

Person Re-id in Appearance Impaired Scenarios

Mengran Gou
<http://www1.coe.neu.edu/~mengran/>

Xikang Zhang
zhangxk@ece.neu.edu

Angels Rates-Borras
ratesborras.a@husky.neu.edu

Sadjad Asghari-Esfeden
www1.coe.neu.edu/~sadjad/

Octavia Camps
camps@coe.neu.edu

Mario Sznaiar
msznaiar@coe.neu.edu

Robust Systems Lab
 Electrical and Computer Engineering
 Northeastern University
 Boston

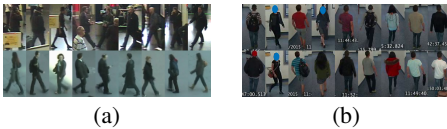


Figure 1: (a) Examples of persons wearing black suits; (b) Examples of images of the same person but wearing different clothing.

Person re-identification is critical in surveillance applications. Current approaches rely on appearance-based features extracted from a single or multiple shots of the target and candidate matches. These approaches are at a disadvantage when trying to distinguish between candidates dressed in similar colors (Figures 1(a)) or when targets change their clothing (Figures 1(b)). In this paper we propose a dynamics-based feature to overcome this limitation. **The main contributions of this paper are:** (i) A novel dynamics-based and Fisher vector encoded feature DynFV for re-id; (ii) Three new challenging “appearance impaired” datasets for re-id performance evaluation; and (iii) A comprehensive evaluation of the effect of choosing different spatio, spatio-temporal, and dynamics-based features on the performance of (unsupervised) re-id methods.

We propose to use soft-biometric characteristics provided by *sets of dense, short trajectories (tracklets)*, which have been shown to carry useful invariants [1]. All tracklets are encoded by using *pyramids of dense trajectories* with *Fisher vector encoding* [2], as illustrated in Figure 2 and described in detail in the paper.

To illustrate the need for dynamic-based features we collected three challenging “appearance-impaired” datasets. Two of them consist of video sequences of people wearing black/dark clothing. They are subsets

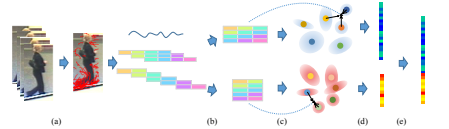


Figure 2: Pipeline of the proposed dynamics-based feature extraction.

of the iLIDS-VID and PRID 2011 datasets and we named them **iLIDSVID BK** and **PRID 2011 BK**, respectively. The third dataset, named the **Train Station dataset (TSD)**, has sequences of persons with different clothing and accessories (Figure 1(b)). We compare unsupervised re-id performance when using different combinations of five different types of features. When combining LDFV and DynFV, the rank-1 accuracies have relative improvements of 142.1% on average for all three new datasets (Figure 3).

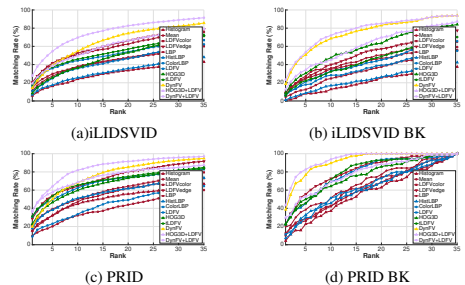


Figure 3: CMC curves for iLIDSVID, PRID and the BK extension datasets

[1] B. Li, O. Camps, and M. Sznaiar. Cross-view activity recognition using hankelets. In *CVPR*, pages 1362–1369, 2012.
 [2] F. Perronnin, J. Sánchez, and T. Mensink. Improving the fisher kernel for large-scale image classification. In *ECCV*, pages 143–156. 2010.