Exploiting Gaussian Mixture Importance for Person Re-identification

Xiangping Zhu Amran Bhuiyan Mohamed Lamine Mekhalfi Vittorio Murino
Pattern Analysis and Computer Vision, Istituto Italiano di Tecnologia, Genova, Italy
xiangping.zhu2010@gmail.com, {amran.bhuiyan,mohamed.mekhalfi,vittorio.murino}@iit.it

Abstract

Person re-identification (ReID) stands for the task of determining the co-occurrence of individuals across a network of cameras with disjoint viewfields. The relevant literature documents a plausible number of contributions so far. KISS metric learning is an effective ReID method. However, as reported in the existing works, KISS metric learning is sensitive to the feature dimensionality and can not capture the multi modes in the dataset. To this end, we propose in this paper a Gaussian Mixture Importance Estimation (GMIE) approach for ReID, which exploits the Gaussian Mixture Models (GMMs) to estimate the observed commonalities of similar and dissimilar person pairs in the feature space. Experiments on three benchmark datasets reveal that our method offers the property of maintaining its efficiency on high-dimensional features. Moreover, the proposed GMIE scores plausible ReID rates as compared to other works.

1. Introduction

Person re-identification aims at determining the correspondence of individuals over a set of cameras with non-overlapping viewfields. It has drawn a great deal of attention in the last few years owing to its significant role in carrying out numerous surveillance tasks such as person tracking and access control. Despite the devoted efforts, ReID remains a rather challenging task due to the nonrigid structure of the human body, the different perspectives in which a pedestrian can be observed, and the highly variable illumination conditions, an exemplar illustration of these factors is depicted in Fig. 1.

Most of the ReID works suggest that, normally, individuals do not change their clothings across a camera network. This assumption inspired the contributions to regard the visual aspect as a main cue in chacarterizing person images. In particular, the mainstream ReID literature can be broadly categorized in two classes, namely *direct* and *learning-based methods*. The former group tends to handcraft, 's-traightforwardly', robust features and potentially their combination thereof. In this regard, a common way is to divide

Figure 1. Challenges characterizing ReID from two datasets in two camera views (first and second rows, respectively).

the person image in hand into horizontal stripes [3, 18, 38], triangular graph [10], regions clustered by color [34], symmetry and assymetrical parts [9], semantic or meaningful parts [6, 2], concentric rings [10] or grid of localized patches [1]. There are numerous features extracted from these regions that have proven to be useful in ReID such as color histograms and their statistics [18, 38, 3, 9, 6, 2], maximally stable color region [9, 6, 2], textures [30, 28, 12], edges [30], Haar-like features [7], interest points [14] and Biologically Inspired Features (BIF) [21].

In the latter group, i.e., *learning based methods*, a dataset of similar and dissimilar persons is used to 'learn' personalized features and/or a metric space where to match them. The underlying assumption is that the knowledge extracted from the training set generalizes to unseen samples. Nearest neighbor classification [12], Partial Least Squares (PLSs) [30] and Rank-SVM [27] have been customized to the ReID issue. Pairwise Constraint Component Analysis (PCCA) [25] and Relaxed pairwise Metric Learning (RPML) [15] learn a projection metric from a high-dimensional input space to low-dimensional space where the distance between the pairs of data points respects the desired constraints. In [29, 17], learning a Mahalanobis metric for the ReID problem has yielded plausible results. In [29], Large Margin Nearest Neighbor Learning (LMNN) [36, 35], Information Theoretic Metric Learning (ITML) [8] and Logistic Discriminant Metric Learning [13] are evaluated for ReID. Such methods often learn a low dimensional subspace that captures as much as discriminative information

978-1-5386-2939-0/17/$31.00 ©2017 IEEE

from the data. Then, the Euclidean distance is applied in this subspace to evaluate the similarities of different samples. In [17], Köstinger *et al.* proposed another kind of metric learning, i.e., KISSME (Keep It Simple and Straightforward Metric Learning). Different from the concept highlighted above, KISSME directly measure the similarity between samples based on the likelihood ratio test. Thus, there is no need to learn a low dimensional discriminative subspace. The effectiveness of KISSME has been reported in many existing works [17, 18, 26, 29].

