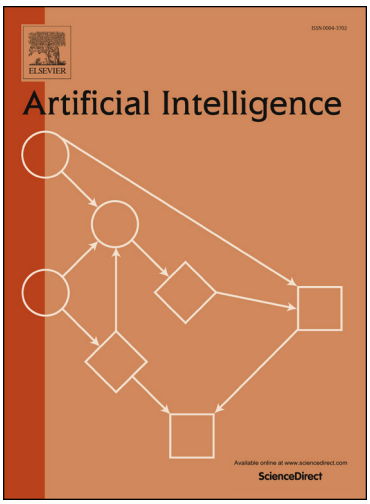
Accepted Manuscript



Discovering visual concept structure with sparse and incomplete tags

Jingya Wang, Xiatian Zhu, Shaogang Gong

PII: S0004-3702(17)30068-1
DOI: http://dx.doi.org/10.1016/j.artint.2017.05.002
Reference: ARTINT 3015

To appear in: *Artificial Intelligence*

Received date: 24 December 2015
Revised date: 7 February 2017
Accepted date: 25 May 2017

Please cite this article in press as: J. Wang et al., Discovering visual concept structure with sparse and incomplete tags, *Artif. Intell.* (2017), http://dx.doi.org/10.1016/j.artint.2017.05.002

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 review of the resulting proof before it is published in its final 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.

ACCEPTED MANUSCRIPT

Discovering Visual Concept Structure with Sparse and Incomplete Tags

Jingya Wang[a], Xiatian Zhu[a], Shaogang Gong[a]

[a]*School of EECS, Queen Mary University of London, Mile End Road, London, E1 4NS, UK*

Abstract

Discovering automatically the semantic structure of tagged visual data (e.g. web videos and images) is important for visual data analysis and interpretation, enabling the machine intelligence for effectively processing the fast-growing amount of multi-media data. However, this is non-trivial due to the need for jointly learning underlying correlations between heterogeneous visual and tag data. The task is made more challenging by inherently sparse and incomplete tags. In this work, we develop a method for modelling the inherent visual data concept structures based on a novel Hierarchical-Multi-Label Random Forest model capable of correlating structured visual and tag information so as to more accurately interpret the visual semantics, e.g. disclosing meaningful visual groups with similar high-level concepts, and recovering missing tags for individual visual data samples. Specifically, our model exploits hierarchically structured tags of different semantic abstractness and multiple tag statistical correlations in addition to modelling visual and tag interactions. As a result, our model is able to discover more accurate semantic correlation between textual tags and visual features, and finally providing favourable visual semantics interpretation even with highly sparse and incomplete tags. We demonstrate the advantages of our proposed approach in two fundamental applications, visual data clustering and missing tag completion, on benchmarking video (i.e. TRECVID MED 2011) and image (i.e. NUS-WIDE) datasets.

Keywords: Visual semantic structure; Tag hierarchy; Tag correlation; Sparse tags; Incomplete tags; Data clustering; Missing tag completion; Random forest.

Email addresses: jingya.wang@qmul.ac.uk (Jingya Wang), xiatian.zhu@qmul.ac.uk (Xiatian Zhu), s.gong@qmul.ac.uk (Shaogang Gong)

Download English Version:

https://daneshyari.com/en/article/4942051

Download Persian Version:

https://daneshyari.com/article/4942051

Daneshyari.com