ISSN: 2278-0181
NCETET - 2016 Conference Proceedings

Minimum Classification Error based Kiss Metric Learning for Personal Reidentification

Remya V
PG Scholar
Computer Science and Engineering
Younus College of Engineering and Technology,
Kollam, India-691010

Arun Raj S
Assistant Professor
Information Technology and Engineering
Younus College of Engineering and Technology,
Kollam, India-691010

***Abstract*— In recent years, person reidentification has received growing attention with the increasing popularity of intelligent video surveillance. This is because person reidentification is critical for human tracking with multiple cameras. Recently, keep it simple and straightforward (KISS) metric learning has been regarded as a top level algorithm for person reidentification. The covariance matrices of KISS are estimated by maximum likelihood (ML) estimation. It is known that discriminative learning based on the minimum classification error (MCE) is more reliable than classical ML estimation with the increasing of the number of training samples. When considering a small sample size problem, direct MCE KISS does not work well, because of the estimate error of small eigenvalues. Therefore, we further introduce the smoothing technique to improve the estimates of the small eigenvalues of a covariance matrix. Our new scheme is termed the minimum classification error-KISS (MCE-KISS). We conduct thorough validation experiments on the VIPeR dataset, which demonstrate the robustness and effectiveness of MCE-KISS for person reidentification.**

Index Terms*—Intelligence video surveillance, LBP Descriptor, HSV histogram, MCE Criterion, Distance metric learning, personal Reidentification

1. INTRODUCTION

Person reidentification is complex but receives intensive attention in the field of intelligent video surveillance (IVS). An aim of person reidentification is to match an instance of a person captured by one camera to another instance of the person captured by different cameras. The person reidentificaiton is a hard problem and still argely unsolved. The main reason may lie in that there usually exists a great variance of a persons appearance in different images obtained by different nonoverlapping cameras. Traditional biometrics, such as face[1][2],iris[3], fingerprint[4], and gait, are not often available, because images are low-quality, variable, and contain motion blur. There are two important stages which need to be considered in the process of person reidentification. They are distance learning, and visual feature extraction and selection.

KISS(Keep It Simple and Straight forward)[5] is tha Top level algorithm for Person Reidentification. KISS metric learning (KISS) based on an assumption that pairwise differences are Gaussian distributed. This has acquired state-of-the-art retrieval performance for real applications, such as person reidentification and face recognition.

The paper introduces the minimum classification error (MCE) criterion to improve KISS distance learning for person reidentification. In particular, eigenvalues of the true covariance matrix are biased, which harms the utilization of the estimated covariance matrix in subsequent operations, such as classification. The covariance matrices of KISS are obtained by maximum likelihood (ML) estimation. With increasing the number of training samples, discriminative learning based on MCE[6] is more reliable than classical ML estimation. However, only introducing the MCE criterion to the training procedure does not work well to estimate the small eigenvalues of the covariance matrices. Therefore, the smoothing technique[7] is required to improve the estimate of the small eigenvalues of a covariance matrix. The improved KISS is termed the minimum classification error-KISS, or simply MCE-KISS.

2. LITERATURE SURVEY

Existing work on personre-identification and appearance modeling can be roughly categorized into three categories: distance learning, local feature selection, and segmentation based matching.

In distance learning[11], a distance metric is learnt as a means of representing the similarity of individuals between camera pairs. In local feature selection, supervised or unsupervised algorithms are designed to select the most relevant features for person re-identification. In segmentation based matching,images of persons are first divided into small blobs and then the correspondences between these blobs are calculated.

In another point of view, existing techniques can be also categorized as single-shot and multiple-shot groups based on their experimental setups. In the single-shot group, an associating pair of images,each containing one instance of the individual,are used for training and testing,respectively.The approaches have to model a person by analyzing the single training image. In contrast to the single-shot group,the multiple-shot approaches train and test the person appearance model using multiple images which are usually obtained through tracking.The person re-identification presents a number of challenges beyond tracking and object detection.

