

Improving Person Re-Identification by Soft Biometrics Based Reranking

Le An, Xiaojing Chen, Mehran Kafai, Songfan Yang, Bir Bhanu

Center for Research in Intelligent Systems, University of California, Riverside

lan004@ucr.edu, xchen010@ucr.edu, mkafai@cs.ucr.edu, syang027@ucr.edu, bhanu@cris.ucr.edu

Abstract—The problem of person re-identification is to recognize a target subject across non-overlapping distributed cameras at different times and locations. The applications of person re-identification include security, surveillance, multi-camera tracking, etc. In a real-world scenario, person re-identification is challenging due to the dramatic changes in a subject’s appearance in terms of pose, illumination, background, and occlusion. Existing approaches either try to design robust features to identify a subject across different views or learn distance metrics to maximize the similarity between different views of the same person and minimize the similarity between different views of different persons. In this paper, we aim at improving the re-identification performance by reranking the returned results based on soft biometric attributes, such as gender, which can describe probe and gallery subjects at a higher level. During reranking, the soft biometric attributes are detected and attribute-based distance scores are calculated between pairs of images by using a regression model. These distance scores are used for reranking the initially returned matches. Experiments on a benchmark database with different baseline re-identification methods show that reranking improves the recognition accuracy by moving upwards the returned matches from gallery that share the same soft biometric attributes as the probe subject.

I. INTRODUCTION

Person re-identification is a recognition task to match the individuals across cameras in disjoint views. Often in surveillance camera systems, person re-identification is desired for practical purposes such as security monitoring. In addition, person re-identification can facilitate other tasks. For instance, tracking people in multi-camera systems may use the output of person re-identification for across-camera track association [1]. In recent years, person re-identification has become a popular research topic [2] [3] [4] [5] [6] [7] [8].

In practice, person re-identification is a challenging topic and high recognition rate is difficult due to but not limited to the following reasons:

- *Low image quality.* Normally the images captured by surveillance cameras are of low resolution and image noise level may be significant.
- *Varying illumination.* Since the images of a subject are captured at different times and locations, the illumination condition may change dramatically, which significantly affects the appearance of the subject.
- *Changing pose.* In different camera views, the poses of a subject may change arbitrarily due to the free movement of the subject.
- *Occlusion.* The accessories associated with a subject such as hat or suitcase may block part of the subject.



Fig. 1. Two sample re-identification results with top 5 returned matches shown. Probe and gallery images are from two different camera views from VIPeR database [2]. Due to large variations in pose, illumination and background, the appearance of the subjects differs significantly in two views. Green bounding boxes show the probe subjects and the returned subjects in red bounding boxes are correct matches. In both cases, the correct matches failed to appear at top 1. However, if soft biometric attributes (*male* for the upper sample and *carrying* for the lower sample) are detected and used for reranking, the correct matches will be moved to top 1.

In general the solutions to person re-identification fall into two categories. In the first category, robust features are designed and extracted from probe and gallery images and matching is based on measuring the distance between the probe and gallery images [2] [3] [4] [5]. However, due to significant appearance change across cameras, extracting features that are invariant to pose, illumination and occlusion is not an easy task. Therefore, accurate recognition is very difficult. For instance, the state-of-the-art performance is less than 30% for rank 1 re-identification rate [9] on VIPeR database [2].

On the other hand, methods in the second category opt to learn the optimal distance measure for the image pairs using metric learning techniques [7] [6] [10] [11] [12] [13]. During the learning process the intra-class distances (image pairs of same subjects) are minimized while the inter-class distances (image pairs of different subjects) are maximized.

