Robust Feature Matching for Remote Sensing Image Registration Based on L_q-Estimator

Jiayuan Li, Qingwu Hu, and Mingyao Ai

***Abstract*— This letter proposes a robust feature matching algorithm for remote sensing images based on l_q-estimator. We start with a set of initial matches provided by a feature matching method such as scale-invariant feature transform and then focus on global transformation estimation from contaminated observations and outliers elimination as well. We use an affine model to describe the global transformation and minimize a new cost function based on l_q-norm. We apply an augmented Lagrangian function and an alternating direction method of multipliers to solve such a nonconvex and nonsmooth optimization problem. Extensive experiments on real remote sensing data demonstrate that the proposed method is effective, efficient, and robust. Our method outperforms state-of-the-art methods and can easily handle situations with up to 90% outliers. In addition, the proposed method is much faster than RANSAC.**

***Index Terms*— Alternating direction method of multipliers (ADMMs), feature matching, image registration, l_q-estimator.**

I. Introduction

IMAGE registration, which refers to aligning two or more images with overlapping regions taken at different times, by different sensors or from different camera positions, is a critical step for many tasks of remote sensing, such as image mosaicking, image fusion, and panorama production. Current image registration methods usually share the same framework: first, establish some control points; then, estimate the geometrical transformation (projection, similarity, or affine); finally, align the image pair into a panoramic map.

Feature-based matching methods (to name a few, scale-invariant feature transform (SIFT) [1] and speeded up robust features (SURF) [2]) are very popular techniques to provide control points between an image pair. Many variants of SIFT and SURF have been proposed for remote sensing image registration, such as bilateral filter scale-invariant feature transform [3] and robust scale-invariant feature transform [4]. The correspondences generated by such a feature matching approach usually contain many outliers. To estimate the geometrical transformation, outliers should be eliminated. The outlier filtering methods can be roughly categorized into two types.

1) Parametric-based methods, including RANSAC [5] and its variants [6], [7]. These methods usually use a hypothesize-and-verify technique. RANSAC alternates between candidate transformation estimation and verification until finding a good solution. Fast sample consensus (FSC) [7] integrates a sampling technique with RANSAC to improve efficiency. These RANSAC-like methods, however, become very time-consuming when the inlier rate is low.
2) Nonparametric-based methods, such as vector field consensus (VFC) [8], identifying correspondence function (ICF) [9], coherent spatial matching (CSM) [10], coherence point drift (CPD) [11], and robust point matching via L_2E [12].

VFC uses a vector field to estimate a consensus of inlier correspondences that follow a nonparametric geometrical transformation. ICF adapts diagnostic technique and support vector machine to learn a correspondence function. CSM chooses the thin-plate spline to parameterize the coherent spatial mapping. CPD uses Gaussian mixture model to formulate point registration problem. Reference [12] introduces L_2E estimator to estimate the transformation between correspondence sets. Although these methods can achieve good results, the computational complexity is usually very high.

In this letter, we propose a robust feature matching method based on l_q-estimator. Unlike RANSAC and its variants, our method directly estimates the transformation from initial correspondences with outliers. We use an affine model to describe the global transformation between an image pair. We formulate a new cost function based on l_q-norm and minimize this nonconvex and nonsmooth problem via augmented Lagrangian function and alternating direction method of multiplier (ADMM) [13].

Locally linear transforming (LLT) [14] is close to the proposed algorithm, which also directly estimates a parametric model from observations with outliers. They formulate the robust feature matching problem as a maximum-likelihood estimation of a Bayesian model and use locally linear as a local geometrical constraint. LLT achieves impressed performance in many situations. As can be seen, the proposed method is different from LLT, including cost function and the solver. We show that the proposed method is more effective, more efficient, and more robust than LLT in Section III.

II. Robust Feature Matching

This section describes our robust feature matching method for satellite or aerial images based on l_q-estimator.

