International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 3, Issue 3, March 2016

A Review on a Person Cross Domain Re Identification Based Adaptive Ranking Support Vector Machines (AdaRSVMs)

[1] V .Hemanth Kumar, [2] Y.Penchalaiah, [3] Pushpalatha
[1] M.Tech Student [DSCE], [2][3] Assistant Professor
Annamacharya Institute of Technology & Sciences (AITS)
[1] hk4390@gmail.com, [2] penchu.yadl@gmail.com [3] pushpalatha_nainaru@rediffmail.com

Abstract: **An adaptive ranking support vector machines (AdaRSVMs) method is used for re identification under target domain cameras without person labels. It addresses a new person re identification problem without label information of persons under non overlapping target cameras. Given the matched (positive) and unmatched (negative) image pairs from source domain cameras, as well as unmatched (negative) and unlabeled image pairs from target domain cameras, To overcome the problems introduced due to the absence of matched (positive) image pairs in the target domain, we relax the discriminative constraint to a necessary condition only relying on the positive mean in the target domain. To estimate the target positive mean, we make use of all the available data from source and target domains as well as constraints in person re identification. Inspired by adaptive learning methods, a new discriminative model with high confidence in target positive mean and low confidence in v target negative image pairs is developed by refining the distance model learnt from the source domain. Experimental results show that the proposed AdaRSVM outperforms existing supervised or unsupervised, learning or non-learning re identification methods without using label information in target cameras. Moreover, our method achieves better re identification performance than existing domain adaptation methods derived under equal conditional probability assumption.**

***Index Terms*— Person re- identification, domain adaptation, Target positive mean, adaptive learning, ranking SVMs.**

I. INTRODUCTION

Person re-identification, or inter-camera entity association, is the task of recognizing an individual across heterogeneous non-overlapping camera views against a background of similar persons. When an individual disappears from one view they need be differentiated from numerous possible alternative people and re-identified in another view, potentially under a different and unknown view angle, pose, lighting conditions and clutter or occlusion (see figure 1 for examples). This is critical to a variety of safety, security and efficiency tasks which require long-term maintenance of consistent identity across space and time. In particular, it is a fundamental capability for long-term tracking across multiple disjoint camera views [11]. Relying on manual re-identification in large camera networks is prohibitively costly and error prone. For these reasons, there has recently been extensive work in the computer vision community on automated re-identification [7, 21, 10] . This is very challenging because of extreme intra-class (person identity) variability in appearance across views with different lighting, pose and occlusion; and limited inter-class variability in appearance among many similarly clothed pedestrians. Existing approaches can be broadly broken down into two complementary categories: those which focus on developing effective feature representations [7, 5], and those which focus on developing learning methods to better discriminate identity using a given representation [21,10]. Feature design approaches [7, 5] suffer from the problem that it is extremely challenging if not impossible to design features that are discriminative enough to distinguish people reliably; while simultaneously being invariant to all the covariates which occur in practice such as, motion blur, view angle and pose change, lighting and occlusion. In contrast, learning approaches [10] try to improve on a given set of features, and focus on discriminative training of models to maximize re-identification performance, for example distance metric learning [4] and support vector machines (svm) [1]. Recently, discriminative approaches have significantly improved state of the art benchmark performance [10, 1] treating re-identification as a binary (same versus different person) rather than multi-class (person identity) problem.

A central limitation of existing discriminative learning approaches is that they are more suited to closed-world benchmark problems than realistic open-world scenarios. In particular they require many pairs of person images annotated by same/different, for each camera pair between which the system is required to operate. This is reasonable for training/testing splits on benchmark dataset ts that are already exhaustively annotated by person identity. However it is highly impractical for real-world use, where there may be very many pairs of cameras in a given network, each requiring exhaustive annotation – making this "calibration" requirement of such a system impossible or prohibitively expensive. Ideally, we would like to deploy a re-identification system between a pair of cameras with minimal calibration/training annotation. What a system learns from annotations of one camera pair should be exploited by another pair without requiring exhaustive annotation in the new pair.