Inspired by the idea of reranking that has been widely applied in information retrieval [14] [15] [16] [17], we propose to use reranking as a post processing step to improve the recognition accuracy upon the results by baseline re-identification methods. In this paper, reranking is based on soft

biometric (SB) attributes. SB attributes can be considered as semantic information and mining semantics or context from image has become a popular topic [18] [19]. For instance, SB attributes are used for face clustering in [20]. In re-identification, although the returned gallery subjects may have similar appearance to a probe subject, SB attributes such as gender can be used to distinguish the correct match from other confusing subjects, thus improving the re-identification accuracy. Figure 1 shows two sample cases in which the attributes *male* and *carrying* help to exclude some top returned matches. In the proposed reranking method, SB attributes are first detected and then SB distance scores are calculated using a trained regression model. The SB distance scores are used to rerank the initial re-identification results.

In the following, Section 2 describes the proposed reranking method. In Section 3 the experimental results are reported. We conclude this paper in Section 4.

A. Related Work

One way to tackle the challenge of re-identification is to extract robust features from different camera views. Chen *et al.* [21] use the pictorial structures to localize the human parts and part-to-part correspondences are searched to match the subjects. Farenzena *et al.* [3] extract features that account for the overall chromatic content, the spatial arrangement and the presence of recurrent local motifs to match individuals with appearance variation. In [22], a model is learned in a covariance metric space to select features based on the idea that different regions for each subject should be matched specifically. Gray *et al.* [23] use AdaBoost to select the most discriminative features instead of using handcrafted features. The re-identification is formulated as a ranking problem with the development of an ensemble RankSVM (ERSVM) in [24]. A two-step method is proposed in [25] by first using a descriptive model to obtain an initial ranking which is refined in the second step by a discriminative model with human feedback. In [26], an attribute-centric and part-based feature representation is learned to better discriminate visual appearance of people in different camera views. Recently, Zhao *et al.* [27] apply an adjacency constrained patch matching scheme to establish dense correspondence between image pairs and human saliency is learned in an unsupervised manner to weigh the matching for re-identification.

Another strategy is to use metric learning techniques. For person re-identification, metric learning methods have been extensively studied in recent years. Hirzer *et al.* [6] propose a relaxed pairwise metric learning (RPLM) based on Mahalanobis distance learning which takes advantages of the structure of the data with reduced computational cost, achieving the state-of-the-art with simple feature descriptors. In [10] a simple yet effective method to learn the distance metric based on a statistical inference perspective is proposed. Zheng *et al.* [7] formulate re-identification as a relative distance comparison (PRDC) problem which aims to maximize the likelihood that the distance between a pair of images of the same person is smaller than a pair of images of different people. The standard metric learning techniques such as Large Margin Nearest

Neighbor (LMNN) [11], Information Theoretic Metric Learning (ITML) [13], and Logistic Discriminant Metric Learning (LDML) [12] are also applicable to person re-identification. In [28] a variant of LMNN is proposed by introducing a rejection option to the unfamiliar matches (LMNN-R). In [29], different visual metrics are learned for different candidate sets instead of having a fixed metric to match all the subjects. Loy *et al.* [30] propose an unsupervised manifold ranking using unlabeled data and it is demonstrated that combining existing metric learning methods with manifold ranking helps to boost the recognition performance. A reference-based approach is developed recently [8]. Instead of comparing two images from different views directly, a reference set is used and reference-based descriptor is generated to measure the similarity between the probe and gallery images, achieving state-of-the-art performance on VIPeR database [2].

The idea of reranking has been widely used in information retrieval to improve the initial ranking performance [14] [15] [16] [17]. For example, web page reranking is used to decide whether the sites returned by search engines are highly linked or highly trusted [14]. In the application of face image retrieval, the returned candidate face images are reranked iteratively based on each candidate's average distance to the reference images [15]. In [16], features representing images from different perspectives are combined to obtain a reranking model to refine the initial rank for content based image retrieval. For video searching, Yang *et al.* [17] explore co-occurrence patterns to rerank the video search results.

II. TECHNICAL APPROACH

Figure 2 shows the system diagram of the proposed approach. The SB attributes are first detected from a probe image using the trained attribute classifiers. In the next step, the SB distances are computed between the probe and gallery images. Using the SB distances, the initially returned ranked matches are reranked to improve re-identification accuracy. Note that the proposed reranking is independent of the choice of specific person re-identification method. Therefore, it can be integrated into any existing framework for person re-identification.

