Key Person Aided Re-identification in Partially Ordered Pedestrian Set

Chen Chen¹
chen.chen@ia.ac.cn

Min Cao¹
caomin2014@ia.ac.cn

Xiyuan Hu¹
xiyuan.hu@ia.ac.cn

Silong Peng¹²
silong.peng@ia.ac.cn

¹ Institute of Automation
Chinese Academy of Sciences
Beijing, China

² Beijing Visytem Co. Ltd.

Abstract

Ideally person re-identification seeks for perfect feature representation and metric model that re-identify all various pedestrians well in non-overlapping views at different locations with different camera configurations, which is very challenging. However, in most pedestrian sets, there always are some outstanding persons who are relatively easy to re-identify. Inspired by the existence of such data deviation, we propose a novel key person aided person re-identification framework based on the re-defined partially ordered pedestrian sets. The outstanding persons, namely “key persons”, are selected by the K -nearest neighbor based saliency measurement. The partial order defined by pedestrian entering time in surveillance associates the key persons with the query person temporally and helps to locate the possible candidates. Experiments conducted on two video datasets show that the proposed key person aided framework outperforms the state-of-the-art methods and improves the matching accuracy greatly at all ranks.

1 Introduction

Person re-identification (re-id) aims to match pedestrians across non-overlapping camera views. Due to the variation in viewpoints, illumination, background and occlusions, the appearances of the same person observed in different camera views are often ambiguous and unreliable for being re-identified. Ideally person re-id research seeks for perfect feature representation and metric model that can re-identify all various query persons in views of different camera configurations from different locations, which is very challenging.

However, we notice that in almost all person re-id datasets, there always are some pedestrians who are relatively easy to be re-identified accurately and some who are not. Persons who are “outstanding” are usually re-identified with high rankings, on the contrary pedestrians who are not unique are usually wrongly matched even using powerful algorithms. Although pedestrians are different from dataset to dataset, there always exists such deviation.

One intuitive idea is first re-identifying the “outstanding” persons; then utilizing those re-identified “outstanding” pedestrians to help re-identify other pedestrians. To realize this idea we need to solve two problems: how to define and select these “outstanding” persons and given the “outstanding” persons and how to associate and re-identify other pedestrians.

The persons that we attempt to find are “outstanding”, “unique”, easily recognized and accurately re-identified, which we call “*key persons*” in this paper. In the sense of vision saliency, the key persons are salient persons among the pedestrian set. The K -nearest neighbor distance is used to measure and select key persons in this paper. In order to associate key persons with other pedestrians, we use the entering time of pedestrian in camera view to define a temporal partial order. The video clip is synopsisized into a pedestrian flow and the pedestrians in flow are associated with each other temporally. The temporal distance between them that measured by the difference of their entering time is used to fine locate the possible candidates.

As illustrated in Fig. 1, the traditional way of representing pedestrian set is re-defined as a pedestrian flow with a temporal partial order, person re-id problem is addressed as element matching among two partially ordered sets using their temporal relations. For instance, the query person who wears a white T-shirt and dark trousers looks similar with many other pedestrians, directly re-identifying her may lead to many wrong matches. Using the proposed idea, we can select the pedestrian with red dress as a key person, re-identify her first and utilize the temporal distance between her and the query person to fine locate possible candidates in gallery set. Re-identifying aided by key persons could reduce many false matches.

In this paper, we introduce the definition of key persons, model the pedestrian set as a partially ordered set with temporal relations and propose a key person aided person re-id framework, which utilized the salient persons to help re-identifying other pedestrians. The following of the paper is organized as follows: section 2 reviews the related works, section 3 presents how to define and select key persons, section 4 describes the proposed key person based pedestrian re-id framework, section 5 shows the conducted experiments and the results and section 6 concludes the paper.

