On the role of logic in image retrieval (Abstract)

Carlo Meghini, Fabrizio Sebastiani and Umberto Straccia
Consiglio Nazionale delle Ricerche
Istituto di Elaborazione dell'Informazione

Introduction

Image retrieval is a relatively new discipline, which has impetuously developed in the last few years. A distinguishing feature of this development is its simultaneous but separate arising in several sectors of computer science, such as Pattern Recognition, Vision, Information Retrieval, and Databases, to mention the most relevant. This fact reveals that there are many different aspects involved in image retrieval, each requiring a specific background and methodology to be successfully tackled, and also that there may be complementary approaches to the same problems, not only within the same discipline (such as different index structures for image data), but also cutting across different disciplines (such as similarity- *versus* semantic-based image retrieval). This richness of paradigms, methods and systems may, on the long run, become a negative attitude, resulting in duplication of effort and loss of potentiality, ultimately prone to slow down progress.

In connection with this problem, we discuss the role of logic in the image retrieval endeavor, arguing that logic has a twofold potentiality. First, it can be used to model *specific forms* of image retrieval. Second, it is a most natural candidate to express a *general* model of image retrieval, encompassing in a unique, formally specified and well-founded framework, the many forms of retrieval that have been studied so far in the areas mentioned above.

Logic for modelling specific kinds of image retrieval

There are two kinds of image retrieval that can be most suitaby approached via logic: that based on the *semantic content* of images, and that based on the *spatial relationships* holding amongst the objects in images.

As far as semantic content-based image retrieval is concerned, the basic question that an image retrieval model should answer is not *how to grasp* the content of an image from its pixel representation, but rather *how to represent and use* such content in order to achieve an effective retrieval functionality. Mathematical logic has a precise answer to this question, by providing (a) languages for representing facts, and (b) an inferential relation (namely, logical implication) for deriving implicit facts from given premises. The suitability of logic stems from the facts that (a) image contents are just facts, and (b) logical implication captures the essence of question answering. In addition, logic is also best suited to the representation of other information that may be very important in retrieval, such as the lexicon of the language being used as content representation medium, and information about the application domain, or *domain knowledge*.

As far as the retrieval of images on the basis of spatial relationships is concerned, a principled approach to the problem would be rooted in a suitable theory of space. Recently, several such theories have been investigated with computational purposes in the context of qualitative spatial reasoning, an area of knowledge representation that aims at studying the common sense view of space. These systems vary depending on the selected ontology (*e.g.* points, points and lines, regions), topology, orientation, distance and size. But they form an enormous body of knowledge from which to draw when specifying systems using spatial properties of objects for retrieval. Clearly, the inferential needs of an image retrieval system are quite simpler than those, say, of an automatic theorem prover for topology. However, models that are firmly grounded on solid theory are usually best adopted as starting points towards the design of effective systems.

Logic for modelling image retrieval in general

Along with the two kinds of image retrieval that have been discussed so far, there are other kinds, the most notable of which is similarity-based retrieval. Similarity-based retrieval is diversified in many sub-kinds, depending on the aspects of images which are lent to retrieval (color, shape, texture), the features used to capture the chosen aspects, the method used to compute such features, the metric used to assess similarity. Expressing this kind of retrieval as a logical inference would bring no particular insight into the subject matter, as the inference that is carried out by a similarity retrieval engine is the least apt to be captured by logic.

However, if one wants to express in a single model all kinds of image retrieval seen so far, then similarity-based retrieval ought to be captured in some way, and the a most appropriate way, we argue, is by means of procedural semantics. That is, the model of retrieval is endowed with two sets of predicate symbols: *mereological* symbols, denoting image parts (such as regions) and their attributes (such as the shape of a region, or its color); *similarity* symbols, denoting similarity criteria between images as a whole or between their components (color histograms, texture) and attributes (shapes). The semantics of these symbols would be given on the basis of an *image algebra*, with two kinds of operations, strictly corresponding to the two kinds of symbols, and giving semantics to them.

Overall, this approach permits to define a logical model of image retrieval, comprising descriptions of the semantics of each image, symbols for denoting image parts and their similarity, and any other feature that may be relevant for the application, such as object spatial relations. A retrieval request would be a formula of the same logic, selecting images on the basis of logical implication.

Uses of the model

The model has both theoretical and practical uses. On the theoretical side, the model unifies a so far disarticulated field of scientific investigation and also sets up a formal framework within which theoretical studies can be carried out. On the practical side, the model is an effective communication medium between two actors of the development of image retrieval systems, namely the user and the system designer. As a tool for rigorously defining in abstract terms *what* retrieval functionalities are offered, it can be used for the specification of operational requirements by the user. If the chosen logic, as many logics, is executable, the specification can also be tested, thus permitting verification. Then, the identification of a corresponding design must be performed, from which implementation proceeds.