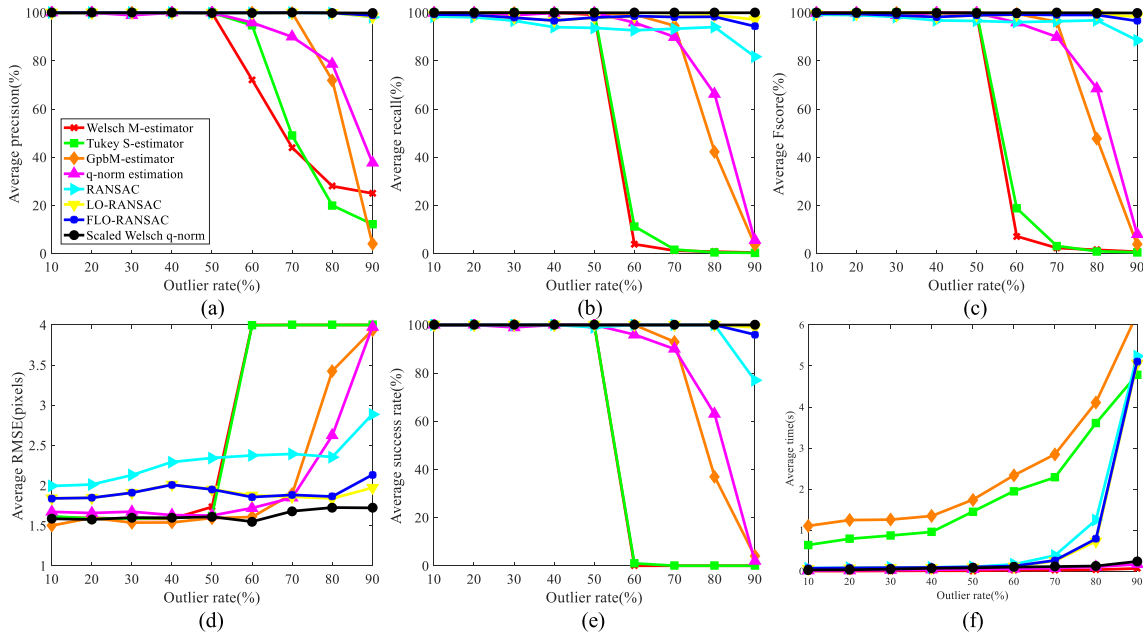

Fig. 4. REO results of the eight compared methods on the simulated data, including (a) precision, (b) recall, (c) F-score, (d) RMSE, (e) success rate, and (f) running time.

Projecting $\{\boldsymbol{Q}_i\}_1^n$ to the image, obtaining the corresponding groundtruth image points $\{\boldsymbol{p}_i^{\text{true}}\}_1^n$

$$d_i \begin{bmatrix} \boldsymbol{p}_i^{\text{true}} \\ 1 \end{bmatrix} = \mathbf{K}\boldsymbol{Q}_i^c. \tag{22}$$

Similarly, we add random noise within $[-2,\ 2]$ pixels to $\{\boldsymbol{p}_i^{\text{true}}\}_1^n$, obtaining $\{\bar{\boldsymbol{p}}_i\}_1^n$. We then randomly select $r_{\text{out}} \times n$ points according to the outlier rate r_{out} to add errors and get the final observed image points $\{\boldsymbol{p}_i\}_1^n$. The set $\{(\boldsymbol{Q}_i, \boldsymbol{p}_i)\}_1^n$ is the 3-D–2-D correspondences contaminated by outliers. In our experiments, we fix the number of inliers to 50 and increase the outlier rate from 10% to 90%.

As mentioned above, the camera projection model is estimated based on the Gauss–Newton method, which requires initial guesses for parameters. In our experiments, the initial angle vector is obtained by randomly adding a 3×1 rotation angle vector between $[-10°,\ 10°]$ to the groundtruth angle vector. The initial translation vector is randomly generated between $[70\% \times \boldsymbol{T},\ 130\% \times \boldsymbol{T}]$. The comparison results are shown in Fig. 4.

From Fig. 4, we can obviously see that there is a breakpoint in the Welsch M-estimator and Tukey S-estimator. They usually fail when the outlier rate is higher than 50%. GpbM-estimator is able to handle cases with less than 70% of outliers. However, it failed completely if the outlier rate reaches 90%. RANSAC-type methods perform much better than typical robust estimators. Their precision accuracies are still close to 100% when the outlier rate is up to 90%. However, their recall accuracies are worse than the proposed method. The performance of the standard q-norm estimation is not as good as in the RFM task. This may be expected. We use the parameters studied in the RFM task for REO evaluation. As mentioned earlier, the standard q-norm estimation is very sensitive to parameters. Thus, the parameters learned from RFM are not the best choice for REO. In contrast, the proposed scaled Welsch q-norm cost achieves the best results. The precision, recall, and F-score are close to 100% even if the outlier ratio is 90%. The proposed method overcomes the drawback of the standard q-norm estimation. Fig. 4(d) reflects the RMSE accuracy of each method. Again, the RMSE accuracy of the proposed method is the smallest. It is almost unaffected by outlier rates. RANSAC-type methods have the lowest RMSE accuracies. They use three-correspondence subsets to solve the solutions, where not all observation data are involved in the optimization stage. Thus, RANSAC-type methods are very sensitive to noise, which make the estimated parameters are not optimal. In this task, we use an iterative Gauss–Newton method to solve the exterior orientation model. Hence, methods with a hypothesize stage become much slower, such as RANSAC-type methods. When the outlier rate is 90%, our method is 20+ times faster than RANSAC-type methods and 25+ times faster than GpbM-estimator.

