

Author's Accepted Manuscript

Person Re-Identification by Unsupervised Video Matching

Xiaolong Ma, Xiatian Zhu, Shaogang Gong, Xudong Xie, Jianming Hu, Kin-Man Lam, Yisheng Zhong



PII: S0031-3203(16)30376-4
DOI: <http://dx.doi.org/10.1016/j.patcog.2016.11.018>
Reference: PR5961

To appear in: *Pattern Recognition*

Received date: 31 March 2016
Revised date: 21 November 2016
Accepted date: 21 November 2016

Cite this article as: Xiaolong Ma, Xiatian Zhu, Shaogang Gong, Xudong Xie, Jianming Hu, Kin-Man Lam and Yisheng Zhong, Person Re-Identification by Unsupervised Video Matching, *Pattern Recognition*, <http://dx.doi.org/10.1016/j.patcog.2016.11.018>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and a review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Person Re-Identification by Unsupervised Video Matching

Xiaolong Ma^{1,4}, Xiatian Zhu², Shaogang Gong², Xudong Xie¹, Jianming Hu¹,
Kin-Man Lam³, Yisheng Zhong¹

Abstract

Most existing person re-identification (ReID) methods rely only on the spatial appearance information from either one or multiple person images, whilst ignore the space-time cues readily available in video or image-sequence data. Moreover, they often assume the availability of exhaustively labelled cross-view pairwise data for every camera pair, making them non-scalable to ReID applications in real-world large scale camera networks. In this work, we introduce a novel video based person ReID method capable of accurately matching people across views from arbitrary *unaligned* image-sequences without any labelled pairwise data. Specifically, we introduce a new space-time person representation by encoding multiple granularities of spatio-temporal dynamics in form of time series. Moreover, a Time Shift Dynamic Time Warping (TS-DTW) model is derived for performing automatic alignment whilst achieving data selection and matching between inherently inaccurate and incomplete sequences in a unified way. We further extend the TS-DTW model for accommodating multiple feature-sequences of an image-sequence in order to fuse information from different descriptions. Crucially, this model does not require pairwise labelled training data (i.e. unsupervised) therefore readily scalable to large scale cam-

Email addresses: goup000@163.com (Xiaolong Ma), xiatian.zhu@qmul.ac.uk (Xiatian Zhu), s.gong@qmul.ac.uk (Shaogang Gong), xdxie@mail.tsinghua.edu.cn (Xudong Xie), hujm@mail.tsinghua.edu.cn (Jianming Hu), kin.man.lam@polyu.edu.hk (Kin-Man Lam), zys-dau@mail.tsinghua.edu.cn (Yisheng Zhong)

¹Tsinghua University, China

²Queen Mary University of London, United Kingdom

³The Hong Kong Polytechnic University, Hong Kong

⁴China Academy of Electronics and Information Technology

Download English Version:

<https://daneshyari.com/en/article/4969656>

Download Persian Version:

<https://daneshyari.com/article/4969656>

[Daneshyari.com](https://daneshyari.com)