

On the Matching of Images Containing Repetitive Patterns by Pairs of Interest Points

Bin Fan, Fuchao Wu and Zhanyi Hu

Abstract—Image matching is a fundamental task for computer vision. However, due to the local ambiguities induced by repetitive patterns in most of man-made objects or scenes, matching images of repetitive patterns is still a challenging problem even the viewpoint changes of them are not very large.

This paper is primarily focused on the problem of matching images containing repetitive patterns and proposes a novel method based on pairs of interest points to solve the problem. It starts from matching pairs of interest points and then obtaining point correspondences from the matched point-pairs based on the low distortion constraint, which is meant that the distortions of point groups should be small across images. By combining pairs of interest points, local ambiguities induced by repetitive patterns can be reduced to some extent since information in a much larger region is used. Moreover, owing to our newly defined compatibility measure between one correspondence and a set of point correspondences, the obtained point correspondences are very reliable. Experimental results have demonstrated the effectiveness of the proposed method as well as its superiority to the existing methods.

I. INTRODUCTION

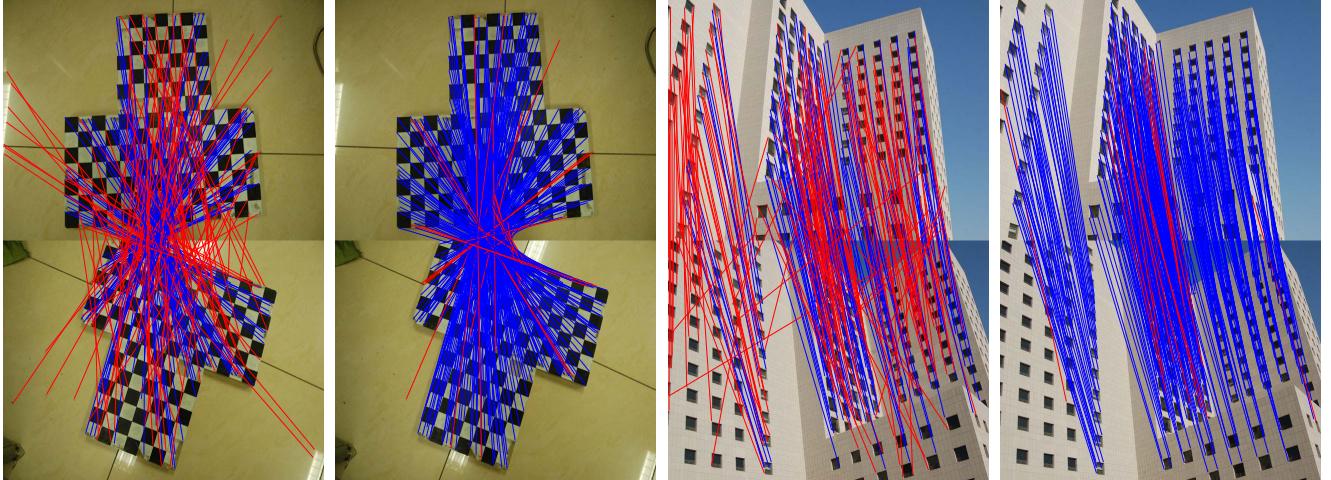
ONE of the core steps in many computer vision tasks is to establish reliable point correspondences between two images, such as object recognition [1], image stitch [2], [3], structure from motion [4], 3D reconstruction [5], [6] and so on. In most existing work, image matching is usually conducted by constructing and comparing local descriptors of interest points. By local descriptors, point matching becomes robust to scale changes, rotation changes, illumination changes as well as viewpoint changes to some extent. Many local descriptors have been proposed for this purpose, such as SIFT [1], GLOH [7], DAISY [8], [9] and so on. However, these methods usually fail when the matching images containing repetitive patterns, cf. Fig. 1(a) and Fig. 1(c). Although such repetitive patterns widely exist in images of man-made objects, such as buildings, these images are hard to be matched due to local ambiguities, even the viewpoint does not change too much. In order to solve the ambiguous problem when images have multiple similar regions, Mortensen et.al [10] proposed to combine the SIFT with global context to augment the performance of SIFT. By incorporating global context to local descriptor, their method can disambiguate the confusion induced by repetitive patterns to some extent.

The authors are with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, China (email: bfan@nlpr.ia.ac.cn).

This work is supported by the National Science Foundation of China (60835003, 61075038).

In order to increase the reliability and discriminative ability of point-to-point matches, Brown and Lowe [11] proposed to match images by invariant features from interest points groups. In their work, group of points is used to estimate a 2D transformation and the feature descriptor is computed by resampling local region to a canonical frame according to the estimated 2D transformation. Then matched features are refined by a Hough transform [12] followed by RANSAC [13] to compute the fundamental matrix as well as to find a final set of point correspondences that are consistent with epipolar geometry. Tell and Carlsson [14] used pairs of interest points for image matching. A pair of interest points is described by the Fourier coefficients of intensity profile between the two points in the pair. Then pairs of interest points are matched by comparing their Fourier coefficients. Finally, a set of point correspondences is obtained from the tentative matched pairs using a voting strategy due to the fact that each interest point can form many different pairs with their neighboring points and thus can be appeared in many matched pairs of interest points. The correspondence with higher votes is considered to be a correct one with higher confidence. Due to the low dimensionality of their descriptor, their work might be less distinctive and would be improved by describing pairs of points with more sophisticated high-dimensional descriptors, such as SIFT [1] or DAISY [8]. What is more, the voting strategy does not take into account the inter-relationship between correct correspondences.