A. Problem Formulation

Suppose we are given a pair of remote sensing images of the same scene. We first perform a feature matching method such as SIFT to determine N initial matching correspondences $S = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{N}$, where $\boldsymbol{x}_n$ and $\boldsymbol{y}_n$ are the 2-D image coordinates of the matched feature points. Typically, $(\boldsymbol{x}_n, \boldsymbol{y}_n)$ satisfies the

Manuscript received July 28, 2016; revised September 1, 2016, September 21, 2016, and October 13, 2016; accepted October 17, 2016. Date of publication November 9, 2016; date of current version December 7, 2016. This work was supported by the National Natural Science Foundation of China under Grant 41271452.

The authors are with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China (e-mail: ljy_whu_2012@whu.edu.cn; huqw@whu.edu.cn; aimingyao@whu.edu.cn).

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

Digital Object Identifier 10.1109/LGRS.2016.2620147

1545-598X © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: Wuhan University. Downloaded on September 17,2020 at 13:33:49 UTC from IEEE Xplore. Restrictions apply.

following relationship if it is an inlier:

$$\boldsymbol{y}_n = T(\boldsymbol{x}_n). \tag{1}$$

The global transformation $T(\cdot)$ can be estimated by least squares method with the K inlier matches

$$\arg\min_{T} \sum_{n=1}^{K} (\boldsymbol{y}_n - T(\boldsymbol{x}_n))^2. \tag{2}$$

Thus, current methods usually use a two-step strategy or a hypothesize-and-verify technique such as RANSAC to solve this problem. Different from these methods, we directly estimate the global transformation from the putative matches with outliers. An impressed method is able to automatically classify the residual vector $\boldsymbol{v} = \{\boldsymbol{y}_n - T(\boldsymbol{x}_n)\}_{n=1}^{N}$ into an outlier set Outlier($v_n||v_n| \gg 0$) and an inlier set Inlier($v_n||v_n| \approx 0$), where v_n is the nth element of $\boldsymbol{v}$. The classical least-squares cost is sensitive to outliers. In contrast, l_0-norm is suitable for solving such a problem because it measures the number of nonzero elements in a residual vector. However, the observations usually contain noise, which makes l_0-norm unreliable. Researchers usually adapt the l_1-norm as the closest convex relaxation of l_0-norm to make a tradeoff. In this letter, we introduce a more robust estimator, i.e., l_q-estimator [15], for robust feature matching. The cost function of the proposed method is

$$\arg\min_{T} \sum_{n=1}^{N} \|\boldsymbol{y}_n - T(\boldsymbol{x}_n)\|_q^q \tag{3}$$

where $\|\cdot\|_q$ is a l_q-norm ($0 < q < 1$) operator.

As known, the ranges of elevations in remote sensing images are very small compared with the flight altitudes of the sensors, especially for satellite and aerial images. As with most of current methods, we use affine transformation as the approximation of $T(\cdot)$

$$T(\boldsymbol{x}_n) = \boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t} \tag{4}$$

where $\boldsymbol{A}$ is a 2×2 affine matrix and $\boldsymbol{t}$ is a 2×1 translation column vector. Consequently, the final cost function of the proposed method is as follows:

$$\arg\min_{\boldsymbol{A},\boldsymbol{t}} \sum_{n=1}^{N} \|\boldsymbol{y}_n - (\boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t})\|_q^q. \tag{5}$$

B. Transformation Estimation

Equation (5) is a nonconvex and nonsmooth function. We introduce auxiliary variables $\boldsymbol{P} = \{\boldsymbol{p}_n\}_{n=1}^{N}$ into (5)

$$\begin{aligned} &\arg\min_{\boldsymbol{A},\boldsymbol{t},\boldsymbol{P}} \sum_{n=1}^{N} \|\boldsymbol{p}_n\|_q^q \\ &\text{s.t. } \boldsymbol{\varepsilon}_n = (\boldsymbol{y}_n - (\boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t})) - \boldsymbol{p}_n = 0. \end{aligned} \tag{6}$$

