Dissecting Person Re-identification from the Viewpoint of Viewpoint Supplementary Material

Xiaoxiao Sun   Liang Zheng

Australian National University

xxsunzrt@gmail.com   liang.zheng@anu.edu.au

In the supplementary material, we include the more information of the PersonX dataset, and detailed numbers of experimental results shown in the paper.

1. PersonX datasets

The details of the datasets used in the paper are shown in Table 1. We combine two different backgrounds as one dataset to study different situations.

Datasets	Description	Backgrounds					
		(1)	(2)	(3)	(4)	(5)	(6)
PersonX$_{123}$	color background	✓	✓	✓			
PersonX$_{12}$	easy situation	✓	✓				
PersonX$_{13}$	hard situation	✓		✓			
PersonX$_{456}$	scene background				✓	✓	✓
PersonX$_{45}$	easy situation				✓	✓	
PersonX$_{46}$	hard situation				✓		✓

Table 1. Subsets of the PersonX dataset. (1) - (3) are three color backgrounds. (4) - (6) are three scene backgrounds PersonX$_{123}$ means the background of images is pure color. PersonX$_{456}$ means the background of images is scene. Excerpt for PersonX$_{123}$ and PersonX$_{456}$ contain three backgrounds respectively, all the other subsets consists of two backgrounds.

Based on the above setting, the results of IDE+, triplet feature and PCB on real-world and synthetic datasets are shown in Table 2 (including the **numerical results of Fig. 4 in the paper**) .

2. Experiments

In experiments of the paper, we have discussed the influence of viewpoint on three aspects. More details of experimental settings and results are as follows.

2.1. How Do Viewpoint Distributions in the Training Set Affect Model Learning?

The settings of the control group and experimental group are shown in Fig 1. We use two data as examples to show selecting half viewpoints (18 out of 36) of each identity for

#	method	IDE+		triplet feature		PCB	
		mAP	R1	mAP	R1	mAP	R1
1	Market-1501	67.3	86.5	68.1	85.6	77.4	92.3
2	Market-1203	66.7	71.0	72.8	75.6	77.8	81.4
3	DukeMTMC-reID	55.1	74.2	57.9	76.3	66.1	81.7
4	PersonX$_{123}$	94.5	99.2	95.3	98.6	97.9	99.6
5	PersonX$_{12}$	94.8	99.5	95.6	99.0	97.8	99.6
6	PersonX$_{13}$	94.6	99.1	94.6	98.6	97.6	99.4
7	PersonX$_{456}$	92.7	98.5	94.3	98.5	97.8	99.7
8	PersonX$_{45}$	94.0	99.4	95.0	99.1	97.7	99.8
9	PersonX$_{46}$	91.5	96.3	92.8	97.0	96.8	99.2
10	PersonX$_{456}$-lr	87.9	96.7	89.8	95.6	94.9	98.4
11	PersonX$_{45}$-lr	90.2	98.4	90.9	97.5	95.1	98.9
12	PersonX$_{46}$-lr	85.9	94.3	86.2	93.2	92.6	96.6

Table 2. Benchmarking the subsets of PersonX. We use mAP (%) and rank-1 (R1, %) accuracy for measurement. "lr" means the frames are low resolution of 512×242 instead of the original resolution 1024×768. As detailed in the text, this table validates the eligibility, purity and sensitivity of PersonX.
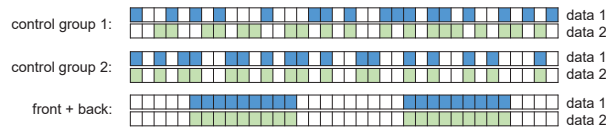

Figure 1. Illustration of experiment. **A:** taking two data as examples to show the difference between control group 1, control group 2 and experimental group. The grids in each row from left to right represent the viewpoints from $0°$ to $350°$.

training. In control group 1, the viewpoints used for training of data 1 and data 2 are different. In control group 2, 18 viewpoints are randomly selected from 36 to train the model, and they are same for the two data. In the experimental group, taking front+back as an example, the used viewpoints are front and back for the two data. Table 3 shows the results of control and experimental groups, which contains the **numerical results of Fig. 5 in the paper**.

