Person Re-Identification using Reduced Dictionary Sparse Representation Based Classifier

M.K.Vidhyalakshmi, E. Poovammal

***Abstract:** Video surveillance has become a necessary tool for monitoring a public places. Automation of the video surveillance is the current need, as it a tough task for the humans to monitor the surveillance video continuously. Also searching a person from a bulk of videos is not an easy job. Person re-identification is a step taken towards to make the video surveillance an automated one. Person re-Identification is a task of matching the identity of a person captured by different cameras in the network at different places and times. The cameras used for surveillance are located at a much higher position than the person so that the conventional method of face recognition is not used for identification of the person. The images of the same person may differ based on the qualities of different cameras (resolution changes), or due to different lighting conditions (variation in illumination) or due to posture changes. Recently sparse based classifier is used in person re-identification and is effective in handling in illumination variation and occlusion. But sparse based decomposition method is not computational efficient. This leads to the development of the person re-identification method based on Reduced Dictionary Sparse representation based Classifier (RDSC). This is done with 2 steps: (i) Similarity score for reduction of the gallery; (ii) Sparse basis expansion of targets in terms of reduced gallery. The proposed method is both computational efficient and creates better outcomes.*

***Keywords** — Video surveillance systems, Person re-identification, Sparse Classifier, Reduced dictionary.*

I. INTRODUCTION

In video surveillance systems, numerous cameras are used to observe a place. The person re-identification process is recognizing the particular person from one camera view with the images captured by other non-overlapping cameras. This task is highly challenging since the images of the same person captured at various places and times vary notably. When re-identification is performed manually, it requires a laborious effort but still remains inaccurate. With the increase in the use of video surveillance in public places, the interest in automated re-identification is growing. In the conventional technique, face recognition is utilized for identifying person, but this method is not possible as it is very difficult to get the details required for extraction of face features. Alternatively, other visual features like clothing, objects associated can be used for re-identification. But, still the visual appearance features remain very weak due to a number of reasons. Firstly, the cameras used for surveillance are fixed at a distance and the environment to be captured is highly uncontrollable. Secondly, when two or more different cameras are used, it is hard to fix the transit time between the cameras, as it varies from person to person. Thirdly, it is difficult to identify the person in terms of features extracted from clothing because of the possibilities of same color clothes wear by many people. In addition, some other significant reasons such as, the variations in lighting conditions, view angles, occlusion, and background clutter. In the person re-identification task, image of the person who needs to be searched across the camera network is called the probe image and the collection of images from which image has to be searched is called the gallery images. Person re-identification is executed using two different strategies such as single and multiple-shot approaches. Based on the number of images known in priori in gallery and probe sets, there are three modalities such as the single-versus- single (SvsS), the multi-versus-single (MvsS), the multi-versus-multi (MvsM). The SvsS modality has the single instance for individual person in the gallery and the single instance for entire person in the probe set. The MvsS modality has single set of many instance for individual person in the gallery, and a single instance of the entire person in the probe set. The MvsM modality has a collection of multiple instances for entire person in the gallery, and a collection of many images of each person in the probe set.

II. RELATED WORKS

A. Submission of the paper

In recent years the field of person re-identification has gained attraction among the researchers working with computer vision. This section briefly describes the survey of the works in the field of person re-identification. The works on person re-identification can be broadly categorized as 2 groups [26].The first group is based on the features and second group is based on learning.

Many research works were done on effective extraction of visual features from the person's image and its representation. The commonly used features are colour histograms of various colour spaces like RGB, HSV [1]-[5] and texture feature [6] – [11]. O. Hamdoun, et al. [12] used SIFT [13] feature and X. Liu,et al. [14] used SURF features for extracting the texture features. Bak, et.al [15] used Haar-like features [16], for person re-identification which are effectively used for face detection.

Revised Manuscript Received on November 22, 2019.

