Pattern Recognition 111 (2021) 107693

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Robust line segment matching via reweighted random walks on the homography graph

Dong Wei [a], Yongjun Zhang [a,*], Chang Li [b]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[b] Key Laboratory for Geographical Process Analysis & Simulation, Hubei Province, and School of Urban and Environmental Science, Central China Normal University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 31 July 2019
Revised 29 September 2020
Accepted 7 October 2020
Available online 8 October 2020

Keywords:
Line segment matching
Epipolar geometry
Reweighted random walks
Graph matching

ABSTRACT

This paper presents a novel method for matching line segments between stereo images. Given the fundamental matrix, the local homography can be over determined with pairwise line segment candidates. We exploit this constraint to initialize the candidate and construct the novel homography graph. Because the constraint between the node is based on the epipolar geometry, the homography graph is invariant to the local projective transformation. We employ the reweighted random walk on the graph to rank the candidate, then, we propose the constrained-greedy algorithm to obtain the reliable match. To the best of our knowledge, this is the first study to embed the epipolar geometry into the graph matching theory for the line segment matching. When evaluated on the 32 image patches, our method outperformed the state of the art methods, especially in the scenes of the wide baseline, steep viewpoint changes and dense line segments. The proposed algorithm is available at https://github.com/weidong-whu/line-match-RRW.

1. Introduction

Matching line segments that have the same pre-image in 3D space is useful for 3D reconstruction, because line segments incorporate more semantic and geometric meanings than points. Generally, there are many line segments in man-made scenes, which enable reconstructing the abstractive 3D structure efficiently via matching the line segment [1]. In addition, matching the line segment is helpful to obtain more accurate and complete 3D reconstructions [2,3]. Ideally, one would like to have a line segment matching algorithm that is able to obtain more matches with a higher accuracy. However, compared with the point matching, the line segment matching remains to be a challenging work due to its inherent difficulties. First, there is no strong epipolar-geometry constraint for line segments in stereo images. Second, the texture along the line segment is poor in general, thus it is difficult to obtain the reliable texture-descriptor. Third, the endpoints of the line segment are indefinite, thus it is difficult to employ the texture and geometry constraint. Many algorithms have been published for the line segment matching and they are of three types: those that match in individual; those that match in group; and those that match with the point correspondence.

Those that match in individual are generally based on the consistency of the geometry or texture. Schmid and Zisserman [4] proposed the classical algorithm to match the individual line segment. In their method, the mapping between the line segment is first calculated via the epipolar geometry through trial, then, the cross-correlation score is calculated to find the correct match. Bay et al. [5] constructed the color histogram along the line segment to obtain initial candidates, then they used the topological filter to remove the wrong candidate. Wang et al. [6] proposed the texture descriptor that is invariant to the rotation and illumination. To make the texture descriptor to be scale invariant, Zhang et al. [7] constructed the descriptor in different scales. Some studies [1,8] used the image sequences to match the line segment, but they required image sequences that are overlapped, e.g., Line 3D ++ [1] requires at least 4 overlapping images.

Those that match in group exploit the geometry and texture information between line segments in the same view. The early studies [9,10] constructed the graph to match line segments, but they are only appropriate for the planar scene. Matching with the pair of line segments can exploit the intersection, cross angle, epipolar line, and local region similarity, etc., which has been a common strategy. OK et al. [11] exploited the Daisy descriptor [12] to initialize the candidate and used six pair-wise constraints to find the correct match. This method is appropriate for the aerial image when the viewpoint change and distortion are small. Al-Shahri et al. [13] used the local homography to obtain the match, but it

* Corresponding author.
E-mail addresses: weidong@whu.edu.cn (D. Wei), zhangyj@whu.edu.cn (Y. Zhang), lichang@mail.ccnu.edu.cn (C. Li).

https://doi.org/10.1016/j.patcog.2020.107693
0031-3203/

requires enough point correspondences to calculate the local homography. López et al. [14] used the context and appearance of adjacent line segments for matching in low-textured Images. Kim et al. [15] exploited the texture in the intersection region of line segments and used the cross-correlation to measure the similarity. Li et al. [16,17] used a SIFT-like method to match the line-to-line junction.

Those that match based on point correspondences use the local affine or projective mapping that is calculated from point correspondences. Fan et al. [18] encoded the local geometric information between the line and its surrounding SIFT points. This method performs well when enough point correspondences exist. Sun et al. [19] calculated the planar homography via the 3D points generated from the 2D point correspondences, which requires both enough point correspondences and calibrated stereo images. Jia et al. [20] exploited the topological adjacency of a point-line to find and filter the line candidate. Recently, Wang et al. [21] matched the line segment in aerial oblique images via exploring the ratio of point-line distance in the affine projection space.

Compared with the point matching, the epipolar geometry is less investigated in the line segment matching. For most of the previous algorithms, the epipolar constraint is used to reduce the candidate before the texture correlation. Thus, they have to deal with the challenges in the correlation: (1) The texture is distorted because of the steep viewpoint change and wide baseline. (2) The texture in dense line segments is quite similar. Matching with point correspondences [18,19] is independent of the texture, but their performances require sufficient point correspondences. More recently, the deep learning technique has been employed to learn the angles between planes using convolutional neural networks (CNNs), and the angle regularization is then used to correct unreliable line matches [22]. But this method requires knowledge of the intrinsic of images.

This paper presents our algorithm that matches line segments based on the epipolar geometry. In our algorithm, candidates are initialized via the homography constraint which is induced from the epipolar geometry. Because the homography constraint is not robust, we build the homography graph, where the node represents the candidate and the edge represents the homography constraint. Then, we employ the reweighted random walk [23] on the homography graph to rank the candidates, and propose the constrained greedy algorithm to obtain the reliable matches. This paper makes two main contributions for the line segment matching.

(1) We propose the homography graph, which is invariant to the local projective transformation and is independent of the texture. The experiments showed that it is able to obtain more correct matches with a higher accuracy when compared with the state of the art methods, especially in the scenes of the wide baseline, steep viewpoint change and dense line segments.

(2) To the best of our knowledge, this is the first study to embed the epipolar geometry to the graph matching method for feature matching, which is able to be a general frame.

The remainder of the paper is organized as follows: Section 2 explains the overview of our algorithm. Section 3 addresses the homography constraint for candidate pairs. Section 4 to Section 6 introduce the details of our algorithm. Section 7 presents our performance evaluation of the proposed algorithm by comparing it with the state of the art matching methods. Section 8 concludes the paper.

2. Overview

Fig. 1 shows the workflow of the proposed method. First, line segments in the two views are extracted and their orientations are recovered [5]. Second, the homography constraint, which is a constraint for candidate pair, is exploited to initialize the candidate and construct the homography graph. Third, the reweighted random walk is employed on the homography graph, which attaches the confidence to the candidate. Finally, because the greedy algorithm [24] cannot be directly used in the matching assignment, we propose the constrained greedy algorithm to find the reliable match.

There are two underlying principles in our algorithm.

(1) Given the fundamental matrix, the homography can be over determined with one constraint via two line segment correspondences. This homography constraint enables initializing the candidate.

(2) If one candidate satisfies the homography constraint with many other candidates, it has a high confidence to be the correct match. This principle is quite similar to the PageRank [25], a website ranking algorithm: the website should be attached more importance if it is linked by many other websites or the important website. Consider each candidate as the node and the homography constraint as the edge between the node, the initial candidates can be ranked via the random walk on the association graph.

3. Homography constraint

Define the scene plane by $\boldsymbol{\pi}^{\mathrm{T}}\mathbf{X} = 0$ with $\boldsymbol{\pi} = (\mathbf{v}^{\mathrm{T}}, 1)^{\mathrm{T}}$; $\mathbf{v}$ is a 3D vector that parameterizes the 3D plane. The point $\mathbf{X}$ on the scene plane can be projected to the first and second view as $\mathbf{x}$ and $\mathbf{x}'$, respectively; the 2D point is denoted as $(x, y, 1)^{\mathrm{T}}$. Based on the scene plane theory [4], the homography induced by the plane is $\mathbf{x}' = \mathbf{H}\mathbf{x}$ with

$$\mathbf{H} = \mathbf{A} - \mathbf{e}'\mathbf{v}^{\mathbf{T}} \tag{1}$$

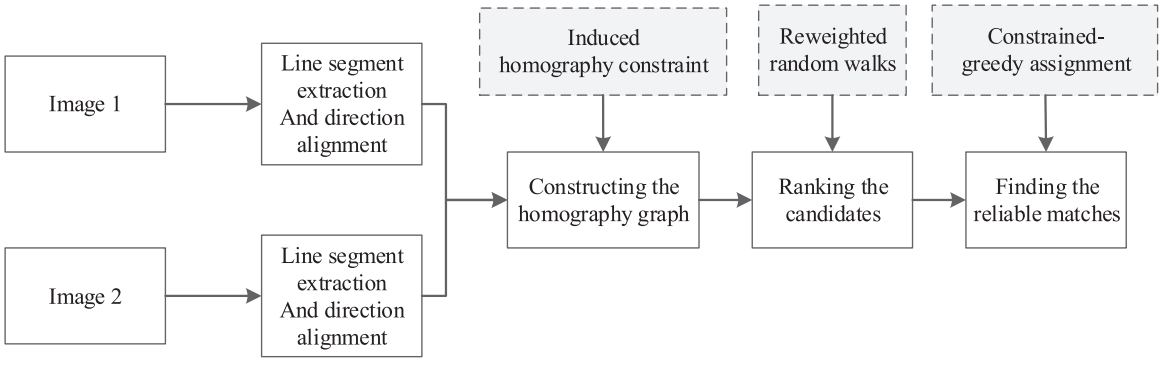


Fig. 1. The workflow of the proposed algorithm.