A. Feature Extraction

For feature extraction we follow the scheme in [24] and [26]. Both color and texture features are extracted from

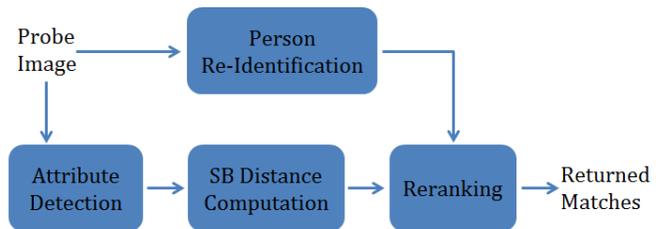


Fig. 2. System diagram of the proposed reranking approach. In parallel to the re-identification process, the soft biometric (SB) attributes are detected from the images. Based on the results of attribute detection, SB distance is computed between a probe image and a gallery image. Using the SB distances, the re-identification results are reranked.



Fig. 3. Examples of the 5 SB attributes used in this paper, including *Backpack*, *Jeans*, *Carrying*, *Short hair*, and *Male*. The subjects in two camera views from VIPeR database [2] are shown.

the original images. Each image is divided into 6 equal sized horizontal strips. For each strip, 8 color channels (RGB, HSV, and YCbCr, V and Y both represent luminance so only one channel is kept) and 21 texture feature channels (13 Schmid filters and 8 Gabor filters) are used. The bin size for each channel is 16, so the feature dimension for one image is $(8 + 21) \times 16 \times 6 = 2784$. The detail of the parameter settings for feature extraction can be found in [24].

B. SB Attributes Detection

Five SB attributes with binary annotations are used, namely *backpack*, *jeans*, *carrying*, *short hair*, and *male*. The ground truth of these attributes for the VIPeR database [2], which contains two camera views (CAM A and CAM B), are provided by the authors of [26]¹. The selection of these attributes is based on their distribution in the VIPeR database. Although other attributes (e.g., *sandals* and *skirt*) are available, due to their highly skewed distribution (i.e., only a small number of subjects wear sandals), they are not used in this paper. Table I summarizes the SB attributes used in this paper and their corresponding counts in the VIPeR database from a total number of 632 subjects. Figure 3 shows some examples of these attributes from both camera views.

Attributes	Backpack	Jeans	Carrying	Short hair	Male
# of subjects	229	221	173	308	309

TABLE I

SOFT BIOMETRICS ATTRIBUTES USED IN THIS PAPER. THE DATABASE CONTAINS A TOTAL NUMBER OF 632 SUBJECTS.

The detection of the SB attributes is formulated as a binary classification problem. Each attribute is associated with binary labels 1 and 0. A value 1 for an attribute indicates the presence of this attribute. We use Support Vector Machine (SVM) to train the attribute classifiers. Linear kernel is used and the slack variable C is chosen by cross-validation. Different classifiers are trained for different SB attributes.

C. SB Distance Computation

Given a probe image and a gallery image, the goal is to compute a score that represents the distance between this image pair in terms of their similarity according to SB

attributes. Although the SB attribute labels are predicted using the trained attribute classifiers, the predictions are prone to error and inaccurate labels may adversely affect the reranking process. Thus, instead of using predicted SB attributes directly, we formulate the SB distance computation as a regression problem.

To train the regression model, the SB distances between pairs of images are calculated as target values using ground truth annotations. We define SB distance between a probe image and a gallery image by the weighted Hamming distance

$$d(X, Y) = \frac{1}{K} \sum_{i=1}^K w_i I(X_i^A, Y_i^B) \quad (1)$$

where X_i^A is the i^{th} SB attribute of a probe image X from CAM A and Y_i^B is the i^{th} SB attribute of a gallery image Y from CAM B. $I(\cdot, \cdot)$ is an indicator function and it equals to 0 when values of an attribute (0 or 1) are the same for two images and 1 otherwise. w_i is a weighting parameter and it is defined as the reciprocal of the training accuracy of the SVM classifier for the i^{th} SB attribute. K is the total number of attributes and in our case $K = 5$. For a pair of images with exactly the same attributes, the SB distance is 0. In such a way, the SB distance between two images using multiple attributes is represented by a single number.