In this paper, we propose a new ReID method, named Gaussian mixture importance estimation (GMIE, for short), which is also based on likelihood ratio test, inspired by KISSME. However, GMIE offers several advantages with respect to KISSME. As described in [17], KISSME uses the Gaussian densities to separately approximate the distributions of intrapersonal and extrapersonal variations. Yet, if the intrapersonal and/or extrapersonal variations are characterized by a multi-modal distribution, which is commonly encountered in practice, the distribution approximation in KISSME would be inaccurate based only on Gaussian densities. Fig. 2 demonstrates an example of approximating a bimodal distribution with Gaussian density in a one projected dimension case, where the estimated (red) curve does not accurately fit the actual (blue) one. In addition, what is important in KISSME is to estimate the covariance matrices of intrapersonal and extrapersonal variations. In practice, since for the small sample size problem in ReID dataset, it is always difficult to accurately estimate these two covariance matrices, especially in high dimensional cases. Although the principal component analysis (PCA) is used for dimension reduction, KISSME remains sensitive to the feature dimension as reported in the existing works [18, 24]. Considering these disadvantages in KISSME, we introduce a novel ReID method based on Gaussian mixture importance estimation, which is invariant to feature dimension. Unlike KISSME, our method directly estimates the ratio of the aforementioned probability densities via GMMs, which maintains the performance as the feature dimension rises; Experiments on three datasets are used to validate the advantages of our approach over existing alternatives on multiple benchmark datasets.

The rest of the paper is organized as follows. Section 2 firstly makes a brief introduction about KISSME, and then describes the proposed approach in detail. Section 3 reports the experimental results, and conclusions are drawn in Section 4.

2. Our Approach

As in KISSME, GMIE is also based on the likelihood ratio between intrapersonal and extrapersonal probability densities. Thus, in the following, a brief description about KISSME will be provided before introducing GMIE.

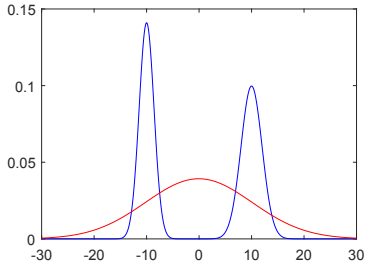

Figure 2. Illustration example of approximating a distribution with two modes using Gaussian density. Blue line is the true probability density and the red line is the approximated probability density.

2.1. KISS Metric Learning

Given a pair of labeled samples $\{\mathbf{z}_i, y_i\}$ and $\{\mathbf{z}_j, y_j\}$, in which y_i and y_j respectively denote the label of person identities associated to the images $\mathbf{z}_i$ and $\mathbf{z}_j$. The difference between these two latter is calculated as $\mathbf{x} = \mathbf{z}_i - \mathbf{z}_j$. If $y_i = y_j$, $\mathbf{x}$ is called the intrapersonal difference, and if $y_i \neq y_j$, $\mathbf{x}$ is called the extrapersonal difference. Let Ω_I denotes the class that includes the intrapersonal differences, while Ω_E denotes the class that holds the extrapersonal differences. In KISSME, the likelihood ratio test is used to determine if $\mathbf{x}$ belongs to Ω_E. The likelihood ratio is:

$$\delta(\mathbf{x}) = \log\left(\frac{p_E(\mathbf{x})}{p_I(\mathbf{x})}\right), \tag{1}$$

where $p_E(\mathbf{x})$ is the probability density for class Ω_E and $p_I(\mathbf{x})$ is for Ω_I. Since the set of sample difference $\mathbf{x}$ is zero mean, the approximations of $p_E(\mathbf{x})$ and $p_I(\mathbf{x})$ with Gaussian densities can be formulated as:

$$p_E(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_E|^{1/2}} \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{\Sigma}_E^{-1}\mathbf{x}), \tag{2}$$

$$p_I(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\mathbf{\Sigma}_I|^{1/2}} \exp(-\frac{1}{2}\mathbf{x}^T\mathbf{\Sigma}_I^{-1}\mathbf{x}), \tag{3}$$

in which $\mathbf{\Sigma}_E$ and $\mathbf{\Sigma}_I$ represent the covariance matrices of Ω_E and Ω_I, respectively, and d is the dimension of $\mathbf{x}$.

Substitute Eqs. (2)-(3) into Eq. (1), the likelihood ratio can be reformulated as:

$$\delta(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T(\mathbf{\Sigma}_I^{-1} - \mathbf{\Sigma}_E^{-1})\mathbf{x} + \frac{1}{2}\log\left(\frac{|\mathbf{\Sigma}_I|}{|\mathbf{\Sigma}_E|}\right). \tag{4}$$

Removing the constant terms, we can get the simplified likelihood ratio:

$$\delta(\mathbf{x}) = \mathbf{x}^T(\mathbf{\Sigma}_I^{-1} - \mathbf{\Sigma}_E^{-1})\mathbf{x}. \tag{5}$$

To ensure the nonnegative value of $\delta(\mathbf{x})$, KISSME further re-projects $\mathbf{M} = \mathbf{\Sigma}_I^{-1} - \mathbf{\Sigma}_E^{-1}$ into the cone of positive semi-definite matrix.

2.2. Gaussian Mixture Importance Estimation

GMIE aims at directly approximating the ratio of intrapersonal and extrapersonal probability densities $p_E(\mathbf{x})$ and $p_I(\mathbf{x})$ using the GMMs. In this way, it avoids explicitly estimating $p_I(\mathbf{x})$ and $p_E(\mathbf{x})$ and effectively improves its robustness to high dimensional features as demonstrated in the experiments. In addition, using GMMs as approximation function, GMIE can also capture the multi-modes existed in $p_I(\mathbf{x})$ and/or $p_E(\mathbf{x})$.

Different from KISSME, we here consider the likelihood of $\mathbf{x}$ belonging to Ω_I. Thus, the ratio used in GMIE between $p_E(\mathbf{x})$ and $p_I(\mathbf{x})$ is:

$$w(\mathbf{x}) = \frac{p_I(\mathbf{x})}{p_E(\mathbf{x})}. \tag{6}$$

As aforementioned, GMIE approximates the ratio $w(\mathbf{x})$ via the GMMs, which combines a number of Gaussian components. Hence, the approximation of $w(\mathbf{x})$ is:

$$\hat{w}(\mathbf{x}) = \sum_{l=1}^{b} \pi_l N(\mathbf{x}|\mathbf{u}_l, \mathbf{\Sigma}_l), \tag{7}$$

where π_l are the weights, and b is the number of Gaussian mixture components. $\mathbf{u}_l$ and $\mathbf{\Sigma}_l$ respectively denote the mean vector and covariance matrix of the l^{th} Gaussian component:

$$N(\mathbf{x}|\mathbf{u}_l, \mathbf{\Sigma}_l) = \frac{\exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{u}_l)^T \mathbf{\Sigma}_l^{-1}(\mathbf{x}-\mathbf{u}_l)\right)}{(2\pi)^{d/2}|\mathbf{\Sigma}_l|^{1/2}}. \tag{8}$$

Thus, the likelihood ratio as in Eq. (5) can be reformulated as:

$$\delta(\mathbf{x}) = \log w(\mathbf{x}) \approx \log \sum_{l=1}^{b} \pi_l N(\mathbf{x}|\mathbf{u}_l, \mathbf{\Sigma}_l). \tag{9}$$

