A Preliminary Study on Optimizing Person Re-identification using Stable Marriage Algorithm

Nik Mohd Zarifie Hashim*‡, Yasutomo Kawanishi*, Daisuke Deguchi†, Ichiro Ide* and Hiroshi Murase*
*Graduate School of Informatics, Nagoya University, Japan
Email: hashimz@murase.is.i.nagoya-u.ac.jp, {kawanishi,ide,murase}@i.nagoya-u.ac.jp, ddeguchi@nagoya-u.jp
†Information Strategy Office, Nagoya University, Japan
Email: ddeguchi@nagoya-u.jp
‡Centre of Telecommunication Research & Innovation (CeTRI),
Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Malaysia

***Abstract*—Person re-identification gains an increasing interest in the surveillance image processing field due to its' importance for security. Most approaches to solve the person re-identification problem match persons one-by-one. However, redundant matching where one of the person is selected for the matching pair several times often occurs. It also degrades the overall image matching performance. To overcome the issue, in this paper, we propose a method which solves the person re-identification problem for multiple persons simultaneously. Instead of one-by-one matching, we consider person re-identification as an instance of the Stable Marriage Problem (SMP). The result of an experiment showed that the proposed method outperforms some of the existing state-of-the-art methods applied to the VIPeR dataset.**

***Index Terms*—Image Matching, Person re-identification, Stable Marriage Problem**

I. Introduction

A surveillance system is important for monitoring people in a wide area. The problem of matching persons across non-overlapping camera-views at different places and times is referred to as *person re-identification*. Usually, person images captured from a certain camera-view are considered as query person images, and person images captured from other camera-views are considered as gallery person images. In case of multiple person tracking across non-overlapping camera-views, person re-identification between adjacent camera pairs is required. In the case, images captured at a camera-view are considered as query images and images captured by the other camera-view as gallery images.

To solve the person re-identification problem, the majority of previous methods find the best-match gallery image for each query person image one-by-one. By following this matching scheme, many researchers have proposed various image features and metric learning methods [1]. Since some of the same gallery images can be matched several times, mismatching often occurs when we match persons by using the matching scheme. An example of matching results between two adjacent camera-views (Cameras A and B) by the scheme is illustrated in Figure 1. We can clearly see both Persons 1 and 3 are matched to the Person 3 in the gallery images. However, Person 2 is mistakenly matched to Person 1. It makes a redundant matching for Person 3 since Person 3 in the gallery is selected twice. This mismatched result will influence the overall image matching performance. The goal of our research is to improve the matching rate by introducing the Stable Marriage Problem (SMP) as an instance to solve such a problem in person re-identification.

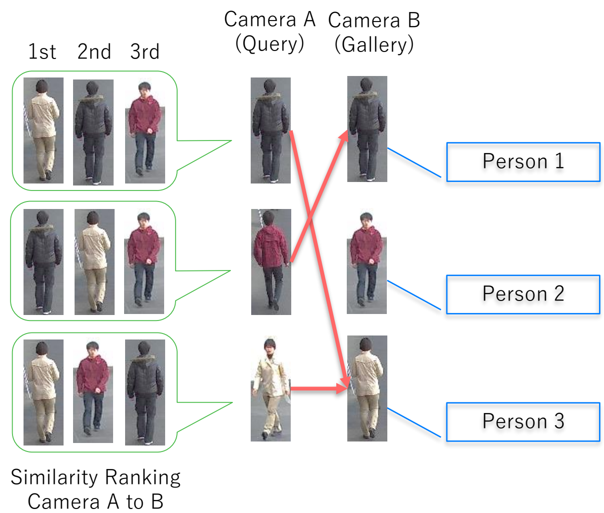

Figure 1. Example of a result of one-by-one person matching. We used the Shinpukan dataset images for example [2]. As for the notation, Person 1 wears a black jacket, Person 2 wears a red jacket, and Person 3 wears a linen white jacket. We can see that there is a redundant matching.

