Information Sciences xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Person Re-identification based on nonlinear ranking with difference vectors

Tianfeng Zhou [a], Meibin Qi [a,c], Jianguo Jiang [a,c], Xin Wang [a], Shijie Hao [a,b,*], Yulong Jin [a]

[a] School of Computer and Information, Hefei University of Technology, Hefei 230009, China
[b] Postdoctoral Research Station of Computer Science and Technology, Hefei University of Technology, Hefei 230009, China
[c] Engineering Research Center of Safety Critical Industrial Measurement and Control Technology, Ministry of Education, Hefei 230009, China

A R T I C L E   I N F O

Article history:
Received 16 July 2013
Received in revised form 19 March 2014
Accepted 8 April 2014
Available online xxxx

Keywords:
Video surveillance
Person re-identification
SVM
Nonlinear ranking with difference vector

A B S T R A C T

Matching people across non-overlapping camera views, a.k.a. the person re-identification problem, is important for video surveillance and gaining increasing attention. In this paper, we propose a re-identification method that uses Nonlinear Ranking with Difference Vectors (NRDVs). Instead of trying to eliminate the differences between cameras or seek more reliable features, our strategy is to make full use of the targets' differences to build a binary classifier. We then achieve re-identification through a ranking approach by employing a support vector machine with a nonlinear kernel based on radial basis function. We also propose to pre-cluster the training images using the affinity propagation clustering algorithm, and select representative images to form negative training instances. In this strategy, the classifier maintains its performance with fewer training samples, and has lower memory requirements. Extensive experiments are conducted on three public benchmark datasets, and the results demonstrate the state-of-the-art performance of the proposed method.

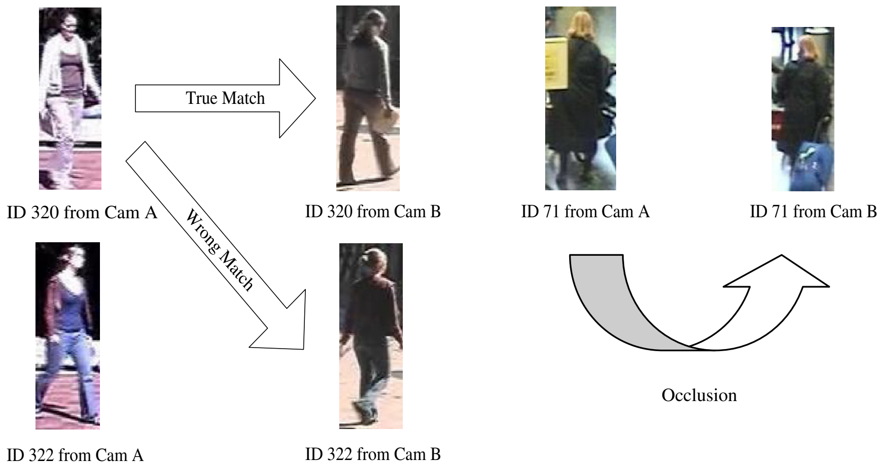
1. Introduction

Given the financial cost and civil privacy concerns, it is impractical to equip an area with an entire camera network that has no blind spots. This makes it difficult to track an individual across different camera views. The technique of person re-identification enables tracking across disjoint cameras. Compared with methods for identifying an individual in a single camera view, the difficulties of matching a person in a large area monitored by disjoint cameras are as follows. First, different individuals might be dressed similarly. Second, the appearance of the same individual may vary due to background occlusions and variations in aspects of view angle, illumination, and pose (Fig. 1). Third, the uncertain transition time of blind spots make it difficult to exploit accurate information about temporal and spatial constraints. These difficulties often result in a mismatch between irrelevant and target individuals. As a result, robust person re-identification techniques are urgently required.