B. Real Experiments

1) RFM Real Experiment: Apart from RANSAC-type methods, many state-of-the-art RFM methods have been presented recently, such as locally linear transforming (LLT) [34], iterative biconvex optimization (IBCO) [61], locality preserving matching (LPM) [62], guided locality preserving matching (GLPM) [63], q-norm estimation [18], locality affine-invariant matching (LAM) [17], and learning for mismatch removal (LMR) [64]. We select LLT, FLO-RANSAC, IBCO, and q-norm estimation for comparison. M-estimators and S-estimators are not included in this section since they completely fail if the outlier rate is higher than 50%. The parameters of each method are fine-tuned to obtain the best performance and are consistent in all experiments. The implementation code of each method is obtained from the author's personal website.

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

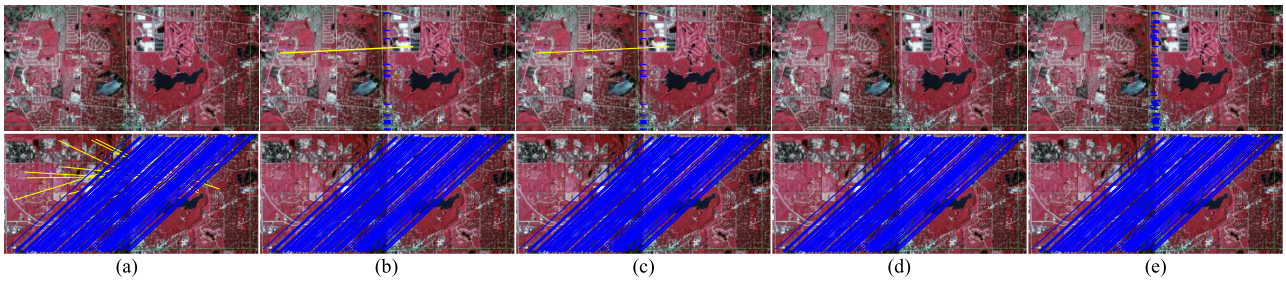

Fig. 5. Qualitative comparisons. First row: result of the third image pair. Second row: results of the last one. Blue lines: correct detected inliers. Yellow lines: false detected inliers. (a) LLT. (b) FLO-RANSAC. (c) IBCO. (d) q-norm estimation. (e) Weighted q-norm estimation.

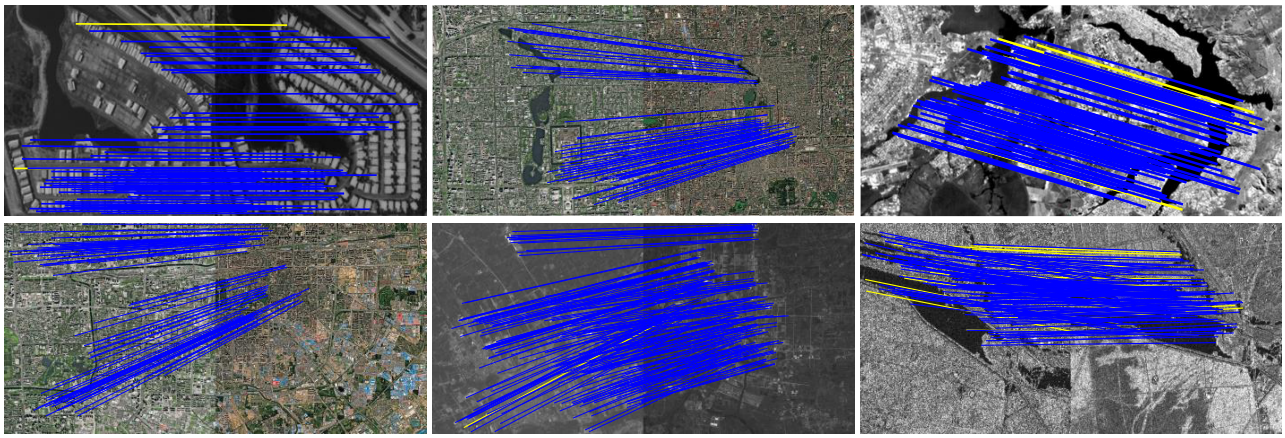

Fig. 6. More qualitative results of the proposed RFM method. Blue lines: correct detected inliers. Yellow lines: false detected inliers.