The features corresponding to a target value for an image pair (X from CAM A and Y from CAM B) are concatenations from the probability outputs of SVM SB attribute classifiers as an indication of the classification confidence

$$F = [X_1^A, X_2^A, \dots, X_K^A, Y_1^B, Y_2^B, \dots, Y_K^B] \quad (2)$$

where for instance X_1^A is the SVM probability output of the first attribute for image X . For training, features from image pairs of the same person are obtained and their SB distance is 0. Then the order of the images from one camera is randomly shuffled to create image pairs of distinct persons with different SB distance. The features and target values are used to learn the regression model.

For re-identification, given a pair of images, first SB attribute detection is performed and the probability outputs are obtained and concatenated as in (2). The SB distance is predicted using the learned regression model. In the implementation we use Support Vector Regression (SVR) to train the

¹Available at http://www.eecs.qmul.ac.uk/~rlayne/#bmvc_attr

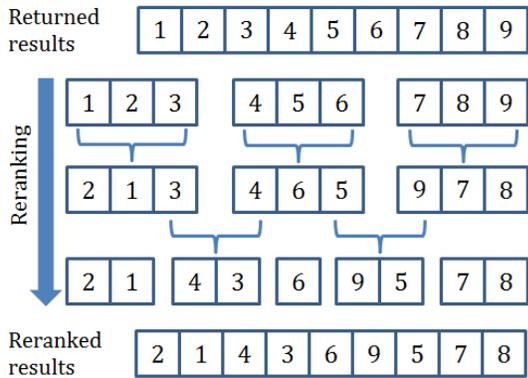


Fig. 4. Illustration of the reranking process. In this case, the returned results are first split into non-overlapping windows, each containing 3 subjects. Reranking is performed using SB distances, subjects with the lower SB distances are moved to front. Adjacent matches from different windows are then reranked to output smooth results.

regression model. Since the feature dimension ($5 + 5 = 10$ in our case) is significantly smaller than the number of samples, we use Radial Basis Function (RBF) kernel to map the lower dimensional features into higher dimensional space for better discrimination.

D. Reranking

The reranking is based on an initially returned result using a baseline re-identification method. Given a probe image, the baseline method returns the top N best matches. The SB distance between the probe image and the top N best matches from gallery are calculated using the learned regression model. Based on the SB distance, reranking is performed in local non-overlapping windows first and then the adjacent matches from neighboring windows are reranked to ensure smoothed reranking output. Figure 4 illustrates how the returned matches are reranked with a reranking window of size 3.

III. EXPERIMENTS

A. Database

The reranking is performed and evaluated using the VIPeR database², which is considered as one of the most challenging benchmark databases for person re-identification [2]. The database contains image pairs of 632 pedestrians. The images were captured by two cameras with significant view change. The view of person in CAM A spans from 0 degree to 90 degree and the view of person in CAM B changes from 90 degree to 180 degree. For each person, a single image is available from each camera view. All of the images in the VIPeR database are normalized to the size of 128×48 . Apart from the view change, other aspects such as changing illumination conditions, cluttered background and occlusions make this database very complicated.

²Available at <http://vision.soe.ucsc.edu/?q=node/178>

B. Parameter Settings

In our experiments we follow the experimental protocols in the previous work [3] [10]. The image pairs are randomly divided into two sets of 316 pairs each. One set is used for training and the other is used for testing. In the testing, the images from one camera are used as gallery data and images from the other camera are the probes. The experiments are performed 10 times and the average results are reported. To train the SVM classifiers for attribute detection, the slack parameter C is set to 1 as the cross-validation results show that the classification accuracy is not sensitive to the value of C . To train the SVR for SB distance prediction, γ in the RBF kernel function is set to 0.1. For reranking, the local window size is set to 3.