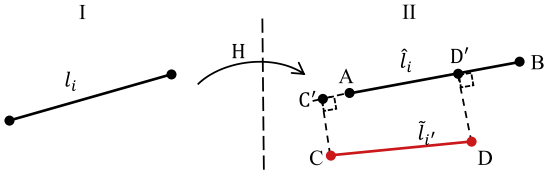


Fig. 2. The illustration of the mapping error and the overlapping rate under the homography.

where $[\mathbf{e}']_\times\mathbf{A}$ is the decomposition of the fundamental matrix $\mathbf{F}$. For the line $ax+by+c=0$, it is denoted as $(a, b, c)^{\mathrm{T}}$. Suppose $\mathbf{x}'$ is on the line $\mathbf{l}'$, then from $\mathbf{l}'^{\mathbf{T}}\mathbf{x}'=0$

$$\mathbf{l}'^{\mathbf{T}}\left(\mathbf{A}-\mathbf{e}'\mathbf{v}^{\mathbf{T}}\right)\mathbf{x}=0 \tag{2}$$

After a short calculation, Eq. (2) can be written as,

$$\mathbf{x}^{\mathrm{T}}\mathbf{v}=\mathbf{x}^{\mathrm{T}}\left[\mathbf{e}'\right]_\times\mathbf{F}\mathbf{l}'/\left(\mathbf{e}'^{\mathrm{T}}\mathbf{l}'\right) \tag{3}$$

Given two candidates denoted as $C_m=(l_i,\tilde{l}_{i'})$ and $C_n=(l_j,\tilde{l}_{j'})$, we can list four equations. Thus $\mathbf{v}$ can be estimated by least squares with one constraint, and $\mathbf{H}$ can be calculated with Eq. (1).

We use $\mathbf{H}$, which is estimated by the candidate pair (C_m and C_n), to validate the candidate pair. Take Fig. 2 for example, l_i can be mapped with $\mathbf{H}$ to the second view as $\hat{l}_i$. Denote the shorter of $\hat{l}_i$ and $\tilde{l}_{i'}$ as l_{CD} and the longer as l_{AB}. If C_m is aligned with $\mathbf{H}$, it should satisfy the homography constraint

$$\begin{cases}\left|\mathrm{CC}'\right| < T_{map}\\ \left|\mathrm{DD}'\right| < T_{map}\\ \left|\mathrm{AD}'\right|/|\mathrm{CD}| > T_{ove}\end{cases} \tag{4}$$

where T_{map} is the threshold of the mapping error and it is set as 2 pixels, and T_{ove} is the threshold of the overlapping rate that is set as 0.5. All the thresholds in our algorithm will be analyzed in the experiments. If both C_m and C_n satisfy the constraint, we say that the candidate pair satisfy the homography constraint.

The homography constraint is effective to confirm the correct match. If C_m and C_n are correct and coplanar, the reconstructed 3D line segments are also coplanar, then C_m and C_n will satisfy the homography constraint (Fig. 3. (a)). Either C_m or C_n is incorrect may fail to induce a scene plane, thus they will not satisfy the constraint (Fig. 3(b)). However, the constraint may bring about false positives. A typical example is shown in Fig. 3(c). Since there are many planes in man-made scenes, we construct the homography graph in the next section to obtain the reliable match.

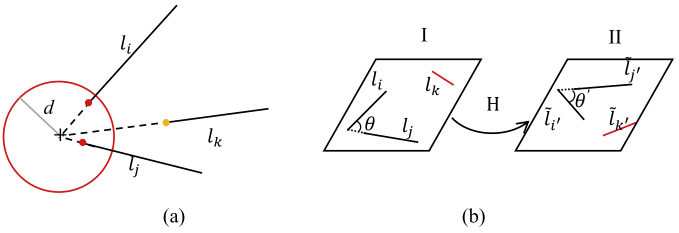


Fig. 4. (a) The intersection of the two line segments must be within T_{int} pixels to its nearest endpoint. Thus, l_k cannot form a pair with others. (b) the individual candidate $(l_k,\tilde{l}_{k'})$ is added based on the homography induced by P_{mn}.

4. Homography graph

4.1. Initial candidates

Given a candidate pair $P_{mn}=(C_m,C_n)$, if P_{mn} satisfy the homography constraint, both C_m and C_n are selected as the initial candidate. A violent method for initialization is to validate all the candidate pairs. However, it is not appropriate in our algorithm for two reasons. First, it is time-consuming because the brute force algorithm has a complexity of $O(n^4)$, where n is the maximum number of the line segments in the two views. Second, the homography constraint is only effective in the plane, thus validating all candidate pairs may bring about many false positives.

The initialization is optimized via setting the geometric constraint on the line segment pair. There are two assumptions for coplanar line segments in man-made scenes. First, their distance should be close. Second, their angle should be large. Thus, only if the intersection of the line segment pair is within T_{int} pixels to the nearest endpoint (Fig. 4(a)), the pair is valid. This kind of line segment pair can be viewed as a subset of the V-junction in [16], and we set T_{int} as 15 (it is 20 in [16]).

Two geometric constraints are employed in the validation to reduce the computation. First, the epipolar constraint is employed as line 3D++ algorithm [1]. Second, the line segment pair in each view forms an angle (Fig. 4. (b)); the angle difference should be smaller than T_{ang} (T_{ang} is set as $\pi/4$). Since some line segments cannot form a pair with others, the individual candidate that is consistent with any induced homography is selected as the initial candidate (Fig. 4). The consistency is validated based on Eq. (4).

4.2. Graph construction

In the homography graph, the node represents the candidate and the edge represents the homography constraint. Denote n^1 as the number of nodes, we use a matrix $\mathbf{W}\in\{0,1\}^{n^1\times n^1}$ to encode

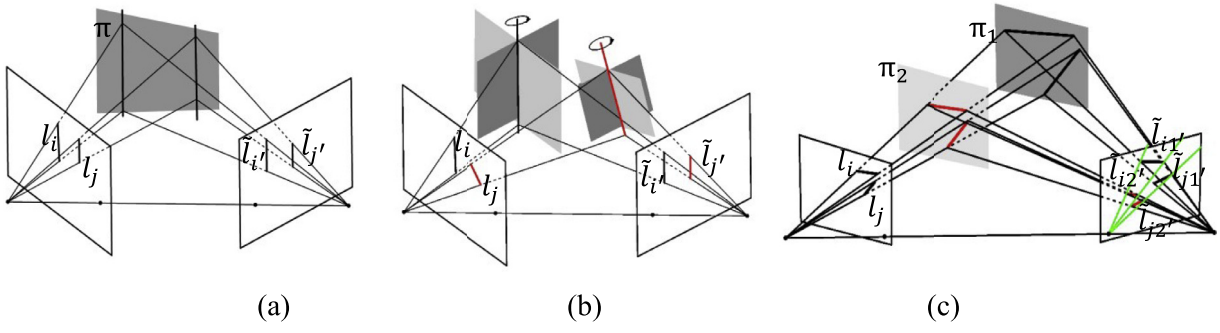


Fig. 3. (a) The candidate pair is correct and coplanar, which induces a world plane and satisfies the homography constraint. (b) the candidate pair is incorrect; thus they fail to induce a plane. (c) $\tilde{l}_{i1'}$ and $\tilde{l}_{j1'}$ are parallel to $\tilde{l}_{i2'}$ and $\tilde{l}_{j2'}$, respectively, moreover, they are all within the epipolar line interval. Thus, if $(l_i,\tilde{l}_{i1'})$ and $(l_j,\tilde{l}_{j1'})$ are correct, $(l_i,\tilde{l}_{i2'})$ and $(l_j,\tilde{l}_{j2'})$ will also satisfy the homography constraint.

