IMAGE MATCHING FOR REPETITIVE PATTERNS

Lu Tian, Xianwei Xu, and Jie Zhou

Tsinghua University
Department of Automation
Beijing, China

ABSTRACT

Repetitive patterns exist widely in real world images, and matching images with plenty of repetitive patterns remains a challenging task. We present in this paper a novel feature matching algorithm of images with notably repetitive patterns, in which a reliable initial correspondence set is established, purified and propagated using a voting strategy, incorporating a local geometrical constraint. Experiments have demonstrated that the proposed algorithm outperforms state-of-the-art feature matching methods on matching images with plenty of repetitive patterns.

Index Terms— Feature correspondence, repetitive pattern

1. INTRODUCTION

Feature point matching is a core problem in many computer vision tasks, such as image registration [1], 3D reconstruction [2], structure from motion [3] and object recognition [4]. A rich and reliable feature matching result greatly contributes to improve the performance of these tasks. Many state-of-the-art matching methods are based on local descriptors. However, repetitive patterns in images, which widely exist in natural and artificial scenes, make the local descriptors ambiguous and therefore bring difficulties to feature matching methods.

One common approach to solve this problem is to reject the match that is probable to be disturbed by ambiguous descriptors [4]. However, this approach may reject potential correct matches in the regions of repetitive patterns, leading to a low number of total matches. Another approach is to take a larger candidate matches set into consideration and refine the match set by various geometrical considerations or by the optimization of the graph matching model [5, 6, 7]. However, the cluster of incorrect matches which are mutually geometrically consensus cannot be rejected by these methods because of the locality of geometric constraints.

Some researches have been taken on repeated pattern detection and matching. Ha *et al.*proposed to detect features of repetitive patterns by clustering, and validate the homography computed from matches of salient features by repetitive patterns in RANSAC iteration [8]. Fan *et al.*proposed to regard a pair of keypoints as a unit, describe them, match them and parse the matching result into point correspondence under the low distortion constraint [9]. However, the low distortion constraint are sensitive to simple geometric transformations such as scale changing or general affine transformation, therefore the constraint will mislead the matching in the situation of these geometric transformation.

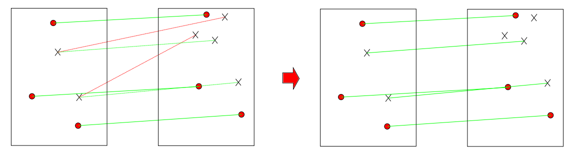


Fig. 1. An illustration of the proposed method. The left image shows the initial matching result. The matching result of non-ambiguous keypoints (shown as red circles) is determined (shown as green solid lines), however, the ambiguous keypoints (shown as black crosses) may be matched to several candidates rather than a single point (shown as dashed lines), in which one of them are correct (shown as gree dashed lines). The right image shows the final matching result, in which all ambiguities are eliminated by correspondence purification and propagation

In this paper, we propose a novel feature matching framework mainly focused on matching images with repeated patterns. In this framework, features outside the repetitive regions are firstly selected and matched, then the matching results are propagated to other features (see Fig. 1). The advantages of our algorithm are summarized as follows. Firstly, the discrimination of ambiguous and not ambiguous (salient) feature points makes the initial match richer and robuster. Secondly, incorporating appearance into matching propagation can decrease the local indistinctness of geometrical information.

The remainder of this paper is organized as follows. The problem is defined in Section 2. Then the proposed method is elaborated in Section 3. Experimentations are reported in Section 4, and Section 5 concludes this paper.

2. PROBLEM DEFINITION

In this paper, we suppose that the features are detected along with a elliptical region and a dominant orientation. Many state-of-art feature detectors such as Hessian-Affine extractors [10] or MSER extractors [11] are of this type.

We denote the keypoints extracted from two images as $V^P = \{P_1, P_2, ..., P_n\}$ and $V^Q = \{Q_1, Q_2, ..., Q_m\}$. For each feature $v_i \in V^P \cup V^Q$, we denote its region as S_i and descriptor as d_i. Our goal is to find as many as possible correct matches in $V^P \times V^Q$.

