FEATURE MATCHING BASED ON TOP K RANK SIMILARITY

Junjun Jiang[1,2]*, Qing Ma*[3]*, Tao Lu*[4]*, Zhongyuan Wang*[5]*, and Jiayi Ma*[5]

[1] CUG, Wuhan, China [2] NII, Tokyo, Japan [3] CCNU, Wuhan, China
[4] WIT, Wuhan, China [5] WHU, Wuhan, China

ABSTRACT

Feature matching plays a key component in many computer vision and pattern recognition tasks. Observing that the spatial neighborhood relationship (representing the topological structures of an image scene) is generally well preserved between two feature points of an image pair, some mismatch removing methods based on maintaining the local neighborhood structures of the potential true matches have been proposed. How to define the local neighborhood structure is an issue of vital importance. In this paper, we propose a robust and efficient method, called Top K Rank Preservation (TopKRP), for mismatch removal from given putative point set matching correspondences. Instead of preserving the intersection of neighbors, TopKRP aims at preserving the top K rank of two feature points. The developed approach is validated on numerous challenging real image pairs for general feature matching, and the experimental results demonstrate that it outperforms several state-of-the-art feature matching methods, especially in case of a large number of mismatches.

Index Terms— Feature matching, mismatch removal, top K rank similarity, local neighborhood structure

1. INTRODUCTION

Feature matching is a fundamental and critical problem in the field of computer vision and pattern recognition, and whose aim is to establish reliable correspondences between two feature sets. It is a very hot topic and has been widely used in stereo vision matching, target recognition and tracking, medical image analysis, image super-resolution, etc. [1, 2, 3, 4, 5, 6]. Due to its combinatorial nature, feature matching is essentially a very complex problem of NPC complex optimization. Specifically, matching N points to another N points would lead to a total of $N!$ permutations [7, 8]. To address this problem, in recent years there are many approaches have been proposed. The common idea is to first obtain a set of putative correspondence based on a local image descriptor (for example, Scale Invariant Feature Transform (SIFT) [9]), and then try to remove the outliers (incorrect match) with additional geometrical constraints. In this paper, we mainly focus on the problem of removing the mismatches from some given putative point correspondences. In the following, we will review some representative methods.

The most representative feature matching method is random sample consensus (RANSAC), which tries to randomly select a set of points that can fit a given geometric model and then calculate the model parameters. This method has been extended by MLESAC [10] and PROSAC [11]. Although the RANSAC algorithm and its variations have achieved great success in the problem of feature matching, they also have some limitations: the algorithm relies on a geometric parameter model; however, when the scene contains non-rigid motion, the geometric relationship between images will not meet any parameter model and the algorithm will no longer be valid [12].

To this end, in recent years some non-parametric interpolation methods [12, 13, 14, 15, 16] have been advocated. They commonly interpolate a non-parametric function based on the assumption that the motion field associated with the feature correspondence is slow-and-smooth. Benefit from graph or hypergraph theory [17, 18], graph matching is another kind of typical feature matching technique. By constructing the affinity matrix of the point set [19], it can obtain the ordered features of the point set based on the graph spectrum. Some representative graph matching based approaches include Dual Decomposition (DD) [20], Spectral Matching (SM) [21], and Graph Shift (GS) [22]. DD algorithm [20] formulas the feature matching into an energy minimization problem, SM [21] uses a high efficiency method to find the correspondence between the two sets of features, and GS [22] constructs an affinity graph based on the SM algorithm [21], where the maximal clique of the graph is regarded as the spatial correspondences. In addition, the approach of [23] uses local neighborhoods for feature description, and graph-based feature matching is presented to alleviate false matches. Although graph matching based algorithms can provide considerable flexibility to the object model and delivers robust matching and recognition, it still suffers from high complexity due to its non-polynomial-hard nature.

Most recently, Ma *et al.* [24] proposed a novel mismatch

The research was supported by the National Natural Science Foundation of China under Grant 61501413, Grant 61773295, Grant 61503288, Grant 61502354, Grant 61671332, and Grant U1404618, the Hubei Province Technological Innovation Major Project under Grant 2017AAA123, and the National Key R&D Project under Grant 2016YFE0202300.

removing framework based on Locality Preserving Matching (LPM). Observing that the absolute distance between two feature points of an image pair may change significantly under viewpoint change or non-rigid deformation, but the spatial neighborhood relationship among feature points (representing the topological structures of an image scene) is generally well preserved [25, 26, 27, 28], they developed a robust and effective mismatch removing method through maintaining the local neighborhood structures of the potential true matches. It can handle both rigid and non-rigid deformations related between two images very efficiently.

