PERSON REIDENTIFICATION USING MCE-KISS METRIC LEARNING WITH MAXIMUM LIKELIHOOD FUNCTION

Sree Vrinda G M
M.Tech student
Department of Computer Science and Engineering
Mohandas College of Engineering
Anad, Trivandrum

Dr. P.Jayaprakash
Head of the Department
Department of Computer Science and Engineering
Mohandas College of Engineering
Anad, Trivandrum

Abstract*—Nowadays, in the area of Intelligent Video Surveillance (IVR), person reidentification receives an intensive attention. Person reidentification aims to match an instance of a person captured by one camera system to the instance of a person captured by another camera system. It is considered as a challenging problem because the appearance of person varies through the scenes, lightning conditions, shadows, different pose of person that has to be searched for. Recently, many algorithms proposed like LMNN, ITML are not suitable for large training samples. This paper introduces Minimum Classification Error (MCE) based KISS metric algorithm with smoothing technique to improve reidentification. Smoothing technique is done with maximum likelihood functions which enlarge small eigenvalues in the estimated covariance matrices.

***Index Terms*— IVR, person reidentification, MCE, metric learning.**

I. Introduction

In recent years, one of the main challenging task in the field of intelligent video surveillance (IVR) is person reidentification. Nowadays, more number of non overlapping camera networks has been set up. These network of cameras helps for monitoring pedestrian activities over a large public area such as the parking lot, airport, metro station, etc. This identification is used to acquire an individual's complete movements along that area. An objective of person reidentification is to verify a person has been already captured by another camera networks. In previous years, traditional biometrics such as face [2], [3], iris [4], fingerprint [5], and gait [6], are used for reidentification, but now they are not using for this purpose because images taken from this are variable, low quality and contain motion blur.

When a person stays within a camera's view, that particular person's position, background effects and lighting conditions are known to the system. The main problems arise when the person moves out from one camera's view and enters into another camera's view. So the system must know that the person seen in one camera is the same person that is already seen in another camera. If there is any issue with the system regarding matching the instance of person, that issue is known as re-identification problem. Person reidentification faces 3 main problems. At first, the segmented and comparison parts should be determined. Second, invariant signatures should be generated for comparing the corresponding parts. And at last compare the signatures by applying appropriate metric. Those steps are depicted in Figure. 1.

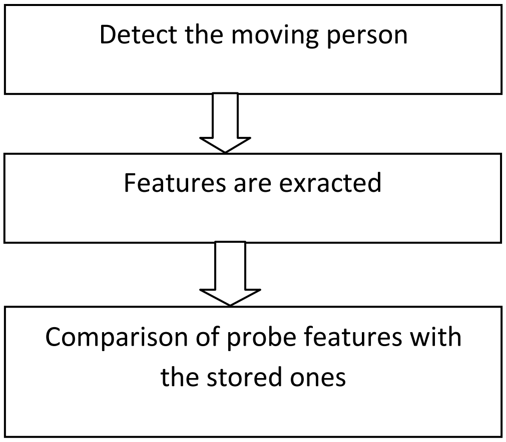

Figure. 1 Steps in reidentification

Reidentification problem has two methods for reidentifying, they are Appearance –based methods and gait-based methods. In appearance-based methods it extracts signatures from color, texture and other appearance properties. In gait-based methods, it extract features from the gait and motion of persons. More importantly, gait based methods are useful because the result obtained through this are not affected by varying lighting conditions between cameras.

Basically, two important stages which need to be focused for person reidentification are distance learning and visual feature extraction and selection. So many exciting studies have been performed on this area for improving person reidentification. Here, we briefly review some representative works. For visual feature extraction, the use of color features such as HSV and RGB color histograms will effectively save computational cost. In feature extraction, the usage of filters for clearance will improve the performance. Gabor [7] and Schmid [8] filters have been added to extraction. The texture extraction process can use Scale Invariant Feature Extraction (SIFT) [9] or Speeded up Robust Features (SURF) [10]. It is better to use LBP descriptor at last for texture classification to exploit person reidentification. This LBP (Local Binary Pattern) [11] descriptor is commonly used for facial image description and used here for getting local geometric structure of an image.