To overcome the redundant matching problem, we introduce a global optimal matching to the simultaneous matching which can avoid the mismatch and redundant matching. In this paper, we propose a new optimal matching method by interpreting the problem as an instance of the SMP. We show an improvement from the graph-based approach proposed by Zhang et al. [3] which solve the person re-identification problem simultaneously for multiple persons matching. We

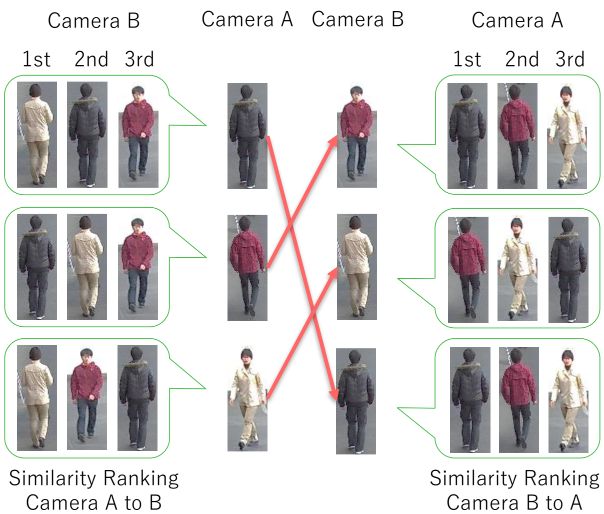

Figure 2. Example of a result of the proposed person re-identification scheme. They are simultaneously matched by the Stable Marriage Algorithm [4].

will introduce the Stable Marriage Algorithm (SMA) which has been proposed by Gale and Shapley [4] to the person re-identification problem. SMA was originally used as a matching tool for finding the most suitable pair for an element from two element's sets. We consider the query images and the gallery images as the two element sets and then apply the matching algorithm to solve the person re-identification problem for all images simultaneously. An example of matching results between the two adjacent camera views by our proposed method is shown in Figure 2. Thanks to the optimal matching algorithm, we can see there are no redundant matchings.

The main contributions of this research can be summarized as follows:

- We introduce a global optimal matching scheme for simultaneous image matching using SMA.
- Our proposed method outperformes some of existing methods. Even though the proposed method uses a simple image feature, it gains competitive result.

The rest of this paper is structured as follows: In Section II, related work that have been previously proposed will be discussed. After that, we will explain the proposed method which uses SMA in detail. Section IV will discuss the evaluation results. Finally, we conclude our paper in Section V.

II. Related Work

In the past few years, many researchers have proposed new methods and approaches to tackle various difficulties in person re-identification such as different view angles, lighting conditions, occlusion with other objects, and cluttered background. The following two topics are the main focus of traditional person re-identification research; image features and metrics. Image feature has been the focused topic from the beginning of person re-identification studies. Researach began by designing the best hand-crafted image feature to achieve better matching. Many researchers have been proposed to describe a person

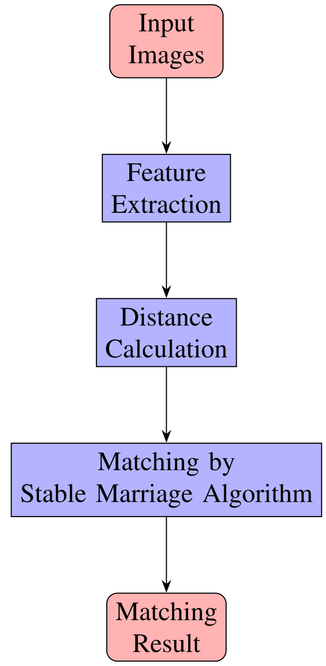

Figure 3. Process-flow of the proposed matching algorithm.

image [5–7]. All the single feature studies were limited to several conditions e.g image viewpoint.

Unfortunately, all these approaches which were focused on new features faced the problem of the dataset bias. Several features were good at some datasets and illumination but not for all datasets. To tackle this problem, another topic which focuses on the metric in a given feature space is also actively developed. Several metric learning methods have been proposed [8–14]. Besides using image features and metric learning in person re-identification, thanks to the recent trend of Deep Learning in computer vision, Deep Learning based approaches have proposed to handle feature modeling and metric learning simultaneously [15]. Ahmed et al. proposed a CNN-based method to achieve this [15]. Their approach with Deep Learning somehow showed a good performance with the help of large scale of datasets. Liu et al. proposed a triplet-loss based CNN model to find good features and metrics [16]. Similar approaches have been introduced in [17–23]. Applying Deep Learning in the person re-identification, however, struggles with the image labeling problem in the training process. It will consume more time as the labeling of training data is done manually. Moreover, in Deep Learning, for achieving a good result, a big number of images is inevitable. Considering this problem, it is still important to have an approach which can utilize a small dataset to realize person re-identification.