To estimate the parameters π_l, $\mathbf{u}_l$, $\mathbf{\Sigma}_l$ and b, the minimization of the Kullback-Leibler divergence from $p_I(\mathbf{x})$ to its approximation $\hat{p}_I(\mathbf{x})$ is used:

$$\text{KL}[p_I(\mathbf{x})||\hat{p}_I(\mathbf{x})] = \int p_I(\mathbf{x}) \log \frac{p_I(\mathbf{x})}{\hat{p}_I(\mathbf{x})} d\mathbf{x}. \tag{10}$$

Take Eqs. (6) and (7) into consideration, the estimation of $\hat{p}_I(\mathbf{x})$ can be formulated as:

$$\hat{p}_I(\mathbf{x}) = \hat{w}(\mathbf{x}) p_E(\mathbf{x}). \tag{11}$$

Substitute Eq. (11) into Eq. (10), the Kullback-Leibler divergence can be reformulated as:

$$\begin{aligned} \text{KL}[p_I(\mathbf{x})||\hat{p}_I(\mathbf{x})] = &\int p_I(\mathbf{x}) \log \frac{p_I(\mathbf{x})}{p_E(\mathbf{x})} d\mathbf{x} \\ &- \int p_I(\mathbf{x}) \log \hat{w}(\mathbf{x}) d\mathbf{x}. \end{aligned} \tag{12}$$

The unknown parameters are only contained in the second term of the equation. Thus, minimizing Kullback-Leibler divergence equals to maximizing the second term as:

$$J = \int p_I(\mathbf{x}) \log \hat{w}(\mathbf{x})\, d\mathbf{x} = \frac{1}{n_I} \sum_{i=1}^{n_I} \log \hat{w}(\mathbf{x}_i^I), \tag{13}$$

where n_I is the number of samples in Ω_I, and $\mathbf{x}_i^I$ denotes the i^{th} sample from Ω_I.

Since $p_I(\mathbf{x})$ is a probability density, thus the following constraint should be held:

$$\begin{aligned} 1 = &\int \hat{p}_I(\mathbf{x})\, d\mathbf{x} = \int \hat{w}(\mathbf{x}) p_E(\mathbf{x})\, d\mathbf{x} \\ \approx &\frac{1}{n_E} \sum_{j=1}^{n_E} \hat{w}(\mathbf{x}_j^E), \end{aligned} \tag{14}$$

where n_E is the number of samples in Ω_E, and $\mathbf{x}_j^E$ denotes the the j^{th} sample from Ω_E.

Consider Eqs. (7), (13) and (14), the optimization problem of GMIE can be formulated as:

$$\begin{aligned} \max_{\{\pi_l, \mathbf{u}_l, \mathbf{\Sigma}_l\}_{l=1}^b} \quad & \sum_{i=1}^{n_I} \log \left(\sum_{l=1}^{b} \pi_l N(\mathbf{x}_j^I|\mathbf{u}_l, \mathbf{\Sigma}_l) \right) \\ \text{s.t.} \quad & \sum_{j=1}^{n_E} \sum_{l=1}^{b} \pi_l N(\mathbf{x}_j^E|\mathbf{u}_l, \mathbf{\Sigma}_l) = n_E \\ & \pi_1, ..., \pi_b \geq 0. \end{aligned} \tag{15}$$

The parameters $\boldsymbol{\mu}_l$, $\mathbf{\Sigma}_l$ and π_l can be estimated by employing the Lagrangian multiplier method on this optimization problem. More detailed estimation procedure can be found in [37].

The parameter b can be determined through the likelihood cross validation method as introduced in [32, 37]. Considering the small sample size problem in ReID, we determine b by providing several candidates and select the one that maximize J in Eq. (13).

