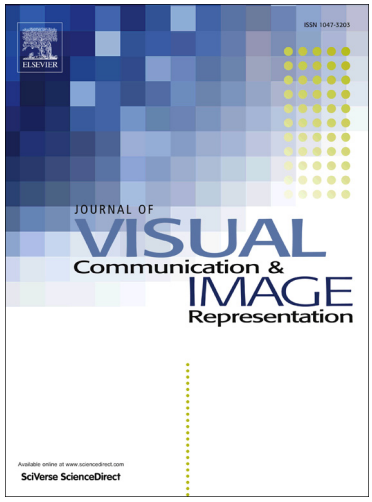

Multi-View Representation Learning for Multi-View Action Recognition

Tong Hao, Dan Wu, Qian Wang, Jin-Sheng Sun

PII: S1047-3203(17)30025-1
DOI: http://dx.doi.org/10.1016/j.jvcir.2017.01.019
Reference: YJVCI 1939

To appear in: *J. Vis. Commun. Image R.*

Received Date: 2 September 2016
Revised Date: 12 December 2016
Accepted Date: 16 January 2017

Please cite this article as: T. Hao, D. Wu, Q. Wang, J-S. Sun, Multi-View Representation Learning for Multi-View Action Recognition, *J. Vis. Commun. Image R.* (2017), doi: http://dx.doi.org/10.1016/j.jvcir.2017.01.019

Multi-View Representation Learning for Multi-View Action Recognition

Tong Hao[a], Dan Wu[a], Qian Wang[a], Jin-Sheng Sun[a,b,*]

[a] *Tianjin Key Laboratory of Animal and Plant Resistance/College of Life Science, Tianjin Normal University, Tianjin 300387, China*
[b] *Tianjin Aquatic Animal Infectious Disease Control and Prevention Center, Tianjin 300221, China*

Abstract

Although multiple methods have been proposed for human action recognition, the existing multi-view approaches can not well discover meaningful relationship among multiple action categories from different views. To handle this problem, this paper proposes an multi-view learning approach for multi-view action recognition. First, the proposed method leverages the popular visual representation method, bag-of-visual-words (BoVW) / fisher vector (FV), to represent individual videos in each view. Second, the sparse coding algorithm is utilized to transfer the low-level features of various views into the discriminative and high-level semantics space. Third, we employ the multi-task learning (MTL) approach for joint action modeling and discovery of latent relationship among different action categories. The extensive experimental results on M^2I and IXMAS datasets have demonstrated the effectiveness of our proposed approach. Moreover, the experiments further demonstrate that the discovered latent relationship can benefit multi-view model learning to augment the performance of action recognition.

Keywords: multi-view learning, multi-task learning, sparse coding, action recognition

1. Introduction

With the rapid development of human computer interaction, action-centric video indexing and retrieval, video surveillance, and so on, human action

*Corresponding author: Jin-Sheng Sun, Email: jinshsun@163.com