2 Related work

Image-based approaches: Researchers usually focus on developing distinctive feature representations [10, 19, 20, 23] and seeking discriminative distance metrics [16, 17, 21, 30] based on a single individual image. For appearance modelling, Farenzena *et al.* [10] proposed to model the human appearance based on different body parts, which are obtained by adopting perceptual principles of symmetry and asymmetry. Matsukawa *et al.* [20] employed both the mean and the covariance information of pixel features via hierarchical Gaussian distribution to describe a local region in an image. For metric learning, Zheng *et al.* [30] introduced a Probabilistic Relative Distance Comparison (PRDC) model to maximize the probability of a pair of true association having a smaller distance than that of a wrong association. A low dimensional subspace is firstly learned by cross-view quadratic discriminant analysis and simultaneously a QDA metric is performed on the subspace in [17]. Recently, deep Learning, as a powerful tool in computer vision, has attracted increasing attention for person re-id [6, 13, 22, 25] and achieved better performance.

video-based approaches: Multiple images of the same person from video have been utilized for person re-identification [1, 3, 11, 24, 27]. Compared with a single individual image,

Figure 1: From the traditional pedestrian set to temporally partially ordered set

multiple images of a person provide more clues to differentiate pedestrians from each other. For example, the appearance features extracted by multiple images were accumulated or averaged into a single signature [10, 9]. The temporal correlation of multiple images was also exploited to build the spatial-temporal appearance representation for person re-id [11, 27].

In addition, the group context of a person can provide visual clues for person re-id, and there is growing interest in utilizing group information to improve re-identification performance [2, 4, 12, 24, 29]. Group is usually defined by sociological theory in behavior and interaction analysis, and group based methods requires three steps: group detection, group feature extraction and/or group matching. More specifically, Zheng *et al.* [29] defined two group features Center Rectangular Ring Ratio-Occurrence Descriptor (CRRRO) which is invariant to the position changes of members in a group and Block based Ratio-Occurrence Descriptor (BRO) addressing variations in illumination and viewpoint on manually segmented group images. Cai *et al.* [9] represented the group images by the covariance descriptor, capturing the appearance and statistical properties of the group image. Li *et al.* [12] computed the grouping possibility by velocity and distance between people and then detected groups by Affinity Propagation (AP) clustering algorithm. For group features they extracted both geometry and visual information of a subject's partners. These methods have shown decent results in terms of accuracies by utilizing context information of other people, however, they work only when groups exist. The proposed key person aided framework is more flexible and practical with or without groups, and the core idea is to efficiently locate possible candidates by key persons.

3 Key person: definition and selection

The K nearest neighbor (K -NN) distance has been used for outlier detection, clutter removal and patch saliency learning [9]. The average k -NN distance can measure how distinct the query point is from the rest points in the set, which represents a saliency measurement for

point set. We utilize the K -NN based saliency measurement to define the key persons in a pedestrian set. In order to select sufficient number of reliable key persons, we introduce a feature bank based key person selection strategy.

3.1 Key person definition

Let $\mathcal{P}$ be a pedestrian set of N persons and f be the feature extraction strategy that maps the person image into feature space, the similarity score between person p_i and p_j in feature space of f using the employed metric is defined as $d_f(p_i, p_j)$. The *saliency score* of person p in set $\mathcal{P}$ is defined as the normalized averaged K -NN distance

$$s_f(p) = \mathcal{N}\left(\frac{1}{K} \sum_{k=1}^K d_f(p, p_{n_k})\right), \quad (1)$$

where $\{p_{n_k}\}_{k=1,\dots,K}$ denotes the K -NNs of person $p \in \mathcal{P}$ and $\mathcal{N}$ denotes a min-max normalization operator to linearly scale the averaged K -NN distance into the range $[0, 1]$. Person p is defined as a *key person* in set $\mathcal{P}$ if the saliency score is larger than a threshold $s_f(p) \geq \rho$, and the *key person set* defined in feature space of f is written as $\mathcal{S} = \{p \in \mathcal{P} | s_f(p) \geq \rho\}$. The saliency threshold ρ should be a relatively large value to ensure the reliability of selected key persons.

We can infer from the definitions that the same person may have different saliency scores and be categorized differently as a key person or not, if employing different feature extraction strategy f to represent pedestrians and using different metric scheme $d_f(\cdot, \cdot)$ to measure the similarity between features.