C. Baseline Methods

We use three methods as baseline re-identification methods including Nearest Neighbor with Euclidean distance (L2), a recently proposed metric learning-based method (KISSME [10]) and a popular multi-view analysis method - Canonical Correlation Analysis (CCA) [31], which has been used for feature transformation in person re-identification recently [9] [8].

D. Results of SB Attributes Detection

The SB attribute detection accuracy is shown in Figure 5. The classification rates for *male* and *short hair* are lower compared to the other three attributes *carrying*, *jeans*, and *backpack*. *Short hair* is difficult to detect due to the small discriminative regions involved and cluttered background that confuses the classifier. From distance with low-resolution, *male* is also a challenging attribute to be distinguished. For appearance-related attributes such as *carrying*, *jeans*, and *backpack*, the image regions accounting for these attributes are larger and visually more noticeable, thus making them less difficult to be detected. *Carrying* is the easiest among five SB attributes to be detected since its salience is very outstanding in the image.

E. Results of Reranking

Table II shows the comparison of re-identification rates with and without reranking at different ranks using different

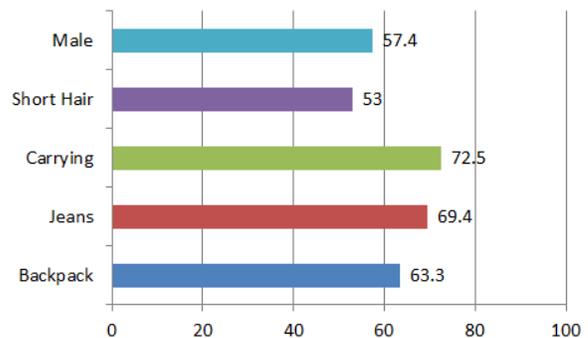


Fig. 5. Attribute detection accuracy for the 5 SB attributes used in this paper (in %).

Rank→	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$	$r = 6$	$r = 7$	$r = 8$	$r = 9$	$r = 10$	$r=25$	$r=50$	$r=75$	$r=100$
L2	7.5	10.4	12.6	14.2	16.1	17.4	19.3	20.8	23.1	24.1	30.7	45.6	56.1	66.7
L2+SB	7.9	11.1	13.6	14.6	16.1	18.4	19	21.9	23.1	24.7	31.4	45.6	56.4	67.1
KISSME [10]	18.6	30.1	38.2	45.5	50.3	53.8	56.6	59.4	60.7	62.3	77.5	90.1	93.3	96.8
KISSME [10]+SB	19.3	31.7	39.2	45.3	50.7	54.2	57.3	60.4	61.1	63.3	78.2	90.6	93.7	97.2
CCA [31]	14.6	22.7	28.8	31.9	37	42.4	44.9	48.1	50.3	52.2	74.1	87.3	91.7	94
CCA [31]+SB	15.5	22.5	29.1	34.2	37.7	42.1	45.9	48.8	51.3	52.2	74.7	87.6	91.7	94.3

TABLE II

THE COMPARISON OF THE TOP RANK RE-IDENTIFICATION RATES ON THE VIPER DATABASE USING DIFFERENT BASELINE METHODS WITH/WITHOUT COMPUTED SB ATTRIBUTES FOR RERANKING (IN %).

baseline methods. Rates at top 10 ranks are shown as well as the rates at higher ranks (25, 50, 75 and 100). At most ranks, using reranking improves the re-identification accuracy for different baseline methods. Note that in this case the SB attribute detection is not specifically optimized and for each attribute the feature representations used for classification remain the same. It is expected that by using more advanced features with customization, the improved SB attributes detection will help more in the reranking process. Figure 6 shows exemplar cases in which reranking leads to improvement.