Generally, this constrained optimization problem can be reformulated as an unconstrained one via augmented Lagrangian function

$$\begin{aligned} L_\rho(\boldsymbol{A},\boldsymbol{t},\boldsymbol{P},\boldsymbol{\Lambda}) &= \sum_{n=1}^{N} \left(\|\boldsymbol{p}_n\|_q^q + \boldsymbol{\lambda}_n^T \boldsymbol{\varepsilon}_n + \frac{\rho}{2}\|\boldsymbol{\varepsilon}_n\|_2^2 \right) \\ &= \sum_{n=1}^{N} \left(\|\boldsymbol{p}_n\|_q^q + \frac{\rho}{2} \left\| \frac{\boldsymbol{\lambda}_n}{\rho} + \boldsymbol{\varepsilon}_n \right\|_2^2 - \frac{1}{2\rho}\|\boldsymbol{\lambda}_n\|_2^2 \right) \end{aligned} \tag{7}$$

where $\boldsymbol{\Lambda} = \{\boldsymbol{\lambda}_n\}_{n=1}^{N}$ are Lagrange multiplier and $\rho > 0$ is a penalty parameter. There are two sets of variables, i.e., affine transformation $(\boldsymbol{A}, \boldsymbol{t})$ and auxiliary variables $\boldsymbol{P}$. We employ ADMM to decompose the function into three subproblems and estimate these two sets of variables alternately

$$\textbf{step 1} : \boldsymbol{P}^{k+1} := \arg\min_{\boldsymbol{P}} L_\rho(\boldsymbol{A}^k, \boldsymbol{t}^k, \boldsymbol{P}, \boldsymbol{\Lambda}^k) \tag{8}$$

$$\textbf{step 2} : (\boldsymbol{A}^{k+1}, \boldsymbol{t}^{k+1}) := \arg\min_{\boldsymbol{A},\boldsymbol{t}} L_\rho(\boldsymbol{A}, \boldsymbol{t}, \boldsymbol{P}^k, \boldsymbol{\Lambda}^k) \tag{9}$$

$$\textbf{step 3} : \left\{ \boldsymbol{\lambda}_n^{k+1} := \boldsymbol{\lambda}_n^k + \rho\boldsymbol{\varepsilon}_n \right\}_{n=1}^{N} \tag{10}$$

where superscript k is an iteration counter. In subproblem 1, $\boldsymbol{P}$ is the only variable and others are constants. In subproblem 2, the affine transformation $(\boldsymbol{A}, \boldsymbol{t})$ is the variable. ADMM alternates between these three subproblems until convergence.

Combining (8) and (9) with (7), respectively, subproblems 1 and 2 are detailed as follows:

$$\begin{aligned} &\arg\min_{\boldsymbol{P}} L_\rho(\boldsymbol{A}^k, \boldsymbol{t}^k, \boldsymbol{P}, \boldsymbol{\Lambda}^k) \\ &= \arg\min_{\boldsymbol{P}} \sum_{n=1}^{N} \left(\|\boldsymbol{p}_n\|_q^q + \frac{\rho}{2}\|\boldsymbol{\delta}_n - \boldsymbol{p}_n\|_2^2 \right) \end{aligned} \tag{11}$$

$$\begin{aligned} &\arg\min_{\boldsymbol{A},\boldsymbol{t}} L_\rho(\boldsymbol{A}, \boldsymbol{t}, \boldsymbol{P}^k, \boldsymbol{\Lambda}^k) \\ &= \arg\min_{\boldsymbol{A},\boldsymbol{t}} \sum_{n=1}^{N} \left(\frac{\rho}{2} \left\| \frac{\boldsymbol{\lambda}_n}{\rho} + (\boldsymbol{y}_n - (\boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t})) - \boldsymbol{p}_n \right\|_2^2 \right) \end{aligned} \tag{12}$$

where $\boldsymbol{\delta}_n = \boldsymbol{\lambda}_n/\rho + \boldsymbol{y}_n - (\boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t})$ in (11).