We select 11 false-color aerial image pairs from the Erdas sample data[1] that were captured over IL, USA. The image sizes of these data are distributed from 1391 × 1374 pixels to 1459 × 1380 pixels. The ground sampling distance is about 0.2 m and the overlapping regions are extremely small. Groundtruth geometric transformations are established for quantitative evaluation. Specifically, we manually select five uniformly distributed matches with sub-pixel accuracy for each image pair and estimate an accurate affine transformation by linear least squares as the approximate groundtruth transformation. We use the scale invariant feature transform (SIFT) [65] method (note that the other feature matching method can also be applied such as speeded up robust features (SURF) [66] and radiation-variation insensitive feature transform (RIFT) [67]) with an nearest neighbor distance ratio (NNDR) of 0.9 to produce initial matches. Then, we apply the established groundtruth transformation to the initial matches and treat the matches with less than three-pixel residual errors as groundtruth inliers. We use the precision and recall metrics as the quantitative evaluation indicators.

Fig. 5 gives the qualitative comparison results on the third and last aerial image pairs. The horizontal overlapping region of the third image pair is less than 5% of the image width. The vertical overlapping region of the last image pair is less than 5% of the image height. Such small overlapping regions result in high proportions of outliers, which pose a huge challenge to image matching problem. As shown, LLT and standard q-norm estimation failed completely on the third image pair and no correspondences are extracted. FLO-RANSAC is obviously superior to the above methods. It achieves good precision accuracy but very poor recall accuracy. IBCO gets similar results to FLO-RANSAC, since it takes the results of FLO-RANSAC as its initial solutions and refines these solutions. In addition, even if the matching of LLT is successful, many false matches are still preserved in the last image pair. Only the proposed scaled Welsch q-norm cost achieves good performance on both the image pairs. There are almost no outliers in the results, which means that the precision accuracies are 100%. In addition, the proposed method extracts many more reliable matches than FLO-RANSAC and IBCO. More qualitative results of the proposed RFM method can be found in Fig. 6.

TABLE VI
INITIAL PARAMETER VALUES

Image ID	φ/°	ω/°	κ/°	Xs/m	Ys/m	Zs/m
1	0	0	0	0	0	500
2	0	0	-30	0	0	500

Fig. 7 reports the quantitative comparison results, where the left figure plots the inlier rate cumulative distribution of the initial matches. As shown, the average outlier rate of this data set is 91.99%. These images are very challenging due to such high outlier rates. The right figure summarizes the matching results, where the dots represent the (precision, recall) pairs. It can be seen that only FLO-RANSAC, IBCO, and the proposed method get high precision performance; and only our proposed method gets high recall performance on all image pairs. The average (precision, recall) pairs of LLT,

[1]http://download.intergraph.com/downloads/erdas-imagine-2013-2014-example-data

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

TABLE VII
REO RESULTS ON TWO LARGE-SCALE REAL IMAGES

Method	Image ID	φ/°	ω/°	κ/°	Xs/m	Ys/m	Zs/m	RMSE/pixels	True/False
LHM	1	-11.98	5.95	-20.39	-154.34	-39.41	-631.59	44.20	False
	2	-19.03	41.64	63.13	-154.34	-39.41	-631.58	256.23	False
EPnP+GN	1	-3.436	0.98	-19.80	19.289	37.57	650.26	1.99	True
	2	17.92	-44.75	-117.56	658.17	161.18	646.31	2.54	True
RPnP	1	-16.67	3.50	-19.82	-205.98	-10.50	-617.84	46.09	False
	2	-19.24	41.62	63.06	627.54	190.29	-684.32	256.90	False
DLS	1	-13.53	5.67	-20.37	-171.50	-35.70	-627.12	43.65	False
	2	-19.53	43.03	62.86	640.38	186.99	-663.04	257.03	False
OPnP	1	-13.93	5.63	-20.35	-176.03	-35.12	-626.10	43.62	False
	2	-19.71	42.91	62.70	647.65	192.06	-672.89	256.51	False
REPPnP	1	-11.29	6.03	-19.33	-148.15	-40.92	-640.06	62.72	False
	2	5.43	3.83	59.19	49.54	-101.96	-705.83	1585.29	False
l_qPnP	1	-3.19	0.60	-19.80	16.50	33.05	650.58	1.05	True
	2	17.98	-44.69	-117.57	657.96	162.18	647.57	1.14	True
Proposed method	1	-3.22	0.58	-19.79	16.86	32.81	650.59	0.82	True
	2	17.96	-44.67	-117.56	657.85	162.09	647.90	0.87	True
PATB	1	-3.21	0.54	-19.80	16.74	32.82	650.60	1.05	True
	2	17.99	-44.69	-117.58	657.72	162.19	646.99	1.19	True

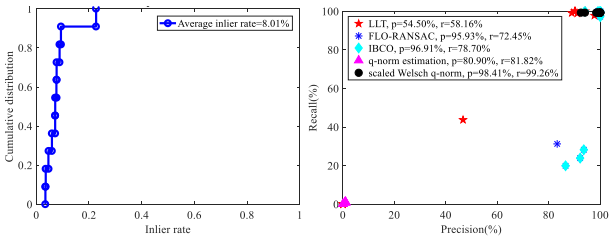

Fig. 7. Quantitative comparison results. (Left) Inlier rate cumulative distribution of the initial matches. (Right) Matching results, where the dots represent the (precision, recall) pairs.