Most of the previous researches are based on an one-by-one person matching scheme, which finds the best match from the gallery for each query image. As we mentioned in the previous section, redundant matching occurres in this scheme. Recently, Zhang et al. proposed a person re-identification approach by utilizing a bipartite graph matching problem [3]

Algorithm 1: Stable Marriage Algorithm (SMA)

initialize all $m \in M$ and $w \in W$ as *free*, M is a list of all available men, and W is a list of all available women;
while $\exists$ *free* man m who still has a woman w to propose to **do**
  w = first woman in m's preference list to whom m has not yet been proposed;
  if w is free **then**
    (m,w) becomes engaged;
  else
    $\exists$ some pair (m',w) has been already engaged;
    if w prefers m to m' **then**
      m' becomes free;
      (m,w) becomes engaged;
    else
      (m',w) remains engaged;
    end
  end
end

which combines the ideas of Taskar et al. and Zhang et al. [24, 25]. Their method, namely, Person Re-identification via Structured Matching (PRISM) considers each person as a node in a bipertite graph, performing two sets of nodes for each query and gallery image sets, respectively. It performs person re-identification by utilizing the simultaneous person matching.

III. Optimal Matching based on Stable Marriage Problem

A. Overview

We propose a person re-identification method, which can avoid the redundant matchings. We introduce a new approach to the simultaneous image matching technique by considering person re-identification as an instance of Stable Marriage Problem (SMP). Our proposed method finds the stable matching, which identifies each query image with its matched gallery image pair simultaneously without any conflict. Compared to PRISM, which is based on a bipartite graph matching, the proposed method only requires similarity order of gallery images for each query image, and find the stable matchings.

The process-flow of the proposed method is shown in Figure 3. Although the method requires image feature extraction and distance calculation, since the matching part is the main issue, we first explain the matching process. After that, we will explain the details of feature extraction and distance calculation in the following subsections.

B. Stable Marriage Problem (SMP)

SMP was proposed by D. Gale and L. S. Shapley [4]. They describe problems related to matching issues in the college admission and marriage as Algorithm 1. SMP defines Definitions 1 and 2 below. In the definitions, we assume that there are man 1, 2, ... and woman 1, 2, ..., and each of them has a preference list of the opposite sex to propose to.

Algorithm 2: Person re-identification using the SMA

initialize all $I_i^A \in I^A$ and $I_i^B \in I^B$ as *free*;
while $\exists$ *free* I_i^A which still has an I_i^B to be matched with **do**
  I_i^B = first rank in I_i^A's list to which I_i^A has not yet been matched;
  if I_i^B is free **then**
    (I_i^A,I_i^B) becomes matched;
  else
    $\exists$ pair (I_i^A',I_i^B) already matched;
    if I_i^B is more similar to I_i^A than I_i^A' **then**
      I_i^A' becomes free;
      (I_i^A,I_i^B) becomes matched;
    else
      (I_i^A',I_i^B) remains matched;
    end
  end
end

Definition 1 *A matching of a marriage will be called unstable if there are two persons, man 1 and man 2 who are assigned to woman 1 and woman 2, respectively, although man 2 prefers woman 1 to woman 2 and woman 1 prefers man 2 to man 1.*

Definition 2 *A stable marriage is called optimal if every person is at least as well off under it as under any other stable matching.*

The Stable Marriage Algorithm (SMA) will give a stable matching pair by rematching an existing pair if it showed less rank preference by either the man or woman who is already existed.

