Some Guidelines for the Design of Multimedia Systems for the Representation of Figurative Artistic Heritage

Luciana Bordoni, ENEA/STUDI, and Attilio Colagrossi, Consiglio dei Ministri, Italy

Abstract

The aim of this work is to furnish the guidelines for the design of a multimedia system for the representation of works of art. A work of art can be interpreted in a number of ways. There have been many theories on analyzing an image as a whole, and as constituent parts of a whole. There have also been many theories on how to interpret its (more or less obvious) contents and relate them to the different contexts in which the image was created and used. This paper provides some guidelines for the design of multimedia tools that combine the specific functions of databases within an essentially graphic environment. The guidelines concern, but are not restricted to, how the design of a tool is furnished and a discussion about some examples of an information access scenario concerning how works of art are presented.

1. Introduction

The means for approaching works of art is one of the most deeply felt issues in the field of the artistic heritage. The classical approach to this issue, in books and guided tours, has been recently enriched by exploiting multimedia applications and virtual reality. The multimedia applications, which are currently available, are usually made up of a text, just like a regular book, and of sophisticated functions such as sound and animation. Viewing a work of art through these applications is most certainly useful, interesting and encourages the natural learning process. It is a well-known fact that the above-mentioned applications provide the end-user with a variety of functions and with the possibility of referring to the issues at hand through various means. Generally speaking, the means and possible combinations for researching subjects are already built into each multimedia application.

The aim of this work is to furnish the guidelines for the design of a multimedia system for the representation of works of art; an approach based on "subject-based structuring" to obtain a semantic structure of a multimedia tool will be proposed. A work of art can be interpreted in a number of ways. There have been many theories on analyzing an image as a whole, and as constituent parts of a whole. There have also been many theories on how to interpret its (more or less obvious) contents and relate them to the different contexts in which the image was created and used.

This paper illustrates some guidelines for the design of multimedia tools that combine the specific functions of databases within an essentially graphic environment. Sound could also be added, although, in this phase, it is not considered strictly necessary for the representation of figurative works. The guidelines concern, but are not restricted to, the design of the following components of the tool:

- the knowledge base, which must include the following information: images of works of art, information concerning the images of works of art (title, size, date, location, ...), specific knowledge essentially comprising biographical facts on works of art (authors, art schools, art movements, ...);

- the query system, which must allows for answering to questions such as: "What are the paintings representing the Fame?", "How many and what are the paintings representing the Infanta Margarita?", etc.

- the multimedia interface, very user friendly, allowing several views of a painting, a gallery of works of art, etc.

We presume, for example, that a professional user is interested in the question of the meaning of a painting. In *The Art of Painting* or *The Atelier* (Vienna Kunthistorisches Museum, 1662-1665) by Vermeer, for example, each element of the painting has a precise meaning and Cesare Ripa's *Iconologia* published in Dutch in 1644, is its iconographic source. In this case the user may request details of the elements or ask for the relations connecting them. The Natural language generation techniques to access the data of the domain are required to automatically process data and to extract information from texts. The fact that it is

possible to connect different images together because they have similar descriptions, or cross references, or links, quotations and descriptions of other authors, constitutes very important reference material for professional users. Some guidelines useful for the design of such systems will be presented and a discussion about some examples of an information access scenario concerning works of art will be furnished.

2. Related Works

A multimedia application mixes speech, audio, image, video and text processing and navigation, in order to offer an improved and transparent interface and provide a communication that is as natural as possible. The traditional functions that can be performed on a multimedia document mainly use its syntactic structure and not its semantic content. More elaborate operations such as a search for a particular event or the generation of a summary of the document require an analysis of the contents of the document. This is a very difficult task because of the variety and complexity of the recognition techniques that are involved (speech recognition, image analysis, natural language understanding). Hence there is a need for tools which facilitate the development, experimentation, combination and integration of various indexing techniques. A major issue in multimedia document processing is the construction of indices that are representative of the content of the document. With the development of multimedia technologies and the wider usage of multimedia documents, a strong emphasis has emerged on the intelligent processing of multimedia documents. Because text indices are easier to manipulate, many approaches use techniques derived from textual Information Retrieval (IR) (Salton, 1989), by processing either the textual component of the document (Ogle & Stonebreaker, 1995), or textual annotation that have been manually added to the document (Weber & Poon 1994). Although it is harder to define indices in non-textual data, such indices are potentially so useful that many projects use indices from a variety of different media, such as the multimedia episodes defined in Gabbe et al. (1994). Srihari (1995) combines natural language processing and image understanding to create an automatic indexing system for captioned pictures of people; Rowe and Frew (1997) use a related approach. Many investigations analyze the possibility to create, manipulate and process indices allowing intelligent processing such as filtering and retrieval. Examples of such systems are presented in Brown et al. (1995), Hauptman and Witbrock (1997),and Manl et al. (1997).