Equation (11) is a nontrivial l_q least squares (l_qLS) function and each variable $\boldsymbol{p}_n$ is independent. For the scalar version of (11) (p and δ are scalar values instead of vectors $\boldsymbol{P}$ and $\boldsymbol{\delta}$, respectively), we have

$$\arg\min_{p} \left(\|p\|_q^q + \frac{\rho}{2}\|\delta - p\|_2^2 \right) = \arg\min_{p} \left(|p|^q + \frac{\rho}{2}(\delta - p)^2 \right) \tag{13}$$

as proven in [15], the solution $\hat{p}$ is

$$\hat{p} = \begin{cases} 0, & \text{if } |\delta| < \tau_a \\ \{0, \operatorname{sgn}(\delta)\beta_a\}, & \text{if } |\delta| = \tau_a \\ \operatorname{sgn}(\delta)\beta_*, & \text{if } |\delta| > \tau_a \end{cases} \tag{14}$$

where

$$\beta_a = \left[\frac{2}{\rho}(1-q) \right]^{\frac{1}{2-q}} \quad \tau_a = \beta_a + \frac{q}{\rho}\beta_a^{q-1} \tag{15}$$

$\operatorname{sgn}(\cdot)$ is the signum function and $\beta_* \in (\beta_a, |\delta|)$ is the larger solution of (16) which has two solutions

$$\beta^{k+1} = f(\beta^k) \text{ where } f(\beta) = |\delta| - \frac{q}{\rho}\beta^{q-1} \tag{16}$$

where k is an iteration counter. With initial guess $\beta^0 = (\beta_a + |\delta|)/2$, this function can yield convergence within two iterations [15].

In subproblem 1, $\boldsymbol{p}_n$ is a 2-D column vector. By performing a minimization operator in (11) along the jth ($j = 1, 2.$) coordinate after completing the square, it becomes a scalar one. This is a coordinatewise minimizer for l_qLS.

Subproblem 2 is a classical linear least squares function and can be easily estimated by solving its normal equations.

Authorized licensed use limited to: Wuhan University. Downloaded on September 17,2020 at 13:33:49 UTC from IEEE Xplore. Restrictions apply.

Algorithm 1 Robust Feature Matching Based on Lq-Estimator

Input: a pair of remote sensing images (X, Y)
Output: correspondences $I = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{K}$, and transformation $T(\cdot)$

1 Perform SIFT on (X, Y) to extract initial matches $S = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{N}$
2 Data normalization
3 Sort S based on the matching scores, and pick the best M matches $Sub = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{M}$
4 Use Sub to estimate $T(\cdot)$ and inliers Sub_{inlier} according to section II A
5 Use linear least squares to refine $T(\cdot)$ based on Sub_{inlier}
6 Apply $T(\cdot)$ on $S = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{N}$ to find $I = \{(\boldsymbol{x}_n, \boldsymbol{y}_n)\}_{n=1}^{K}$

After determination of the global transformation $T(\cdot)$, outliers inside the initial matches can be removed based on the residual vector $\boldsymbol{v} = \{\boldsymbol{y}_n - (\boldsymbol{A}\boldsymbol{x}_n + \boldsymbol{t})\}_{n=1}^{N}$ and the image pair can be aligned together.

C. Algorithm Details

The proposed method has six degrees of freedom, i.e., four in 2×2 affine matrix and two in translation vector, which can be theoretically solved by three noncollinear inlier matches. In practice, more inliers are used for estimation because redundant observations may improve the accuracy and robustness. Feature matching methods such as SIFT usually can provide hundreds of initial matches. However, too much matches will decrease the efficiency while do not improve the accuracy. Therefore, we sample $M (M < N)$ ($M = 100$ in our experiments) initial matches to compute the affine transformation and then apply the transformation to all the initial matches. Random sample is a simple and popular strategy. It, however, treats all initial matches equally. In our implementation, we make use of the matching scores of initial correspondences. We sort the initial matches based on matching scores, and the best M matches are used for estimation. Algorithm 1 summarizes the details of the proposed method. We set $q = 0.2$ fixed in our experiments; initial penalty parameter $\rho = 3 \times 10^{-4}$ changes by a multiplying factor $\alpha = 1.65$ along with iterations.

