International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-5 Issue-6, August 2016

Adaptive Minimum Classification Error based KISS Metric Learning for Person Re-identification

Jasher Nisa A J, Sumithra M D

Abstract*— *Person re-identification becoming an interesting research area in the field of video surveillance and is taken as the area of intense research in the past few years. It is the task of identifying a person from a camera image, who is already been tracked by another camera image at different time at different location. Manual re-identification in large camera network is costly and mostly of inaccurate due to large number of camera that he had to simultaneously operate. In a crowded and unclear environment, when cameras are at a lengthy distance, face recognition is not possible due to insufficient image quality. So, visual features based on appearence of people, using their clothing, objects carried etc. can be exploited more reliably for re-identification. A person's appearence can change between different camera views, if there is large changes in view angle, lighting, background and occlusion, so visual feature extraction is not possible accurately. For solving a person re-identification problem, have to focus on "developing feature representations which are discriminative for identity,but invarient to view angle and lighting". Recently, Minimum Classification Error (MCE) based KISS metric learning is considered as one of the top level algorithm for person re-identification. It uses VIPeR feature set as input, which contains the extracted features. MCE-KISS is more reliable with increasing the number of training samples. It uses the smoothing technique and MCE criteria to improve the accuracy of estimate of eigen values of covarience metrics. The smoothing technique can compensate for the decrease in performance which arose from the estimate errors of small eigenvalues. Here, the value of average number of small eigen values of the covarience metrics is set as a constant. So it does not work well for a large number of samples. In such situation, introduce a new method to find the value of average of such small eigen values by maximizing the likelihood function. The new scheme is termed as Adaptive MCE-KISS and conduct validation experiments on VIPeR feature dataset.

Index Terms: ***reidentification, matric learning, covarience matrics, likelihood method.***

I. INTRODUCTION

Video surveillance is an important technique for criminology, security and forensic applications. Because of the increased volume of surveillance videos today, it is very difficult to searching objects from such a massive video. Person reidentification is very useful for such situations. Person reidentification is becoming a hot research area today due to its importance in variety of applications[1]. The task of person re-identication is to determine which person in a gallery has the same identity to a probe image. This task basically assumes that the subject of the probe image belongs to the gallery, that is, the gallery contains this person. State-of art person re-identification methods provide robust person matching through combining various feature vectors like RGB, HSV and LBP histograms. Among them, RGB & HSV are used to extract color information & LBP for texture distribution of the information. It uses color features with an assumption that person donot change their clothes as they move between cameras. In a crowded and unclear environment, when cameras are at a lengthy distance, conventional biometric like face recognition is not possible due to insufficient image quality. So, visual features based on appearence of people, using their clothing, objects carried etc. can be exploited more reliably for re-identification. But, visual appearence is weak for matching people, that means a person's appearence can change between different camera views, if there is large changes in view angle, lighting, background and occlusion, that is different people appearing alike than the same person. So we use person reidentification methods for obtaining the robust features.

For obtaining robust features for an image, dimension reduction is important to retain the most effective features for subsequent matching. This is because a combination of the previously mentioned selected features is usually deployed, and dimension reduction results in a an effective representation of the sample [2]. Dimension reduction algorithms have received increasing attentions in recent years. Principal component analysis (PCA) [3] is a representative classical linear algorithm. Distance learning can significantly improve the performance of retrieval applications[2]. In this paper, we aim to update the retrieval applications by applying robust distance learning. In previous studies, several approaches which have achieved top-level performance in image retrieval applications perform poorly for person reidentification. It is important to mention that MCE-KISS metric learning is both efficient and effective [2]. However, it is assumed that pairwise differences are sampled from a Gaussian distribution, has the small sample size problem for estimating the covariance, and therefore results in retrieval precision not always performing robustly in practice.

Manuscript published on 30 August 2016.
***** Correspondence Author (s)
Jasher Nisa A J, M.Tech Scholar, Department of Computer Science and Engineering, LBS Institute of Technology for Women, Thiruvananthapuram, India.
Sumithra M D, Assistant Professor, Department of Computer Science and Engineering, LBS institute of Technology for Women, Thiruvananthapuram, India.