3. Semantic Characterization of Representation of Works Of Art

A generic user spends most of his resources to study the semantic content of the works of art. We are going to talk about "subject-based structuring", in which by combining text and image it is possible to obtain a semantic structure of the multimedia object being analyzed. Although text, image and sound are the fundamental components for multimedia information, text, as Takeshita has pointed out (Takeshita, 1997), is the only tool that allows us to carry out a semantic analysis. The visual analysis of a painting can be carried out in many ways, and on different levels. Starting from a certain level, which can be considered the basic level, the first step we decided to make was to highlight the areas of a painting that are considered essential to the work itself, i.e. its constituents, and to connect adequate information to these parts. These elements that contain a semantic meaning for the work itself are extracted, in other words the structures of the image are selected.

Many cases have been studied and will be detailed in further accounts. For the time being, we shall just say that the following "subject-based structuring" was analyzed. It may be considered a global structuring as it belongs to a specific class of works:

- *historic-literary* (identifying historic or literary figures present in the painting, e.g.: *The School of Athens, The Parnassus* by Raphael, *Las Meninas* by Velazquez, etc.);
- *iconological* (interpretation of the meaning of the painting, e.g.: *The Art of Painting, The Allegory of Faith*, by Vermeer, etc.);
- *exploratory* (highlighting characteristic details of models and/or objects present in the painting, e.g.: *Davide with the head of Golia* by Caravaggio, *Old Woman Frying Eggs* by Velazquez, etc.).

A local "subject-based structuring" follows the previous one. The objects that characterize the subject are extracted from the image. Putting the global and local activities together, makes it possible to extract significant information from the image being analyzed.

Consider for instance *Las Meninas* by Velazquez (Madrid, Prado Museum) painted in 1656. The first denomination of this great painting can be traced back to the seventeenth century: *The Family of Philip IV*. Its contents are undoubtedly intrinsic to

the painting and they imply dynastic meanings in relation to the intimate representation of the painting on behalf of the royal family. In fact, they are the ones reflected in the mirror while they are viewing themselves (Marini, 1994). It is thus possible to carry out a global historic "subject-based structuring" on this painting, and to attach a historic reference to each of the identifiable people present in the painting, through a local "subject-based structuring." The same goes for all the objects present in the painting: the mirror, the dog, the paintings. It is possible to explore both its physical and informative characteristics, such as the representation contained in the paintings (*Minerva and Arachne* by Rubens, *Apollo and Marsia* by Jacob Jordaens).

In *The Art of Painting* or *The Atelier* (Vienna Kunthistorisches Museum, circa 1662-1665) by Vermeer, each element of the painting has a precise meaning and Cesare Ripa's *Iconologia* published in Dutch in 1644, may be its iconographic source (Danesi Squarzina, 1990).

The trumpet and the laurel, which the painter is sketching on the canvas and that is intertwined in the model's hair, are attributes of Fame. The book the model is holding in her hands, represents history. The plaster mask lying on the table is an allusion to art's fictitiousness. Even the *Allegory of Faith* (New York The Metropolitan Museum, 1670) is brimming with contents. The church chalice, the crucifix, the book, the bitten apple, the snake squashed by a rock, can all be found in Cesare Ripa's *Iconologia*. Furthermore, Vermeer's work often contains paintings by other artists to which a meaning is attached.

The methodology we propose, in reference to these examples, will allow users to communicate both visually and in formal-iconographic and semantic terms with the image.