(a) (b)

Fig. 8. Test images for real experiment. Red circles: control points. (The right-hand side image is rotated for better visualization.). (a) Central image. (b) Oblique image.

FLO-RANSAC, IBCO, q-norm estimation, and the proposed method are (54.5%, 58.16%), (95.93%, 72.45%), (96.91%, 78.70%), (80.90%, 81.82%), and (98.42%, 99.26%), respectively. The proposed method achieves (1.50%, 20.56%) growth rates compared with IBCO.

2) REO Real Experiment: Two large-scale aerial images are used for real experimental evaluation, which are obtained by Si Wei digital camera (SWDC) camera system over the Pingdingshan, Henan, China. The SWDC camera system contains four oblique cameras and a central camera. As shown in Fig. 8, the left-hand side image is acquired by the central camera with a focal length of 12102.1 pixels. Its image size is 5406 × 7160 pixels. The right one with an image size of 7160 × 5406 pixels is acquired by an oblique camera with a focal length of 14671.5 pixels. We measure 12 and 15 3-D ground control points by using the GPS-RTK technique for Fig. 8(a) and (b), respectively. We manually find the corresponding 2-D image points of these 3-D control points, obtaining the 3-D–2-D correspondences. The location precision of these 2-D image points is better than 0.3 pixels.

We use GPS-RTK to obtain 2-D–3-D correspondences, which are exact and do not contain outliers. Thus, we compare our method with several camera exterior orientation methods (both robust methods and nonrobust methods), including LHM [68], EPnP+GN [69], RPnP [70], DLS [71], OPnP [72], REPPnP [73], and l_qPnP [19] (the standard q-norm estimation). All the implementation codes except for l_qPnP are provided by Ferrazs PnP MATLAB toolbox. The default parameters of each method are used.

The proposed algorithm is an iterative optimization algorithm, which requires initial parameter guesses. The initial external orientation parameters (include three angle parameters $[\varphi, \omega, \kappa]$ and three location parameters $[Xs, Ys, Zs]$) are given in Table VI. The results are summarized in Table VII.

From the results, we can see that only four of the total nine methods, i.e., EPnP+GN, PATB, standard q-norm estimation, and the proposed method, are able to correctly recover the camera orientation parameters. The other five methods all get the wrong solutions, whose coordinates in the elevation direction are negative. The reasons may be two-fold: first, our cases are large-scale (approximately 450 m × 450 m × 30 m), while most of the compared PnP methods only evaluated on very small-scale cases in their articles. Second, 3-D points in our cases can be approximately regarded as locating on a plane (near-planar case). Hence, closed-form methods may not

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

be as reliable as the iterative methods such as EPnP+GN and our method. Among these four successful methods, the proposed method achieves the best reprojection accuracy. The RMSEs are only 0.82 and 0.87 pixels, which are even better than the commercial software PATB. Moreover, it can be found that when the observations do not suffer from outliers, the proposed algorithm has low dependence on the initial parameter values. For instance, the true orientation parameters of the right image are [17.964, −44.673, −117.563, 657.875, 162.093, 647.846]. However, the given initial values are only [0, 0, −30, 0, 0, 500].

To verify the robustness of the proposed method on the real data, 30 randomly generated outliers are added to the 3-D–2-D correspondences. The outlier rates are 71.4% and 66.7% for the left-hand side and the right-hand side test images, respectively. The average RMSEs of the proposed method are 1.22 and 1.25 pixels on 100 independent tests, respectively.

VI. Conclusion

This article proposes a novel robust cost function for geometric model estimation and outlier removal, which is called scaled Welsch q-norm ($0 < q < 1$). We integrate a scaled Welsch weight function into a q-norm framework to improve the robustness to outlier rates. We adopt ALF to reformulate the new cost function and use ADMM to decompose the ALF into two subproblems, namely, an l_qLS problem and a WLS estimation problem. We introduce a coarse-to-fine strategy into the IRLS method, i.e., we change the weight function by decreasing its scale parameter along iterations. This strategy can largely avoid that the solver gets stuck in local minimums. We also introduce the proposed cost into typical remote sensing applications and develop three algorithms, including RFM, REO, and RAO. These three algorithms are very important for geometric processing in remote sensing. Note that the proposed cost is also suitable for other geometric model estimation tasks. Based on the exhaustive experimental evaluations, we can make some conclusions. First, the proposed cost is robust with respect to gross errors and can tolerate high outlier rates, which overcomes the limitation of the classical robust estimators. Second, our method is less sensitive to parameters compared with the standard q-norm estimation, which makes it more flexible in real-life applications.

The limitations of the proposed method are twofold.

1) The proposed method cannot absolutely guarantee that the globally optimal solution can be achieved. Although we use a coarse-to-fine optimization strategy in IRLS, the proposed method still cannot absolutely avoid getting stuck in local minimums, especially good initial guesses for parameters are required (such as the REO problem).
2) The proposed cost cannot handle cases with extremely high outlier rates such as 95%. In addition, although the proposed method is less sensitive to parameters compared with the standard q-norm cost, the parameters cannot be set arbitrarily.