C. Stable Marriage Problem in Person Re-identification

We formulated person re-identification across two cameras as an instance of SMP by using given person images I^A and I^B detected at two camera-views (Cameras A and B), respectively. As the SMA, we obtain a similarity ranking list for each person image.

$$I^A = \{I_1^A, I_2^A, \ldots, I_N^A\} \quad (1)$$

$$I^B = \{I_1^B, I_2^B, \ldots, I_N^B\} \quad (2)$$

Tables I and II show three images from Camera A and three images from Camera B regarding to Figure 2 as an example. Here, we assume that all the persons are detected only once for each camera-view without redundancy. For matching these three images, first, images in I^B and I^A will be ranked based on the distance with each image in I^A and I^B in ascending order. The proposed algorithm, SMA for person re-identification is shown in Algorithm 2. In SMA, matching will

Table I
EXAMPLE OF IMAGE SIMILARITY FROM I^A TO I^B BASED ON FEATURE SIMILARITY

I^A Image List			
Query Image	First Rank	Second Rank	Third Rank
I_1^A	I_3^B	I_1^B	I_2^B
I_2^A	I_1^B	I_3^B	I_2^B
I_3^A	I_3^B	I_2^B	I_1^B

Table II
EXAMPLE OF IMAGE SIMILARITY FROM I^B TO I^A BASED ON FEATURE SIMILARITY

I^B Image List			
Query Image	First Rank	Second Rank	Third Rank
I_1^B	I_1^A	I_2^A	I_3^A
I_2^B	I_2^A	I_3^A	I_1^A
I_3^B	I_1^A	I_2^A	I_3^A

be considered by looking up both Tables I and II alternately. First, matching I_1^A and I_3^B is performed as I_3^B is the best choice for I_1^A and I_1^A is also the first rank for I_1^B. Then, I_2^A is to be matched with I_1^B. As I_2^A is the second rank in I_1^B and it is still not matched with any image yet, they are matched. Then I_3^A is to be matched with I_3^B, but here I_3^B is already matched with I_1^A. In this case, I_3^A remains not matched in the first loop and I_3^B is removed from the rank list of I_3^A. In the second loop, I_3^A which is not matched yet is to be matched with I_2^B. Finally, we have three matching images pairs which are $\{I_1^A\ ,\ I_3^B\}$, $\{I_2^A\ ,\ I_1^B\}$, and $\{I_3^A\ ,\ I_2^B\}$.

Here, we can see the matching for a query image has been matched to all gallery images simultaneously by considering the similarity rank of each image. It brings us a new versatile approach compared to the conventional person re-identification which just implements a normal one-by-one image matching for the targeted person image. Applying the SMA in person re-identification directed us to optimize our image matching by avoiding redundant image matches.

D. Feature and Similarity Metric

As the objective of our study is to optimize person re-identification based on SMA, we are not focusing to design a new image feature. Thus, here we use a simple HSV color histogram as the image feature in order to handle the visual ambiguity and illumination issue [26]. A histogram intersection method similar to [27] is applied for image comparison to calculate the distance between the images accordingly. To concentrate only on the distinctive area of an image, we crop all images as a pre-processing at the center half of height and width and several other center positions as illustrated in Figure 4. We studied the effect of frame selection in having a good performance of image matching using SMA. An M-bin HSV color histogram is calculated from each image as $[\mathbf{x}_1^A, \mathbf{x}_2^A, \ldots, \mathbf{x}_N^A]$ and $[\mathbf{x}_1^B, \mathbf{x}_2^B, \ldots, \mathbf{x}_N^B]$ before the similarity calculation between all images as N, where

$$\mathbf{x}_i^A = f(I_i^A), \quad i = 1, 2, ..., N, \tag{3}$$

$$\mathbf{x}_j^B = f(I_j^B), \quad j = 1, 2, ..., N. \tag{4}$$

The image feature similarity of two images will be calculated by the similarity function $s(I_i^A, I_j^B)$ which is defined as

$$s(I_i^A, I_j^B) = s_f(x_i^A, x_j^B), \tag{5}$$

where $s_f(x_i^A, x_j^B)$ is the histogram-intersection defined as

$$s_f(x_i^A, x_j^B) = \sum_{k=1}^{M} \min(x_{ik}^A, x_{jk}^B). \tag{6}$$

Using this appearance similarity, we sort the images and then apply SMA to find the matched image pairs by using simultaneous image matching.

