Florentine Coats of Arms on the Web:

Experimenting retrieval based on text or image content

Fabrizio Falchi [1]; Fausto Rabitti[1];
Werner Schweibenz[2]; Jan Simane[2]

[1.] ISTI-CNR, Pisa, Italy
[2] Kunsthistorisches Institut in Florenz, Florence, Italy

Abstract
The paper describes two different ways of accessing a collection of Florentine coats of arms, which was digitised and made available on the Web. One way is a traditional approach: textual description and indexing. As heraldry uses a specific and complex terminology the textual approach seemed alright for experts but insufficient for laypersons. Therefore an additional retrieval mode was chosen which is based on the visual specification by users of the content of images representing the coat of arms. This access mode allows users who are not familiar with heraldry to access the collection.

Introduction

Since 1901 the *Kunsthistorisches Institut in Florenz (KHI), Max-Planck-Institut*, is in possession of more than 2,800 coloured drawings of Florentine coat of arms. This unique collection is a very useful instrument for identifying coat of arms on palace facades, tomb monuments, altar pieces etc. The high importance and usefulness of this collection were the motifs for the decision to digitize the material and make it accessible on the Web (the current version is a prototype and prove of concept). Another reason was to improve the access to the collection which was indexed only rudimentary by index cards. In addition to traditional textual indexing of the coats of arms we wanted to try out retrieval based on image content as we considered the heraldic terminology to be a problem for laypersons.
In this application context, we decided to explore the possibility for the users to search the content of images representing the coat of arms using visual specification of this content. We believe that a very important direction towards this kind of solutions is to support *content-based retrieval* to the image database through *feature based similarity access*. A feature (or *content-representative metadata*) is a set of characteristics of images, such as color, texture, and shapes. Similarity based access means that the user specifies

some characteristics of the wanted images, and the system retrieves the most relevant images with respect to the given characteristics, i.e., the images *most similar* to the query. Such approach assumes the ability to measure the distance (with some kind of metric) between the query and the database images. This search paradigm:
generalizes the information retrieval approach for textual archives, where search is based on key terms and the retrieved documents can be ranked with respect to the relevance of these key terms in the document text;
generalizes the web searching mechanism, as used by today's commercial systems, where searching is based on heuristics on links and domains and ranking is based on adapted text information retrieval techniques.

Description of the Material

The collection consists of two parts, on the one hand the coat of arms of Florentine families and on the other hand those of Florentine churches, hospitals and confraternities. The biggest part of the collection comprises coat of arms of noble families, 2,512 colour drawings and 22 sketches. These drawings, which were made by an anonymous draughtsman in the 19th century, were acquired by the Institute's library in 1901. A much smaller number of drawings refers to coats of arms of Florentine churches, hospitals and confraternities. The 317 colour drawings were created by Otto Wenzel, the KHI's librarian from 1902 to 1915, based on the Priorista di Luca Chiari.
In the library's inventory of 1901 the drawings are described as „Sammlung von Wappen florentiner Familien u.a." [Collection of coats of arms of Florentine families etc.], while the other part is mentioned first in the inventory of 1909 as „Sammlung von Wappen florentiner Kirchen, Hospitälern und Bruderschaften nach Priorista di Luca Chiari" [Collection of coats of arms of Florentine churches, hospitals and brotherhoods according to the Priorista di Luca Chiari].
Each coat of arms is drawn as a triangular shield in the centre of a sheet of light brown paper of ca. 16 x 23 cm; some of the sheets have watermarks. Most of the drawings are made by pencil and water colour, a few of them are only sketched with pencil. Below each coat of arms, the name of the family is written in black ink. Sometimes commentaries in pencil are added. The whole collection is made accessible by a handwritten card index. This index combines the names of the families with the corresponding heraldic elements. For decades, both parts of the collections were stored in three boxes, one of them containing the index. During the 1980s the sheets were stored in 65 black ring binders while the index cards remained in boxes.