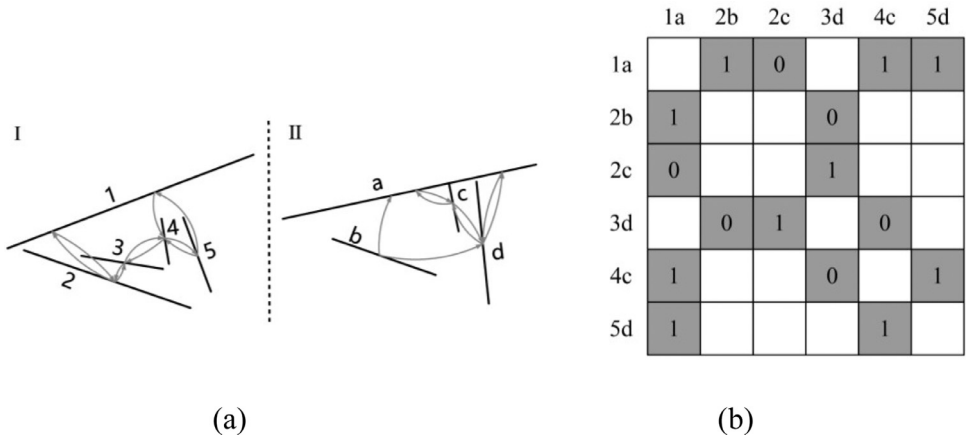


Fig. 5. The workflow of constructing W, the initial candidates are (1,a), (2,b), (2,c), (3,d), (4,c), (4,d), (5,c), and (5,d). (a) Edges are connected between line segments in each view. (b) Only the candidate pair in gray cells are validated.

the homography graph, whose element is computed as follows:

$$\mathbf{W}_{i,i';j,j'} = \begin{cases} 0, & if\ i = j\ and\ i' = j',\ or\ C_m\ and\ C_n\ cannot \\ & satisfy\ the\ homography\ constraint \\ 1, & else \end{cases} \tag{5}$$

There are two steps to construct **W**. First, in each view, edges are generated for each line segment by connecting to its T_{nei} nearest neighbors (Fig. 5(a)). The edge between l_i and l_j is denoted as $e_{i,j}$. Second, validate $(l_i, \tilde{l}_{i'})$ and $(l_j, \tilde{l}_{j'})$ if both $e_{i,j}$ and $e_{i',j'}$ are exist (Fig. 5(b)).

In graph matching [26], if the matching is with deformations, T_{nei} is set as a small value to capture the local property, otherwise, T_{nei} is set as a big value to capture the global property. Since the homography constraint is only effective in the plane, our algorithm is a type of the former. We follow [26] and [27] to set T_{nei} as 20, which we think is enough to construct a robust graph.

5. Ranking the candidate

We rank the node via the reweighted random walk (RRW) [23] on the homography graph. RRW is a kind of the random walk (RW) but different from the general RW in two aspects:

(1) To reduce the adverse effect of the incorrect nodes, instead of normalizing **W** via $\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}$, where **D** is a diagonal matrix with entries $\mathbf{D}_{ii} = d_i = \sum_j \mathbf{W}_{i,j}$, it normalizes **W** by adding the absorbing node $\boldsymbol{x}_{abs}^{(n)}$, and formulates the transition matrix **P** and the Markov chain as

$$\mathrm{P} = \begin{pmatrix} \mathbf{W}/d_{\max} & 1 - d/d_{\max} \\ 0^T & 1 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{x}^{(n+1)T} & x_{abs}^{(n+1)T} \end{pmatrix} = \alpha \begin{pmatrix} \mathbf{x}^{(n)T} & x_{abs}^{(n)T} \end{pmatrix} \mathbf{P} \tag{6}$$

(2) To strengthen the effect of the correct nodes, the personalized jump is employed via the inflation and bistochastic normalization [28]. Finally, the reweighted random walk is formulated by,

$$\begin{pmatrix} \mathbf{x}^{(n+1)T} & x_{abs}^{(n+1)T} \end{pmatrix} = \alpha \begin{pmatrix} \mathbf{x}^{(n)T} & x_{abs}^{(n)T} \end{pmatrix} \mathbf{P} + (1-\alpha) \left(f_C \left(\mathbf{x}^{(n)T} \mathbf{W} \right)^T \right) \tag{7}$$

where α is the parameter that adjusts the effect of the last walk, and f_C is the reweighting function.

When the random walk reaches the stationary state, i.e., $\mathbf{x}^{(n+1)} = \mathbf{x}^n$, the probability of the candidate is calculated by the conditional distribution,

$$\bar{\mathbf{x}}_i^{(n)} = P\left(\boldsymbol{X}^{(n)} = v_i | \boldsymbol{X}^{(n)} \neq v_{abs}\right) = \mathbf{x}_i^{(n)} / \left(1 - x_{abs}^{(n)}\right) \tag{8}$$

In fact, $\bar{\mathbf{x}}$ is the optimal solution of the mapping constraints,

$$\bar{\mathbf{x}} = argmax\left(\mathbf{x}^{\mathbf{T}} \mathbf{W} \mathbf{x}\right), \quad \text{s.t.} \quad \sum_{i=1}^{n} \mathbf{x}_i = 1 \tag{9}$$

We choose RRW to rank the candidates in the homography graph because it is robust when there exist many wrong nodes.

6. Constrained-greedy assignment

Various algorithms have been proposed to assign matches with $\bar{\mathbf{x}}$ that is optimized via Eq. (9), e.g. the greedy algorithm [24] and Hungarian algorithm [29]. However, these methods cannot be used directly in our algorithm. First, these methods are under the one-to-one matching constraint, while a line segment in the first view may have more than one match in the second view and vice versa. Second, the incorrect candidate that is not contradictory with the correct match cannot be rejected via these methods, whereas the homography graph has many wrong nodes.

Fig. 6. Natural stereo images with various transformations. (1) Rotation, (2) Occlusion, (3) Scale, (4) Light, (5) Viewpoint, (6) Viewpoint.

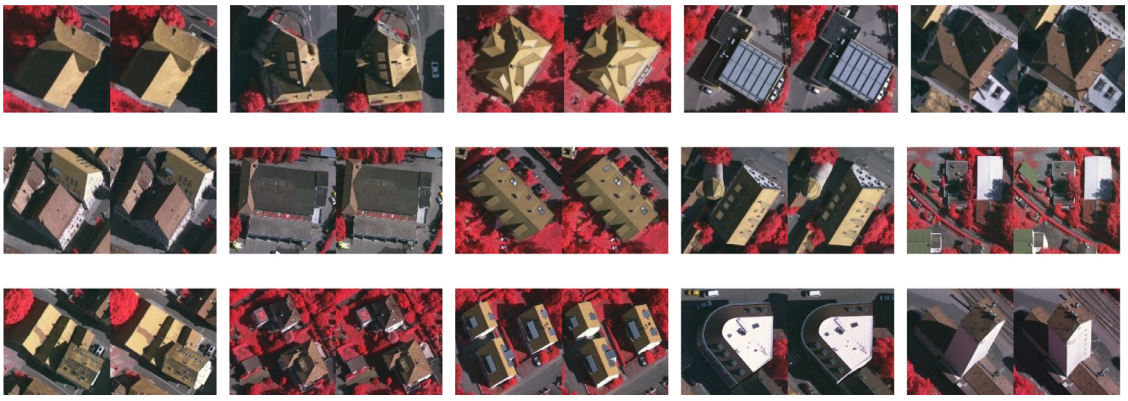


Fig. 7. The aerial image dataset provided in [11].

We use the greedy algorithm with the homography constraint to confirm the correct match. For convenience, $\bar{\mathbf{x}}$ is transformed into the form of $\mathbf{X} \in [0,1]^{m^1 \times m^2}$ where m^1 and m^2 are the number of line segments in the two views. The details are as follows:

(1) Finding the maximum $\mathbf{X}_{i,i'}$ in $\mathbf{X}$, if no cell $c \in \{\mathbf{X}_{i,*} \cup \mathbf{X}_{*,i'}\}$ except for $\mathbf{X}_{i,i'}$ that satisfies $c > 0.5 * \mathbf{X}_{i,i'}$, $C_m = (l_i, \tilde{l}_{i'})$ is selected as the correct match and $\mathbf{X}_{i,*}$ and $\mathbf{X}_{*,i'}$ are then set as 0. C_m is robust because its contradictory match has a much smaller confidence.

(2) Step (1) is repeated until $\mathbf{X} = 0^{m^1 \times m^2}$. The first T_{fir}matches of $\boldsymbol{C} = \{C_1, C_2, \ldots, C_n\}$ are selected as the reliable matches (T_{fir} is set as 0.7).

(3) Greedy mapping with the homography constraint. The greedy algorithm [24] is employed for the remaining candidates, during which the candidate with the maximum confidence should satisfy at least one homography, which is induced by its T_{nei} nearest reliable matches. T_{nei} has been introduced in Section 4.2.