This is an issue in transfer learning [19, 6, 12]. Transfer learning is already important for many classical vision problems such as object recognition with multiple classes or domains. However it is critically important for re-identification because the number of domains (camera pairs) is quadratic in the number of cameras. Therefore obtaining exhaustive training data for each domain is even more impractical than for conventional vision applications, and transfer learning becomes critical. Nevertheless, no prior reidentification studies have addressed this issue, relying solely on benchmark datasets with sufficient annotated data in each camera pair of interest. In this paper we relax the practically unrealistic assumption of exhaustive training data within each domain by generalizing recent ideas in learning re-identification [1] and SVM transfer learning [12]. Specifically, we consider re-identification based on binary relation learning [1, 13], and show how to generalize this approach to to achieve effective cross-domain learning by combining nonlinear decision boundaries from source domains to create a more accurate target domain re-id classifier. In this way we are able to improve on within-domain learning both for sparse and even nonsparse training data volumes. Moreover we show how to achieve this while systematically avoiding negative transfer, even when there are multiple and irrelevant source domains.

II. RELATED WORK

O. Javed, Z. Rasheed, K. Shafique,[1] and M. Shah, "Tracking across multiple cameras with disjoint views," in Proc. 9th IEEE Int. Conventional tracking approaches assume proximity in space, time and appearance of objects in successive observations. However, observations of objects are often widely separated in time and space when viewed from multiple non-overlapping cameras. To address this problem, we present a novel approach for establishing object correspondence across non-overlapping cameras. Our multi camera tracking algorithm exploits the redundance in paths that people and cars tend to follow, e.g. roads, walk-ways or corridors, by using motion trends and appearance of objects, to establish correspondence. Our system does not require any inter-camera calibration, instead the system learns the camera topology and path probabilities of objects using Parzen windows, during a training phase. Once the training is complete, correspondences are assigned using the maximum a posteriori (MAP) estimation framework. The learned parameters are updated with changing trajectory patterns. Experiments with real world videos are reported, which validate the proposed approach.

N. Gheissari, t. B. Sebastian, [2] and r. Hartley, "person reidentification using spatiotemporal appearance," in proc. Ieee conf. Comput. Vis. Pattern recognit., in many surveillance applications it is desirable to determine if a given individual has been previously observed over a network of cameras. This is the person reidentification problem. 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, and (ii) generate signatures that are invariant to variations in illumination, pose, and the dynamic appearance of clothing. A novel spatiotemporal 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, and (ii) an interest point operator approach that nominates a large number of potential correspondences which are evaluated using a region growing scheme. Finally, the approaches are evaluated on a 44 person database across 3 disparate views.

C. Madden, e. D. Cheng, and [3] m. Piccardi, "tracking people across disjoint camera views by an illumination-tolerant appearance representation," mach. Vis. Appl., vol. 18, nos. Tracking single individuals as they move across disjoint camera views is a challenging task since their appearance may vary significantly between views. Major changes in appearance are due to different and varying illumination conditions and the deformable geometry of people. These effects are hard to estimate and take into account in real-life applications. Thus, in this paper we propose an illumination-tolerant appearance representation, which is capable of coping with the typical illumination changes occurring in surveillance scenarios. The appearance representation is based on an online *k*-means colour clustering algorithm, a data-adaptive intensity transformation and the incremental use of frames. A similarity measurement is also introduced to compare the appearance representations of any two arbitrary individuals. Post-matching integration of the matching decision along the individuals‘tracks is performed in order to improve reliability and robustness of matching. Once matching is provided for any two views of a single individual, its tracking across disjoint cameras derives straightforwardly. Experimental results presented in this paper from a real surveillance camera network show the effectiveness of the proposed method.