***** Correspondence Author

Mrs.M.K.Vidhyalakshmi, Pursuing Part Time Research In Srm Institute Of Science And Technology, Kattangulathur

Dr. E. Poovammal, Professor Department Of Computer Science And Engineering At Srm Institute Of Science And Technology

The other group of researchers concentrated in building models which are based on metric learning. The learning based methods aims to minimize intraclass variation and maximize the interclass variation. Kostinger , et al. [17] proposed Keep it simple and straight forward (KISS) metric learning which is a effective technique for person re-identification. But this method is not very efficient when the dataset is small as estimating the co variance matrix is performed using maximum likelihood estimation. To overcome this problem Dapeng , et al. [10] proposed a method namely MCE-KISS (Minimum Classification Error Based Keep It Simple and Straight Forward) which performs better than [17]. Zheng , et al. [18] proposed RDC (Relative Distance Comparison) which finds relative distance among the images. The matching image pairs will have relatively small distance when compared to wrong match pair of images. Liu , et. al [19] proposed a new metric learning algorithm which changes the adjustable margin if the scale changes.

Sparse based representation is gaining popularity because of its successful performance in computer vision task like face recognition, image restoration and object tracking. Masi, et.al [20] introduced a sparse based re-identification technique which ranks all the targets using reweighting approach. Mirmahboub, et.al. [21] introduced a novel person re-identification technique based on sparse representation and manifold learning. This technique extracts the discriminant feature from the images of a person captured in different shots. Kim, et.al. [22] introduced sparse representation using colour name as descriptor. He used colour histogram matching for gallery reduction and reduced amount of calculations significantly. Li, et.al. [29] used sparse method for multishot person reidentification .The work explores the knowledge about similar probe images having same identity and are coded in the same dictionary.

III. PROPOSED PERSON RE-IDENTIFICATION METHOD

The proposed Reduced Dictionary Sparse representation based Classifier (RDSC) based person reidentification algorithm consists of two steps: (i) Similarity score for reduction of the gallery; (ii) Sparse basis expansion of targets in terms of reduced gallery. Fig. 1 illustrates the flow chart of proposed person re-identification algorithm.

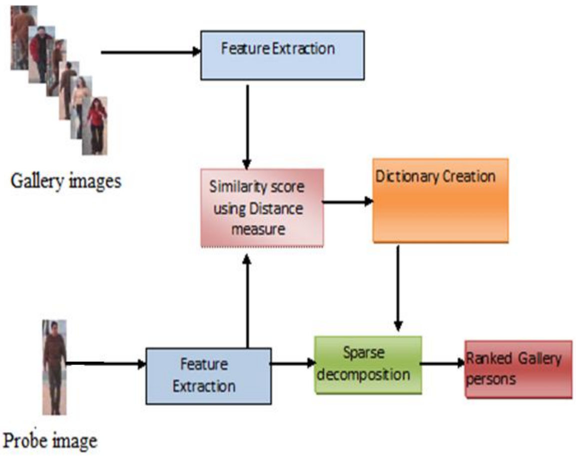

Figure 1: The proposed re-identification architecture

A. Feature extraction

We have used a descriptor as specified in [20]. The input target image is measured to a size of 64 x 128 pixels, and in order to develop a spatial pyramid the image is separated into overlapping horizontal stripes of 16 pixels in height. Hue-Saturation (HS) and RGB histograms are extracted from each stripe. Then concatenate a Histogram of Oriented Gradient (HOG) descriptor. The HS histogram contains 8 x 8 bins and RGB has 4 x 4 x 4 bins. The HS and RGB histograms are gauged for the 15 levels of the pyramid. Totally it has 1,920 Color histogram bins. HOG has several blocks where each block has 2 X 2 grid cells of 8 X 8 pixels. For every cell the gradient histogram is computed over four angular bins for each HOG block. The final descriptor is of size 2,960.

