



Working Without Words - Image based Retrieval as a Model to overcome Language Hurdles

Jan Simane
Head Librarian
Kunsthistorisches Institut in Florenz
Max-Planck-Institut
Florence, Italy

Meeting:

79 — Tackling the challenges of multilingualism in the arts: catalogues, databases, digital collections and other material in the global context — Art Libraries Section

Abstract:

In art history images are the primary medium of interest. In a scientific context, multi media objects, including digital images, are represented in form of an intellectually or automatically generated semantic code. Consequently, the wide options for retrieving information in digital networks are primarily text-based. Using an image database of Florentine coats of arms an experimental feature has been developed allowing non-verbal retrieval of digital images. This is based on the comparison of structural elements of portions of the image. This is an example how problems of multilingualism can be resolved, albeit in a very special case.

Studying art and examining art works is a strange discipline. Visual arts are made for the eyes; the forms, the material, the styles are complementary elements of figurative expressions. And the artists address their messages first and foremost to the visual sense of the spectator. But all our information and knowledge, our descriptions and valuations, all types of systematization, classification and any approach of analyzing art and artistic expressions is based on words. The two different media, the non-verbal nature of the object on one hand and its verbal nomination on the other represent the specific dualism of our system of encoding visual phenomena for scientific purposes. As a consequence of this not particularly surprising fact we are accustomed to deal with verbal representations of visual objects as the only way to name, to differentiate and to contextualize what we see. This fact has obtained relevance in context of the availability of masses of images on the internet or specific data bases respectively, when efficient retrieval instruments became a matter of importance. Not only millions of new photographs are going online daily, but also whole collections of analogue images have been scanned and restored in extensive databases. The access to this important material now is much easier and quicker and we do also use the manifold options of cross-linking and networking in order to join together collections located in different places and to make new findings by getting more and different results thanks to

simultaneous meta-search tools. However, in every case, the search tools are based on textual and therefore “limited” meta-data. These limits are a result of heterogeneous ways of description or, in the age of global communication, due to the mere plurality of languages. The above mentioned incongruence between the visual structure of the image and its verbal representation, the so called semantic gap, is, under these circumstances, impossible to avoid. This problem is unknown when we deal exclusively with texts. Here, the encoding of the single characters is the decisive condition to find identical and diverse elements. In this respect the structure of a term, which is a sequence of characters, and its digital code are a unit. Could this principle be transferred to images? The consequence would be a non-verbal, form-based retrieval, a semantic identity between search and aiming visual object. This seems to be an auspicious imagination. And an interesting model to hurdle the problem of multilingualism in verbal search procedures.

In the following I will present you a project which is running at the library of the Kunsthistorisches Institut in Florence and in which the above mentioned solution of a non-verbal, image-based retrieval model has been realized. Let me first describe briefly what the Kunsthistorisches Institut is today. The Institute’s activity is dedicated to basic research in the field of Italian art history. Even if located in Florence (Italy), it is one of the leading German research institutes, specialized in Italian art, founded at the end of the 19th Century. Its central ‘laboratories’ are the library with a collection of about 330.000 volumes and a photographic archive with about 600.000 images, both primarily focused on Italian art. The Library of the Kunsthistorisches Institut does not only provide specialized literature, but also important primary sources for the study of Italian art history. Furthermore, the library is also active in a rather experimental area. And this is the topic of this presentation. The Library owns a collection of about 2,800 colour drawings of coats of arms of noble or important Florentine families, confraternities and hospitals completed by an appendix, which consists in an index of names. The drawings, created by an anonymous draughtsman from the 19th century, were acquired in 1901 and became an important instrument of reference. Each coat of arms is drawn on a ca. A5 sized paper sheet with the name of the family written below. For many decades, scholars who wanted to identify a coat of arms depicted on artworks or related to architecture checked this significant collection. As we all know, sometimes coat of arms give evidence not only for the correct dating but also to identify the owner or patron of an artwork or even a building. As an ‘emblem’ on a tomb or on documents or “charters” a coat of arms is also an authentic and reliable witness, which may contain important information on the status and the social position of a personality or/and the family. Although the description of coats of arms does not require a broad vocabulary, the number of different variants and combination of terms is rather high and follows a complicated system. Coats of arms are no artworks, but insignia. They have to be decoded precisely in an unequivocal manner. For this reason, the heraldry developed a very concise terminology, completely different from our everyday language as well as from the language that we use for the description of artworks. For non experts searching for coats of arms in handbooks or indexes is often difficult, especially when the owner of arms is unknown. Our *stemmario*-project – the term refers to the Italian word *stemma* and means something like “collection of arms” – has the purpose to facilitate the search tools in a new database and to open it to an international circle of users, which may not be familiar with the technical terms and their specific variants in different languages.

