

# Towards More Reliable Matching for Person Re-identification

Xiang Li<sup>1</sup>, Ancong Wu<sup>1</sup>, Mei Cao<sup>2</sup>, Jinjie You<sup>1,3</sup>, and Wei-Shi Zheng<sup>\*1</sup>

<sup>1</sup>School of Information Science and Technology, Sun Yat-sen University, China

<sup>2</sup>School of Finance and Statistics, East China Normal University, China

<sup>3</sup>School of Mechanical, Electrical and Information Engineering, Shandong University, China

lixiang651@gmail.com, caomeibb0406@gmail.com, wuancong@mail2.sysu.edu.cn,  
youjinjie9@gmail.com, wszheng@ieee.org

## Abstract

Person re-identification is an important problem of matching persons across non-overlapping camera views. However, the re-identification is still far from achieving reliable matching. First, many existing approaches are whole-body-based matching, and how body parts could affect and assist the matching is still not clearly known. Second, the learned similarity measurement/metric is equally used for each pair of probe and gallery images, and the bias of the measurement is not considered. In this paper, we address the above two problems in order to conduct a more reliable matching. More specifically, we propose a reliable integrated matching scheme (IMS), which uses body parts to assist matching of the whole body. Moreover, a sparsity-based confidence is also presented for regulating the learned metric to improve the matching reliability. The experiments conducted on three publicly available datasets confirm that the proposed scheme is effective for person re-identification.

## 1. Introduction

Person re-identification has been seen a lot of developments in recent years. To address the re-identification problem of finding the correct matching (the same person/class) for an unknown probe person image among a large number of gallery images captured from different camera views, many existing approaches use images of the whole body to match. These methods exploit either cross-view invariant features [4, 5, 21, 3, 13, 7, 9, 28, 24] or cross-view robust metrics [17, 6, 15, 30, 19, 16, 11, 10, 25, 14, 12].

For these practices, an intuitive doubt is that – is using whole body always reliable for matching in person re-identification? In a surveillance system, the inter-class variation of persons varies drastically and there are often severe

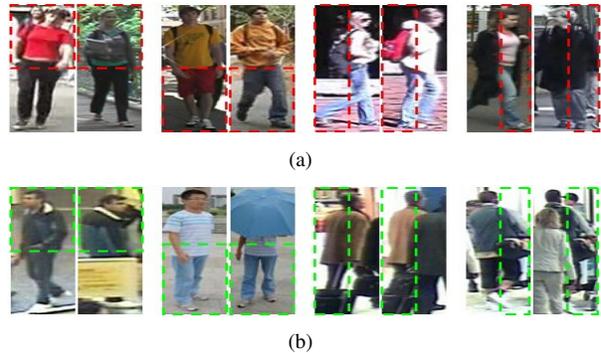


Figure 1: Illustration of person images in a surveillance system. (a) shows the obvious discriminative body parts with red dotted boxes between images of different persons, and (b) shows the obvious similar body parts with green dotted boxes between images of the same person.

overlaps between classes in a feature space due to similar appearance (e.g. clothing) of different people. In Fig. 1(a), it is clear that one body part with a highlighted red dotted box (e.g. upper-/lower-/left-/right- body) certainly has more discriminative information than the whole body for distinguishing different persons. Meanwhile, the intra-class variation can also be large, caused by large and unpredictable condition changes. From Fig. 1(b), we can observe that one body part with a highlighted green dotted box is also more effective than the whole body for matching persons. Therefore, we believe that some body parts sometimes are more reliable than whole body.

Another doubt is that – is the learned metric which is used equally for all the probe images after training always reliable for matching in person re-identification? In an uncontrolled environment, different persons/classes have different large visual appearance changes. This may lead to an overfitting problem during training process. An illustration of person re-identification using a distance metric learned by LFDA [16] is shown in Fig. 2. Since the metric treats

\*Corresponding Author.



