An Approach to Indexing Annotated Images

Panos Constantopoulos and Martin Doerr
Institute of Computer Science
Foundation for Research and Technology -Hellas
Heraklion, Crete, Greece
(panos/martin) @ics.forth.gr

Abstract

The various relevant aspects for the retrieval of images present familiar documentation issues as well as some entirely new requirements. The Semantic Index System (SIS) has already been successfully employed in maintaining an evolving body of knowledge about the cultural and historical contexts of museum artifacts and multimedia documents (CLIO SYSTEM). Its distinct features of runtime meta-modelling and classification of references satisfy the multiple classification needs for images and their parts and allow to access simultaneously through detailed annotations. Consequently, it integrates access methods through generic and specific indices and the relevant context. In this paper we present the principles and architecture of such a system as it is currently being implemented.

Introduction

It is commonplace that images in any form, from sketches to photography to photo-realistic representations, are important means of rendering human expression and documenting knowledge. Recent computer technology, especially the availability of high bandwidth global networks, is allowing access to a rapidly growing, eventually overwhelming amount of digitized images.

On one hand there is the opportunity to achieve on the computer screen the quality of presentation so far only available in good books and exhibitions, reaching a potentially vast public. On the other hand, the ease of publication and broad dissemination goes along with the apparent paradox of increasing confrontation with irrelevant or poor material. This is because traditional authoring and editorial work, concerned with producing a coherent, targeted, high quality document, is partly giving way to information retrieval functions performed directly by the user. Hence the challenge is to develop retrieval tools for selecting with high recall and precision not only the relevant items, but also the "best," the prototypical, the beautiful - a task traditionally performed by a human specialist. Of course this

raises interesting questions of objectivity and of how the multitude of human opinions can be represented in a machine.

An image can be interpreted at different levels. It can be seen as "raw data"([Oomoto93], [Chan92]); sensory information prior to any conceptualization; a geometric pattern; a scene where visual elements have semantics; an entity which can be assigned a meaning or aesthetic attribute as a whole. Accordingly, there are several access structures and retrieval techniques for selecting images from an image base, the choice depending on the application. Lower levels of interpretation are more amenable to automatic processing and can be effectively used in some narrow domains with fixed, commonly agreed semantics. More abstract levels of interpretation are often less amenable to automation ([Chan92]) and can depend a lot on context and the intentions of the involved parties.

Annotation and keyword indexing are common image access methods, of seemingly general applicability. However, it appears that their effectiveness is largely determined by such factors as context and intended use.

In this paper we review the intentional aspects of image indexing and annotation and propose a semantic indexing approach that allows the required flexibility, combined with discipline, for effectively retrieving images by conceptual content. The pictures referred in the text can be found in the appendix.

Image content and retrieval

In this work, we regard an image as a two-dimensional composition of colors [Megh95]. If this composition is defined over a discrete, connected grid, then the image is *digitized*. As such, we distinguish an image from its physical carrier, or original, which is unique, and has physical substance. Without going into details of quality and limits of reproduction, we understand an image as the respective information commonly reproduced in books, pictures, or electronically in computers. In this sense, the image is immaterial, and exists beyond the physical destruction of the original. Digitized images can be reproduced without losses and we can assume, for practical purposes, the images in a digital archive to be of adequate quality. The physical carrier may be a photograph, print, drawing, painting, applied on some object, such as paper, canvas, vase etc, or even an electronic carrier.

Images accompany a series of states in human cognition, thinking and feeling, which range from photography of reality to abstract patterns to reality-like presentations of imaginary scenes. Understanding and indexing image content has been an important research theme in the area of image analysis, pictorial data bases and multimedia information systems.

The *content* of an image is understood ("extracted") through a process of interpretation which, in some applications, may be (partially) automated. If no interpretation is made, the image is a

two-dimensional composition of colored regions and content-based retrieval involves matching visual features, such as color, texture, size and shape. A stage of geometric feature extraction precedes matching. This view of content may be adequate in some narrow application fields, where the suppliers and users of the images have a precise, common way of interpreting the visual features, such as in specialized medical image data bases (e.g. brain tomography). On the other hand, a process of recognizing objects, spatial relationships and other conceptual properties of the objects or the image as a whole, leads to a mental reconstruction of the depicted scene and a set of global judgements as representing the image content. Then the content assessment may vary from interpreter to interpreter (human or machine) depending on the context in which the interpretation is performed. Such would be the case of a data base of paintings or documentary photography.