The stage distance learning can significantly improve the performance of retrieval applications. In this paper, robust distance learning is applied to update the retrieval precision. Traditionally several approaches are used for image retrieval applications but still they shouldn't produce any good result for person reidentification. Among other approaches, KISS metric learning is efficient and effective. In KISS metric learning, the results are estimated by evaluating covariance which will produce more accurate results robustly.

In this paper, we introduce the minimum classification error (MCE) [1] based KISS metric distance. The eigenvalues of the covariance matrices are biased in KISS metric algorithm. Because of biased matrices, reidentification will lead in poor performance. The covariance matrices of KISS are estimated using maximum likelihood (ML) estimation. But with increasing number of training samples, MCE criterion is more preferable than classical ML estimation. The MCE criterion technique does not work well with small eigenvalues in the covariance matrices. Therefore, the smoothing technique is required to improve the small eigenvalues of a covariance matrix. Here, we introduced maximum likelihood function in a logarithmic way for smoothing technique. Maximum likelihood function selects a parameter from a set of parameters based on probability.

The main steps for MCE-KISS-based person reidentification can be summarized by the following steps:

1) Images are partitioned into a regular grid of size 8 x 4, and the color features and texture features are extracted

2) Concatenating the feature descriptors together

3) Conducting PCA to obtain dimensional representation

4) Training MCE-KISS with smoothing technique; and

5) Finally finding the matching rank.

The main contributions of MCE-KISS are;

- The proposed MCE-KISS algorithm integrates MCE criterion and smoothing technique to improve the performance of KISS metric learning.
- KISS metric integrates with maximum likelihood function for smoothing technique.
- The newly proposed MCE-KISS exploits a discriminative learning procedure to effectively adjust the parameters of Gaussian density model.

II. Related Work

In Section I, we briefly reviewed the techniques used in person reidentification. It is worth noting the importance of distance learning schemes, which have been receiving increasing attention because the retrieval quality is known to be highly dependent on distance metrics. Porikli [12] proposed a new distance learning algorithm for the color calibration problem. This method is mainly used only on multicamera systems. Weinberger [13] and Saul proposed a large margin nearest neighbor metric (LMNN) which is used to improve the performance of the classical *k*NN (k-nearest neighbor) classification. But still this technique is time consuming. From the perspective of information theoretic, Davis *et al*. [14] proposed information theoretic metric learning (ITML), which is built on the Mahalanobis distance metric.

Recently, Zheng *et al*. [15] relative distance comparison (RDC). In this scheme, they assumed that wrong and right matches correspond to large and small distances, respectively. In contrast to consider modeling the similarity globally, local distance metric is proposed by Yang *et al*. [16] to improve the performance of retrieval and classification accuracy. However, most methods may perform poorly when the view conditions change greatly and the training samples are insufficient. In addition to distance metric learning-based matching schemes, researchers have exploited other schemes to improve retrieval precision. By utilizing subspace learning, Javed *et al*. [17] introduced a brightness transfer function to reduce the illumination changes. Prosser *et al*. [18] introduced Rank support vector machines (RankSVM) to person reidentification and proposed ensemble RankSVM to handle the scalability issue.

III. Proposed System

A. KISS Metric Learning

KISS metric learning is proposed recently for the best retrieval performance in real world applications such as person reidentification, face recognition etc. It is based on the assumption that pair wise differences are Gaussian distributed.

Consider a set of image pair samples for person reidentification problem and extract feature descriptors from these. It is known that both texture features and color histograms are useful for person reidentification. After extraction, all the feature descriptors are concatenated together and these feature vector pairs are split into two sets as test and train sets using random permutation.