Figure 2: Illustration of person re-identification using the metric learned by LFDA [16]. The left-most is the probe image; images in the middle are the top 7 matched gallery images, with a red box highlighting the correct match, and the right-most shows the ground truth.

all matches between probe image and each gallery image equally, the correct matching in gallery is not the top 1 rank for the probe person, even though the visual appearance of them are similar. This shows that the similarity measurement sometimes is not reliable, and therefore one needs to take the confidence of the similarity measurement during the matching.

In this work, we investigate how to conduct a more reliable matching in person re-identification. Firstly, rather than only utilizing the whole body to match, we propose exploiting body parts to assist the whole body to re-identify a probe image. Secondly, for each pair of probe image and gallery image, we compute the confidence of each pair to match in order to reduce the bias of the similarity measurement learned from training data. For this purpose, a sparsity-based confidence is introduced to regulate the metric. Finally, we apply the two reliable matching strategies to re-identification and propose an integrated matching scheme (IMS). This scheme consists of two components: whole body matching and body part matching, and each component uses the confidence to improve the reliability.

Several existing re-identification methods also consider using body parts. The pictorial model is a typical part-body-based model for part-to-part matching in person re-identification [2]. Zhao et al. also utilize body parts to match by salience learning [27, 26]. However, these models only rely on body parts for re-identification, and they neglect the effect of whole body in matching.

Extensive experiments are conducted on three person re-identification datasets, including VIPeR [5], i-LIDS [20, 29] and 3DPES [1]. The results demonstrate that the proposed scheme is much more effective for re-identification.

## 2. Proposed Scheme

### 2.1. Body Parts Assist Whole Body to Match

The problem of person re-identification is that a probe person image is required to be matched with gallery person images. To address this problem, we first consider utilizing body parts to enhance re-identification. The basic idea is that, some body parts may contain discriminative details for each individual, which are more reliable than the whole

body to match. For each unknown probe person image, it is both costly and difficult to search the discriminative body parts precisely. Therefore, these body parts are selected by an overall evaluation in this paper (see details in Sec. 3.3).

One is able to use the body parts to assist the whole body to determine the identity of an unknown person. The effectiveness of the body parts is demonstrated in our experiments. The combination of the discriminative body parts and whole body achieves higher matching rate than using body parts and whole body separately, for details please see Sec.3.5. The discriminative body parts are less affected by various changes including person-specific and environment-specific changes, but they contain less information than the whole body. In contrast, using the whole body enforces spatial layout consistency, but it suffers greatly from un-controlled condition. Therefore, they are complementary to each other, and the fusion of them is effective for matching.

### 2.2. Confidence Regulates Metric to Match

In person re-identification, different persons have different large visual appearance changes. This may lead to an overfitting problem during the training process, and thus the learned metric sometimes is not reliable for re-identification. In this section, for a probe image (no matter using whole body or body parts), we consider that each person/class in gallery should have a confidence to estimate the probability that how likely the probe image belongs to this class. So the confidence can be used to regulate the metric in order to make it more reliable. Wright et al. [23] show that representing a probe image by a sparse linear combination of gallery images is very effective for classification. In this paper, we propose to apply this method to compute the confidence.

Firstly, we construct the gallery person dictionary. For each gallery image, the generic features are extracted (see Sec.3.2 for details). Suppose  $k_c$  person images, say,  $g_c^1, g_c^2, \dots, g_c^{k_c}$ , are obtained for class  $c$  in the gallery ( $c = 1, 2, \dots, C$ ). Then corresponding  $k_c$  feature vectors are denoted by  $\mathbf{x}_{g_c^1}, \mathbf{x}_{g_c^2}, \dots, \mathbf{x}_{g_c^{k_c}}$ , where each feature vector has the same dimensionality (assume  $M$ ). Let  $\mathbf{D}_c = [\mathbf{x}_{g_c^1}, \mathbf{x}_{g_c^2}, \dots, \mathbf{x}_{g_c^{k_c}}]$ . So the gallery person dictionary for all the  $C$  classes is built as

$$\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_C]. \quad (1)$$

Obviously,  $\mathbf{D}$  has a total of  $K = \sum_{c=1}^C k_c$  images, resulting in an  $M \times K$  dictionary.