3.2 Feature bank based key person selection

Researchers have proposed various efficient feature representations and metrics re-identifying pedestrians descriptively and discriminatively across views. Based on a feature bank $\mathcal{F} = \{f^m, m = 1, \dots, M\}$, M key person sets $\mathcal{S}^m = \{p \in \mathcal{P} | s_{f^m}(p) > \rho^m\}$ are obtained in the feature space projected by feature f^m . And for the pedestrian set, the whole key person set is the union of every key person set $\mathcal{S} = \cup_{m=1}^M \mathcal{S}^m$ with threshold $\rho = \{\rho_1, \rho_2, \dots, \rho_m\}$. These M key person sets may (partly) overlap, since some key persons may be salient in multiple feature space. The non-overlapping part suggests the whole key person set $\mathcal{S}$ is selected by complementary features in feature bank, which shows the effectiveness and necessity of introducing feature bank based key person selection.

4 Key person aided pedestrian re-identification

In the traditional way, the probe (or gallery) set represents only a number of pedestrian images within bounding boxes, and person re-id means matching an image in probe set with the whole set of images in gallery. However, with the temporal information from videos, the probe (or gallery) set can be defined as a temporal sequence with a partial order, and the key persons can help to locate possible candidates and reduces false matches.

Figure 2: Overview of the proposed key person aided person re-id framework with four steps: 1) find the nearest key person of query person; 2) re-identify key person with the top match in gallery set; 3) use temporal constrain to locate the possible candidate of query person; 4) weight the possible candidate and rank candidates in gallery set. Better viewed in color.

4.1 From traditional pedestrian set to partially ordered set

Let p denote a person in probe set P and T denote the starting frame of person p appeared in a camera view, a partial order for the probe set is defined by the relation of the starting frame T of pedestrian, let $p_i \leq p_j$ if $T_i \leq T_j$. Then the probe set P from a camera view is defined as a partially ordered set

$$\mathcal{P} = \{p | p \leq p_i \text{ if } T \leq T_i, p \in P, p_i \in P\}. \quad (2)$$

The temporal distance between elements p_i and p_j in $\mathcal{P}$ is measured by $\Delta T = T_i - T_j$. Similarly, the traditional gallery set G is re-defined as a partially ordered set $\mathcal{G}$. In a vivid description, the pedestrian set is now re-defined as a temporally ordered pedestrian flow, and the relative temporal distance between each other is defined by the temporal partial order.

People appeared in the same scene usually walk with similar velocity. Under such assumption, the partially ordered pedestrian set keep the internal temporal relation; when the flow passes by two adjacent cameras views, most pedestrians in the flow keep their relative partial orders. The stability of relative order usually ensures the resistance of the temporal relations across adjacent camera views.

Consider more practical cases in real world, when pedestrian velocity vary greatly and their relative temporal partial order is hardly kept, the probe set can split into subsets due to velocity contains as $\mathcal{P} = \cup_i \mathcal{P}_{\mathbf{v}_i}$, and

$$\mathcal{P}_{\mathbf{v}_i} = \{p \in \mathcal{P} | \|\mathbf{v}\| \in [\|\mathbf{v}_i\| - \varepsilon, \|\mathbf{v}_i\| + \varepsilon], \arccos(\frac{\mathbf{v} \cdot \mathbf{v}_i}{\|\mathbf{v}\| \|\mathbf{v}_i\|}) < \theta\}, \quad (3)$$

where $\mathbf{v}_i$ denotes the main velocities and $\mathcal{P}_{\mathbf{v}_i}$ denotes the subset of pedestrians with similary velocity both in speed magnitude and walking direction. Then within each subset $\mathcal{P}_{\mathbf{v}_i}$ the proposed key person based person re-id framework can still apply.

4.2 Key person based person re-identification framework

The framework of key person based person re-id consists four steps: 1) in probe set, find the nearest key person of the query person by temporal distance; 2) Match that key person with candidates in gallery set and obtain the top match; 3) Using temporal constrain provided

by key person's top match to locate the possible candidates of query person; and 4) assign different weights to the possible candidates and rest pedestrians in gallery set and rank them.

