Super-resolution Person Re-identification with Semi-coupled Low-rank Discriminant Dictionary Learning

Xiao-Yuan Jing\textsuperscript{1,3}, Xiaoke Zhu\textsuperscript{1}, Fei Wu\textsuperscript{1,3}, Xinge You\textsuperscript{2}, Qinglong Liu\textsuperscript{1}, Dong Yue\textsuperscript{3}, Ruimin Hu\textsuperscript{4}, Baowen Xu\textsuperscript{1}

\textsuperscript{1}State Key Laboratory of Software Engineering, School of Computer, Wuhan University, China
\textsuperscript{2}School of Electronic Information and Communications, Huazhong University of Science and Technology, China
\textsuperscript{3}College of Automation, Nanjing University of Posts and Telecommunications, China
\textsuperscript{4}National Engineering Research Center for Multimedia Software, School of Computer, Wuhan University, China

Abstract

Person re-identification has been widely studied due to its importance in surveillance and forensics applications. In practice, gallery images are high-resolution (HR) while probe images are usually low-resolution (LR) in the identification scenarios with large variation of illumination, weather or quality of cameras. Person re-identification in this kind of scenarios, which we call super-resolution (SR) person re-identification, has not been well studied. In this paper, we propose a semi-coupled low-rank discriminant dictionary learning (SLD\textsuperscript{2}L) approach for SR person re-identification. For the given training image set which consists of HR gallery and LR probe images, we aim to convert the features of LR images into discriminating HR features. Specifically, our approach learns a pair of HR and LR dictionaries and a mapping from the features of HR gallery images and LR probe images. To ensure that the converted features using the learned dictionaries and mapping have favorable discriminative capability, we design a discriminant term which requires the converted HR features of LR probe images should be close to the features of HR gallery images from the same person, but far away from the features of HR gallery images from different persons. In addition, we apply low-rank regularization in dictionary learning procedure such that the learned dictionaries can well characterize intrinsic feature space of HR and LR images. Experimental results on public datasets demonstrate the effectiveness of SLD\textsuperscript{2}L.

1. Introduction

Person re-identification is a fundamental task in automated video surveillance and has been widely researched in recent years. Given an image/video of a person taken from one camera, re-identification is the process of identifying the person from images/videos taken from a different camera \cite{2}. Many methods have been presented for person re-identification \cite{4, 33, 13, 15, 17, 21}. These methods can be roughly classified into two categories: (1) Methods on feature representation \cite{1, 6, 23, 31}: they focus on seeking a distinct and robust feature representation for matching. The Literature \cite{9} seeks a more distinctive representation by exploiting the class information to overcome the large intra-class appearance variations. Literature \cite{24} is presented to solve the feature selection problem. Literature \cite{22} is proposed to learn the most discriminative attribute that characterizes a particular individual, in which saliency detection is utilized to drive automatically the PTZ camera to focus on certain parts of a human body. (2) Methods on distance learning \cite{10, 14, 19, 35, 56}: they focus on seeking an optimal distance metric for person re-identification. In \cite{10}, Hirzer et al. learned a metric from pairs of samples belonging to different cameras using discriminative Mahalanobis metric learning, which can be efficiently solved after some relaxations. In \cite{14}, a distance metric is learned based on equivalence constraints from a statistical inference perspective. In \cite{56}, Zheng et al. learned a Mahalanobis distance metric with a probabilistic relative distance comparison (RDC) method. Recently, dictionary learning methods have gained increasing attention and achieved promising results. However, these methods are usually designed for LR probe images and may not perform well for HR gallery images.
Acknowledgement

The work described in this paper was supported by the National Nature Science Foundation of China under Project Nos. 61272273, 61233011, 61231015, 61171273, 61272203, 61374055, the National High Technology Research and Development Program of China (863 Program) under Project Nos. 2015AA016306, 2013AA014602, the Internet of Things Development Project of Ministry of industry in 2013 (No. 25), and the Research Project of NJUPT (XJKY14016).

References