Then, given a feature vector  $\mathbf{x}_p$  of a probe person image  $p$ , the sparse coding problem is formulated as the following  $L_1$ -minimization problem

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \|\alpha\|_1, \text{ s.t. } \mathbf{x}_p = \mathbf{D}\alpha, \quad (2)$$

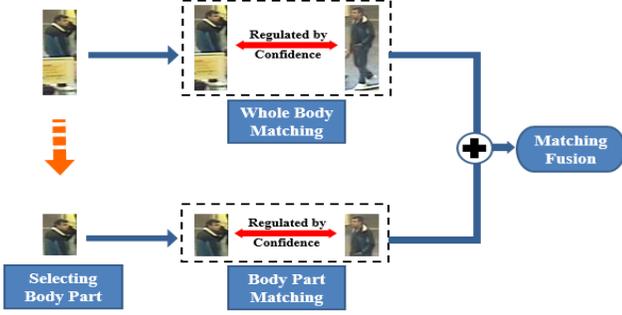


Figure 3: The proposed integrated matching scheme for person re-identification.

where  $\alpha \in R^K$  is the sparse coefficient vector,  $\|\cdot\|_1$  denotes the  $L_1$ -norm of a vector. Numerous efficient fast  $L_1$ -minimization algorithms can be used to solve Eq.(2). In this paper, we employ the feature-sign search algorithm [8] to obtain the sparse coefficient vector  $\alpha$ .

Finally, for the probe image  $p$ , the reconstruction residual of  $c$ -th class in the gallery is computed as follows

$$r_{p,c} = \|\mathbf{x}_p - \mathbf{D}_c \delta_c(\hat{\alpha})\|_2, \quad c = 1, 2, \dots, C, \quad (3)$$

where  $\delta_c$  is a function that selects the coefficients associated with the  $c$ -th class. We define the confidence by a Gaussian-shaped kernel  $G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{x^2}{2\sigma^2})$ , where  $\sigma$  is a bandwidth parameter. Specifically, for the probe image  $p$ , the confidence of a gallery image  $g_c$  of  $c$ -th class is computed as

$$s_{p,c} = G_\sigma(r_{p,c}), \quad c = 1, 2, \dots, C. \quad (4)$$

We call  $s_{p,c}$  as the sparsity-based confidence. A higher sparsity-based confidence means a higher similarity (a smaller distance). Therefore, we wish to obtain a reliable distance metric such that it is negatively correlated to the confidence. Given a feature vector  $\mathbf{x}_{g_c}$  of the gallery image  $g_c$  of the  $c$ -th class, suppose a distance metric is denoted as  $dist(\mathbf{x}_p, \mathbf{x}_{g_c})$  between the probe image  $p$  and the gallery image  $g_c$ , then the corresponding new and reliable distance is computed as  $\frac{dist(\mathbf{x}_p, \mathbf{x}_{g_c})}{s_{p,c}}$ .

### 2.3. The Integrated Matching Scheme

Two strategies towards reliable matching for person re-identification have been presented in the previous sections. In this section, an effective integrated matching scheme is developed, which consists of two matching components: the whole body matching and body part matching. Meanwhile, each component uses the sparsity-based confidence to improve the reliability of matching. Without loss of generality, we present a formulation which uses one body part to enhance re-identification (see Fig. 3). A multi-parts fusion formulation can be readily generalized.