Correspondingly, a wide variety of image retrieval methods have been developed, ranging from pattern matching to semantic information processing. At the pattern matching end of the spectrum, users would submit images as queries and the system would perform a matching operation using some visual similarity criterion. At the other end, queries involve semantic properties of the objects depicted in the images, or even of the images as a whole. At an intermediate level of abstraction, queries can involve spatial (geometric) relationships between objects depicted in the images.

In current image retrieval systems the techniques mainly used to represent conceptual image content are captions, keywords, iconic indexing and hypermedia annotation. A *caption* is a piece of text associated with an image, which describes the depicted scene and/or contains general statements about the image. Thus captions are free text annotations and they support content-based image retrieval through text retrieval. *Keywords* are terms drawn from a fixed vocabulary and assigned to images, so as to form an index to the image base. In *iconic indexing* images showing objects in symbolic form along with their spatial relationships and descriptive (object and/or image) attributes are used to access the image base [Chan87, Cons91, Kato91]. An important stream of research effort in developing indexing structures for spatial and temporal data is reviewed in [Salz94], see also [Pras94]. In *hypermedia annotation* an arbitrary image region is annotated by a piece of text, structured attribute, or even another image or audio recording. The relationship is established by a reference (*hypermedia link*) originating in the image region (*anchor*) and pointing to the annotation. Apart from these content descriptors, image retrieval systems use external identifiers and attributes (e.g., creator, time and place of creation, owner, etc.).

An interesting attempt to incorporate the various forms of image retrieval in a single model based on classical logic is presented in [Megh95], where it is argued that the content representation schemes currently used fail to capture image content in a satisfactory way and the use of semantic data modeling is advocated. Focusing on scene representation and on establishing the mapping between "form" (visual content) and "content" (conceptual content), Meghini employs a closed world assumption whereby all the depicted objects and their conceptual relationships are known. However, the closed world assumption is too restrictive and artificial. A strong counterexample is provided by video frame sequences, the contents of which can only be described in a dynamic, incremental way [Oomo93].

Furthermore, limiting contents description to scene understanding overlooks a concurrent process of comprehensive interpretation which addresses such important aspects as impression, expression, prototypicality, etc. The separate assessment and recording of such interpretations would be beneficial not only in terms of completeness and usability, but also of performance. In fact, a detailed scene analysis will result in data overkill if adequate access can be ensured by some global image attributes in the context of a particular application.

For cultural documentation purposes, a closed world, purely analytical view seems to be too narrow, for at least two reasons. First, we are interested in several objective features that are not bound to discrete objects, such as use of color, light, etc.. Second, the main interest in an image may lie in impressions, or human psychological reactions in general (see pic.1, pic.5). We can assume that certain combinations of features result in mental associations, some of which can be articulated. According to our interest, possibly unconscious, we select a few of them, remember them or consider them as worth communicating. The purpose of an index is then to make the link between those associations and the image accessible to other people. Annotations can further link specific features of an image with associations (see pic.3-4). This is especially helpful if the relation is not obvious, e.g. boundaries of houses or walls of different layers in a picture of an archeological site. Associations need not be expressible in words, they may rather refer geometric patterns (define the bourbon lily in words !), similarity with other images, scenes etc.

Subscribing to an undisciplined, informal open world approach, on the other hand, would also be ineffective. Lack of formality in the representation of semantics can result in large collections of hypermedia or free text annotations that are hard to search and cause serious disorientation and cognitive overhead. For this reason, [Hjel94] already categorizes free text annotations into "person-, location-, and event-annotation," without intercorrelations however. Similarly, a growing keyword index can prove hard to maintain and use if not supplied with appropriate semantic abstractions [Sven89],[Cons95]. Furthermore, certain limitations, or requirements, on an index should already be apparent from the above brief discussion. Mental associations must be expressible, and their expression must be communicated to other people. One step back, these associations must have been somehow established (by a human or mechanical agent). In turn, such an action could only be part of the pursuit of explicit or implicit intentions and goals of the various actors involved in the generation and use of an image or multimedia document base. In [Pate94], Patel claims, that semantics exist only with respect to an agents goals, and cannot be established from the independent nature of the agents environment.