Appendix A
ADMM Decomposition and Problem $\boldsymbol{part1}$ Optimization

ADMM decomposes (9) into two main parts

$$\begin{aligned}
\boldsymbol{RFM_{part1}}: \ & \arg\min_{\mathbf{M}} L_\rho \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \begin{pmatrix} \| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \frac{\boldsymbol{\lambda}_i}{\rho} \\ +w_i^{k-1}(\mathbf{A}^{k-1}\boldsymbol{x}_i + \boldsymbol{t}^{k-1} - \boldsymbol{y}_i) - \boldsymbol{m}_i \|_2^2 \end{pmatrix} \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \boldsymbol{\beta}_i - \boldsymbol{m}_i \|_2^2 \right) \qquad \text{(A.1)} \\
\boldsymbol{RFM_{part2}}: \ & \arg\min_{\mathbf{A},\boldsymbol{t},W} L_\rho \\
= \ & \arg\min_{\mathbf{A},\boldsymbol{t},W} \sum_{i=1}^{n} \begin{pmatrix} \| \boldsymbol{m}_i^{k-1} \|_q^q + \frac{\rho}{2} \| \frac{\boldsymbol{\lambda}_i}{\rho} \\ +w_i(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{y}_i) - \boldsymbol{m}_i^{k-1} \|_2^2 \end{pmatrix} \\
= \ & \arg\min_{\mathbf{A},\boldsymbol{t},W} \sum_{i=1}^{n} \frac{\rho}{2} \left\| w_i \left[\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \left(\boldsymbol{y}_i + \frac{\boldsymbol{m}_i^{k-1}}{w_i} - \frac{\boldsymbol{\lambda}_i}{\rho w_i} \right) \right] \right\|_2^2 \\
\to \ & \arg\min_{\mathbf{A},\boldsymbol{t},W} \sum_{i=1}^{n} \frac{\rho}{2} \left\| w_i \left[\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \left(\boldsymbol{y}_i + \frac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}} \right) \right] \right\|_2^2 \\
= \ & \arg\min_{\mathbf{A},\boldsymbol{t},W} \sum_{i=1}^{n} \frac{\rho}{2} \left\| w_i(\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{\gamma}_i) \right\|_2^2 \qquad \text{(A.2)}
\end{aligned}$$

where the superscript k is an iteration counter. Note that w_i^{k-1} is a known value while w_i is an unknown to be estimated. In $\boldsymbol{part2}$, $w_i = e^{-(\|\mathbf{A}\boldsymbol{x}_i + \boldsymbol{t} - \boldsymbol{\gamma}_i\|/u)^2}$.

For $part1$, $(\mathbf{A}, \boldsymbol{t}, W)$ is fixed and $\mathbf{M}$ is the variable. Each element in $\boldsymbol{m}_i$ can be optimized independently by adapting a shrinkage algorithm, i.e., performing the minimization on the separated (10) along the jth ($j = 1, 2$) coordinate. Consequently, $part1$ is simplified to the scalar version

$$\arg\min_{m} \left(|m|^q + \frac{\rho}{2}(\beta - m)^2 \right). \qquad \text{(A.3)}$$

Then, the optimal solution $\hat{m}$ is given by [56]

$$\hat{m} = \begin{cases} 0, & \text{if} |\beta| \le \tau \\ \text{sgn}(\beta)\phi, & f|\beta| > \tau \end{cases} \qquad \text{(A.4)}$$

where $\tau = \varpi + (q/\rho)\varpi^{q-1} \ \ s.t. \ \ \varpi = [2/\rho(1-q)]^{1/(2-q)}$; $\text{sgn}(\cdot)$ is the signum function; $\phi \in (\varpi, \ |\beta|)$ is computed from the following iteration with initial guess $\phi^0 = (\varpi + |\beta|)/2$:

$$\phi^{k+1} = |\beta| - \frac{q}{\rho}(\phi^k)^{q-1}. \qquad \text{(A.5)}$$

The shrinkage operator acts as a classifier which automatically classifies the residual vector into a possible inlier set and an outlier set.

Appendix B
Subproblems of REO via ADMM Decomposition

With the ADMM decomposition, the REO problem can be simplified into two main subproblems. Then, (10)–(12) are

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

detailed as follows:

$$\begin{aligned}
\boldsymbol{REO}_{\boldsymbol{part1}}: \quad & \arg\min_{\mathbf{M}} L_\rho \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\begin{array}{c} \| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \left\| \frac{\boldsymbol{\lambda}_i}{\rho} - \boldsymbol{m}_i \right. \\ \left. + w_i^{k-1} \left(\frac{\mathbf{P}_{1:2}^{k-1} [\, \boldsymbol{Q}_i \ 1 \,]^T}{\mathbf{P}_3^{k-1} [\, \boldsymbol{Q}_i \ 1 \,]^T} - \boldsymbol{p}_i \right) \right\|_2^2 \end{array} \right) \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \boldsymbol{\beta}_i - \boldsymbol{m}_i \|_2^2 \right) \qquad \text{(B.1)} \\
\boldsymbol{REO}_{\boldsymbol{part\ 2}}: \quad & \arg\min_{\mathbf{R},\, \boldsymbol{T},\, W} L_\rho \\
\rightarrow \ & \arg\min_{\mathbf{R},\, \boldsymbol{T},\, W} \sum_{i=1}^{n} \left(\begin{array}{c} \frac{\rho}{2} \left\| w_i \left[\frac{\mathbf{P}_{1:2} [\, \boldsymbol{Q}_i \ 1 \,]^T}{\mathbf{P}_3 [\, \boldsymbol{Q}_i \ 1 \,]^T} \right. \right. \\ \left. \left. - \left(\boldsymbol{p}_i + \frac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}} \right) \right] \right\|_2^2 \end{array} \right) \\
= \ & \arg\min_{\mathbf{R},\, \boldsymbol{T},\, W} \sum_{i=1}^{n} \frac{\rho}{2} \left\| w_i \left(\frac{\mathbf{P}_{1:2} [\, \boldsymbol{Q}_i \ 1 \,]^T}{\mathbf{P}_3 [\, \boldsymbol{Q}_i \ 1 \,]^T} - \boldsymbol{\gamma}_i \right) \right\|_2^2 \qquad \text{(B.2)} \\
& \begin{cases} \boldsymbol{\beta}_i = \dfrac{\boldsymbol{\lambda}_i}{\rho} + w_i^{k-1} \left(\dfrac{\mathbf{P}_{1:2}^{k-1} [\, \boldsymbol{Q}_i \ 1 \,]^T}{\mathbf{P}_3^{k-1} [\, \boldsymbol{Q}_i \ 1 \,]^T} - \boldsymbol{p}_i \right) \\ \boldsymbol{\gamma}_i = \boldsymbol{p}_i + \dfrac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \dfrac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}} \end{cases} . \qquad \text{(B.3)}
\end{aligned}$$

APPENDIX C
SUBPROBLEMS OF RAO VIA ADMM DECOMPOSITION

With the ADMM decomposition, the RAO problem can also be simplified into two main subproblems. Then, (10)–(12) are detailed as follows:

$$\begin{aligned}
\boldsymbol{RAO}_{\boldsymbol{part1}}: \quad & \arg\min_{\mathbf{M}} L_\rho \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\begin{array}{c} \| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \left\| \frac{\boldsymbol{\lambda}_i}{\rho} - \boldsymbol{m}_i \right. \\ \left. + w_i^{k-1} (s^{k-1} \mathbf{R}^{k-1} \boldsymbol{X}_i + \boldsymbol{T}^{k-1} - \boldsymbol{Y}_i) \right\|_2^2 \end{array} \right) \\
= \ & \arg\min_{\mathbf{M}} \sum_{i=1}^{n} \left(\| \boldsymbol{m}_i \|_q^q + \frac{\rho}{2} \| \boldsymbol{\beta}_i - \boldsymbol{m}_i \|_2^2 \right) \qquad \text{(C.1)} \\
\boldsymbol{RAO}_{\boldsymbol{part\ 2}}: \quad & \arg\min_{s, \mathbf{R},\, \boldsymbol{T},\, W} L_\rho \\
\rightarrow \ & \arg\min_{s, \mathbf{R}, \boldsymbol{T}, W} \sum_{i=1}^{n} \frac{\rho}{2} \left\| w_i \left[s\mathbf{R}\boldsymbol{X}_i + \boldsymbol{T} - \left(\boldsymbol{Y}_i + \frac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \frac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}} \right) \right] \right\|_2^2 \\
= \ & \arg\min_{s, \mathbf{R}, \boldsymbol{T}, W} \sum_{i=1}^{n} \frac{\rho}{2} \| w_i (s\mathbf{R}\boldsymbol{X}_i + \boldsymbol{T} - \boldsymbol{\gamma}_i) \|_2^2 \qquad \text{(C.2)} \\
& \begin{cases} \boldsymbol{\beta}_i = \dfrac{\boldsymbol{\lambda}_i}{\rho} + w_i^{k-1} (s^{k-1} \mathbf{R}^{k-1} \boldsymbol{X}_i + \boldsymbol{T}^{k-1} - \boldsymbol{Y}_i) \\ \boldsymbol{\gamma}_i = \boldsymbol{Y}_i + \dfrac{\boldsymbol{m}_i^{k-1}}{w_i^{k-1}} - \dfrac{\boldsymbol{\lambda}_i}{\rho w_i^{k-1}}. \end{cases} \qquad \text{(C.3)}
\end{aligned}$$

REFERENCES

[1] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[2] J. Li, Q. Hu, M. Ai, and R. Zhong, "Robust feature matching via support-line voting and affine-invariant ratios," *ISPRS J. Photogramm. Remote Sens.*, vol. 132, pp. 61–76, Oct. 2017.

[3] R. Feng, Q. Du, X. Li, and H. Shen, "Robust registration for remote sensing images by combining and localizing feature- and area-based methods," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, pp. 15–26, May 2019.

[4] X. Li, N. Hui, H. Shen, Y. Fu, and L. Zhang, "A robust mosaicking procedure for high spatial resolution remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 109, pp. 108–125, Nov. 2015.

[5] J. Li, Q. Hu, and M. Ai, "Optimal illumination and color consistency for optical remote-sensing image mosaicking," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 11, pp. 1943–1947, Nov. 2017.