In the whole body matching, for the probe image  $p$  and each gallery image  $g_c$  of  $c$ -th class ( $c = 1, 2, \dots, C$ ), we first obtain the whole body feature vectors  $\mathbf{x}_p^w$  and  $\mathbf{x}_{g_c}^w$ , respectively. Then we compute the whole body distance  $dist(\mathbf{x}_p^w, \mathbf{x}_{g_c}^w)$  (utilizing existing re-identification methods, e.g. LFDA [16]; seeing detail in Sec. 3.1) and the corresponding sparsity-based confidence  $s_{p,c}^w$ . Similarly, in the body part matching, we can also obtain the body part feature vectors  $\mathbf{x}_p^b$  and  $\mathbf{x}_{g_c}^b$ , and further compute the body part distance  $dist(\mathbf{x}_p^b, \mathbf{x}_{g_c}^b)$  and the confidence  $s_{p,c}^b$ . Finally, the integrated matching scheme is formulated as follows

$$dist_F(\{\mathbf{x}_p^w, \mathbf{x}_p^b\}, \{\mathbf{x}_{g_c}^w, \mathbf{x}_{g_c}^b\}) = \beta \frac{dist(\mathbf{x}_p^w, \mathbf{x}_{g_c}^w)}{s_{p,c}^w} + (1 - \beta) \frac{dist(\mathbf{x}_p^b, \mathbf{x}_{g_c}^b)}{s_{p,c}^b}, \quad c = 1, 2, \dots, C, \quad (5)$$

where  $dist_F(\{\mathbf{x}_p^w, \mathbf{x}_p^b\}, \{\mathbf{x}_{g_c}^w, \mathbf{x}_{g_c}^b\})$  is the fusion distance between the probe image  $p$  and the gallery image  $g_c$ , and  $\beta$  is a parameter for regulating the effect of two components. In our experiments, we denote the above scheme as IMS (integrated matching scheme) for convenience.

## 3. Experiments

### 3.1. Datasets and Settings

**Datasets.** Three publicly available person re-identification datasets, VIPeR [5], i-LIDS [20, 29] and 3DPES [1] were used for evaluation. The VIPeR dataset is composed of 1264 images of 632 individuals, with 2 images per individual. View angle change was the most significant cause of appearance change (see Fig. 4(a)). In the i-LIDS dataset, which was captured indoor in a busy airport arrival hall, there are 119 people with total 476 images captured by multiple non-overlapping cameras. The images have fair amount of occlusion caused by people and luggage (see Fig. 4(b)). The 3DPES dataset includes 1011 images of 192 individuals captured from 8 outdoor cameras with significantly different viewpoints. All the images were collected during various times of the day, resulting in strong variations of lighting conditions (see Fig. 4(c)).

**Settings.** In our experiments, all datasets were randomly divided into training set and testing set by half so that there are 316, 60 and 96 individuals in the testing set of VIPeR, i-LIDS and 3DPES respectively. Each testing set was composed of a gallery set and a probe set. Following the most popular testing protocol [30], the gallery set consisted of one image for each person, and the remaining images were used as the probe set. This procedure was repeated 10 times.

**Evaluation.** To evaluate our scheme, we first focused on a representative distance metric in IMS. The metric was defined by local fisher discriminant analysis (LFDA) [18] which has been proven to be useful for re-identification



Figure 4: Illustration of person images on three datasets.

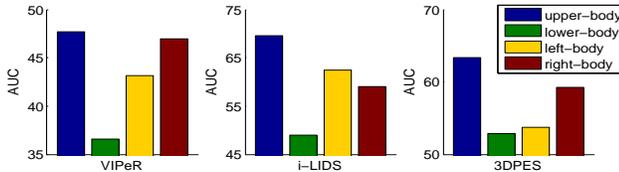


Figure 5: Performance using different body parts.

