Chapter 20
An Investigation in Multi-Feature Query Language Based Classification in Image Retrieval:
Introduction

Raoul Pascal Pein
University of Huddersfield, UK

Joan Lu
University of Huddersfield, UK

Wolfgang Renz
Hamburg University of Applied Sciences, Germany

ABSTRACT

With rapid development of digital technologies, building an efficient and reliable image retrieval system is always challenging in computing science and related application disciplines. This book part presents an investigation in how "Content-Based Image Retrieval (CBIR)" queries could be designed in order to achieve an extensible language understandable by both humans and machines. The query language used applies concepts from established text search and image retrieval engines. The question of whether such a query language can be sufficiently expressive to formally describe certain real-life concepts is investigated. Sets of images from different classes are used to build "descriptor" queries that are supposed to capture a single concept.

DOI: 10.4018/978-1-4666-1975-3.ch020

INTRODUCTION

Modern technology allows people to easily create and store huge amounts of digital images as well as other multimedia content. As storage space is getting rapidly cheaper, the challenge of efficiently managing the generated content becomes a growing challenge, not only for companies maintaining huge repositories of digital images, but also for amateur users. This chapter highlights the main issues of Multimedia Retrieval and the motivation for this investigation.

With rapid development of digital technologies, building an efficient and reliable image retrieval system is always challenging in computing science and related application disciplines (Garber & Grunes, 1992, Flickr, 2006, Google, Inc., 2008b, Schietse, Eakins, & Veltkamp, 2007, Müller et al., 2005, Blaser & Egenhofer, 2000, Fatto, Paolino, & Pittarello, 2007, Pein, Amador, Lu, & Renz, 2008, Bosma, Veltkamp, & Wiering, 2006). A typical application area of CBIR is multimedia publishing and design. Often, an image with specific properties is required for a certain layout. Garber and Grunes (1992) describe a typical layout task, where somebody needs to pick some images from a huge repository. Similarly, the growing amount of personal digital image collections (like Flickr (Flickr, 2006) or Picasa (Google, Inc., 2008b)) could benefit from a CBIR system. Another suggestion is the use of CBIR techniques to retrieve copyright infringements (Schietse et al., 2007). Further more, in medical environments, the use of 3D body scanners with high resolution results in an immense amount of visual data (Müller et al., 2005). Space agencies and GIS companies taking pictures with high-resolution cameras on satellites also produce large amounts of pixel images, depicting planet surfaces, surface features and other content (Blaser & Egenhofer, 2000, Fatto et al., 2007). Even more general cases in MIR have been studied, such as the powerful supporting tools in the retrieval of 3D models for engineering companies (Pein, Amador, et al., 2008) or similar sound files for musicians (Bosma et al., 2006). It follows that this key technology, CBIR, always plays a significant role in the application search engines, though the development of consistent theory in CBIR is still rudimental.

Many basic issues in CBIR have been collected by Eakins and Graham (Eakins & Graham, 1999) in 1999. The most severe problems of image retrieval identified in their report remain unsolved. The key is "bridging the semanticgap" between low-level image content (pixels, as seen by machines) and its high-level meaning (semantics, as seen by humans) (Zhao & Grosky, 2002). Though many new technologies and methods have been continually improving the quality of MIR (Wang, Boujemaa, & Chen, 2007) and several real-life applications have been developed in many areas. A summary of recent challenges in MIR is provided by Lew et al. (Lew, Sebe, Djeraba, & Jain, 2006).

MIR systems mostly apply the fv paradigm because the documents to be retrieved are typically very large. Further, the information contained in each document exhibits a considerable amount of redundancy and fuzziness. For example, several default colour space models in computing are mapped to 24 bits which corresponds to 2^{24}=16777216 different colours for each pixel. Modern hardware is capable of generating images with a resolution expressed in megapixels. Therefore, a single image can easily contain millions of pixels, encoding some real or synthetically generated information. Matching any set of two high-resolution images directly would consume a very high amount computing resources, rendering a naive matching based retrieval for thousands or even millions of images useless with the currently available technology. To greatly reduce the retrieval complexity, MIR systems perform the most time consuming part of data analysis only once and generate a so-called "fv" for each document during the indexing phase. These fv are designed to contain a highly condensed piece of information representing the original data, without losing too much of the relevant characteristics.