In the past few years, many person re-identification methods have been proposed. These can be categorized into three groups: appearance representation methods, metric learning methods, and feature relationship modeling methods. Appearance representation methods [3,8,10,17,18,21,32] aim to seek distinctive and invariable features. However, it is a

* Corresponding author at: School of Computer and Information, Hefei University of Technology, Hefei 230009, China. Tel.: +86 13855147539; fax: +86 055182901392.
E-mail address: hfut.hsj@gmail.com (S. Hao).

http://dx.doi.org/10.1016/j.ins.2014.04.014



(a) VIPeR dataset (b) i-LIDS dataset

Fig. 1. Several images from the VIPeR dataset [9] and i-LIDS dataset [34].

sophisticated and difficult problem to extract reliable features when one camera catches a frontal view of a person, while another catches a side view. Metric learning methods [4,15,16,19,34] attempt to reduce the differences in disjoint camera images. This is done by using learning metrics that maximize the distance between mismatched image pairs and minimize the distance between correctly matched ones. Metric learning methods achieve good performance, but can be computationally expensive. In addition, when the training samples are small in number or imbalanced, the performance is likely to decline. Feature relationship modeling methods [1,14,24,25] build models by mining the relationship between inter-camera features, e.g., the brightness transfer function [24] and binary relation models [1,25]. Although these methods generally achieve promising performance, their rank 1 matching rate is unsatisfactory. To address this problem, this paper proposes a Nonlinear Ranking with Difference Vectors (NRDVs) model. NRDV not only obtains a high rank 1 matching rate, but can also be applied when there are few training data.

In the NRDV model, we formulate image pairs (each is formed by two images, i.e., one from camera A and the other from camera B) and calculate a difference vector between the two images. The difference vectors are split into positive difference vectors (positive samples) and negative difference vectors (negative samples) based on whether the two images are of the same person (see Section 3.1). We then learn a binary classifier under a support vector machine (SVM) framework with a nonlinear kernel. Similar to most multimedia retrieval systems, the re-identification is finally realized through a ranking method. To maintain the performance of our model, we also introduce a pre-clustering stage based on affinity propagation. This selects fewer, more representative samples. Extensive experiments are carried out on the VIPeR dataset [9], i-LIDS dataset [34], and ETHZ dataset [6,26] to validate the efficiency and effectiveness of the NRDV model. Our NRDV model requires simpler features than most related methods, and outperforms them in terms of the matching rate. The main contributions of our work are as follows: (1) We employ a very simple 180-dimensional HSV histogram feature for our framework. Using this feature, we formulate positive/negative difference vectors to build a binary classification model. (2) Re-identification is realized by ranking target candidates with a nonlinear kernel based on radial basis function (RBF). (3) We introduce the idea of pre-clustering to select representative negative samples for the training stage, thus maintaining the system's accuracy and reducing the computational costs.

2. Related work

Matching people across disjoint camera views is a challenging task that has gained increasing attention in recent years. Gray and Tao [10] performed viewpoint invariant pedestrian recognition with an ensemble of localized features by using the AdaBoost algorithm, while Farenzena et al. [8] selected salient parts of the body by exploiting perceptual principles of symmetry and asymmetry. Cheng et al. [3] focused on body parts and used pictorial structures to realize part-to-part correspondences. These methods have proved useful and creative, but variations in illumination or pose still greatly impact their performance. In [19], a keep it simple and straightforward metric (KISSME) was introduced to learn a distance metric from equivalence constraints. This is fast and can be used for real-time tracking. However, a singular matrix is likely to occur when there are insufficient training samples. In [34], Zheng et al. proposed a relative distance comparison (RDC) algorithm based on a logistic function in order to determine an optimal matrix, and this maximizes the intra-person similarities and inter-person dissimilarities according to the strength of metric learning. Although groundbreaking, this method is still computationally expensive. In [25], the person re-identification problem was formulated as a ranking problem in the Ensemble Rank-SVM algorithm. Avraham et al. [1] implicitly filtered out background noise by training a binary classifier with concatenated

Please cite this article in press as: T. Zhou et al., Person Re-identification based on nonlinear ranking with difference vectors, Inform. Sci. (2014), http://dx.doi.org/10.1016/j.ins.2014.04.014