In this paper, we introduce the Adaptive minimum classification error (AMCE) criterion to improve MCE-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 MCE-KISS are obtained by maximum likelihood (ML) estimation. With increasing the number of training samples, discriminative learning based on MCE is more reliable than classical ML estimation. In addition, the MCE criterion is widely used in the field of machine learning [4]. So, here smoothing technique [5] is required to improve the estimate of the small eigenvalues of a covariance matrix. But, in this technique the average number of small eigen values are set as a constant and is not ewll suited for large applications. So, we introduce a method by maximizing the likelyhood function to improve this. Such a new scheme is termed as Adaptive Minimum Classification Error based KISS matric learning or simply AMCE-KISS.

The procedure for MCE-KISS-based person reidentification can be summarized by the following steps: 1) partitioning the image into a grid of size 8 × 4 and overlapping block of size 8×8, 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 AMCE-KISS; and 4) finally finding the matching rank according to the query target. Furthermore, the Cumulative Matching Characteristic (CMC) curve is used to illustrate the identification accuracy versus the number of samples retrieved.

The rest of the paper is organized as follows: in Section II, review related works for improving distance learning for person reidentification. We detail the proposed AMCE-KISS in Section III. Section IV shows the experiment results on the VIPeR feature dataset.. We conclude the paper in Section V.

II. RELATED WORK

Here, we review the the different techniques used in person re-identification. It is very importance for noting the distance learning schemes, which has an increased attention [6], [7] because the retrieval quality is highly dependent on distance metrics. Weinberger and Saul [6] proposed a large margin nearest neighbor metric (LMNN) to improve the performance of the classical *k*NN classification. However, the computational processing of *k* closest within-class samples is time-consuming.

Yang *et al*. [8] proposed local distance metric 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. Prosser *et al*. [9] introduced Rank support vector machines (RankSVM) to person reidentification and proposed ensemble RankSVM to handle the scalability issue. Dapeng Tao and Lianwen proposed Regularized smoothing KISS (RSKISS) [10] by integrating smoothing and regularization techniques for robustly estimating covariance matrices. RS-KISS is superior to KISS, because RS-KISS can enlarge the underestimated small eigenvalues and can reduce the overestimated large eigenvalues of the estimated covariance matrix in an effective way. However, retraining RS-KISS on all the available examples in a straightforward way is time consuming. Dapeng Tao, Lianwen Jin again proposed an another scheme called MCE-KISS[2] uses the smoothing technique to improve the estimates of the small eigenvalues of a covariance matrix. But, here the number of small eigen values are set as a constant and is not well suited for large samples.

III. ADAPTIVE MINIMUM CLASSIFICATION ERROR-BASED KISS METRIC LEARNING

A. *KISS Metric Leraning Review*

It has been reported that the KISS metric learning (KISS) [11] has obtained the state of the art performance for person reidentification on the VIPeR dataset [12]. Given a feature vector pair $\mathbf{x_i}$ and $\mathbf{x_j}$ referring to two people, respectively, let H_0 denote the hypothesis that the feature vector pair is dissimilar ($\mathbf{x}_i$ and $\mathbf{x_j}$ are different people), and H_1 denote the hypothesis that the feature vector pair is similar ($\mathbf{x}_i$ and $\mathbf{x}_j$ are the same person). The logarithm of the ratio between the two posteriors is

$$\delta\,(x_i, x_j) = log(\, p(H_0|x_i,x_j\,) \,/\, \; p(H_1|x_i,x_j)) \qquad (1)$$

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

$$\delta(x_{ij}) = log\,(p(x_{ij}|H_0) \,/p\,(x_{ij}|H_1)) \qquad (2)$$

which can be rewritten as

$$\delta(x_{ij}) = log\,(f\,(x_{ij}|\theta_0) \,/f\,(x_{ij}|\theta_1)) \qquad (3)$$

where f ($\mathbf{x}ij|\theta$) is the probability density functions with parameteθ for hypothesis H. After assuming the difference space is a Gaussian structure, we have

$$f\,(x_{ij}|\theta_k) = 1/(2\pi^{d/2}|\textstyle\sum_K|^{1/2})\, \exp(-\frac{1}{2}\, x_{ij}{}^T \textstyle\sum_K{}^{-1} x_{ij}) \qquad (4)$$

where $k \in \{0, 1\}$, d is the dimensionality of the feature vector, and $\sum k$ is the covariance matrix of $\mathbf{x}ij$. Note that for specific i and j, since both $\mathbf{x}_{ij}$ and $\mathbf{x}_{ji}$ belong to the pairwise difference set, we have $\sum_{ij} x_{ij}$=0,i.e., zero mean and $\Theta_1 = (0,\sum_1)$ and $\Theta_0 = (0,\sum_0)$