Intentional aspects of image bases

It appears worthwhile to pay increased attention to the intentions and goals that users may have when querying an image base. First, this will help defining the desirable descriptive structures and retrieval mechanisms. Second, it will enable understanding what is out of the scope of an indexing system.We

can distinguish three actor groups involved, each of them having different goals: the creator of the image, the person in charge. of the classification, i.e. the documentalist, and the searching end user. In the simplest case, the goal of the creator conforms with that of the end user, e.g. a photo reporter working for a TV station. In other cases, predominant goals of end users are obvious, e.g. in medical imaging. In general, it is the difficult role of the documentalist to mediate between the creator and the potential interests of the end user. In the following we attempt a classification of images according to goals of the creator and the end user. We do not claim completeness, nor do we want to compete with related work in iconography. Rather, the aim is to convey an approach to bringing a retrieval system in correspondence with the human goals involved. This classification is subject of a cooperation with the Benaki Museum, Athens and will be incorporated in the CLIO cultural documentation system (see fig. 1).

From the creator's side we may distinguish documentary and non-documentary images. Documentary images can be photos, or photo-realistic paintings, copper, steel (see pic.7) or wood engravings, etc. They may show either a real world scene, i.e. an arbitrary "snapshot" of reality (e.g. historical events, tourist photos, archeological sites, landscapes, human settlings, etc.), or depict a real world object, often deliberately isolated from its environment (see pic.6), but without intentional deviation from the original (e.g. portraits, museum object or product photos, scientific images, etc.). We should also include in this category images produced by technical means beyond human vision, and images of things that were planned to become real (e.g. architectural or car realistic renderings, etc.).

Documentary images can also be abstract, conveying structural informations about real things, such as maps, technical drawings and sketches.

Non-documentary images, including photographs, could be art objects themselves. They may either show parts of reality more or less modified and abstracted (see pic.8), with intended expression beyond rendering information on the world behind, give a more or less reality-like appearance to imaginary things and scenes (e.g. mythological subjects, see also pic.5), or be plainly abstract.

Non-documentary images further include patterns, such as icons, symbols, crests, or decorative patterns. They may represent a certain meaning, as language neutral instructions and Han characters, refer to a certain dynasty or person, or have magical value (see pic.2). They may be traditional decorations of which the initial meaning is often lost. The derivation from reality may be apparent (e.g. a plant motive), or not (e.g. a mandala). They are characterized by a standardized abstract shape, which may be re-enriched (e.g. initial letters in medieval handwriting).

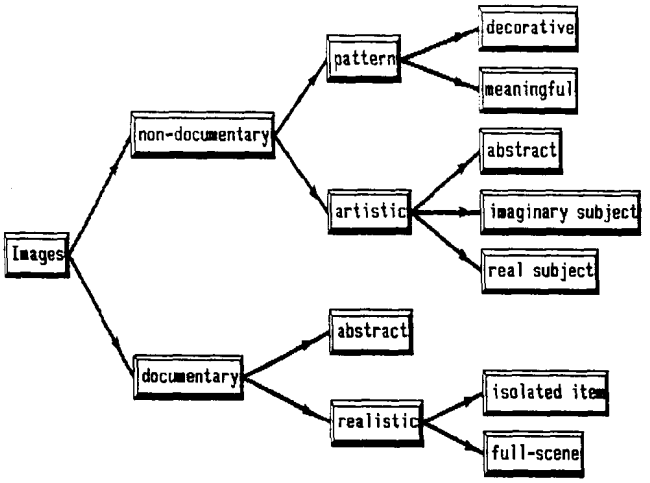


Figure 1: Classification hierarchy of images by not mutual exclusive creator goals

Obviously, the above categories are not mutually exclusive. They differ strongly in the number and heterogeneity of potential aspects of interest.

From the point of view of the end user we broadly distinguish studies, where the user wants to get an impression or understanding of the image, from actual usages, which imply transferring a copy to another environment (or, perhaps, buying the original).

One kind of studies are those , where the focus is on understanding what the image shows, not on the image itself. The image is just one kind of source of information, it illustrates its subject. The interest in image quality is restricted to the fidelity, by which the relevant features of the reality behind are rendered. Typically, the user is then interested in several images on the same topic.