2.1. Using quaternionic local binary pattern

Person reidentification is to identify the persons observed in nonoverlapping camera networks. Most existing methods usually extract features from the red, green, and blue color channels of images individually. They, however, neglect the connections between each color component in the image. To overcome this problem, a novel quaternionic local binary pattern (QLBP)[14] is proposed for person reidentification in this paper. The quaternion, introduced by Hamilton in 1843, is a four-dimensional generation of the complex number with one real part and three imaginary parts. It can be represented in a complex form as follows: q = a + ib + jc + kd, where a, b, c and d are real numbers, i, j and k are complex operators, In the proposed QLBP, each pixel in a color image is represented by a quaternion so that we can handle all color components in a holistic way and then extract LBP features[8] based on the quaternionic representation. Consequently, the obtained features integrate advantages of both the quaternionic representation and LBP such that they can be used to address the person reidentification problem well. A novel pseudo-rotation of quaternion (PRQ) is proposed to rank two quaternions. After a QLBP coding,the local histograms are extracted and used as features.

2.2. Using Spatiotemporal Appearance

This paper focuses on reidentification algorithms that use the overall appearance of an individual as opposed to passive biometrics such as face and gait. Person reidentification approaches have two aspects:(i) establish correspondence between parts.(ii) generate signatures that are invariant to variations in illumination, pose, and the dynamic appearance of clothing.A novel spatiotemporal [15] segmentation algorithm is employed to generate salient edgels that are robust to changes in appearance of clothing. The invariant signatures are generated by combining normalized color and salient edgel histograms.Two approaches are proposed to generate correspondences:(i)a model based approach that fits an articulated model to each individual to establish a correspondence map.(ii)an interest point operator approach that nominates a large number of potential correspondences which are evaluated using a region growing scheme.

2.3. Using Attribute-restricted latent topic model

This paper, proposes a novel Attribute-Restricted Latent Topic Model(ARLTM)[12] to encode targets into semantic topics.Although the appearance of the target varies in different poses and from different view points, humans recognize them all asthe target is wearing a black and gray patched jacket and dark blue uniform pants.This semantic description is invariant with pose, view, and other such changes in appearance.Thus the paper propose to use such intermediate invariable descriptions as a reliable repre sentation for person re-identification.

Background subtraction and human tracking are first carried out inorder to capture the human targets in continuous video streams.For the targets captured,we then extract color and texture features such as localHSVhistogram. These local features are then clustered into sets of visual words,thus called codebooks. This intermediate representation encodes high level semantic information for human appearance modeling in multi-camera surveillance systems. The values of pre-defined attributes are then manually assigned to the human target. ARLTM is then trained under the restriction of the human-specific attributes. ARLTM, bridges the gap between human-specific individual attributes and self-organized topic models, and can be easily generalized and applied to other attribute-based situations.

2.4. Using Haar-based and DCD-based Signature

A person can be recognized in one camera if his/her visual signature has been previously extracted in another camera. Here the re-identification algorithm was tested with manually annotated people in two non-overlapping scenes to validate the method. The algorithm uses Histograms of Oriented Gradients (HOG) [9]to automatically detect and track humans. Each detected human is tracked in order to accumulate images with person of interest. From these images we extract human blobs using foreground-background separation technique. Finally, sets of human blobs are used by AdaBoost scheme to create a reliable visual signature. The AdaBoost scheme is applied to haar-like features[13] and to dominant color descriptor. For each detected and tracked human a visual signature is generated. All such created visual signatures from different scenes are stored in one human signature database. The performance evaluation of our re-identification algorithms is based on querying the human signature database by extracted signatures. The results are analyzed using cumulative matching characteristic (CMC) curve.

3. MINIMUM CLASSIFICATION ERROR-BASED KISS METRIC LEARNING

3.1. KISS Metric Learning Review

Kostinger proposed KISS metric learning[5] (KISS) based on an assumption that pairwise differences are Gaussian distributed. This has acquired state-of-the-art retrieval performance for real applications, such as person reidentification and face recognition.

