Machine Vision and Applications (2013) 24:637–649
DOI 10.1007/s00138-012-0412-x

ORIGINAL PAPER

A robust non-iterative method for similarity transform estimation

Li Yinan · Yang Lingyu · Shen Gongzhang

Received: 17 June 2011 / Revised: 18 October 2011 / Accepted: 19 January 2012 / Published online: 19 February 2012

Abstract A non-iterative and robust method—direct outliers remove (DOR) is proposed, which efficiently estimates the similarity transform based on a data set containing both correct and incorrect correspondences. Unlike hypothesize-and-test methods such as Random Sample Consensus algorithm and its variants, DOR removes mismatches by exploring all the correspondences only once, using two invariant features of similarity transform. One is the angles between two vectors and the other is the length ratios of corresponding vectors. Given two images related by similarity transform, experiments demonstrate that all the mismatches introduced in matching stage could be detected and removed. Without losing computational accuracy, DOR is faster compared with several hypothesize-and-test algorithms, especially when the percentage of correct correspondence is relatively low.

Keywords Similarity invariance · Transform model estimation · Non-iterative · Robust

1 Introduction

Transform model estimation is a critical step in image registration and many computer vision problems. It is usually applied after the detection and matching of features from different scenes or images. So far, several feature detectors and descriptors have been developed. Among them are the famous scale and affine invariant Harris detector [1], MSER [2] and SIFT [3]. Although most of the feature points are invariant to image grayscale, scale and rotation, due to image distortion, cluttered background, noise infection, repetitive patterns and some other disturbing situations, erroneous correspondences are hardly avoided. Referring to related researches, the correct correspondences are defined as *Inliers* while the ones that do not match the real model are *Outliers*. Since the outliers would have great influence on estimation accuracy, there is an urgent demand for robust estimation methods to automatically exclude them.

Random sample consensus (RANSAC) [4,5] has proven very successful in solving such problems. Standard RANSAC consists of two steps which repeat under the hypothesize-and-test framework. A minimum number of samples are randomly selected to compute the model parameters in the first step. In the second step, all the samples that are consistent with the model estimated in the first step are collected in the Consensus set (CS). The iterative loop with these two steps terminates when the probability of finding a larger CS falls below a pre-determined threshold. However, two major disadvantages of RANSAC—sensitive to the setting of the appropriate noise threshold and low efficiency caused by randomly selecting minimum samples—are also pointed out.

To compensate for the first unexpected effect, Torr and Zisserman [6] proposed MSAC (M-estimator SAmple Consensus) and MLESAC (Maximum Likelihood Estimation SAmple Consensus). Instead of maximizing the cardinality of CS, MSAC minimizes the cost function based on M-estimators, and MLESAC evaluates the likelihood of the hypothesis by applying a mixture model of error distribution. As an extension to MLESAC, Guided-MLESAC [7] was also developed to speed up the procedure of estimation according to prior information. To minimize computational time,

L. Yinan (✉) · Y. Lingyu · S. Gongzhang
School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics (Beihang University), 37 Xue Yuan Road, Hai Dian District, Beijing 100191, People's Republic of China
e-mail: mmchong.yinan@gmail.com

PROSAC (Progressive SAmple Consensus) was proposed in [8]. In this approach, samples are sorted and selected according to a similarity function $q(.)$. NAPSAC (N adjacent points sample consensus) [9] and GASAC (Generic algorithm sample consensus) [10] are the other two representative estimators designed to promote efficiency in dealing with high-dimensional problems or stereo matching problems. GroupSAC [11] upgrades the efficiency in the presence of low inlier ratios by group sampling. And it is also based on the heuristic information of natural groupings.

Similarity transform is one of the frequently used mapping functions and only consists of rotation, translation and scaling between two images [12]. In many applications such as face recognition, global motion estimation, and scene matching navigation, either purely similarity transform exists or similarity transform approximation is satisfied. From Reddy [13] to Tzimiropoulos [14], numerous researches have been conducted to estimate the model parameters.

Since this kind of transform is relatively simple—only two correspondences are required to calculate four model parameters, the efficiency of most of the improved RANSAC algorithms will drop to the level of RANSAC. Therefore, our objective is to design an approach that is well suited to estimating similarity transform with higher efficiency. Under the scheme of DOR, we remove the erroneous correspondences directly by investigating two invariant properties of similarity transform—the angle between two line segments and the length ratio of them. Following this strategy, we are able to avoid the hypothesize-and-test framework which slows down RANSAC and its variants and thus save computational time, especially in low percentage of inliers. Experimental results show the validity and efficiency of the proposed method.

The rest of the paper is organized as follows: Sect. 2 presents the related similarity invariance and the overall scheme of DOR. Section 3 discusses several problems in implementation. Experiments are provided to illustrate the key features of the proposed method in Sect. 4. Finally, Sect. 5 summarizes this whole work.

2 Method

The overall structure of DOR consists of two steps. First, remove outliers from the tentative data set. Second, compute model parameters based on identified inliers adopting LSE (least square estimation). The procedure is similar to RANSAC but in a non-iterative way. One distinctive difference from RANSAC is the approach to filter the original data set. Rather than picking up a subset with minimum samples and then expanding it to the maximum size, DOR removes outliers directly from the original sample set. This section aims to illustrate its principle.

2.1 Invariant properties of similarity transform

Shapes will not change under similarity transform. If two images G_1, G_2 are related by translation (t_x, t_y), rotation θ and scale factor s, the transform model can be described as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta & t_x \\ -\sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{1}$$

where $(u \quad v)^T$ and $(x \quad y)^T$ are the corresponding locations of a feature point in G_1 and G_2, respectively. $[\theta \quad s \quad t_x \quad t_y]^T$ is the parameter vector of the 4-DOF model.

Let $\mathbf{l}^{(1)} = [\Delta u \quad \Delta v]^T$ be a vector connecting two feature points in G_1, and $\mathbf{l}^{(2)} = [\Delta x \quad \Delta y]^T$ be the corresponding vector in G_2, then

$$\left\| \mathbf{l}^{(1)} \right\| = \|s\mathbf{R}\| \bullet \left\| \mathbf{l}^{(2)} \right\| = s \bullet \left\| \mathbf{l}^{(2)} \right\| \tag{2}$$

$$\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = s\mathbf{R} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{3}$$

where $\| \bullet \|$ is Euclidean norm. $[\Delta u, \ \Delta v]^T$ and $[\Delta x, \ \Delta y]^T$ are the locations of the points described in Cartesian coordinate system. Superscripts in this paper indicate the image number.

It can be easily derived from (2) that the magnitude ratio of any corresponding vectors is a constant factor s. We denote the direction angle of $\mathbf{l}^{(1)}$ by $\beta^{(1)}$ and the direction angle of $\mathbf{l}^{(2)}$ by $\beta^{(2)}$. Interpreting (3) with the direction angle gives

$$\begin{aligned} \tan\beta^{(2)} &= \frac{\Delta y}{\Delta x} = \frac{-s\Delta u \sin\theta + s\Delta v \cos\theta}{s\Delta u \cos\theta + s\Delta v \sin\theta} \\ &= \frac{-\tan\theta + \tan\beta^{(1)}}{1 + \tan\theta \tan\beta^{(1)}} = \tan(\beta^{(1)} - \theta) \end{aligned} \tag{4}$$

From (4), it is evident that the angle is unchanged if its both sides rotate the same degree θ in the same direction.

2.2 Removing outliers

The first step of DOR is to discard outliers and collect inliers directly using invariance of the similarity transform, which synthesizes two approaches: the sequence-based approach and the inlier-based approach.