[6] X. Li, R. Feng, X. Guan, H. Shen, and L. Zhang, "Remote sensing image mosaicking: Achievements and challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 4, pp. 8–22, Dec. 2019.

[7] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, Oct. 2007.

[8] G. Vivone *et al.*, "A critical comparison among pansharpening algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2565–2586, May 2015.

[9] C. Wu, S. Agarwal, B. Curless, and S. M. Seitz, "Multicore bundle adjustment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 3057–3064.

[10] S. Vassilopoulou *et al.*, "Orthophoto generation using IKONOS imagery and high-resolution DEM: A case study on volcanic hazard monitoring of Nisyros Island (Greece)," *ISPRS J. Photogramm. Remote Sens.*, vol. 57, nos. 1–2, pp. 24–38, Nov. 2002.

[11] C. Lee, J. Oh, C. Hong, and J. Youn, "Automated generation of a digital elevation model over steep terrain in antarctica from high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1186–1194, Mar. 2015.

[12] T.-J. Chin and D. Suter, "The maximum consensus problem: Recent algorithmic advances," *Synth. Lect. Comput. Vis.*, vol. 7, no. 2, pp. 1–194, Feb. 2017.

[13] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Statist.-Theory Methods*, vol. 6, no. 9, pp. 813–827, Jan. 1977.

[14] R. A. Maronna, R. D. Martin, V. J. Yohai, and M. Salibián-Barrera, *Robust Statistics: Theory and Methods*. Hoboken, NJ, USA: Wiley, 2006.

[15] P. J. Huber, *Robust Statistics*. New York, NY, USA: Wiley, 1981.

[16] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[17] J. Li, Q. Hu, and M. Ai, "LAM: Locality affine-invariant feature matching," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 28–40, Aug. 2019.

[18] J. Li, Q. Hu, and M. Ai, "Robust feature matching for remote sensing image registration based on L_q -estimator," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1989–1993, Dec. 2016.

[19] J. Li, Q. Hu, R. Zhong, and M. Ai, "Exterior orientation revisited: A robust method based on l_q-norm," *Photogramm. Eng. Remote Sens.*, vol. 83, no. 1, pp. 47–56, Jan. 2017.

[20] D. P. Bertsekas, "Nonlinear programming," *J. Oper. Res. Soc.*, vol. 48, no. 3, p. 334, 1997.

[21] S. Boyd *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011.

[22] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, vol. 589. Hoboken, NJ, USA: Wiley, 2005.

[23] X. Li, H. Shen, L. Zhang, H. Zhang, and Q. Yuan, "Dead pixel completion of Aqua MODIS band 6 using a robust M-estimator multiregression," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 4, pp. 768–772, Apr. 2014.

[24] D. J. Peter, V. K. Govindan, and A. T. Mathew, "Nonlocal-means image denoising technique using robust M-estimator," *J. Comput. Sci. Technol.*, vol. 25, no. 3, pp. 623–631, May 2010.

[25] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.

[26] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: A localized similarity measure," in *Proc. IEEE Int. Joint Conf. Neural Netw. Proc.*, Oct. 2006, pp. 4919–4924.

[27] R. He, W.-S. Zheng, B.-G. Hu, and X.-W. Kong, "A regularized correntropy framework for robust pattern recognition," *Neural Comput.*, vol. 23, no. 8, pp. 2074–2100, Aug. 2011.

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.

[28] R. He, W.-S. Zheng, and B.-G. Hu, "Maximum correntropy criterion for robust face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1561–1576, Aug. 2011.

[29] R. He, B.-G. Hu, W.-S. Zheng, and X.-W. Kong, "Robust principal component analysis based on maximum correntropy criterion," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1485–1494, Jun. 2011.

[30] B. Chen, L. Xing, H. Zhao, N. Zheng, and J. C. Príncipe, "Generalized correntropy for robust adaptive filtering," *IEEE Trans. Signal Process.*, vol. 64, no. 13, pp. 3376–3387, Jul. 2016.

[31] L. Wang and C. Pan, "Robust level set image segmentation via a local correntropy-based K-means clustering," *Pattern Recognit.*, vol. 47, no. 5, pp. 1917–1925, May 2014.

[32] S. Mittal, S. Anand, and P. Meer, "Generalized projection-based M-estimator," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 12, pp. 2351–2364, Dec. 2012.

[33] H. Chen, "Robust regression with projection based M-estimators," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 878–885.

[34] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, 2015.

[35] X. Chen, F. Xu, and Y. Ye, "Lower bound theory of nonzero entries in solutions of ℓ_2-ℓ_p minimization," *SIAM J. Sci. Comput.*, vol. 32, no. 5, pp. 2832–2852, Jan. 2010.

[36] G. Marjanovic and V. Solo, "On l_q optimization and matrix completion," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5714–5724, Nov. 2012.

[37] Z. Xu, X. Chang, F. Xu, and H. Zhang, "$L_{1/2}$ regularization: A thresholding representation theory and a fast solver," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, Jul. 2012.

[38] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, "Weighted schatten p-norm minimization for image denoising and background subtraction," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4842–4857, Oct. 2016.

[39] P. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, Apr. 2000.