F. Discussion

The effectiveness of reranking depends on the accuracy of SB attributes detection. As shown in Figure 5, the performance of attribute classifiers is limited by the lowest classification accuracy as 53% for *short hair* (slightly better than random guess) and the highest classification accuracy for *carrying* less than 75%. In this case, the inaccuracy of SB attribute detection prevents larger performance gain in the reranking process. If the annotated SB attributes are provided and used for reranking, the performance improvement after reranking would be more significant. Table III shows the reranking results for top 5 ranks on the same database using the same baseline methods. In this case the SB distance is computed as the Hamming distance between the SB attributes of probe and gallery images using the annotated attribute labels. As can be seen from Table III, the performance gain after reranking is more significant. Especially for more advanced methods (KISSME [10] and CCA [31]), the re-identification rate at rank 1 nearly doubles. The improvements from both Table II and Table III suggest that either the predicted SB attributes or human annotated SB attributes can be used to improve the re-identification system performance. To further improve the reranking results, it is desirable to select highly discriminative SB attributes and to improve the SB attributes detection using more robust features or classifiers.

IV. CONCLUSIONS

Person re-identification is an inherently difficult recognition problem due to the significant appearance change of individuals captured by different cameras. In this paper, a reranking method is proposed to improve the initially returned re-identification results using existing re-identification approaches. For reranking, soft biometric (SB) attributes are used. As a higher level semantic descriptor, SB attributes can

Rank→	$r = 1$	$r = 2$	$r = 3$	$r = 4$	$r = 5$
L2	7.6	10	12.9	13.9	16.3
L2+SB	12	12.9	15.7	18.1	20.9
KISSME [10]	17.9	28.2	34.8	40.1	45.6
KISSME [10]+SB	32	34.5	42.9	49.2	49.4
CCA [31]	15	23.6	29.1	35.2	39.4
CCA [31]+SB	27.7	29.1	36.6	41.3	41.8

TABLE III

THE COMPARISON OF THE TOP 5 RE-IDENTIFICATION RATES ON VIPER DATABASE USING DIFFERENT BASELINE METHODS WITH/WITHOUT HUMAN ANNOTATED SB ATTRIBUTES FOR RERANKING(IN %).

be used to lower the ranks of the returned matches that are visually similar to the probe subject but possessing different SB attributes such as gender. The SB attributes are detected using trained SVM classifier and then SB distance is computed between a probe and a gallery image with a trained regression model. The reranking is performed using the computed SB distances. The proposed reranking approach is independent of the re-identification method, thus, it is compatible and can be integrated into any existing re-identification system. Experiments on benchmark database show that the reranking process helps to improve the re-identification rates at different ranks. Future work involves improving the SB attribute detection accuracy and selecting most discriminative SB attributes for reranking.

ACKNOWLEDGMENTS

This work was supported in part by ONR grants N00014-12-1-1026, N00014-09-C-0388 and NSF grant 0905671. The contents and information do not reflect the position or policy of the U.S. Government.

REFERENCES

- [1] W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, "Principal axis-based correspondence between multiple cameras for people tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 663–671, April 2006.
- [2] D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in *10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, Sept. 2007.
- [3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2010, pp. 2360–2367.
- [4] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, "Shape and appearance context modeling," in *IEEE International Conference on Computer Vision (ICCV)*, Oct. 2007, pp. 1–8.



Fig. 6. Examples of the improved results after reranking. Probes are shown in green bounding boxes and the correct matches from gallery are shown in red bounding boxes.