(4) Finding more individual matches. The individual line segment that has not been matched is validated by its T_{nei} nearest homographies. Only those that are aligned with at least T_{hom} homographies (it is set as 3 in our algorithm) are selected as the correct match.

7. Experiments and analysis

Thirty-two image patches have been used to evaluate the proposed algorithm. Table 1 shows the details of the image dataset in each section. The fundamental matrix is estimated via SIFT matching [30] and RANSAC [31]. Our algorithm is mainly compared with the two state of the art methods, LPI [17] and LJL [18]. The classical method proposed by Schmid et al. [4] is also compared in Section 7.3.

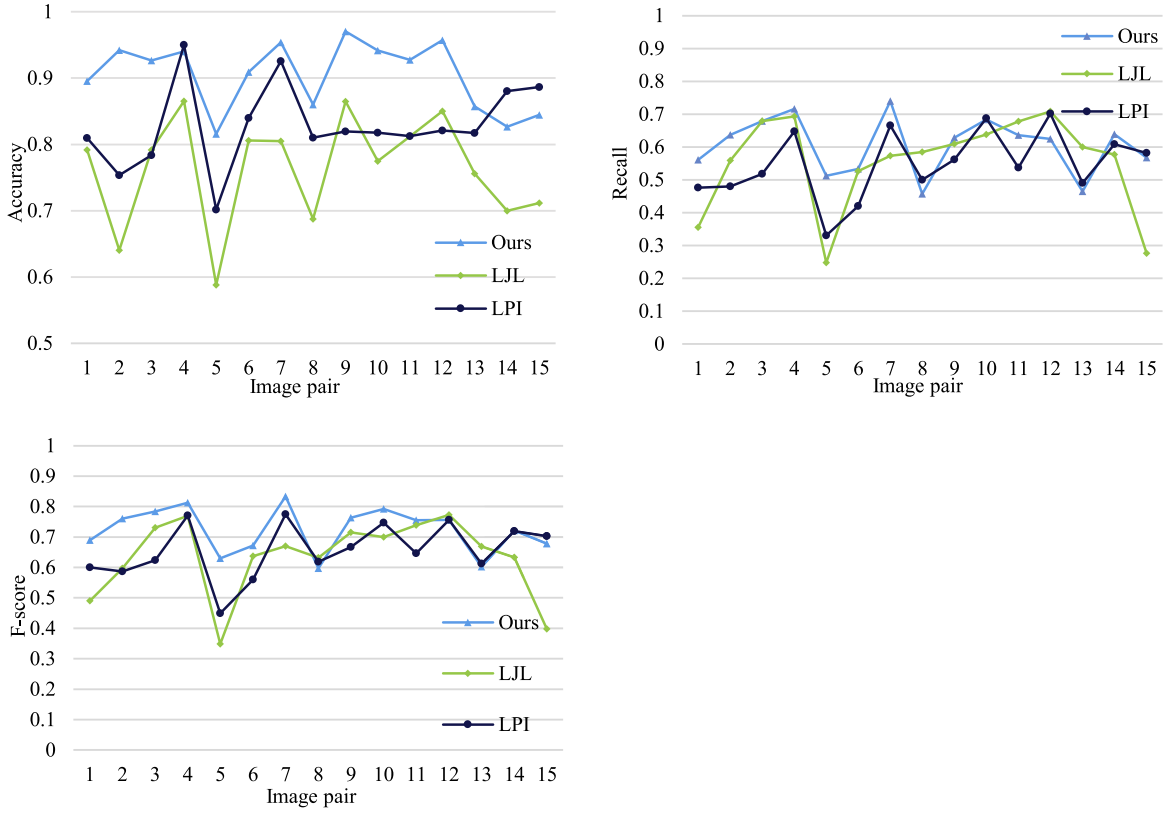


Fig. 8. Experimental results of the aerial stereo images in Fig. 7.

Table 1
The details of the image dataset in each section.

Section	Image type	Stereo number	Image source	Line segment	Fundamental matrix	
7.1	Natural	6	[4,18]	EDPF [33]	SIFT+RANSAC	RANSAC parameters:
7.2, 7.5	Aerial	15	[11]	Provided in [11]	SIFT+RANSAC	number of random trials, 9999;
7.3	Natural	18	[4]	EDPF	Provided in [4]	distance threshold, 0.01;
7.4	Aerial	3	Ourselves	EDPF	SIFT+RANSAC	desired confidence, 99%.

Table 2
The matching result (correct matches, accuracy) of the stereo images in Fig. 6. The number in bold represents the best.

	(1)	(2)	(3)	(4)	(5)	(6)	Summary
Ours	**226**,**99.6%**	**82**,**98.8%**	**420**,98.1%	344,**98.3%**	**560**,98.4%	363,**89.6%**	**1995**,**96.8%**
LPI	224,97.5%	79,97.5%	411,**98.3%**	312,98.1%	556,**99.5%**	**369**,85.8%	1954,96.1%
LJL	226,99.1%	79,97.5%	328,85.7%	**367**,95.2%	549,94.3%	414,86.9%	1963,91.7%

For the dataset in Section 7.2, the ground truth is provided by OK et al. [11] and the algorithms are evaluated by precision, recall, and F-score. The correct matches of other datasets are confirmed through one-by-one visual inspection. Since the line segments extracted by line detectors are often broken and with position errors, for the match denoted as $(l_c, \tilde{l}_{c'})$, we first find the complete line segments $(l_t, \tilde{l}_{t'})$ in stereos that have the same pre-image in 3D space, note that l_t and $\tilde{l}_{t'}$ are identified by our judgement but not the line detector. If l_c and l_t satisfy Eq. (4), and so do $\tilde{l}_{c'}$ and $\tilde{l}_{t'}$, $(l_c, \tilde{l}_{c'})$ is considered as the correct match. In our inspection, the distance error (T_{map} in Eq. (4)) is less than 3 pixels and the overlapping rate (T_{ove} in Eq. (4)) is more than 40%. Having confirmed the correct matches, the algorithms are evaluated by the number of total matches (TM), the number of correct matches(CM), and the accuracy are calculated as CM/TM. The details including the wrong label can be found on the website [32].

7.1. Matching for natural images under specific transformations

In this section, experiments were employed on six image pairs with five specific transformations. The transformations include rotation, scale illumination, occlusion, and viewpoint change. Fig. 6 shows the image pairs and Table 2 shows the results.

In the first five image pairs, all the three algorithms obtained a high accuracy, which exceeded 98%. Although the sixth image pair has the same transformation as the fifth, its scene structure is more complex, therefore, all the accuracies were below 90%. Compared with LPL and LJL, our algorithm ranked first in either the correct match or the accuracy for each image pair. In the summary our algorithm ranked first in both the correct match and the accuracy.

7.2. Matching for aerial images with man-made scenes

The aerial stereo dataset [11] was used to evaluate the algorithms. the dataset comprises fifteen image pairs (Fig. 7) and the author has provided the line segment and the ground truth. The algorithms were evaluated by the ratio of the number of correct matches and the number of ground truth matches (Recall), the ratio of the number of correct matches and total matches (Accuracy) and the F-score=$\frac{2\times Recall\times Accuracy}{Recall+Accuracy}$. Fig. 8 shows the evaluation results.