The painting entitled *Old Woman Frying Eggs* by Velazquez (Edinburgh, National Gallery of Scotland, 1618) can be considered the perfect example of how the curiosity of a person approaching the painting for the first time, can be satisfied by an exploration based not only on the image but primarily on the characteristics of the elements it contains:

- the objects in the boy's hands, i.e. a melon and a jug of red wine;
- old woman, Maria del Paramo wife of Francisco Pacheco (artist and father in law of Velazquez).

The possibility to perform in a very friendly way the operations of inputting, modifying, erasing textual information is the aim of this methodology. The same work of art can, in fact, be characterized by different interpretations and found through queries based on the characterization given to the painting. Even in a didactic environment, the teacher can control the correct interpretation of the work of art taken into consideration, through queries.

4. Scenario Description

The goal of our work is to furnish some guidelines for the design of a framework for developing content-based multimedia retrieval tool according to the "subject-based structuring" approach. The system should allow professional users in historic of art to quickly construct a retrieval application that supports automatic processing to extract content, semantic queries and views of the semantic content. The framework must be also provide support for a wide variety of multimedia data types. Although we will focus on image data in this paper, these techniques can be applied to other multimedia data types as well.

The data model underlying the knowledge base can be designed as a network of objects from the art history discipline that are interconnected by semantic relations. Examples of object types included in this network are person, profession, nationality, work of art, movement, art style, school, place, etc. They may be connected by relations, such has, for example, "has profession" connecting persons with professions, "is teacher of" connecting students, and so on. With this approach must be also considered facts: facts comprise a spatial and a temporal dimension events. They are represented by relation tuples. We say that knowledge consists of facts, such as "Francisco Pacheco is father in law of Velazquez" or "Las Meninas was painted in 1656". In this way the knowledge base is a network of n-ary relation instances.

The system we want to design must make it possible to treat/process the following information:

a) a group of images;

b) for each image belonging to (a), information concerning the image;

c) for each image belonging to (a), a group of constituent images;

d) for each image belonging to (a), information and reference to other images of (a);

Here is a general diagram that intultively represents how the elements are related to each other:

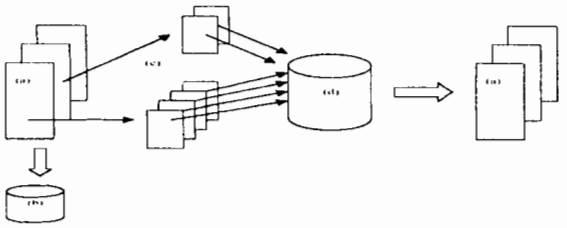

Fig. 1

Here is an explanation of the diagram and of points (a) to (d):

Group (a) is made up of images concerning certain paintings, for example by Vermeer (*The Atelier*) or by Diego Velazquez (*Las Meninas, The portraits of the Infanta Margarita, Portrait of Francisco Pacheco)*;

Group (b) is made up of information (on each work of art taken into consideration) such as, for example: title, size, date, location;

Group (c), in the case of *Las Meninas* for example, is made up of the following images: little girl, woman, dog, mirror;

Group (d), in conjunction with what has been said about group (c) contains the following information: the Infanta Margarita, Doña Maria Augustina de Sarmiento, Doña Isabella de Velasco, Castilian mastiff, mirror that reflects Queen Marianne of Austria and King Philip IV.

Basic functions in order to allow a user to perform the operations above briefly described must also be provided.

5. A Prototypal System and Experiments.

A prototype system, for experimental purposes only, has been developed (Bordoni & Colagrossi, 1999) using the graphic and storage functions provided by a low cost geographical information system (GIS), suited for storing images, selecting parts of images, associating structured information to images and parts of images.

From a conceptual point of view, the prototypal system consists in a multimedia database (images and text) containing the components briefly highlighted in the previous diagram. The database's conceptual diagram used in the prototype is as follows (the points marked with a "*" indicate dominions made up of images):

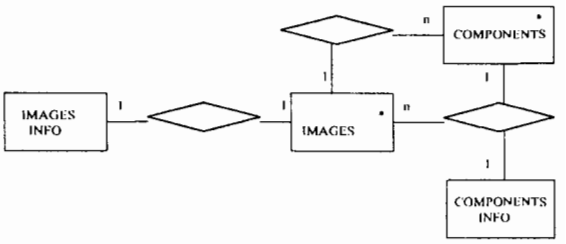

Fig. 2

Further, the prototypal system makes available the functions to perform the following necessary operations:

- acquisition of a raster image and memorization of the image on the hard disk;
- association of each image with a description in form of structured data (using the relational data base management system underlying the GIS);
- selection of parts of an image and association of each image part with a description in form of structured data (using the relational data base management system underlying the GIS);
- retrieval of a set of images;
- consultation of the information associated with a retrieved single image.