Given (4), (3) can be rewritten as

$$\delta(x_{ij}) = \frac{1}{2} x_{ij}{}^T(\textstyle\sum_1{}^{-1}-\sum_0{}^{-1})x_{ij} + \frac{1}{2} log(\sum_1 | \sum_0) \qquad (5)$$

By dropping the constant terms, we have

$$\delta(x_{ij}) = x_{ij}{}^T(\textstyle\sum_1{}^{-1}-\sum_0{}^{-1})\, x_{ij} \qquad (6)$$

Define yij as the indicative variable of $\mathbf{x}i$ and $\mathbf{x}j$: y_{ij} =1 if $\mathbf{x}i$ and $\mathbf{x}j$ are the same person, otherwise y_{ij} =0. Let $N1$ denote the number of similar feature vector pairs, while $N0$ denotes the number of dissimilar feature vector pairs. The covariance matrices are estimated as follows:

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

$$\sum_0=\frac{1}{N_0}\sum_{y_{ij}=0} x_{ij}x_{ij}^T=\frac{1}{N_0}\sum_{y_{ij}=0}(x_i-x_j)(x_i-x_j)^T$$
$$\sum_1=\frac{1}{N_1}\sum_{y_{ij}=1} x_{ij}x_{ij}^T=\frac{1}{N_1}\sum_{y_{ij}=1}(x_i-x_j)(x_i-x_j)^T \qquad (7)$$

Let KISS project $\sum_1^{-1}-\sum_0^{-1}$onto the cone of the positive a semi-definite matrix M, so we have

$$\delta(x_{ij}) = x_{ij}^T M x_{ij} \qquad (8)$$

where M is the KISS metric matrix.

B. MCE-KISS Metric Learning Review

Although KISS improved the accuracy of person reidentification, there is a way to improve efficiency and stability. It is necessary to estimate the covariance matrices in (6) accurately to improve performance for person reidentification. The model of Gaussian distribution have estimate error given limited training samples. It is very critical 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. To obtain robust estimations, a large number of techniques have been proposed. In this paper, the smoothing technique [13], and the MCE criterion [14], are introduced to improve the accuracy of estimates of covariance matrices in KISS. By increasing 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. The covariance matrix $\sum i$ in (6) is first diagonalized and can be written as

$$\sum i=\Theta_i\Lambda_i\Theta_i^T \qquad (9)$$

where Λ_i = diag [$\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{id}$] with λ_{ij} being an eigenvalue of $\sum i$, Φi = [$\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{id}$] with φ_{ij} being an eigenvector of $\sum i$, and eigenvalues in Λ_i are arranged in a descending order. By substitiing (9) into (6) we can explain that small eigen values significantly affect score of metric. Next, we replace the small eigenvalues of the covariance matrix with a small constant β_i.

$$\Lambda_i=diag[\lambda_{i1},\lambda_{i2},\ldots\ldots,\lambda_{iN},\underbrace{\beta_i,\ldots\ldots,\beta_i}_{k}] \qquad (10)$$

where d is the dimensionality of training samples. Taking into account the smoothing technique, the constant βi is set to the value of the average of all the small eigenvalues

$$\beta_i=\frac{1}{d-k}\sum_{n=k+1}^{d}\lambda_{in} \qquad (11)$$

According to the MCE criterion, we need to optimize the parameters of covariance matrices by utilizing the gradient descent method. The parameters in KISS include the eigenvectors and eigenvalues of $\sum 0$ and $\sum 1$, i.e., λ_{1n}, β_1, λ_{0n}, β_0, φ_{1n} and φ_{0n}. In MCE-KISS, we minimize the empirical loss L by means of adjusting these parameters via gradient descent. Then

$$\sum_0=\lambda_{0n}*\beta_0*\varphi_{0n} \qquad (12)$$
$$\sum_1=\lambda_{1n}*\beta_1*\varphi_{1n} \qquad (13)$$

Let KISS project $\sum_1^{-1}-\sum_0^{-1}$onto the cone of the positive a semi-definite matrix M, so we have

$$\delta(\mathbf{x}_{ij}) = \mathbf{x}_{ij}^T M\mathbf{x}_{ij} \qquad (14)$$

where M is the KISS metric matrix.

