ABSTRACT

This tutorial focuses on challenges and solutions for content-based image annotation and retrieval in the context of online image sharing and tagging. We present a unified review on three closely linked problems, i.e., tag assignment, tag refinement, and tag-based image retrieval. We introduce a taxonomy to structure the growing literature, understand the ingredients of the main works, clarify their connections and difference, and recognize their merits and limitations. Moreover, we present an open-source testbed, with training sets of varying sizes and three test datasets, to evaluate methods of varied learning complexity. A selected set of eleven representative works have been implemented and evaluated. During the tutorial we provide a practice session for hands-on experience with the methods, software and datasets. For repeatable experiments all data and code are online at http://www.micc.unifi.it/tagsurvey.

Categories and Subject Descriptors

H.3.3 [INFORMATION STORAGE AND RETRIEVAL]: Information Search and Retrieval; H.3.1 [INFORMATION STORAGE AND RETRIEVAL]: Content Analysis and Indexing—Indexing Methods

General Terms

Algorithms, Experimentation, Performance

Keywords

Content-based image retrieval, social tagging, tag relevance, tag assignment, tag refinement, tag retrieval

*Equal contribution and corresponding authors.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author(s). Copyright is held by the owner/author(s).

MM’15 October 26-30, 2015, Brisbane, Australia.
ACM 978-1-4503-3459-4/15/10.
DOI: http://dx.doi.org/10.1145/2733373.2807419.
Table 1: Methods evaluated in this tutorial. The media column characterizes what essential information a specific method exploits, while the learning SemanticField depicts how such information is exploited.

<table>
<thead>
<tr>
<th>Method</th>
<th>Media</th>
<th>Learning</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>TagCooccur [9]</td>
<td>tag</td>
<td>Instance</td>
<td>Python</td>
</tr>
<tr>
<td>TagRanking [6]</td>
<td>tag + image</td>
<td>Instance</td>
<td>C + Python</td>
</tr>
<tr>
<td>KNN [7]</td>
<td>tag + image</td>
<td>Instance</td>
<td>C + Python</td>
</tr>
<tr>
<td>TagVote [4]</td>
<td>tag + image</td>
<td>Instance</td>
<td>Python</td>
</tr>
<tr>
<td>TagCooccur+ [4]</td>
<td>tag + image</td>
<td>Instance</td>
<td>Python</td>
</tr>
<tr>
<td>TagProp [2]</td>
<td>tag + image</td>
<td>Model based</td>
<td>C + Matlab + Python</td>
</tr>
<tr>
<td>TagFeature [1]</td>
<td>tag + image</td>
<td>Model based</td>
<td>C + Python</td>
</tr>
<tr>
<td>RelExample [3]</td>
<td>tag + image</td>
<td>Model based</td>
<td>C + Python</td>
</tr>
<tr>
<td>RobustPCA [10]</td>
<td>tag + image</td>
<td>Transduction</td>
<td>C + Matlab + Python</td>
</tr>
<tr>
<td>TensorAnalysis [8]</td>
<td>tag + image + user</td>
<td>Transduction</td>
<td>C + Matlab + Python</td>
</tr>
</tbody>
</table>

by various research groups. A selected set of eleven representative works, i.e., SemanticField [11], TagRanking [6], KNN [7], TagVote [4], TagProp [2], TagCooccur [9], TagCooccur+ [4], TagFeature [1], RelExample [3], RobustPCA [10], TensorAnalysis [8], have been implemented and evaluated on the test bed for tag assignment, refinement, and/or retrieval. An overview of the methods is given in Table 1. The interested reader is referred to [5] for a comprehensive comparison between these methods.

During the tutorial, we also provide a practice session for hands on experience with the methods, software, and datasets. For each method a front-end pipeline is implemented, allowing users to conduct tag relevance learning from scratch, obtain tag ranks and image ranks accordingly, and report multiple performance metrics including image-centric Mean image Average Precision (MiAP), tag-centric Mean Average Precision (MAP), and Normalized Discounted Cumulative Gain (NDCG). In addition, Python wrappers for C and Matlab code are given for the ease of cross-platform use.

We conclude the course with our perspective on the many challenges and opportunities ahead for the multimedia community.


Acknowledgments. This research is supported by Natural Science Foundation of China (No. 61303184), the Fundamental Research Funds for the Central Universities and the Research Funds of Remin University of China (No. 14XNLQ01), SRFDP (No. 2013004120006), the Dutch national program COMMIT, the STW STORY project, Telecom Italia PhD grant funds, and the AQUIS-CH project granted by the Tuscany Region (Italy). L. Ballan acknowledges also the support of the EC’s FP7 under the grant agreement No. 623030 (Marie Curie IOF).

2. REFERENCES


1 Results provided by the method developers.