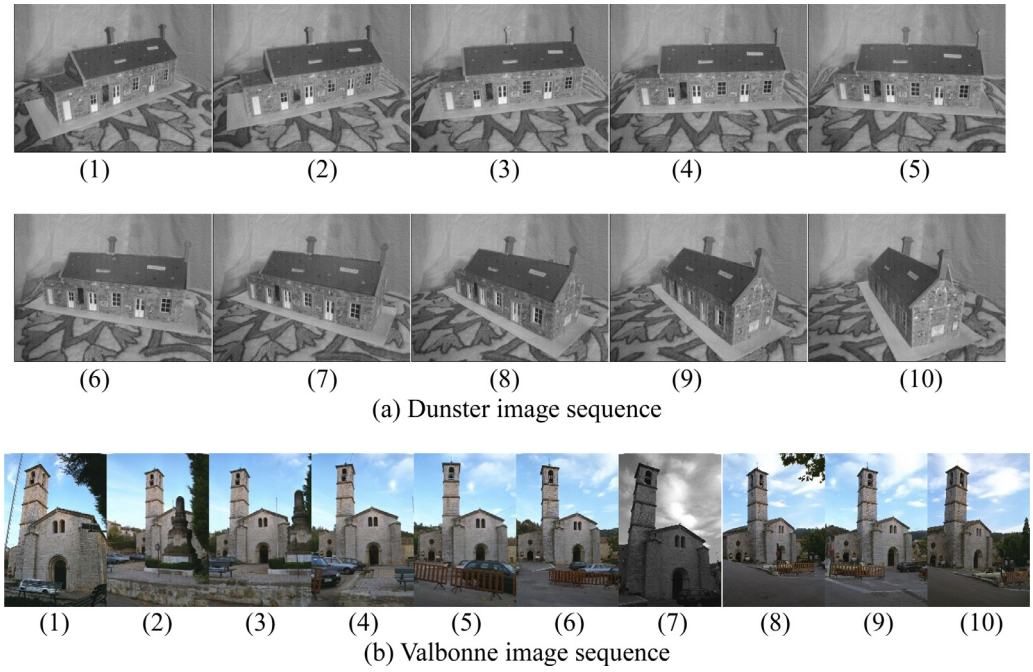
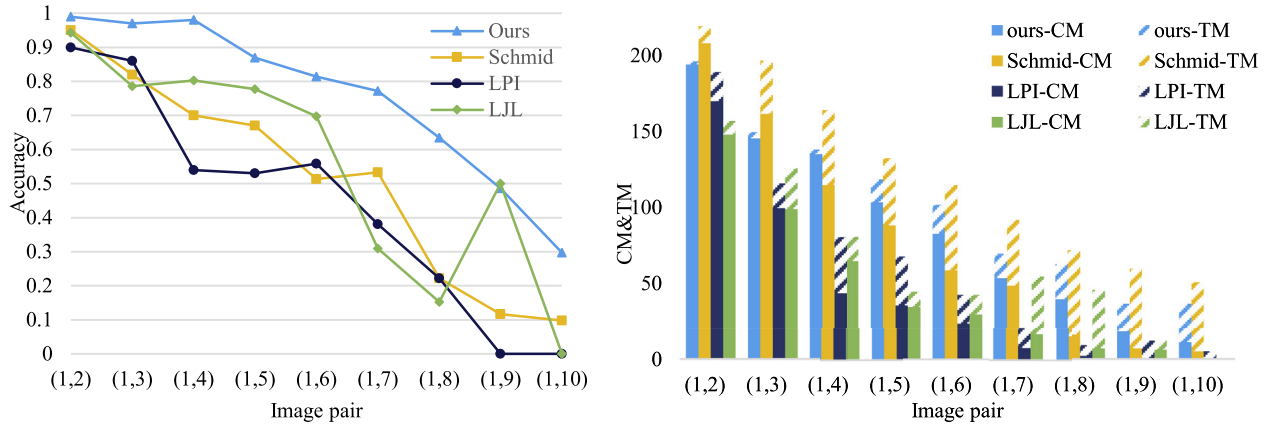
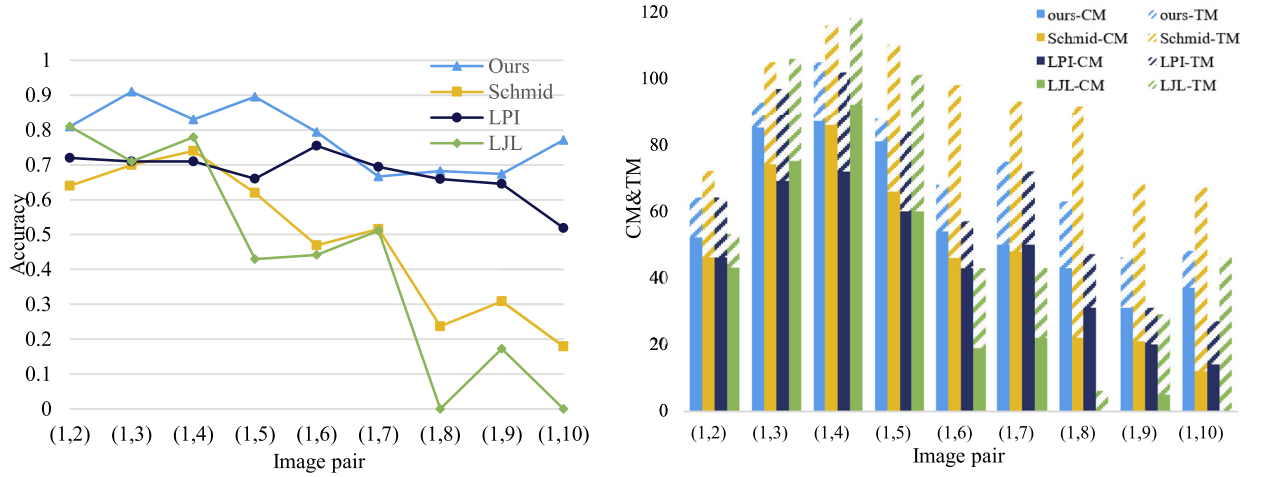


Fig. 9. The two image sequences that are provided in [4]. The baseline becomes wider and the viewpoint change becomes steeper for the first view. (a) Dunster image sequence, (b) Valbonne image sequence.



(a) Experimental results of Dunster image sequence (CM, correct matches; TM, total matches)



(b) Experimental results of Valbonne image sequence

Fig. 10. Experimental results of the two image sequences in Fig. 9. (a) Experimental results of Dunster image sequence (CM, correct matches; TM, total matches), (b) Experimental results of Valbonne image sequence.

For the accuracy, our algorithm ranked first in thirteen image pairs, and the average accuracy was higher than LPI and LJL by 10%. For the accuracy of the last two image pairs, our method ranked second and was lower than the first for 6%. LPI ranked first in the last two pairs. LJL achieved the last place in most of the image pairs.

For the recall, our method ranked first in half of the image pairs. LJL performed better than LPI in most of the image pairs. This result was expected. LPI relied on the point match, and in most cases, fewer point matches will lead to fewer line segment matches. LJL used the junction of the line segment pair to construct the descriptor, and fewer junctions would lead to fewer matches. Our algorithm relied on the line segment and the epipolar geometry, thus obtained more correct matches than LPI and LJL in most pairs.

The F-score is a combination of the recall and accuracy. Because our method obtained a good performance in both the recall and accuracy, it ranked first in eleven image pairs, and ranked second in other image pairs.

7.3. Matching with baseline and viewpoint changes

We used two image sequences that have been calibrated in Schmid's work [4] to evaluate the algorithms. Also, the Schmid's algorithm was evaluated. As shown in Fig. 9, these images are variant in depth, viewpoint, scale, and occlusion. To evaluate the four algorithms in different baseline and viewpoint, every image in the sequence was matched with the first image.

Fig. 10 shows the evaluation results. For all the four algorithms, both the accuracy and the correct match were declined with the increase of the baseline, but our method was superior to others. For the accuracy, our method ranked first in seventeen pairs, and the average accuracy was higher than Schmid's method, LJL, and LPI by 27%,27% and 21%, respectively. For the correct matches, our method ranked the first place in 18 pairs, and the average score was higher than Schmid's method, LJL and LPI by 9, 32 and 28, respectively.

The proposed algorithm was robust in matching with the change of the baseline and viewpoint. Because it relied on the **F** matrix, which is invariant to the projective transformation and is independent of the texture. Fig. 11 shows the local matching results for Dunster (8,10), where there is steep viewpoint change and various local homographies. Our method obtained most of the correct matches in the different scene planes. Schmid's method used the texture-correlation score to confirm the correct match, which failed when the texture changed steeply. The performance of LPI was unsatisfactory because the point correspondence was not sufficient. LJL used the texture to initialize the candidate, which may fail when the texture distortion is large. Therefore, LJL found only a few matches.