S. B¸ak, e. Corvée, f. Brémond, [4] and m. Thonnat, "boosted human re-identification using riemannian manifolds," image vis. Comput., this paper presents an appearance-based model to address the human re-identification problem. Human re-identification is an important and still unsolved task in computer vision. In many systems there is a requirement to identify individuals or determine whether a given individual has already appeared over a network of cameras. The human appearance obtained in one camera is usually different from the ones obtained in another camera.
In order to re-identify people a human signature should handle difference in illumination, pose and camera parameters. The paper focuses on a new appearance model based on mean riemannian covariance (mrc) patches extracted from tracks of a particular individual. A new similarity measure using riemannian manifold theory is also proposed to distinguish sets of patches belonging to a specific individual. We investigate the significance of mrc patches based on their reliability extracted during tracking and their discriminative power obtained by a boosting scheme. Our method is evaluated and compared with the state of the art using benchmark video sequences from the ethz and the i-lids datasets. Re-identification performance is presented using a cumulative matching characteristic (cmc) curve. We demonstrate that the proposed approach outperforms state of the art methods. Finally, the results of our approach are shown on two further and more pertinent datasets.

S. B¸ak, e. Corvee, f. Bremond, [5] and m. Thonnat, "person re-identification using haar-based and dcd-based signature," in proc. 7th ieee int. Conf. Adv. In many surveillance systems there is a requirement to determine whether a given person of interest has already been observed over a network of cameras. This paper presents two approaches for this person re-identification problem. In general the human appearance obtained in one camera is usually different from the ones obtained in another camera. In order to re-identify people the human signature should handle difference in illumination, pose and camera parameters. Our appearance models are based on hoar-like features and dominant color descriptors. The adaboost scheme is applied to both descriptors to achieve the most invariant and discriminative signature. The methods are evaluated using benchmark video sequences with different camera views where people are automatically detected using histograms of oriented gradients (hog). The re-identification performance is presented using the cumulative matching characteristic (cmc) curve.

III.EXISTING WORK

TABLE I

Compare rank-n identification rates (%) with other published single-shot results on viper the gallery size is 316.

METHODS	TOP 1	TOP 10	TOP 25	TOP 50
OURS	**29.6**	**69.3**	**88.7**	**96.8**
KISSME [19]	**19.6**	**62.2**	**80.7**	**91.8**
PS[2]	**21.8**	**57.2**	**76.0**	**88.1**
SDALF[32]	**19.9**	**49.4**	**70.5**	**84.8**
PRDC[32]	**15.7**	**53.9**	**76**	**87**
LDML[12][19]	**10.4**	**31.3**	**44.6**	**60.4**
LMNN-R[4]	**23.7**	**68**	**84**	**93**
MCC[8][32]	**15.2**	**57.6**	**80**	**91**
PCCAX^2_{RBF} [17]	**19.3**	**64.9**	**83**	**96**

TABLE II

Compare rank-n identification rates (%) with other published single-shot results on viper the gallery size is 512.

METHODS	TOP 1	TOP 10	TOP 25	TOP 50
OURS	**12.90**	**30.30**	**42.73**	**58.02**
PRDC[32]	**9.12**	**24.19**	**34.40**	**48.55**
PCCAX^2_{RBF} [17]	**9.27**	**24.89**	**37.43**	**52.89**
MCC[8][32]	**5.00**	**16.32**	**25.92**	**39.64**

All the random partitions described below repeat for 100 times. Two protocols on viper were used in the past: randomly splitting the whole dataset into 316 persons for training and the remaining 316 for test; and randomly splitting into 100 persons for training and 532 for test. We evaluate both. Table 1 and 2 compare with results previously published on viper with the same protocol. For cuhk02, we choose view pair p1 for evaluation. It has 971 persons, which are split to 485 for training and 486 for test. Each person has two images in each view. They are also randomly selected. Cavirr has a small number of persons, so we did not split the persons. It is also to be consistent with existing protocol. If a person has images in both camera views, we randomly select two pairs of images in different views for training. One query image and one gallery image are randomly selected from the remaining images per person. Table 3 compares with results previously published on caviar. Experimental results show that our method significantly outperforms other learning approaches and achieves the best results on the two public datasets. Cca does not work very well since it assumes the feature transforms to be uni-modal while the three datasets are much more complicated. Kernel cca alleviates the problem, but its performance is still not good as ours after tuning the kernel.

IV. PERSON RE-IDENTIFICATION

In order to ensure that feature representation of the person image is less sensitive to large inter-camera variations, many existing re-identification methods focus on extracting robust features. Popular ones include sift, color distribution, space-time methods and pictorial structures. Besides feature extraction, discriminative distance learning methods are proposed to further improve the re-identification performance. In, person re-identification was formulated as a ranking problem and the rank svm model is learnt by assigning higher confidence to the positive image pairs and vice versa. Denote xi as the feature vector.

