

Chapter 8.8

Discovering Semantics from Visual Information

Zhiyong Wang

University of Sydney, Australia

Dagan Feng

University of Sydney, Australia & Hong Kong Polytechnic University, China

ABSTRACT

Visual information has been immensely used in various domains such as web, education, health, and digital libraries, due to the advancements of computing technologies. Meanwhile, users realize that it has been more and more difficult to find desired visual content such as images. Though traditional content-based retrieval (CBR) systems allow users to access visual information through query-by-example with low level visual features (e.g. color, shape, and texture), the semantic gap is widely recognized as a hurdle for practical adoption of CBR systems. Wealthy visual information (e.g. user generated visual content) enables us to derive new knowledge at a large scale, which

will significantly facilitate visual information management. Besides semantic concept detection, semantic relationship among concepts can also be explored in visual domain, other than traditional textual domain. Therefore, this chapter aims to provide an overview of the state-of-the-arts on discovering semantics in visual domain from two aspects, semantic concept detection and knowledge discovery from visual information at semantic level. For the first aspect, various aspects of visual information annotation are discussed, including content representation, machine learning based annotation methodologies, and widely used datasets. For the second aspect, a novel data driven based approach is introduced to discover semantic relevance among concepts in visual domain. Future research topics are also outlined.

DOI: 10.4018/978-1-60960-818-7.ch8.8

1. INTRODUCTION

In the last decades we have witnessed tremendous growth in visual information such as images and videos, due to the advancements of computing technologies. It has never been easier than today to take images through digital cameras or video shots through camcorders, ranging from personal collections to professional archives such as news. And the rapid development of the Web has further accelerated this process by allowing any users to publish or share their visual contents conveniently. There are about 3 billion images are hosted by Flickr¹ and about 12.7 billion videos were watched in a month by America Internet users only². And millions of images are uploaded every day to popular photo sharing web sites like Flickr. Therefore, efficient access to such enormous amount of visual information has emerged as a challenging issue.

Since 1980s, visual information retrieval has been a very active research topic with the endeavor from two communities, database management and computer vision (Tamura & Yokoya, 1984) (Chang, Shi, & Yan, 1987) (Chang, Yan, Dimitroff, & Arndt, 1988). These two communities study image retrieval from different aspects, one being text or alphanumeric based and the other visual based, respectively. Unlike textual information which can be characterized in terms of its semantic primitives (i.e. terms), visual information lacks such primitives even with the state-of-the-art techniques in computer vision. Therefore, in order to leverage the success of relational database, visual information in general is manually annotated with textual descriptions for retrieval purpose. However, manual annotation suffers from the follow issues:

- Manual annotation is very time consuming and labor intensive, which is not scalable to the dramatically growing visual information.

- Manual annotation is subjectively dependent on annotators and not general for all the possible front-end users, especially for the users using different languages.
- It is very difficult and challenging to describe image content with only several keywords. As said a picture is worth a thousand words. Sometimes some aspects of image content such as texture are even beyond words.

In order to overcome these problems, it would be more ideal to characterize visual content with perceptual attributes (i.e. visual features) such as color, shape, texture, and motion. In the early 1990s, content based image retrieval (CBIR) was proposed to allow users to search for target visual information in terms of its true content represented with visual features by making use of techniques from image processing and computer vision domains. CBIR has been an interesting and promising research area for decades and many query strategies have been proposed, such as query by example (QBE) and query by sketch (Veltkamp & Tanase, 2000). As reviewed in (Gupta & Jain, 1997) (Idris & Panchanathan, 1997) (Loncaric, 1998) (Brunelli, Mich, & Modena, 1999) (Y. Rui, Huang, & Chang, 1999) (Smeulders, Worring, Santini, Gupta, & Jain, 2000) (Antani, Kasturi, & Jain, 2002) (D. Feng, Siu, & Zhang, 2003) (C. G. M. Snoek & Worring, 2005) (Datta, Li, & Wang, 2005) (Lew, Sebe, Djeraba, & Jain, 2006) (Y. Liu, Zhang, Lu, & Ma, 2007) (Datta, Joshi, Li, James, & Wang, 2008) (Ren, S. Singh, M. Singh, & Y. S. Zhu, 2009), most of works have focused on the following issues, feature extraction to characterize visual content, feature transformation including dimension reduction and feature selection to achieve compact and optimal representation, similarity measurement for efficient matching (Santini & Jain, 1999), high dimensional indexing for efficient search (Valle, Cord, & Philipp-Foliguet, 2008), and relevance

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/discovering-semantics-visual-information/56238

Related Content

Wolf-Swarm Colony for Signature Gene Selection Using Weighted Objective Method

Prativa Agarwalla and Sumitra Mukhopadhyay (2019). *Nature-Inspired Algorithms for Big Data Frameworks* (pp. 170-195).

www.irma-international.org/chapter/wolf-swarm-colony-for-signature-gene-selection-using-weighted-objective-method/213035

A Dualism Based Semantics Formalization Mechanism for Model Driven Engineering

Yucong Duan (2009). *International Journal of Software Science and Computational Intelligence* (pp. 90-110).

www.irma-international.org/article/dualism-based-semantics-formalization-mechanism/37490

Cognitive Memory: Human Like Memory

Bernard Carlos Widrow and Juan Aragon (2012). *Breakthroughs in Software Science and Computational Intelligence* (pp. 84-99).

www.irma-international.org/chapter/cognitive-memory-human-like-memory/64604

A Collaborative Pointing Experiment for Analyzing Bodily Communication in a Virtual Immersive Environment

Divesh Lala and Toyooki Nishida (2012). *International Journal of Software Science and Computational Intelligence* (pp. 1-19).

www.irma-international.org/article/collaborative-pointing-experiment-analyzing-bodily/76267

Classifier Ensemble Based Analysis of a Genome-Wide SNP Dataset Concerning Late-Onset Alzheimer Disease

Lúcio Coelho, Ben Goertzel, Cassio Pennachin and Chris Heward (2010). *International Journal of Software Science and Computational Intelligence* (pp. 60-71).

www.irma-international.org/article/classifier-ensemble-based-analysis-genome/49132