Consider S_a and S_b represents the samples of feature vector pair. Here two hypotheses are mentioned as H_1 and H_2, where H_1 can assume that the feature vector pair is dissimilar, i.e., S_a and S_b are sampled from different people, and the hypothesis H_2 can assume that the feature vector pair is similar, i.e., S_a and S_b are sampled from same person.

$$\delta(S_a, S_b) = \log\left(\frac{p(S_a, S_b|H_1)}{p(S_a, S_b|H_2)}\right) \quad (1)$$

Equation (1) defines the logarithm of ratio between the two hypotheses. For metric learning, a small $\delta(S_a , S_b)$ indicates the two samples represent same person, while large $\delta(S_a , S_b)$ indicates the two samples represent different people. Define X_{ab} as an indicative variable of S_a and S_b : $X_{ab}=1$ if S_a and S_b are the same person, otherwise $X_{ab}=0$.

Let P_1 denote the number of similar feature vector pairs and P_0 denotes the number of dissimilar feature vector pairs. The covariance matrices are estimated as:

$$\sum\nolimits_0 = \frac{1}{P_0}\sum_{X_{ab}=0}(S_a - S_b)(S_a - S_b)^T \quad (2)$$

$$\sum\nolimits_1 = \frac{1}{P_1}\sum_{X_{ab}=1}(S_a - S_b)(S_a - S_b)^T \quad (3)$$

Equation shows that the eigenvalues of $\sum_0$ and $\sum_1$ are positive.

Finally the KISS metric matrix M is calculated by, $\sum_1^{-1} - \sum_0^{-1}$

B. Adaptive MCE - KISS Metric Learning

KISS has largely improved the accuracy of person reidentification, but still there is a lot to improve its efficiency and stability. The result of matching can be improved by improving the accuracy of covariance matrices. Sometimes small eigenvalues must appear in covariance matrices and this will cause estimate errors. To avoid this errors, smoothing technique and MCE criterion are introduced to improve the accuracy of estimate of covariance matrices in KISS. By enlarging the estimated 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.

The covariance matrix $\sum_i$ is first diagnalized and can be written as

$$\sum\nolimits_a = \varphi_a \wedge_a \varphi_a^T \quad (4)$$

where $\wedge_a = diag[\lambda_{a1}, \lambda_{a2}, ..., \lambda_{an}]$ with λ_{ab} being an eigenvalue of $\sum_a$, $\varphi_a = [\varphi_{a1} , \varphi_{a2} , . . . , \varphi_{an}]$ with φ_{ab} being an eigenvector of $\sum_a$.

Recently, Tao et al. [1] proposed MCE-KISS by applying smoothing technique and in that the small eigenvalues of covariance matrix are replaced with a constant value. This constant is set to the value of the average of all the small eigenvalues. Here we used smoothing technique with maximum likelihood function. Likelihood function estimates a parameter in an adaptive way from a set of statistics based on probability. In maximum likelihood function, it selects the set of values of the parameters that maximizes the likelihood function. Most importantly, the usage of logarithm achieves the maximum value from the likelihood function.

According to MCE metric learning, we need to minimize the empirical loss by updating the parameters via gradient descent method. We can compute the empirical loss by using (5)

$$E = \frac{1}{N}\sum_{n=1}^{N} l_{ci}\,(X_i) \qquad (5)$$

where 'ci' is the class information, N is the maximum number of training samples. The loss of misclassification can be estimated by using the following equation.

$$l_{ci}(x) = \frac{1}{1+e^{-\xi M_{ci}(x)}} \qquad (6)$$

where ξ is a trade-off parameter and is selected in the range of (0, +∞]. Here we have the evaluation of misclassification of a sample x belonging to class 'T'

$$M_T\,(x) = max_T\ \delta(x,\ x_T\,) - min_r\ \delta(x,\ x_r\,) \qquad (7)$$

where x_c is a sample of the class c, and x_r is the closest interclass sample. First element in (7) represents the distance between x and the farthest intraclass sample and second element represents the distance between x and the closest interclass sample.