Recently, the inter-relationship of correct correspondences has been shown to be important in rejecting mismatches in many literatures since it is reasonable to assume that the geometric relationship of correct correspondences should be preserved across images. In [15], Leordeanu and Hebert proposed to use a spectral technique to solve the correspondence problem using pairwise constraints. The basic idea of their method is to construct a graph with each potential correspondence to be a node and the weights are represented by pairwise agreements between linked nodes, i.e. pairwise agreements between potential correspondences. Since correct correspondences are likely to establish strong links among each other, while incorrect ones are only accidentally establish links with other correspondences, therefore a main strongly connected cluster can be formed by correct correspondences so that the correspondence problem can be solved by the spectral technique. Olivier Duchenne et.al [16] used tensors to generalize the idea of spectral method, thus higher-order constraints (such as relationship between tuples) can be incorporated in the correspondence



(a) 118/214 correct matches (55.1%) (b) 270/293 correct matches (92.2%) (c) 113/259 correct matches (43.6%) (d) 344/360 correct matches (95.6%)

Fig. 1. Comparison of matching results of image pairs with repetitive patterns. Blue lines indicate correct matches while red lines indicate incorrect ones. (a)(c) are matching results by local descriptor based method(DAISY [8]), while (b)(d) are Matching results by our proposed method.

problem. In Tell and Carlsson's work [17], topological constraints of corresponding points were incorporated and results were improved. Berg et.al [18] used low distortion correspondences for shape matching and object recognition. Low distortion correspondences are meant that the geometric relationship (in their paper are distance and angle) between correct correspondences should not change too much across images. They casted the correspondence problem to an integer quadratic programming problem, which is further approximated by a number of linear programming problems.

This paper is primarily focused on matching of images containing repetitive patterns. To this end, it presents a novel method based on matching pairs of interest points and defines a new compatibility measure in order to reliably obtain corresponding points from the tentative matched pairs of points. As shown in Fig. 2, the proposed method starts from matching point-pairs and then establishes consistent point correspondences from the tentative matched point-pairs under the low distortion constraint. Since the main challenge in matching images with repetitive patterns is the local ambiguities induced by repetitive patterns, describing pairs of interest points can utilize information in a much larger region, thus the local ambiguities may be reduced to some extent compared with the previous works based on local descriptors. In addition, owing to our newly defined compatibility measure between one correspondence and a set of point correspondences based on the low distortion constraint, consistent point correspondences are established reliably from the tentative matched point-pairs. As can be seen from Fig. 1(b) and Fig. 1(d), our proposed method can match images containing repetitive structures successfully. The effectiveness of our proposed method and its superiority to the state-of-the-art methods have been tested on images containing repetitive patterns (structures and textures) of both planar and 3D objects (indoor and outdoor).

The remainder of this paper is organized as follows.

Section 2 elaborates how to match pairs of interest points, followed by the algorithm to obtain point correspondences from the tentative matched pairs of interest points in Section 3. Experiments are reported in Section 4 and Section 5 concludes this paper.

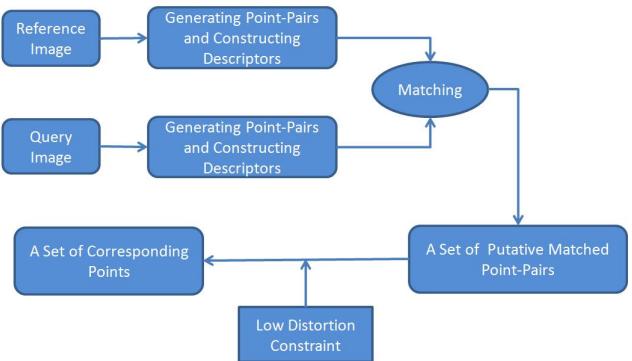


Fig. 2. Workflow of the proposed method.

II. MATCHING PAIRS OF INTEREST POINTS

As shown in Fig. 2, our proposed method starts from matching pairs of interest points, and then obtaining the final set of point correspondences from the tentative matched pairs of points based on the low distortion constraint. Therefore, firstly we need to generate, describe and match pairs of interest points successively for a pair of images.

A. Generating Pairs of Interest Points

Given a set of interest points in an image, intuitively, any two of them can form one pair. However, such a way to form pairs of points is redundant and will produce too many pairs since the number of pairs is quadratic in the number of points. If the number of points is large, then the computational burden of matching will be high. Therefore,

the number of pairs should be limited. Actually, as we will describe in the next subsection, the orientation decided by a pair of points is used for describing the pair in order to be rotation invariant. Thus the two points formed a pair should not too far from each other since the rotation invariance holds locally in real images. In the other aspect, if the two points of a pair are near, they do not contain more information than using only one of them, thus may reducing the distinctive ability of pairs of interest points. As a result, for each interest point X_i , pairs are formed by X_i and the other interest points $X_j \in \text{SubR}(X_i)$, in which $\text{SubR}(X) = \{Y : 50 \leq \|X - Y\|_2 < 100\}$. Therefore, for a given set of interest points $\{X_i, i = 1, 2, \dots, n\}$, pairs of interest points are formed as: $\{(X_i, X_j) : X_j \in \text{SubR}(X_i), i, j = 1, 2, \dots, n\}$ in which n is the number of interest points.

