



# Person Re-identification with Patch-Based Local Sparse Matching and Metric Learning

Bo Jiang<sup>(✉)</sup>, Yibing Lv, Aihua Zheng, and Bin Luo

School of Computer Science and Technology, Anhui University, Hefei, China  
{jiangbo, ahzheng214, luobin}@ahu.edu.cn, yibing95@foxmail.com

**Abstract.** Recently, patch based matching has been demonstrated effectively to address the spatial misalignment issue caused by camera-view changes or human pose variations in person re-identification (Re-ID) problem. In this paper, we propose a novel local sparse matching model to obtain a reliable patch-wise matching for Re-ID problem. In particular, in the training phase, we develop a robust Local Sparse Matching model to learn more precise corresponding relationship between patches of positive sample image pairs. In the testing phase, we adopt a local-global distance metric learning for Re-ID task by considering global and local information simultaneously. Extensive experiments on four benchmarks demonstrate the effectiveness of our approach.

**Keywords:** Person re-identification · Graph matching · Metric learning

## 1 Introduction

Person re-identification (Re-ID) is an active research problem in computer vision and visual surveillance. The aim of Re-ID is to identify a specific probe person image from a set of gallery images captured from cross-view cameras. Many existing methods [12, 15, 28, 30] generally first extract a kind of global feature representation for person images and then utilize some metric learning methods to conduct a holistic comparison between test images for Re-ID.

One main challenge for Re-ID is to deal with the misalignment issue between image pair caused by large variations in camera views or human poses. Obviously, traditional global-based methods generally ignore spatial misalignment. To alleviate this issue, one kind of popular ways is to use part (or patch) based metric learning methods [16, 20, 23, 29, 33, 36]. These methods generally first partition each person image into a set of local patches. Then, they aim to conduct online patch-wise matching to obtain the spatial correspondences between patches of different images. Finally, the computed patch-wise matching is combined with local patch features to generate a robust metric learning for person Re-ID. However, one main issue for this online patch-wise matching is that it may lead to some mismatching among patches due to (1) lacking spatial and visual context information among local patches and (2) existing similar patch appearances or occlusions.

To overcome this issue, recent works [16, 23, 36] propose to develop some matching learning strategies for Re-ID. These methods generally first obtain a kind of reliable patch-wise matchings between training images in the training phase. Then, these

learned matchings are utilized or transferred to guide the robust patch-wise matching between test images in the testing phase.

However, the graph matching methods they used generally do not explicitly consider the one-to-one matching constraint and the impact of outlier patches which may lead to inaccurate correspondence relationships. In this paper, we propose a novel patch graph matching model for Re-ID problem. The aim of the proposed matching model is that it can obtain a robust one-to-one matching solution for the patches of two images. In particular, in the training phase, we first use the proposed matching model to learn an optimal correspondence relationship between positive sample pairs. In the testing phase, we then select the former  $R$  pairs references by pose pairs similarity, and use their correspondence relationship for the new testing image pair. Finally, we adopt the local-global distance metric for Re-ID problem. Overall, this paper makes the following contributions.

- In order to reduce the impact of outliers and obtain robust patch-wise correspondence relationships, we propose a novel graph matching model to make the spatial misalignment problem better solved.
- We propose a novel person Re-ID approach by employing both visual context information and spatial correspondence relationship learning simultaneously to avoid the limitations caused by local information and misalignment issues.
- Experimental results demonstrate that the proposed Re-ID method outperforms the other state-of-the-art approaches, validating the effectiveness of the method.

## 2 Related Work

Here, we briefly review some related works that are devoted to spatial misalignment for person Re-ID. Oreifej et al. [20] propose to utilize Earth Movers Distance (EMD) to obtain the whole similarity based on similarities between extracted patches. However, it ignores the spatial context of patches which may lead to mismatching for the patches with similar appearance or occlusions. Cheng et al. [9] propose to alleviate the influence of misalignment based on body part detection. The effectiveness of this approach generally relies on the detection result of the body part, which may be less effective in presence of occlusion. Some recent works [2, 17, 33] also propose to explore saliency or body prior to guiding the patch-wise matching between image pair.