1) $p^A \rightarrow p^{keyA}$: Given the probe set $\mathcal{P}$ and the gallery set $\mathcal{G}$ with a partial order defined by the starting frame of pedestrian appeared in video, select the key person set $\mathcal{S} = \cup_{m=1}^M \mathcal{S}^m$ in $\mathcal{P}$ with a feature bank $\mathcal{F} = \{f^1, f^2, \dots, f^M\}$. Let a query person be p^A with the starting frame T^A , find the nearest key person $p^{keyA} \in \mathcal{S}$ with feature f^m by the temporal distance, and compute the temporal distance between them,

$$p^{keyA} = \arg \min_{p_i \in \mathcal{S}} \|T^A - T_i^A\|, \quad (4)$$

2) $p^{keyA} \rightarrow p^{keyB}$: Match the key person p^{keyA} with the candidate $p^B \in \mathcal{G}$, compute the similarity score d^{key} in feature space of f^m and normalized into (0,1) as

$$d^{key} = \tilde{d}_{f^m}(p^{keyA}, p^B), \quad (5)$$

and find the top match p^{keyB} , where p^{keyB} does not denote a key person in gallery set by the proposed definition but refers a short expression for the key person p^{keyA} 's top match,

$$p^{keyB} = \arg \min_{p^B \in \mathcal{G}} \tilde{d}_{f^m}(p^{keyA}, p^B) \quad (6)$$

3) $p^{keyB} \rightarrow \mathcal{G}^*$: Compute the temporal distance between the key person and query person $\Delta T^A = T^A - T^{keyA}$, and let T^{*B} denote the starting frame of the query person's correspondence p^{*B} in $\mathcal{G}$, and $\Delta T^{*B} = T^{*B} - T^{keyB}$. If the temporal order of pedestrian flow across camera is strictly kept, $\Delta T^A = \Delta T^{*B}$ holds, i.e. $T^{*B} = T^{keyB} + \Delta T^A$. In practice, the possible candidates p^B should appear with large probability around the time T^{*B} with a tolerance interval parameterized by τ ,

$$\mathcal{G}^* = \{p_j^{*B} | p_j^{*B} \in \mathcal{G}, T_j^{*B} \in [T^{keyB} + (1 - \tau)\Delta T^A, T^{keyB} + (1 + \tau)\Delta T^A]\}. \quad (7)$$

4) $p^A \rightarrow \mathcal{G}$: Match the query person p^A with candidate p^B in gallery set $\mathcal{G}$ by assigning the original similarity scores with different weights. The original similarity score $d_{f^{base}}$ is computed using the baseline method in feature space of f^{base} . We assign weights to the possible candidates in $\mathcal{G}^*$, compute the new similarity score for all candidates in $\mathcal{G}$ and rank them,

$$d^*(p^A, p^B) = \omega \cdot d_{f^{base}}(p^A, p^B), \quad (8)$$

where

$$\omega = \begin{cases} d^{key}, & \text{if } p^B \in \mathcal{G}^*, \\ 1, & \text{otherwise.} \end{cases} \quad (9)$$

denotes the weight and d^{key} is defined in Eq. (5).

In order to increase the matching accuracy for the query person, we usually take multiple key persons close to the query person to locate the possible candidate set. Let p_l^{keyA} be the l_{th} nearest key person defined by temporal distance, $l = 1, 2, \dots, L$, and let p_l^{keyB} be the top match of p_l^{keyA} in $\mathcal{G}$ and the similarity score be d_l^{key} , the possible candidate set for query person is $\mathcal{G}_l^*$. As a result, the assigned weight for similarity score in Eq. (8) is written as

$$\omega = \begin{cases} d_l^{key}, & \text{if } p^B \in \mathcal{G}_l^*, \\ 1, & \text{otherwise.} \end{cases} \quad (10)$$

Then the rankings of all $p^B \in \mathcal{G}$ are computed based on the new similarity score $d^*(p^A, p^B)$.

Figure 3: Example of CYBJ-G dataset: two cropped person images and the corresponding video frames in different camera views.

5 Experiments and Results

We first provide a proof of concept for key person, then evaluate the proposed framework on two datasets and show the comparisons with the state-of-the-art methods. Due to the limited video dataset, we evaluate our approach on a public dataset: PRID2011 dataset and a new dataset proposed by us: CYBJ-G dataset.

PRID2011 dataset [12] includes video and person images recorded from two cameras and full surveillance videos. 385 and 749 persons were recorded in camera views, respectively. We use the single images of the first 200 pedestrians who appear in both cameras.