Considering the person reidentification problem, a feature vector pair X_i and X_j represents two samples. The hypothesis H_0 can assume that the feature vector pair is dissimilar, i.e., X_i and X_j are sampled from different people, and the hypothesis H_1 denotes that the feature vector pair is similar, i.e., X_i and X_j are sampled from the same person. Equation (1) defines the logarithm of the ratio between the two posteriors.

$$\delta(X_i, X_j) = \log\left(\frac{p(X_i, X_j \mid H^0)}{p(X_i, X_j \mid H_1)}\right) \quad (1)$$

For metric learning, a large $\delta(X_i, X_j)$ indicates X_i and X_j represent different people, while a small $\delta(X_i, X_j)$ indicates X_i and X_j represent a same person. We denote the difference of the feature vector pair by $X_{ij} = X_i - X_j$, and thus we have

$$\delta(X_{ij}) = \log\left(\frac{p(X_{ij} \mid H^0)}{p(X_{ij} \mid H_1)}\right) \quad (2)$$

which can be rewritten as

$$\delta(X_{ij}) = \log\left(\frac{f(X_{ij} \mid \theta^0)}{f(X_{ij} \mid \theta_1)}\right) \quad (3)$$

where $f(X_{ij})$ is the probability density function with parameter θ for hypothesis H. After assuming the difference space is a Gaussian structure, we have

$$f(X_{ij} \mid \theta_k) = \frac{1}{2\pi^{d/2} \mid \Sigma^{1/2} \mid} \exp\left(-\frac{1}{2} X_{ij}^T \Sigma_k^{-1} X_{ij}\right) \quad (4)$$

where $k \in \{0, 1\}$, d is the dimensionality of the feature vector, and Σ_k is the covariance matrix of X_{ij}. Note that for specific i and j, since both X_{ij} and X_{ji} belong to the pairwise difference set, we have $\sum_{i,j} X_{ij} = 0$ i.e., zero mean and $\theta_1 = (0, \Sigma_1)$ and $\theta_0 = (0, \Sigma_0)$. Given(4), (3) can be written as

$$\delta(X_{ij}) = \frac{1}{2} X_{ij}^T (\Sigma_{11} - \Sigma_{01}) X_{ij} + \frac{1}{2}\log\left(\frac{|\Sigma_1|}{|\Sigma_0|}\right) \quad (5)$$

By dropping the constant terms we have

$$\delta(X_{ij}) = X_{ij}^T (\Sigma_{11} - \Sigma^{01}) X_{ij} \quad (6)$$

Define y_{ij} as the indicative variable of X_i and X_j. $y_{ij} = 1$ if X_i and X_j are the same person, otherwise $y_{ij} = 0$. Let N_1 denote the number of similar feature vector pairs, while N_0 denotes the number of dissimilar feature vector pairs. The covariance matrices are estimated as follows:

$$\Sigma_0 = \frac{1}{N_0} \sum_{y_{ij}=0} X_{ij} X_{ij}^T = \frac{1}{N_0} \sum_{y_{ij}=0} (X_i - X_j)(X_i - X_j)^T$$
$$\Sigma_1 = \frac{1}{N_1} \sum_{y_{ij}=0} X_{ij} X_{ij}^T = \frac{1}{N_1} \sum_{y_{ij}=0} (X_i - X_j)(X_i - X_j)^T \quad (7)$$

Equation (7) shows that the eigenvalues of Σ_0 and Σ_1 are positive. Let KISS project $\Sigma^{11} - \Sigma^{01}$ onto the cone of the positive a semi-definite matrix M, so we have

$$\delta(X_{ij}) = X_{ij}^T M X_{ij} \quad (8)$$

where M is the KISS metric matrix.

3.2. MCE-KISS Metric Learning

Although KISS has largely improved the accuracy of person reidentification, there is a lot of room to improve efficiency and stability. It is critical to estimate the covariance matrices in (3.6) accurately to improve performance for person reidentification. It is known that the model of Gaussian distribution suffers from estimate error given limited training samples. Specifically, it is laborious and tedious to get a large number of labeled samples in real applications, to overcome the estimate error of the small eigenvalues of the covariance matrices which arose through the problem of small sample size. In statistics, to obtain robust estimations, a large number of techniques have been proposed. In this paper, the smoothing technique[7] , and the MCE criterion[6], are introduced to improve the accuracy of estimates of covariance matrices in KISS. By enlarging the estimate to the small eigenvalues of a covariance matrix, the smoothing technique can compensate for the decrease in performance which arose from the estimate errors of small eigenvalues. On the other hand, the covariance matrices of KISS are estimated by ML estimation. It is known that the ML estimation for Gaussian density model is imprecise, the discriminative learning procedure based on MCE aims to adjust the parameters of Gaussian density model and improves the generalization ability by increasing the number of training samples.