The tables, which correspond to IMAGES INFO and to COMPONENTS INFO of the diagram previously described, contain the following information:

- IMAGES INFO: <title of the work art>, <author>, <date of execution>, <size>, <location where it is conserved>, <short description>.
- COMPONENTS INFO: <subject>, <age>, <synthetic description>, <name>, <role>, <appearance>, <attitude>, <color of the dress>, <color of the hair>, <ornamental objects>, <nearby to>.

Possible values for the fields of the COMPONENTS INFO table are the following:

<subject>: man, woman, animal, thing, etc.;

<age>: young, child, old, etc.;

<role>: princess, king, servant, etc.;
<appearance>: normal, dwarf, one-armed, bald, etc.;
<attitude>: sitting, standing, running, sleeping, etc.;
< nearby to >: person, dog, tree, etc.

An example of use of the prototypal system is shown in Fig. 3:

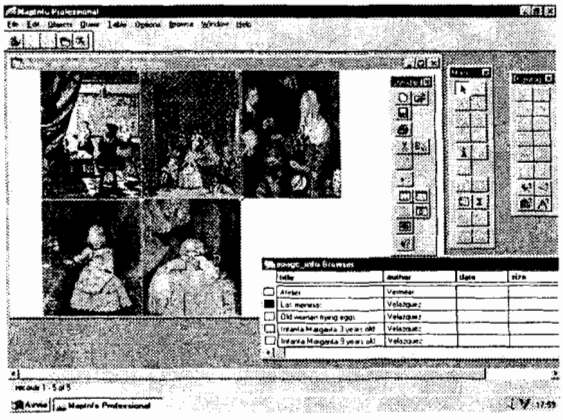

Fig. 3

There, a Gallery consisting in five works of art is represented. The Gallery has been built by using the very simple tools provided by the GIS. The paintings in the Gallery are, in the order: *The Atelier* by Vermeer, *Las Meninas*, *Old Woman Frying Eggs*, *The Portrait of the Infanta Margarita 3 years old*, *The Portrait of the Infanta Margarita 9 years old* by Velazquez. In the figure the operation of selecting the work of art *Las Meninas* is shown. Then, according to the local "subject-based structuring" approach, we can extract from the image all the objects that characterize this subject, such as, for example, the elements Infanta Margarita and the dog. Such elements, once selected and extracted by using the functions provided by the GIS, can be described by associating to them some information (structured as a table or a complex database). In fact, it is possible to attach an historic reference to each of these selected subjects and carry out a local historic "subject-based structuring" of this painting.

6. Conclusions

Some guidelines in order to design a system for a personalized multimedia representation of figurative artistic heritage have been presented. Several aspects of such a task have been explored: starting from the meaning of 'personalized representation', passing for strategic elements of design (such as data analysis), concluding with the presentation of a prototypal system and experiments performed with it. The guidelines underline a methodology that allows personalization of a tool in order to create multimedia applications in which the user's subjectivity is the fundamental element. We have designed a framework, based on the research approach mentioned above, in which the identification of the semantic content is a crucial aspect. This enterprise was motivated by limited design guidelines for the design of a multimedia system for the representation of works of art based on "subject-based structuring". Through our experiments we have identified several kind of deficits in this design. The nature of these deficits enabled us to formulate additional and more concrete guidelines for this design. Finally, we have argued that there are wide range of areas for beneficially incorporating intelligent methods, including filtering by user modelling, and adaptive hypertext. Techniques originating from these areas promise to contribute to the avoidance or, at least, to the reduction of a variety of the deficits observed.

7. References

Bach, J.R., Fuller, C., Gupta, A., Hampapur, A., Horowitz, B., Humphrey, R., Jain, R., & Shu, C. "The Virage Image Search Engine: an open framework for image management" in I.K. Sethi and R.C. Jain (Eds.) *Storage and Retrieval for Still Image and Video Database IV*, (SPIE 1996 v.2670 p. 76-87. San Jose, CA: International Society for Optical Engineering, 1996.

