Multi-Shot Human Re-Identification Using Adaptive Fisher Discriminant Analysis

Yang Li yangli625@gmail.com Ziyan Wu ziyan@alum.rpi.edu Srikrishna Karanam karans3@rpi.edu Richard J. Radke rjradke@ecse.rpi.edu

Human re-identification (re-id) in surveillance videos has been widely studied over the past few years. The research effort was mainly focused on building appearance models [1, 3] and learning a suitable metric to determine whether or not an image pair belongs to the same person[5, 8]. In terms of the available number of images per person in each camera view, the re-id problem can be categorized into single- [7] and multi-shot [4] scenarios. In this paper, we propose a novel algorithm based on the Fisher criterion to learn a representative and discriminative feature subspace from image sequences to perform re-id.

From the practical perspective of video surveillance applications, each person should have an image sequence available from tracking algorithms. To re-identify a person, we want to learn a feature space where images belonging to the same person stay close while images belonging to different people are far apart, which can be achieved by Fisher Discriminant Analysis (FDA) [2], defined as

$$J = \operatorname{Tr}\left((\boldsymbol{T}^{\top} \boldsymbol{S}^{w} \boldsymbol{T})^{-1} \boldsymbol{T}^{\top} \boldsymbol{S}^{b} \boldsymbol{T} \right)$$
(1)

where T is the linear transformation matrix that projects data samples onto a low-dimensional subspace. S^w and S^b are the within-class scatter matrix and between-class scatter matrix respectively.

If we want to describe a person using a number of images, intuitively we need to capture as many different states (i.e., visual aspects) of a person as possible. However, it may be that in an image sequence, a majority of frames of the person are in a single state, which may cause a biased result in the FDA. That is, features reflecting other important states will be ignored. Hence, it is necessary to select representative samples which can cover the diversity of the person by clustering, as shown in Figure 1, before conducting FDA. On the other hand, the performance of clustering depends on the feature space. Only important and discriminative features should be used in clustering. Either conducting clustering before or after FDA, the results will be suboptimal.



Figure 1: A tracking image sequence is clustered into 5 groups.

In this paper we propose an Adaptive Fisher Discriminant Analysis (AFDA) algorithm to mitigate this issue. We apply the Fisher criterion in two different ways. First, considering the inherent difference between camera views, Local Fisher Discriminant Analysis (LFDA) [6] is adapted to learn a feature subspace that preserves the local structure of each person, while maximizing inter-class distance and minimizing intra-class difference. Then, we hierarchically cluster the sample data from each image sequence under this subspace, again based on the Fisher criterion. New sample points are selected from the clustering results and used to learn a new LFDA subspace. Thus, by iteratively updating the representative samples and the discriminative feature space, the algorithm can determine a better feature subspace before feeding into a metric learning method to establish the relationship between cameras. Finally, since LFDA pre-

Department of Electrical, Computer, and Systems Engineering Rensselaer Polytechnic Institute 110 8th St. Troy, NY USA

serves local structures, a metric learning step is added to compensate for the difference caused by viewpoint variation.



Figure 2: The algorithm flowchart

A flowchart of the proposed algorithm is shown in Figure 2. The proposed algorithm effectively mitigates the problem of biased sample distributions; thus, more discriminative features can be selected and more robust classifiers can be trained. It is tailored for the multi-shot scenario and shows significant improvement compared with state-of-the-art techniques.

- [1] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani. Person re-identification by symmetrydriven accumulation of local features. In *CVPR*, 2010.
- [2] Ronald A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2):179–188, 1936.
- [3] Douglas Gray and Hai Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*, 2008.
- [4] Yang Li, Ziyan Wu, and Richard J. Radke. Multi-shot reidentification with random-projection-based random forests. In WACV, 2015.
- [5] Alexis Mignon and Frédéric Jurie. PCCA: A new approach for distance learning from sparse pairwise constraints. In CVPR, 2012.
- [6] Masashi Sugiyama. Local Fisher discriminant analysis for supervised dimensionality reduction. In *ICML*, 2006.
- [7] Rui Zhao, Wanli Ouyang, and Xiaogang Wang. Unsupervised salience learning for person re-identification. In *CVPR*, 2013.
- [8] Wei-Shi Zheng, Shaogang Gong, and Tao Xiang. Person reidentification by probabilistic relative distance comparison. In *CVPR*, 2011.