C. Adaptive MCE-KISS Metric Learning

Although MCE-KISS has largely improved the accuracy of person reidentification, there is problem dealing with number of small eigen values. In MCE-KISS tha average number of small eigen values are set as a constant. It is not god for a large number of samples . So, instead of value k in (10) , we use maximum likelyhood function to find value of k.

$$L = log(\sum(\sum_1))- log(\sum(\sum_0)) \qquad (15)$$
$$K = max(L) \qquad (16)$$

Algorithm 1 : Adaptive Minimum Classification Error-KISS

Step 1: The initial $\sum_0^{-1}$ and $\sum_1^{-1}$ are calculated .
Step 2: Find the value of number of small eigen values using maximum likelihood function.
Step 3: Smooth technique: To find the estimation errors of small eigenvalues of $\sum_0^{-1}$ and $\sum_1^{-1}$
Step 3: MCE technique: to optimize the parameters of $\sum_0^{-1}$ and $\sum_1^{-1}$
Step 4: The distance metric is calculated by using (6).

IV. EXPERIMENTAL RESULT

In this section, we used a challenging dataset VIPeR [15] to demonstrate the effectiveness of the proposed Adaptive MCE-KISS method. All images from the dataset taken with a size of 128 × 48. In general, it leads to shape distortion , which has limited effect on human visual systems. For each image, we concatenated the extracted LBP descriptor [16] and RGB and HSV color histograms [17] to form the VIPeR feature vector.

In our experiments, all samples of *pts* subjects are selected to form the test set and the remaining *ptr* was used for 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 $\sum_1$ *and* $\sum_1^{-1}$ according to 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 person's 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 (*n* is a small value).

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

A. VIPeR Dataset

The VIPeR dataset was collected by Gray *et al*. [18] and contains 1264 outdoor images obtained from two views of 632 subjects. Intraperson image pairs may contain a viewpoint change of 90◦. Other variations are also considered, such as lighting conditions, shooting locations, and image quality. We set *pts* = 316 to evaluate the matching performance of different algorithms. We repeated the process 10 times, and the average CMC curves were drawn and is shown in fig 2.

B. 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 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. Fig. 1 shows the process of feature extraction. We conducted PCA to obtain a 40-dimensional representation, to suppress the Gaussian noise.

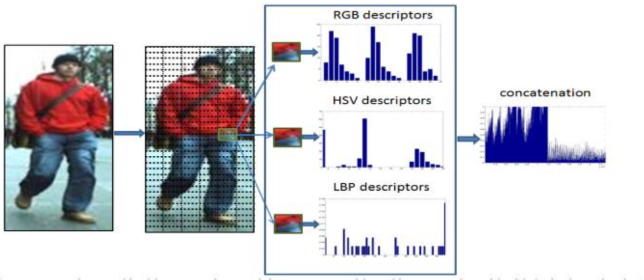

Fig. 1. Process of feature extraction used in this paper. First, each image was partitioned into a regular grid with 8 pixel spacing in the horizontal direction and 4 pixel spacing in the vertical direction. Second, from the grid, the LBP descriptor, HSV histogram, and RGB histogram were extracted from overlapping blocks of size 8 × 8. Third, all the feature descriptors were concatenated together 2].

C. Experiment Results

Fig. 2 shows the comparison of our proposed ADAPTIVE MCE-KISS metric learning with KISS and MCE-KISS on the VIPeR dataset. In each subfigure, the *x*-coordinate is the rank score, and the *y*-coordinate is the matching rate. Only the top 300 ranking positions are shown in the figure. Table 2 shows the comparison of average training time of different metric learning with proposed method. Table 2 reports the performance of all the algorithms within the scope of the first 300 ranks.

Table 1 : Comparison of average training time of existing systems with proposed one.

KISS	MCE-KISS	ADAPTIVE MCE-KISS
0.344582 S	5.181491 S	5.386925 S

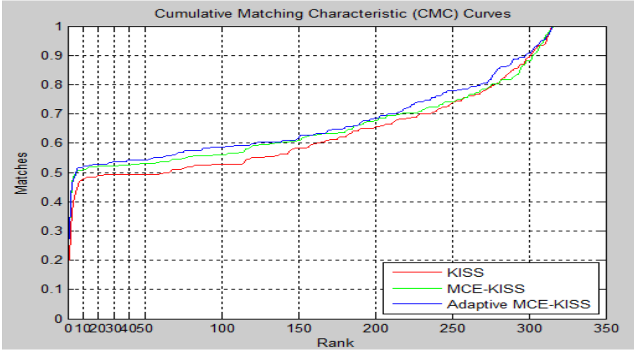

Fig. 2. Performance comparison using CMC curves, the *x*-coordinate is the rank score and the *y*-coordinate is the matching rate. We compare Adaptive MCE-KISS with KISS and MCE-KISS on the VIPeR dataset.