[40] O. Chum, J. Matas, and J. Kittler, "Locally optimized RANSAC," in *Proc. Joint Pattern Recognit. Symp.*, 2003, pp. 236–243.

[41] R. Raguram, O. Chum, M. Pollefeys, J. Matas, and J.-M. Frahm, "USAC: A universal framework for random sample consensus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 2022–2038, Aug. 2013.

[42] E. Brachmann *et al.*, "DSAC—Differentiable RANSAC for camera localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognition*, Jul. 2017, pp. 6684–6692.

[43] J. Li, Q. Hu, and M. Ai, "Robust feature matching for geospatial images via an affine-invariant coordinate system," *Photogramm. Record*, vol. 32, no. 159, pp. 317–331, Sep. 2017.

[44] M. T. El-Melegy, "Random sampler M-estimator algorithm with sequential probability ratio test for robust function approximation via feed-forward neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 7, pp. 1074–1085, Jul. 2013.

[45] H. Li, "Consensus set maximization with guaranteed global optimality for robust geometry estimation," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 1074–1080.

[46] M. E. Pfetsch, "Branch-and-cut for the maximum feasible subsystem problem," *SIAM J. Optim.*, vol. 19, no. 1, pp. 21–38, Jan. 2008.

[47] T.-J. Chin, Y. H. Kee, A. Eriksson, and F. Neumann, "Guaranteed outlier removal with mixed integer linear programs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 5858–5866.

[48] T.-J. Chin, P. Purkait, A. Eriksson, and D. Suter, "Efficient globally optimal consensus maximisation with tree search," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 758–772, Apr. 2017.

[49] P. Hart, N. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Trans. Syst. Sci. Cybern.*, vol. SSC-4, no. 2, pp. 100–107, Jul. 1968.

[50] J. Matoušek, "On geometric optimization with few violated constraints," *Discrete Comput. Geometry*, vol. 14, no. 4, pp. 365–384, Dec. 1995.

[51] A. P. Bustos and T.-J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2868–2882, Dec. 2018.

[52] A. Dalalyan and R. Keriven, "Robust estimation for an inverse problem arising in multiview geometry," *J. Math. Imag. Vis.*, vol. 43, no. 1, pp. 10–23, May 2012.

[53] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi, "Practical low-rank matrix approximation under robust L_1-norm," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1410–1417.

[54] F. Wen, D. Zou, R. Ying, and P. Liu, "Efficient outlier removal for large scale global structure-from-motion," 2018, *arXiv:1808.03041*. [Online]. Available: https://arxiv.org/abs/1808.03041

[55] C. Olsson, A. Eriksson, and R. Hartley, "Outlier removal using duality," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1450–1457.

[56] G. Marjanovic and V. Solo, "l_q sparsity penalized linear regression with cyclic descent," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1464–1475, Mar. 2014.

[57] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 707–710, Oct. 2007.

[58] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[59] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 5, no. 7, p. 1127, Jul. 1988.

[60] K. Lebeda, J. Matas, and O. Chum, "Fixing the locally optimized RANSAC–full experimental evaluation," in *Proc. Brit. Mach. Vis. Conf.*, 2012, pp. 1–11.

[61] Z. Cai, T.-J. Chin, H. Le, and D. Suter, "Deterministic consensus maximization with biconvex programming," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 685–700.

[62] J. Ma, J. Zhao, J. Jiang, H. Zhou, and X. Guo, "Locality preserving matching," *Int. J. Comput. Vis.*, vol. 127, no. 5, pp. 512–531, May 2019.

[63] J. Ma, J. Jiang, H. Zhou, J. Zhao, and X. Guo, "Guided locality preserving feature matching for remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4435–4447, Aug. 2018.

[64] J. Ma, X. Jiang, J. Jiang, J. Zhao, and X. Guo, "LMR: Learning a two-class classifier for mismatch removal," *IEEE Trans. Image Process.*, vol. 28, no. 8, pp. 4045–4059, Aug. 2019.

[65] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[66] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008.

[67] J. Li, Q. Hu, and M. Ai, "RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform," *IEEE Trans. Image Process.*, vol. 29, pp. 3296–3310, 2020.

[68] C.-P. Lu, G. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 6, pp. 610–622, Jun. 2000.

[69] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *Int. J. Comput. Vis.*, vol. 81, no. 2, pp. 155–166, Feb. 2009.

[70] S. Li, C. Xu, and M. Xie, "A robust O(n) solution to the perspective-n-point problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 7, pp. 1444–1450, Jan. 2012.

[71] J. A. Hesch and S. I. Roumeliotis, "A direct least-squares (DLS) method for PnP," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 383–390.

[72] Y. Zheng, Y. Kuang, S. Sugimoto, K. Astrom, and M. Okutomi, "Revisiting the PnP problem: A fast, general and optimal solution," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 2344–2351.

[73] L. Ferraz, X. Binefa, and F. Moreno-Noguer, "Very fast solution to the PnP problem with algebraic outlier rejection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 501–508.

Authorized licensed use limited to: Wuhan University. Downloaded on January 31,2021 at 12:26:15 UTC from IEEE Xplore. Restrictions apply.