B. Reduced Dictionary Sparse representation based Classifier (RDSC)

The ultimate purpose of sparse modelling is to construct efficiently a dictionary which is learned from the training data. Let there are be adequate number of training samples, A=[$a_{i,1}; a_{i,2}; \ldots . a_{i,ni}$] belonging to some class i. If Y is the test sample belonging to the same class, it should be lying in the linear span of the given training samples, the solution of which may be exactly as given by (1).

$$y = a_{i,1}x_{i,1} + a_{i,2}x_{i,2} + \ldots\ldots + a_{i,,ni}; x_{i,ni} \quad (1)$$

$$= \sum_{i=1}^{ni} A_{i,j}\, x_{i,j} \quad (2)$$

$$= A_i x_i \quad (3)$$

In Eq. (1), (2), and (3), A_i represents the matrix of basis vectors for class i and x_i represent the vector of reconstruction coefficients related to class i. The sparse vector x* can be estimated by l_1 regularised least squares problem [23] , [24] as follows:

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$x^* = \arg\min_{x}\{\|y - Ax\|_2^2 + \lambda\|x\|_1\} \quad (4)$$

where x* represents the modified sparse vector, λ is a regularization factor, $||.||_2$ represents the l_2 norm operator, and $||.||_1$ represents l_1 norm operator. In (4), x* ensures the optimal reconstruction as the first term, since the error between input signal **y** and recovered version **Ax** is minimized. Also, x* satisfies the sparsity condition as the second term is l_1 -norm. In terms of person re-identification, the developed technique searches the gallery using a reconstructed probe from the estimated co-efficient and dictionary A of gallery images. A is the dictionary which consist of all gallery images. When A is large in case of real time databases, this optimization problem becomes computationally intensive. It is highly desirable to minimise the size of Dictionary A. We propose a Reduced Dictionary sparse representation based classifier (RDSC) in which the dictionary is reduced using similarity score based on Euclidean distance measurement.

This technique initially reduces the size of a gallery by checking the features of a probe person and gallery people based on Euclidean distance. Then in the next step, it similar to the probe person with reduced gallery based on sparse representation method. The major advantages of the proposed algorithm are that it can control the re-identification error with the help of a reduced gallery as well as the computation complexity. In addition, since sparse representation is used, it can be robust for illumination and occlusion.

The normalised reconstruction error is calculated by the following formula:

$$e_i = \frac{||y - Px_i||_2}{||y||_2} \quad (5)$$

For ranking the gallery images the value of the normalised reconstruction error is observed with $e_i < 1$. When e_i value is the smallest, the gallery image is given first rank. The gallery image with first rank most likely represents the correct image. The reconstruction error is utilized for ranking the other gallery images.

Algorithm 1 shows the procedure of the proposed method Reduced Dictionary sparse representation based classifier (RDSC):

Algorithm 1: Reduced Dictionary sparse representation based classifier (RDSC)

1. ***Input: Given training data of gallery images*** **A ∈ R ^{mxn} belonging to i classes, test sample of probe image y ∈ R ^{m}**
2. ***Compute the reduced dictionary P*** **∈ R ^{dxm} using Euclidean distance measurement.**
3. **Use the reduced dictionary for sparse representation:**

$$x^* = \arg\min_{x}\{\|y - Px\|_2^2 + \lambda\|x\|_1\}$$

4. **Compute the residuals:**

$$e_i = \frac{||y - Px_i||_2}{||y||_2} \quad \textit{for } i=1,2....c;$$

5. **Output: label (y) = arg min e_i (y)**

IV. EXPERIMENTAL RESULTS

4.1. VIPeR Dataset:

We used, the viewport-invariant pedestrian recognition (VIPeR) dataset [25] for our experiments. It contains 632 person images which are captured with the help of two non-overlapping cameras. Here, viewpoint changes can be observed up to 180 degrees from image pairs and illumination changes that result in large intra-class variations. In the datasets, two samples of each person is maintained (one from each view); this is utilized for SvsS re-identification process. The images are normalised to 128 x 48 pixels. The total images are 1264 which are varied in lighting conditions, viewpoints and back ground clutter, thus making it a challenging dataset. The total number of images is randomly divided into two groups for training and testing. The set 'cam-a' contains testing images (Probe images) and the set 'cam-b' contains training images (Gallery images).