Meanwhile, we show some re-ID results on Market-1203 in Fig. 2. From the first three rows, we can find that both the two models can work well at query viewpoint of front or back. The difference is the ranking of

Group	Viewpoint / Direction	Img#	PersonX$_{12}$ mAP (%)	R1 (%)	PersonX$_{13}$ mAP (%)	R1 (%)	PersonX$_{45}$ mAP (%)	R1 (%)	PersonX$_{46}$ mAP (%)	R1 (%)	Market-1203 Img#	mAP (%)	R1 (%)
Oracle	$[0°, 350°]$	36	97.8	99.6	97.6	99.4	97.7	99.8	96.8	99.2	1-8	77.8	81.4
CG 1	$[0°, 350°]$	27	97.8	99.6	97.2	99.5	97.6	99.9	96.8	98.8	1-6	72.8	75.9
CG 2	reserve 27 viewpoints	27	97.7	99.7	97.2	99.5	97.4	99.9	96.4	98.8	1-6	69.0	73.3
	right+back+front	27	97.4	99.6	97.1	99.4	96.9	99.9	96.2	98.6	1-6	66.6	69.4
EG	left+right+back	27	97.2	99.6	96.7	99.5	96.6	99.9	96.1	98.9	1-6	62.1	65.4
	left+front+back	27	97.5	99.5	97.3	99.2	97.3	99.9	96.4	98.8	1-6	64.0	67.8
	left+front+right	27	97.1	99.5	96.8	99.4	96.7	100	96.1	98.6	1-6*	59.0	62.6
CG 1	$[0°, 350°]$	18	97.6	99.6	97.0	99.4	97.0	99.9	96.3	98.7	1-4	50.2	53.9
CG 2	reserve 18 viewpoints	18	97.3	99.6	96.8	99.3	96.8	99.8	96.1	98.8	1-4	49.9	53.3
	left+right	18	96.2	99.5	95.7	99.4	95.5	99.9	95.3	98.5	1-4	44.4	47.9
EG	front+back	18	95.4	99.5	94.5	99.2	94.8	99.6	94.2	98.1	1-4	42.2	45.2
	front+right (left)	18	96.4	99.5	96.0	99.2	95.9	99.8	95.1	98.2	1-4	41.0	45.4
	back+right (left)	18	96.4	99.6	95.8	99.3	95.4	99.8	95.0	98.4	1-4	45.0	45.7
CG 1	$[0°, 350°]$	9	97.1	99.6	96.3	99.4	96.5	99.6	95.3	98.1	1-2	–	–
CG 2	reserve 9 viewpoints	9	95.7	99.5	95.1	99.2	95.1	99.6	94.3	98.1	1-2	–	–
EG	left	9	94.7	99.6	93.7	99.1	93.9	99.6	93.2	98.1	1-2	–	–
EG	front	9	87.7	99.4	86.5	98.8	87.0	99.5	86.6	97.4	1-2	–	–
EG	right	9	94.3	99.5	93.6	95.8	93.7	99.7	93.1	97.9	1-2	–	–
EG	back	9	88.1	99.4	85.9	98.6	87.8	99.3	87.4	96.3	1-2	–	–

Table 3. Performance using different viewpoints in the training set. "CG" and "EG" denote the control group and experiment group, respectively. #1 Oracle means all 36 viewpoints are contained by training data. "Viewpoint / Direction" shows the viewpoints contained in the training set. "Img#" is the number of training images for each identity.

Figure 2. Sample re-ID results on Market-1203. Images in the first column are queries. The retrieved images are sorted according to their similarity to the query (high to low) from left to right. True matches and false matches are in green (or orange) and red rectangles, respectively. **A:** The similarity is calculated by using feature extracted from PCB model trained on front+back orientations. **B:** The similarity is calculated by using feature extracted from PCB model trained on left+right orientations.

the true matches: model trained on front+back would rank true matches of front/back viewpoints in a higher position, while true matches of left/right viewpoints will be ranked higher by the model trained on left+right (see images in orange rectangles). On the other hand, from the fourth to sixth rows, we can observe that the model **B** can perform better than model **A** when query viewpoint is left or right.

This observation is consistent with the analysis (the plausible reason) in the paper.