III. Experimental Results

We compare our method with five state-of-the-art algorithms, i.e., LLT [14], VFC [8], CSM [10], FSC [7], and RANSAC [5]. The parameters of each method are set according to the authors' suggestion and all fixed throughout the experiments. The implementations except FSC are obtained from the authors' Web sites. FSC is implemented by the authors. All experiments are performed on a laptop PC with an Intel Core i5-3210M, 2.5 GHz CPU, 8 GB of RAM, and MATLAB code.[1]

A. Data Sets and Settings

Two data sets are used to evaluate the proposed method. The first test set is obtained from Erdas example data, http://download.intergraph.com/downloads/[2] which was taken over Illinois, USA. It contains 11 image pairs formed by color-infrared aerial images with extremely small overlapping regions. The image sizes of these images are from 1391×1374 pixels to 1459×1380 pixels (Data set 1). Another one is generated by a preregistered Worldview-2 RGB-infrared image[3] pair with different rotations, which was taken over Guangzhou, China. The spatial resolution of this data set is 0.5 m and image size is 1000×1000 pixels. In detail, we rotate the infrared image by angles 15°, 30°, 45°, 60°, 75°, and 90° to form six image pairs with the original RGB image (Data set 2). The exact ground truth transformation of Data set 2 can be computed easily. To establish the approximate ground truth transformation of Data set 1, we find five distinct corner correspondences artificially for each image pair. These correspondences are selected at subpixel accuracy (absolutely less than 0.5 pixels) by enlarging their local regions. We then estimate the affine transformation by linear least squares method. We use SIFT algorithm implemented by open-source library VLFEAT toolbox [16] to provide initial matches for all experiments. The matching ratio of SIFT is set to be 1.2. The ground truth transformation is then applied to the initial matches. Matches whose residual errors less than a certain threshold τ are chosen as ground truth inliers. For Data set 1, $\tau = 3$ pixels. $\tau = 2$ pixels for Data set 1 because the ground truth transformation is very exact. We use precision and recall to evaluate the experimental results on Data set 1, where the precision is the ratio of the detected inlier number and the detected match number, and the recall is the ratio of the detected inlier number and the ground truth inlier number. For Data set 2, we use two more metrics, i.e., root-mean-square error and max error (ME), to evaluate the accuracy of the extracted matches.

TABLE I
Quantitative Evaluation on **Data Set 2**

Method	Ours	LLT	VFC	CSM	FSC	RANSAC	GT
Precision(%)	**94.33**	88.2	88.04	87.22	89.18	89.23	100
Recall (%)	**100**	**100**	**100**	**100**	**100**	97.22	100
RMSE(pixels)	**1.12**	1.55	1.59	2.08	1.41	1.42	1.0
ME (pixels)	**2.84**	7.64	7.8	15.48	4.89	5.48	1.97
Time (s)	**0.04**	0.28	1.66	24.08	0.34	5.28	-

Note that GT is the short for ground truth.

B. Results on Data Set 1

1) Qualitative Comparison: Figs. 1 and 2 show the results of LLT, VFC, CSM, FSC, RANSAC, and our proposed method on the first and tenth image pairs. The first and tenth image pairs have small overlap areas along horizontal direction and vertical direction, respectively. The matching problem is very difficult because of such small overlapping regions. As can be seen, LLT and VFC are not reliable because they both fail to detect any correct matches in one of the two image pairs. In addition, there are still many outliers preserved in their successful results. The detected matches of CSM are all outliers. RANSAC and its variant FSC are much better than above-mentioned methods. Although most detected matches are correct inliers, some outliers are still preserved. The proposed method achieves the best performance on these two image pairs. There are no outliers in the results, and our method detects many more correct inliers than RANSAC and FSC.