The core idea of LPM [24] is to preserve the topological structures of two feature points from image pairs. However, it defines the topological similarity between feature points through calculating their intersection of neighbors, which ignores the differences of neighbors and is difficult to exploit the true topology. In order to more effectively exploit the local neighborhood structures of feature points, in this paper we apply the top K rank similarity to measure the differences of topological structures of two feature points. We coin our proposed method Top K Rank Preservation (TopKRP) based feature matching. Our proposed method could take into consideration that the differences between neighbors (different neighbors have different rankings), thus exploit the truth topological similarity. In particular, we first search the K nearest neighbors (K-NN) for the feature point, and obtain its ranking list. By this way, we transform the data from the feature space to the ranking list space, which is much more robust to the raw feature extraction approaches [29, 30]. Therefore, the topological structure similarity of two feature points can be simply calculated through comparing the top K ranking lists.

2. PROPOSED METHOD

2.1. Top K Rank Similarity Measurement

To solve the problem of (i) local descriptor will inevitably lead to a number of false matches and (ii) global transformation model requires solving a special parametric or non-parametric optimization problem, Ma *et al.*'s [24] proposed a novel method by preserving the local neighborhood structures of potential true matches, based on the assumption that if two feature points are the correct match, they will share similar local neighborhood structure. They defined the local neighborhood structure similarity between two feature points as the overlap of their K-NN sets. It treats the K-NN equally, however, cannot exploit the true local neighborhood structure.

In contrast to using the intersection (overlap) to define the similarity between feature points, in this paper we introduce the top K rank similarity to measure two putative feature correspondences, $\mathbf{x}$ and $\mathbf{y}$, extracted from two given images. Denote $\sigma(\mathbf{x})$ and $\sigma(\mathbf{y})$ the ranking lists of $\mathbf{x}$ and $\mathbf{y}$, respectively. The top K rank similarity between two feature points $Sim(\sigma(\mathbf{x}), \sigma(\mathbf{y}))$ [1] can be defined by the weighted Spearman's footrule distance [31],

$$Sim(\sigma(\mathbf{x}), \sigma(\mathbf{y})) = \frac{\sum_{k=1}^{K} \phi_k}{\phi(\sigma(\mathbf{x}), \sigma(\mathbf{y}))}, \tag{1}$$

where $\phi(\sigma(\mathbf{x}), \sigma(\mathbf{y}))$ is the weighted Spearman's footrule measure between two full ranked lists $\sigma(\mathbf{x})$ and $\sigma(\mathbf{y})$:

$$\phi(\sigma(\mathbf{x}), \sigma(\mathbf{y})) = -2K + 2z \sum_{k=1}^{K} 1/k. \tag{2}$$

In Eq. (1), the contribution of item k to $\phi(\sigma(\mathbf{x}), \sigma(\mathbf{y}))$ is

$$\phi_k = \begin{cases} \frac{||\sigma_k(\mathbf{x}) - \sigma_k(\mathbf{y})||_1}{\min\{\sigma(\mathbf{x}), \sigma(\mathbf{y})\}}, & k \in \sigma(\mathbf{x}) \cap \sigma(\mathbf{y}), \\ \frac{||\sigma_k(\mathbf{x}) - z||_1}{\min\{\sigma_k(\mathbf{x}), z\}}, & \text{otherwise}, \end{cases} \tag{3}$$

where $\sigma_k(\mathbf{x})$ is the ranking of item k in the list $\sigma(\mathbf{x})$, $\sigma_k(\mathbf{y})$ is the ranking of item k in the list $\sigma(\mathbf{y})$, and z is defined as,

$$z = \frac{K - 4\lfloor K/2 \rfloor + 2(K+1)\sum_{k=1}^{\lfloor K/2 \rfloor} 1/k}{\sum_{k=1}^{K} 1/k}. \tag{4}$$

Here, the operator $\lfloor \cdot \rfloor$ rounds the elements of a to the nearest integers less than or equal to a.

2.2. Objective Function

Given a set of N putative feature correspondences $S = \{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^{N}$ extracted from two given images, where $\mathbf{x}_i$ and $\mathbf{y}_i$ are 2D column vectors denoting the spatial positions of feature points (our approach is not limited by the dimension of the input data, which can be directly applied to 3D matching problems), the aim of our proposed method is to remove the outliers from a putative correspondence set S to establish accurate correspondences.