The parameters in MCE-KISS metric algorithm include the eigenvalues, eigenvectors and the updated constant values through smoothing technique i.e., λ_{1n}, λ_{0n}, β_0, β_1, φ_{1n} and φ_{0n}. Before updating we need to make sure that eigenvalues are positive, so we define

$$\begin{cases}\lambda_{an} = e^{\sigma_{an}} \\ \beta_a = e^{\tau_a}\end{cases} \qquad (8)$$

$$\begin{cases}\sigma_{an} = \ln \lambda_{an} \\ \tau_a = \ln \beta_a\end{cases} \qquad (9)$$

The parameters of covariance matrices $\sum_0^{-1}$ *and* $\sum_1^{-1}$ are optimized using (10), (11) and (12).

$$\begin{cases}\dfrac{\partial\delta(X,X_b)}{\partial\sigma_{1n}} = -e^{-\sigma_{1n}}\,[\,\varphi_{1n}^T\,(X - X)]^{\,2} \\ \dfrac{\partial\delta(X,X_b)}{\partial\sigma_{0n}} = -e^{-\sigma_{0n}}\,[\,\varphi_{0n}^T\,(X - X_b)]^{\,2}\end{cases}$$

(10)

$$\begin{cases}\dfrac{\partial\delta(X,X_b)}{\partial\tau_1} = -e^{-\tau_1}\,[\,||X - X_b\,||^{\,2} - \sum_{n=1}^{C}[\,\varphi_{1n}^T\,(X - X_b)]^{\,2}] \\ \dfrac{\partial\delta(X,X_b)}{\partial\tau_0} = -e^{-\tau_0}\,[\,||X - X_b\,||^{\,2} - \sum_{n=1}^{C}[\,\varphi_{0n}^T\,(X - X_b)]^{\,2}]\end{cases}$$

(11)

$$\begin{cases}\dfrac{\partial\delta(X,X_b)}{\partial\varphi_{1nl}} = 2(e^{-\sigma_{1n}} - e^{-\tau_1})[\,\varphi_{1n}^T(X - X_b)](X - X_b)_{\,l} \\ \dfrac{\partial\delta(X,X_b)}{\partial\varphi_{0nl}} = 2(e^{-\sigma_{0n}} - e^{-\tau_0})[\,\varphi_{0n}^T(X - X_b)](X - X_b)_{\,l}\end{cases}$$

(12)

Based on the above discussions, we can compute the distance metric by $\sum_1^{-1} - \sum_0^{-1}$

IV. RESULT ANALYSIS

In our experiments, 632 image pairs are taken as samples. From total samples half of the samples were selected to form test set and the rest were used for model training. Both texture features and color histograms are extracted using RGB histogram and LBP descriptor. During training, covariance matrices are estimated for similar and dissimilar feature vector pairs. The test set is divided into gallery set and probe set during training. Reidentification aims to identify a person's photo in the probe set by comparing it with images of several individuals in the gallery set.

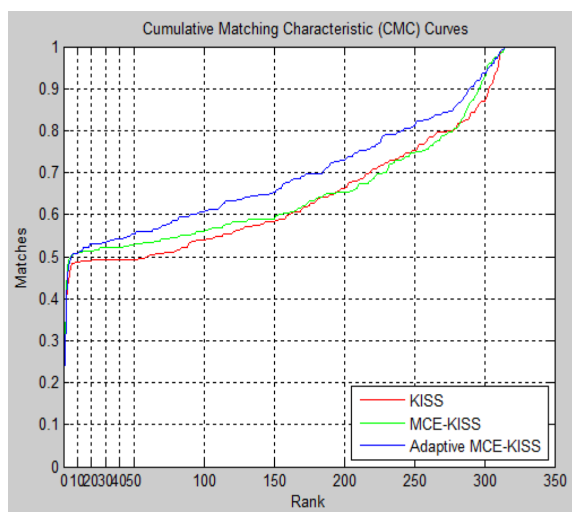

Figure 1. Performance comparison using CMC curve

The above figure shows the performance comparison of our proposed Adaptive MCE-KISS metric learning with MCE-KISS and KISS. Here, we evaluated the performance using cumulative match characteristic (CMC) curves. In the figure, x-coordinate represents the ranking score and y-coordinate represents the matching rate. The figure shows only the top 320 ranking positions.