- [5] O. Javed, K. Shafique, Z. Rasheed, and M. Shah, "Modeling inter-camera space-time and appearance relationships for tracking across non-overlapping views," *Computer Vision and Image Understanding*, vol. 109, no. 2, pp. 146 – 162, 2008.
- [6] M. Hirzer, P. M. Roth, M. Köstinger, and H. Bischof, "Relaxed pairwise learned metric for person re-identification," in *European conference on Computer Vision (ECCV)*, 2012, pp. 780–793.
- [7] W.-S. Zheng, S. Gong, and T. Xiang, "Reidentification by relative distance comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 653–668, 2013.
- [8] L. An, M. Kafai, S. Yang, and B. Bhanu, "Reference-based person re-identification," in *IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, 2013.
- [9] W. Li and X. Wang, "Locally aligned feature transforms across views," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
- [10] M. Köstinger, M. Hirzer, P. Wohlhart, P. Roth, and H. Bischof, "Large scale metric learning from equivalence constraints," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2012, pp. 2288–2295.
- [11] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *Journal of Machine Learning Research*, vol. 10, pp. 207–244, Jun. 2009.
- [12] M. Guillaumin, J. Verbeek, and C. Schmid, "Is that you? metric learning approaches for face identification," in *IEEE International Conference on Computer Vision (ICCV)*, 2009, pp. 498–505.
- [13] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, "Information-theoretic metric learning," in *International Conference on Machine Learning*, 2007, pp. 209–216.
- [14] P. Massa and C. Hayes, "Page-rerank: using trusted links to re-rank authority," in *IEEE/WIC/ACM International Conference on Web Intelligence*, 2005, pp. 614–617.
- [15] Z. Wu, Q. Ke, J. Sun, and H.-Y. Shum, "Scalable face image retrieval with identity-based quantization and multireference reranking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 10, pp. 1991–2001, 2011.
- [16] C. Xu, Y. Li, C. Zhou, and C. Xu, "Learning to rerank images with enhanced spatial verification," in *IEEE International Conference on Image Processing (ICIP)*, 2012, pp. 1933–1936.
- [17] Y.-H. Yang and W. Hsu, "Video search reranking via online ordinal reranking," in *IEEE International Conference on Multimedia and Expo*, 2008, pp. 285–288.
- [18] T. Meng and M.-L. Shyu, "Leveraging concept association network for multimedia rare concept mining and retrieval," in *IEEE International Conference on Multimedia and Expo*, Melbourne, Australia, July 2012, pp. 860–865.
- [19] L. Zhang, D. Kalashnikov, S. Mehrotra, and R. Vaisenberg, "Context-based person identification framework for smart video surveillance," *Machine Vision and Applications*, pp. 1–15, 2013.
- [20] L. Zhang, D. V. Kalashnikov, and S. Mehrotra, "A unified framework for context assisted face clustering," in *ACM International conference on multimedia retrieval*, 2013, pp. 9–16.
- [21] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in *British Machine Vision Conference (BMVC)*, 2011, pp. 68.1–68.11.
- [22] S. Bak, G. Charpiat, E. Corvee, F. Bremond, and M. Thonnat, "Learning to match appearances by correlations in a covariance metric space," in *European Conference on Computer Vision (ECCV)*, 2012, pp. 806–820.
- [23] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *European Conference on Computer Vision (ECCV)*, 2008, pp. 262–275.
- [24] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang, "Person re-identification by support vector ranking," in *British Machine Vision Conference (BMVC)*, 2010, pp. 21.1–21.11.
- [25] M. Hirzer, C. Beleznaï, P. M. Roth, and H. Bischof, "Person re-identification by descriptive and discriminative classification," in *Scandinavian Conference on Image Analysis*, 2011, pp. 91–102.
- [26] R. Layne, T. Hospedales, and S. Gong, "Person re-identification by attributes," in *British Machine Vision Conference (BMVC)*, 2012, pp. 24.1–24.11.
- [27] R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
- [28] M. Dikmen, E. Akbas, T. S. Huang, and N. Ahuja, "Pedestrian recognition with a learned metric," in *Asian Conference on Computer Vision (ACCV)*, 2011, pp. 501–512.
- [29] W. Li, R. Zhao, and X. Wang, "Human reidentification with transferred metric learning," in *Asian Conference on Computer Vision (ACCV)*, 2012, pp. 31–44.
- [30] C. C. Loy, C. Liu, and S. Gong, "Person re-identification by manifold ranking," in *IEEE International Conference on Image Processing (ICIP)*, Barcelona, 2013.
- [31] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. pp. 321–377, 1936.