Assume that G_2 and G_1 are related by translation, rotation and scaling. Figure 1 shows seven pairs of correspondences which are found in the image pre-processing stage for illustration. The inliers are consistent with (1) within a certain error threshold while the outliers are not. Feature points in G_1 with the index 1 through 7 are the matching points in G_2 with the same index. Among all the correspondences, 4 is apparently an outlier.

The sequence-based approach is to connect the points in ascending order according to the indices, and finally connect the first one with the last one, generating two sequences of

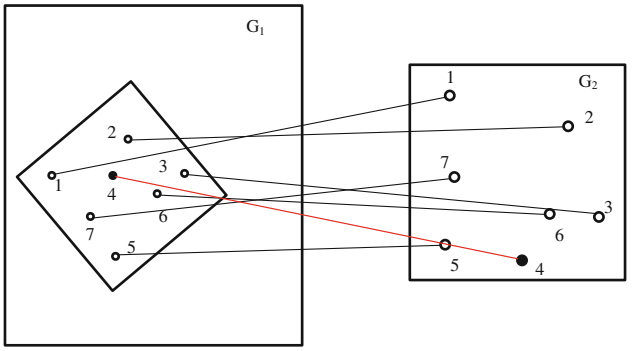

Fig. 1 Two given images G_1 and G_2 with seven correspondences. Feature points are presented in *circles* along with corresponding indices. 1, 2, 3, 5, 6, 7 are inliers while 4 is an outlier. Matched points in these two images are connected in *solid lines*

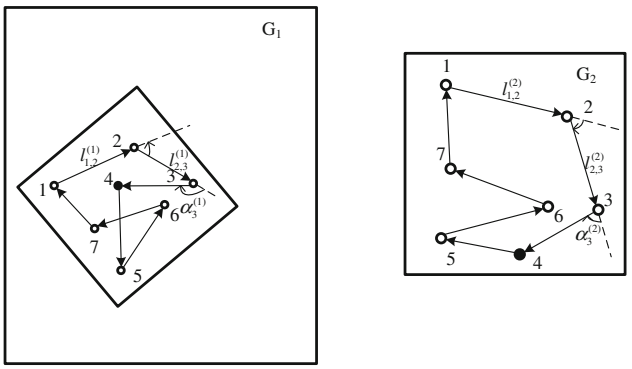

Fig. 2 Images with sequences of vectors. According to the storing order, feature points in G_1 and G_2 are connected, respectively. Influenced by the outlier 4, $l_{43}^{(1)}/l_{43}^{(2)}$ and $l_{54}^{(1)}/l_{54}^{(2)}$ are not equal or approximate to the scale factor s, and the angle equation also does not holds at point 3, 4, 5

vectors $\{\mathbf{l}_{2,1}^{(1)}, \mathbf{l}_{3,2}^{(1)}, \ldots \mathbf{l}_{N,N-1}^{(1)}, \mathbf{l}_{1,N}^{(1)}\}$ and $\{\mathbf{l}_{2,1}^{(2)}, \mathbf{l}_{3,2}^{(2)}, \ldots \mathbf{l}_{N,N-1}^{(2)}, \mathbf{l}_{1,N}^{(2)}\}$, where N is the overall number of correspondences. The subscripts of the elements of each sequence indicate the initial point and terminal point. For an easier implementation, the indices are assigned in accordance with the storing order of these matches.

Without loss of generality, we denote the feature points in $G_k(k = 1, 2)$ by $F_N^{(k)}$, the correspondence by $(f_i^{(1)}, f_i^{(2)})$, $f_i^{(1)} \in F_N^{(1)}$, $f_i^{(2)} \in F_N^{(2)}, i \in [1, N]$. Let $I_M^{(k)}$ be the set which only consists of inliers, and $O_{N-M}^{(k)}$ be the set of outliers. M is the number of inliers found in the tentative data. Furthermore, we describe each vector in magnitude $l_{i,i-1}^{(k)}$ and direction $\beta_{i,i-1}^{(k)}$ separately. $\alpha_i^{(k)}$ is defined as the angle between two adjacent vectors, where the subscript denotes the mutual point of these two vectors. For example, $\alpha_3^{(k)}$ is the angle between vector $\mathbf{l}_{2,3}^{(k)}$ and $\mathbf{l}_{3,4}^{(k)}$. These vectors with their magnitudes and directions are marked in Fig. 2.

When both of the two endpoints are inliers, the magnitude equation (5) holds.

$$s = l_{i,i-1}^{(1)}/l_{i,i-1}^{(2)} \tag{5}$$

If all the three points associated with vectors $\mathbf{l}_{i,i-1}^{(k)}$ and $\mathbf{l}_{i+1,i}^{(k)}$ are inliers, the angle between them should satisfy the angle equation

$$\alpha_i^{(1)} = \alpha_i^{(2)} \tag{6}$$

Once all the feature points are connected in this way, each correspondence will be determined either as an inlier or an outlier according to Criterion 1 which integrates (5) and (6).

Criterion 1 If correspondences $f_{i-2}^{(k)}, f_{i-1}^{(k)}$ are inliers, and $l_{i,i-1}^{(k)}, \alpha_{i-1}^{(k)}$ satisfy the magnitude and angle equations, respectively, then $f_i^{(k)}$ can be considered as an inlier. Otherwise, it is an outlier.

Proof $f_{i-1}^{(1)}$ and $f_{i-1}^{(2)}$ are represented in Cartesian Coordinate System as $(u_{i-1} \quad v_{i-1})^T$ and $(x_{i-1} \quad y_{i-1})^T$. And they satisfy

$$\begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} = s \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \tag{7}$$

$\beta_{i-1}^{(k)}(k = 1, 2)$ is the direction of the vector connecting $f_{i-2}^{(k)}$ and $f_{i-1}^{(k)}$, then it could be obtained that $\beta_{i-1}^{(1)} - \beta_{i-1}^{(2)} = \theta_{i-1}$. In addition, from the known conditions, we get $l_{i,i-1}^{(1)} = s l_{i,i-1}^{(2)}, \alpha_{i-1}^{(1)} = \alpha_{i-1}^{(2)} = \alpha_{i-1}$, where s is the scale factor. Then

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = l_{i,i-1}^{(2)} \begin{bmatrix} \cos(\alpha_{i-1} - \beta^{(2)}) \\ \sin(\alpha_{i-1} - \beta^{(2)}) \end{bmatrix} + \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} \tag{8}$$

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = l_{i,i-1}^{(1)} \begin{bmatrix} \cos(\alpha_{i-1} - \beta^{(1)}) \\ \sin(\alpha_{i-1} - \beta^{(1)}) \end{bmatrix} + \begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} \tag{9}$$

(8) and (9) yield

$$\begin{aligned}
& s \begin{bmatrix} \cos\theta_{i-1} & \sin\theta_{i-1} \\ -\sin\theta_{i-1} & \cos\theta_{i-1} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \\
&= s l_{i,i-1}^{(2)} \begin{bmatrix} \cos\theta_{i-1} & \sin\theta_{i-1} \\ -\sin\theta_{i-1} & \cos\theta_{i-1} \end{bmatrix} \begin{bmatrix} \cos(\alpha_{i-1} - \beta^{(2)}) \\ \sin(\alpha_{i-1} - \beta^{(2)}) \end{bmatrix} \\
&\quad + s \begin{bmatrix} \cos\theta_{i-1} & \sin\theta_{i-1} \\ -\sin\theta_{i-1} & \cos\theta_{i-1} \end{bmatrix} \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \\
&= l_{i,i-1}^{(1)} \begin{bmatrix} \cos(\alpha_{i-1} - \beta_{i-1}^{(2)} - \theta_{i-1}) \\ \sin(\alpha_{i-1} - \beta_{i-1}^{(2)} - \theta_{i-1}) \end{bmatrix} + \begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} \\
&= l_{i,i-1}^{(1)} \begin{bmatrix} \cos(\alpha_{i-1} - \beta_{i-1}^{(1)}) \\ \sin(\alpha_{i-1} - \beta_{i-1}^{(1)}) \end{bmatrix} + \begin{bmatrix} u_{i-1} \\ v_{i-1} \end{bmatrix} \\
&= \begin{bmatrix} u_i \\ v_i \end{bmatrix}
\end{aligned}$$

Therefore, $f_i^{(k)}$ fits the model (1) and should be determined as an inlier. □

The inlier-based approach starts by connecting the current correspondence with two different inliers, yielding two vectors correspondingly. Then check whether Eqs. (5) and (6) hold for all the vectors. The one with both of the equations satisfied shall be viewed as an inlier.

