PROJECT MUSE®

Image Indexing and Retrieval: Challenges and New
Perspectives / Indexation et repérage d'images : défis et
nouvelles perspectives

Elaine Ménard

Canadian Journal of Information and Library Science, Volume 34,
Number 3, September/septembre 2010, pp. 245-248 (Article)

Published by University of Toronto Press

DOI: <https://doi.org/10.1353/ils.2010.0002>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/390924>

Image Indexing and Retrieval: Challenges and New Perspectives

Indexation et repérage d'images : défis et nouvelles perspectives

Elaine Ménard
School of Information Studies
McGill University
elaine.menard@mcgill.ca

The Web constitutes a gigantic source for the image hunter who is looking for either illustrative or learning material. However, given the possibility of multiple interpretations of the visual resources available, many difficulties tend to complicate retrieval. Current research on image indexing and retrieval focuses on numerous areas, but from machine methods for image searching to cognitive aspects of image perception and understanding, there are still a number of theoretical and practical questions to be addressed.

Over the years, information specialists have learned to look at images and to collect and annotate them, in order to provide the “best” access possible. Nevertheless, it has certainly not been an easy journey. One mistake to avoid making is to treat images as text or speech. Another is to think independently of any text. The image, they say, is misleading. True, but is text so different from images? The caption may lie more easily than the image. The image is risky, but the image is also indolent. Nothing mobilizes more memory or puts in motion more neurons than the recognition of an image. Nothing is more difficult to understand and manipulate. Criticism of the image is still very rough and has been delayed even more by the approach of searchers who believe that the image can be reduced to a few words. The image has a mind of its own. And this is why information specialists have a special responsibility when it comes to images. We cannot search out any images without knowing where they come from, who made them, why, and for whom. This is most certainly why image retrieval is still a complex exercise. It must respect the fact that an image has the sense that everyone gives. Image and language are incompatible but complementary.

Another difficulty is that the image is never a solitary object. It lives in colonies, called collections, databases, and digital libraries. The common user now employs image search engines in everyday life. The relevance of these engines is often questionable because the design of such tools is hampered by two major obstacles. First, we encounter an enormous discrepancy between text and image. We are used to searching from textual data. Almost all search tools at our disposal use text as a reference search. So, whether we are looking for a definition in a dictionary, information in a database, or a number in a directory, the text will always be the element initiating the search. Yet, for an image, there are many other search methods, such as search by visual content (with the colour, the form, and the texture, for instance). These search modes that use the content of the image, content-based image retrieval (CBIR), rather than the context are comparative forms (by similarity). If CBIR systems represent an interesting alternative, their integration in the Web is still very hypothetical. Meanwhile, image search engines generally propose the use of a textual query to access a visual resource. But *a priori*, the image does not contain any text. The solution is to combine the text in the form of keywords to an image to be retrieved. There are three kinds of solutions: the information is contained in the image itself; the information is provided by the textual environment of the image; or the information comes from an external database associated with the image. A search engine like Google Images will be satisfied with the second solution. Therefore, the text content of the page that displays the image will reference it. This solution is convenient to index a significant number of images (more than 1 billion images of all kinds on Google Images), but it also raises obvious problems. On the one hand, an image is not always used in the context of what it represents; on the other hand, search engines often use the file name, which is generally part of the surrounding text associated with images. Nevertheless, this is extremely risky, especially when you consider that most images directly extracted from digital cameras carry a code name defined by the device itself and that has absolutely no meaning or use for retrieval.

Most image databases (to differentiate them from image search engines) use the third option. Each image in the database is submitted to an indexer who manually or semi-automatically enters keywords associated with the photo. The list of keywords defined for the photo is then a static list. These terms can be extracted from controlled vocabularies (of a general nature or created especially for image indexing). The use of folksonomies, or tags, is now also considered another possible avenue.

Users can freely assign keywords to documents, including visual resources, in order to improve the retrieval. The assignment of these “free” keywords has gained popularity over the years. This type of indexing adds a new layer in providing a particular indexing vocabulary, sometimes very different from the vocabulary that indexing specialists would assign in the same context. The introduction of a more collaborative dimension in the image description, based on the instantaneous perception of the image, is often considered to be more in direct link with the way the common user perceives images. Another significant advantage of free tagging is the possibility of assigning keywords that come from one language or several foreign languages. Everyone can indeed contribute to the improvement of the image description by adding tags in his or her own language and these keywords will be taken into account in future searches. Actually when we perform a Web image search query with our preferred engine, the principle is always the same. We enter one or more keywords and the engine returns a list of links to pages that are supposedly relevant to the subject of our search. However, it always takes a certain skill (or habit) to perform an efficient and satisfactory search. And this is not the end of the difficulties. When we use a textual query with an image search engine, it will indeed take a few seconds to distinguish relevant from irrelevant materials, making the retrieval task extremely difficult and unrewarding. Many factors can overwhelm the image searcher who is trying to retrieve images, including an overload of available images or images indexed with a vocabulary that is incomprehensible or too specialized to be useful, and search tools that offer well-designed functionalities but are not particularly adapted to the current behaviours of image searchers.

Throughout the world, searchers continue their quest to develop and promote issues and research involving visual resource representation. It is with great pleasure that the *Canadian Journal for Information and Library Science* presents this special issue on image indexing and retrieval. The study by Youngok Choi and Ingrid Hsieh-Yee examines the characteristics of user search query terms and query formulation strategies in order to determine how effective the Library of Congress subject headings are in matching image queries in an online catalogue and how effective subject description notes are in matching image queries in an online catalogue. Hsin-liang Chen, Thomas Kochtanek, and Rick Shaw test the use of metadata elements for photojournalism images on the Web, collaborating with end users on organizing and accessing photojournalism images, and collecting evidence to support system enhancement of the

Pictures of the Year International website test bed. Brian Stewart proposes an exploratory investigation into indexers' understanding and indexing of the subject content of historic photographs that contributes to a larger investigation into subject access to historic photographs. Finally, Diane Neal investigates the communicative roles played by the text, image, and social interaction present in high- and low-relevance ranked Flickr photographic documents with an emotion-based tag.

The four articles included in this special issue illustrate the vastness and diversity of research in image indexing and retrieval. These few lines of thought also offer a list of emerging and recurring themes: organizing multidisciplinary information, increasing exchanges, pooling resources and skills, reflecting on purposes, defining tasks, developing tools, building bridges with industry to assess systems, creating new interfaces and thereby working on methods of indexation, and increasing training and awareness. In conclusion, we hope this special issue will highlight the importance of carrying on the work on image indexing and retrieval, given the immense scientific, cultural, and economic significance and value of visual resources.