One main limitation for the above online matching is that it may lead to some mismatching among patches due to (1) lacking using spatial context information among patches and (2) similar patch appearances or occlusions. To alleviate this limitation, Zhou et al. [36] recently propose to use a graph matching technique to obtain the optimal correspondence between each image pair during the training phase. Then, it transfers the learned patch-wise correspondence directly to the test image pair based on pose-pair configuration. This approach generally relies on image-level matching results obtained in the training phase. Lin et al. [16, 23] propose to learn a correspondence structure via a boosting based approach for each camera (pose) pair in the training phase. The learned correspondence structure is then utilized to guide the robust patch matching between test images. However, this method lacks considering spatial context

information of patches in the matching process, which may be less effective in the presence of similar patch appearances.

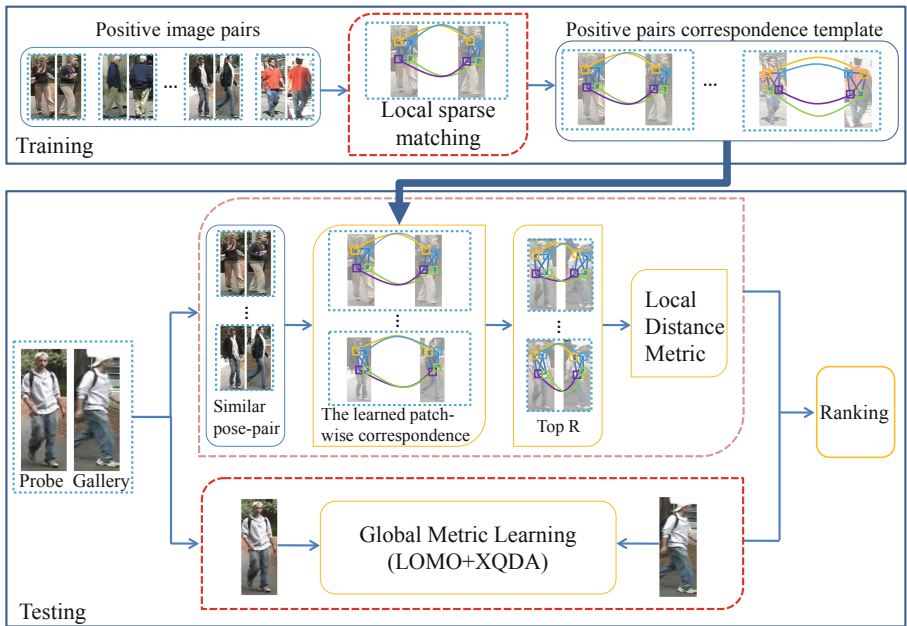


Fig. 1. Framework of the proposed approach.

### 3 The Proposed Model

In this section, we propose our patch matching model, followed by an effective update algorithm to compute it. We present our complete Re-ID approach in Sect. 4.

#### 3.1 Model Formulation

Given a positive image pair  $I$  and  $I'$ , we first divide them into several overlapping patches  $P = (p_1, p_2 \dots p_n)$  and  $P' = (p'_1, p'_2 \dots p'_m)$ , respectively. Then, we extract feature descriptor for each patch of the image. Our aim is to find the correspondence relationship between patches of two images. In order to do so, we construct an attributed relation graph  $G = (V, E, A, R)$  for image  $I$ , where nodes  $V$  represent patches  $P$  and edges  $E$  denote the relationship among patches. Each node  $v_i \in V$  has an associated attribute vector  $\mathbf{a}_i \in A$  and each edge  $e_{ih} \in E$  has a weight value  $\mathbf{r}_{ih} \in R$ . Similarly, we can construct a graph  $G' = (V', E', A', R')$  for  $I'$ . Based on this graph representation, the above patch matching problem can be reformulated as finding the correspondences between nodes of two graphs. Let  $\mathbf{Z} \in \{0, 1\}^{n \times m}$  denote the correspondence

solution between two graphs, in which  $\mathbf{Z}_{ij} = 1$  implies that node  $v_i \in G$  corresponds to node  $v'_j \in G'$ , and  $\mathbf{Z}_{ij} = 0$  otherwise. To obtain the optimal  $\mathbf{Z}$ , we define an affinity matrix  $\mathbf{K}$ . The diagonal term  $\mathbf{K}_{ij,ij}$  of  $\mathbf{K}$  represents the unary affinity  $f_a(\mathbf{a}_i, \mathbf{a}_j)$  that measures how well node  $v_i \in V$  matches node  $v'_j \in V'$ . The non-diagonal element  $\mathbf{K}_{ij,hk}$  contains the pair-wise affinity  $f_r(\mathbf{r}_{ih}, \mathbf{r}_{jk})$  that measures how compatible the nodes  $(v_i, v_h)$  in  $G$  are with the nodes  $(v'_j, v'_k)$  in  $G'$ . We can obtain the optimal  $\mathbf{Z}$  by optimizing the following objective function,