2.1. Geometry of affine feature regions

Because every feature region is elliptical, its shape and dominant orientation can be represented by a 3×3 affine transformation matrix T_i, which transform the ellipse into the normalized patch(a round patch with unit radius and dominant orientation at y-axis). Generally the T_i has the following form

$$T_i = \left[\begin{array}{cc} A_i & x_i \\ 0 & 1 \end{array} \right], \tag{1}$$

where x_i is the coordinate of the feature region center in the image and A_i is the 2×2 matrix which represent the transformation from the ellipse feature region (centered at the origin point) to the normalized patch.

2.2. Geometrical compatibility valuation

In the proposed approach, measuring the geometrical compatibility of a pair of matches is usually needed. In practise we use the Standard Transfer Error (STE) [5] to measure it. Let $P_i, P_j \in V^P$ and $Q_i, Q_j \in V^Q$, the STE are defined as the following:

$$\begin{aligned} STE((P_i, Q_i), (P_j, Q_j)) = \frac{1}{4}(d_{P_iQ_i|P_jQ_j} + d_{P_jQ_j|P_iQ_i} \\ +d_{Q_iP_i|Q_jP_j} + d_{Q_jP_j|Q_iP_i}), \end{aligned} \tag{2}$$

where

$$d_{ab|cd} = \|x_d - \rho(T_b^{-1} T_a \begin{bmatrix} x_c \\ 1 \end{bmatrix})\|, \tag{3}$$

where ρ means the conversion from the homogeneous coordinates to 2-D coordinates.

2.3. Correspondence purification

Formally, we denote the match set before purification as $M = \{(P_1, Q_{\pi(1)}), (P_2, Q_{\pi(2)}), \ldots, (P_k, Q_{\pi(k)})\}$. We score each match by its STE with neighbour matches and eliminate the low-score matches.

The pseudo code of this part is shown in Alg. 1.

Algorithm 1 Correspondence Purification

Require: A set of correspondence $M = \{(P_{i_1}, Q_{\pi(i_1)}), (P_{i_2}, Q_{\pi(i_2)}), \ldots, (P_{i_k}, Q_{\pi(i_k)})\}$

Ensure: C, a subset of M which includes correct correspondences.

```
1: for all (P_i, Q_π(i)) ∈ M do
2:   for all P_j ∈ k-NN(P_i; {P_i1, P_i2, ..., P_ik}) do
3:     E_ij ← STE((P_i, Q_π(i)), (P_j, Q_π(j)))
4:   end for
5:   Select A such that it contains [k/2] smallest elements
     in {E_ij||1 ≤ j ≤ k}
6:   Score(i):=mean(A);
7: end for
8: C ← Select the best 70% matches
```

3. PROPOSED METHOD

The proposed method is divided into the following steps: initial matching, correspondence purification and correspondence propagation. These steps are elaborated in the following subsections. The flow chart of the algorithm is shown in Fig. 3.

3.1. Initial matching

Firstly we partition the feature point set into two sets: We define the distance between the two descriptor vectors d_i and d_j as $\theta_{ij} = \arccos d_i \cdot d_j$. For each feature $P_i \in V^P$ We find the nearest descriptor in V^Q which has the least angle with P_i, denoted as $n_i = \arg\min_j \theta_{ij}$.

All the feature descriptors whose distance with v_i is less than $\rho\theta_{i,n_i}$ are selected as its candidate matching set.

$$A_i = \{v_j : d_i \cdot d_j < r \cos \theta_{i,n_i}\}, \tag{4}$$

where r is an preset ratio.

We eliminate the keypoints whose nearest neighbor distance is too large, and the ambiguity of a feature point is determined by the size of its candidate set. The salient feature point set is defined as

$$S^P = \{v_i || A_i| = 1 and \theta_{i,n_i} < t_1\}, \tag{5}$$

The initial matches is set as $M = \{(P_i, Q_{n_i}) | v_i \in S^P\}$

3.2. Correspondence propagation

After initial matching, we get the initial match set $M = \{(P_{i_1}, Q_{\pi(i_1)}), (P_{i_2}, Q_{\pi(i_2)}), \ldots, (P_{i_k}, Q_{\pi(i_k)})\}$ associating with the keypoints in the two salient keypoint sets. We divide the whole image into several Voronoi regions and propagate the match to the keypoints in the same region using the same reference matches. To fully present the algorithm, we present the pseudo code of this part in Algorithm 2.