B. Constructing Descriptors for Pairs of Interest Points

Since local descriptors are designed to be distinctive while robust to photometric and geometric transformations, we use local descriptors to describe pairs of interest points in this work. Two descriptors of the two points in a pair are concatenated together to form the descriptor of this pair. Recently, DAISY [8] is proposed and reported with better performance than SIFT as well as other local descriptors [19], [9]. In addition, its source code is available on the Internet¹. Therefore, we choose DAISY to describe pairs of interest points, although any other good local descriptor can be embedded in our method.

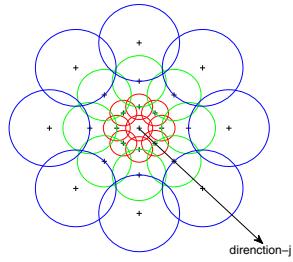


Fig. 3. The DAISY descriptor [8].

Similar to SIFT, DAISY is obtained by accumulating distributions of gradient orientation in the local neighborhood of an interest point. As shown in Fig. 3, the crosses represent sample points in the neighborhood of the interest point. For each sample point, a vector is obtained by convolving its neighboring points' gradient orientation maps with a Gaussian kernel. Then all the vectors of sample points are concatenated together as the DAISY descriptor of the interest point. The circles around the crosses are proportional to the standard deviation of their Gaussian kernels. In order to be rotation invariant, the DAISY descriptor is computed according to a direction. The reader is referred to [8], [9] for more details about DAISY.

As we said before, a direction is required to be assigned for each interest point in order to make the computed descriptor

be rotation invariant. Such a direction is usually assigned by histogramming technique [1]. In brief, the orientation corresponding to the largest peak in the orientation histogram within a region around the interest point is taken as the direction of this point. Any other local peak that is within 80% of the highest peak is also used as direction for descriptor extraction. Different from one single point, a pair of points supplies a direction intrinsically, i.e. the direction from the first point to the second point. Therefore, such direction is used for constructing descriptor of the pair. Note that here the DAISY descriptors of the two points in a pair is not identical to those obtained by the original DAISY method due to different direction assignment. The direction assigned according to the two points of a pair is more reliable than the direction assigned by histogramming technique according to our experiments.

C. Matching Pairs of Interest Points

Similar to point matching based on local descriptors, point-pairs are matched by the nearest neighbor distances of their descriptors. Although with a threshold on the nearest neighbor distance one can reject many mismatches, more correct matches will be rejected too. As we will introduce in the next section, an algorithm is proposed to obtain reliable point correspondences from all the tentative matched point-pairs based on the low distortion constraint. Most of mismatches will be rejected by such an algorithm, thus we do not set such a threshold and just take all the nearest neighbor matches as the tentative matched pairs of interest points. The set of tentative matched pairs is denoted as $\mathcal{P} = \{(M_{i1}, M_{i2}, s_i), i = 1, 2, \dots, N\}$ in which N is the number of matched pairs. In each matched pair of interest points, M_{i1}, M_{i2} are formed by the points in the matched pair and either is a tentative correspondence of points between two images. s_i is the nearest neighbor distance ratio(NNDR) [7] of this matched pair. The s_i of each matched pair can be considered as its confidence to be a correct one. The smaller the s_i is, the higher confidence of this matched pair to be a correct one. We choose the NNDR but not the nearest neighbor distance because the NNDR is more reliable and has the ability to discard ambiguous matches to some extent. Such a set of tentative matched pairs is used to obtain point correspondences in the subsequent step.

III. OBTAINING POINT CORRESPONDENCES FROM THE TENTATIVE MATCHED PAIRS OF INTEREST POINTS

Generally speaking, in each matched point-pair, it contains two correspondences of points formed by the points in the matched pair. However, we found in experiments that point correspondences existed in all the tentative matched pairs of points usually contain large amount of outliers. Therefore, we presents a method to obtain point correspondences reliably, which explores the inter-correspondence geometric relationships.

¹<http://cvlab.epfl.ch/~tola/daisy.html>



Fig. 4. Illustration of low distortion correspondences.

A. The Low Distortion Constraint

As shown in Fig. 4, (X_i^1, X_i^2) and (X_k^1, X_k^2) are two correspondences of points, the geometric relationship (such as distance) between X_i^1 and X_k^1 in the left image does not deviate too much from the relationship between X_i^2 and X_k^2 in the right image. Such an observation is called low distortion constraint. More specifically, given two point correspondences, $M_i = (X_i^1, X_i^2)$ and $M_k = (X_k^1, X_k^2)$, they are said to be compatible mutually if they are low distortion correspondences, i.e. they satisfy the following constraints:

$$\|X_i^1 - X_k^1\|_2 - \|X_i^2 - X_k^2\|_2 \leq t, \quad (1)$$

$$\|X_i^1 - X_k^1\|_2 \leq t_n \text{ or } \|X_i^2 - X_k^2\|_2 \leq t_n, \quad (2)$$

where $\|X_i^1 - X_k^1\|_2 - \|X_i^2 - X_k^2\|_2$ measures the deformation across images between two correspondences of points. The t controls the sensitivity on deformations. The larger t is, the more inter-image deformations can accommodate, also incorrect point correspondences may more likely to be considered as compatible. It is worth noting that (2) imposes the constraint of low distortion on neighboring correspondences (at least in one image), because the assumption of low distortion may not hold when point correspondences are far from each other. In all of our experiments, t is set to 15 pixels while t_n is set to 50 pixels. For convenience, we define the indicator function of compatibility as:

$$g(M_i, M_k) = \begin{cases} 1, & \text{if (1) (2) hold} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