2) Quantitative Evaluation: The quantitative comparisons on Data set 1 are shown in Fig. 3. The left figure reports the

[1]The code of our method is available: https://github.com/sunrainljy/lqRFM.
[2]erdas-imagine-2013-2014-example-data.
[3]The data set is obtained from: https://www.digitalglobe.com.

Authorized licensed use limited to: Wuhan University. Downloaded on September 17,2020 at 13:33:49 UTC from IEEE Xplore. Restrictions apply.

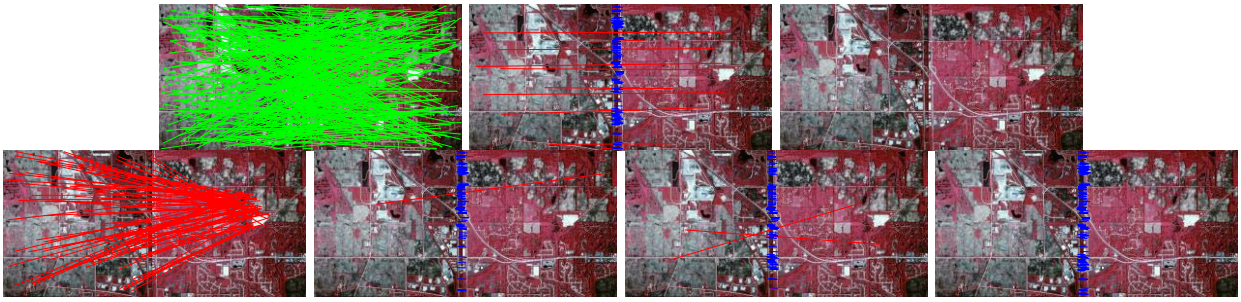

Fig. 1. Qualitative results of the first image pair in Data set 1. First row, from left to right: initial matches, LLT result, and VFC result. Second row, from left to right—CSM result, FSC result, RANSAC result, and our result. The green lines represent the initial matches, the blue lines represent correct detected inliers, and the red lines represent false detected inliers.

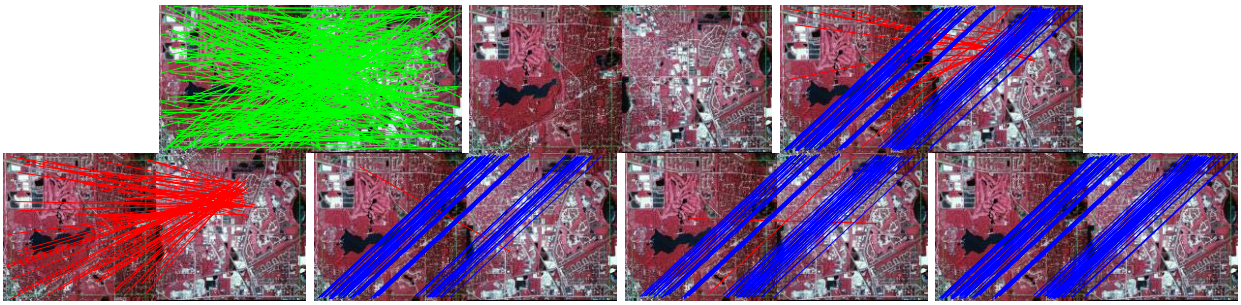

Fig. 2. Qualitative results of the tenth image pair in Data set 1. First row, from left to right: initial matches, LLT result, and VFC result. Second row, from left to right—CSM result, FSC result, RANSAC result, and our result. The green lines represent the initial matches, the blue lines represent correct detected inliers, and the red lines represent false detected inliers.