[16]. We called such a model as IMS-LFDA in our experiments. We compared IMS-LFDA with LFDA and other five existing re-identification methods, that is, two distance learning methods LMNN [22] and KISSME [6], a non-learning distance based L1-norm, a ranking based model RDC [30], and a part-body-based salience learning method SDC [27]. In addition, an effective face recognition model SRC [23] was also compared. We would also apply the proposed IMS to two other metrics (i.e. RDC and L1-norm) for full comparison in Sec. 3.5. We used the average cumulative match characteristic (CMC) curves over 10 trials to show the ranked matching rates. A rank  $r$  matching rate indicates the percentage of the probe images with correct matches found in top  $r$  ranks against the gallery images.

### 3.2. Feature Representation

In our experiments, all whole body images were normalized to  $128 \times 48$  pixels according to the common practice. We obtained overlapping patches of size  $16 \times 16$  from each image, defined with every 8 pixels in both the horizontal and vertical directions. Thus, a total of 75 regions were selected in this feature set. We then extracted features in each patch and concatenated them to form the final feature of an image. The patch feature vectors consist of 16-bins histogram of 8 color channels (RGB, YCbCr, HS), uniform LBP histograms and HOG descriptors. So each patch was represented by a 484-dimensional feature vector, resulting in a 36300-dimensional vector for each image. Moreover, we define four representative body parts, i.e. upper-/lower-/left-/right- body. These body parts are obtained from whole body automatically (1/2 scale difference), and the similar feature representations were computed for different body parts respectively.

### 3.3. Selecting Discriminative Body Part

For each unknown probe person image, it is still hard to search the most discriminative body part precisely. Therefore, we investigate the overall re-identification performance using four representative body parts, including upper-/lower-/left-/right- body. Specifically, LFDA model is applied to match on the VIPeR, i-LIDS, and 3DPES datasets, and AUC (area under CMC curve ( $r \leq 20$ )) is utilized to measure the performance. The overall results are presented in Fig. 5. It is clear that the upper-body is superior to other body parts on all three datasets. Based on this fact, we focus on seeing how the upper-body of each person assist the whole body matching and enhance re-identification in our following experiments.

### 3.4. Evaluation of the Proposed Scheme

Firstly, we compare IMS-LFDA with LFDA [16]. As shown in Fig. 6 and Table 1, IMS-LFDA outperforms LFDA notably on all three datasets. The advantage is particularly significant on the more challenging VIPeR and i-LIDS datasets. For instance, on i-LIDS dataset, IMS-LFDA achieves about 10% improvement over LFDA at rank-1. Then we compare IMS-LFDA against several typical re-identification methods, including LMNN [22], KISSME [19], L1-norm and RDC [30]. Our results (Fig. 6 and Table 1) show clearly that IMS-LFDA obtains better matching significantly over these methods. More specifically, on VIPeR dataset, the rank-1 matching rate is 26.27% for IMS-LFDA, whilst 11.20% for LMNN, 21.55% for KISSME, 9.62% for L1-norm, and 17.78% for RDC. The SDC model [27] which only uses body parts to match was also adopted for comparison. One can observe that our model always outperforms SDC. Finally, SRC [23] is applied for matching. It is evident that our model still yields overall better performance than SRC. These results highlight the effectiveness of the proposed scheme.

### 3.5. Further Evaluations

In this section, we further evaluate the proposed IMS in the following four aspects.

**Contributions of Individual Components in IMS.** Our scheme consists of two matching components, i.e., whole body matching and body part matching. In our experiments, the two components in IMS-LFDA model are denoted as  $IMS-LFDA_{whole}$  and  $IMS-LFDA_{part}$ , respectively. In Table 2, we evaluate the two components on how they contribute to the full scheme. The results show that the best performance is achieved when the two components are combined. This suggests the body part could help the whole body to re-identify a probe person. The body part matching is less affected than the whole body matching by various changes including person-specific and environment-specific