B. Compatibility Measure Between One Correspondence and a Set of Point Correspondences

Having the definition of compatibility between two point correspondences, the compatibility between one correspondence and a set of point correspondences can be further defined, which leads to our proposed algorithm of obtaining a set of point correspondences from the tentative matched pairs. Given an initial set of point correspondences, reliable point correspondences are successively added to the set from the tentative matched pairs according to their compatibilities to the set. Suppose we have one corresponding points $M = (X^1, X^2)$ and a set of point correspondences, denoted as $\mathcal{C} = \{M_i, i = 1, 2, \dots, m\}$, then define

$$g_c(M, \mathcal{C}) = \begin{cases} 1, & \text{if } f_c(M, \mathcal{C}) \geq 0.85 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

as the indicator function of whether M is compatible with \mathcal{C} . In Eq. (4),

$$f_c(M, \mathcal{C}) = \frac{\sum_{i=1}^m g(M_i, M)}{\sum_{i=1}^m f(M_i, M)} \quad (5)$$

where

$$f(M_i, M) = \begin{cases} 1, & \text{if } \|X_i^1 - X^1\|_2 < t_n \text{ or } \|X_i^2 - X^2\|_2 < t_n \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

indicates whether M_i and M are neighboring correspondences.

Eq. (5) is the ratio of the number of correspondences that are compatible with M to the number of M 's neighboring correspondences. It can be considered as a qualitative measure of how much the point correspondence M is compatible with the set \mathcal{C} . If $f_c(M, \mathcal{C}) \geq 0.85$, M is said to be compatible with the set \mathcal{C} . In other words, M is compatible with \mathcal{C} if at least 85% of M 's neighboring correspondences in \mathcal{C} are compatible with M .

C. The Algorithm

Based on the definition of compatibility between one correspondence and a set of correspondences as in Eq. (4), an algorithm is proposed in order to reliably obtain point correspondences from the tentative matched pairs which is outlined in Algorithm 1. The input is a set of tentative matched pairs \mathcal{P} while the output is a set of point correspondences \mathcal{C} . Firstly, the set of point correspondences \mathcal{C} is initialized by one matched pair (such as the matched pair with the highest confidence, assumed to be the j th matched pair), that is $\mathcal{C} = \{M_{j1}, M_{j2}\}$. Once \mathcal{C} has been initialized, selecting and removing the matched pair (M_{i1}, M_{i2}, s_i) from \mathcal{P} with the highest confidence (i.e. the lowest s_i), then checking its compatibility with \mathcal{C} . If both M_{i1} and M_{i2} are compatible with \mathcal{C} , they are added into \mathcal{C} . Here we would note that since the compatibility between one correspondence M and a set of correspondences \mathcal{C} is defined on the basis of correspondences belonging to \mathcal{C} that are located in the t_n -neighborhood of M , there may be no such correspondence at all. In this case, Eq. (4) is meaningless. Therefore, if either two point correspondences in the selected matched pair (i.e. M_{i1} and M_{i2}) does not have any t_n -neighboring correspondence in \mathcal{C} , (M_{i1}, M_{i2}, s_i) is added to another auxiliary set \mathcal{U} . After all matched pairs in \mathcal{P} have been processed, the procedure is repeated with the matched pairs in \mathcal{U} until convergence. It typically needs less than 5 iterations to converge according to our experiments. If non or very few point correspondences are added to \mathcal{C} after initialization, it is very possible that the \mathcal{C} is initialized with incorrect correspondences. Therefore, we discard it (line 15-17 in Algorithm 1). Meanwhile, since such a greedy strategy tends to generate point correspondences in the same cluster as the initialized correspondences, we continue the process to obtain point correspondences from the remaining matched pairs (line 18-20 in Algorithm 1). As

Algorithm 1 ParseMatchedPair(\mathcal{P})