3. Experiments

We assess our introduced method on three datasets, i.e., VIPeR dataset [11], GRID dataset [19] and PRID 450S dataset [29]. In the experiments, b is determined using five candidates, i.e., $b = [1, 2, 3, 4, 5]$, and the one that maximizes J is selected as mentioned earlier. The evaluation procedure in every dataset is repeatedly carried out 10 times. It is to note that PCA is employed for feature dimension reduction. We report the results in terms of ReID rate by means of Cumulative Matching Characteristic (CMC) curve, which accumulates the ReID accuracy as the rank index increases. For the analysis, the performance of our method is evaluated from three different aspects, i.e. robustness to subspace dimensions, effect of different descriptors

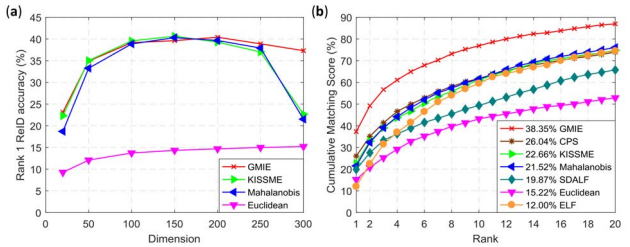

Figure 3. Experimental results on VIPeR dataset with GOG feature: (a) dimensionality influence; (b) CMC curves with 300 dimensions.

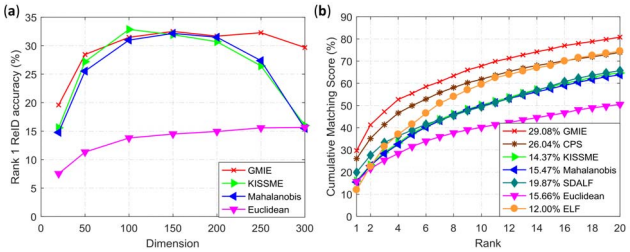

Figure 4. Experimental results on VIPeR dataset with LOMO feature: (a) dimensionality influence; (b) CMC curves with 300 dimensions.

and ReID rate. The details and discussions on each dataset are presented in following three subsections.

3.1. Experiments on VIPeR dataset

VIPeR dataset [11] is widely used for evaluating the performance of ReID methods. It contains 632 person image pairs captured from two different cameras. Large variations of viewpoint and illumination in images make VIPeR be a challenging dataset. In compliance with common evaluation scenario, 316 person image pairs are randomly selected for training, and the rest of 316 image pairs are retained for testing.

Fig. 3 shows evaluation results on VIPeR dataset by adopting the Gaussian Of Gaussian (GOG) feature [24]. Its sub-figure (a) presents the trend of rank 1 scores with increasing the subspace dimension. For comparison, the KISSME, Mahalanobis distance trained with genuine pairs [17] and Euclidean distance are implemented. It can be found in the figure that KISSME and Mahalanobis distance work well only within the range from 100^{th} to 250^{th} dimensions. For the dimension larger than 250, their rank 1 scores decrease dramatically since the estimated covariance matrices become inaccurate in high dimensional subspace. However, the proposed GMIE method estimates the parameters based on the Kullback-Leibler divergence instead of a direct inference from samples. This renders it robust to the dimension increase which is also demonstrated in the Fig. 3 (a). Although the Euclidean distance is also robust to dimensionality, its ReID accuracy is relatively low. We further compare our method for a dimension of 300 with other methods, i.e., CPS [5], SDALF [9], and ELF [12]. The CMC curves for these considered methods are displayed in Fig. 3 (b). It can be seen that GMIE works better than the considered methods, often by a large margin.

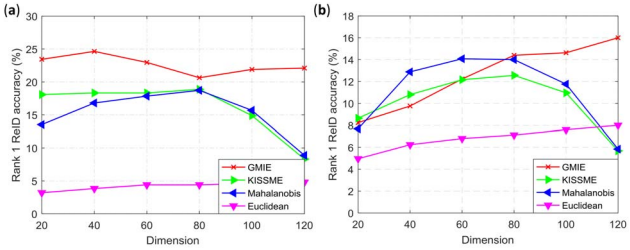

Figure 5. Experimental results of dimensionality influence on grid dataset with (a) GOG feature and (b) LOMO feature.