4.2 State-of-the-art Comparison

We compared our results with the other methods and obtained better results.Fig.2 exhibits the results obtained from our experiment. It shows the query probe image on the left and the corresponding top 10 ranked gallery images on the right .From the Fig.2 we could see the correct matches for the probe image in the gallery images were found in the first few ranks. Fig.3 provides the results of our proposed algorithm RDSC under different challenging conditions.Fig.3. (a) Provides the results our proposed algorithm under various illumination conditions. Fig .3. (b) Provides the results for case of occlusion. Our algorithm finds the correct match within top 10 ranks.

Figure 2: The left column displays the probe image and to the right are the 10 top-ranked candidates. Highlighted box represents the correct match.

(a)

(b)

Figure 3: Results for proposed method for 10 top-ranked candidates under challenging conditions. (a)Illumination changes (b) Occlusion

The proposed re-identification methods are compared with existing person re-identification methods SDALF [26], ELF [25], ISR [20], PRDC [27], and PRSVM [28]. From the experimental comparisons by cumulative match curve Fig [4], we infer that the proposed method provides better results than other methods. Our method outperforms state of art methods within 50 ranks. Table 1 enables the comparison of the ranking methods namely SDALF, ELF, PRSVM, PRDC, ISR.

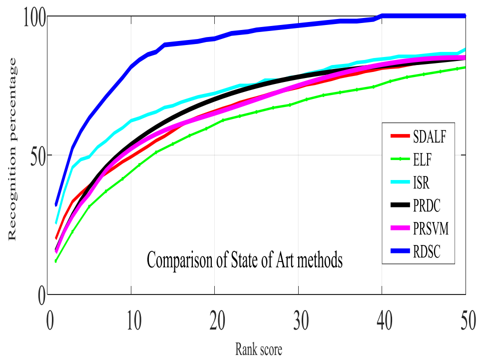

Figure 4: Person Reidentification performance comparisons on VIPeR dataset.

Table 1: Comparison of proposed method with state of art methods.

Rank	1	5	10	20	50
SDALF	19.87%	38.42%	49.37%	65.73%	85%
ELF	12%	31%	44%	60.2%	81.3%
ISR	25.32%	49.37%	62.34%	72.15%	87.97%
PRDC	15%	38.42%	53.8%	70.09%	85%
PRSVM	15.6%	36%	52.28%	65%	85%
RDSC	31.96%	63.29%	81.65%	91.77%	100%

4.3 Comparison with different dictionary sizes:

We also performed experiment to compare the performance of proposed method RDSC using dictionary sizes 25%, 50% and 10% of total number of features. The dataset VIPER contains 632 training images. Other sparse based method [20] used all training image features for dictionary but our proposed method obtained better result with dictionary size of 316 features which is 50% of total training features .These results are shown in Fig. [5].

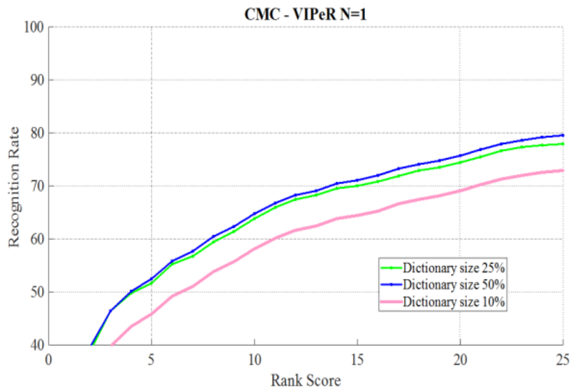

Figure 5: Performance comparisons on VIPeR dataset using proposed method with different dictionary size

IV. CONCLUSIONS

In this paper, we introduced a new method for person re-identification known as Reduced Dictionary Sparse representation based classifier (RDSC). The proposed method provides good quality performance, as robust descriptors are used as a dictionary, but with reduced computation cost. Also, our method finds the correct match for the probe image in gallery images within first few ranks even for partially occluded and illumination variations. The dictionary is efficiently constructed using similarity score based on Euclidean distance. Therefore, the proposed "Person Re-identification Based on Reduced Dictionary Sparse Classifier" method can be applied in a large scale person re-identification database to search for a specific person.