Owing to the fact that both of the approaches work through formulating vectors between correspondences, they all rely on the selection of initial points set in some degree. But the difference is that the inlier-based approach is not able to start without the known set of inliers while the sequence-based approach could solve this problem by properly choosing the starting point (this issue will be discussed in Sect. 3). During the process of the sequence-based approach, some of the inliers would probably be deemed as outliers because of improper thresholds, which calls for the need of a second inliers selection. Hence, we propose to adopt the sequence-based approach in the first round to obtain the initial set of inliers and then enlarge this set using the inlier-based approach for its simpler procedure.

2.3 Least square estimation

The least square estimation upon collected inliers is applied to compute the transform parameters in the second step.

Gathering all the unknown parameters into a column vector, model (1) can be rewritten as

$$A\hat{M} = b \tag{10}$$

where

$$A = \begin{bmatrix} x_1 & y_1 & 1 & 0 \\ y_1 & -x_1 & 0 & 1 \\ & \cdots & \cdots & \\ x_i & y_i & 1 & 0 \\ y_i & -x_i & 0 & 1 \\ & \cdots & \cdots & \end{bmatrix}, \quad b = \begin{bmatrix} u_1 \\ v_1 \\ \cdots \\ u_i \\ v_i \\ \cdots \end{bmatrix},$$

$$\hat{M} = \begin{bmatrix} \hat{m}_1 & \hat{m}_2 & \hat{m}_3 & \hat{m}_4 \end{bmatrix}^T = \begin{bmatrix} \hat{s}\cos\hat{\theta} & \hat{s}\sin\hat{\theta} & \hat{t}_x & \hat{t}_y \end{bmatrix}^T$$

$(u_i \quad v_i)^T$, $(x_i \quad y_i)^T$ are the coordinates of correspondences in G_1 and G_2.

With the purpose of minimizing the distances between the projected model locations and their corresponding feature locations, the least square solution for the parameter vector M can be computed by the equation

$$\hat{M} = (A^T A)^{-1} A^T b$$

Then the four parameters can be determined by

$$\hat{s} = \sqrt{\hat{m}_1^2 + \hat{m}_2^2}, \quad \hat{\theta} = \tan^{-1}\frac{\hat{m}_2}{\hat{m}_1}, \quad \hat{t}_x = \hat{m}_3, \quad \hat{t}_y = \hat{m}_4.$$

3 Implementation

The method is simple in principle, but some details in implementation should be discussed.

3.1 Simplification of Criterion 1

Neither magnitude nor angle equation would hold for an outlier in most cases, but the angle condition is more vulnerable to outliers. The determination of the current correspondence would be influenced by its previous two matched pair of points. Moreover, computing the angle between two vectors involves three consecutive points while calculating the magnitude only involves two. Therefore, the Criterion 1 can be simplified by first considering the angle equation and then the magnitude ratios. In addition, since the scale factor is unknown a priori, the magnitude equation could not be used directly, which limits the usage of length ratio.

Note that noise and deformation exist through actual image acquisition and processing. The angles between two vectors in G_1 andG_2 are not strictly equal, so threshold T_g is set as the upper bound of the angle difference. Equation (6) holds if the angle difference $\Delta\alpha$ is lower than T_g.

If neither of the previous two correspondences of the current inlier fit the similarity transform model, the associated angle and magnitude would not agree with Criterion 1. We call the inlier of this kind a "false outlier". Take the situation shown in Fig. 2 as an example. Because of the outlier 4, (5) does not hold for $l_{43}^{(k)}, l_{54}^{(k)}$, and (6) not for $\alpha_3^{(k)}, \alpha_4^{(k)}, \alpha_5^{(k)}$. Hence, correspondence 5 and 6 shall be determined as "false outliers". It is also clear that a correspondence will have influence on its adjacent ones: an outlier will lead to the misjudgment of the next two correspondences, and only the pair of feature points whose two previous points are inliers can be determined correctly.

To lower the incidence of "false outliers" at greatest effort, the solution is to remove the outliers once they are found by evaluating the angle property. Examining all the correspondences from 1 to 7, the angle equation does not hold for point 4. Thus, point 4 should be considered as an outlier and be removed immediately from the set of matches. As a result of this operation, the magnitudes of the vector connecting 5 and its previous point and angle $\alpha_5^{(k)}$ should be recomputed. The procedure goes on until all the correspondences have been checked. In this way, the influence of the outlier 4 is eliminated. As an illustration, this operation is shown in Fig. 3.

Since each correspondence is checked in sequence and the correctness of the determination relies on how the previous points are judged, the quality of the starting point is particularly important. The solution, pursued here, is to look for the correspondence which fits both the angle and magnitude equation the best, rather than considering angle

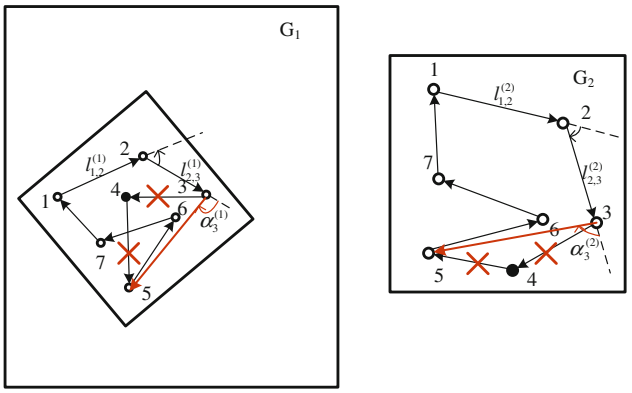

Fig. 3 The elimination of an outlier once it is found

equation solely. In this work, the correspondence with the highest quality is defined as

$$\begin{cases} \Delta\alpha_i < T_g \\ \arg\min_{f_i^{(k)} \in F_N^{(k)}} |\frac{s_i - s_{i-1}}{s_i}| \,, \quad s_i = \frac{l_{i,i-1}^{(1)}}{l_{i,i-1}^{(2)}} \,, \; s_{i-1} = \frac{l_{i-1,i-2}^{(1)}}{l_{i-1,i-2}^{(2)}} \,. \end{cases} \tag{11}$$

After examining angle differences, the magnitude ratios of corresponding vectors should be considered to make the overall procedure more convincing. However, the magnitude equation cannot be used directly because the scale factor is untold. It is assumed here that inliers contaminated by noises follow a normal distribution with unknown mean μ and standard deviation σ. To keep the inliers with higher accuracy, the outliers are defined as the points that deviate from the mean value more than one or two standard deviation. All the points in $I_{M'}^{(k)}$ are sorted either in ascending or descending order according to their related magnitude ratio and the value of the median point is considered to be $\hat{\mu}$ which is an approximation of μ. For a normal distribution, about 68% of values drawn from it are within one standard deviation σ away from the mean μ; about 95% of the values are within two standard deviations. So we select the matches that lie within $\sigma \sim 2\sigma$ from the mean as the inliers. Let P_l denotes the proportion, we get

$$\begin{aligned} I_M^k &= \left\{ f_i^k \in SI_{M'}^k | i \in [N_l + 1, M' - N_l], N_l \right. \\ &\left. = \left[\frac{(1 - P_l) \times M'}{2} \right] \right\} \end{aligned} \tag{12}$$

where $SI_{M'}^{(k)}$ is the sorted sequence of inliers.