To assess the performance of our method in other feature descriptors, we also repeat the experiments on VIPeR dataset using the LOMO feature [18]. The experimental results are depicted in Fig. 4. From the figure, the same behavior is observed. It is to note from these outcomes that GOG seems to incur more robustness on the KISSME and Mahalanobis distance.

3.2. Experiments on GRID Dataset

The GRID dataset comprises 250 person image pairs. Each pair consists of two images of the same person but are seen from different camera views. In addition, there are another 775 person images that do not belong to any of the 250 paired persons. In the experiment, 125 person image pairs are randomly selected for training. The remaining 125 person image pairs and 775 unpaired person images are used for testing. Finally, in the testing procedure, there are 125 probe images and 900 gallery images. Different images in GRID have different sizes. In the experiment, we resize all the images to 128×48. Compared with VIPeR dataset, GRID dataset is more challenging given its smaller sample size and extra gallery person images.

In the experiment, the KISSME, Mahalanobis and Euclidean distance are also included for comparison as in VIPeR dataset. To evaluate the effectiveness of GMIE on different feature descriptors, the experiments are performed with both GOG and LOMO features. The results are presented in Fig. 5. It can be viewed from the figure that KISSME and Mahalanobis distance are not robust to the dimension increase. In the high dimensional subspace (the dimension greater than 80), their ReID accuracies decrease largely, especially at 120^{th} dimension of LOMO feature, the ReID accuracies of KISSME and Mahalanobis distance are even lower than the Eluclidean's. However, GMIE always works well in the high subspace dimension both in GOG and LOMO cases as demonstrated in Fig. 5.

To make a comparison with the state of the art results reported on GRID dataset, we calculate the ReID accuracy in the subspace dimension with which the related method has

Table 1. Experimental results on GRID dataset of different methods comparing with our GMIE approach (%).

Methods	r=1	r=5	r=10	r=15	r=20
MRank-RankSVM [20]	12.24	27.84	36.32	42.24	46.56
MtMCML [22]	14.08	34.64	45.84	52.88	59.84
PolyMap [4]	16.30	35.80	46.00	52.80	57.60
SSDAL+XQDA [31]	22.40	39.20	48.00	–	58.40
KEPLER [23]	18.40	39.12	50.24	57.04	61.44
NLML [16]	24.54	35.86	43.53	–	55.25
DR-KISS [33]	20.60	39.30	51.40	–	62.60
SCSP [4]	24.24	44.56	54.08	–	59.68
GOG+KISS	18.32	39.36	51.84	58.48	63.44
GOG+XQDA [24]	**24.70**	47.00	58.40	–	69.00
GOG+GMIE (ours)	24.64	**49.52**	**63.84**	**69.48**	**73.32**

the highest ReID accuracy, and the final results are listed in Tab. 1. The table shows that for the rank 1, ReID accuracy of GMIE (24.64%) almost equals to the highest state of art result (24.70%) reported by XQDA with using GOG feature [24]. However, for the rank 5, 10, 15 and 20, GMIE has the best accuracies compared with other considered methods. The main reason of the high ReID accuracy of GMIE traces back to the fact that GMIE is based on GMMs, which makes it more accurate in capturing the multi-modal properties in the densities $p_I(\mathbf{x})$ and $p_E(\mathbf{x})$ in Equ. (6).

3.3. Experiments on PRID 450S Dataset

PRID 450S dataset contains 450 single-shot image pairs that depict the walking persons captured in two spatially disjoint camera views. All of the persons are manually annotated by bounding boxes form the original surveillance images. The dataset also offers the person part-level segmentation results. More description about the dataset can be found in [29].