Require:A set of tentative matched pairs $\mathcal{P} = \{(M_{i1}, M_{i2}, s_i), i = 1, 2, \dots, N\}$.**Ensure:**A set of point correspondences $\mathcal{C} = \{M_i, i = 1, 2, \dots, m\}$.

```
1: Initialization:  $\mathcal{U} = \emptyset$ ,  $\mathcal{C} = \{M_{j1}, M_{j2}\}$ ,  $\mathcal{P} = \mathcal{P} - (M_{j1}, M_{j2}, s_j)$ .
2: while  $\mathcal{P} \neq \emptyset$  do
3:   Select  $(M_{i1}, M_{i2}, s_i)$  such that  $s_i \leq s_k, \forall (M_{k1}, M_{k2}, s_k) \in \mathcal{P}$ .
4:    $\mathcal{P} = \mathcal{P} - (M_{i1}, M_{i2}, s_i)$ .
5:   if  $\sum_{M \in \mathcal{C}} f(M_{i1}, M) > 0$  and  $\sum_{M \in \mathcal{C}} f(M_{i2}, M) > 0$  then
6:     if  $g_c(M_{i1}, M) = 1$  and  $g_c(M_{i2}, M) = 1$  then
7:        $\mathcal{C} = \{\mathcal{C}, M_{i1} : M_{i1} \notin \mathcal{C}\}$ ,  $\mathcal{C} = \{\mathcal{C}, M_{i2} : M_{i2} \notin \mathcal{C}\}$ 
8:     end if
9:   else
10:     $\mathcal{U} = \{\mathcal{U}, (M_{i1}, M_{i2}, s_i)\}$ 
11:  end if
12: end while
13:  $\mathcal{P} = \mathcal{U}$ ,  $\mathcal{U} = \emptyset$ 
14: Repeat 2-13 until  $\mathcal{P} = \emptyset$  or the size of  $\mathcal{C}$  does not change any more.
15: if The size of  $\mathcal{C}$  is less than 15 then
16:    $\mathcal{C} = \emptyset$ 
17: end if
18: if  $\mathcal{P} \neq \emptyset$  then
19:    $\mathcal{C} = \mathcal{C} \cup \text{ParseMatchedPair}(\mathcal{P})$ 
20: end if
21: return  $\mathcal{C}$ 
```

the NNDR of a matched pair reflects its confidence to be a correct one, it is reasonable to initialize \mathcal{C} with the matched pair whose NNDR is very low. To reduce the risk that the matched pair with the lowest NNDR is actually wrong, each of the first 5 lowest matched pairs is used to initialize \mathcal{C} separately. The matching result with the largest number of obtained point correspondences is served as the final set of point correspondences.

IV. EXPERIMENTS

A. Experimental Setup

We have conducted experiments to validate the effectiveness of our proposed method. As shown in Fig. 5, six pairs of images are used for experimental evaluation. All of them contain either repetitive structures or repetitive textures. Half of them are planar scenes while the rest are general 3D scenes. In our experiments, we used the Harris interest point [20] with the threshold set as 0.03. We systematically compare our method with the following state-of-the-art methods:

- **Daisy**: The point matching method based on the DAISY descriptor for each individual point. The matching strategy used here is NNDR (Nearest Neighbor Distance Ratio) [7].
- **Tell's method** [14]: Another matching method that uses pairs of interest points. It uses the Fourier coefficients of the profile between two points of a pair to describe the pair. Then the point correspondences are obtained from the tentative matched pairs of points by voting.

- **Daisy+GC**: Similar to [10], describing an interest point with both its local and global information. The local information is encoded in DAISY while the global information is encoded in GC(Global Context)². The matching strategy is NNDR, the same one as that used in the Daisy method.

In all of these methods, an one-to-one correspondence constraint has been imposed: if two or more points are matched to a single point in another image, we just keep the best one and discard the others. The matching results are sorted from the best to the worst. In our method, the point correspondence that is earlier added to the set of point correspondences is considered to be better. In Tell's method, the point correspondence that has more votes is considered to be better while in Daisy and Daisy+GC, the better correspondence is the one with smaller NNDR score.

In our experiments, we tested the performance of different methods under planar scenes and 3D scenes. For planar scenes, a correct point correspondence is related by a homography. However, since the coordinates of corresponding points are often corrupted by noise, they usually do not strictly satisfy the homography. In our experiments, if one point transformed by the homography is within 3 pixels of its corresponding point, then they are considered to

²In our experiments we have found that the Euclidean distance performs better than the χ^2 measure proposed in [10] when measuring the distance between two GC vectors. Therefore, in our experiments, the Euclidean distance is used.

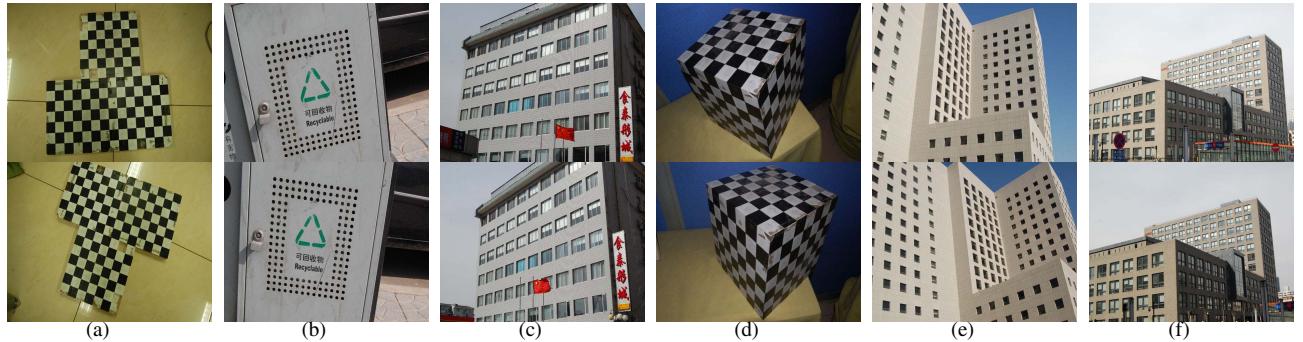


Fig. 5. Image pairs used for experimental evaluation.

be a correct match. Handcrafted point correspondences are used for estimating the ground-truth homographies in our experiments. Whereas, for 3D scenes, the epipolar geometry is used to check whether a pair of corresponding points is correct or not. Given a correspondence of points, if both of them lie in the corresponding epipolar lines, then they are regarded as a correct correspondence. Since the coordinates of corresponding points are usually corrupted by noise, they usually do not strictly satisfy the epipolar constraint. In our work, if the distance from a point to its corresponding epipolar line is within 3 pixels, it is considered to be on the epipolar line. The ground-truth fundamental matrices are estimated from handcrafted point correspondences for each tested pair. Note that such an epipolar relationship is just a necessary condition for a correct correspondence, some false correspondences may pass through. However, since all the methods are evaluated with the same criterion, it is still a meaningful indicator.