In the implementation, there's no need to do the sorting additionally. An inlier can be inserted into the sequence of inliers with respect to the related magnitude ratio as soon as it is found satisfying the angle equation.

3.2 Removing repeated points

Whatever detectors and descriptors are used, repeated feature points (defined as the feature points in the same image location) frequently occur. When these feature points are stored one next to another, the vectors connecting them will be zero. Thus, the error of division by zero will occur when computing magnitude ratio. Moreover, the direction of the zero vectors hold whatever angle the image rotates and therefore the rotational invariance of corresponding angles will be destroyed. Hence, there is a necessity to remove repeated feature points before the estimation starts.

3.3 Enlarge the set of inliers

It is inevitable to yield false outliers if the starting point and the thresholds are not properly chosen. But from another perspective, these false outliers give us a second chance to enlarge the set of inliers. After the previous processing, a non-empty set $I_{M'}^{(k)}$ could be obtained and therefore can be used to recognize those false outliers.

The detailed solution is to connect the current correspondence in the set of outliers to two of the inliers, forming two vectors $v_1^{(k)}, v_2^{(k)} (k = 1, 2)$. Let the angle between $v_1^{(k)}$ and $v_2^{(k)}$ be $\alpha^{(k)}$, and the magnitude ratio of the corresponding vectors be s_1, s_2, then we get $s_1 = v_1^{(1)}/v_1^{(2)}, s_2 = v_2^{(1)}/v_2^{(2)}$ and the angle difference $\Delta\alpha$. If s_1, s_2 and $\Delta\alpha$ agree with the magnitude and angle equation, respectively, the current correspondence shall be re-determined as an inlier. Every correspondence in $O_{N-M'}^{(k)}$ should be examined following the above steps. In this solution, we choose the median two correspondences in the sorted sequence of $SI_{M'}^{(k)}$ to re-select inliers with relatively high accuracy.

3.4 Periodicity ambiguity

In computing the direction of vectors β and the angles between vectors α under polar coordinate system, there is a periodicity ambiguity. Two cases regarded in determining α are:

(a) if $\beta_1^{(k)} \le \beta_2^{(k)}$, then $\alpha^{(k)} = \beta_2^{(k)} - \beta_1^{(k)}$
(b) if $\beta_1^{(k)} > \beta_2^{(k)}$, then $\alpha^{(k)} = 360° - (\beta_1^{(k)} - \beta_2^{(k)})$, $(k = 1, 2)$

As to the difference of corresponding angles, another two situations should be considered, and they are illustrated in Fig. 4.

(a) if $\alpha^{(1)}$ and $\alpha^{(2)}$ locate on different sides of the zero direction, then $\Delta\alpha = 360° - |\alpha^{(2)} - \alpha^{(1)}|$
(b) if $\alpha^{(1)}, \alpha^{(2)}$ locate on the same side, then $\Delta\alpha = |\alpha^{(2)} - \alpha^{(1)}|$

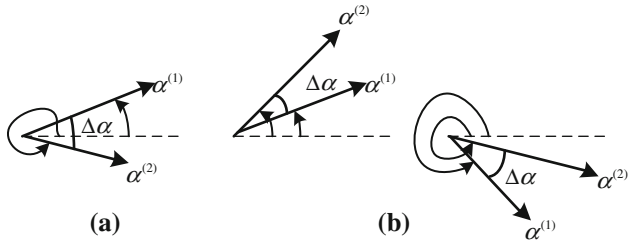

Fig. 4 Two situations of angle distributions

INPUT:
$F_{No}^{(k)}$ (k=1, 2): tentative correspondences drew from G_1, G_2
T_g, P_l : the threshold of angle difference and magnitude ratio

1. Initialization
$I_M^{(k)} \leftarrow \varnothing$, $O_{N-M}^{(k)} \leftarrow \varnothing$, $F_N^{(k)} \leftarrow \text{REmoveREpeat}(F_{No}^{(k)})$
2. Formation of a Sequence of Vectors
$\{f_N^{(k)}, f_1^{(k)}\} \leftarrow \text{SeqVector}(F_N^{(k)})$
3. Computation of Magnitude and Angle
$\{s_i, \Delta\alpha_i\} \leftarrow \text{MagAlg}(F_N^{(k)})$
4. Determination of the Starting Point $\rightarrow$Equation (11)
$f_{sp}^{(k)} \leftarrow \text{FindStart}(s_i, \Delta\alpha_i)$
5. Remove Outliers $\rightarrow$**Criterion 1** and Equation (12)
$\{I_{M'}^{(k)}, O_{N-M'}^{(k)}\} \leftarrow \text{RemoveOutliers}(f_{sp}^{(k)}, s_i, \Delta\alpha_i, T_g, P_l)$
6. Enlarge the Set of Inliers
$\{I_{M'}^{(k)}, O_{N-M'}^{(k)}\} \leftarrow \text{EnlargeInliers}(O_{N-M'}^{(k)}, T_g, P_l)$
7. Least Square Estimation on Set of Inliers
$\{s, \theta, t_x, t_y\} \leftarrow \text{LSE}(I_{M'}^{(k)})$
RETURN $\{s, \theta, t_x, t_y\}$

Fig. 5 Pseudo-code of DOR

3.5 DOR procedure

With the above analysis and the tentative correspondences obtained using SIFT, the overall procedure of DOR is presented in Fig. 5.

4 Experiments and analysis

The performance of our proposed method has been tested on both forged data sets and a widely-used database of real images [15]. All the experiments aim to evaluate the performance of removing outliers as well as the efficiency of this method.

4.1 Performance of removing outliers

4.1.1 Determination of thresholds

Thresholds T_g and P_l play important roles in the procedure of DOR. Not only do they prevent outliers from being calculated in the estimation stage, but also discard the inliers that deviate largely from the mean value. As stated in the previous section, P_l is less decisive than T_g, and thus is assigned a value between 0.8 and 0.95 to be consistent with normal distribution. T_g, which is the threshold of the angle difference between two related images, needs to be chosen carefully. On one hand, a relatively small value would generate a relatively precise result, but DOR is likely to fail if T_g is set too small. On the other hand, too large T_g would allow bad inliers or even outliers to be considered as inliers and thus to ruin the final estimation results.

To determine T_g, we have done multiple tests on forged data sets with Gaussian inliers and outliers known in prior. The total number of correspondences is 500, and the standard deviation of Gaussian noise added is varied. Figure 6 shows the experimental results in which T_g and standard deviation σ are varied.

Some of the points and lines are incomplete because outliers have been erroneously collected as a result of the large setting of T_g. Those incorrect predictions are displayed in Fig. 6. Concluded from the results, it is not applicable to make T_g larger than 3, and the setting is even more stringent in dealing with low fraction of inliers. Since the features of a real image are rarely located in absolute accuracy, lower threshold would collect less inliers. Just as the experimental results show, the performance of removing outliers are better when T_g is 1, rather than 0.5, for that less inliers are deemed as outliers incorrectly. To summarize, a proper choice for T_g is around 1, and T_g is kept in 1 in the following experiments.