There may be object studies, where the image replaces the availability of the object, the mobility of the user (see pic.2, pic.6), or the limitations of human vision. Domains can be any science, or even detective or jurisdictional work. Here "object" may denote something dead or alive, or a group or class of objects. Studies may further be about events (historical, scientific, etc.), i.e. about non-repeatable phenomena, where an image may be the only objective information available at all. They may concern processes, natural or social (see pic.3-4), evolutions, such as changes of landscapes, buildings, etc., or repeatable phenomena, such as seasonal changes. In this case series of images are needed documenting the stages of a process (see, e.g., [PICQUERY+]). The focus of annotation then shifts from the image to the relation between images (see pic.3-4), a case mostly not foreseen in image bases.

Another group of studies are concerned with the image itself, the presentation in contrast to the representation. Such are iconographic studies and may deal with such aspects as harmony, light, color, impression, expression or induced feelings (pic.5). The interest in the depicted subject is usually at a conceptual level, and often associated with a characteristic way of presentation. Now the creator becomes an item of interest, and the importance of the reality behind may reduce to emotions about it.

A number of usage goals can be identified, such as illustration, instruction, decoration, or even psychological testing.

Illustrations may have technical, psychological or spiritual intentions, as scientific publications, advertising, religious images (pic.8) etc. The interest shifts from completeness of the material retrieved to *prototypicality*, as in pic.6, (a also reason why sometimes a drawing may be preferred to a photograph), completeness of the desired features in one image, and additional artistic or psychological qualities. Decorations present images for the sake of beauty or psychological value in general, in private or public places, where the harmonic relation to the environment in subject, structure and color plays a major role. Instructional usage of images is highly focused and easily supported. A usage like psychological testing, however, may be hard to capture. In any case, even in the complete absence of any sensible criteria, the simple and often accessible information about the past usages of an image can be valuable for future requests. Humans are apt at developing search strategies through various associations, to support unforeseen or subjective criteria.

The analysis below suggests that we can identify groups of images supporting specific predominant goals. The goals of the creators are mostly a relevant subset of those. Some groups of images support a rather limited number of goals (e.g. a map), whereas others are obviously rich and complex, far beyond what the creator had in mind. Consequently indexing criteria cannot be restricted to scene recognition. Moreover, a complete scene analysis might even be irrelevant. Most user goals must be seen in a context, where the relation between things, concepts and theories/beliefs is of primary interest (pic.3-4), possibly not evident from the scene itself. Beyond objective criteria, subjective impressions and psychological values are characteristic for certain user/usage groups and should be explicitly accounted for. [Kato91] e.g. refers a successful experiment, to normalize subjective impressions on colors against automated criteria.

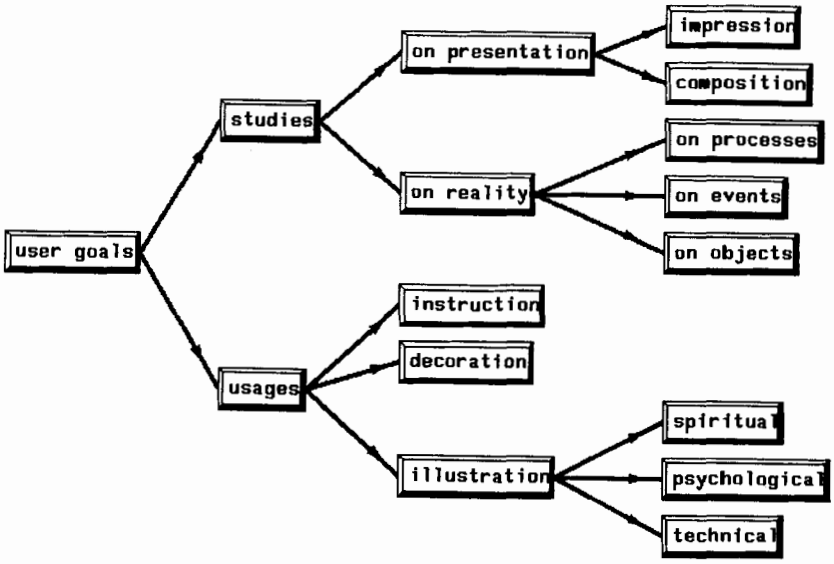


Figure 2: Classification hierarchy of end-user goals for image retrieval.