B. Experimental Results

The comparative results are shown in Fig. 6. From Fig. 6(a) to Fig. 6(c) are results of planar scenes while the results of 3D scenes are shown in Fig. 6(d), Fig. 5(e) and Fig. 5(f). It can be seen that due to repetitive patterns, the traditional local descriptor based method can not achieve satisfactory results, as shown by plots of Daisy in Fig. 6. By incorporating global context to local descriptor or matching pairs of interest points, Daisy+GC, Tell's method and our method can achieve better results than the Daisy method. In fact, matching pairs of interest points also incorporates information in a much larger region than matching local descriptors of individual interest points. This is one of the reasons why our proposed method works well. Although both the Tell's method and our proposed method are based on the pairs of interest points, our proposed method significantly outperforms Tell's method. This is because that the descriptor constructed for point-pairs in this work is much more distinctive than that used in Tell's method. What is more, in order to obtain point correspondences from the tentative matched point-pairs, our newly defined compatibility measure takes into consideration of the geometrical relationship of correspondences which is much more reliable than the voting strategy used in Tell's method. It is clearly that our

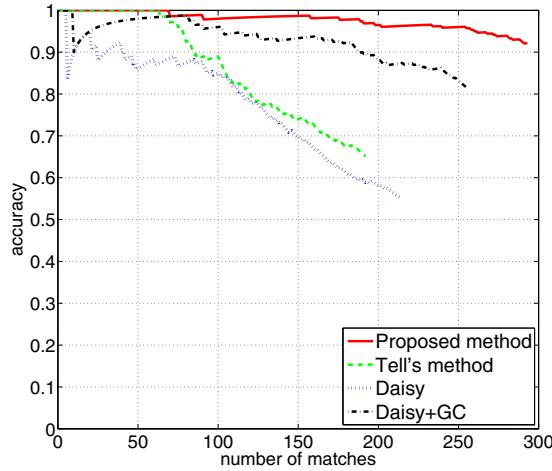
proposed method outperforms the state-of-the-art method as can be seen in Fig. 6. Fig. 7 and Fig. 8 show concrete matching results by the four tested methods for image pairs in Fig. 5(a) and Fig. 5(e) respectively. It is clear that our method not only has more correct correspondences but also with higher accuracy than other methods.

V. CONCLUSION

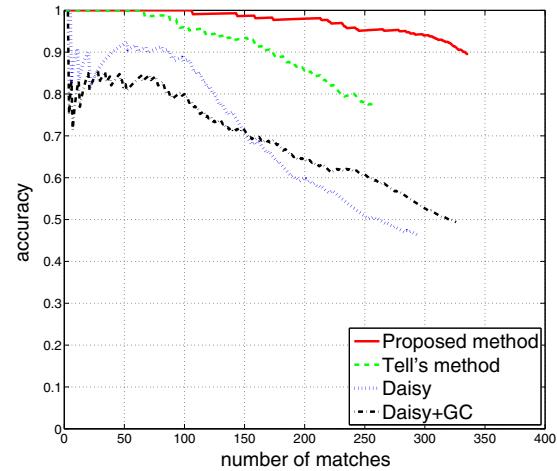
Although image matching methods based on local descriptors have been well studied, they are usually fail to match images of repetitive patterns even the viewpoint changes between these images are not very large. In this paper, we presents a novel method for matching images containing repetitive patterns, which are very common in images of man-made scenes or objects. Since pairs of interest points not only have higher distinctive ability but also can encode more information in a much larger region than local descriptors of individual points, the proposed method is based on pairs of interest points. In addition, the final set of point correspondences are obtained from the tentative matched point-pairs based on a newly defined compatibility measure between one correspondence and a set of correspondences under the low distortion assumption. Experimental results on images of both planar and 3D scenes including comparisons with other methods have demonstrated the effectiveness of our proposed method.