4.1.2 Reliability test on artificial models

One important issue of our proposed method is to what extent the effect of the established sequence of matching points on its performance of removing outliers will be. If the inliers distribute too sparsely and the sequence is dominated by outliers, the probability of successfully detecting inliers would be greatly reduced. To examine its reliability, substantial tests on artificial models have been conducted. These models cover a wide range of parameters—the scale factor s ranges from 0.5 to 2, and the rotation angle θ from 0 to 180°. For each of them, multiple conditions were included with fractions of inliers from 20 to 95%, and all the correspondences were randomly sorted. The wide range of parameters is used to simulate various transform situations, and aim of multiple inlier ratios with random distributions is to test the influence of the sequence. The statistical results are listed in Table 1.

A total number of 504 models, 56 models for each fraction of inliers, have been used in this experiment. All results are correct under the condition that the percentage of inliers is above 40%. But when the inlier ratio is low enough, i.e. lower than 30%, DOR fails to discard all the outliers. There are 7 models out of 56 that could not be estimated accurately.

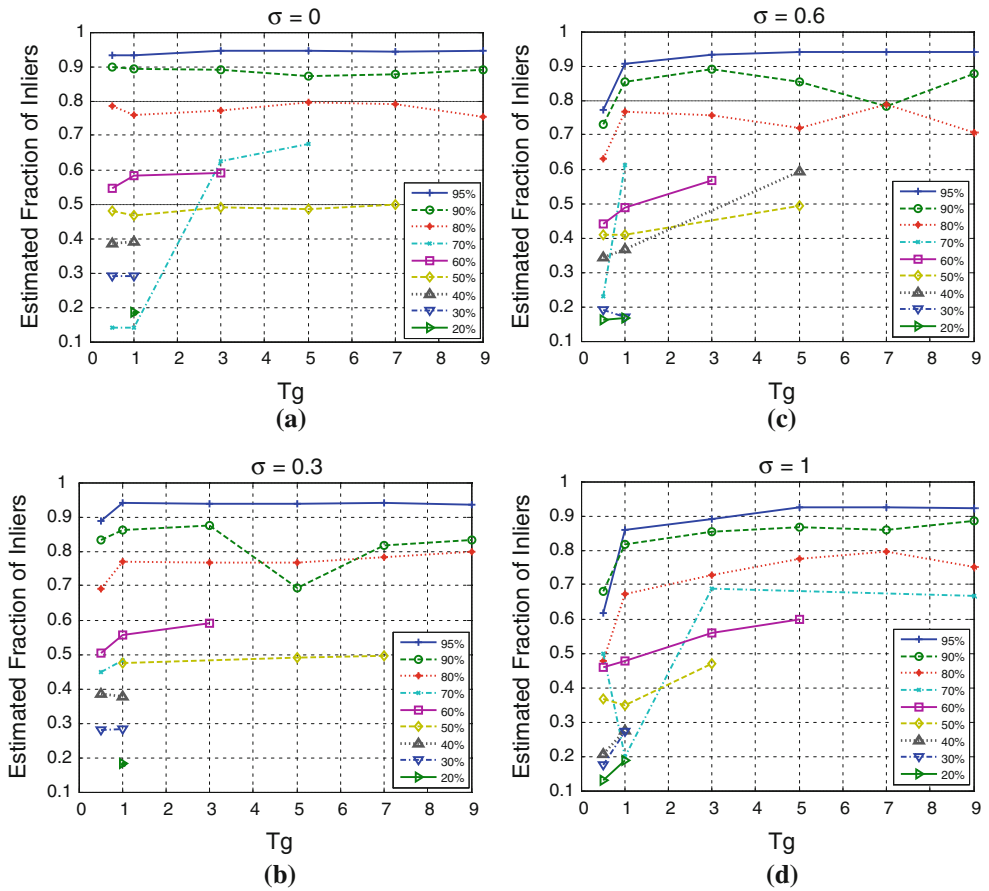

Fig. 6 The percentage of inliers found by DOR varies with T_g and σ. The legends indicate the true percentage. **a**–**d** Display the results when σ are set to be 0, 0.3, 0.6, 1, respectively

Table 1 The number of models whose outliers have been correctly eliminated under different percentages of inliers

Fraction of inliers (%)	20	30	40	50	60	70	80	90	95
Accuracy rate	24/56	49/56	56/56	56/56	56/56	56/56	56/56	56/56	56/56

And the accuracy rate drops as the inlier ratio decline. The results reveal, in some degree, that inliers are scattered by a large amount of outliers, which makes it harder to establish a sequence containing at least three successive inliers. Therefore, DOR had poor performances under these conditions. However, randomly distributed inliers and outliers are somewhat a worse case since it assumes that there is no particular structure and storing order of the tentative correspondences while most of the natural features are actually grouped. Hence, **t**he results would be a slightly better in application to real images, which could also be spotted in Table 2 in Sect. 4.2. As a conclusion, DOR has a high reliability when the inlier ratio is above 40%. Moreover, analyzing the results in Table 2, it is not difficult to find that the fraction of inliers falls between 40 and 80% in most real image registrations, which promises good performances of DOR.

4.1.3 Inspections on selected results

Among the tests running on artificial models, one of the intuitive results after removing outliers is presented in Fig. 7. As observed, none of the outliers is calculated and the inliers contaminated seriously by the Gaussian noise, are also eliminated.

Examinations have also been conducted on real images. Real image "Topview" (shown in Fig. 8) is chosen to illustrate the performance of removing outliers for its simple distributions of matched feature points. We have applied SIFT-based method [3] in the stage of features extraction and matching. Different fractions of inliers have been obtained by setting matching threshold to various values. Three of the results are presented in Fig. 9.

In case (a), 3 out of 32 points are outliers intuitively, that is, the proportion of inliers is around 91.6%. 25 inliers are

Table 2 Time consumptions on real image pairs

Image pairs	Number of features in two images	Number of matches	Fraction of inliers (%)	Time consumption (ms)			
				RANSAC	MSAC	PROSAC	DOR
"ASTRIX"	(762, 1,301)	254	45.3	3.0642	3.446	3.39	0.966
"BELLEDONNE"	(2,008, 3,488)	177	73.4	1.0682	2.468	1.064	0.725
"BOAT"	(1,979, 2,500)	283	47.7	3.2004	7.3594	3.585	1.107
"EAST PARK"	(1,411, 1,519)	296	29.7	8.4994	13.8408	8.505	1.214
"EAST SOUTH"	(3,043, 2,912)	537	75.4	3.7024	7.5392	3.419	2.276
"INRIA"	(1,316, 1,440)	288	56.2	2.6008	2.6186	2.467	1.157
"INRIA MODEL"	(2,394, 1,059)	115	30.4	2.8868	4.2108	3.36	0.479
"LAPTOP"	(2,192, 2,713)	672	73.4	4.5376	6.1096	4.559	2.866
"MARS"	(3,164, 2,715)	1,525	94.2	8.3088	8.2886	8.147	8.803
"MONET"	(1,290, 1,224)	538	84.8	3.1368	2.8944	2.261	2.242
"NewYork"	(2,273, 3,103)	1,064	81.3	6.3844	6.5672	6.315	4.793
"RESID"	(1,339, 1,507)	404	49	4.2442	4.809	4.175	1.569
"UBC"	(1,741, 1,566)	445	73.5	4.249	4.0388	4.180	1.918

