Chapter XV
Semantic Classification and Annotation of Images

Yonghong Tian
Peking University, China

Shuqiang Jiang
Chinese Academy of Sciences, China

Tiejun Huang
Peking University, China

Wen Gao
Peking University, China

ABSTRACT

With the rapid growth of image collections, content-based image retrieval (CBIR) has been an active area of research with notable recent progress. However, automatic image retrieval by semantics still remains a challenging problem. In this chapter, the authors will describe two promising techniques towards semantic image retrieval—semantic image classification and automatic image annotation. For each technique, four aspects are presented: task definition, image representation, computational models, and evaluation. Finally, they will give a brief discussion of their application in image retrieval.

INTRODUCTION

With the advance of multimedia technology and growth of image collections, content-based image retrieval (CBIR) is therefore proposed, which finds images that have low-level visual features (e.g., color, texture, shape) similar to those of the query image. However, retrieving images via low-level features has proven unsatisfactory since low-level visual features cannot represent the high-level semantic content
Semantic Classification and Annotation of Images

Figure 1. Examples of images for semantic image categorization.

<table>
<thead>
<tr>
<th></th>
<th>Coast</th>
<th>Forest</th>
<th>Highway</th>
<th>Inside City</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mountain</td>
<td>Open Country</td>
<td>Street</td>
<td>Tall Building</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Examples of images for automatic image annotation.

<table>
<thead>
<tr>
<th>Images</th>
<th>Original manual annotations</th>
<th>Automatic annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>sky jet Plane</td>
<td>sky jet Plane smoke</td>
<td>sky jet plane water tree</td>
</tr>
<tr>
<td>smoke</td>
<td>bear close-up face polar</td>
<td>bear snow polar grass water</td>
</tr>
<tr>
<td></td>
<td>cat rocks tiger water</td>
<td>water cat tiger people swimmers</td>
</tr>
<tr>
<td></td>
<td>flowers porcupine rodent</td>
<td>leaf plants flowers garden</td>
</tr>
<tr>
<td></td>
<td>sky sun tree water</td>
<td>sunset water horizon sun clouds</td>
</tr>
</tbody>
</table>

of images. To reduce the so-called semantic gap (Smeulders, Worring, Santini, Gupta, & Jain, 2000), a variety of techniques have been developed. In this chapter, we discuss promising techniques on two important aspects of CBIR — (a) semantic image classification, and (b) automatic image annotation. Each component plays an important role in the greater semantic understanding of images.

General speaking, the semantics of images can be categorized into four levels from the lowest to the highest (Wang, Li, & Wiederhold, 2001):

- **semantic types** (e.g., landscape photograph, clip art),
- **object composition** (e.g., a bike and a car parked on a beach, a sunset scene),
- **abstract semantics** (e.g., people fighting, happy person, objectionable photograph), and
- **detailed semantics** (e.g., a detailed description for a given picture).

In (Song, Wang, & Zhang, 2003), the image semantics can be further grouped into local semantic level and thematic level (or global semantic level). Currently, most works in image classification and annotation have been done in the first two levels. Good progress has been made at the first level, which corresponds closely to image’s physical attributes such as indoor vs. outdoor. The second level of
Related Content

Elitist and Ensemble Strategies for Cascade Generalization
www.igi-global.com/article/ellitist-ensemble-strategies-cascade-generalization/3359?camid=4v1a

Does Protecting Databases Using Perturbation Techniques Impact Knowledge Discovery?
www.igi-global.com/chapter/does-protecting-databases-using-perturbation/4369?camid=4v1a

The Expert's Opinion
www.igi-global.com/article/expert-opinion/51100?camid=4v1a

On the Representation of Temporal Dynamics
www.igi-global.com/chapter/representation-temporal-dynamics/4340?camid=4v1a