REFERENCES

1. M. Swain and D. Ballard, "Indexing via Color Histograms," Proc. IEEE Int'l Conf. Computer Vision, pp. 390-393, 1990
2. B. Funt and G. Finlayson, "Color Constant Color Indexing," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 5, pp. 522-529, May 1995.
3. T. Gevers and A. Smeulders, "Color-Based Object Recognition," J. Pattern Recognition Soc., vol. 32, no. 3, pp. 453-464, 1999.
4. U. Park, A. Jain, I. Kitahara, K. Kogure, and N. Hagita, "Vise: Visual Search Engine Using Multiple Networked Cameras," Proc. Int'l Conf. Pattern Recognition, pp. 1204-1207, 2006.
5. Kviatkovsky, Igor, Amit Adam, and Ehud Rivlin. "Color invariants for person re-identification." *IEEE Transactions on pattern analysis and machine intelligence* 35, no. 7 (2012): 1622-1634.

6. B. Prosser, W.-S.Zheng, S. Gong, and T. Xiang, “Person re-identification by support vector ranking,” in Proc. BMVC, Aberystwyth, U.K., Aug. 2010, pp. 1–11.
7. W.-S. Zheng, S. Gong, and T. Xiang, “Re-identification by relative distance comparison,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 3, pp. 653–668, Mar. 2013.
8. C. Liu, S. Gong, C. C. Loy, and X. Lin, “Person re-identification: What features are important?” in Proc. ECCV, Firenze, Italy, Oct. 2013, pp. 391–401.
9. M. Hirzer, P. M. Roth, and H. Bischof, “Person re-identification by efficient impostor-based metric learning,” in Proc. IEEE AVSS, Sep. 2012, pp. 203–208.
10. Dapeng Tao, LianwenJin, YongfeiWang , and Xuelong Li, “Person Re-identification by Minimum Classification Error-Based KISS Metric Learning”, IEEE Transactions On Cybernetics, Vol. 45, No. 2, February 2015, pp. 242 -252.
11. Yimin Wang, RuiminHu, Chao Liang ,Chunjie Zhang, and Qingming Leng , “Camera Compensation Using a Feature Projection Matrix for Person Re-identification”, IEEE Transactions On Circuits And Systems For Video Technology, Vol. 24, No. 8, August 2014, pp.. 1350-1361.
12. O. Hamdoun, F. Moutarde, B. Stanciulescu, and B. Steux, “Person reidentification in multi-camera system by signature based on interest point descriptors collected on short video sequences,” in Proc. ICDSC, Stanford, CA, USA, 2008, pp. 1–6.
13. D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.
14. X. Liu et al., “Attribute-restricted latent topic model for person reidentification,” Pattern Recognit., vol. 45, no. 12, pp. 4204–4213, 2012.
15. S. Bak, E. Corvee, F. Bremond, and M. Thonnat, “Person reidentification using Haar-based and DCD-based signature,” in Proc. AVSS, Boston, MA, USA, 2010, pp. 1–8.
16. R. Lienhart and J. Maydt, “An extended set of Haar-like features for rapid object detection,” in Proc. ICIP, 2002, pp. 900–903.
17. Kostinger, M., Hirzer, M., Wohlhart, P., Roth, P. M., and Bischof, H., “Large Scale Metric Learning from Equivalence Constraints,” *Proceedings on IEEE Conference on Computer Vision and Pattern Recognition*, Providence, Rhode Island, USA, pp. 2288–2295, June.2012.
18. Zheng, Wei-Shi, Shaogang Gong, and Tao Xiang. “Reidentification by relative distance comparison,” *IEEE transactions on pattern analysis and machine intelligence* 35, no. 3 (2012): 653-668.
19. Hao Liu, Meibin Qi, and Jianguo Jiang , “ Kernelized Relaxed Margin Components Analysis for Person Re-identification ”, IEEE Signal Processing Letters, Vol. 22, No. 7, July 2015, pp . 910 -914.
20. Lisanti, G., Masi, I., Bagdanov, A. D., and Del Bimbo, A., “Person Re-Identification by Iterative Re-Weighted Sparse Ranking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp.1629-1642, Aug. 2015.
21. Mirmahboub, Behzad, Hamed Kiani, Amran Bhuiyan, Alessandro Perina, Baochang Zhang, Alessio Del Bue, and Vittorio Murino. "Person re-identification using sparse representation with manifold constraints." In *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 774-778. IEEE, 2016.
22. Kim, Miri, Jaehoon Jung, Hyuncheol Kim, and Joonki Paik. "Person re-identification using color name descriptor-based sparse representation." In *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 1-4. IEEE, 2017.
23. Tibshirani, Robert. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* (1996): 267-288.
24. Kim, S. J., Koh, K., Lustig, M., Boyd, S., & Gorinevsky, D. , “An Interior-Point Method for Large-Scale l_1-Regularized Least Squares. *IEEE Journal of selected topics in signal processing*, *1*(4), pp. 606-617, Dec 2007.
25. Gray, D., and Tao, H.,"Viewpoint Invariant Pedestrian Recognition with An Ensemble of Localized Features," *European Conference on Computer Vision,* Marseille, France, pp.262-275, Oct. 2008.
26. Farenzena, M., Bazzani, L., Perina, A., Murino, V., & Cristani, M.,“Person Re-identification By Symmetry-Driven Accumulation of Local Features,” *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*,California, USA, pp. 2360–2367, June. 2010.
27. Zheng, W. S., Gong, S., and Xiang, T., “Re-Identification by Relative Distance Comparison,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.35, no. 3, pp.653-668, March. 2013.
28. Prosser, B., Zheng, W. S., Gong, S., Xiang, T., and Mary, Q., “Person Re-Identification By Support Vector Ranking,” in *Proceedings of British Machine Vision Conference*, Aberystwyth, Wales, vol. 2, no. 5, pp. 1–11, Sep. 2010.
29. Li, Xiaobao, Lijun Guo, Rong Zhang, and Yuanyuan Zhang. "Multishot person reidentification using joint group sparse representation." *Journal of Electronic Imaging* 27, no. 6 (2018): 063012.