CYBJ-G dataset consists 194 pedestrians captured by two surveillance cameras from the frontal and back view in a residential area. For each view, data for each pedestrian consists one cropped image and the sequential images of the corresponding video clip. The cropped images of persons have been resized to 384×144 pixels. Some examples are shown in Fig. 3.

The feature bank employs three classic and state-of-the-art feature SDALF [13], GOG [20] and DNS [26]. The baseline person re-id method adopts GOG+XQDA [20] method. For all experiments, each dataset was randomly split into half for training and half for testing. All experiments were repeated for 10 trials and the results were averaged for better stability. The results are shown in standard Cumulated Matching Characteristics (CMC) curves.

5.1 Proof of concept for key person

Key person selection The saliency threshold ρ should ensure most selected key persons are reliable matching across view. According to the key person definition, the saliency threshold ρ should not be too small, otherwise many key persons that are not reliable enough will be selected; while ρ should not be too large, otherwise very few key persons are selected so that the temporal distance with query person will be large and the re-id accuracy will be harmed by larger noise. Figure.4 shows the matching accuracy σ and the number of key persons N_s vary due to the change of ρ for three features on two datasets. As a trade off between high accuracy and reasonable size, we set the saliency thresholds $[\rho_{GOG}, \rho_{DNS}, \rho_{SDALF}] = [0.7, 0.9, 0.99]$ for PRID2011 dataset and $[0.9, 0.6, 0.99]$ for CYBJ-G dataset.

Key person illustration In Fig. 5, we illustrate the key persons in PRID2011 with saliency score $s = 1$ using GOG, DNS and SDALF, and the rankings of their true candidates are all top 1. Most shown key persons appearance with bright color or unique accessories, which matches with the processing of human visual saliency.

Fig. 6 illustrate the key persons selected from a trial from experiments on PRID2011 with their feature labels, saliency scores and the rankings of their true matches in gallery set.

Figure 4: Reference curves for saliency threshold setting: (a) the matching accuracy σ and (b) number of key persons N_S change due to the saliency threshold ρ varies for PRID2011 dataset; (c-d) for CYBJ-G dataset.

Figure 5: Illustration of the key persons with saliency scores $s = 1$ in different feature space on PRID2011 dataset.

We can see that a) sometimes some key person is salient in multiple feature space and the key person sets overlap, *e.g.* person No. 67 is selected with both GOG and DNS feature; b) the key persons selected in different feature space are mostly different, which verifies the complementary of the feature space; c) the key persons are reliable and their rankings of true matches are all top 1 except one key person. The high ranking rate ensures the correct matching of key persons as the foundation for key person based person re-id framework.

5.2 Comparison with state-of-art methods

The statistics of pedestrian velocity shows that walking speeds in both datasets are in a reasonable interval, and walking directions mainly pointing in opposite two directions. According to Eq.(3), we split the probe set into two subsets due to the walking directions. To perform the proposed framework, we also need to set the time interval parameter τ and the number of nearest key persons L , which rely on the stability of the temporal order among pedestrians. We set $\tau = 0.3, L = 4$ for PRID2011 and $\tau = 0.1, L = 2$ for CYBJ-G since

Person Image
Person ID	107	61	67	63	98	194
Feature Score	GOG 0.8388	GOG 0.7047	GOG 0.7987 DNS 1.0	GOG 0.7724	SDALF 1.0	GOG 1.0
Re-ID Rank	1	2	1	1	1	1

Figure 6: The illustration of key persons selected in one trial on PRID2011 dataset.

Table 1: Comparison with the classic and existing state-of-the-art methods on PRID2011 and CYBJ-G. The best and second best results (%) are respectively shown in red and blue. Better viewed in colour.