It makes no sense to devise a "complete" set of descriptors that would apply to all images, in the spirit of relational data modeling. For one thing, few would be worth applying in most cases. We are confronted with two tasks. The first is to determine major relevant aspects at a very high level, and assign descriptions directly (e.g. "funny," pic.1), or indirectly through reference to concepts implying further aspects (e.g. "deer"). Thus we achieve at least a significant reduction of the search space, without loss of recall. The lesson from applications of faceted classification is that a reasonably limited number of agreed high level aspects is feasible and useful. Here we see the role of the initial documentalist, who should better be a generalist enabling access by specialists. Suitable guidelines may help to achieve a reasonable coverage. The second task, to be undertaken by specialists, is the assignment of specific descriptions. This task is not limited to a contribution of established special knowledge. Rather, it involves recording the output of ongoing research, hence it is necessary to be able not only to update the values of the descriptors, but also to modify and extend the set of descriptors. In this context it seems worthwhile to invite a larger user group to contribute in a controlled manner to an evolving body of knowledge. Notions of relevance can be expected to change gradually, following general trends in cultural and scientific movements.

A semantic indexing approach

We propose to adopt a formal semantic data modeling approach to image indexing under an open world assumption that will allow for incomplete and evolving image content representations. This approach will also support a multiplicity of points of view and user intentions. Achieving model-based

representational openness and flexibility requires allowing a virtually unlimited number of classification structures on objects, concepts and relationships, as well as a multilevel organization of these structures. The multiplicity of classification structures will enable capturing different points of view (including those of users).

Multilevel organization implies various levels of abstractions, thus the ability to express properties and conditions in a manner closer to human thinking. In an indexing and retrieval system, abstractions serve several purposes. To formulate a valid query, the user must learn the index itself, as pointed out in [Cons95]. Abstractions serve as an index on the index itself. Besides by scope notes, the meaning of a term is clarified by its position in a hierarchy. Semantic indexes are highly structured, and need particular indexes on available categories or attributes, especially when they are user definable (see [Oomoto93]). Multilevel abstractions further mediate between the level of detail used in annotations (e.g. "early Maya vase"), and the level of interest of the user (e.g. "Middle American Art"). If suitably set up, they allow also to classify consistently incomplete information (e.g."Maya vase").

This approach can be implemented (and, in fact, is under way of implementation) using an appropriate set of tools, such as the Semantic Index System and the CLIO cultural documentation system, briefly introduced below. Although a semantic data model can be designed to encompass spatial relationships, our main interest here lies in representing conceptual content in a manner subsuming captions, keyword indexing and hypermedia annotations. The Semantic Index System (SIS) is a tool for describing and documenting large evolving varieties of highly interrelated data, concepts and complex relationships,as opposed to large homogeneous populations in fixed formats (handled by traditional DBMS). It was developed at the Institute of Computer Science, FORTH, in the framework of a number of European projects, from 1990 to 1994 [Cons93, Cons94b, Cons95].

The SIS consists of a persistent storage mechanism based on an object-oriented semantic network data model, and a generic interactive user interface to insert and retrieve information in various ways. Several interfaces to other systems and customized user interfaces are provided.The SIS offers significantly richer referencing mechanisms than relational or ordinary object oriented systems. Together with the very high query speed along references, these mechanisms allow to keep data and schema redundance-free; each notion has one identifier in the system, be at the schema, data, or field level. Fast built-in inheritance mechanisms gather, at query time, the relevant information per node from its environment. The importance of powerful inheritance mechanisms is pointed out in [Oomo93,Megh95].

The SIS employs the structural data model of the frame-based knowledge representation language Telos [Mylo9O]. Its distinct features are the user-definable metaschema, allowing not only to treat the schema as data, but also to see the data itself under the view of metaentities; a multiple instantiation mechanism, i.e. the capability to describe the same data item in more than one table ("class"); and the classification of relations as objects in their own right.

This flexibility allows to import without loss of semantics data from relational, object-oriented and many knowledge representation models. As such, the SIS serves for logical integration of data from automated knowledge acquisition tools and indexes originating in human assessment. It is compatible e.g. with models as [Megh95,Oomo93,Card93] and others. It also provides the capacity and performance to deal with the associated large amount of data. The implementation of the SIS allows to extend an application schema at run-time. The discipline imposed by the metamodel ensures consistency and validity of the existing access methods. Once the fundamental notions are established at the metalevel, i.e. the facets and their correlations, as "actors," "events," "object," "part of," "participates," "creates", an extensible semantic index can be built, which conforms with the open-endedness of image annotations in particular. Especially the problem, how the end-user becomes informed and makes actually use of a dynamically growing list of attributes can efficiently be solved by use of metacategories.