AUTHORS PROFILE

Mrs.M.K.Vidhyalakshmi has completed her B.E degree in Electronics and communication Engineering from Bharathidasan University in the year 2002and M.E degree in Communication Sytems from Trichy Anna University in the year 2011. She has more than nine years of teaching experience in the field of Engineering. Currently she is pursuing part time research in SRM Institute of Science and Technology, Kattangulathur .She has published more than six papers in International journals .Her area of interest is image processing and computer vision.She is the life member of professional bodies like ISTE, ISRO.

Dr. E. Poovammal is the Professor in the Department of Computer Science and Engineering at SRM Institute of Science and Technology She joined SRM in the year 1996. She obtained her B.E. Degree in Electrical and Electronics Engineering from Madurai Kamaraj University in the year 1990, M.E degree in Computer Science and Engineering from Madras University in the year 2002 and Ph.D. degree in Computer Science and Engineering from SRM University in 2011. Her research interests include Association Rule Mining, Web Mining, Text Mining, Image and Video mining, Privacy and Security in the Cloud as well as Big Data Analytics. She currently guides six research scholars. Since last 4 years she is certified as Adjunct Faculty for the two courses, by Institute of software Research, Carnegie Mellon University, Pittsburgh, USA. She has published more than 30 referred journals and presented various international and national conferences. She is the Recipient of “Best Academic Dean award”, by Association of Scientists, Developers and Faculties (ASDF), 2015 and Recipient of “Women Engineer award”, by IET-CLN, 2013. She is the University III-Rank holder, Madras University, M.E (CSE) 2002. She has received the National Merit Scholarship from 1984 to 1990.She is the Fellow IE (I). She is the life member of ISTE and Indian Science Congress. She is also the member of other professional bodies IET, IEEE, ACM and CSI.