We also use the practically used image feature for person re-identification, which is Symmetry Driven Accumulation of Local Features (SDALF) proposed by Farenzena et al. [28]. In case of SDALF, it is difficult to be extracted from cropped image because the method first detect the symmetry axes of the target person, we just use the whole image without cropping. The method accumulates chromatic (histograms), region-based (blobs), and edge-based (contours, textures) information.

IV. EVALUATION

A. Dataset

In person re-identification study there are many datasets available publicly [2], [29], and [30]. We compare the matching rate performance with the existing method in person re-identification using the most exciting and competitive dataset which is Viewpoint Invariant Pedestrian Recognition (VIPeR) dataset [31]. The VIPeR dataset consists images captured from two different camera-views. The dataset contains 632 individual, and each of them has one image taken from each of the two cameras. The dataset was collected in order to test viewpoint invariant pedestrian recognition. The cameras have different viewpoints and illumination variations in the pair images taken. The images are then cropped and scaled to the size of 128×48 pixels. This is one of the most challenging datasets in the person re-identification study.

B. Comparative Methods

We used Symmetry-driven Accumulation of Local Features (SDALF) as the baseline method [28] in our comparison study. In addition, we compare with PRISM as it is the state-of-the-art method of simultaneous image matching. Apart from these two methods mentioned, we also compare with other state-of-the-art methods that take the one-by-one matching approach; eSDC [32], Mid-level features [33], and Mid-level LADF [33]. In this paper, we compare several image features for the proposed method; HSV + SMA , SDALF + SMA, and Crop HSV + SMA. For the HSV color histogram feature, we used an M=16-bins histogram. Here, the distance between images was compared by histogram intersection.

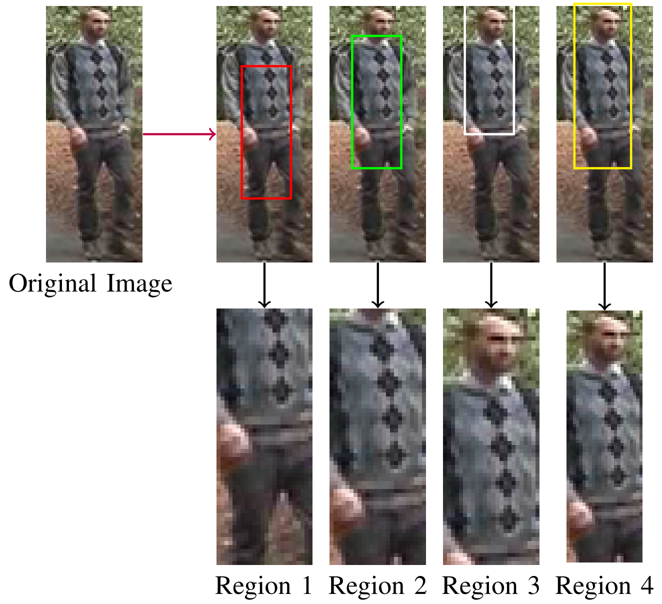

Figure 4. Example of image cropping with four different positions for a person.

Table III
COMPARISON OF MATCHING RATES WITH EXISTING METHODS

Matching Method	Comparison Method	Matching Rate
One-by-one Image Matching	SDALF[28]	19.87 %
	eSDC[32]	26.31 %
	Mid-Level[33]	29.11 %
	Mid-Level LADF[33]	43.39 %
Simultaneous Image Matching	PRISM-I[3]	35.76 %
	PRISM-II[3]	36.71 %
	PRISM-III[3]	35.44 %
	Proposed HSV + SMA	33.48 %
	Proposed SDALF + SMA	**40.20 %**

Table IV
COMPARISON ON DIFFERENT CROPPING METHODS USING SMA + HSV METHOD

Comparison Method	Matching Rate (%)
Region 1	45.57 %
Region 2	**46.33 %**
Region 3	44.72 %
Region 4	45.52 %
Original Image	33.48 %

C. Evaluation Settings

We follow the experimental set-up described in [28] using the VIPeR dataset images. We randomly selected 316 person images for evaluation. The evaluation was performed ten times and their results were averaged. The matching rate was calculated from the ratio of successfully matched image pairs from two cameras at the first rank identification after applying our proposed SMA.

