Neurocomputing 167 (2015) 280–289

Contents lists available at ScienceDirect

Neurocomputing

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

Locality based discriminative measure for multiple-shot human re-identification

Wei Li [a,*], Yang Wu [b], Masayuki Mukunoki [c], Yinghui Kuang [d], Michihiko Minoh [e]

[a] School of Instrument Science and Engineering, Southeast University, 2 Sipailou, Nanjing 210096, China
[b] Institute for Research Initiatives, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma 630-0192, Japan
[c] Faculty of Engineering, University of Miyazaki, 1-1 Gakuen Kibanadai Nishi, Miyazaki 889-2192, Japan
[d] Chien-Shiung Wu College, Southeast University, 2 Sipailou, Nanjing 210096, China
[e] Academic Center for Computing and Media Studies, Kyoto University, Yoshida Nihonmatsu-cho, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 7 November 2014
Received in revised form
20 February 2015
Accepted 7 April 2015
Communicated by Liang Lin
Available online 9 May 2015

Keywords:
Multiple-shot human re-identification
Locality
Discriminative measure
Set-based matching

ABSTRACT

Multiple-shot human re-identification is an important issue in both academia and industry. It addresses the problem of building correspondences among object instances appearing in a camera network using biometric cues. Among all possible cues, face is a typical one that has long been investigated, while the whole body has become a recent trend. This problem is challenging because of small intra-class similarities and inter-class dissimilarities caused by the variations of human appearance in real scenarios. Existing methods mainly involve designing a representation and/or devising a measure to explore the within-class compactness and between-class separation. Although encouraging progress has been made, the results are still far from satisfactory. In this paper, we propose a novel set-based matching model, "Locality Based Discriminative Measure", to re-identify the human body when a set of test samples for each person are available. A new set-to-set dissimilarity is crafted considering both majorities and minorities of samples from each pair of sets. The discriminability of this dissimilarity is then further exploited by the local metric field; it can thereby serve as a more capable low-level measure to support the high-level measure for the final matching. Extensive experiments on widely used benchmarks demonstrate that our proposal remarkably outperforms state-of-the-arts.

1. Introduction

To determine the re-appearance of a person who has been previously observed in deployed cameras, human re-identification is a valuable but challenging visual surveillance task for both academia and industry. Human re-identification has a wide range of applications, such as off-line video retrieval, on-line tracking, and others. Among biometric cues, face is a crucial one and is commonly used to distinguish human identities [1,2]. However, in some real-world situations, the low quality of the images makes face re-identification impossible; therefore, whole body based re-identification has been given increasing attention. Nevertheless, unavoidable pose variations, illumination changes, viewpoint alterations, occlusions, and possibly similar body shapes and clothing styles present significant challenges to this approach.

This paper addresses multiple-shot human re-identification in which there are multiple body images available in query and corpus domains for each identity in question. The relevant methodologies used can be categorized in two paradigms. Some of the methodologies pay attention to reliable representation [3–8]. Although a sound representation can characterize the human appearance, the complexity of the real-world situation and the subjectivity of the feature hand-designing process inevitably impede the performance enhancement. Other solutions rely on robust measure for the representation to address these problems [9–14]. Their inspiring results have been achieved because of improved intra-class compactness and inter-class separation. However, a lack of significant progress in this direction has also rapidly slowed the performance improvement.

Feature representations of human images can be considered as points in the topological space. We define a group of features extracted from multiple-shot images of the same person to be one whole set. Based on this definition, the key problem of multiple-shot human re-identification involves determining a suitable dissimilarity measure for accurately matching the sets between query and corpus domains. This type of measure can be explored from two aspects: one is to adapt the point-based distance to the set level; the other is to seek the underlying metrics for the sets. Exploration and collaboration of these two aspects could lead to a significant advancement in this important yet unresolved area in computer vision. Based on this objective, we propose a novel

* Corresponding author.
E-mail address: seuliwei@126.com (W. Li).

http://dx.doi.org/10.1016/j.neucom.2015.04.068
0925-2312/

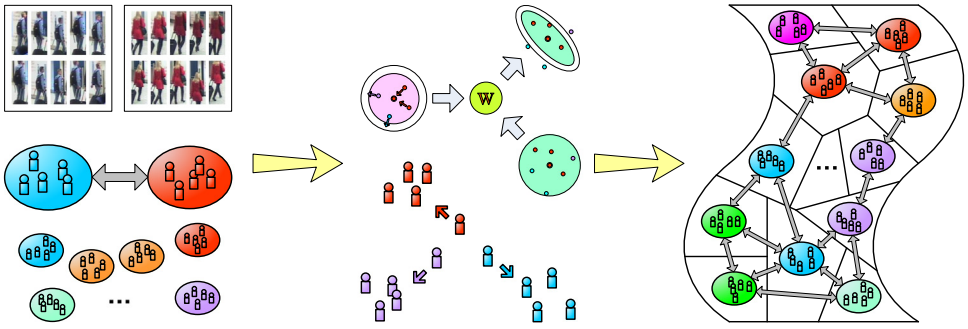

Fig. 1. Locality-based discriminative measure (LBDM) comprises three primary steps: set-to-set dissimilarity crafting (diagram left), local metric field constructing (diagram middle), and set-based matching (diagram right).