2.2. How Does Query Viewpoint Affect Retrieval?

In the paper, we have shown the results of training a model on the original training set comprised of every viewpoint and the query viewpoints are modified in Fig. 7 (A
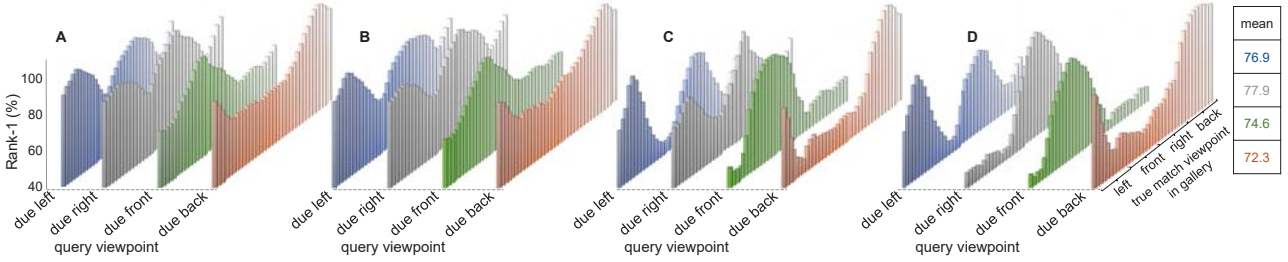
Figure 3. The impact of query viewpoint on system performance on PersonX$_{45}$. Four viewpoints are evaluated, *i.e.*, due left ($0°$), due right ($180°$), due front ($90°$) and due back ($270°$). In the gallery, there is only one true match for each query. The true match viewpoint varies from $0°$ to $350°$ (deep axis). Under each query viewpoint, we report 36 rank-1 scores obtained by the query to retrieve 36 types of true match viewpoints. **A, B, C and D:** the R1 are testing on PersonX$_{45}$ by using the models trained on left, right, front and back orientations, respectively. On the left, we show the average rank-1 accuracy for each query viewpoint when testing on four models.

#	Training Data	Oracle		CG 1-1		EG 1		CG 1-2		CG 2-1		EG 2		CG 2-2	
		mAP	R1	mAP	R1	mAP ($\downarrow$)	R1 ($\downarrow$)	mAP	R1	mAP	R1	mAP ($\downarrow$)	R1 ($\downarrow$)	mAP	R1
1	PersonX$_{45}$	97.7	99.8	97.6	99.8	**97.4** (0.3)	**99.4** (0.4)	97.7	99.8	97.4	99.8	**96.6** (1.1)	**98.3** (1.5)	97.8	99.8
2	PersonX$_{45}$ 3/4	97.6	99.9	97.5	99.9	**97.2** (0.4)	**98.9** (1.0)	97.6	99.9	97.3	99.8	**96.7** (0.9)	**98.4** (1.5)	97.7	99.9
3	PersonX$_{45}$ 2/4	97.0	99.9	96.9	99.9	**96.5** (0.5)	**98.9** (1.0)	97.0	99.9	96.9	99.9	**95.5** (1.5)	**97.4** (2.5)	97.1	99.9
4	PersonX$_{45}$ 1/4	96.5	99.6	96.3	99.6	**95.9** (0.6)	**98.3** (1.3)	96.4	99.6	96.0	99.4	**94.8** (1.7)	**96.6** (3.0)	96.4	99.6
5	PersonX$_{46}$	96.8	99.2	96.7	99.1	**96.4** (0.4)	**98.2** (1.0)	96.8	99.2	96.4	99.0	**95.5** (1.3)	**97.1** (2.1)	96.8	99.2
6	PersonX$_{46}$ 3/4	96.8	98.8	96.7	98.8	**96.4** (0.4)	**98.1** (0.7)	96.8	98.8	96.4	98.7	**95.5** (1.3)	**96.7** (2.1)	96.8	98.8
7	PersonX$_{46}$ 2/4	96.3	98.7	96.2	98.7	**95.8** (0.5)	**97.1** (1.6)	96.3	98.7	95.8	98.5	**94.8** (1.5)	**95.8** (2.9)	96.3	98.7
8	PersonX$_{46}$ 1/4	95.3	98.1	95.1	98.1	**94.6** (0.7)	**96.6** (1.5)	95.1	98.1	95.2	98.1	**93.4** (1.9)	**94.9** (3.2)	94.7	97.8
9	PersonX$_{46}$-lr	92.6	96.6	92.4	96.4	**91.8** (0.8)	**95.9** (0.7)	92.6	96.6	91.7	96.1	**90.2** (2.4)	**94.3** (2.3)	92.7	96.6
10	PersonX$_{46}$-lr 3/4	92.5	96.0	92.2	95.7	**91.7** (0.8)	**95.3** (0.7)	92.4	96.0	91.6	95.9	**90.0** (2.5)	**94.0** (2.0)	92.5	96.0
11	PersonX$_{46}$-lr 2/4	91.7	95.7	91.4	95.6	**90.8** (0.9)	**94.6** (1.1)	91.6	95.6	90.8	95.4	**89.0** (2.7)	**93.4** (2.3)	91.7	95.6
12	PersonX$_{46}$-lr 1/4	90.6	95.6	90.2	95.4	**89.6** (1.0)	**94.3** (1.3)	90.5	95.6	89.5	95.0	**87.5** (3.7)	**92.7** (2.9)	90.6	95.6
13	Market-1203	77.8	81.4	77.2	79.4	**74.7** (3.1)	**74.4** (7.0)	77.2	79.4	79.3	79.7	**74.1** (3.7)	**71.0** (10.4)	83.1	83.0
14	Market-1203 3/4	72.8	75.9	71.6	73.0	**68.8** (4.0)	**67.9** (8.0)	76.3	76.4	71.5	68.6	**68.5** (4.3)	**64.9** (11.0)	80.3	79.3
15	Market-1203 2/4	50.2	53.9	48.8	50.0	**44.7** (5.5)	**43.5** (10.4)	56.3	55.7	48.7	42.7	**44.5** (5.7)	**41.1** (12.8)	62.3	60.1