$$\begin{aligned} \max_{\mathbf{Z}} \quad & Q_{\text{IQP}} = \text{vec}(\mathbf{Z})^T \mathbf{K} \text{vec}(\mathbf{Z}), \\ \text{s.t.} \quad & \forall i \sum_{j=1}^n \mathbf{Z}_{ij} = 1, \forall j \sum_{i=1}^m \mathbf{Z}_{ij} \leq 1, \mathbf{z}_{ij} \in \{0, 1\} \end{aligned} \quad (1)$$

It is known that the above problem is an Quadratic Assignment Problem (QAP) which is a NP-hard problem. Therefore, relaxation models are required to find some approximate solutions. For the person image matching problem, an ideal matching relaxation model should be satisfied with the following two aspects. (1) An one-to-one matching constraint should be imposed in the final matching results, i.e., each patch in image  $I$  should correspond to at most one patch in  $I'$ . (2) There may exist outlier patches in both  $I$  and  $I'$ . Thus, the matching process should perform robustly to the outlier patches. To address these issues, we propose to obtain the optimal matching  $\mathbf{Z}$  by solving the following novel sparse relaxation matching problem,

$$\max_{\mathbf{Z}} \quad \text{vec}(\mathbf{Z})^T \mathbf{K} \text{vec}(\mathbf{Z}) - \beta \|\mathbf{Z}\|_2^2 \quad \text{s.t.} \quad \|\mathbf{Z}\|_{1,2} = 1, \mathbf{Z} \geq 0 \quad (2)$$

where  $\|\mathbf{Z}\|_{1,2} = (\sum_i (\sum_j |\mathbf{Z}_{ij}|)^2)^{1/2}$  is used to encourage local sparse and thus one-to-one matching constraint [3]. The  $\ell_2$ -norm regularization term is used to control the compactness of all inlier, which make the model perform robustly to the outlier patches, as discussed in work [26].

### 3.2 Computational Algorithm

The proposed patch matching model can be solved effectively via a simple multiplicative update algorithm. Starting from  $\mathbf{Z}^{(0)}$ , the proposed algorithm conducts the following update until convergence.

$$\mathbf{z}_{ij}^{(t+1)} = \mathbf{z}_{ij}^{(t)} \sqrt{\frac{\mathbf{M}_{ij}^{(t)}}{\lambda \sum_j \mathbf{z}_{ij}^{(t)}} + \beta \mathbf{z}_{ij}^{(t)}} \quad (3)$$

where matrix  $\mathbf{M}^{(t)} \in \mathbb{R}^{m \times n}$  is the matrix form of the vector  $[\mathbf{K}^{(t)} \text{vec}(\mathbf{Z}^{(t)})]$ , and  $\lambda$  is computed as,

$$\lambda = \text{vec}(\mathbf{Z}^{(t)})^T \mathbf{K} \text{vec}(\mathbf{Z}^{(t)}) - \beta \text{vec}(\mathbf{Z}^{(t)}) \quad (4)$$

**Theoretical Analysis.** The optimality and convergence of the algorithm are guaranteed by Theorems 1 and 2, respectively.

**Theorem 1.** *Update rule of Eq. (3) satisfies the first-order Karush-Kuhn-Tucker (KKT) optimality condition.*

**Theorem 2.** *Under the update rule Eq. (3), the Lagrangian function  $\mathcal{L}(\mathbf{X})$  Eq. (5) is monotonically increasing.*

$$\mathcal{L}(\mathbf{Z}) = \text{vec}(\mathbf{Z})^T \mathbf{K} \text{vec}(\mathbf{Z}) - \beta \|\mathbf{Z}\|_2^2 - \lambda \left( \sum_i \left( \sum_j \mathbf{Z}_{ij} \right)^2 - 1 \right) \quad (5)$$

The proof of them can be similarly derived from work [3] which is omitted here due to limited space.

## 4 Person Re-ID

In this section, we describe our Re-ID approach based on the proposed patch matching model. The complete process is shown in Fig. 1.