The first step was the complete digitization of all 2.800 sheets. They have been scanned as TIFF-files and the images have been restored in a database. As a second step we described in German language all coats of arms on the basis of the standards of heraldic vocabulary. Additionally we also conceived a category scheme, a tool which permits a systematic

research on the basis of single elements represented on the coat of arms, even if the owner is unknown. So far, the project provides a sort of conventional image database, which has to deal with the already mentioned disadvantages of verbal indexing and of the complexity of the terminology. Finally, we started a more experimental phase in order to eliminate or at least to limit the risk of an incomplete number of hits. For this purpose we took advantage from our excellent contacts to the network of multimedia information system of the Istituto di Scienza e Tecnologie dell'informazione of the Consiglio Nazionale di Ricerca at Pisa (Italy). From the beginning our colleagues in Pisa had been interested in our *stemmario*-database as a sort of experimental ground and we needed their knowhow and the technologies developed in Pisa to generate practical ideas in order to resolve the problems mentioned previously. The result is a merely image based research tool without any verbal description. Its functional principle will be explained briefly below: In the case of a non verbal, image based search for objects, all necessary information has to be extracted directly from the image objects. This happens in form of an automatically conducted semantic object analysis able to determine the formal characteristics of an image surface. Recently, the software possibilities to isolate elements from images and for the indexing of image structures have been improved and refined. There are two different options for the information retrieval. The first and the most common one, which was developed during the 1990's, is based on the paradigm of a general reconciliation with an example, that has been chosen as a starting point. The software program is searching for objects with similar total structures. The results, as illustrated by an example which represents one of the stages during the developing process of our *stemmario*-Website, are not satisfying and cannot be used for art history research in general and particularly not for the heraldry. Another possibility is to search for similar images, as it is recently used by Google Images. In this case, verbal metadata are connected to an optical similarity search tool. The results reveal a similarity in respect of content, but from a formal point of view, they have nothing in common. Both methods are not able to overcome the semantic gap in the case of our collection of coat of arms.

The development of applications, that permit to search for single elements in a picture, signifies a substantial progress, because in this case the user can choose by himself which elements of the image are more and which are less important for his search. Isolating the relevant parts of image objects by an automatically generated analysis of the semantic structure of the image refines the search and leads to more precise results. Because of this advantage we decided to choose this method for the database of Florentine coats of arms. The relevance of coats of arms as images with a meaningful dimension is based on the repertory of symbols and their arrangement on the shield. Since the same element exists in many variations, the description and classification of coat of arms essentially consists in nominating the single elements. Therefore, the identification of arms depends on the correct classification of the all elements depicted on the shield. The starting point of the semantic analysis is the segmentation, or the decomposition of the whole object in homogenous entities (or units). Our partners in Pisa decided to use a double stage method of segmentation: the first step consists in the division of the surface according to its colour-morphology. The segments are not necessarily divided in compact zones of one and the same colour, but they are generated automatically according to the intensity of a colour, independently from their position on the shield. Afterwards the isolated fields are subject to a second segmentation, based on the so called K-Means-Algorithm, which represents the isolated objects as clusters; in our case that means a representation as closed and compact subdivisions, reflecting not only the different intensity of the various subdivisions, but also their position on the surface. The results of the segmentation are shown to the user as image objects. Now, the user can pursue his research

strategy by deciding about the relevance of the single elements. By clicking on the single segments he can search in the database for other elements with a similar structure.

To illustrate the whole on an example: the so called Capo d'Angiò chief is a red label (a horizontal strap with three or sometimes even four pendants, usually called points or more rarely drops). This heraldic symbol was introduced by Charles of Anjou in 1246. As a gulf he defended the interests of the pope battling against the Staufer in Sicily. Consequently his heraldry symbol appears on the coats of arms of numerous noble families in order to stretch their loyalty to the family and its politic position. A refined search for an image symbol isolated through segmentation brings to more than 25 hits of arms with very different and complex shield compositions, but they all have been brought together in a non verbal way because of the semantic context, that they all have in common as described above.

Conclusion: undoubtedly the method of searching for images on the basis of isolated segments from other images is a very particular solution for special cases like the formally far reaching homogeneous heraldry. In collections of very heterogeneous images the results would certainly be much more disappointing. However, the project has revealed a way to avoid complicated verbal descriptions of coats of arms and to get useful results even without any knowledge about their historical or semantic relationship. And no language specification is needed, of course. With regard to information science the project succeeded to close the semantic gap mentioned above by isolating all constitutional elements of an image, by representing them as images and by surpassing the traditional secondary medium of verbal description. Since the mass of available images for the arts is growing steadily and related data bases appear more and more in supranational if not global contexts, the presented solution could help to overcome the language problems. It is obvious that such fully automated methods will play a decisive role in future.

About the author:

Born in 1959. Studied the history of art, archaeology and philosophy at the Universities of Heidelberg and Vienna. Master of Art (1988) and Ph.D. (1991) at the University of Heidelberg. From 1991 to 1994 assistant at the Hessisches Landesmuseum in Darmstadt. 1994/1995 University for applied sciences (Göttingen / Frankfurt am Main): master in scientific librarianship. From 1996 onwards head librarian at the Kunsthistorisches Institut in Florenz, Max-Planck-Institut. Member of the Steering committee of the international project artlibraries.net. Member of the Standing Committee of the Art Libraries Section of IFLA, Chair since 2009.