The collection has been frequently consulted by scholars and researchers for identifying unknown coats of arms on buildings and artworks or for looking up a certain family's coat of arms. In this way it has become an important resource for art historians.

Reasons for the Digitisation

For several reasons, the KHI decided to digitise this collection and make it accessible on the Web. First, there were conservation issues as over time the drawings were battered to a certain extend due to the extensive use. Therefore it seemed reasonable to digitise the material and provide access to digital facsimiles instead of the originals. Then, there were restrictions in physical access. The original drawings are kept in a special room with restricted office hours. The goal was to increase the accessibility in-house and extend it to the public on the Internet. The third reason was to improve the access to the collection and the retrieval of individual coats of arms.

Project Description

The project started in 2005. The first step was to provide a formal verbal description (blazon) of each individual coat of arms in German, according to the international standards of heraldry. As heraldry applies a complex specialist terminology, the KHI decided to seek the advice of a specialist.

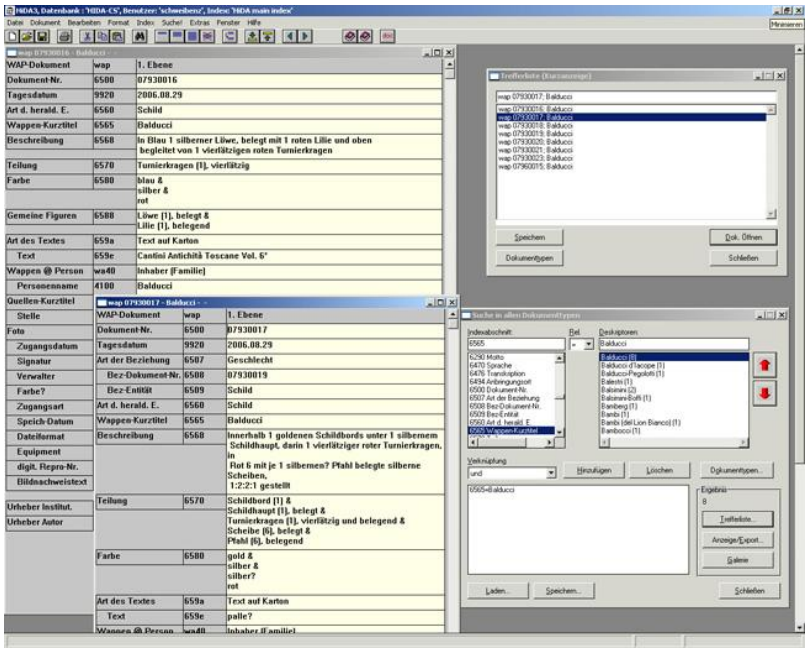

Figure 1. A heraldic blazon for a coat of arms in the HiDA database (family name: Balducci, heraldic charge: lion)

Harald Drös, Head of the *Research Institute for German Inscriptions* at the *Heidelberg Academy of Science*, took up the responsibility of supervising this work. The information from the heraldic blazon for each coat of arms was entered into the KHI's art information database (HiDA). Inside the database, a variety of indices was created for relevant fields such as family names, colours, heraldic ordinaries and charges, etc. in order to allow combined searches for the content (cf. Figure 1).
In parallel, all the coats of arms were scanned in-house. The master scans were manufactured on a high-end scanner in RGB colour, 16 bit colour depth, 600 pixel per inch. The height for each scan was approximately 5,500 pixel, the width approximately 4,000 pixel. This resulted in a file size of about 125 MB for the master. Each master was reduced to a working copy in lower quality with 8 bit colour depth and 300 pixel per inch for further use (e.g. print). For the display on the Web JPEGs with 350 pixel width are used.
The content of the art information database and the digitised images were integrated into an open-source web database (ZOPE) which was provided by the KHI's technical partner, the Fafalter GmbH in Düsseldorf. The same product has been successfully used for the KHI's digital photo library [Bieber, Schweibenz 2004, 2005]. In addition to the traditional text-based searches the KHI decided to explore new means of access. The idea was to analyse the pictorial content of the individual coats of arms and provide a query by image content. In order to implement this feature the KHI looked for a partner with profound competence in this field. This was the *Istituto di Scienza e Tecnologie dell'Informazione/Consiglio Nazionale di Ricerca* (ISTI/CNR) in Pisa.

