Chapter VI

Automatic and Semi-Automatic Techniques for Image Annotation

Biren Shah, University of Louisiana at Lafayette, USA
Ryan Benton, University of Louisiana at Lafayette, USA
Zonghuan Wu, University of Louisiana at Lafayette, USA
Vijay Raghavan, University of Louisiana at Lafayette, USA

Abstract

When retrieving images, users may find it easier to express the desired semantic content with keywords than with visual features. Accurate keyword retrieval can occur only when images are described completely and accurately. This can be achieved either through laborious manual effort or automated approaches. Current methods for automatically extracting semantic information from images can be classified into (a) text-based methods that use metadata such as ontological descriptions and/or text associated with images to assign and/or refine annotations, and (b) image-based methods that focus on extracting semantic information directly from image content. The focus of this chapter is to create an awareness and understanding of research and advances in this field by introducing them to basic concepts and theories and then by classifying, summarizing, and describing works from the published literature. It also will identify unsolved problems and offer suggestions for future research directions.
Introduction

A picture is worth a thousand words. As human beings, we are able to tell a story from a picture based upon what we see and our background knowledge. Can a computer program discover semantic concepts from images, build models from them, and recognize them based on these models? The short answer is yes. However, there is no perfect solution, and there are several challenges that need to be addressed.

Automatic image annotation is important to both video and image retrieval and computer object recognition. It potentially can be applied to many areas, including biomedicine, commerce, military, journalism, education, digital libraries, advertising, and Web searching. Decades of research has shown that it is extremely difficult to design a set of generic computer algorithms that can learn concepts from images to automatically translate their contents to linguistic terms.

The mission of this chapter is to create an awareness and understanding of research and advances in the field of automatic image annotation by introducing the basic concepts and theories and then by classifying, summarizing, and describing published works. The chapter also will identify unsolved problems and offer suggestions for future research directions. The chapter will address the following questions:

• What is a semantic gap?
• Why is there a need for automatic techniques for image annotation?
• What are the various automatic annotation methods in the literature?
• How does one assess the quality of annotations?

The rest of the chapter is organized as follows. The next section provides background and related work. The text-based and image-based methods for automatic annotation are discussed in the third and fourth sections, respectively. In the fifth section, we provide several metrics for evaluating the quality of annotations. Finally, we conclude the chapter with a general discussion on future and emerging trends.

Background and Related Work

“Given a new image, find similar images in the current database that best describe its content. Given a large image database, find images that have tigers.” These are examples of queries encountered by many current image indexing and retrieval systems. Effective techniques are needed to model and search the content of large digital image archives.

One solution is to use content-based image retrieval (CBIR) systems (Flickner et al., 1995; Shah, Raghavan, Dhatric, & Zhao, 2006) that compute relevance based on the visual similarity of low-level image features such as color, texture, shape, and so forth. However, visual
A Broader View on Context Models towards Supporting Business Process Agility
www.irma-international.org/chapter/broader-view-context-models-towards/60069/