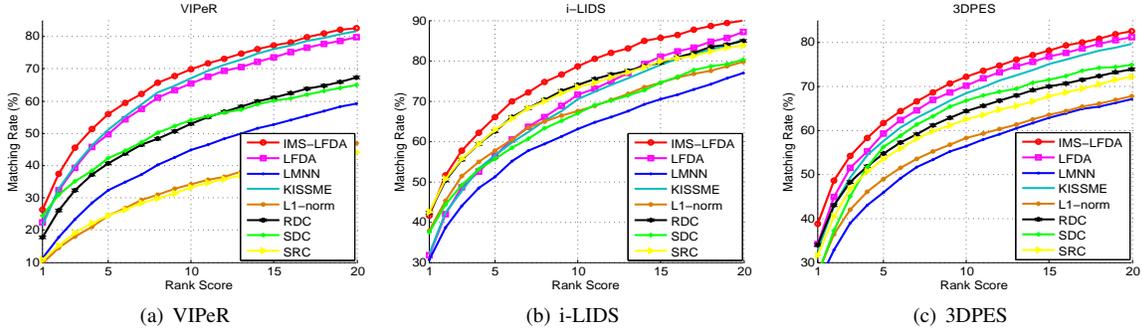


Figure 6: Matching rates (%) of different methods on VIPeR, i-LIDS, and 3DPES datasets.

Methods	VIPeR				i-LIDS				3DPES			
	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$
IMS-LFDA	<b>26.27</b>	<b>55.95</b>	<b>69.78</b>	<b>82.66</b>	41.66	<b>66.27</b>	<b>78.67</b>	<b>90.15</b>	<b>38.71</b>	<b>61.64</b>	<b>72.22</b>	<b>82.54</b>
LFDA [16]	21.58	49.78	65.38	79.75	31.67	56.44	71.52	87.04	34.24	59.31	70.13	81.12
LMNN [22]	11.20	32.34	44.75	59.27	30.69	51.14	62.98	76.87	25.12	45.92	56.39	67.03
KISSME [6]	21.55	50.98	67.09	81.65	32.22	56.63	70.48	85.24	32.76	57.65	68.51	79.71
L1-norm	9.62	24.56	34.21	46.93	37.80	57.76	67.42	79.64	27.43	48.86	58.22	67.75
RDC [30]	17.78	40.66	52.88	67.18	42.29	62.48	73.99	85.00	34.04	54.74	64.37	73.84
SDC [27]	24.58	42.37	54.14	64.92	37.41	55.74	66.98	80.23	27.25	56.16	66.82	74.88
SRC [23]	10.89	24.63	33.16	44.15	<b>42.47</b>	62.79	73.49	83.83	31.69	53.47	62.48	72.19

Table 1: Top ranked matching rate (%) at  $r=1,5,10,20$  on VIPeR, i-LIDS, and 3DPES.

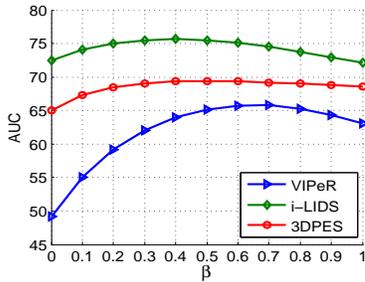


Figure 7: Performance comparison using different parameter  $\beta$  on three datasets.

changes, but the body part contains less spatial layout information than the whole body. So they are complementary to each other and fusion of them is reliable and effective.

**Effects of the Proposed Confidence in IMS.** The sparsity-based confidence is proposed for regulating the learned distance metric between two person images to match. In Table 2,  $IMS-LFDA_{w/o}$  denotes IMS-LFDA without using the confidence. Compared to LFDA (see Table 1),  $IMS-LFDA_{w/o}$  still has more accurate matching. This suggests that the body part matching is helpful for re-identification. Furthermore, it is also evident that IMS-LFDA consistently outperforms  $IMS-LFDA_{w/o}$ . The advantage of the confidence is validated by this experimental result.