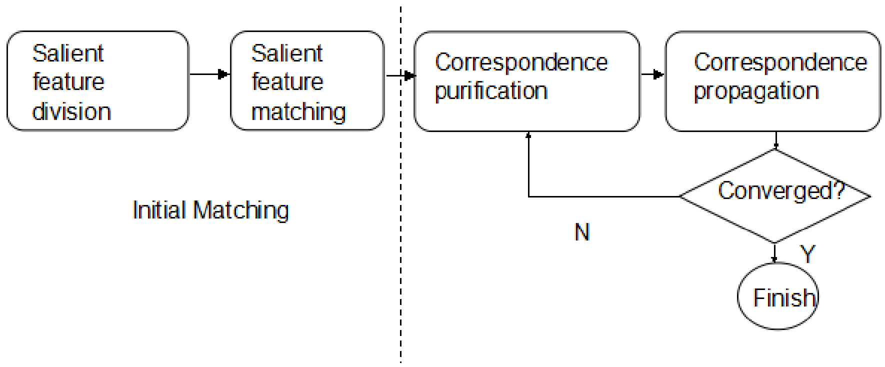
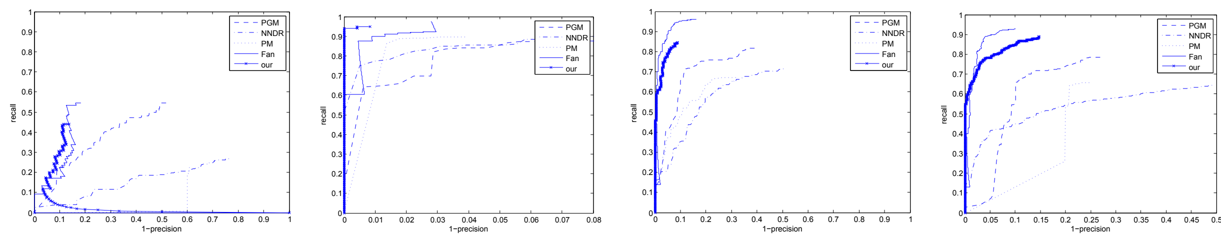


Fig. 2. The flow chart



(a) (b) (c) (d)

Fig. 3. Precision recall curve in CAS dataset

In the algorithm,

$$v(Q_i) = d_{P_j} \cdot d_{Q_i} + \overline{exp(-STE((P_{ni}, Q_{\pi(ni)}), (P_j, Q_i))/\sigma)}, \tag{6}$$

in which the average is performed on all possible assignment of P_{ni} which satisfies the condition that $Q_i \in kNN(Q_j i; V_N^Q)$.

4. EXPERIMENTS

In our experiments, Hessian-Affine feature regions with SIFT descriptor is used for keypoint extraction and description. We have systematically compared our method with several state-of-the-art feature matching methods, including Progressive Graph Matching [5], NNDR [4], Fan's method [9], Pairwise Matching [7].

4.1. Evaluation criterion

To judge the correctness of each match, we use the criterion proposed by Mikolajczyk and Schmid [12]. The definition of $recall$ and $precision$ are as follows:

$$\begin{aligned} recall &= \frac{\#correct\ matches}{\#correspondences}, \\ 1 - precision &= \frac{\#false\ matches}{\#allmatches}, \end{aligned} \tag{7}$$

where $\#$ means the cardinal of a set, and $correspondences$ is the total number of ground truth matches between two images.

4.2. Results on planar scenes

We test all methods on the 4 image pairs referenced in the article [9]. This database is constructed by Chinese Academy of Science. In each pair, the "Left" image can be transformed into the "Right" one under a homography. The precision-recall curve of all methods on the 4 image pairs are shown in Fig. 3. From the precision-recall curves it can be seen that our method performs well among all method taken into consideration, especially in the high-precision range.

4.3. Results on real building image pairs

We select some image pairs of real buildings in the ZuBuD dataset to test the performance of our method. Fig. 4 shows the matching result of several methods in comparison in one of the image pairs. In order to evaluate the performance of all methods, for each method we counted the number of correct matches in 150 best-scored matches. The ratio of correct matches are shown under the figures. It can be observed that our method gives more accurate correspondences.