Bordoni, L. and Colagrossi A. "A Multimedia Personalized Representation of Figurative Artistic Heritage by a GIS-based Methodology" in *Proceedings of the IEEE International Conference on Multimedia Computing and Systems Multimedia Systems*, Vol. 2 p. 189-193. IEEE Computer Society, 1999.

Brown, M.; Foote, J., Jones, G., Sparck-Jones, K. And Young , S. "Video Mail Retrieval by Voice: An Overview of the Cambridge/Olivetti Retrieval System" ACM Multimedia Conference, *Workshop on Multimedia Database Management Systems*. p.47-55, 1994.

Danesi Squarzina, S. Vermeer. *Art e Dossier* n.45, Giunti Editore. 1990.

Gabbe, J.D., Ginsberg, A.; and Robinson, B.S. "Towards Intelligent Recognition of Multimedia Episodes in Real-time Application" in *ACM Multimedia Conference* p.227-235. San Francisco: ACM Press, 1994.

Hauptmann, A. G. and Witbrock, M.J. "Informedia: News-on-Demand Multimedia Information Acquisition and Retrieval" in Mark T. Maybury (Ed.) *Intelligent Multimedia Information Retrieval* p. 215-239. Cambridge, Massachusetts:The MIT Press, 1997.

Kamps, T., Huser, C., Mohr, W. and Schmidt, I. "Knowledge-Based Information Access for Hypermedia Reference Works: Exploring the Spread of the Bauhaus Movement" in M. Agosti and A. Smeaton (Eds.) *Information Retrieval and Hypertext* p.225-256. Kluwer Academic Publishers. 1996.

Krause, M.G. "Intellectual problems of indexing picture collections" in *Audiovisual Librarian, 14*(2), 73-81, 1988.

Mani, I., House, D., Maybury, M. and Green, M. "Towards Content-Based Browsing of Broadcast News Video" in Mark T. Maybury (Ed.) *Intelligent Multimedia Information Retrieval* p. 241-258. Cambridge, Massachusetts: The MIT Press, 1997.

Marini, M. Velazquez. *Art e Dossier* n.94, Giunti Editore, 1994.

Ogle, V., Stonebreaker, M. Chabot: "Retrieval from a Relational Database of Images" in *IEEE Computer* 28(9), 40-48, 1995.

Ohlgren, T. "Subject indexing of visual resources: a survey". *Visual Resources, 1*(1), 67-73, 1980.

Plaisant, C., Shneiderman B. and Mushlin R. "An information architecture to support the visualization of personal histories" in *Information Processing & Management* 34 (5), 581-597, 1998.

Rowe, N. C. and Frew, B. "Automatic Retrieval of Objects in Captioned Depictive Photographs" in Mark T. Maybury (Ed.) *Intelligent Multimedia Information Retrieval* (pp. 65-79). Cambridge, Massachusetts: The MIT Press, 1997.

Salton, G. *Automatic Text Processing*. Reading, MA: Addison-Wesley, 1989.

Srihari, R.K. "Automatic Indexing and Content-based Retrieval of Captioned Images" in *IEEE Computer* 28(9), 49-56, 1995.

Takeshita, A., Takafumi Inoue and Kazuo Tanaka. "Topic-based Multimedia Structuring" in Mark T. Maybury (Ed.) *Intelligent Multimedia Information Retrieval* p. 259-277. Cambridge, Massachusetts: The MIT Press, 1995.

Turner, J.M. "Subject access to pictures :considerations in the surrogation and indexing of visual documents for storage and retrieval" in *Visual Resources, 9*(3), 241-271, 1993.

Weber, K. and Poon, A. Marquee. "A Tool for Real-time Video Logging." in Proceedings of ACM CHI'94 *Conference on Human Factors in Computing Systems*, volume 2 of PAPER ABSTRACTS: Multimedia in Use, 1994.

White, C. "Integrating Document Image Management with GIS" in *ESRI International User Conference*, 1998.

Williams, B. "Kansas Cultural Resources GIS" in. *ESRI International User Conference*, 1998.