The CLIO cultural documentation system [Chri94, Cons94a] was built on top of SIS. The analysis of requirements for CLIO was carried out in close cooperation with the Benaki Museum and the Historical Museum of Crete. The system is currently installed in the former. The development was partly funded by the STRIDE program.

CLIO serves as a scientific catalogue of museum artifacts, as opposed to the basic documentation and administrative purposes served by usual collections management systems.It supports artifact descriptions concerning temporal, geographical, cultural, historical, style, technique, usage, and physical information under a metamodel for cultural documentation. Besides structural information in the sense of database records, objects and notions can be annotated by free text and illustrated by images. As such, it covers the correlation of depicted items and scenes with the relevant context, and supports further associations originating in the nature of what is shown. E.g., the image of a musket can be found through the term weapon, or through any information associated with its historical owner etc. At present CLIO supports browsing, and a powerful extensible set of predefined queries on direct and indirect relations. We are currently implementing an ooSQL-style query language, that will allow for free query formulation in a user-friendly manner [Hill95].

Example

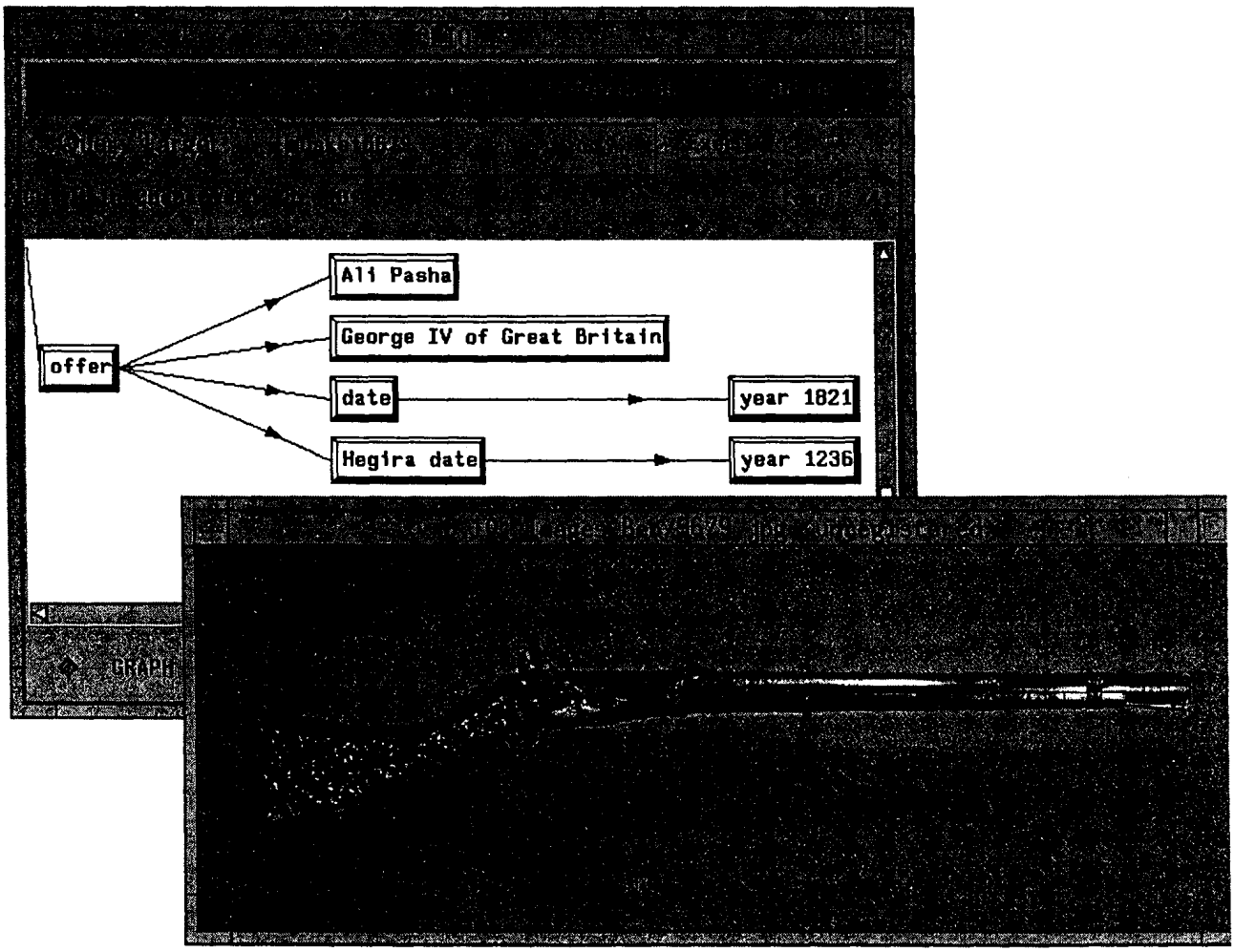


Figure 3: Screen dump of CLIO, retrieving descriptions of musket 5679.

Fig. 3 shows part of the descriptions hold in CLIO about a musket in the property of the Museum Benaki. Besides standard technical data, a record was entered in the system, stating that the musket with museum number 5679 was an offer of George IV of Great Britain to Ali Pasha, a ruler of the Greek province of Epirus. The record is shown in reduced graphical form in fig.3. The full record reads as:

offer of object 5679
  type:
    offer
  date:
    Julian Date: 1821CE
    Hegira Date: 1236
  subject:
    George IV of Great Britain
  acceptor:
    Ali Pasha
  offered objects:
    musket5679
end

The rulers are further analyzed by their realms and ruling periods, life data etc. The relevant story behind is, that this musket is a token of the international relations between Epirus and Great Britain in the early 19th century, and that Great Britain became aware of this Ali Pasha. The annotation correlates the object with the relation between countries. It lies not in the nature of the involved parties. This seems to be characteristic for many annotations.

