Book Review

Mark T. Maybury (ed.) (The MITRE Corporation, Bedford, MA)
Intelligent Multimedia Information Retrieval. AAAI Press / The MIT Press, Menlo Park, CA / Cambridge, MA, 1997, xxxiii+478pp., $31.95, ISBN 0-262-63179-2.

Categories and subject descriptors: H.3.1 [Information storage and retrieval]: Content analysis and indexing - *Indexing methods*; H.3.3 [Information storage and retrieval]: Information search and retrieval.

General terms: design, human factors

Multimedia information retrieval is nowadays more of a Holy Grail than a reality. Large amounts of information are now available electronically in various non-textual forms, such as graphics, images, spoken or non-spoken audio, animations, video, and the very availability of this information corpora poses the problem of how to automate the location of precisely the information relevant to a user's need. When text is concerned, this is the subject domain of information retrieval, a discipline whose history goes back to the very birth of digital computers, and that has evolved, slowly but steadily, into a solid body of concepts, tools and techniques for tackling large corpora of text documents. But the transition from the textual medium to other media poses far more problems than can be solved by merely switching from the text processing techniques developed by statistical linguistics and natural language processing, to the techniques for the analysis of non-textual media developed within digital signal processing. The key problem lies with the fact that central to information retrieval, irrespective of the considered medium, is the quest for documents whose semantic content matches the semantic content of the user's information need, and that natural language is, for humans, the most immediate and natural means of conveying content. Because of this immediacy, the very task of detecting the "content" of a natural language text, although still an unsolved problem, is much more manageable (i.e. admits of far less unsatisfactory approximate solutions) than the problem of detecting the "content" of a picture, a task that is even problematic to define, let alone to solve. Although "an image is worth a thousand words", detecting what it is worth is not so immediate.

It is only natural then that the various problems related to the detection of meaning within non-textual media should be attacked by means of a combination of tools borrowed from the various disciplines that revolve around the analysis, synthesis, and management of non-textual documents. *Intelligent Multimedia Information Retrieval* reports on a number of research efforts aimed towards the design of tools for indexing, searching, summarising, presenting, and browsing multimedia artifacts, achieved by the application of techniques from fields as diverse as digital signal processing, computational linguistics, human-computer interaction, scientific visualization, and still others. The 23 papers contained in this collection extend the basic picture of textual information retrieval not only by addressing new types of data (images, graphics, audio, speech, video) and non-classical types of information seeking behaviour (browsing), but also by tackling the novel problems that accrue from the co-existence of the various media, such as architectural and evaluation problems.

The title of the book is, unfortunately, slightly misleading, as it conveys the idea that the topic dealt with is the tackling of multimedia information retrieval *by artificial intelligence techniques* (an impression reinforced by the fact that this book originated at a satellite workshop of IJCAI95). Artificial intelligence plays, however, a decidedly minor role in this book, unless one wants to gather under the (admittedly hospitable) umbrella of AI a number of disciplines, such as human-computer interaction, speech recognition, computer graphics and scientific visualization, computational linguistics, that lead instead a happy life of their own. On the contrary, key subdisciplines of AI, such as knowledge representation and reasoning, machine learning, reasoning under uncertainty, planning, are either marginally present or absent *tout court*. This does not mean that the book is disappointing, though, as this is instead likely to be an indication that AI is not yet contributing massively to multimedia information retrieval, and that other fields are making instead a stronger impact on it. Should we conclude that most of the techniques that have been developed within AI proper, although relevant in principle, are not yet ready to scale to the problem sizes of nowadays multimedia information retrieval? Possibly. The application of AI techniques to information retrieval in the textual case has a history of mixed successes, and only recently have a few selected subdisciplines of AI (machine learning, to name one) made a durable impact on it. It is thus possible that we are seeing a rerun of past episodes, and that the application potential of AI for multimedia information retrieval has yet to fully emerge.

However, letting aside the word *intelligent* in the title, this book gives quite a faithful picture of multimedia information retrieval, of its various aspects, issues, problems, and shows that the field is nowadays at the center of a series of concentrated efforts from numerous disciplines. One of the aspects that makes this field exciting, as it emerges from this book, is precisely its multidisciplinary aspect. This aspect is also one of the strengths of the book, in that it succeeds in bringing together quality material by authors who, because of their largely different backgrounds and interests, tend otherwise to publish on widely scattered conferences and journals.

Karen Sparck Jones starts her foreword to this volume by saying that "Intelligent multimedia information retrieval is an exciting idea. With this timely book, displaying the state of the art, the reader can judge how far the idea is a reality." The overall picture that results from this book is of a discipline in tumultuous development, in which the urge for effective solutions is spawning new concepts and novel applications of consolidated ones, but one in which, at the same time, open problems still outnumber consolidated solutions. Unifying key concepts are still needed for multimedia information retrieval to blossom, and they will emerge only as a result of the joint effort of many contributing disciplines. This book makes a valid contribution towards this goal.

Fabrizio Sebastiani
Pisa, Italy