set-based matching model, referred to as "Locality Based Discriminative Measure (LBDM)", to re-identify the human images.

As depicted in Fig. 1, LBDM comprises three primary steps: set-to-set dissimilarity crafting, local metric field constructing, and set-based matching. The first step involves designing a novel set-to-set dissimilarity for multiple-shot human images. This dissimilarity transfers the sample based distance to the set level, while providing a tool to determine the neighboring sets of each set needed in the local metric learning framework. Local metric learning is intended to pull closer together samples of the same set and those in its neighborhood farther away to ensure a more reliable set-to-set dissimilarity measure. By enabling a more reliable set-to-set dissimilarity measure, the set-level oriented common-near-neighbor modeling can be fully leveraged for effective matching.

In sum, the main contributions of this paper are as follows:

- We craft a novel set-to-set dissimilarity as a low-level measure. By considering the local minorities of samples in each set, the measure can preserve within-set variability; by determining the global distance between majorities of samples from paired sets, the measure is robust to the irregular outliers (Section 3.1).
- We introduce a new local metric learning model. It constructs the local metric field,[1] in which the discriminability of the new set-to-set distance is greatly enhanced into a middle-level measure (Section 3.2).
- We present an effective set-based matching framework. It propagates the effectiveness of the middle-level dissimilarity measure to the high-level dissimilarity measure presented by set-level common-near-neighbor modeling for final matching (Section 3.3).

The above contributions are based on but extend our previous work [15,16], with further scope of elaboration and experimentation.

2. Proposed method: problem definition and overview

Generally speaking, the issue of person re-identification can be divided into two research directions. One is single-shot re-identification; the other is multiple-shot re-identification. For the single-shot case, there is only one single image for each identity in query and corpus sides. While for the multiple-shot case, there are several images, which form an image set, for each identity in query and corpus domains. The set concept is very important for multiple-shot re-identification. Although methods for single-shot re-identification can also be applied to the multiple-shot problem by breaking the human image set into multiple single-shot images, this brute way is inevitably detrimental to the performance of multiple-shot re-identification.

In our study, we reformulate the multiple-shot re-identification problem into a single-set versus single-set matching problem between cameras based on the set-integrity assumption: for each camera domain, the images within each set have the same identity label (i.e., they are relevant) though the label is unknown, and images from different sets have different identity labels (i.e., they are irrelevant). Such set-integrity information is commonly considered by multiple-shot re-identification solutions more or less [5,9,10].

It should be noted that the single set based matching problem is one of the most important, common, and fundamental cases for the issue of multiple-shot re-identification. It is reasonable for multiple-shot re-identification, because it is normal and relatively convenient to collect the image sequence of each identity within one fixed camera view to form one single set.

Let us suppose that the query sets contain images acquired from one camera while the corpus sets are comprised of images from another camera. Our objective is to determine the correct correspondences between the query and corpus sets to see whether they belong to the same person or not. For each query set, a ranking list can be generated by sorting the dissimilarities between itself and all corpus sets. The matching can therefore be evaluated by these ranking results according to the ground-truth correspondences.

To address the above, we employ the proposed LBDM model. We measure the new dissimilarity between the query and corpus sets, which provides a method to statistically transfer the measure from the sample-based distance into the set-based dissimilarity. We then construct the local metric field to further exploit the discriminability of this dissimilarity, which searches the suitable underlying local metric to ameliorate the inner structure and neighboring distribution for each set. Finally, we match the query and corpus sets by propagating the effectiveness of this exploited distance to the high-level measure, which models the set-level relative neighborhood structure comparison for each set pair.

3. Locality based discriminative measure

3.1. Set-to-set distance crafting

To correctly match the sets, an opportune set-to-set distance is important. Most previous methods have spotlighted minority-based distance, while claiming the effectiveness of this strategy. Minority-based distance takes within-set variability into account by measuring the closest local minorities of each paired sets. Two exemplary methods are Minimum Point-wise Distance (MPD) and

[1] In this paper, we treat possibly different metrics at local areas as a field of local metrics.