### 4.1 Training Stage

Given a positive image pair  $I_p$  and  $I_g$ , we decompose them into many overlapping patches at the start. Then, we construct a graph for these patch features and learn the patch-wise correspondence  $\mathbf{Z}$  of them via the proposed matching model. Of course, the part of graph matching can also be replaced by some classic graph matching methods such as [4, 11].

In Re-ID,  $\mathbf{Z}_{ij} = 1$  means the  $i^{\text{th}}$  patch in  $I_p$  semantically corresponds to the  $j^{\text{th}}$  patch in  $I_g$ . And the graph matching model was detailed in the previous section. In this stage, we can obtain satisfactory patch-wise matching results for Re-ID, and these correspondence relationships will be used to distance measure.

### 4.2 Testing Stage

We use the local-global pattern for the final distance metric learning due to local information is one-sided and global information is easily lead to misalignment. We compute the final distance  $D$  as follows,

$$D(I'_p, I'_g) = \alpha D_l + (1 - \alpha) D_g \quad (6)$$

where  $D_g$  and  $D_l$  represent global and local distance between test images  $I'_g$  and  $I'_p$  respectively, and  $\alpha$  is a balance parameter.

**Local Distance Metric.** When we have learned the patch-wise correspondence relationships of all trained positive sample pairs from the training stage, we can get the most similar  $R$  pairs of positive references to test pair by comparing the pose similarity. The local distance compute is as follows,

$$D_l(I'_p, I'_g) = \sum_{i=1}^R w_i \sum_{j=1}^{S_i} \zeta(f'_p, f'_g) \quad (7)$$

where  $f'_p$  and  $f'_g$  represent features of patches in probe image  $I'_p$  and gallery image  $I'_g$ , and  $\zeta(\cdot)$  denotes the KISSME metric [12]. We use  $\Phi = \{\phi_i\}_{i=1}^R$  to represent the  $R$  templates selected, where each template  $\phi_i = \{z_{ij}\}_{j=1}^{S_i}$  contains a total of  $S_i$  patch-wise correspondences, and each correspondence  $z_{ij}$  denotes the positions of matched patches calculated by  $\mathbf{Z}_{ij}$ , which is consistent with the original method [36]. The difference with the method lies in that we normalized weighted and summed the selected  $R$  pairs of references, so that the more similar the pose is, the higher the weight is, where the weight is represented by  $w_i$ .

**Global Distance Metric.** For the test pair  $I'_p$  and  $I'_g$ , we adopt the LOMO+XQDA [15] to supplement the global distance compute. We combine global information into patch-wise feature distances between each correspondence of selected references, so we can calculate the local and global distances between the test image pairs. In all our experiments, we use Local Maximal Occurrence features [15] for the patch features representation, whether local or global.

## 5 Experiments

### 5.1 Datasets

**VIPeR Dataset:** The VIPeR dataset [10] includes 1264 images of 632 pedestrians, and each pedestrian has two images collected from cameras A and B. Each image is adjusted to the size of  $128 \times 48$ . The dataset is characterized by a diversity of perspectives and lighting.

**Road Dataset:** This dataset [23] is captured from a crowd road scene by two cameras and consists of 416 image pairs. It is very challenging due to the large variation of human pose and camera view.

**PRID450S Dataset:** This dataset [22] consists of 450 image pairs from two camera views. The low image qualities and camera viewpoint changes make it very challenging for person re-identification.

**CUHK01 Dataset:** This dataset [13] consists of 971 individuals captured from two disjoint camera views. The images on this dataset are of higher resolutions. We also adopt the commonly utilized 485/486 setting for person re-identification evaluation.

### 5.2 Evaluation Settings

**Parameter Setup.** We follow previous methods and perform experiments under the half-training and half-testing setting. All images are scaled to  $128 \times 48$ . The patch size is set to  $32 \times 24$ . The stride size between neighboring patches is 6 horizontally and 8 vertically for probe images and gallery images. More specifically, for each stripe in the probe image, patch-wise correspondences are established between the corresponding gallery stripe within the search range in the gallery image.

**Evaluation.** On all datasets, both the training/testing set partition and probe/gallery set partition are performed 10 times and average performance is reported. The performance

is evaluated by using the Cumulative Matching Characteristic (CMC) curve, which represents the expected probability of finding the correct match for a probe image in the top  $r$  matches in the gallery list [36]. Tables 1, 2, 3 and 4 show the CMC results of different methods on four datasets and Fig. 2 shows the CMC curves on three datasets.