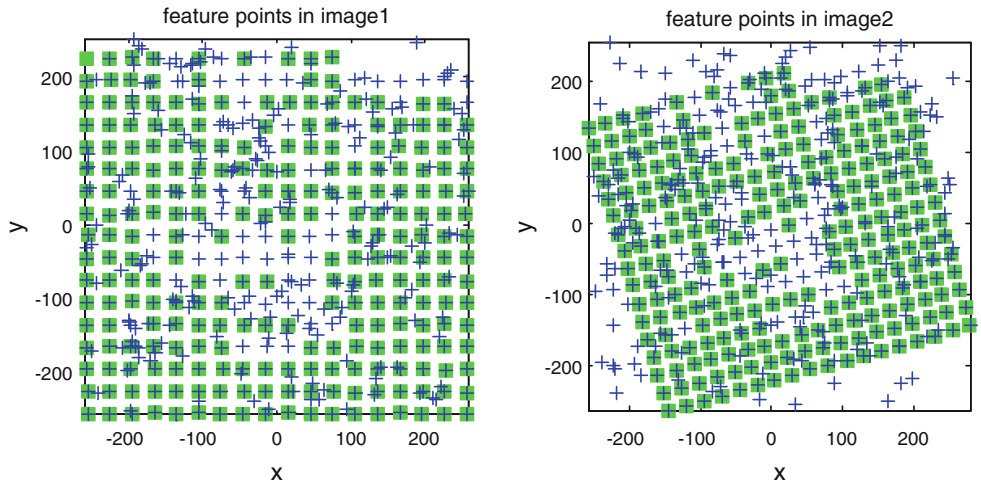

Fig. 7 Forged data set with 500 matches, and 60% of Gaussian inliers. The standard deviation is 0.6. Data painted *green* are inliers found by DOR (color figure online)

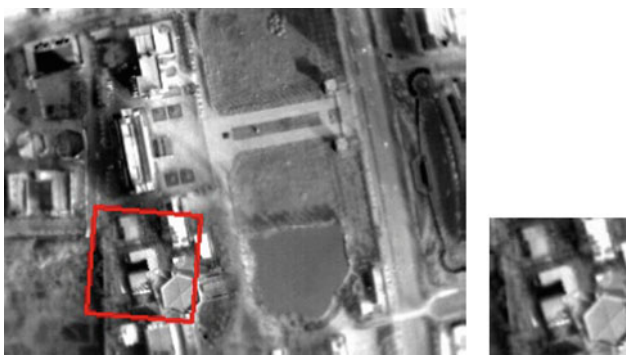

Fig. 8 Real image pair "Topview"

found finally and 3 outliers with large position deviations (point 3, 5 and 14) are detected and removed. The rest four outliers cannot be noticed without accurate computation but are beyond the error limit of inliers. (b) shows 31 inliers and 18 outliers in images 1 and 2 (fraction of inliers is 63.2%), and we find 26 inliers with highest positional precision. In (c), there are 76 tentative correspondences with only 28 inliers (the percentage of inliers is 38.4%), 27 inliers with high quality are collected to estimate model as a result.

4.2 Efficiency and accuracy evaluation

According to the overall procedure, DOR is largely affected by the total number of correspondences and a low percentage of inliers will save time in sorting. That is to say, DOR can run fast even if the proportion of inliers is small. In this case, however, RANSAC would cost a great number of iterations before finally converging into an optimal set of inliers and therefore it is time-consuming in an average case. In this

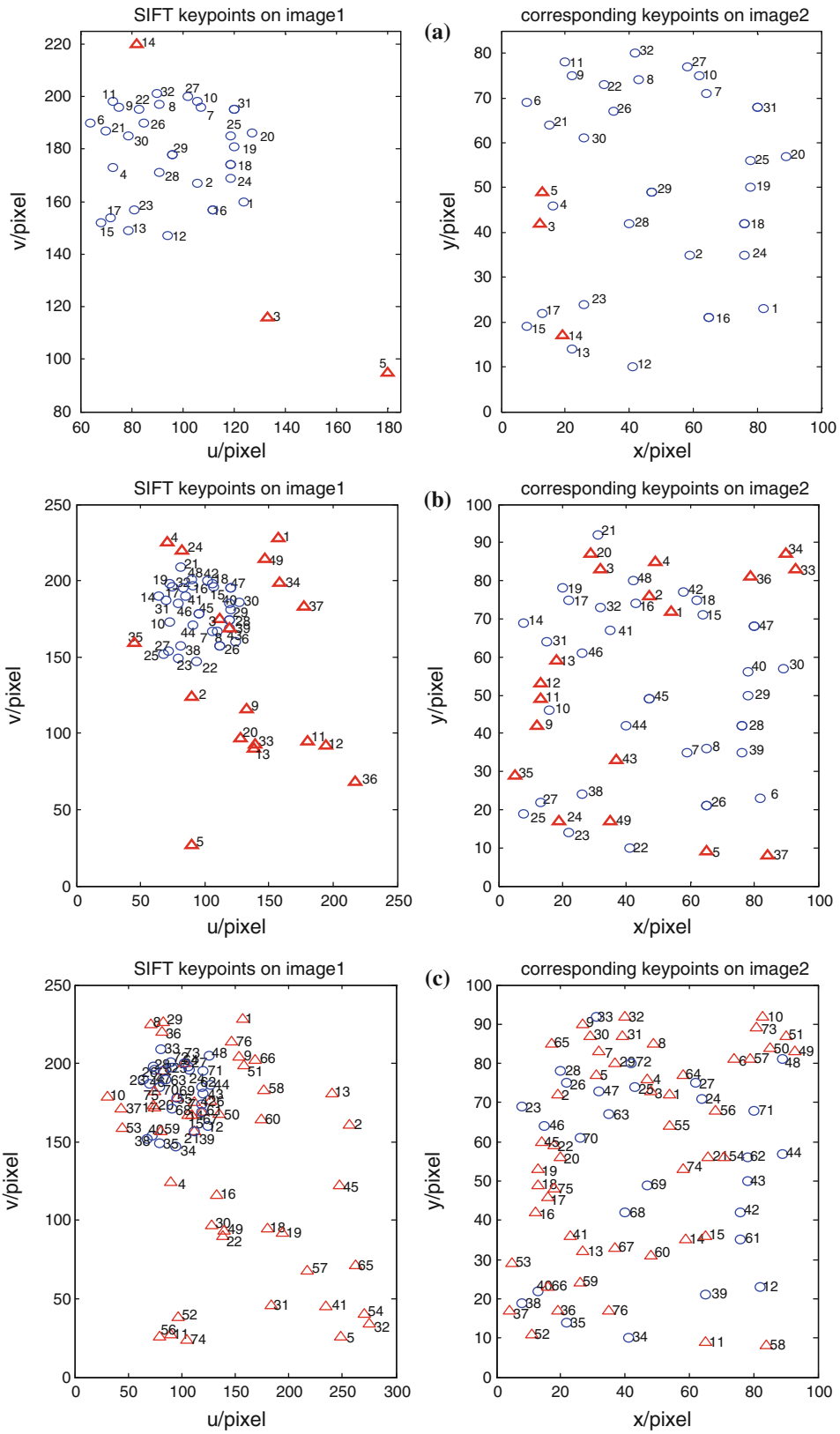

Fig. 9 Feature points are presented in the plane of image 1 and 2 correspondingly. The numbers next to the points are assigned according to storing sequence. *Circles* indicates the inlier and *triangles* indicates the outlier. **a** shows the distribution of tentative correspondences when the matching threshold is set to be 0.8. **b** and **c** are the cases when the threshold is 0.9 and 0.95