Text-based Retrieval in the Coats of Arms Database

The text-based retrieval relies on the indices created in the art information database HiDA (Hierarchical Data Administrator) in German language. By entering information in certain HiDA fields a controlled vocabulary for different subjects is set up, e.g. for tinctures (colours) such as "gold", "silver" or "blue", for heraldic ordinaries such as "bend", "cross" or "chief" or for heraldic charges such as "lion" or "eagle". Other fields in HiDA such as the blazon information or the text information contained on the sheet of paper are automatically indexed full-text. All indices are transferred to the ZOPE database and can be searched by entering a term in the search field of the Web interface (cf. Figure 2) or in the advanced search (cf. Figure 3).

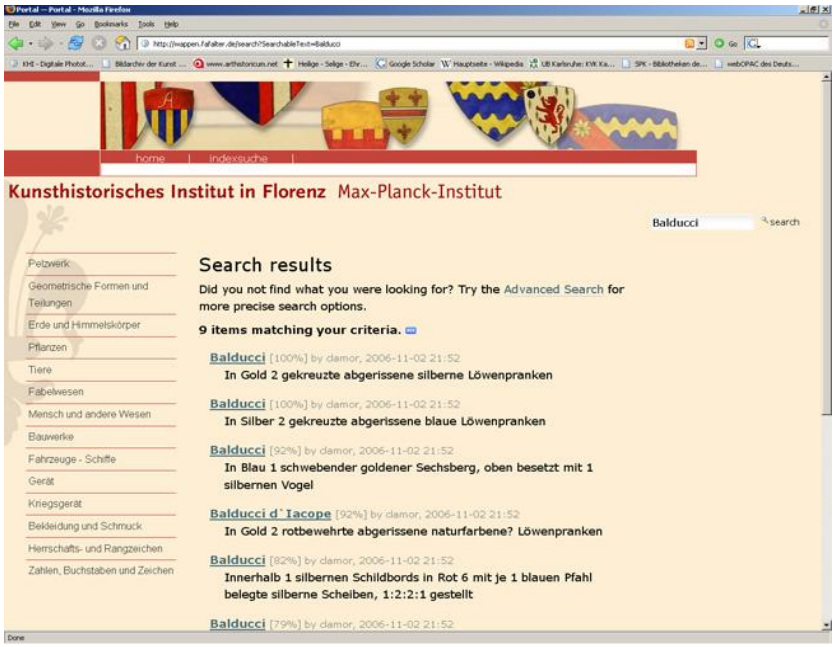

Figure 2. Live Search for a heraldic charge (lion)