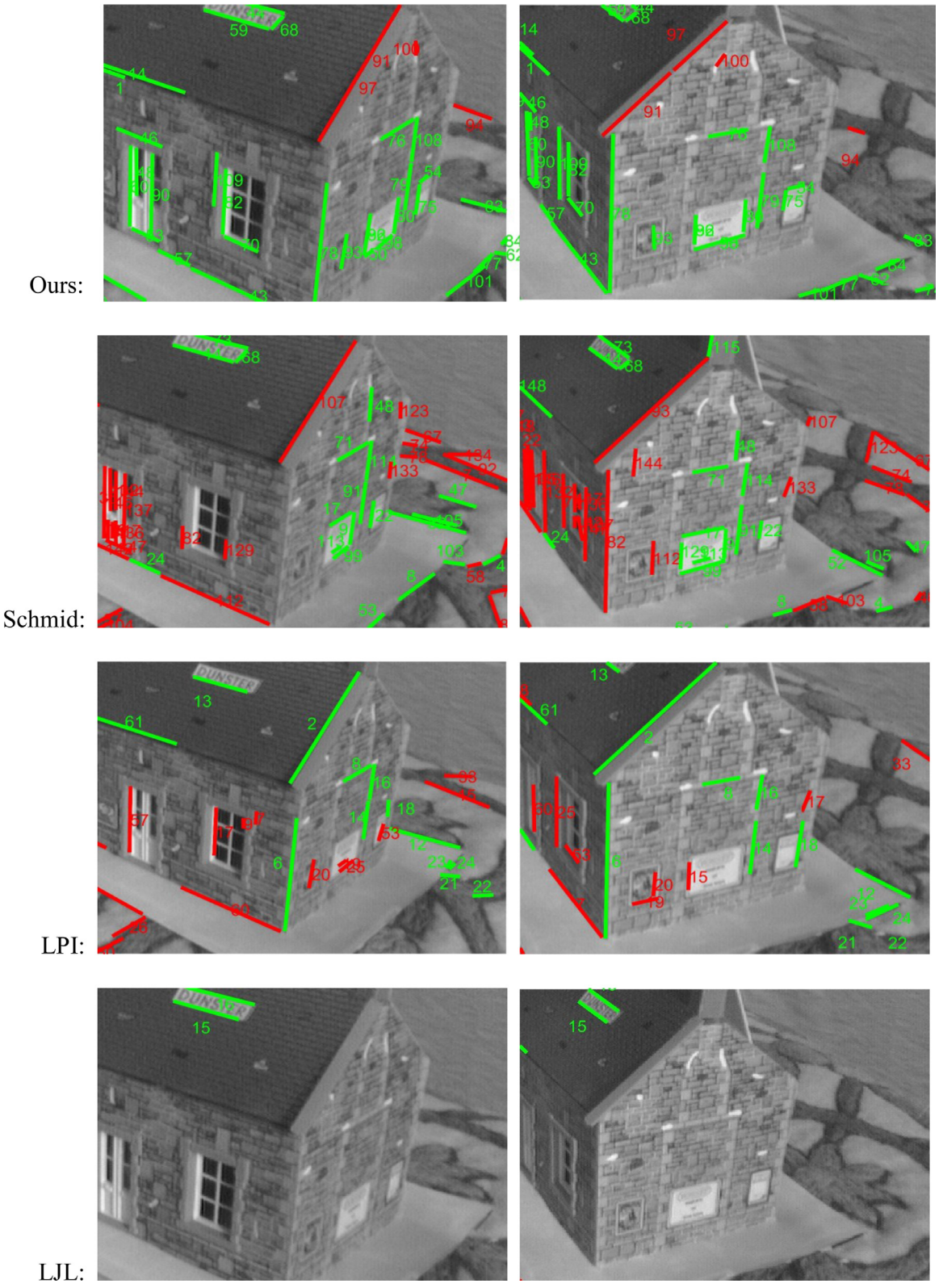


Fig. 11. Local matching results of Dunster (8,10). The wrong match is plotted in red.

7.4. Matching with dense line segments

Dense line segments in high-resolution images can make troubles to match line segments. First, the distance between line segments is quite close, thus the geometry constraint may be invalid. Second, textures in dense line segments are similar to each other, thus the texture correlation may fail. Three image pairs (Fig. 12) were used to evaluate our method in matching with dense line segments. To identify the correct match, we first calculated the homography of the local plane, then, we mapped the line segment in the first view to the second view for overlap analysis (Fig. 13).

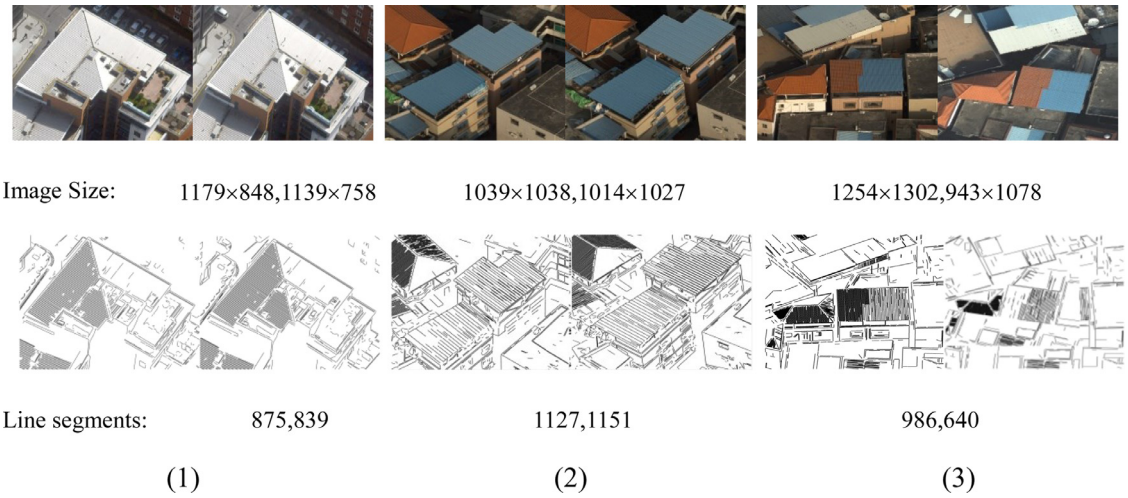


Fig. 12. The aerial image pairs and the dense line segments.

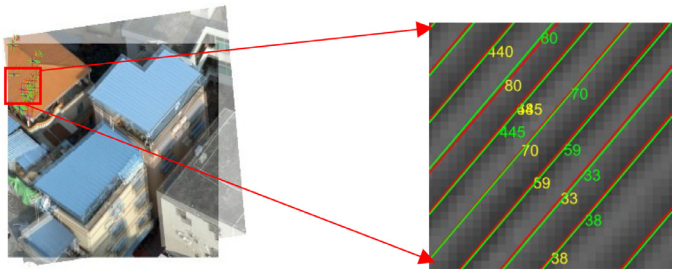


Fig. 13. To check the correct match, the line segment in the first view was mapped to the second view by the local homography.

Table 3 shows the quantitative evaluation and Fig. 14 plots the result. The first image pair has a slight viewpoint change. Thus, our algorithm achieved the accuracy of 98.2% and LJL achieved the accuracy of 95.4%. LPI ranked third in both the accuracy and the correct match. The second image pair has a larger viewpoint change and its scene structure is more complicated. All algorithms were declined in accuracy. However, our method still obtained a high accuracy of around 95%. The third image pair has obvious changes in viewpoint and scale. Thus, all the three methods obtained an unsatisfied result. Our method obtained the highest accuracy, moreover, the correct matches of LPI and LJL were lower than our method.

Table 3
Evaluation results (correct matches, accuracy) of the image pairs in Fig. 12. The number in bold represents the best.

	(1)	(2)	(3)
Ours	**484,98.2%**	**560,94.9%**	**171,54.8%**
LPI	393,85.4.5%	335,78.1%	55,36.7%
LJL	474,95.4%	488,79.0%	21,12.1%

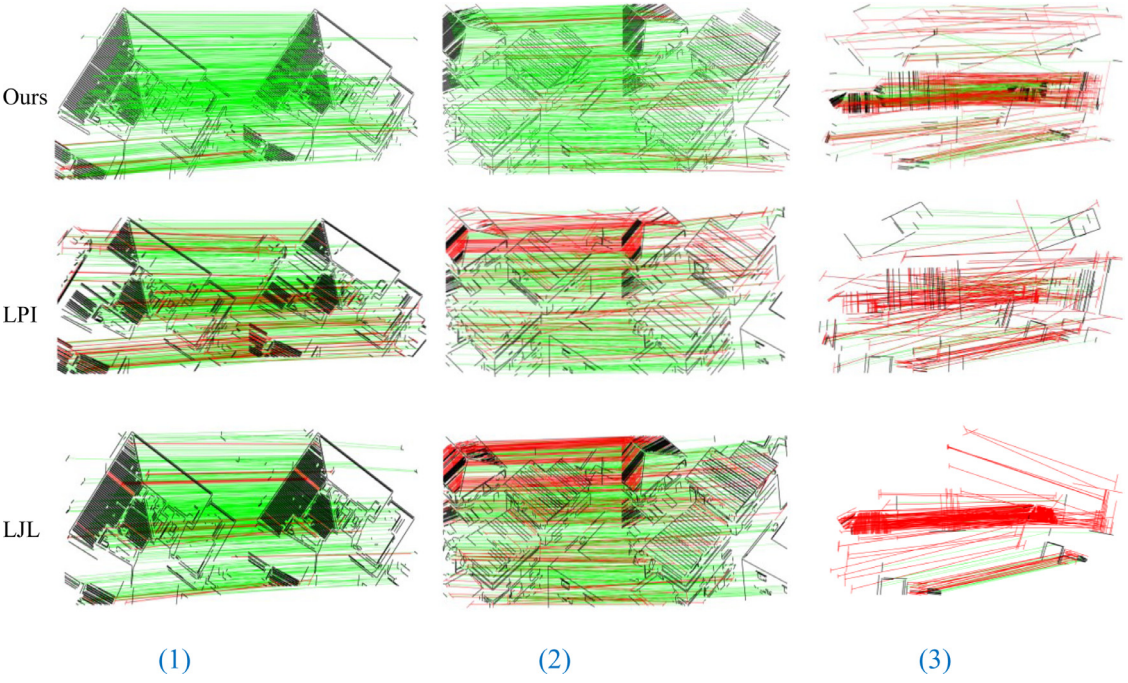


Fig. 14. The result of matching with dense line segments. The incorrect match is plotted in red and the correct is plotted in green.