Table 2: Person reidentification top ranked matching rate on the viper dataset

Pts=316				
RANK	**0**	**100**	**200**	**300**
KISS	0.2	0.52	0.65	0.90
MCE-KISS	0.3	0.56	0.68	0.90
ADAPTIVE MCE-KISS	**0.35**	**0.59**	**0.69**	**0.92**

V. Conclusion

Distance metric is very important for the video surveillance applications related to person re-identification. Thus, it is very important to find a proper distance metric learning algorithm to boost the performance of person re-identification. There are many distance metric learning algorithms have been developed, such as information theoretic metric learning (ITML) ,metric learning for large margin nearest neighbor (LMNN), k-nearest neighbour algorithm, but they are not suitable for person re-identification. So, developed KISS metric learning for person re-identification. To obtain robust estimations in small sample size, introduce a new learning algorithm MCE-KISS, in which the number of small eigen values of the covarience metric is a constant, so introduce a new concept Adaptive MCE-KISS which introduces maximul likelihood function. Therefore, MCE-KISS significantly improves KISS for person reidentification.

ACKNOWLEDGMENT

We are greatly indebted to our principal, Dr. JAYAMOHAN J, Dr. V. GOPAKUMAR, Professor, Head of the Department of Computer Science and Engineering, Mrs. SUMITHRA M.D, Assistant Professor, Department of Computer Science and Engineering, LBS Institute of Technology for Women who have been instrumental in keeping my confidence level high and for being supportive in the successful completion of this paper. We would also extend our gratefulness to all the staff members in the Department; also thank all my friends and well-wishers who greatly helped me in my endeavor. Above all, we thank the Almighty God for the support, guidance and blessings bestowed on us, which made it a success.

REFERENCES

1. Vezzani, R., Baltieri, D., Cucchiara, R.: People reidenti_cation in surveillance and forensics: A survey. ACM Computing Surveys (CSUR) 46(2) (2013) 29.
2. Dapeng Tao, Lianwen Jin, Member, IEEE, Yongfei Wang, and Xuelong Li, Fellow, IEEE “Person Reidentification by Minimum Classification Error-Based KISS Metric Learning”, ieee transactions on cybernetics, vol. 45, no. 2, february 2015.
3. H. Hotelling, “Analysis of a complex of statistical variables into principal components,” J. Educ. Psychol., vol. 24, no. 6, pp. 417–441, 1933.
4. 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. 203–223, Jan. 2007.
5. 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. 149–153, Jan. 1987.
6. 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.
7. 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.
8. L. Yang, R. Jin, R. Sukthankar, and Y. Liu, “An efficient algorithm for local distance metric learning,” in Proc. AAAI, 2006, pp. 543–548.
9. B. Prosser, W.-S. Zheng, S. Gong, T. Xiang, and Q. Mary, “Person re-identification by support vector ranking,” in Proc. BMVC, 2010.
10. D. Tao, L. Jin, Y. Wang, Y. Yuan, and X. Li, “Person re-identification by regularized smoothing KISS metric learning,” IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 10, pp. 1675–1685, Oct. 2013.
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. 2288–2295.
12. M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof “Large scale metric learning from equivalence constraints,” in Proc. IEEE Conf. Comput. Vision Pattern Recogn., Jun. 2012, pp. 2288–2295.
13. 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. 149–153, Jan. 1987.
14. 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. 257–265, May 1997.
15. D. Gray, S. Brennan, and H. Tao, “Evaluating appearance models for recognition, reacquisition, and tracking,” in Proc. 10th PETS, 2007.
16. 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.
17. Shamik Sural, Gang Qian and Sakti Pramanik, “segmentation and histogram generation using the hsv color space for image retrieval”.
18. D. Gray, S. Brennan, and H. Tao, “Evaluating appearance models for recognition, reacquisition, and tracking,” in Proc. 10th PETS, 2007.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)