The covariance matrix Σ_i in (6) is first diagonalized and can be written as

$$\Phi_i \Lambda_i \Phi^i \quad (9)$$

where $\Lambda_i = \mathrm{diag}\,[\lambda_{i1}, \lambda_{i2}, \ldots \lambda_{id}]$ with λ_{ij} being an eigenvalue of Σ_i, $\Phi_i = [\varphi_{i1}, \varphi_{i2} \ldots \varphi_{id}]$ with φ_{ij} being an eigenvector of Σ_i, and eigenvalues in Λ_i are arranged in a descending order.

we substitute (9) into (6)

$$\delta(X_{ij}) = X_{ij}^T (\Sigma_{11} - \Sigma^{01}) X_{ij}$$
$$= X_{ij}^T (\Phi_1 \Lambda^{11} \Phi^1 - \Phi_0 \Lambda^{01} \Phi^0) X_{ij} \quad (10)$$
$$\delta(X_{ij}) = \sum_{n=1}^{d} \frac{1}{\lambda_{1n}} (\varphi_{1n} X^{ij})^2 - \sum_{n=1}^{d} \frac{1}{\lambda_{0n}} (\varphi_{0n} X^{ij})^2$$

Through (10), we can explain that the small eigenvalues significantly affect the score of metric. Next, we replace the small eigenvalues of the covariance matrix with a small constant β_i

$$\Lambda_i = \mathrm{diag}[\lambda_{i1}, \lambda_{i2} \ldots \lambda_{ik}, \beta_i, \ldots \beta_i] \quad (11)$$

Let there be d-k number of β_i, where d is the dimensionality of training samples. Taking into account the smoothing technique, the constant β_i is set to the value of the average of all the small eigenvalues.

$$\beta_i = \frac{1}{d-k}\sum_{n=k+1}^{d} \lambda in \qquad (12)$$

Thus (10) can be written as

$$\delta(X_{ij}) = \sum_{n=1}^{k} \frac{1}{\lambda_{1n}} (\varphi 1n X^{ij})^2 + \sum_{n=k+1}^{d} \frac{1}{\beta_1} (\varphi 1n X^{ij})^2 - \sum_{n=1}^{k} \frac{1}{\lambda_{0n}} (\varphi 0n X^{ij})^2 - \sum_{n=k+1}^{d} \frac{1}{\beta_0} (\varphi 0n X^{ij})^2$$

$$\delta(X_{ij}) = \sum_{n=1}^{k} \frac{1}{\lambda_{1n}} (\varphi 1n X^{ij})^2 + \frac{1}{\beta 1} \left(\|X_{ij}\|^2 - \sum_{n=1}^{k} (\varphi^{1n} X_{ij})^2\right) - \sum_{n=1}^{k} \frac{1}{\lambda_{0n}} (\varphi 0n X^{ij})^2 - \frac{1}{\beta 1} \left(\|X_{ij}\|^2 - \sum_{n=1}^{k} (\varphi^{0n} X_{ij})^2\right) \qquad (13)$$

According to the MCE criterion, we need to optimize the parameters of covariance matrices by utilizing the gradient descent method. Then, we have the evaluation of misclassification of a sample x belonging to class c.

$$hc(x) = \max\delta(X, Xc) - \min\delta(X, Xr) \qquad (14)$$

where X_c is a sample of the class c, and X_r is the closest interclass sample.Equation (14) considers two aspects: 1) the distance between x and the farthest intraclass sample and 2) the distance between x and the closest interclass sample.