KISS	0.82s
MCE-KISS	11.79s
Adaptive MCE-KISS	10.92s

Table 1. Training time

In Table 1, we compared the training time of Adaptive MCE-KISS metric learning with MCE-KISS and KISS.

The main observations from the comparisons are:

- Adaptive MCE-KISS integrates maximum likelihood function for smoothing technique and MCE criterion for getting precise covariance matrix estimation.
- Figure 1. shows that KISS performs poorly and also illustrates that it can work with limited training samples.
- Table 1. suggests Adaptive MCE-KISS metric learning as it needs less training time. This illustrates that the proposed algorithm is more reliable than the previous ones.

V. CONCLUSION

In recent years, distance metric algorithms are developed for effective person reidentification. ITML and LMNN algorithms have been developed but still these are suitable only for reidentifying limited training samples. Here we introduce MCE-KISS metric algorithm which works well in large training samples. In KISS algorithm, the estimated covariance matrices are biased because of small number of training samples. The proposed technique exploits the smoothing technique in maximum likelihood function to enlarge small eigenvalues in the estimated covariance matrix. Hence, MCE-KISS significantly improves the performance of KISS for person reidentification.

REFERENCES

[1] D. Tao, L. Jin, Y. Wang, and X. Li, "Person reidentification by minimum classification error-based KISS metric learning," *IEEE Trans. Cybern.*, 2014, 10.1109/TCYB.2014.2323992.

[2] 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. 674–680, Nov. 1999.

[3] F. Dornaika and A. Bosaghzadeh, "Exponential local discriminant embedding and its application to face recognition," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 921–934, Jun. 2013.

[4] R. M. Da Costa and A. Gonzaga, "Dynamic features for iris recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 4, pp. 1072–1082, Aug. 2012.

[5] 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. 956–962, Jun. 2012.

[6] I. Venkat and P. De Wilde, "Robust gait recognition by learning and exploiting sub-gait characteristics," *Int. J. Comput. Vis.*, vol. 91, no. 1,pp.7–23,2011.

[7] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biol.Cybern.*,vol.61,no.2,pp.103–113,1989.

[8] C. Schmid, "Constructing models for content-based image retrieval," in *Proc. IEEE CVPR*, 2001, pp. 39–45.

[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[10] X. Liu *et al.*, "Attribute-restricted latent topic model for person reidentification,"*Pattern Recognit.*, vol. 45, no. 12, pp. 4204–4213, 2012.

[11] 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. 971–987, Jul. 2002.

[12] F. Porikli, "Inter-camera color calibration by correlation model function," in *Proc. ICIP*, 2003, pp. 133–136.

[13] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, pp.207–244,Feb.2009.

[14] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon, Informationtheoretic metric learning," in *Proc. ICML*, Corvallis, OR, USA, 2007, pp. 209–216.

[15] W.-S. Zheng, S. Gong, and T. Xiang, "Reidentification by relative distance comparison," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 653–668, Mar. 2013.

[16] L. Yang, R. Jin, R. Sukthankar, and Y. Liu, "An efficient algorithm for local distance metric learning," in *Proc. AAAI*, 2006, pp. 543–548.

[17] O. Javed, K. Shafique, Z. Rasheed, and M. Shah, "Modeling intercamera space–time and appearance relationships for tracking across non-overlapping views," *Comput. Vis. Image Understand.*, vol. 109, no.2,pp.146–162,2008.

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