**Is IMS Useful for Other Metrics?** The proposed IMS uses the metric learned by LFDA to re-identify in all previous experiments. One may ask whether the IMS can be useful for other metrics. To answer this question, we introduce two other metrics to IMS, which are defined by L1-

norm and RDC, respectively. We call the two models as IMS-L1-norm and IMS-RDC. We performed experiments on the three datasets and the results are reported in Table 3. It shows that IMS-L1-norm and IMS-RDC also give a notable improvement over L1-norm and RDC, respectively. This demonstrates the wide application of our scheme.

**Influence of Parameter.** In our experiments, we implement our scheme by selecting optimal parameter  $\beta$  (see Eq.(5)) in the range of  $[0,1]$ . In this section, we vary the parameter  $\beta$  to evaluate IMS. Specifically, we show results of IMS-LFDA as an example here. The AUC ( $r \leq 20$ ) of different  $\beta$  on three datasets are plotted in Fig. 7. It can be seen that when  $\beta$  is around 0.7 and 0.4 for VIPeR and other datasets, respectively, IMS-LFDA achieves the best performance. Since more variations including person-specific and environment-specific changes exist on i-LIDS and 3DPES, the body part matching is more useful (smaller optimal  $\beta$ ) on the two datasets than that on the VIPeR.

## 4. Conclusion

In this work, we have proposed two strategies to make matching in person re-identification more reliable. Based on these two strategies, we propose an integrated matching scheme for re-identification. This scheme utilizes body parts to assist whole body to match, and a sparsity-based confidence is introduced to our scheme to regulate the learned metric. Extensive experiments are conducted to evaluate the proposed scheme on three different person re-identification datasets. The results show clearly that our scheme is very effective for person re-identification.

Methods	VIPeR				i-LIDS				3DPES			
	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$
IMS-LFDA	<b>26.27</b>	<b>55.95</b>	<b>69.78</b>	<b>82.66</b>	<b>41.66</b>	<b>66.27</b>	<b>78.67</b>	<b>90.15</b>	<b>38.71</b>	<b>61.64</b>	<b>72.22</b>	<b>82.54</b>
IMS-LFDA <sub>whole</sub>	23.61	52.58	67.03	81.23	38.73	62.08	74.57	88.89	36.08	60.80	71.54	82.16
IMS-LFDA <sub>part</sub>	14.43	37.88	51.61	67.41	37.25	62.40	75.59	88.18	32.56	56.74	68.01	78.79
IMS-LFDA <sub>w/o</sub>	23.92	52.88	66.87	80.59	35.96	62.55	76.15	88.51	36.27	60.76	70.81	81.63

Table 2: Evaluation on VIPeR, i-LIDS, and 3DPES using LFDA in our scheme.

Methods	VIPeR				i-LIDS				3DPES			
	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$	$r=1$	$r=5$	$r=10$	$r=20$
IMS-L1-norm	<b>12.18</b>	<b>28.26</b>	<b>38.96</b>	<b>51.49</b>	<b>43.01</b>	<b>63.27</b>	<b>72.83</b>	<b>83.11</b>	<b>30.98</b>	<b>52.62</b>	<b>62.02</b>	<b>72.09</b>
L1-norm	9.62	24.56	34.21	46.93	37.80	57.76	67.42	79.64	27.43	48.86	58.22	67.75
IMS-RDC	<b>20.32</b>	<b>42.63</b>	<b>55.22</b>	<b>68.10</b>	<b>44.99</b>	<b>66.44</b>	<b>75.55</b>	<b>86.97</b>	<b>35.73</b>	<b>57.23</b>	<b>66.06</b>	<b>75.68</b>
RDC	17.78	40.66	52.88	67.18	42.30	62.48	73.99	85.00	34.04	54.74	64.37	73.84

Table 3: Evaluation on VIPeR, i-LIDS, and 3DPES using L1-norm and RDC in our scheme.

## Acknowledgment

This research was supported by the National Natural Science of Foundation of China (Nos. 61102111, 61472456).