Methods	PRID2011				CYBJ-G			
	r=1	r=5	r=10	r=20	r=1	r=5	r=10	r=20
SDALF [10]	6.4	24.3	32.7	44.8	32.2	57.2	69.8	79.9
Saliency [28]	25.8	43.6	52.6	62.0	40.8	64.2	75.5	83.4
LOMO+XQDA [17]	39.0	68.0	83.0	91.0	67.2	87.4	91.7	95.2
SCSP [8]	12.7	32.7	51.0	66.0	21.7	39.1	50.0	67.4
DNS [26]	38.6	66.6	78.0	91.1	58.2	84.2	91.5	94.1
GOG [20]	59.2	79.6	89.7	95.6	76.6	94.2	96.9	98.4
DTDL [13]	41.0	70.0	78.0	86.0	-	-	-	-
PaMM [9]	45.0	72.0	85.0	92.5	-	-	-	-
STA [18]	64.1	87.3	89.9	92.0	-	-	-	-
Ours	81.4	96.2	98.7	99.7	89.4	98.8	99.9	100.0

pedestrians in CYBJ-G dataset show better temporal stability than PRID2011 dataset.

In Table 1, we reported the comparison of the proposed method with the classic and existing state-of-the-art person re-id methods including SDALF [10], Saliency [28], LOMO+XQDA [17], SCSP [8], DNS [26] and GOG [20] which are based on single static image, as well as DTDL [13], PaMM [9] and STA [18] based on videos.

As shown Table 1, the proposed method outperforms all the classic and state-of-art methods and the performance enhancement has been achieved at all ranks $r = 1, 5, 10, 20$. For PRID2011 dataset, Rank 1 accuracy is 81.4% with an improvement of 22.2% compared with the baseline method GOG and 17.3% compared with the second best result by STA; Rank 20 accuracy is 99.7% which means the true correspondence is almost 100% in top 20 matches. For CYBJ-G dataset, the Rank 1 accuracy is 89.4%, which is 12.8% greater than the baseline (and the second best) method GOG; the Rank 5, 10 and 20 accuracies are 98.8%, 99.9% and 100.0% respectively, nearly 100% accuracy for top 5 matches and truly 100% for top 20 matches. The proposed framework improves the person re-id accuracy greatly compared with the state-of-art methods.

6 Conclusion

In this paper, we proposed a novel key person based person re-id framework. It tackles the person re-id difficulty in significantly different way from the existing methods: 1) it models the pedestrian set as temporally ordered pedestrian flow, and the temporal orders between pedestrians keep relatively stable when the flow passes adjacent camera views by pedestrian velocity constrain; 2) key persons selected and first re-identified, then help to locate the possible candidates of the query person using their temporal distance defined by the difference of their entering time. The experiments show the proposed framework greatly improve the re-id accuracy and outperforms all the existing state-of-the-art person re-id methods.

Acknowledgment

We would like to thank Dr. Shijun Wang and Dr. Jianbin Jia for the fruitful discussion, and Zhenfeng Fan for carefully proof reading. This work is supported by the National Key R&D Program of China under Grant 2017YFC0803505, China Postdoctoral Science Foundation, and National Natural Science Foundation of China with Grant No.61571438.

References

- [1] Loris Bazzani, Marco Cristani, Alessandro Perina, and Vittorio Murino. Multiple-shot person re-identification by chromatic and epitomic analyses. *Pattern Recognition Letters*, 33(7):898–903, 2012.
- [2] A Bialkowski, P Lucey, Xinyu Wei, and S Sridharan. Person re-identification using group information. In *International Conference on Digital Image Computing: Techniques and Applications*, pages 1–6, 2013.
- [3] S Bk, Awomir, Corv, . Etienne, E., Francois Mond, and Monique Thonnat. Boosted human re-identification using riemannian manifolds. *Image & Vision Computing*, 30(6-7):443–452, 2012.
- [4] Yinghao Cai, Valtteri Takala, and Matti Pietik?inen. Matching groups of people by co-variance descriptor. In *International Conference on Pattern Recognition*, pages 2744–2747, 2010.
- [5] Dapeng Chen, Zejian Yuan, Badong Chen, and Nanning Zheng. Similarity learning with spatial constraints for person re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1268–1277, 2016.
- [6] De Cheng, Yihong Gong, Sanping Zhou, Jinjun Wang, and Nanning Zheng. Person re-identification by multi-channel parts-based cnn with improved triplet loss function. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1335–1344, 2016.
- [7] Yeong Jun Cho and Kuk Jin Yoon. Improving person re-identification via pose-aware multi-shot matching. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1354–1362, 2016.