**Table 1.** Comparisons of top  $r$  matching rate using CMC on VIPeR dataset. The best results are marked in bold.

Methods	Rank = 1	Rank = 5	Rank = 10	Rank = 20	Reference
KISSME	27.3%	55.3%	69.0%	82.7%	CVPR2012
SVMML	30.0%	64.7%	79.0%	91.3%	CVPR2013
SalMatch	30.2%	52.3%	65.5%	79.2%	T-PAMI2017
ELS	31.3%	62.1%	75.3%	86.7%	TIP2015
KLFDA	32.4%	65.9%	79.8%	90.8%	ECCV2014
CSL	34.8%	68.7%	82.3%	91.8%	ICCV2015
IDLA	34.8%	64.5%	-	-	CVPR2015
JLR	35.8%	68.0%	-	-	CVPR2016
LOMO + XQDA	40.0%	68.1%	80.5%	91.1%	CVPR2015
Semantic	41.6%	71.9%	86.2%	95.1%	CVPR2015
Single-KMFA	43.3%	77.4%	88.2%	92.7%	TIP2017
LMF + LADF	43.4%	73.0%	84.9%	93.7%	CVPR2014
DCSL	44.6%	73.4%	82.6%	-	IJCAI2016
TCP	47.8%	74.7%	84.8%	91.1%	CVPR2016
TMA	48.2%	-	87.7%	95.5%	ECCV2016
GCT	49.4%	77.6%	87.2%	94.0%	AAAI2018
LOMO-fusing	51.2%	82.1%	90.5%	95.9%	CVPR2016
Multi-mamu-KMFA	51.4%	82.5%	91.2%	<b>96.0%</b>	TIP2017
<b>Our</b>	<b>59.7%</b>	<b>83.2%</b>	<b>91.4%</b>	<b>96.0%</b>	

**Table 2.** Comparisons of top  $r$  matching rate using CMC on Road dataset. The best results are marked in bold.

Methods	Rank = 1	Rank = 5	Rank = 10	Rank = 20	Reference
eSDC-knn	52.4%	74.5%	83.7%	89.9%	CVPR2013
SVMML	57.2%	85.2%	92.1%	96.1%	CVPR2013
MFA	58.3%	85.5%	92.4%	96.2%	ECCV2014
kLFDA	59.1%	86.5%	91.8%	95.8%	ECCV2014
CSL	61.5%	91.8%	95.2%	98.6%	ICCV2015
Single-KMFA	71.4%	92.5%	95.6%	97.1%	TIP2017
Multi-manu-KMFA	73.1%	92.8%	96.4%	97.8%	TIP2017
GCT	88.8%	<b>96.7%</b>	98.4%	99.6%	AAAI2018
<b>Our</b>	<b>88.9%</b>	96.6%	<b>99.0%</b>	<b>100%</b>	

### 5.3 Results

To evaluate the effectiveness of the proposed method, we evaluate the proposed Re-ID method by comparing it with some other methods including KISSME [12], SVMML [14], SalMatch [35], ELS [6], MFA [28], kLFDA [28], IDLA [1], JLR [27], LOMO+XQDA [15], Semantic [25], LMF+LADF [34], DCSL [32], deCPPs+MER [24], TCP [9], TMA [18], LOMO-fusing [31], SCNCD [30], Mirror-KMFA [8], eSDC-knn [33], CSL [23], single-KMFA [16], multi-manu-KMFA [16], DeepRanking [7], ME [21], GOG [19], SalMatch [35], CSBT [5], and GCT [36].

**Table 3.** Comparisons of top  $r$  matching rate using CMC on PRID450S dataset. The best results are marked in bold.

Methods	Rank = 1	Rank = 5	Rank = 10	Rank = 20	Reference
KISSME	33.0%	-	71.0%	79.0%	CVPR2012
SVMML	42.8%	69.7%	79.2%	86.6%	CVPR2013
CSL	44.4%	71.6%	82.2%	89.8%	ICCV2015
MFA	44.7%	71.9%	80.2%	86.9%	ECCV2014
Semantic	44.9%	71.7%	77.5%	86.7%	CVPR2015
kLFDA	46.9%	73.1%	81.7%	87.2%	ECCV2014
TMA	54.2%	73.8%	83.1%	90.2%	ECCV2016
Mirror-KMFA	55.4%	79.3%	87.8%	93.9%	IJCAI2015
GCT	58.4%	77.6%	84.3%	89.8%	AAAI2018
Single-KMFA	63.2%	82.7%	90.5%	94.8%	TIP2017
Multi-manu-KMFA	65.1%	84.9%	91.2%	95.4%	TIP2017
<b>Our</b>	<b>75.6%</b>	<b>91.2%</b>	<b>95.1%</b>	<b>97.5%</b>	