Following [24], we also assume that the local neighborhood structure should be preserved, especially when the spatial relationship between the image pair is a simple rigid transformation. In LPM method [24], they ignore the differences of neighbors and cannot exploit the true local neighborhood structure. To this end, in this paper we consider the differences between neighbors (the closer the neighbors, the greater the weights), and develop a top K rank similarity preservation method. Mathematically, denoting $\mathcal{I}$ the unknown inlier set, our objective function can be written as follows,

$$\mathcal{I}^* = \arg\min_{\mathcal{I}} C(\mathcal{I}; S, \lambda), \tag{5}$$

where C is the cost function:

$$C(\mathcal{I}; S, \lambda) = \sum_{i \in \mathcal{I}} \big(1 - Sim(\sigma(\mathbf{x}_i), \sigma(\mathbf{y}_i))\big) + \lambda(N - |\mathcal{I}|), \tag{6}$$

[1] The returned value of Sim is between 0 and 1, where the lists are identical if Sim = 1 and are completely disjoint if Sim = 0.

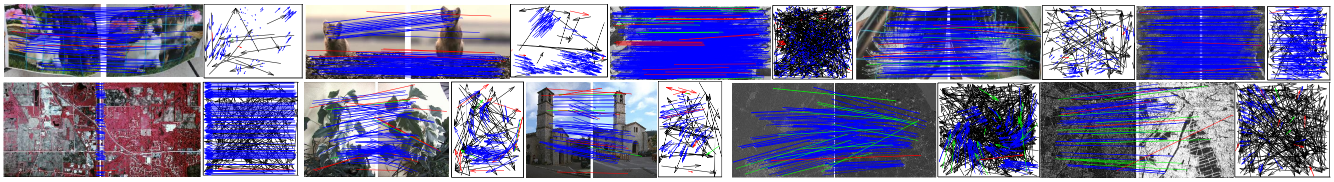

Fig. 1. Feature matching results of our proposed TopKRP method on *Dog*, *Fox*, *RS01*, *Peacock*, *RS02*, *RS03*, *Tree*,*Church*, *RS04*, and *RS05* (from left to right and top to down). The ratio of outliers in the 10 image pairs are 17.7%, 15.48%, 30.25%, 28.39%, 26.28%, 59.19%, 43.71%, 45.24%, 68.38%, and 79.57%. The head and tail of each arrow in the motion field correspond to the positions of feature points in two images (blue = true positive, black = true negative, green = false negative, red = false positive). For visibility, in the image pairs, at most 100 randomly selected matches are presented, and the true negatives are not shown.

where $\sigma(\mathbf{x}_i)$ denotes the top K ranking list of $\mathbf{x}_i$, and $Sim(\sigma(\mathbf{x}_i), \sigma(\mathbf{y}_i))$ is top K rank similarity between $\mathbf{x}_i$ and $\mathbf{y}_i$, and $|\cdot|$ denotes the cardinality of a set. In this cost function, the first term penalizes any match which does not preserve the local ranking list of a point pair, the second term discourages the outliers, and the parameter $\lambda > 0$ controls the tradeoff between these two terms. Ideally, the optimal solution should achieve zero penalty, *i.e.*, the first term of C should be zero.

Mathematically, we introduce an $N \times 1$ binary vector $\mathbf{p} = [p_1, p_2, \cdots, p_N]$ to associate the putative set S, where $p_i \in \{0, 1\}$ denotes the match correctness of the i-th correspondence $(\mathbf{x}_i, \mathbf{y}_i)$. Specifically, $p_i = 1$ indicates inlier while $p_i = 0$ points to outlier. Therefore, the cost function in Eq. (6) can be rewritten as:

$$C(\mathbf{p}; S, \lambda) = \sum_{i=1}^{N} p_i \big(1 - Sim(\sigma(\mathbf{x}_i), \sigma(\mathbf{y}_i))\big) + \lambda \left(N - \sum_{i=1}^{N} p_i\right). \tag{7}$$

Let $d_i = 1 - Sim(\sigma(\mathbf{x}_i), \sigma(\mathbf{y}_i))$, which could be seen as the distance (dissimilarity) of local structure between $\mathbf{x}_i$ and $\mathbf{y}_i$, Eq. (7) can be rewritten as,

$$C(\mathbf{p}; S, \lambda) = \sum_{i=1}^{N} p_i (d_i - \lambda) + \lambda N. \tag{8}$$

The problem of removing outliers and establishing accurate feature correspondences is transformed to solving an optimization problem (8).