- [8] R. T. Collins, Weina Ge, and R. B. Ruback. Vision-based analysis of small groups in pedestrian crowds. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 34 (5):1003–1016, 2012.
- [9] M. P. Dubuisson and A. K. Jain. A modified hausdorff distance for object matching. In *Iapr International Conference on Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing*, pages 566–568 vol.1, 2002.
- [10] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2360–2367, 2010.
- [11] Niloofar Gheissari, Thomas B Sebastian, and Richard Hartley. Person reidentification using spatiotemporal appearance. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1528–1535, 2006.
- [12] Martin Hirzer, Csaba Beleznaï, Peter M Roth, and Horst Bischof. Person re-identification by descriptive and discriminative classification. In *Scandinavian Conference on Image Analysis*, pages 91–102, 2011.
- [13] Srikrishna Karanam, Yang Li, and Richard J. Radke. Person re-identification with discriminatively trained viewpoint invariant dictionaries. In *IEEE International Conference on Computer Vision*, pages 4516–4524, 2015.
- [14] Wei Li and Shishir K. Shah. Subject centric group feature for person re-identification. In *Computer Vision and Pattern Recognition Workshops*, pages 28–35, 2015.
- [15] Wei Li, Rui Zhao, Tong Xiao, and Xiaogang Wang. Deepreid: Deep filter pairing neural network for person re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 152–159, 2014.
- [16] Shengcai Liao and Stan Z. Li. Efficient psd constrained asymmetric metric learning for person re-identification. In *IEEE International Conference on Computer Vision*, pages 3685–3693, 2016.
- [17] Shengcai Liao, Yang Hu, Xiangyu Zhu, and Stan Z. Li. Person re-identification by local maximal occurrence representation and metric learning. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2197–2206, 2015.
- [18] Kan Liu, Bingpeng Ma, Wei Zhang, and Rui Huang. A spatio-temporal appearance representation for video-based pedestrian re-identification. In *IEEE International Conference on Computer Vision*, pages 3810–3818, 2015.
- [19] Bingpeng Ma, Qian Li, and Hong Chang. *Gaussian Descriptor Based on Local Features for Person Re-identification*. Springer International Publishing, 2014.
- [20] Tetsu Matsukawa, Takahiro Okabe, Einoshin Suzuki, and Yoichi Sato. Hierarchical gaussian descriptor for person re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1363–1372, 2016.
- [21] Sakrapee Paisitkriangkrai, Chunhua Shen, and Anton Van Den Hengel. Learning to rank in person re-identification with metric ensembles. pages 1846–1855, 2015.

- [22] Hailin Shi, Yang Yang, Xiangyu Zhu, Shengcai Liao, Zhen Lei, Weishi Zheng, and Stan Z. Li. Embedding deep metric for person re-identification: A study against large variations. 2016.
- [23] Zhiyuan Shi, Timothy M. Hospedales, and Tao Xiang. Transferring a semantic representation for person re-identification and search. In *Computer Vision and Pattern Recognition*, 2015.
- [24] Norimichi Ukita, Yusuke Moriguchi, and Norihiro Hagita. People re-identification across non-overlapping cameras using group features. *Computer Vision & Image Understanding*, 144(C):228–236, 2016.
- [25] Rahul Rama Vavior, Mrinal Haloi, and Gang Wang. *Gated Siamese Convolutional Neural Network Architecture for Human Re-identification*. Springer International Publishing, 2016.
- [26] Li Zhang, Tao Xiang, and Shaogang Gong. Learning a discriminative null space for person re-identification. pages 1239–1248, 2016.
- [27] Wei Zhang, Bingpeng Ma, Kan Liu, and Rui Huang. Video-based pedestrian re-identification by adaptive spatio-temporal appearance model. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, PP(99):1–1, 2017.
- [28] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Unsupervised salience learning for person re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3586–3593, 2013.
- [29] Wei Shi Zheng, Shaogang Gong, and Tao Xiang. Associating groups of people. In *British Machine Vision Conference, BMVC 2009, London, UK, September 7-10, 2009. Proceedings*, 2009.
- [30] Wei Shi Zheng, Shaogang Gong, and Tao Xiang. Person re-identification by probabilistic relative distance comparison. In *The IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, Colorado Springs, Co, Usa, 20-25 June*, pages 649–656, 2011.