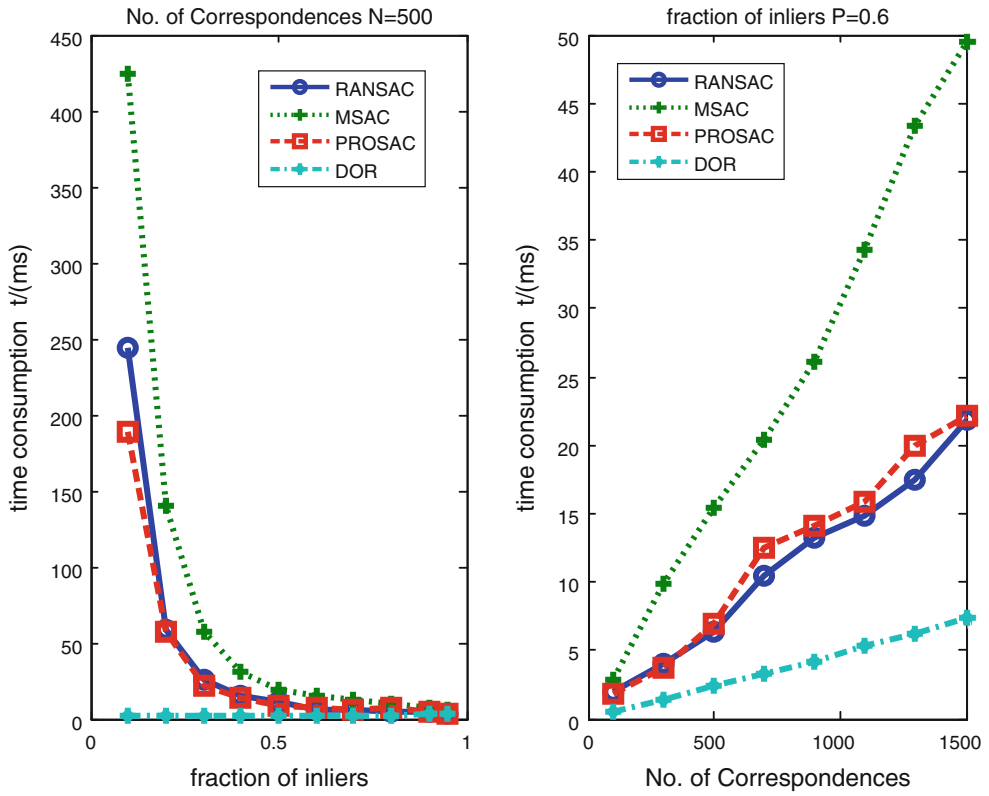

Fig. 10 The *left side* shows the result of the test on a set of 500 correspondences with percentage of inliers from 20 to 95%. The *right side* is the curves of the time consumption in varying the cardinality of the data sets

section, we provide a discussion on the complexity of DOR and the related experiment.

The complexity of connecting correspondences, computing the magnitude ratios and angle differences are all $O(N)$, where N denotes the total number of input matches. In the worst case of finding staring point $f_{sp}^{(k)}$, all the matches should be traversed and therefore the complexity is $O(N)$. Insertion sort are used in sorting the inliers in $I_{M'}^{(k)}$ which is obtained by single angle examination. Hence the complexity in this part is $O(M'^2)$, where M' denotes the number of inliers. Based on the above analysis, the complexity of the overall procedure of removing outliers can approximate to $O(N)$ when the fraction of inliers is much below 50%. If most of the correspondences are inliers, the complexity can be considered to be $O(N^2)$. Therefore, DOR is more efficient in dealing with conditions of high percentage of outliers.

To verify it in practice, we have run DOR along with RANSAC and some of its variants on a standard PC with an Intel Pentium(R) 2.4 GHz processor. Likewise, tests have been conducted on forged data sets and real images. Due to the randomness of RANSAC, the computational time has been obtained by taking the average of multiple running results.

Figure 10 is the visualized result of the time used in estimating the transform models of different artificial data. It shows that DOR obviously outperforms several other widely-used hypothesize-and-test methods in the aspect of efficiency. Moreover, the time consumption of RANSAC and its variants surge up exponentially as the inliers decrease, while DOR maintains a relatively low level. This little contribution of the variation of inliers to the total time consumption is consistent with its principle. Although the computational time of all the methods compared here increase as the number of correspondences grows, DOR is still the most advantageous since the computational load of it grows slowest among all the algorithms. As to PROSAC, which is accepted as a rather fast method, its efficiency declines to the degree that is equivalent to RANSAC in the simple situation of estimating similarity transform. In the stage of hypothesis, only two correspondences are required to be collected, which guarantees the performance of randomly selecting approach and thus makes the efficiencies of PROSAC and RANSAC almost equal. Nevertheless, PROSAC regains its advantage when the fraction of inliers is low.

Efficiency tests have also been carried out on extensive real images, and the statistical results are listed in Table 2. For all the image pairs, the feature correspondences have been extracted by SIFT with 0.8 as the matching threshold. As observed, DOR still exceeds other methods in performance comparison. The advantage is especially evident in the case with low fraction of inliers.

Table 3 Estimated model parameters and ground truths

Image pairs	Estimation results [listed in the order (s, θ/deg, t_x/pixel, t_y/pixel)]				
	Ground truth	RANSAC	MSAC	PROSAC	DOR
"ASTRIX"	0.5546	0.5805	0.5805	0.5805	0.5806
	0	0.01	−0.005	−0.005	−0.11
	109.960520	106.5549	106.5505	106.5505	106.5635
	81.051620	79.9339	79.8841	79.8841	79.5947
"BELLEDONNE"	0.56397	0.5632	0.5634	0.5634	0.5628
	0	−0.4987	−0.4919	−0.4839	−0.5087
	88.1718	88.4298	88.4284	88.4272	88.5996
	138.829	139.1187	139.0671	139.0967	139.2393
"BOAT"	0.62517	0.6234	0.6228	0.6228	0.6228
	51.0914	−50.9949	−51.0577	−51.0538	−51.0538
	248.189	247.5448	247.8163	247.8087	247.8087
	0.25124	0.6825	0.7237	0.7570	0.7570
"EAST PARK"	0.6237	0.6237	0.6232	0.6232	0.6233
	54.9736	54.9518	54.9165	54.9165	54.9786
	56.47059	56.6144	56.8030	56.8030	56.7094
	241.71765	241.6540	241.4560	241.4560	241.7729
"EAST SOUTH"	0.7089	0.7096	0.7099	0.7099	0.7095
	35.7829	35.6200	35.6010	35.6048	35.6457
	−1.5108	−1.9760	−2.0791	−2.0666	−2.0738
	161.6029	161.7160	161.6509	161.6814	161.8402
"LAPTOP"	0.7008	0.7074	0.7070	0.7070	0.7070
	29.7451	−29.8996	−29.9065	−29.9065	−29.9206
	290.172	284.7572	284.8176	284.8176	284.8456
	−30.7085	−33.2641	−33.0967	−33.0967	−33.1800
"MARS"	1.0061	1.005	1.005	1.005	1.005
	31.9287	31.8093	31.8087	31.8087	31.8164
	−93.8337	−94.1703	−94.1584	−94.1584	−94.0599
	176.4983	175.4372	175.4363	175.4363	175.4792
"MONET"	1.0053	1.0039	1.0039	1.0039	1.0047
	23.3595	23.2897	23.2905	23.2905	23.2869
	−77.9914	−78.1575	−78.1658	−78.1658	−78.4059
	124.5620	124.0893	124.0831	124.0831	124.8241
"New York"	0.9952	0.9995	0.9995	0.9995	0.9991
	49.5767	49.9458	49.9477	49.9477	49.9424
	−135.7898	−134.6786	−134.6787	−134.6787	−134.5730
	256.8285	257.8925	257.8948	257.8948	257.8457
"RESID"	0.8693	0.7416	0.7476	0.7476	0.7425
	−77.6837	−75.6289	−75.6671	−75.6671	−75.6870
	358.8141	355.4347	355.6147	355.6147	355.7702
	59.1548	58.4654	58.4987	58.4987	58.3729

Meanwhile, the parameters of similarity transform estimated as well as ground truths have been recorded in Table 3 to evaluate the estimation accuracy. The results of all the three types of hypothesis-and-test methods and DOR are quite close, albeit a certain amount of errors occurs when compared to true value. These errors are mainly introduced in the stage of feature extraction and localization, but they also affect the final estimation results. Therefore, there is a little difference in the aspect of accuracy among these estimation methods, which is guaranteed by the accuracy of LSE.