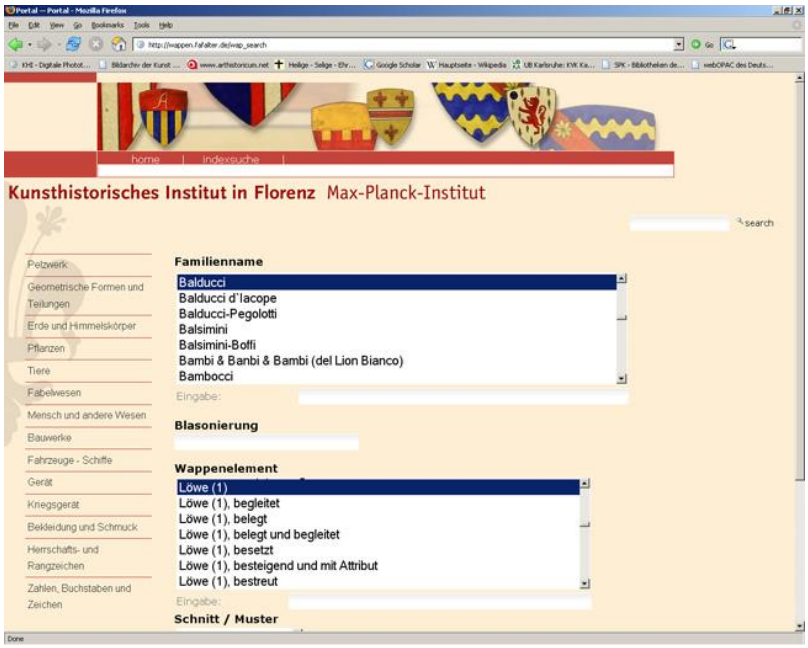

Figure 3. Combined search for family name (Balducci) and heraldic charge (lion)

As already mentioned above, the terminology of heraldry is very complex. Therefore one cannot expect the users to be familiar with all the specific terms. For this reason, apart from the simple and advanced search, a number of categories were added. These categories, based on a German heraldry book

called “Wappenbilderordnung”, for example “animals,” “plants” or “mythical creatures”, allow users to approach the coats of arms starting from very general terminology and getting more and more specific. Each category branches out, e.g. animals into mammals and birds, and finally leads to a list of family names with an accompanying blazon. The categories are based on the indices provided by HiDA and are set up by the Content Management System PLONE using a service being called “smart folders”. In this way users who are non-specialists can access the coats of arms in a convenient way without having to know or use heraldic terminology. Another way of access for non-specialists is the query by image content.

Content-based Searching in the Coats of Arms Image Database

We had different options in the implementation of the retrieval application based on the visual specification by users of the content of images representing the coat of arms.

The earliest and most common approach is the global query-by-example paradigm. It consists in retrieving images whose visual appearance is globally similar to a selected example image. Initially proposed by Swain and Ballard [Swain, Ballard 1999], it was then adopted by a vast majority of content-based image retrieval systems (e.g., IBM QBIC, VisualSEEk, Virage’s VIR Image Engine, and Excalibur’s Image RetrievalWare). However this paradigm has limited application in our case, since it would allow users to retrieve only coat of arms similar to a given one.

The partial query by example paradigm was introduced later on. This approach allows the user to explicitly select a visual component which is relevant for the query and retrieves images which contain a similar visual component. This approach proved to be more selective, hence more precise than the global query-by-example paradigm [Ma, Manjunath 1999]. This is really what is needed for our application: searching coat of arms images based on some contained visual element.

In our approach, bases on similarity search on images, is possible to apply the relevance feedback mechanism, inspired from text retrieval, in refining the image search [Huang, Rui, Mehrotra 1997]. Among the retrieved images, the user specifies the ones which are relevant and reiterates the search. By refining the similarity measure, the searched image can be reached more efficiently, taking into account the subjective preference of the user.

A basic decision in the approach, is the adoption of MPEG-7 standard to represent the visual metadata, i.e. features, extracted from the coat of arms images. MPEG-7 [ISO/IEC 15938], also called the Multimedia Content Description Interface, aims to cover the need for searching and retrieving