D. Comparison of Matching Rate

The evaluation results are shown in Table III. This table presents two categories of image matching schemes: one-by-one person matching in the upper-half and the simultaneous person matching in the lower-half of the table. The numbers shown in Table III are the matching rates for all available images matched by each method. In this evaluation, image features are extracted from the whole image without cropping. We achieved the best matching rate in the simultaneous person matching category with 40.20%. Applying the SMA was the main factor rewarding the high performance of the simultaneous person matching. While in the one-by-one person matching category, we ranked in the second after the Mid-Level LADF method [33].

E. Effect of the Image Cropping

Next, we studied the effect of image cropping for HSV color feature. We applied four types of cropping area with the original image as illustrated in Figure 4. In this experiment, we used the HSV + SMA method as to compare all image frames. VIPeR dataset images have mostly been captured with noise in the background such as green trees and black background, which gave bad influence to most of their images. Region 2 avoided this affected area and produced a good matching rate. We can conclude that by cropping images especially with Region 2, we can achieve the best matching rate compared to other state-of-the-art methods in Table IV. It might also be a good alternative to have a good matching rather than using the traditional person re-identification approaches.

V. CONCLUSION

We proposed a new method by interpreting the person re-identification problem as an instance of the Stable Marriage Problem (SMP). Our proposed method using the Stable Marriage Algorithm (SMA) in person re-identification achieved the best performance among the simultaneous person matching methods successfully. It avoids conflict matching pairs in one-by-one person matching. We showed that our approach outperforms the other methods through an experiment using a public dataset. For future improvement, since the image feature utilized by PRISM [3] might be helping them to achieve a better result in their case, we will reconsider using it.

ACKNOWLEDGMENT

The authors would like to thank Minister of Higher Education of Government of Malaysia and Universiti Teknikal Malaysia Melaka. Parts of this research were supported by MEXT, Grant-in-Aid for Scientific Research.

REFERENCES

[1] A. Bedagkar-Gala and S. K. Shah, "A survey of approaches and trends in person re-identification," *Image and Vision Computing*, vol. 32, no. 4, pp. 270–286, Apr. 2014.

[2] Y. Kawanishi, Y. Wu, M. Mukunoki, and M. Minoh, "Shinpuhkan2014: A multi-camera pedestrian dataset for tracking people across multiple cameras," in *Proc. 20th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, vol. 5, Feb. 2014, pp. 1–6.

[3] Z. Zhang and V. Saligrama, "Prism: Person reidentification via structured matching," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 27, no. 3, pp. 499–512, Mar. 2017.

[4] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, Jan. 1962.

[5] O. Javed, Z. Rasheed, K. Shafique, and M. Shah, "Tracking across multiple cameras with disjoint views," in *Proc. 9th IEEE Conf. on Computer Visions*, vol. 2, Oct. 2003, pp. 952–957.

[6] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, "Shape and appearance context modeling," in *Proc. 11th IEEE Conf. on Computer Vision*, Oct. 2007, pp. 1–8.

[7] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *Proc. 10th European Conf. on Computer Vision*, Oct. 2008, pp. 262–275.

[8] M. Hirzer, P. M. Roth, and H. Bischof, "Person re-identification by efficient impostor-based metric learning," in *Proc. 2012 IEEE Conf. on Advanced Video and Signal-Based Surveillance*, Sept. 2012, pp. 203–208.

[9] M. Koestinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in *Proc. 2012 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2012, pp. 2288–2295.

[10] F. Xiong, M. Gou, O. Camps, and M. Sznaier, "Person re-identification using kernel-based metric learning methods," in *Proc. 13th European Conference on Computer Vision*, Sept. 2014, pp. 1–16.

[11] D. Tao, L. Jin, Y. Wang, Y. Yuan, and X. Li, "Person re-identification by regularized smoothing kiss metric learning," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 23, no. 10, pp. 1675–1685, Mar. 2013.