Kuo *et al.*proposed an online-learnt appearance affinity model to decrease the required number of labeled samples under some specific assumptions. On the other hand, an adaptive feature weighting method was proposed in [38] under the observation that the universal model may not be good for all individuals. Different from traditional per-individual identification scheme, zheng *et al.* [39] addressed a watch list (set) based verification problem and proposed to transfer the information from non-target person data to mine the discriminative information for the target people in the watch list.

V. DOMAIN ADAPTATION

The main objective of domain adaptation approach is to adapt the classification model learnt from the source domain to target domain without serious deterioration of recognition performance. The target domain refers to data from the target task usually without or with only a small amount of labeled training data, while there is plenty of labeled training data in the source domain. In the last decade, many algorithms have been proposed to solve the joint distribution mismatch problem, i.e. Prs (y, z) = pr$t$ (y, z). For unsupervised domain adaptation, the instance re-weighting or covariate shift approach learns the target classification model by re-weighting the labeled samples in the source domain to minimize the approximated empirical classification error in the target domain. To estimate the sample weights calculated by pr$s$ (z) dividing pr$t$ (z), many density ratio estimation methods have been proposed. Besides instance re-weighting, the feature representation methods construct feature vectors to reduce the difference between features in the source and target domains. Blitzer *et al.* *P*roposed a structural correspondence learning algorithm by selecting pivot features for natural language processing, while other method try to learn a mapping , s.t. Pr$s((z)) \approx$ pr$t((z))$. Without label information in the target domain, these methods assume that the conditional probabilities are equal to each other in the source and target domains. And it was shown in [6] that the empirical classification error can be very small under this assumption. However, this assumption may not be valid, so that the recognition performance may deteriorate. For supervised domain adaptation with target labeled data, existing methods learn an informative prior using the source domain data and estimate the target model based on such prior. Based on the assumption that the recognition tasks in the source and target domains are related, multi-task learning methods [2], [3] can be employed to discover the task relationship and learn the classification models in the source and target domains simultaneously. Unlike supervised domain adaptation techniques, unlabeled data in the target domain are considered together with the labeled data to learn the target classification model for better performance in [9]–[5]. However, labeling person images for each camera is expensive, especially in large-scale camera networks applications. Thus, existing supervised or semi-supervised domain adaptation algorithms cannot be employed directly.

VI. CONCLUSIONS

In this paper, propose a novel adaptive ranking support vector machines (adarsvm) method to deal with the problem that label information of persons is not available under target cameras. Without positive image pairs generated by the label information of persons, we relax the discriminative constraint to a necessary condition, which only relies on the mean of positive pairs. In order to estimate the positive mean in the target domain, we make use of the labeled data from the source domain, the negative and unlabeled data from the target domain. With two estimations of the target positive mean, the optimal combination is determined by the training data. The target distance model is trained by adapting the source domain distance model to target domain.

REFERENCES

[1] O. Javed, Z. Rasheed, K. Shafique, and M. Shah, "Tracking across multiple cameras with disjoint views," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, vol. 2. Oct. 2003, pp. 952–957.

[2] N. Gheissari, T. B. Sebastian, and R. Hartley, "Person reidentification using spatiotemporal appearance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2006, pp. 1528–1535.

[3] C. Madden, E. D. Cheng, and M. Piccardi, "Tracking people across disjoint camera views by an illumination-tolerant appearance representation," *Mach. Vis. Appl.*, vol. 18, nos. 3–4, pp. 233–247, 2007.

[4] S. B¸ak, E. Corvée, F. Brémond, and M. Thonnat, "Boosted human re-identification using Riemannian manifolds," *Image Vis. Comput.*, vol. 30, nos. 6–7, pp. 443–452, Jun. 2010.

[5] S. B¸ak, E. Corvee, F. Bremond, and M. Thonnat, "Person

[6] re-identification using Haar-based and DCD-based signature," in *Proc. 7th IEEE Int. Conf. Adv. Video Signal Based Surveill.*, Aug./Sep. 2010, pp. 1–8.