Furthermore, the loss of the misclassification can be written as

$$lc(x) = \frac{1}{1 + e^{-\xi h} c^{(x)}} \qquad (15)$$

where ξ is a trade-off parameter and is selected in the range of $(0,+\infty]$. Given the training samples $\{X_n \mid n = 1, 2, ..., N\}$, and the label of each sample $\{C_i \mid i = 1, 2, ..., M\}$, we can compute the empirical loss by using (16)

$$L = \frac{1}{N} \sum_{n=1}^{N}\sum_{i=1}^{M} l_i(X_n) I(X_n \in C_i) \qquad (16)$$

$$\begin{aligned} I(X_n \in C_i) &= 1, \text{if} X_n \in C_i \\ I(X_n \in C_i) &= 0, \text{if} X_n \in C_i \end{aligned} \qquad (17)$$

And (16) can be further deduced to

$$L = \frac{1}{N} \sum_{n=1}^{N} lc(Xn) \qquad (18)$$

where c is the class information. According to (18), we expect that the distance between x and the farthest intraclass sample are as small as possible and the distance between x and the closest interclass sample are as large as possible. The parameters in KISS include the eigenvectors and eigenvalues of Σ_0 and Σ_1, i.e.,λ_{1n}, β_1, λ_{0n}, β_0, φ_{1n}, φ_{0n}. In MCE-KISS, we minimize the empirical loss L by means of adjusting these parameters via gradient descent. Let θ denote the parameters, according to gradient descent, we can get a general update rule of parameters.

$$\begin{aligned} \theta(t+1) &= \theta(t) - \varepsilon(t) \frac{\partial L}{\partial\theta} \\ &= \theta(t) - \varepsilon(t) \frac{\partial lc(X)}{\partial\theta} \end{aligned} \qquad (19)$$

$$\frac{\partial l_c(X)}{\partial\theta} = (-1) . \frac{1}{(1 + e^{-\xi h} c^{(X)})^2} . (-\xi) e^{-\xi hc(X)} \frac{\partial hc(X)}{\partial\theta} = \xi l_c(X)(1 - l_c(X)) \frac{\partial hc(X)}{\partial\theta} \qquad (20)$$

$$\frac{\partial h_c(X)}{\partial\theta} = \left(\frac{\partial\delta(X, X^c)}{\partial\theta} - \frac{\partial\delta(X, X^r)}{\partial\theta}\right) \qquad (21)$$

According to (19) to (21) we have

$$\theta(t+1) = \theta(t) - \varepsilon(t)\xi l_c(X)(1 - l_c(X)) \left(\frac{\partial\delta(X, X^c)}{\partial\theta} - \frac{\partial\delta(X, X^r)}{\partial\theta}\right) \qquad (22)$$

In the learning process, we need to guarantee eigenvalues are positive, so we further define

$$\lambda_{in} = e^{\sigma}in , \beta_{in} = e^{\tau}i \qquad (23)$$

We rewrite (23) to

$$\sigma_{in} = \ln_{\lambda} in , {}^{\tau}i = {}^{\ln}\beta in \qquad (24)$$

Based on (22), we convert the parameter updating to the computation of partial derivatives (25)-(28).

$$\begin{cases} \frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\tau_1} = -e^{-\tau_1}\left[\|\mathbf{x} - \mathbf{x}_j\|^2 - \sum_{n=1}^{k}\left[\phi_{1n}^T(\mathbf{x} - \mathbf{x}_j)\right]^2\right] \\ \frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\tau_0} = e^{-\tau_0}\left[\|\mathbf{x} - \mathbf{x}_j\|^2 - \sum_{n=1}^{k}\left[\phi_{0n}^T(\mathbf{x} - \mathbf{x}_j)\right]^2\right] \end{cases} \qquad (25)$$

We can optimize the parameters of Σ_0 and Σ_1 by using (25)-(28) and the distance metric $\delta(x_{ij})$ satisfied theneeds of the MCE.