REFERENCES

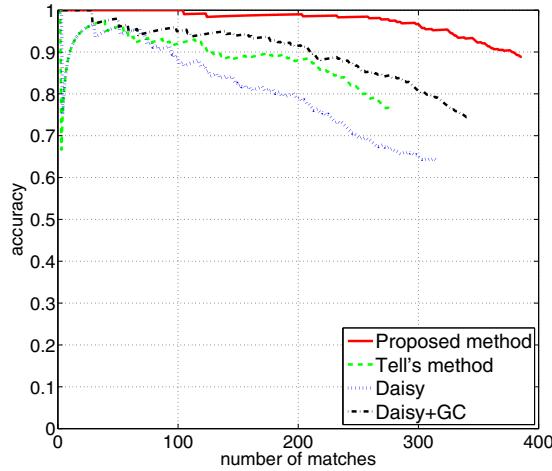
- [1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [2] M. Brown and D. Lowe, "Recognising panoramas," in *International Conference on Computer Vision*, 2003.
- [3] R. Szeliski, "Image alignment and stitching: A tutorial," *Foundations and Trends in Computer Graphics and Vision*, vol. 2, pp. 1–104, 2006.
- [4] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: Exploring photo collections in 3D," *ACM Transactions on Graphics (TOG)*, vol. 25, pp. 835–846, 2006.
- [5] S. Agarwal, N. Snavely, I. Simon, S. M. Seitz, and R. Szeliski, "Building rome in a day," in *International Conference on Computer Vision*, 2009.
- [6] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 99, 2009.
- [7] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.



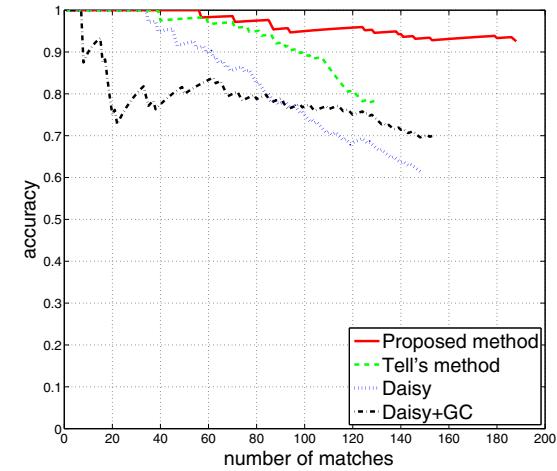
(a) Results for image pair in Fig.5(a).



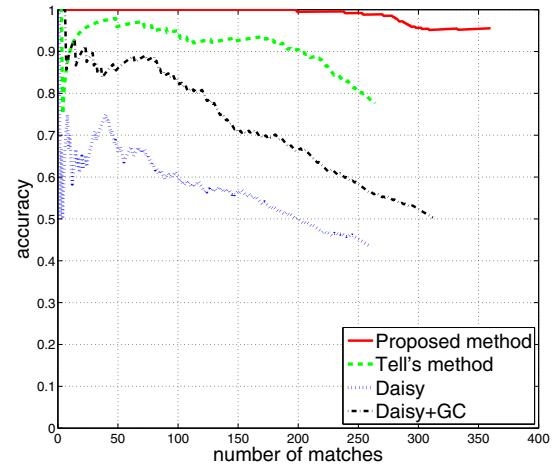
(b) Results for image pair in Fig.5(b).



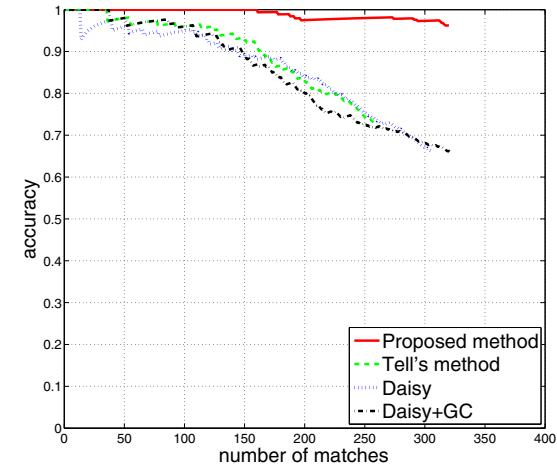
(c) Results for image pair in Fig.5(c).



(d) Results for image pair in Fig.5(d).



(e) Results for image pair in Fig.5(e).



(f) Results for image pair in Fig.5(f).

Fig. 6. Experimental results for image pairs shown in Fig.5. The number of detected corners in each pair of images are: (a) 352 and 334, (b) 494 and 499, (c) 495 and 493, (d) 197 and 231, (e) 500 and 500, (f) 495 and 489.