[12] S. Paisitkriangkrai, C. Shen, and A. van den Hengel, "Learning to rank in person re-identification with metric ensembles," in *Proc. 2015 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2015, pp. 1846–1855.

[13] W. Li, Y. Wu, and J. Li, "Re-identification by neighborhood structure metric learning," *Pattern Recognition*, vol. 61, pp. 327–338, 2017.

[14] C. Sun, D. Wang, and H. Lu, "Person re-identification via distance metric learning with latent variables," *IEEE Trans. Image Processing*, vol. 26, no. 1, pp. 23–34, Oct. 2017.

[15] E. Ahmed, M. Jones, and T. K. Marks, "An improved deep learning architecture for person re-identification," in *Proc. 2015 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2015, pp. 3908–3916.

[16] J. Liu, Z.-J. Zha, Q. Tian, D. Liu, T. Yao, Q. Ling, and T. Mei, "Multi-scale triplet CNN for person re-identification," in *Proc. 24th ACM Multimedia Conf.*, Oct. 2016, pp. 192–196.

[17] T. Xiao, H. Li, W. Ouyang, and X. Wang, "Learning deep feature representations with domain guided dropout for person re-identification," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2016, pp. 1249–1258.

[18] L. Zhang, T. Xiang, and S. Gong, "Learning a discriminative null space for person re-identification," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2016, pp. 1239–1248.

[19] D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng, "Person re-identification by multi-channel parts-based CNN with improved triplet loss function," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2016, pp. 1335–1344.

[20] W. Yang, W. Ouyang, H. Li, and X. Wang, "End-to-end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation," in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2016, pp. 3073–3082.

[21] S.-Z. Chen, C.-C. Guo, and J.-H. Lai, "Deep ranking for person re-identification via joint representation learning," *IEEE Trans. Image Processing*, vol. 25, no. 5, pp. 2353–2367, Mar. 2016.

[22] T. Xiao, S. Li, B. Wang, L. Lin, and X. Wang, "Joint detection and identification feature learning for person search," in *Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition*, July 2017, pp. 3415–3424.

[23] Y. Liu, H. Li, and X. Wang, "Learning deep features via congenerous cosine loss for person recognition," *arXiv preprint arXiv:1702.06890*, 2017.

[24] B. Taskar, V. Chatalbashev, D. Koller, and C. Guestrin, "Learning structured prediction models: A large margin approach," in *Proc. 22nd Int. Conf. on Machine Learning*, Aug. 2005, pp. 896–903.

[25] Z. Zhang, Y. Chen, and V. Saligrama, "A novel visual word co-occurrence model for person re-identification," *arXiv preprint, arXiv:1410.6532*, 2014.

[26] S. Sural, G. Qian, and S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval," in *Proc. IEEE 2002 Int. Conf. on Image Processing*, vol. 2, Sept. 2002, pp. 589–592.

[27] D. N. T. Cong, C. Achard, L. Khoudour, and L. Douadi, "Video sequences association for people re-identification across multiple non-overlapping cameras," in *Proc. 15th Int. Conf. on Image Analysis and Processing*, Nov. 2009, pp. 179–189.

[28] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *Proc. 2010 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2010, pp. 2360–2367.

[29] A. Ess, B. Leibe, and L. Van Gool, "Depth and appearance for mobile scene analysis," in *Proc. 11th IEEE Int. Conf. on Computer Vision*, Oct. 2007, pp. 1–8.

[30] W. Li, R. Zhao, and X. Wang, "Human reidentification with transferred metric learning," in *Proc. 11th Asian Conf. on Computer Vision*, Nov. 2012, pp. 31–44.

[31] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *Proc. 2007 IEEE Workshop on Performance Evaluation for Tracking and Surveillance*, vol. 3, no. 5, Oct. 2007, pp. 1–7.

[32] R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in *Proc. 2013 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2013, pp. 3586–3593.

[33] R. Zhao and W. Ouyang and X. Wang, "Learning mid-level filters for person re-identification," in *Proc. 2014 IEEE Conf. on Computer Vision and Pattern Recognition*, June 2014, pp. 144–151.