(a)

(b)

(c)

(d)

(e)

Fig. 4. Feature matching results of different algorithm on a real building image pair. (a)Our method (140/150); (b)Fan's method (132/150); (c)PGM (144/150); (d)Pairwise Matching (67/150); (e)NNDR (33/150)

Algorithm 2 Correspondence Propagation

Require: feature sets of two images, $V^P = \{P_1, P_2, ..., P_n\}$, $V^Q = \{Q_1, Q_2, ..., Q_m\}$ $M = \{(P_{i_1}, Q_{\pi(i_1)}), (P_{i_2}, Q_{\pi(i_2)}), \ldots, (P_{i_k}, Q_{\pi(i_k)})\}$

Ensure: C, A set of feature correspondence including some features that are in V^P but not in M

$C \leftarrow \emptyset$

$V_M^P \leftarrow \{P_{i_1}, P_{i_2}, \ldots, P_{i_k}\}, V_M^Q \leftarrow \{Q_{\pi(i_1)}, Q_{\pi(i_2)}, \ldots, Q_{\pi(i_k)}\}$

$V_N^P \leftarrow V^P - V_M^P, V_N^Q \leftarrow V^Q - V_M^Q$

$G \leftarrow$ Delaunay triangle net generated by V_M^P

for all $P_j \in V_N^P$ **do**

  $P_n = \arg\min_{P_i \in V_M^P} \|P_i - P_j\|^2$;

  $DN(P_n) \leftarrow$ neighbors of P_n in graph G

  for all $P_{ni} \in DN(P_n)$ **do**

    $H_n i = T_{Q_{\pi(ni)}}^{-1} T_{P_{ni}}$;

    $Q_j i = H_n i(P_j)$

  end for

  $N^Q = \cup kNN(Q_j i; V_N^Q)$;

  for all $Q_i \in N^Q$ **do**

    Compute $v(Q_i)$ by Eq. 6

  end for

  $Q_{im} = \arg\max_{Q_i \in N^Q}$

  if $v(Q_{im}) > th$ **then**

    $C \leftarrow C \cup \{(P_j, Q_{jm})\}$

  end if

end for

---

5. CONCLUSION

In feature point matching, repetitive patterns often disturb the matching result. A novel feature matching method is proposed for matching images with plenty of repetitive patterns. We start by establishing reliable correspondences based on salient features. Then the correspondence set is iteratively refined and enriched to increase its precision. Experiments demonstrated that the proposed method outperforms several state-of-the-art methods.

6. ACKNOWLEDGEMENT

This work is supported by the National Natural Science Foundation of China under Grants 61225008, 61373074 and 61373090, the National Basic Research Program of China under Grant 2014CB349304, the Ministry of Education of China under Grant 20120002110033, and the Tsinghua University Initiative Scientific Research Program.

7. REFERENCES

[1] Barbara Zitova and Jan Flusser, "Image registration methods: a survey," *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[2] Yasutaka Furukawa and Jean Ponce, "Accurate, dense, and robust multiview stereopsis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 8, pp. 1362–1376, 2010.

[3] Noah Snavely, Steven M Seitz, and Richard Szeliski, "Photo tourism: exploring photo collections in 3d," *ACM transactions on graphics (TOG)*, vol. 25, no. 3, pp. 835–846, 2006.

[4] David G Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[5] Minsu Cho and Kyoung Mu Lee, "Progressive graph matching: Making a move of graphs via probabilistic voting," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 398–405.

[6] Mohammad Izadi and Parvaneh Saeedi, "Robust weighted graph transformation matching for rigid and nonrigid image registration," *Image Processing, IEEE Transactions on*, vol. 21, no. 10, pp. 4369–4382, 2012.

[7] Ee Sin Ng and Nick G Kingsbury, "Robust pairwise matching of interest points with complex wavelets," *Image Processing, IEEE Transactions on*, vol. 21, no. 8, pp. 3429–3442, 2012.

[8] Seong Jong Ha, Sang Hwa Lee, and Nam Ik Cho, "Discrimination and description of repetitive patterns for enhancing the performance of feature-based recognition," *Image and Vision Computing*, vol. 30, no. 11, pp. 817–828, 2012.

[9] Bin Fan, Fuchao Wu, and Zhanyi Hu, "Towards reliable matching of images containing repetitive patterns," *Pattern Recognition Letters*, vol. 32, no. 14, pp. 1851–1859, 2011.

[10] Tinne Tuytelaars and Krystian Mikolajczyk, "Local invariant feature detectors: a survey," *Foundations and Trends® in Computer Graphics and Vision*, vol. 3, no. 3, pp. 177–280, 2008.

[11] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.

[12] Krystian Mikolajczyk and Cordelia Schmid, "A performance evaluation of local descriptors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pp. 1615–1630, 2005.