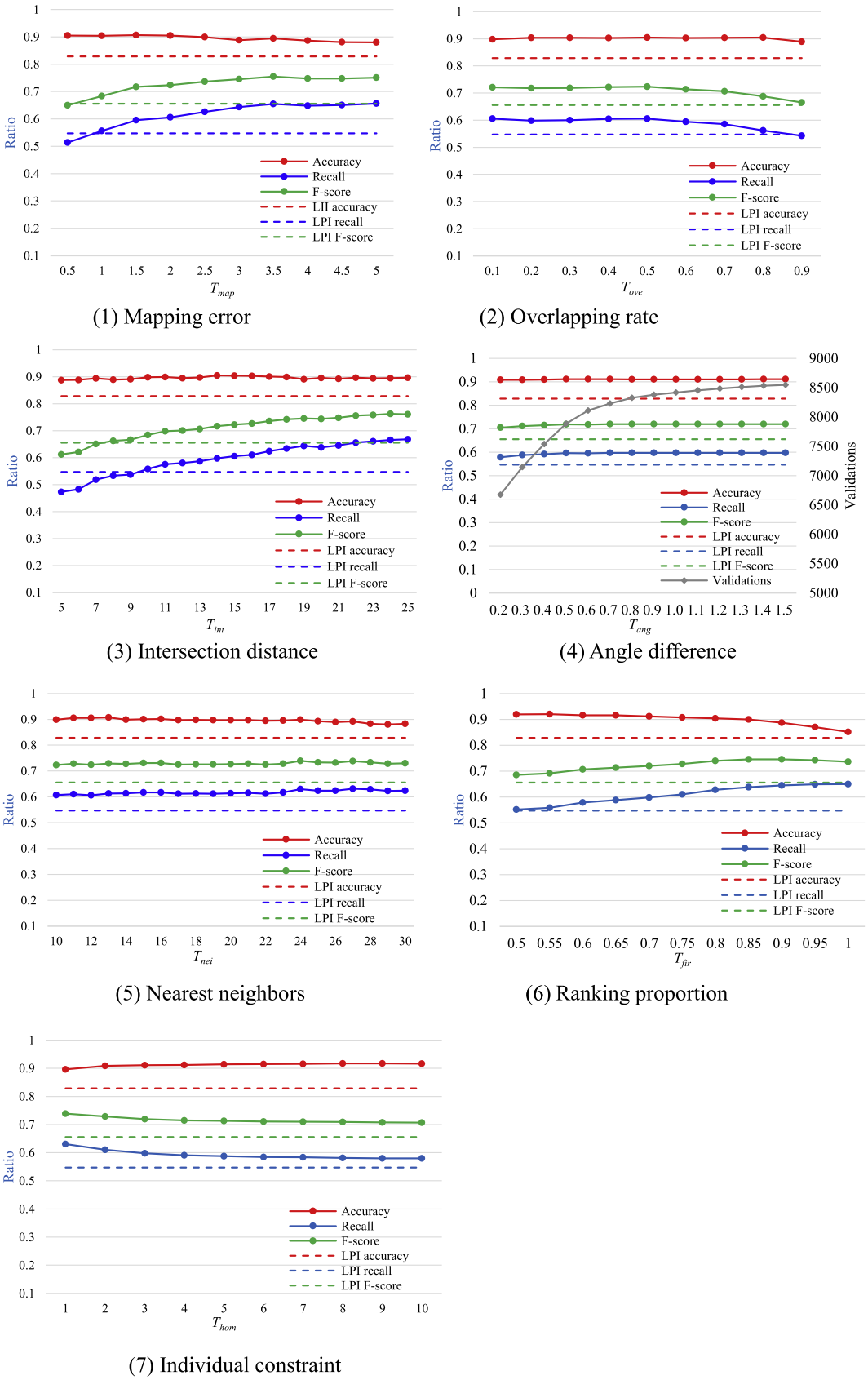


Fig. 15. Evaluation results of the image dataset [11] with threshold changes. (1) Mapping error, (2) Overlapping rate, (3) Intersection distance, (4) Angle difference, (5) Nearest neighbors, (6) Ranking proportion, (7) Individual constraint.

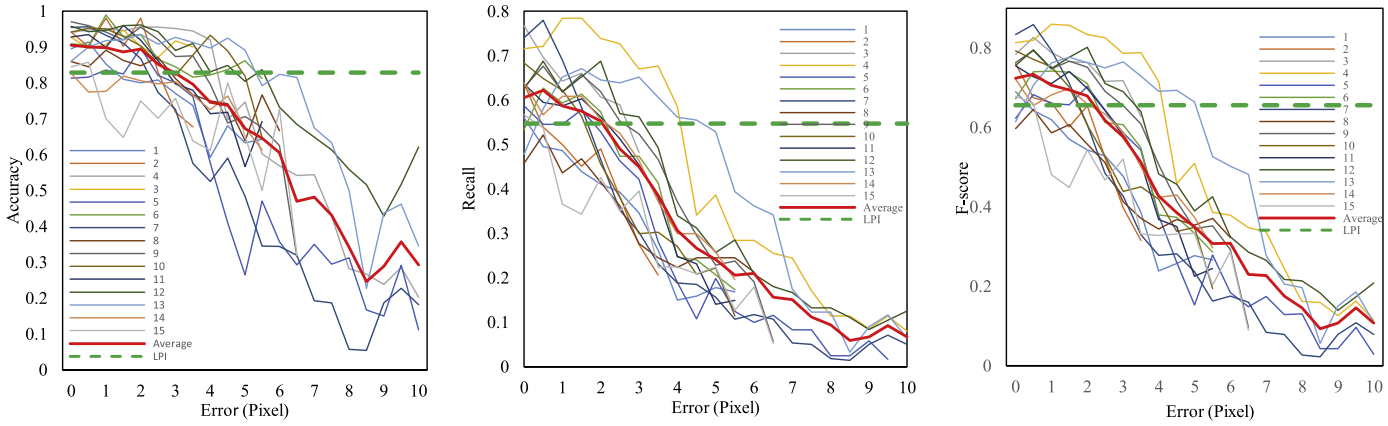


Fig. 16. Evaluation of the image dataset [11] with errors in F matrix.

Table 4
All the thresholds in our algorithm.

Threshold	Symbol	Section	Unit	Value	Evaluation range
Mapping error	T_{map}	3	Pixel	2	0.5–5.0
Overlapping rate	T_{ove}	3	None	0.5	0.1–0.9
Intersection distance	T_{int}	4.1	Pixel	15	5–25
Angle difference	T_{ang}	4.1	Radian	$\pi/4$	$\pi/18$-$\pi/2$
Nearest neighbors	T_{nei}	4.2	Line segment	20	10–30
Ranking proportion	T_{fir}	6	None	0.7	0.5–1
Individual constraint	T_{hom}	6	Homography	3	1–10

The proposed algorithm is appropriate for dense line segments matching. First, our algorithm relies on the epipolar geometry, which will not be influenced by the ambiguous texture. Second, dense line segments generally occur in scene planes, which enable constructing the robust homography graph to rank the candidate.

7.5. Matching with the threshold changes

This section analyzes the influences of all the internal thresholds. Table 4 shows the details of the thresholds in our algorithm. We evaluated their influences via changing these internal thresholds and tested them on the aerial image dataset introduced in Section 7.2. The evaluation results are shown in Fig. 15. The performance of LPI is plotted in dotted line for compare.

The evaluation results showed that our algorithm is robust to the change of the threshold. First, the result changed in a slow trend with the variation of the threshold. On the other hand, although the threshold had changed, our method remained to be better than both LPI and LJL in most cases. The detailed analyses are as follows.

(1) With the increase of T_{map}, our algorithm produced more initial candidates, while brought about more incorrect nodes. Thus, the recall and F-score raised at the first time and then kept steady.
(2) With the increase of T_{ove}, some matches were rejected because of the fraction of the line segment. Thus, both the recall and F-score went down.
(3) Changing T_{int} was similar to change T_{map}, because a larger intersection threshold reduced the false negative while increased the false positive.
(4) T_{ang} has a slight influence to the result, but the number of the validations of the homography constraint raised with the increase of T_{ang}.
(5) With the increase of T_{nei}, more neighbors that were not coplanar were added in, thus the accuracy went down slowly.
(6) With the increase of T_{fir}, more candidates were obtained without the homography constraint, thus the accuracy went down and the recall raised.
(7) Finally, with the increase of T_{hom}, the constraint for matching the individual line segment became stricter, thus the accuracy raised slightly and the recall went down.