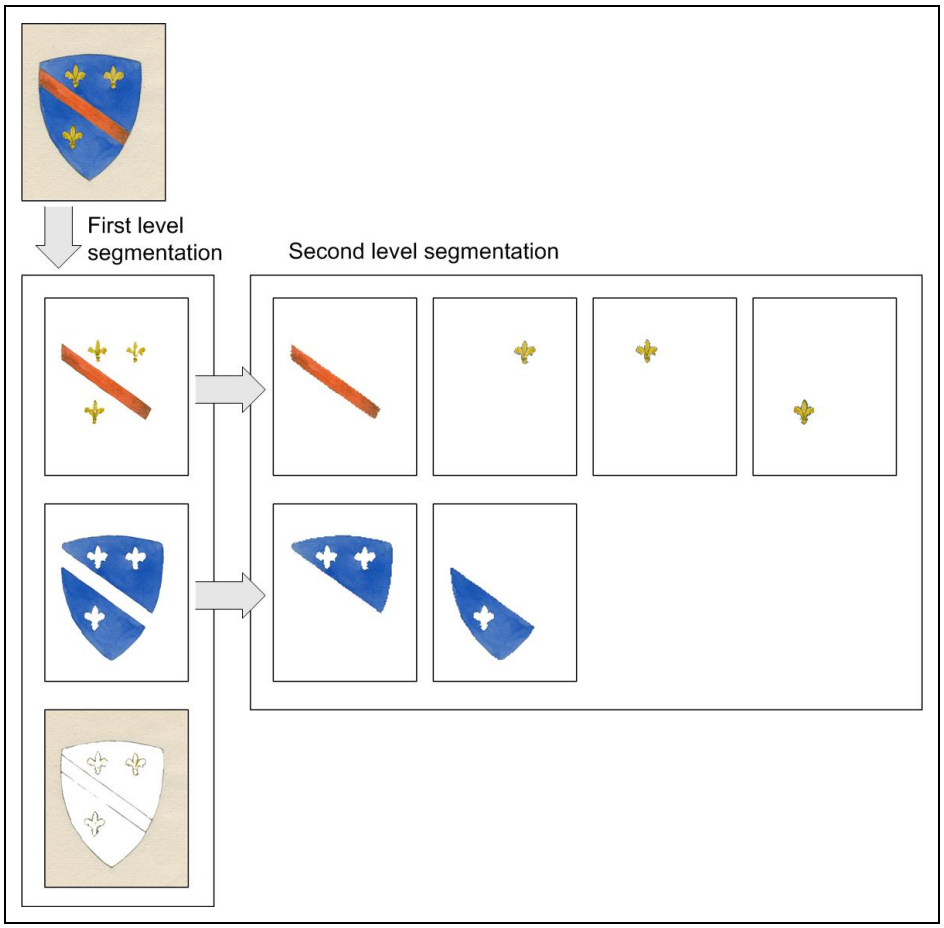

Figure 4. Two levels segmentation

multimedia information, by describing the content of audiovisual objects in a standardized way. The approach of MPEG-7 is to specify a set of standardized *descriptors*, that is, entities that contain the syntax and semantics of audiovisual content. A descriptor could for example refer to the color of an object by specifying the name of the attribute color (e.g. Color) and the type of the value (e.g. a string or three integer values for RGB colors).

Image Analysis and Segmentation

Before utilizing the images of the database in the retrieval process, it is necessary to analyze their content, identifying the relevant components. This process is called segmentation.

Image segmentation refers to partitioning the image into homogeneous regions. We decided to use two levels of segmentation using two different segmentation algorithms. Basically the first one divides the image in regions considering color gradient. The outputs of this first level segmentation are not necessarily connected regions of similar color. These regions are then

segmented again using another algorithm based on K-Means which is an algorithm to cluster objects based on attributes trying to minimize total intra-cluster variance. The outputs of this second level segmentation are connected regions. All regions obtained by both segmentation algorithms are used.
For the first step segmentation we used an implementation, made by Dorin Comaniciu, of the algorithm described in [Comaniciu, Meer 1997]. The technique is based on the mean shift algorithm, a simple nonparametric procedure for estimating density gradients.
For the second level segmentation we used a novel variant of the well known K-Means-with-connectivity-constraint algorithm (KMCC), a member of popular K-Means family, which is described in [Mezaris, Kompatsiaris Strintzis 2002, 2004]. The KMCC algorithm classifies the pixels into regions taking into account not only the intensity information associated with each pixel but also the position of the pixel, thus producing connected regions rather than set of chromatically similar pixels.