2.3. Optimization

In practice, the obtained putative set contains not only correct matches, but also a large number of incorrect matches. Given that we know the correct correspondences, we can obtain the ranking lists of $\mathbf{x}$ and $\mathbf{y}$ in the correct correspondence set, then we can calculate the cost values $\{d_i\}_{i=1}^N$ beforehand. Therefore, there is only one variable that needs to be optimized in Eq. (8), *i.e.*, the match correctness of p_i. Under the constraints that $p_i \in \{0, 1\}$ for $i = 1, 2, \cdots, N$, the solution of Eq. (8) is not difficult to get: (i) when the distance d_i is larger than λ, we prefer to set the value of p_i to 0, thus avoid increasing the cost function. (ii) when the distance d_i is smaller than λ, we prefer to set the value of p_i to 1, thus decreasing the cost function. Therefore, the optimal solution of $\mathbf{p}$ that minimizes Eq. (8) is determined by the following simple criterion:

$$p_i = \begin{cases} 1 & d_i \le \lambda \\ 0 & d_i > \lambda \end{cases}, \ i = 1, \cdots, N. \tag{9}$$

And hence, the optimal inlier set $\mathcal{I}^*$ is determined by:

$$\mathcal{I}^* = \{i \mid p_i = 1, i = 1, \cdots, N\}. \tag{10}$$

As we discussed above, the solution of Eq. (8) is determined given that the correct correspondences are known beforehand. In this paper, we propose an iterative optimization strategy to update the putative set S and the cost values $\{d_i\}_{i=1}^N$. In our experiments, after three iterations, our proposed method can converge to a stable result.

3. EXPERIMENTAL RESULTS

In this section, to evaluate the performance of the proposed TopKRP method, we experimentally validate it on several pairs of real images. The open source VLFEAT toolbox [33] is employed to determine the putative correspondence of SIFT [9] and to search the K nearest neighbors using K-D tree. To demonstrate the advantage of the proposed method, all of the experimental results are compared with those of three state-of-the-art feature matching methods, such as ICF [12], GS [22], and LPM [24]. RANSAC [32] is taken to be the baseline for comparison. For these comparison methods, we fine tune their parameters to obtain best performances. As for the proposed TopKRP, we use a grid searching of parameter settings for parameter optimization, $K = [3:2:33]$, and $\lambda = [0.1:0.05:0.9]$, and set $K = 23, 9, 5$ and $\lambda = 0.45, 0.2, 0.2$ for the first to the third iteration, respectively.

As shown in Fig. 1, we show the feature matching results of the proposed TopKRP on ten public image pairs,

Table 1. Precision (%) and recall (%) of RANSAC [32], ICF [12], GS [22], LPM [24], and our proposed TopKRP method on the ten image pairs in Fig. 1. For each result in the bracket, the left is the precision and the right is the recall. For comparison, we mark the values below 90% in blue.

	RANSAC [32]	ICF [12]	GS [22]	LPM [24]	TopKRP
Dog	(100.0, 95.70)	(92.19, 63.44)	(97.70, 91.40)	(96.88, 100.0)	(97.87, 98.92)
Fox	(100.0, 89.31)	(100.0, 48.85)	(98.48, 99.24)	(92.25, 100.0)	(97.01, 99.24)
RS01	(97.94, 94.36)	(100.0, 60.69)	(100.0, 75.54)	(96.01, 100.0)	(98.12, 97.92)
Peacock	(98.59, 82.84)	(99.12, 68.86)	(99.32, 86.98)	(89.42, 100.0)	(98.82, 98.82)
RS02	(100.0, 98.70)	(100.0, 63.91)	(100.0, 98.26)	(94.63, 99.57)	(99.13, 98.70)
RS03	(96.52, 100.0)	(100.0, 80.18)	(100.0, 72.97)	(62.71, 100.0)	(99.11, 100.0)
Tree	(97.10, 71.28)	(92.75, 68.09)	(97.62, 87.23)	(80.70, 97.87)	(90.82, 94.68)
Church	(94.74, 78.26)	(91.67, 63.77)	(91.78, 97.10)	(94.74, 78.26)	(90.41, 95.65)
RS04	(99.46, 99.46)	(100.0, 65.76)	(100.0, 32.07)	(79.04, 98.37)	(97.28, 97.28)
RS05	(89.33, 100.0)	(95.59, 97.01)	(81.97, 74.63)	(47.10, 97.01)	(92.75, 95.52)

five general (named as "*Dog*", "*Fox*", "*Peacock*", "*Tree*", and "*Church*") images and five remote sensing images (named as "*RS01*", "*RS02*", "*RS03*", "*RS04*", and "*RS05*"). For each group, the left image pair schematically shows the matching result, and the right motion field provides the decision correctness of each correspondence in the putative set. The ratio of outliers of the below five image pairs is higher than that of the above five ones. From the results, we learn that our TopKRP approach can always achieve reasonable results even through there exists many outliers. Please see the images of "*RS03*", "*RS04*", and "*RS05*", whose outlier ratios are 59.19%, 68.38%, and 79.57%, respectively.