**Table 4.** Comparisons of top  $r$  matching rate using CMC on CUHK01 dataset. The best results are marked in bold.

Methods	Rank = 1	Rank = 5	Rank = 10	Rank = 20	Reference
Semantic	32.7%	51.2%	-	76.3%	CVPR2015
kLFDA	32.8%	59.0%	69.6%	-	ECCV2014
IDLA	47.5%	71.5%	80.0%	-	CVPR2015
DeepRanking	50.4%	75.9%	84.1%	-	TIP2016
ME	53.4%	76.3%	84.4%	-	CVPR2015
GOG	57.8%	79.1%	86.2%	-	CVPR2016
SalMatch	28.5%	46.0%	-	67.3%	T-PAMI2017
CSBT	51.2%	76.3%	-	91.8%	CVPR2017
TCP	53.7%	84.3%	91.0%	96.3%	CVPR2016
GCT	61.9%	81.9%	87.6%	92.8%	AAAI2018
<b>Our</b>	<b>82.7%</b>	<b>93.6%</b>	<b>96.9%</b>	<b>98.6%</b>	



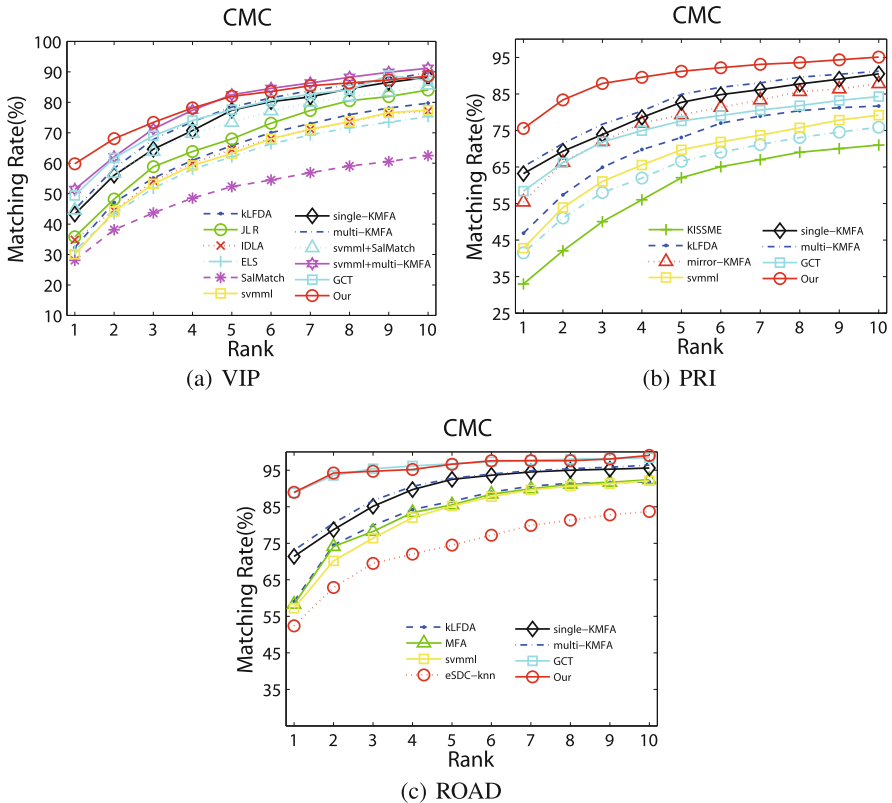


Fig. 2. CMC scores comparison on different datasets.

Tables 1, 2, 3 and 4 summarize the comparison results. Here, we can note: (1) Our approach performs better than graph correspondence transfer (GCT) [36], which demonstrates the effectiveness and robustness of the proposed patch matching model by reducing the impact of outliers. (2) Our method also outperforms many other Re-ID methods and obtains the best performance on all datasets, which indicates the effectiveness of the proposed Re-ID approach.