MPEG-7 Descriptors Extraction and Searching

Feature extraction was performed employing an application we built upon the MPEG-7 [ISO/IEC 15938] experimentation model [ISO/IEC 15938-6:2003] of MPEG-7. The software can extract all MPEG-7 image Visual Descriptors [ISO/IEC 15938-3:2003]. The same application has been used on MILOS [MILOS, Amato, Gennaro, Rabitti, Savino 2004] and particularly on the PhotoBook [Amato, Bolettieri, Debole, Falchi, Rabitti, Savino2006, Amato, Bolettieri, Debole, Falchi, Rabitti Savino 2006], which is a Multimedia Digital Library Application (http://milos.isti.cnr.it) we built over MILOS for online photo sharing.
For both whole coat of arms images and automatic selected regions we extracted 4 MPEG-7 descriptors: *ScalableColor*, *ColorStructure*, *ColorLayout*, *DominantColor*, *EdgeHistogram*, *HomogeneousTexture*.
Because of the fact that they are concerned about shape, 2 MPEG-7 descriptors were extracted only for regions: *RegionShape*, *ContourShape*.
Two more descriptors about color were used for whole images: *ColorStructure*, *ColorLayout*. These descriptors were not used for regions because they are meaningless for non rectangular images.
We now give a brief description of the MPEG-7 descriptor used. *ScalableColor*, *ColorStructure*, *ColorLayout* and *DominantColor* are all about colors. *ScalableColor* is a color histogram in the HSV Color Space. *ColorStructure* captures both color content and information about the spatial arrangement of the colors. *ColorLayout* represents the spatial layout of color images. *EdgeHistogram* describes the spatial distribution of five types of

Figure 5. Research interface

edges. *HomogeneousTexture* characterizes the properties of texture in an image. *RegionShape* expresses pixel distribution within a region. *ContourShape* is based on the Curvature Scale-Space representation of the contour.

The result of the extraction process is an XML document. The online application developed for this project performs a linear scan of the descriptors present in the document. Given the limited number of images in the database, just a few thousands, the use of an index structure for similarity search is not necessary and a linear search of the MPEG-7 descriptors is feasible. However, in MILOS we use M-Tree [Salembier, Sikora 2002] as an index structure for similarity search: this allows the system to be scalable, searching tens of million of MPEG-7 image descriptors in a few seconds.

For comparing the descriptors coming from two images, we used the functions proposed in [Swain, Ballard 1991]. In case the user want to search using a combination of descriptors, distances are combined summing single feature results using predefined (heuristically) weights.

Online image content search application

Once an image is selected (e.g., from the textual description database), the application asks the user to select the region (even the whole image) he wants to use as query searching for similar ones. Then first 10 results are reported in

order of decreasing similarity using *EdgeHistogram* and *RegionShape* that we found to be usually the most useful descriptors for this kind of images.
From the result page it's possible to select which descriptors to use for the search and then ask for the relative results. Clicking on a result region, the search is performed again using the selected one as query region.
Under each result region there is a small image of the entire coat-of-arm from which the region comes from. Using the Info link that can be found over each region present in the result page, the user can access the textual metadata inserted in the database.
The result page also reports, on the upper left, the query coat-of-arms and the region used as query.

Conclusion and Future Research

The project combines two different ways of accessing a collection of Florentine coats of arms, one being traditional and text-based, the other being experimental image searching-by-content. With this combination we hope to satisfy the needs of both experts and laypersons in the field of heraldry.
The text-based approach is based on controlled vocabulary from heraldic vocabulary (colours, ordinaries, charges, etc.) and vocabulary from free-text fields such as blazon information and text information. The index search allows to combine entries in selected fields such as family name, blazon, ordinaries and charges. To support the users, the correct terminology for the fields is shown in a scroll field. Nevertheless, we think it difficult for non-experts to practice efficient retrieval based on text alone.
An alternative approach which does not need terminological knowledge or knowledge of the German language on the side of the users is an image-based search. We have described the implementation of an image searching-by-content application where a user can retrieve images of coat of arms indicating some of their visual components. We will have to investigate how this approach is accepted by users, on the ground of future experimentations, and comparing it to the text-based approach. In any case, the main advantage of this approach is that the procedure (segmentation, MPEG-7 feature extraction, content indexing) is completely automatic, once the images are digitalized, and so can be applied to huge image databases with minimal effort.