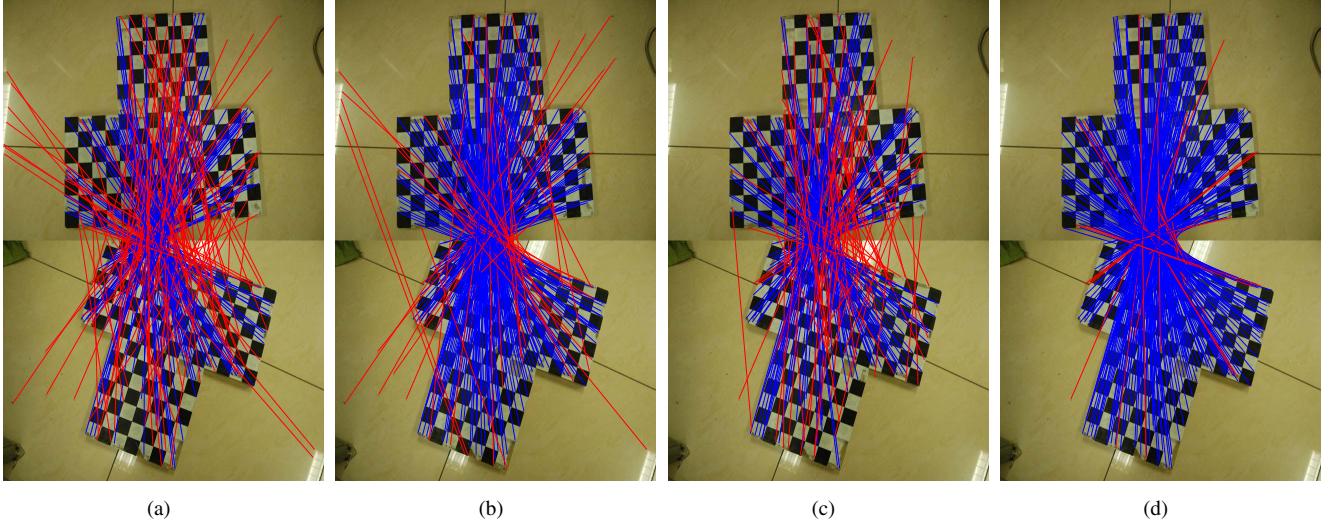


Fig. 7. Matching results for images of planar scene containing repetitive patterns. The blue lines are correct matches, while the incorrect ones are marked by red lines. (a) 118/214 correct matches with **Daisy** method (55.1%); (b) 208/256 correct matches with **Daisy+GC** method (81.3%); (c) 125/192 correct matches with **Tell's method** (65.1%); (d) 270/293 correct matches with our proposed method (92.2%).

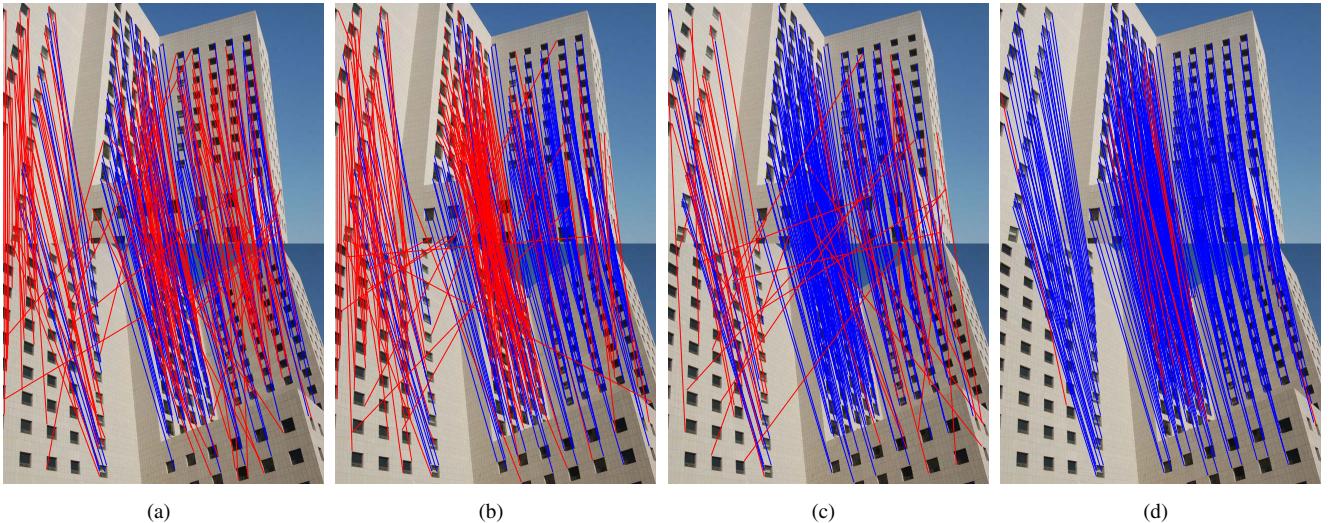


Fig. 8. Matching results for images of 3D scene containing repetitive structures. The blue lines are correct matches, while the incorrect ones are marked by red lines. (a) 113/259 correct matches with **Daisy** method (43.6%); (b) 157/312 correct matches with **Daisy+GC** method (50.3%); (c) 205/264 correct matches with **Tell's method** (77.7%); (d) 344/360 correct matches with our method (95.6%).

- [8] E. Tola, V. Lepetit, and P. Fua, "A fast local descriptor for dense matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [9] S. Winder, G. Hua, and M. Brown, "Picking the best DAISY," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.
- [10] H. D. E. Mortensen and L. Shapiro, "A SIFT descriptor with global context," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [11] M. Brown and D. Lowe, "Invariant features from interest point groups," in *British Machine Vision Conference*, 2002.
- [12] D. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, pp. 111–122, 1981.
- [13] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, pp. 381–385, 1981.
- [14] D. Tell and S. Carlsson, "Wide baseline point matching using affine invariants computed from intensity profiles," in *European Conference on Computer Vision*, 2000.
- [15] M. Leordeanu and M. Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *International Conference on Computer Vision*, 2005.
- [16] O. Duchenne, F. Bach, I. Kweon, and J. Ponce, "A tensor-based algorithm for high-order graph matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.
- [17] D. Tell and S. Carlsson, "Combining appearance and topology for wide baseline matching," in *European Conference on Computer Vision*, 2002.
- [18] A. C. Berg, T. L. Berg, and J. Malik, "Shape matching and object recognition using low distortion correspondences," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.
- [19] S. Winder and M. Brown, "Learning local image descriptors," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [20] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conference*, 1988, pp. 147–151.