[7] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 2360–2367.

[8] M. Bauml and R. Stiefelhagen, "Evaluation of local features for person re-identification in image sequences," in *Proc. 8th IEEE Int. Conf. Adv. Video Signal-Based Surveill.*, Aug./Sep. 2011, pp. 291–296.

[9] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in *Proc. Brit. Mach. Vis. Conf.*, 2011, pp. 68.1–68.11.

[10] G. Doretto, T. Sebastian, P. Tu, and J. Rittscher, "Appearance-based person reidentification in camera networks: Problem overview and current approaches," *J. Ambient Intell. Humanized Comput.*, vol. 2, no. 2, pp. 127–151, 2011.

[11] K. Jungling and M. Arens, "View-invariant person re-identification with an implicit shape model," in *Proc. 8th IEEE Int. Conf. Adv. Video Signal- Based Surveill.*, Aug./Sep. 2011, pp. 197–202.

[12] L. Bazzani, M. Cristani, A. Perina, and V. Murino, "Multiple-shot person re-identification by chromatic and epitomic analyses," *Pattern Recognit. Lett.*, vol. 33, no. 7, pp. 898–903, May 2012.

[13] S. B¸ak, G. Charpiat, E. Corvée, F. Brémond, and M. Thonnat, "Learning to match appearances by correlations in a covariance metric space," in *Proc. Eur. Conf. Comput. Vis.*, Oct. 2012, pp. 806–820.

[14] B. Ma, Y. Su, and F. Jurie, "BiCov: A novel image representation for person re-identification and face verification," in *Proc. Brit. Machive Vis. Conf.*, 2012, pp. 1–11.

[15] B. Ma, Y. Su, and F. Jurie, "Local descriptors encoded by Fisher vectors for person re-identification," in *Proc. Int. Workshop Re-Identificat. Conjunct. Eur. Conf. Comput. Vis.*, Oct. 2012, pp. 413–422.

[16] I. Kviatkovsky, A. Adam, and E. Rivlin, "Color invariants for person reidentification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 7, pp. 1622–1634, Jul. 2013.

[17] R. Zhao, W. Ouyang, and X. Wang, "Unsupervised salience learning for person re-identification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3586–3593.

[18] Y. Xu, L. Lin, W.-S. Zheng, and X. Liu, "Human re-identification by matching compositional template with cluster sampling," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 3152–3159.

[19] D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *Proc. Eur. Conf. Compt. Vis.*,Oct. 2008, pp. 262–275.

[20] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang, “Person re-identification by support vector ranking,” in *Proc. Brit. Mach. Vis. Conf.*, 2010, pp. 21.1–21.11.

[21] T. Avraham, I. Gurvich, M. Lindenbaum, and S. Markovitch, “Learning implicit transfer for person re-identification,” in *Proc. Int. Workshop Re- Identificat. Conjunct. Eur. Conf. Comput. Vis.*, Oct. 2012, pp. 381–390.

[1]**V.Hemanth Kumar**, did his Bachelor of Technology in Electronics and Communication Engineering at Vemu institute of technology, Chittoor and doing Master of Technology Digital Electronics and Communication Systems in Annamacharya Institute of Technology & Science, Karakambadi, Tirupati, Chittoor, Andhra Pradesh, India.

[2] **Y.Penchalaiah** completed his B.Tech at JNTU Hyderabad in 2005 and M.Tech at S.V.University, Tirupati in 2010. Presently He is working as Assistant Professor of ECE, in Annamacharya Institute of Technology and Sciences, Tirupati since 2011. He has guided many B.Tech & M.tech projects. His Research area includes Signal Processing.

[3]**Miss.N.Pushpalatha** completed her B.Tech at JNTU, Hyderabad in 2004 and M.Tech at A.I.T.S., Rajampet in 2007. Presently she is working as Assistant Professor of ECE, Annamacharya Institute of Technology and Sciences, Tirupati since 2006. She has guided many B. Tech projects. Her Research area includes Data Communications and Ad-hoc Wireless Sensor Networks.