## 6 Conclusion

In this paper, we propose a new model to solve the problem of cross-view spatial misalignment of person Re-id. We first propose a novel local sparse matching model to learn the corresponding relationship between patches of image pair in the training stage. Then, in the testing phase, we adopt local-global distance measure method to make the learning result of person images more accurate. Extensive experimental results on several benchmarks demonstrate the effectiveness of the proposed Re-ID method.

**Acknowledgment.** This work was supported by the National Natural Science Foundation of China (61602001), Natural Science Foundation of Anhui Province (1708085QF139), Open Project Program of the National Laboratory of Pattern Recognition (NLPR) (201900046).

## References

1. Ahmed, E., Jones, M., Marks, T.K.: An improved deep learning architecture for person re-identification. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 3908–3916 (2015)
2. Bak, S., Carr, P.: Person re-identification using deformable patch metric learning. In: *International Workshop on Applications of Computer Vision*, pp. 1–9 (2016)
3. Jiang, B., Tang, J., Ding, C., Luo, B.: A local sparse model for matching problem. In: *American Association for Artificial Intelligence*, pp. 3790–3796 (2015)
4. Jiang, B., Tang, J., Ding, C., Luo, B.: Binary constraint preserving graph matching. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 4402–4409 (2017)
5. Chen, J., Wang, Y., Qin, J., Liu, L., Shao, L.: Fast person re-identification via cross-camera semantic binary transformation. In: *Proceedings of Computer Vision and Pattern Recognition*, p. 1 (2017)
6. Chen, J., Zhang, Z., Wang, Y.: Relevance metric learning for person re-identification by exploiting listwise similarities. *IEEE Trans. Image Process.* **24**(12), 4741–4755 (2015)
7. Chen, S.Z., Guo, C.C., Lai, J.H.: Deep ranking for person re-identification via joint representation learning. *IEEE Trans. Image Process.* **25**(5), 2353–2367 (2016)
8. Chen, Y.C., Zheng, W.S., Lai, J.: Mirror representation for modeling view-specific transform in person re-identification. In: *International Joint Conference on Artificial Intelligence*, pp. 3402–3408 (2015)
9. Cheng, D., Gong, Y., Zhou, S., Wang, J., Zheng, N.: Person re-identification by multi-channel parts-based cnn with improved triplet loss function. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 1335–1344 (2016)
10. Gray, D., Brennan, S., Tao, H.: Evaluating appearance models for recognition, reacquisition, and tracking. In: *Proceedings of International Workshop on Performance Evaluation for Tracking and Surveillance*, pp. 1–7 (2007)
11. Jiang, B., Tang, J., Ding, C., Gong, Y., Luo, B.: Graph matching via multiplicative update algorithm. In: *Proceedings of Neural Information Processing Systems*, pp. 3187–3195 (2017)
12. Koestinger, M., Hirzer, M., Wohlhart, P., Roth, P.M., Bischof, H.: Large scale metric learning from equivalence constraints. In: *Computer Vision and Pattern Recognition*, pp. 2288–2295 (2012)
13. Li, W., Zhao, R., Wang, X.: Human reidentification with transferred metric learning. In: Lee, K.M., Matsushita, Y., Rehg, J.M., Hu, Z. (eds.) *ACCV 2012*. LNCS, vol. 7724, pp. 31–44. Springer, Heidelberg (2013). [https://doi.org/10.1007/978-3-642-37331-2\\_3](https://doi.org/10.1007/978-3-642-37331-2_3)
14. Li, Z., Chang, S., Liang, F., Huang, T.S., Cao, L., Smith, J.R.: Learning locally-adaptive decision functions for person verification. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 3610–3617 (2013)
15. Liao, S., Hu, Y., Zhu, X., Li, S.Z.: Person re-identification by local maximal occurrence representation and metric learning. In: *Proceedings of Computer Vision and Pattern Recognition*, pp. 2197–2206 (2015)
16. Lin, W., et al.: Learning correspondence structures for person re-identification. *IEEE Trans. Image Process.* **26**(5), 2438–2453 (2017)
17. Ma, L., Yang, X., Xu, Y., Zhu, J.: A generalized emd with body prior for pedestrian identification. *J. Vis. Commun. Image Represent.* **24**(6), 708–716 (2013)