The experimental results are illustrated in Fig. 6. From its Fig. 3(a), it can be found that KISS and Mahalanobis distance perform well only in the certain subspace dimension range (between 50^{th} and 150^{th} dimension) as in VIPeR and GRID datasets. For the low or high dimensions, their ReID accuracies decrease. However, GMIE still works well in the high dimensional case. It can be seen from the figure that the GMIE's trend of rank 1 ReID accuracy consistently keeps stable from 50^{th} dimension. By setting the subspace dimension to 220, the CMC curves for these four methods are plotted in Fig. 6 (b). As shown, Euclidean distance works better than both KISS and Mahalanobis distance in high dimensional subspace. GMIE always has the highest ReID accuracy compared with other considered methods.

From the experimental results on these three datasets, it can be found that KISS and Mahalanobis distance are sensitive to the subspace dimension. In particular their performances degrade significantly in the high dimensional case. Thus, in order to use the KISS and Mahalanobis distance for ReID task, a cross validation procedure may be necessary

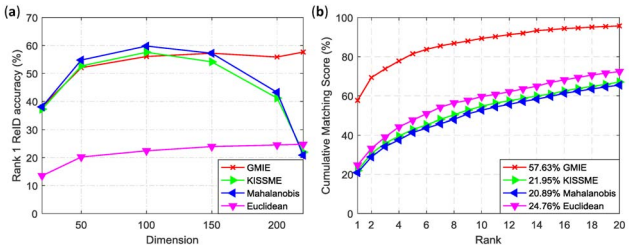

Figure 6. Experimental results on PRID 450S dataset with GOG feature: (a) dimensionality influence; (b) CMC curves with 220 dimensions.

in the training stage in order to find an appropriate subspace dimension which will come at the cost of more complexity. However, Our approach GMIE always performs well in the high dimensional subspace on all three datasets. In addition, GMIE also has high ReID accuracy, especially on the GRID datasets.

4. Conclusions

In this paper, we proposed a GMIE person ReID method. Unlike KISSME, it directly approximates the density ratio between the intrapersonal and extrapersonal variations. By adapting the Kullback-Leibler divergence technique, GMIE can maintain its ReID performance even in the high dimensional case, which is difficult for KISSME. In addition, thanks to the GMMs used for approximating the density ratio, GMIE is also capable of capturing the multi modal properties existed in the underlying densities of intrapersonal and extrapersonal variations. These advantages of GMIE have been validated with detailed experiments on three datasets. The ReID accuracy of GMIE is also satisfactory compared with other works.

References

[1] S. Bak, E. Corvee, F. Bremond, and M. Thonnat. Multiple-shot human re-identification by mean riemannian covariance grid. In *AVSS*, 2011.

[2] A. Bhuiyan, A. Perina, and V. Murino. Person re-identification by discriminatively selecting parts and features. In *ECCV-WK*, 2014.

[3] N. D. Bird, O. Masoud, N. P. Papanikolopoulos, and A. Isaacs. Detection of loitering individuals in public transportation areas. *TITS*, 2005.

[4] D. Chen, Z. Yuan, G. Hua, N. Zheng, and J. Wang. Similarity learning on an explicit polynomial kernel feature map for person re-identification. In *CVPR*, pages 1565–1573, 2015.

[5] D. S. Cheng and M. Cristani. Person re-identification by articulated appearance matching. In *Person Re-Identification*. Springer, 2014.

[6] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino. Custom pictorial structures for re-identification. In *BMVC*, 2011.

[7] E. Corvee, F. Bremond, M. Thonnat, et al. Person re-identification using spatial covariance regions of human body parts. In *AVSS*, 2010.

[8] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon. Information-theoretic metric learning. In *ICML*, 2007.

[9] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. In *CVPR*, 2010.

[10] N. Gheissari, T. B. Sebastian, and R. Hartley. Person reidentification using spatiotemporal appearance. In *CVPR*, 2006.

[11] D. Gray, S. Brennan, and H. Tao. Evaluating appearance models for recognition, reacquisition, and tracking. In *PETS*, 2007.

[12] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*, pages 262–275. Springer, 2008.

[13] M. Guillaumin, J. Verbeek, and C. Schmid. Is that you? metric learning approaches for face identification. In *ICCV*, 2009.