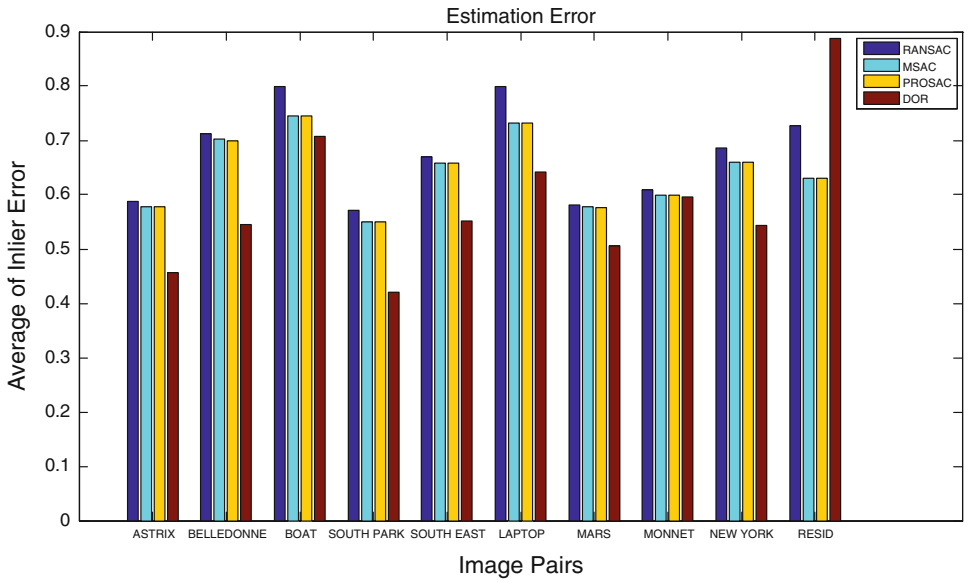

Fig. 11 AIE results of multiple image pairs

To evaluate the accuracy more precisely, the average of inlier error (AIE) [16] is taken into account. The AIE calculated here is

$$\mathrm{AIE}(M;\, D_{in}) = \frac{1}{N(D_{in})} \sum_{d \in D_{in}} \sqrt{\mathrm{Err}(d_i;\, M)^2} \tag{13}$$

where M is the estimated similarity transform model, D_{in} is the set of inliers, d_i is an inlier belong to D_{in}, N is the number of inliers.

AIEs of the four methods are collected and compared in Fig. 11. For all the image pairs evaluated in this experiment, except "RESID", DOR presents slightly better results. However, they are all in the same order of magnitude.

The experiments in this section show that DOR saves computational time greatly due to its non-iterative scheme. However, this scheme is limited to similarity transform.

5 Conclusions

In this work, a non-iterative robust method—direct outliers remove (DOR) is introduced to estimate similarity transform efficiently. Inliers are achieved by traversing all the correspondences once instead of iterative testing as in RANSAC. After connecting correspondence and creating vectors in the image plane, DOR then finds out the outliers and inliers according to the criterion based on the similarity invariance—magnitude ratio of corresponding vectors and the angle between two vectors.

In comparison to RANSAC and its variants, DOR is more efficient in computational time due to the non-iterative approach while preserving good accuracy, especially in the case that inliers only take up a low fraction.

Acknowledgments This work is supported by Graduate Student Innovation Fund from Advanced Aerospace Vehicle Research Base of Beihang University.

References

1. Mikolajczyk, K., Schmid, C.: Scale and affine invariant interest point detectors. IJCV **60**(1), 63–86 (2004)
2. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proceedings of BMVC, pp. 384–393 (2002)
3. Lowe David, G.: Distinctive image features from scale-invariant keypoints. IJCV **60**(2), 91–110 (2004)
4. Fuchao, W.: Mathematical Methods in Computer Vision. Scientific Press, Beijing (2008)
5. Fischler, M., Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. CACM **24**(6), 381–395 (1981)
6. Torr, P.H.S., Zisserman, A.: MLESAC: a new robust estimator with application to estimating image geometry. CVIU **78**, 138–156 (2000)
7. Tordoff, B., Murray, D.: Guided sampling and consensus for motion estimation. In: Proceedings of the 7th ECCV, vol. 1, pp. 82–96. Springer, Berlin (2002)
8. Chum, O., Matas, J.: Matching with PROSAC—progressive sample consensus. In: Conference of CVPR (2005)
9. Myatt, D.R., Torr, P.H.S, Nasuto, S.J., Bishop, J.M., Craddock, R.: NAPSAC: High noise, high dimensional robust estimation-it's in the bag. In: Proceedings of the 13th BMVC, pp. 485–467 (2002)
10. Rodehorst, V., Hellwich, O.: Genetic Algorithm SAmple Consensus (GASAC)—a parallel strategy for robust parameter estimation. In: Proceedings of the CPVR Workshop (CVPRW) (2006)
11. Kai, N., Hailin, J., Dellaert, F.: GroupSAC: efficient consensus in the presence of groupings. In: Proceedings of ICCV, pp. 2193–2200 (2009)
12. Zitova, B., Flusser, J.: Image registration method: a survey. Image Vis. Comput. **21**, 977–1000 (2003)
13. Reddy, B.S., Chatterji, B.N.: An FFT-based technique for translation, rotation and scale-invariant image registration. IEEE Trans. Image Process. **5**(8), 1266–1271 (1996)

14. Tzimiropoulos, G., Argyriou, V., Zafeiriou, S., Stathaki, T.: Robust FFT-based scale-invariant image registration with image gradients. IEEE Trans. Pattern Anal. Mach. Intell. **32**(10), 1899–1906 (2010)
15. Image Database. http://lear.inrialpes.fr/people/mikolajczyk/ (2011)
16. Choi, S., Kim, J.-H.: Robust regression to varying data distribution and its application to landmark-based localization. In: Proceedings of SMC (2008)

Author Biographies

Li Yinan is a Masters student in the School of Automation Science and Electrical Engineering, Beihang University. She received the B.E. degree from Beihang University in 2009. Her research interests include scene matching navigation, pattern recognition, and integrated vehicle management system.

Yang Lingyu received his Ph.D. in Navigation, Guidance and Control from Beihang University, P. R. China, 2007. He is currently an assistant instructor in the School of Automation Science and Electrical Engineering in Beihang University. His main field of research is control allocation and flight control of aerospace vehicles. His research interests include integrated control system, modeling, nonlinear control of aerial vehicles and scene matching navigation, etc.

Shen Gongzhang was born in 1945 in China. Since the 1990s, he has been a professor at the School of Automation Science and Electrical Engineering, Beihang University, P. R. China. His research interests include integrated control technology, integrated vehicle management system, nonlinear modeling, and control, etc.