$$\begin{cases} \frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\sigma_{1n}} = -e^{-\sigma_{1n}}\left[\phi_{1n}^T(\mathbf{x}-\mathbf{x}_j)\right]^2 \\ \frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\sigma_{0n}} = e^{-\sigma_{0n}}\left[\phi_{0n}^T(\mathbf{x}-\mathbf{x}_j)\right]^2 \end{cases} \quad (26)$$

$$\frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\phi_{1nl}} = \frac{\partial}{\partial\phi_{1nl}}\left\{\begin{array}{l} \sum_{n=1}^{k} e^{-\sigma_{1n}}\left[\sum_{l=1}^{d}\phi_{1nl}(\mathbf{x}-\mathbf{x}_j)_l\right]^2 \\ +e^{-\tau_1}\left[\left\|\mathbf{x}-\mathbf{x}_j\right\|^2\right. \\ \left.-\sum_{n=1}^{k}\left(\sum_{l=1}^{d}\phi_{1nl}\cdot(\mathbf{x}-\mathbf{x}_j)_l\right)^2\right] \\ -\sum_{n=1}^{k} e^{-\sigma_{0n}}\left[\sum_{l=1}^{d}\phi_{0nl}(\mathbf{x}-\mathbf{x}_j)_l\right]^2 \\ -e^{-\tau_0}\left[\left\|\mathbf{x}-\mathbf{x}_j\right\|^2\right. \\ \left.-\sum_{n=1}^{k}\left(\sum_{l=1}^{d}\phi_{0nl}\cdot(\mathbf{x}-\mathbf{x}_j)_l\right)^2\right] \end{array}\right\} \quad (27)$$

$$\begin{aligned} &= e^{-\sigma_{1n}}\cdot 2\left[\phi_{1n}^T(\mathbf{x}-\mathbf{x}_j)\right]\cdot(\mathbf{x}-\mathbf{x}_j)_l \\ &\quad +e^{-\tau_1}.(-1).2\left[\phi_{1n}^T(X-X_j)\right].(X-X_j)_l \\ &= 2\left(e^{-\sigma_{1n}}-e^{-\tau_1}\right)\left[\phi_{1n}^T(X-X_j)\right](X-X_j)_l \end{aligned}$$

$$\frac{\partial\delta(\mathbf{x},\mathbf{x}_j)}{\partial\phi_{0nl}} = -2\left(e^{-\sigma_{0n}}-e^{-\tau_0}\right)\left[\phi_{0n}^T(\mathbf{x}-\mathbf{x}_j)\right](\mathbf{x}-\mathbf{x}_j)_l. \quad (28)$$

Based on the above discussions, we summarize MCE-KISS as:

Algorithm for Minimum Classification Error-KISS

Step 1: The initial Σ^{01} and Σ^{11} are calculated by using (7) and (8).
Step 2: Smooth technique: By using (12) to amend for the estimation errors of small eigenvalues of Σ^{01} and Σ^{11} Step 3: MCE technique: By using (25), (26), (27), and (28) optimize the parameters of Σ^{01} and Σ^{11}
Step 4: The distance metric is calculated by using (6).

The procedure for MCE-KISS-based person reidentification can be summarized by the following steps:
1) partitioning the image into a regular grid of size 8 4 and overlapping block of size 88, and the color features and texture features are extracted from the overlapping blocks; 2) concatenating all the feature descriptors together and conducting PCA to achieve a robust feature representation for each sample;
3) training MCE-KISS;
4) finally finding the matching rank according to the query target.

The main contribution of this paper include the following.
1) The newly proposed MCE-KISS by seamlessly integrating MCE criterion and smoothing technique to significantly improve the performance of KISS metric learning.
2) A rough comparison between MCE KISS and KISS to demonstrate its robust and effectiveness.

4. EXPERIMENTAL RESULTS

The widely used yet challenging dataset, VIPeR by Gray et al [10] was used to demonstrate the effectiveness of the proposed MCE-KISS method. All images from the two datasets were normalized to a standard size of 128 x 48. In general, this manipulation causes shape distortion which has limited effect on human visual systems. For each image, we concatenated the extracted LBP descriptor and some color features into a high dimensional feature vector.

In experiments, all samples of p_{ts} subjects were selected to form the test set, while the rest p_{tr} were used for model training. During training, we used intraperson image pairs as similar pairs, and generated interperson image pairs (by randomly selecting two images from different subjects) as dissimilar pairs.The image pairs are used to estimate Σ^{01} and Σ^{11} according Algorithm 1. During testing, the test set were divided into two parts, i.e., a gallery set and a probe set. We randomly chose one sample of each subject to comprise the gallery set. The rest were used for the probe set. Person reidentification aims to identify a persons photo in the probe set by comparing it with images of several individuals stored in the gallery set.