Table 4. Evaluation of viewpoint during testing. Oracle means database containing 36 viewpoints as true matches of each identities. 3/4, 2/4 and 1/4 in "Training Data" represent the using 3/4 (2/4 and 1/4) amount of data in original dataset for training (corresponding to the CG1 in Table 3). ($\downarrow$) represents the discrepancy of result compared with the result in column of Oracle.

probe image is set to the *due left*, *due front*, *due right* and *due back* to represent different sides of person. During retrieval, the true match contains the same person as the query, and its viewpoint is between $0°$ and $350°$, respectively. Viewpoints of the distractor gallery images are images of all other person.). And we find that *the query viewpoint of left/right generally leads to higher re-ID accuracy than front/back viewpoints*. To further prove this conclusion, more experiments are conducted based on the four models trained on left, front, right and back orientations, respectively. The **similar results with Fig. 7** are shown in Fig. 3. The consistent conclusion can be summarized that due left/right results higher re-ID accuracy than the due front/back.

2.3. How Do True Match Viewpoints in the Gallery Affect Retrieval?

Similar to the paper, the viewpoint of a query and its true match are defined as θ_q and θ_t, respectively. The experimental groups for PersonX subsets are designed as below.

- Control group 1-1. Three true matches are randomly removed from the gallery.

- Control group 1-2. Three true matches whose $\theta_t \in [\theta_p + 180 \pm 10°]$ are removed, because these images can be considered as having a big distance with the query images.

- Experimental group 1. The three true matches whose $\theta_t \in [\theta_q \pm 10°]$ are removed (set as "junk").

- Control group 2-1. Three true matches are randomly removed from the gallery.

- Control group 2-2. Three true matches whose $\theta_t \in [\theta_p + 180 \pm 40°]$ are removed.

- Experimental group 2. The nine true matches whose $\theta_t \in [\theta_q \pm 40°]$ are removed (set as "junk").

Experimental results are shown in Table 4, some of which has been shown in **Fig. 8 of the paper**. Furthermore, there are two newly added control groups (CG 1-2 and CG 2-2), in which the matches of dissimilar viewpoint to query are removed. We find the final results have no obvious changes compared to Oracle (the improvements on Market-1203 dataset is due to the reduction of retrieval difficulty).

3