Table 1 tabulates the objective results in terms of precision and recall[2] indexes of five methods on the ten image pairs. The first five image pairs have relative small outlier ratios, while the rest five image pairs have large outlier ratios. From these results, we can clearly see that all methods can achieve very good performances when the inlier ratio is high. ICF [12] cannot work well for *Dog* and *Fox* pairs, and this can be explained by that the introduced slow-and-smooth prior of ICF [12] will fail if the motion field involves large depth discontinuity or motion inconsistency. GS [22] usually has high precision and low recall. When the outlier ratio is high, *i.e.*, there are many incorrect correspondences in the putative set, the performances (especially the precision) of all the comparison methods begin to fall sharply. Even there is only 20.43% (67 correct correspondences in 328 putative correspondences) inliers, our method can still obtain a very good results, 92.75% and 95.52% in terms of precision and recall, respectively. In contrast, the proposed TopKRP method shows the tradeoff between precision and recall for large outlier ratios. This demonstrates the generality and its ability of TopKRP to handle various matching problems.

[2]The precision is defined by the ratio between the number of correct matches found between the number of total matches found, while the recall can calculated by the ratio between number of correct matches found and the number of right correspondences

4. CONCLUSIONS

In this work we present a novel and robust feature matching approach through Top K Rank Preservation (TopKRP) based on mismatch removal. In TopKRP, we first use the Top K ranking list to denote the topology of data point to be matched. Then, the weighted Spearmans footrule distance is introduced to measure the similarity between two Top K ranking lists of one image pair. And then, we formulate the feature matching as an iterative optimization problem that has a very efficient solution. Experimental results have demonstrated that our proposed TopKRP method can achieve better results on ten public image pairs with typical scenes compared with the state-of-the-art feature matching algorithms. Specially, when there are many incorrect correspondences in the putative set, our method can get a well-balanced result between precision and recall.

5. REFERENCES

[1] Jiayi Ma, Huabing Zhou, Ji Zhao, Yuan Gao, Junjun Jiang, and Jinwen Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, 2015.

[2] Song Bai, Xiang Bai, and Qi Tian, "Scalable person re-identification on supervised smoothed manifold," in *CVPR*, 2017.

[3] Junjun Jiang, Ruimin Hu, Zhongyuan Wang, Zhen Han, and Shi Dong, "Manifold regularized sparse support regression for single image super-resolution," in *ICASSP*, 2013, pp. 1429–1433.

[4] Jiayi Ma, Ji Zhao, Jinwen Tian, Xiang Bai, and Zhuowen Tu, "Regularized vector field learning with sparse approximation for mismatch removal," *Pattern Recognition*, vol. 46, no. 12, pp. 3519–3532, 2013.

[5] Jiayi Ma, Ji Zhao, Yu Zhou, and Jinwen Tian, "Mismatch removal via coherent spatial mapping," in *ICIP*, 2012, pp. 1–4.

[6] Junjun Jiang, Jican Fu, Tao Lu, Ruimin Hu, and Zhongyuan Wang, "Locally regularized anchored neighborhood regression for fast super-resolution," in *ICME*, 2015, pp. 1–6.

[7] Jiayi Ma, Ji Zhao, Jinwen Tian, Zhuowen Tu, and Alan L Yuille, "Robust estimation of nonrigid transformation for point set registration," in *CVPR*, 2013, pp. 2147–2154.

[8] Chao Wang, Lei Wang, and Lingqiao Liu, "Progressive mode-seeking on graphs for sparse feature matching," in *ECCV*, 2014, pp. 788–802.

[9] David G Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.

[10] Philip HS Torr and Andrew Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.

[11] Ondrej Chum and Jiri Matas, "Matching with PROSAC - progressive sample consensus," in *CVPR*, 2005, pp. 220–226.

[12] Xiangru Li and Zhanyi Hu, "Rejecting mismatches by correspondence function," *Int. J. Comput. Vis.*, vol. 89, no. 1, pp. 1–17, 2010.