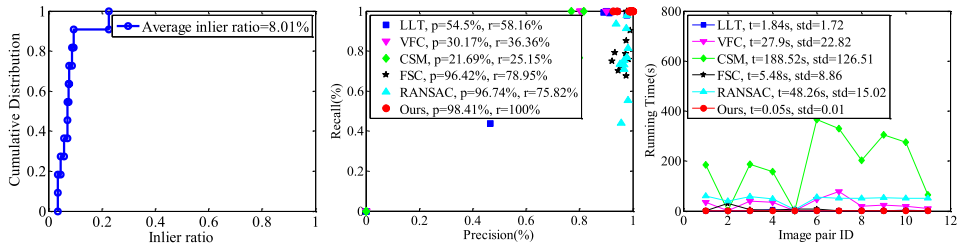

Fig. 3. Comparisons of our method with LLT, VFC, CSM, FSC, and RANSAC on Data set 1. The left plot shows the cumulative distribution of initial matching inlier ratio in Data set 1, the middle plot shows the precision-recall statistics of the six methods, and the right plot shows is the running time statistics of the six methods.

cumulative distribution of the initial matching inlier ratio. The average inlier ratio of this data set is 8.01%, which means that the average outlier ratio is up to 91.99%. This data set is quite challenging because of such low inlier ratios (lowest is 3.46%). The accuracy performance is summarized in the middle figure, where each dot stands for a precision-recall pair. From the figure, only FSC, RANSAC, and our method achieve high precision accuracy on all image pairs, and only our proposed method achieves high recall accuracy on Data set 1. The average precision-recall pairs are (54.5%, 58.16%), (30.17%, 36.36%), (21.69%, 25.15%), (96.42%, 78.95%), (96.74%, 75.82%), and (98.41%, 100%) for LLT, VFC, CSM, FSC, RANSAC, and ours. As shown, the proposed method achieves (1.67%, 24.18%) growth rates compared with the second best algorithm (ranked by precision), i.e., RANSAC.

3) Running Time Comparison: The right figure of Fig. 3 plots the running time statistics of the six methods. As shown, our method, FSC, and LLT are very fast; VFC and RANSAC need dozens of seconds; CSM is the slowest. The average (running time, standard deviation) pairs are (1.84, 1.72), (27.9, 22.82), (188.52, 126.51), (5.48, 8.86), (48.26, 15.02), and (0.05, 0.01) for LLT, VFC, CSM, FSC, RANSAC, and our method. Our method is 30+ times faster than LLT, 100+ times faster than FSC, and 950+ times faster than RANSAC on this challenging data set. In addition, the running time of the proposed method is quite stable that the standard deviation is 0.01s.

C. Results on Data Set 2

We also compare the RMSE and the ME of the detected matches on Data set 2. Fig. 4 displays the comparisons of precision, recall, RMSE, and ME in the leftmost, left middle, right middle, and rightmost plots, respectively. All the six methods perform well on this data set because the average inlier ratio (26.11%) of this data set is much higher than Data set 1. On all the six image pairs, our method achieves better performance than other methods in all the four metrics. Table I summarizes the average performance of the six image pairs. As can be seen, the proposed method achieves the best performance in all the five metrics compared with LLT, VFC, CSM, FSC, and RANSAC.

Authorized licensed use limited to: Wuhan University. Downloaded on September 17,2020 at 13:33:49 UTC from IEEE Xplore. Restrictions apply.

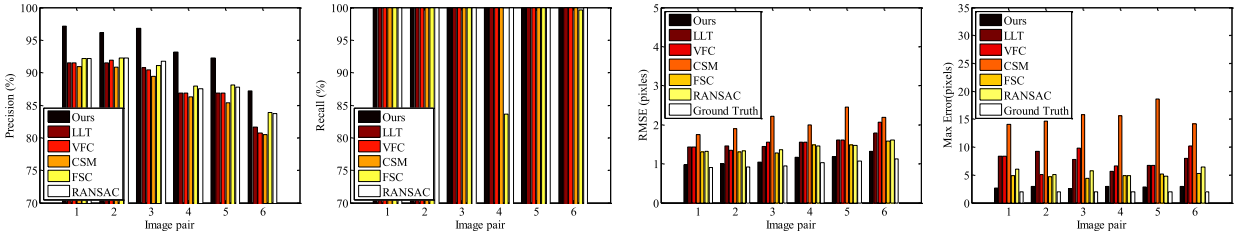

Fig. 4. Comparisons of our method with LLT, VFC, CSM, FSC, and RANSAC on Data set 2. From left to right: precision, recall, RMSE, and ME comparisons on each image pair in Data set 2.