7.6. Influence of the error in fundamental matrix

Since the proposed algorithm relies on **F** matrix to establish the homography constraint, the error in the matrix will inevitably influence the matching result. The fifteen stereos introduced in Section 7.2 were employed to analyze the limitation of the dependence of our algorithm on the known **F** matrix. Three steps were employed to add error to **F** matrix: (1) eight point correspondences were generated based on the original **F** matrix; (2) the position errors were added to the eight point correspondences; (3) the **F** matrix with intentional errors was estimated via normalized eight-point algorithm. The error in **F** matrix was evaluated by the average distance of 100 points and their epipolar lines, and the point correspondences were generated randomly by the original **F** matrix.

Fig. 16 presents the evaluation of the dependence of our algorithm on **F** matrix. Obviously, it had degenerated with the increase of the error in **F** matrix. When the average distance of the point to the epipolar line was within 3 pixels, our algorithm obtained a satisfactory result. But the performance degenerated quickly when the distance was over 3 pixels, and some stereos could not be matched. It has manifested that our algorithm requires a reliable **F** matrix to obtain the robust match.

8. Conclusions

This paper presents a novel line segment matching algorithm based on the epipolar geometry. The method initializes the candidate based on the homography constraint. The homography graph is then constructed, and the reweighted random walk method is employed on this graph to rank the candidates. Finally, the correct matches are obtained with the constrained greedy algorithm. The proposed approach relies purely on the epipolar geometry which is invariant to the local projective transformation. The experiments demonstrate the robustness of our algorithm when compared with the state of the art methods and the classical method. The proposed algorithm is easy to use because it requires only the fundamental matrix of image pair, which can be obtained via the advanced point matching algorithm, or the pre-calibrated binocular camera. Also, we must admit the weaknesses of our algorithm. First, since the homography constraint is applied to the

scene plane, it will fail when the scene has few coplanar 3D line segments. Second, when the fundamental matrix is not available or not reliable, the proposed algorithm cannot work.

Finally, there are two aspects which we are currently investigating. The first is to embed multiple-view geometries of different features such as line and point across multiple views into the homography graph. Thus, both the line and point candidate can be ranked by considering the candidate in not only the same stereo and also other stereos. The second is to speed up the algorithm with GPU because the construction and ranking of the graph require large computations, and running in real time will make it more practical.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by The National Key Research and Development Program of China under Grant 2017YFB0503004, and in part by the National Natural Science Foundation of China under Grant 41871368. We would like to thank the authors Schmid et al. [4], Fan et al. [18] and Li et al. [17] to share their codes or test datasets, thank the Author Ok et al. [11] for their kind reply and providing the aerial image dataset with the ground truth. They are quite helpful for our work in this paper.

References

[1] M. Hofer, M. Maurer, H. Bischof, Efficient 3D scene abstraction using line segments, Comput. Vision Image Underst. 157 (2016) 167–178.
[2] X. Liu, Y. Zhang, X. Ling, et al., TopoLAP: topology recovery for building reconstruction by deducing the relationships between linear and planar primitives, Remote Sens. (Basel) 11 (11) (2019) 1372.
[3] C. Li, Y. Zhang, Z. Zhang, Automatic keyline recognition and 3D reconstruction for quasi-planar facades in close-range images, Photogrammetric Rec. 31 (153) (2016) 29–50.
[4] Schmid, C. and A. Zisserman. Automatic line matching across views. CVPR, 1997.
[5] Bay, H., V. Ferrari, and L.J.V. Gool. Wide-baseline stereo matching with line segments. CVPR. 2005.
[6] Z. Wang, F. Wu, Z. Hu, MSLD: a robust descriptor for line matching, Pattern Recognit. 42 (5) (2009) 941–953.
[7] L. Zhang, R. Koch, J. Jvcir, An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency, J. Vis. Commun. Image Represent. 24 (7) (2013) 794–805.
[8] Jain, A., C. Kurz, T. Thormählen, et al. Exploiting global connectivity constraints for reconstruction of 3D line segments from images. CVPR. 2010.
[9] N. Ayache, B. Faverjon, Efficient registration of stereo images by matching graph descriptions of edge segments, Int. J. Comput. Vis. 1 (2) (1987) 107–131.
[10] W.J. Christmas, Structural matching in computer vision using probabilistic reasoning, IEEE Trans. Pattern Anal. Mach. Intell. 17 (8) (1995) 749–764.
[11] A.O. Ok, J.D. Wegner, C. Heipke, et al., Matching of straight line segments from aerial stereo images of urban areas, ISPRS J. Photogrammetry Remote Sens. 74 (6) (2012) 133–152.
[12] E. Tola, V. Lepetit, P. Fua, DAISY: an Efficient Dense Descriptor Applied to Wide-Baseline Stereo, IEEE Trans Pattern Anal Mach Intell 32 (5) (2010) 815–830.
[13] M. Al-Shahri, A. Yilmaz, Line matching in wide-baseline stereo: a top-down approach, IEEE Trans. Image Process. 23 (9) (2014) 4199–4210.
[14] J. López, R. Santos, X.R. Fdez-Vidal, et al., Two-view line matching algorithm based on context and appearance in low-textured images, Pattern Recognit. 48 (7) (2014) 2164–2184.
[15] K. Hyunwoo, L. Sukhan, L. Yeonho, Wide-baseline stereo matching based on the line intersection context for real-time workspace modeling, J. Opt. Soc. America Optics Image Sci. Vision 31 (2) (2014) 421–435.
[16] L. Kai, Y. Jian, Line segment matching and reconstruction via exploiting coplanar cues, ISPRS J. Photogrammetry Remote Sens. 125 (2017) 33–49.
[17] K. Li, J. Yao, X. Lu, et al., Hierarchical line matching based on line-junction–line structure descriptor and local homography estimation, Neurocomputing 184 (2016) 207–220.
[18] B. Fan, F. Wu, Z. Hu, Robust line matching through line–point invariants, Pattern Recognit. 45 (2) (2012) 794–805.
[19] Y. Sun, L. Zhao, S. Huang, et al., Line matching based on planar homography for stereo aerial images, ISPRS J. Photogrammetry Remote Sens. 104 (2015) 1–17.
[20] X. Jia, X. Huang, F. Zhang, Y. Gao, C. Yang, Robust line matching for image sequences based on point correspondences and line mapping, IEEE Access 7 (2019) 39879–39896.
[21] Q. Wang, W. Zhang, X. Liu, et al., Line matching of wide baseline images in an affine projection space, Int. J. Remote Sens. 41 (2) (2020) 632–654.
[22] W. Wang, W. Gao, H. Cui, et al., Reconstruction of lines and planes of urban buildings with angle regularization, ISPRS J. Photogrammetry Remote Sens. 165 (2020) 54–66.
[23] Cho, M., J. Lee, and K.M. Lee. Reweighted random walks for graph matching. ECCV. 2010.
[24] Leordeanu, M. and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. ICCV. 2005. 1482–1489.
[25] Page, L., The pagerank citation ranking: bringing order to the web. Stanford Digital Library Technologies Project, 1998.
[26] O. Duchenne, F.R. Bach, I.S. Kweon, et al., A tensor-based algorithm for high-order graph matching, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2383–2395.
[27] Lee, J., M. Cho, and K.M. Lee. Hyper-graph matching via reweighted random walks. CVPR. 2011. 1633–1640.
[28] R. Sinkhorn, A relationship between arbitrary positive matrices and doubly stochastic matrices, Ann. Math. Stat. 35 (2) (1964) 876–879.
[29] X. Bai, L.J. Latecki, Path similarity skeleton graph matching, IEEE Trans. Pattern Anal. Mach. Intell. 30 (7) (2008) 1282–1292.
[30] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[31] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Commun. ACM 24 (6) (1981) 381–395.
[32] The implementation of our algorithm. Available from: https://github.com/weidong-whu/line-match-RRW.
[33] C. Akinlar, C. TOPAL, EDPF: a real-time parameter-free edge segment detector with a false detection control, Int. J. Pattern Recognit. Artif. Intell. 26 (01) (2012) 1255002–1255022.

Dong Wei was born in 1992. He received the B.S. and M.S. degrees from Central China Normal University, Wuhan, China, in 2015 and 2018, respectively. He is now studying for the doctoral degree of photogrammetry and remote sensing in the School of Remote Sensing and Information Engineering, WHU.

Yongjun Zhang was born in 1975. He received the B.S., M.S., and Ph.D. degrees from Wuhan University (WHU), Wuhan, China, in 1997, 2000, and 2002, respectively. He is currently a Professor of photogrammetry and remote sensing with the School of Remote Sensing and Information Engineering, WHU.

Chang LI was born in 1982. He received the B.S., M.S., and Ph.D degrees from Hubei University, China University of Geosciences (Wuhan) and Wuhan University, China, in 2004, 2006, 2009 respectively. He is currently a Associate Professor of geographic information science with College of Urban and Environmental Science, Central China Normal University, Wuhan, China.