Bibliography

[Amato, Gennaro, Rabitti, Savino 2004] G. Amato, C. Gennaro, F. Rabitti, and P. Savino. Milos: A multimedia content management system for digital library applications. In Europeean Conference on Digital Libraries, ECDL 2004, Bath, UK, September 12-17 2004, 2004.

[Amato, Bolettieri, Debole, Falchi, Rabitti, Savino2006] G. Amato, P. Bolettieri, F. Debole, F. Falchi, F. Rabitti and P. Savino: Using MILOS to Build a Multimedia Digital Library Application: The PhotoBook Experience. ECDL 2006: 379-390

[Amato, Bolettieri, Debole, Falchi, Rabitti Savino 2006] G. Amato, P. Bolettieri, F. Debole, F. Falchi, F. Rabitti and P. Savino: Using MILOS to build an on-line photo album: the PhotoBook. SEBD 2006: 233-240

[Bieber, Schweibenz 2005] C. Bieber and W. Schweibenz.. Images of Italian Art on the Web. The Database of the Photo Library of the German Art Institute in Florence. In: New Library World 106 (9/10) 2005. 430-440.

[Bieber, Schweibenz 2004] C. Bieber and W. Schweibenz. Digitalisierung und Langzeitarchivierung von Fotonegativen. In: Information zwischen Kultur und Marktwirtschaft. Proceedings des 9. Internationalen Symposiums für Informationswissenschaft (ISI 2004) Chur 6.-8. Oktober 2004. Hrsg. von Bernard Bekavac, Josef Herget, Marc Rittberger. (Schriften zur Informationswissenschaft 42). Konstanz: UKV. 125-140.

[Comaniciu, Meer 1997] D. Comaniciu and P. Meer. Robust Analysis of Feature Spaces: Color Image Segmentation. IEEE Conf. on Comp. Vis. and Pattern Recognition, pages 750--755, Puerto Rico 1997.

[Huang, Rui, Mehrotra 1997] T. Huang Y. Rui and S. Mehrotra. Content-based image retrieval with relevance feedback in mars. IEEE International Conference on Image Processing (ICIP), 1997.

[ISO/IEC 15938] ISO/IEC. Information technology - Multimedia content description interfaces. 15938.

[ISO/IEC 15938-6:2003] ISO/IEC. Information technology - Multimedia content description interfaces. Part 6: Reference Software. 15938-6:2003.

[ISO/IEC 15938-3:2003] ISO/IEC. Information technology -Multimedia content description interfaces. Part 3: Visual. 15938-3:2002.

[Ma, Manjunath 1999] W. Y. Ma and B. S. Manjunath. Netra: A toolbox for navigating large image databases. Multimedia Systems, 7(3):184{198, 1999.

[Mezaris, Kompatsiaris Strintzis 2004] V. Mezaris, I. Kompatsiaris and M. G. Strintzis, "Still Image Segmentation Tools for Object-based Multimedia Applications", International Journal of Pattern Recognition and Artificial Intelligence, vol. 18, no. 4, pp. 701-725, June 2004.

[Mezaris, Kompatsiaris Strintzis 2002] V. Mezaris, I. Kompatsiaris, and M. G. Strintzis, "A Framework for the Efficient Segmentation of Large-format Color Images", Proc. IEEE International Conference on Image Processing (ICIP 2002), vol. 1, pp. 761-764, Rochester, NY, USA, September 2002.

[MILOS] MILOS: Multimedia dIgital Library for On-line Search. http://milos.isti.cnr.it/.

[Salembier, Sikora 2002] P. Salembier and T. Sikora. Introduction to MPEG-7: Multimedia Content Description Interface. John Wiley & Sons, Inc., New York, NY, USA, 2002.

[Swain, Ballard 1991] M. Swain and D. Ballard. Color indexing. International Journal of Computer Vision (IJCV), 7(1):11{32, 1991.