TABLE II
STUDY OF PARAMETER Q ON SIMULATION AND REAL EXPERIMENTS

Experiments	q	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Simulation	Success rate (%)	98.4	95.9	92.6	90.7	85.0	71.7	54.0	19.5	1.4
Dataset1	Precision (%)	90.8	98.4	89.9	72.3	63.1	71.3	44.7	27.0	26.9
Dasaset2	Precision (%)	94.3	94.3	94.3	94.3	94.3	94.3	94.3	94.3	94.3

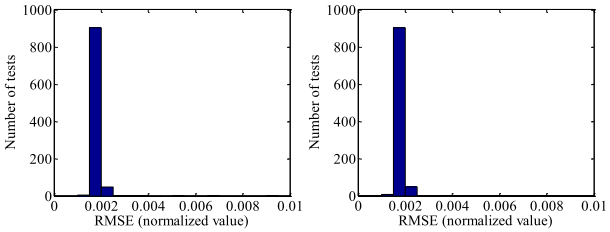

Fig. 5. Statistics of RMSE in two groups of tests. The success rates of these two groups are 95.8% and 95.9%, respectively.

D. Convergence Analysis and Parameter Study

We perform a simulation experiment to analyze the convergence and parameters of our method. We randomly generate 100 reference feature points normalized to [0, 1] and an affine transformation. The correspondences (target feature points) of these points are then computed by the transformation and the reference points. We add Gaussian noise with zero-mean and deviation $\sigma = 0.002$ to the target points. Then, random errors between $[-0.5, 0.5]$ are added to 50% of the noisy target points. We use the reference points and the contaminative target points as the input of our method and estimate the transformation. The proposed method is success if the RMSE is smaller than 0.003. This process is performed by 1000 times at one group. We do four groups of tests and the success rate is about 95.8% (see Fig. 5). We then change q from 0.1 to 0.9. For each setting q, 1000 independent tests are performed and the success rate is computed. We also calculate the precision of our method for each setting q on Data set 1 and Data set 2. The results are shown in Table II. We find that small values of q achieve better success rates or precision than large values of q on simulation experiment and Data set 1. For relative less challenging data set (Data set 2), parameter q almost has no influence to the results.

IV. CONCLUSION

In this letter, we propose a robust feature matching method for remote sensing image registration based on l_q-estimator. We formulate a novel cost function and minimize this nonconvex and nonsmooth problem via augmented Lagrangian function and ADMM. Our method directly estimates the transformation from initial correspondences with outliers. The evaluations on two remote sensing data sets demonstrate that the proposed method outperforms the state-of-the-art methods, i.e., LLT, VFC, CSM, FSC, and RANSAC, significantly.

REFERENCES

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
[2] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.
[3] S. Wang, H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 649–653, Jul. 2012.
[4] Q. Li, G. Wang, J. Liu, and S. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 287–291, Apr. 2009.
[5] M. A. Fischler and R. 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.
[6] O. Chum and J. Matas, "Matching with PROSAC—Progressive sample consensus," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 220–226.
[7] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 43–47, Jan. 2015.
[8] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
[9] X. Li and Z. Hu, "Rejecting mismatches by correspondence function," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, 2010.
[10] J. Ma, J. Zhao, Y. Zhou, and J. Tian, "Mismatch removal via coherent spatial mapping," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep./Oct. 2012, pp. 1–4.
[11] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
[12] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu, "Robust L_2E estimation of transformation for non-rigid registration," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1115–1129, Mar. 2015.
[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[14] 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, Dec. 2015.
[15] 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.
[16] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," in *Proc. 18th ACM Int. Conf. Multimedia*, 2010, pp. 1469–1472.

Authorized licensed use limited to: Wuhan University. Downloaded on September 17,2020 at 13:33:49 UTC from IEEE Xplore. Restrictions apply.