Chapter 10
Soft Computing Techniques in Content-Based Multimedia Information Retrieval

G. Castellano
University of Bari, Italy

A. M. Fanelli
University of Bari, Italy

M. A. Torsello
University of Bari, Italy

ABSTRACT

Due to the diffusion of multimedia databases and new ways of communication, there is an urgent need for developing more effective search systems capable of retrieving information by specifying directly in user queries elements strictly related to the multimedia content. This is the main rationale behind the flourishing area of Content-Based Multimedia Information Retrieval (CB-MIR), that finds in Soft Computing (SC) techniques a valid tool to handle uncertainty and vagueness underlying the whole information retrieval process. The main reason for this success seems to be the synergy resulting from SC paradigms, such as fuzzy logic, neural networks, rough sets and genetic algorithms. Each of these computing paradigms provides complementary reasoning and searching methods that allow the use of domain knowledge and empirical data to solve complex problems. In this chapter, the authors emphasize the potential of SC techniques, also combined in hybrid schemes, for the development of effective CB-MIR systems. As an example, the authors describe a content-based image retrieval system that employs SC techniques in its working scheme.

INTRODUCTION

In the last years, the growing dissemination of digital information and the consequent need to improve the retrieval of this information have given rise to a strong interest for the research in Information Retrieval (IR), a field that is expanding along with the challenges related to the explosion of the Web and the development of information and communication technologies. The diffusion of new ways of com-
munication and information sharing has led to the so-called digital convergence. Today, the key role of multimedia documents combined with the tools offered by digital technologies have promoted the creation of multimedia databases characterized by a complexity level higher than traditional databases. This is mainly due to the need to equip the modern multimedia databases with integrated methods of management and access to heterogeneous documents through specific systems for indexing, search and automatic extraction of data representing the complex content of multimedia documents (images, videos, audios, …) which support the traditional manual systems of analysis and indexing term-based (terminological) of textual or audiovisual documents. Whereas in databases of textual documents a search based on key words encoding the textual descriptors extracted from documents seems to be sufficient and appropriate, in multimedia databases the assignment of textual descriptions to multimedia contents reveals to be very limited and subjective.

The retrieval approach for multimedia databases needs to be improved through the definition of queries based on multimedia content, including not only textual data but also visual and audio data. Today, to achieve a good precision level in the retrieval from a multimedia database, IR systems are directing towards the possibility to combine term-based queries with content-based queries. In this way, term-based queries could be conveniently used to preliminary select a portion of the huge quantity of documents that can be potentially retrieved from a multimedia collection and to hit the search according to data such as membership domains, types, classes, titles, authors, etc. Successively, content-based queries can improve and refine search results. In this way, the traditional textual database interfaces that allow the search on the basis of an index exclusively composed of keywords extracted from documents or also inserted in textual metadata are replaced with new interfaces that enable the formulation of queries in different dimensions, not only on the basis of textual key words but also specifying multimedia contents of the same nature of the information to be searched. In this way, it is possible to realize a search based on heterogeneous indexes composed of texts extracted from captions or spoken language, key images from a sequence, simple figures, melodies, shapes, colours and sounds, without excluding the importance of textual data, descriptive or classificatory, that can be related to aspects not specifically audiovisual of the document.

A wider criterium of IR has been defined reflecting the evolution of a more visual culture, leading to so-called Content-Based Multimedia Information Retrieval (CB-MIR) (Lew et al., 2006). In CB-MIR, each kind of digital document is managed, stored and retrieved through the elements of the language, or metalanguage, own of the specific digital object, hence referred to the content of the object itself. In the last years, CB-MIR has been investigated in a growing number of research projects and applications at international level, showing that various communities working in the field of multimedia document management are interested in the new chances of exploitation of archives and professional activities: from biomedicine applications to architecture and archaeology, from film and video collections to journalism, from cultural heritage to design, from education to e-commerce and entertainment, from GIS (Geographical Information Systems) to remote sensing systems, from medical sector to identification and surveillance systems. However, due to their augmented potential, CB-MIR systems require a more advanced consideration about the intrinsic characteristics of digital documents, the semantic and the structuration of the elements included in a document, the way of interaction and presentation of query results to the user. Hence, today, the content-based approach underlying CB-MIR, represents a border field arising many challenges that the current research has only partially addressed up to now.