[13] Ji Zhao, Jiayi Ma, Jinwen Tian, Jie Ma, and Dazhi Zhang, "A robust method for vector field learning with application to mismatch removing," in *CVPR*, 2011, pp. 2977–2984.

[14] Jiayi Ma, Ji Zhao, Jinwen Tian, Alan L Yuille, and Zhuowen Tu, "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, 2014.

[15] Jiayi Ma, Junjun Jiang, Chengyin Liu, and Yansheng Li, "Feature guided gaussian mixture model with semi-supervised em and local geometric constraint for retinal image registration," *Information Sciences*, vol. 417, pp. 128–142, 2017.

[16] Gang Wang, Qiangqiang Zhou, and Yufei Chen, "Robust non-rigid point set registration using spatially constrained gaussian fields," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1759–1769, 2017.

[17] Junchi Yan, Xu-Cheng Yin, Weiyao Lin, Cheng Deng, Hongyuan Zha, and Xiaokang Yang, "A short survey of recent advances in graph matching," in *ICMR*, 2016, pp. 167–174.

[18] Junchi Yan, Changsheng Li, Yin Li, and Guitao Cao, "Adaptive discrete hypergraph matching," *IEEE Trans. Cybernetics*, vol. 48, no. 2, pp. 765–779, 2018.

[19] Junchi Yan, Minsu Cho, Hongyuan Zha, Xiaokang Yang, and Stephen M Chu, "Multi-graph matching via affinity optimization with graduated consistency regularization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 6, pp. 1228–1242, 2016.

[20] Lorenzo Torresani, Vladimir Kolmogorov, and Carsten Rother, "Feature correspondence via graph matching: Models and global optimization," in *ECCV*, 2008, pp. 596–609.

[21] Marius Leordeanu and Martial Hebert, "A spectral technique for correspondence problems using pairwise constraints," in *ICCV*, 2005, pp. 1482–1489.

[22] Hairong Liu and Shuicheng Yan, "Common visual pattern discovery via spatially coherent correspondence," in *CVPR*, 2010, pp. 1609–1616.

[23] Man Hee Lee, Minsu Cho, and In Kyu Park, "Feature description using local neighborhoods," *Pattern Recognit. Lett.*, vol. 68, pp. 76–82, 2015.

[24] Jiayi Ma, Ji Zhao, Hanqi Guo, Junjun Jiang, Huabing Zhou, and Yuan Gao, "Locality preserving matching," in *International Joint Conference on Artificial Intelligence (IJCAI-17)*, 2017, pp. 4492–4498.

[25] L. Torresani, V. Kolmogorov, and C. Rother, "A dual decomposition approach to feature correspondence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 2, pp. 259–271, Feb 2013.

[26] Shanmin Pang, Jianru Xue, Qi Tian, and Nanning Zheng, "Exploiting local linear geometric structure for identifying correct matches," *Computer Vision and Image Understanding*, vol. 128, no. Supplement C, pp. 51 – 64, 2014.

[27] Jianwei Fan, Yan Wu, Ming Li, Wenkai Liang, and Qiang Zhang, "Sar image registration using multiscale image patch features with sparse representation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 4, pp. 1483–1493, 2017.

[28] Jiayi Ma, Ji Zhao, Junjun Jiang, and Huabing Zhou, "Non-rigid point set registration with robust transformation estimation under manifold regularization," in *AAAI*, 2017, pp. 4218–4224.

[29] Mang Ye, Chao Liang, Yi Yu, Zheng Wang, Qingming Leng, Chunxia Xiao, Jun Chen, and Ruimin Hu, "Person reidentification via ranking aggregation of similarity pulling and dissimilarity pushing," *IEEE Transactions on Multimedia*, vol. 18, no. 12, pp. 2553–2566, 2016.

[30] Mang Ye, Andy J Ma, Liang Zheng, Jiawei Li, and Pong C Yuen, "Dynamic label graph matching for unsupervised video re-identification," in *ICCV*, 2017.

[31] Ronald Fagin, Ravi Kumar, and Dakshinamurthi Sivakumar, "Comparing top k lists," *SIAM Journal on Discrete Mathematics*, vol. 17, no. 1, pp. 134–160, 2003.

[32] Martin A Fischler and Robert C Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[33] Andrea Vedaldi and Brian Fulkerson, "VLFeat - An open and portable library of computer vision algorithms," in *MM*, 2010, pp. 1469–1472.