Multilevel Access to our Cultural Heritage (MACH): Egon L. van den Broek en Thijs Kok

Most search engines utilize keywords to retrieve information (e.g., Google). However, when one accesses a specific domain such as art and is a layperson in this field, how to retrieve information? Without the knowledge of artist's name, the period he lived in, the art movement he was affiliated with, the material he used, and the name of his pieces of art, a text-based query cannot be conducted. All this knowledge is not required to search using the content of the image; i.e., by way of using an example image as the query for the system. The use of content-based querying enables laypersons in the field of art to access the online collection of the Rijksmuseum.

The newly developed, online Multimedia for Art ReTrieval (M4ART) system uses both text and content-based techniques, which provide entrance to the digitized collection of the National Gallery of the Netherlands (the Rijksmuseum). The current online system of the Rijksmuseum is text-based and requires expert knowledge concerning the work searched for, else it fails in retrieving it. M4ART extends this system with querying by an example image that can be either uploaded or its address can be provided to the system or can be selected through browsing the collection. The global color distribution and (optionally) a set of texture features of the example image are extracted and compared with those of the images in the collection. Hence, based on either text or content-based features, the collection can be queried. For both manners of querying, a standard and an advanced user interface is available. The standard text-based interface provides the means to enter a number of keywords. The standard content-based interface provides the means to select an image of interest. With the advanced text-based interface access is provided to specific database fields, terms can be included or excluded, and AND or OR operators can be used in combination with the keywords. The advanced content-based interface lets the user select a vector model (i.e., a combination of a color space with its quantization scheme), a distance measure, and provides the choice to include texture or not. When texture, next to the global color analysis, is included for the analysis of the image material, parallel-sequential texture analysis is conducted. Another feature of M4ART is that its matching process can be inspected. With the latter feature, M4ART not only integrates the means to inspect collections by both experts and laypersons in one system but also provides the means to let the user to understand its working. These characteristics make M4ART a unique system to access, enhance, and retrieve the knowledge available in digitized art collections.

The M4ART system is build using PHP, MySQL, HTML, and Javascript and is validated using the W3C validator. Consequently, M4ART runs on all platforms within all modern browsers (e.g., Internet Explorer 6+ and Firefox). In the near future we extend it with mechanisms of relevance feedback, more advanced content-analysis in general, and with artist profiling by way of visual characteristics of his paintings. In parallel, research is conducted toward common dimensions of description in classic annotations and content-analysis of the art material.

Acknowledgments: This research was supported by the Netherlands Organization for Scientific Research (NWO) under project numbers 634,000,001 and 634,000,018. Further, we thank Xenia Henny and Kees Schoemaker for their cooperation. They provided us the database of the Rijksmuseum. Theo E. Schouten, Eduard Hoenkamp, Peter J. Petiet, and Louis G. Vuurpijl are gratefully acknowledged for their support in various ways.