By using the average cumulative match characteristic (CMC) curves, we evaluated the performance of the proposed algorithm. Because the complexity of the reidentification problem, the top n-ranked matching rate was considered.

4.1. VIPeR Dataset

The VIPeR dataset was collected by Gray et al [10] and contains 1264 outdoor images obtained from two views of 632 subjects. Intraperson image pairs may contain a viewpoint change of 90 degree celsius. Other variations are also considered, such as lighting conditions, shooting locations, and image quality. Thus, it is challenge to conduct image-based person reidentification on the VIPeR dataset. Example images are shown in Fig. 4.1.

We set p_{ts} = 316 and p_{ts} = 632, respectively to evaluate the matching performance of different algorithms. We repeated the process and the average CMC curves were depicted.

4.2. Feature Descriptors

It is known that both texture features and color histograms are useful for person reidentification. In our experiments, each image was partitioned into a regular

Figure 1. Some typical samples from the VIPER dataset

grid with 8 pixel spacing in the horizontal direction, and 4 pixel spacing in the vertical direction. From the grid, the LBP descriptor, HSV histogram, and RGB histogram were extracted from overlapping blocks of size 8 8. The HSV and RGB histograms encoded the different color distribution information in the HSV and RGB color spaces, respectively. The texture distribution information was modeled effectively by LBP descriptor. All the feature descriptors were concatenated together. Figure shows the process of feature extraction. We conducted PCA to obtain a 40-dimensional representation, to suppress the Gaussian noise.

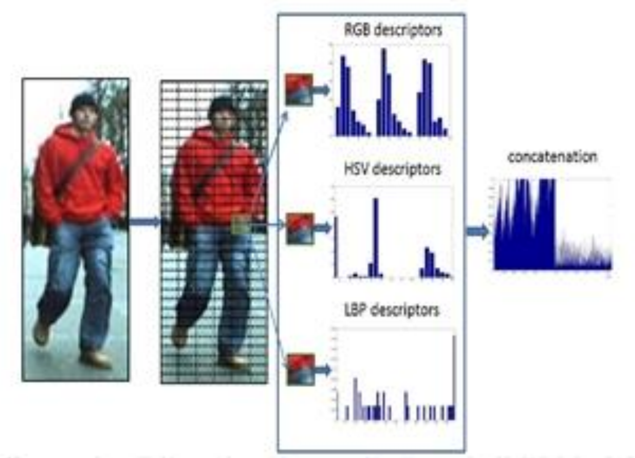

Figure 2. Feature extraction

5. RESULT ANALYSIS

The graph shows the comparison of our proposed MCEKISS metric learning with KISS. The x-coordinate is the rank score, and the y-coordinate is the matching rate. MCE-KISS integrates the smoothing technique and the MCE criterion for precise covariance matrix estimation.

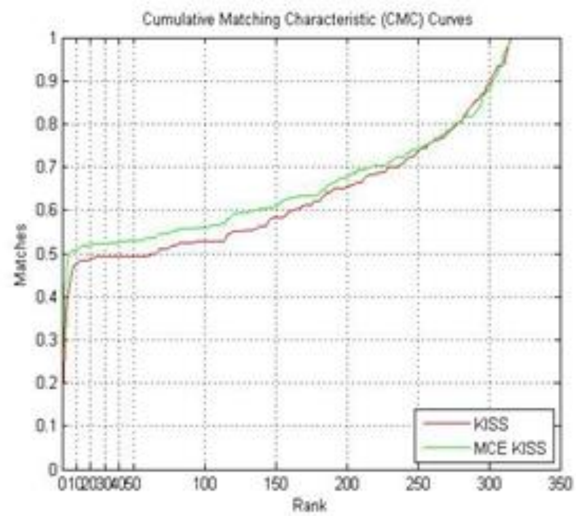

Figure 3. Comparison between MCE KISS and KISS

6. CONCLUSION

The distance metric is critically important for effective person reidentification in surveillance tasks. Thus, it is rational to find a suitable distance metric learning algorithm to boost the performance of person reidentification.The problem in this is there are only limited training image pairs to learn a metric in person reidentification.