18. Martinel, N., Das, A., Micheloni, C., Roy-Chowdhury, A.K.: Temporal model adaptation for person re-identification. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9908, pp. 858–877. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-46493-0\\_52](https://doi.org/10.1007/978-3-319-46493-0_52)
19. Matsukawa, T., Okabe, T., Suzuki, E., Sato, Y.: Hierarchical gaussian descriptor for person re-identification. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1363–1372 (2016)
20. Oreifej, O., Mehran, R., Shah, M.: Human identity recognition in aerial images. In: Computer Vision and Pattern Recognition, pp. 709–716 (2010)
21. Paisitkriangkrai, S., Shen, C., Van Den Hengel, A.: Learning to rank in person re-identification with metric ensembles. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1846–1855 (2015)
22. Roth, P.M., Hirzer, M., Köstinger, M., Beleznai, C., Bischof, H.: Mahalanobis distance learning for person re-identification. In: Gong, S., Cristani, M., Yan, S., Loy, C. (eds.) Person Re-Identification. ACVPR, pp. 247–267. Springer, London (2014). [https://doi.org/10.1007/978-1-4471-6296-4\\_12](https://doi.org/10.1007/978-1-4471-6296-4_12)
23. Shen, Y., Lin, W., Yan, J., Xu, M., Wu, J., Wang, J.: Person re-identification with correspondence structure learning. In: Proceedings of International Conference on Computer Vision, pp. 3200–3208 (2015)
24. Sheng, H., Huang, Y., Zheng, Y., Chen, J., Xiong, Z.: Person re-identification via learning visual similarity on corresponding patch pairs. In: Zhang, S., Wirsing, M., Zhang, Z. (eds.) KSEM 2015. LNCS (LNAI), vol. 9403, pp. 787–798. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-25159-2\\_73](https://doi.org/10.1007/978-3-319-25159-2_73)
25. Shi, Z., Hospedales, T.M., Xiang, T.: Transferring a semantic representation for person re-identification and search. In: Proceedings of Computer Vision and Pattern Recognition, pp. 4184–4193 (2015)
26. Suh, Y., Adamczewski, K., Lee, K.M.: Subgraph matching using compactness prior for robust feature correspondence. In: Proceedings of Computer Vision and Pattern Recognition, p. 1 (2015)
27. Wang, F., Zuo, W., Lin, L., Zhang, D., Zhang, L.: Joint learning of single-image and cross-image representations for person re-identification. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1288–1296 (2016)
28. Xiong, F., Gou, M., Camps, O., Sznai, M.: Person re-identification using kernel-based metric learning methods. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8695, pp. 1–16. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-10584-0\\_1](https://doi.org/10.1007/978-3-319-10584-0_1)
29. Yang, Y., Wen, L., Lyu, S., Li, S.Z.: Unsupervised learning of multi-level descriptors for person re-identification. In: American Association for Artificial Intelligence, p. 1 (2017)
30. Yang, Y., Yang, J., Yan, J., Liao, S., Yi, D., Li, S.Z.: Salient color names for person re-identification. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014. LNCS, vol. 8689, pp. 536–551. Springer, Cham (2014). [https://doi.org/10.1007/978-3-319-10590-1\\_35](https://doi.org/10.1007/978-3-319-10590-1_35)
31. Zhang, L., Xiang, T., Gong, S.: Learning a discriminative null space for person re-identification. In: Proceedings of Computer Vision and Pattern Recognition, pp. 1239–1248 (2016)
32. Zhang, Y., Li, X., Zhao, L., Zhang, Z.: Semantics-aware deep correspondence structure learning for robust person re-identification. In: International Joint Conference on Artificial Intelligence, pp. 3545–3551 (2016)
33. Zhao, R., Ouyang, W., Wang, X.: Unsupervised salience learning for person re-identification. In: Proceedings of Computer Vision and Pattern Recognition, pp. 3586–3593 (2013)

34. Zhao, R., Ouyang, W., Wang, X.: Learning mid-level filters for person re-identification. In: Proceedings of Computer Vision and Pattern Recognition, pp. 144–151 (2014)
35. Zhao, R., Oyang, W., Wang, X.: Person re-identification by saliency learning. IEEE Trans. Pattern Anal. Mach. Intell. **39**(2), 356–370 (2017)
36. Zhou, Q., et al.: Graph correspondence transfer for person re-identification, p. 1. arXiv preprint. [arXiv:1804.00242](https://arxiv.org/abs/1804.00242) (2018)