## References

- [1] D. Baltieri, R. Vezzani, and R. Cucchiara. 3dpes: 3d people dataset for surveillance and forensics. In *joint ACM workshop on Human gesture and behavior understanding*, 2011.
- [2] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino. Custom pictorial structures for re-identification. In *BMVC*, 2011.
- [3] M. Farenzena, L. Bazzani, A. Perina, M. Cristani, and V. Murino. Person re-identification by symmetry-driven accumulation of local features. In *CVPR*, 2010.
- [4] N. Gheissari, T. Sebastian, and R. Hartley. Person reidentification using spatiotemporal appearance. In *CVPR*, 2006.
- [5] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*, 2008.
- [6] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof. Large scale metric learning from equivalence constraints. In *CVPR*, 2012.
- [7] I. Kviatkovsky, A. Adam, and E. Rivlin. Color invariants for person reidentification. *IEEE TPAMI*, 35(7):1622–1634, 2013.
- [8] H. Lee, A. Battle, R. Raina, and A. Y. Ng. Efficient sparse coding algorithms. In *NIPS*, 2006.
- [9] W. Li and X. Wang. Locally aligned feature transforms across views. In *CVPR*, 2013.
- [10] W. Li, R. Zhao, T. Xiao, and X. Wang. Deepreid: Deep filter pairing neural network for person re-identification. In *CVPR*, 2014.
- [11] Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao, and J. R. Smith. Learning locally-adaptive decision functions for person verification. In *CVPR*, 2013.
- [12] G. Lisanti, I. Masi, A. Bagdanov, and A. Del Bimbo. Person re-identification by iterative re-weighted sparse ranking. *IEEE TPAMI*, 2014.
- [13] B. Ma, Y. Su, and F. Jurie. Local descriptors encoded by fisher vectors for person re-identification. In *ECCV Workshop*, 2012.
- [14] L. Ma, X. Yang, and D. Tao. Person re-identification over camera networks using multi-task distance metric learning. *IEEE TIP*, 2014.
- [15] A. Mignon and F. Jurie. Pcca: A new approach for distance learning from sparse pairwise constraints. In *CVPR*, 2012.
- [16] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *CVPR*, 2013.
- [17] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang. Person re-identification by support vector ranking. In *BMVC*, 2010.
- [18] M. Sugiyama. Local fisher discriminant analysis for supervised dimensionality reduction. In *ICML*, 2006.
- [19] D. Tao, L. Jin, Y. Wang, Y. Yuan, and X. Li. Person re-identification by regularized smoothing kiss metric learning. *IEEE TCSVT*, 23(10):1675–1685, 2013.
- [20] UK. Home Office i-LIDS multiple camera tracking scenario definition. 2008.
- [21] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu. Shape and appearance context modeling. In *ICCV*, 2007.
- [22] K. Q. Weinberger, J. Blitzer, and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. In *NIPS*, 2005.
- [23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE TPAMI*, 31(2):210–227, 2009.
- [24] Z. Wu, Y. Li, and R. Radke. Viewpoint invariant human re-identification in camera networks using pose priors and subject-discriminative features. *IEEE TPAMI*, 2014.
- [25] F. Xiong, M. Gou, O. Camps, and M. Szaier. Person re-identification using kernel-based metric learning methods. In *ECCV*, 2014.
- [26] R. Zhao, W. Ouyang, and X. Wang. Person re-identification by saliency matching. In *ICCV*, 2013.
- [27] R. Zhao, W. Ouyang, and X. Wang. Unsupervised saliency learning for person re-identification. In *CVPR*, 2013.
- [28] R. Zhao, W. Ouyang, and X. Wang. Learning mid-level filters for person re-identification. In *CVPR*, 2014.
- [29] W.-S. Zheng, S. Gong, and T. Xiang. Person re-identification by probabilistic relative distance comparison. In *CVPR*, 2011.
- [30] W.-S. Zheng, S. Gong, and T. Xiang. Re-identification by relative distance comparison. *IEEE TPAMI*, 35(3):653–668, 2013.