[14] O. Hamdoun, F. Moutarde, B. Stanciulescu, and B. Steux. Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences. In *ACM*, 2008.

[15] M. Hirzer, P. M. Roth, M. Köstinger, and H. Bischof. Relaxed pairwise learned metric for person re-identification. In *ECCV*, 2012.

[16] S. Huang, J. Lu, J. Zhou, and A. K. Jain. Nonlinear local metric learning for person re-identification. *arXiv preprint arXiv:1511.05169*, 2015.

[17] M. Köstinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof. Large scale metric learning from equivalence constraints. In *CVPR*, 2012.

[18] S. Liao, Y. Hu, X. Zhu, and S. Z. Li. Person re-identification by local maximal occurrence representation and metric learning. In *CVPR*, 2015.

[19] C. Liu, S. Gong, and C. C. Loy. On-the-fly feature importance mining for person re-identification. *Pattern Recognition*, 2014.

[20] C. C. Loy, C. Liu, and S. Gong. Person re-identification by manifold ranking. In *ICIP*, pages 3567–3571. IEEE, 2013.

[21] B. Ma, Y. Su, and F. Jurie. Bicov: a novel image representation for person re-identification and face verification. In *BMVC*, 2012.

[22] L. Ma, X. Yang, and D. Tao. Person re-identification over camera networks using multi-task distance metric learning. *IEEE Transactions on Image Processing*, 23(8):3656–3670, 2014.

[23] N. Martinel, C. Micheloni, and G. L. Foresti. Kernelized saliency-based person re-identification through multiple metric learning. *IEEE Transactions on Image Processing*, 24(12):5645–5658, 2015.

[24] T. Matsukawa, T. Okabe, E. Suzuki, and Y. Sato. Hierarchical gaussian descriptor for person re-identification. In *CVPR*, 2016.

[25] A. Mignon and F. Jurie. Pcca: A new approach for distance learning from sparse pairwise constraints. In *CVPR*, 2012.

[26] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *CVPR*, 2013.

[27] B. Prosser, S. Gong, and T. Xiang. Multi-camera matching using bi-directional cumulative brightness transfer functions. In *BMVC*, 2008.

[28] B. Prosser, W.-S. Zheng, S. Gong, T. Xiang, and Q. Mary. Person re-identification by support vector ranking. In *BMVC*, 2010.

[29] P. M. Roth, M. Hirzer, M. Koestinger, C. Beleznai, and H. Bischof. Mahalanobis distance learning for person re-identification. In *Person Re-Identification*, pages 247–267. Springer, 2014.

[30] W. R. Schwartz and L. S. Davis. Learning discriminative appearance-based models using partial least squares. In *2009 XXII Brazilian Symposium on Computer Graphics and Image Processing*, 2009.

[31] C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian. Deep attributes driven multi-camera person re-identification. In *ECCV*, pages 475–491. Springer, 2016.

[32] M. Sugiyama, S. Nakajima, H. Kashima, P. V. Buenau, and M. Kawanabe. Direct importance estimation with model selection and its application to covariate shift adaptation. In *Advances in neural information processing systems*, 2008.

[33] D. Tao, Y. Guo, M. Song, Y. Li, Z. Yu, and Y. Y. Tang. Person re-identification by dual-regularized kiss metric learning. *IEEE Transactions on Image Processing*, 25(6):2726–2738, 2016.

[34] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu. Shape and appearance context modeling. In *ICCV*, 2007.

[35] K. Q. Weinberger and L. K. Saul. Fast solvers and efficient implementations for distance metric learning. In *ICML*, 2008.

[36] K. Q. Weinberger and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. *JMLR*, 2009.

[37] M. Yamada and M. Sugiyama. Direct importance estimation with gaussian mixture models. *IEICE transactions on information and systems*, 2009.

[38] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li. Salient color names for person re-identification. In *ECCV (ECCV)*, 2014.