Given a small number of training samples, we observe that covariance matrices estimated by KISS are biased. Therefore, we present the MCE-KISS. The proposed MCE-KISS algorithm exploits the smoothing technique to enlarge the small eigenvalues of the estimated covariance matrix, and discriminative learning based on MCE which is more reliable than classical ML estimation. The employed two statistical techniques effectively enlarge the underestimated small eigenvalues and better estimate the covariance matrix. Therefore, MCE-KISS significantly improves KISS for person reidentification.

7. FUTURE ENHANCEMENT

The algorithm demands as input, the number of small eigen values to be replaced by its average. Thus making the algorithm to take the number automatically according the input data, can be considered as a future work.

REFERENCES

[1] V. Chatzis, A. G. Bors, and I. Pitas, Multimodal decision-level fusion for person authentication, IEEE Trans. Syst., Man, Cybern. A, Syst., Humans, vol. 29, no. 6, pp. 674680, Nov. 1999

[2] F. Dornaika and A. Bosaghzadeh, Exponential local discriminant embedding and its application to face recognition, IEEE Trans. Cybern., vol. 43, no. 3, pp. 921934, Jun. 2013.

[3] R. M. Da Costa and A. Gonzaga, Dynamic features for iris recognition, IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 42, no. 4, pp. 10721082, Aug. 2012.

[4] R. Cappelli, M. Ferrara, and D. Maio, A fast and accurate palmprint recognition system based on minutiae, IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 42, no. 3, pp. 956962, Jun. 2012.

[5] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, Large scale metric learning from equivalence constraints, in Proc. IEEE CVPR, Providence, RI, USA, 2012, pp. 22882295.

[6] B.-H. Juang, W. Hou, and C.-H. Lee, Minimum classification error rate methods for speech recognition, IEEE Trans. Speech Audio Process., vol. 5, no. 3, pp. 257265, May 1997.

[7] F. Kimura, K. Takashina, S. Tsuruoka, and Y. Miyake, Modified quadratic discriminant functions and the application to Chinese character recognition, IEEE Trans. Pattern Anal. Mach. Intell., vol. 9, no. 1, pp. 149153, Jan. 1987.

[8] T. Ojala, M. Pietikainen, and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 7, pp. 971987, Jul. 2002.

[9] R. Lienhart and J. Maydt, An extended set of Haar-like features for rapid object detection, in Proc. ICIP, 2002, pp. 900903.

[10] D. Gray, S. Brennan, and H. Tao, Evaluating appearance models for recognition, reacquisition, and tracking, in Proc. 10th PETS, 2007.

[11] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, Large scale metric learning from equivalence constraints, in Proc. IEEE CVPR, Providence, RI, USA, 2012, pp. 22882295.

[12] Liu, Xiao, et al. "Attribute-restricted latent topic model for person re-identification." Pattern recognition 45.12 (2012): 4204-4213.

[13] Bak, Slawomir, et al. "Person re-identification using haar-based and dcd-based signature." Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on. IEEE, 2010.

[14] Lan, Rushi, et al. "Person reidentification using quaternionic local binary pattern." Multimedia and expo (ICME), 2014 IEEE international conference on. IEEE, 2014.

[15] Gheissari, Niloofar, Thomas B. Sebastian, and Richard Hartley. "Person reidentification using spatiotemporal appearance." Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 2. IEEE, 2006.

[16] Ma, Andy J., et al. "Cross-domain person reidentification using domain adaptation ranking svms." Image Processing, IEEE Transactions on 24.5 (2015): 1599-1613.

[17] Li, Annan, et al. "Clothing attributes assisted person reidentification." Circuits and Systems for Video Technology, IEEE Transactions on 25.5 (2015): 869-878.

[18] E. McDermott, T. J. Hazen, J. Le Roux, A. Nakamura, and S. Katagiri, Discriminative training for large-vocabulary speech recognition using minimum classification error, IEEE Trans. Audio, Speech, Lang. Process., vol. 15, no. 1, pp. 203223, Jan. 2007.