CLIO provides various abstractions of the above notions. At metalevel, 'offer' is an event type,' 'ruler' an 'actor type,' the categories 'subject' and 'acceptors' are of kind 'history,' 'participants,' and 'date' is of kind 'temporal'. A more detailed discussion of the notions of occurrence in CLIO can be found in [Chri95]. An international relation reads as a linked chain country-representative-event-representative-country. This granularity of description, and the explicit modeling of events, is conformant with the principles of the latest work of the ICOM data standards [ICOM94]. It provides a redundance-free formal representation of the above fact for all involved individuals : both rulers, both countries, and the object itself. The resolution of a notion like "international relation" however cannot be expected to be obvious to an untrained end-user. For that sake, CLIO provides predefined queries as deduction mechanism. Up to now, our experience suggests, that queries by metacategories lead to a significant reduction of the necessary queries. Further research will be done by our group on this field.

Further work

Planned further developments in the CLIO model include:

(i) Modeling relations between the concrete, conceptual and imaginary, e.g. between a byzantine iconographic class of a saint and the historical person, between a unicorn and a mammal, or between "an African" and a certain king of Benin.

(ii) Modeling overall properties of the image itself, e.g. composition, coloring etc.

(iii) Modeling qualities, subjective impressions, and image classes by user goals.

For point (i), our discussions so far suggest the following: An iconographic class should carry a class attribute, e.g. "shows: Saint Paulus" (see pic.8), which is inherited by all instances of this class. Thus consistency with non-formalized representations of Saint Paulus is achieved, and redundant repetitions are avoided. The imaginary may be attached in hierarchy at the real term, it is derived from, and simultaneously classified in a metaclass of type "imaginary":

```
unicorn
  type:
      mythological animal type
  derived from:
      horse
end
```

Thus a duplication of all real concepts is avoided. The inheritance properties of "derived from" in contrast to broader term relations has to be studied. A user may ask for "mythological mammals". A more complex case is that of the king of Benin. On one side, from its origin cannot be concluded, that an image shows him as a characteristic native of Benin. From the latter, it should however be concluded, that it shows a characteristic native of Africa. Nativeness in general can be concluded from the origin. If we do not want to repeat all topological hierarchies for their natives, we must introduce links to the topological hierarchies with the respective inheritance semantic, similar to the unicorn case. These are some examples of hierarchies, which have the potential to generate parallel hierarchies of different nature.

Point (ii) we intend to base mainly on respective work in iconography, whereas (iii) requires extensive experimentation on both the documentalist and end user side.

Conclusions

Images should not be regarded as self-consistent and self-explanatory. Rather, they should be understood as a medium playing certain roles in human conceptualization, understanding and communication on the world around and in us. The specifities of these roles seem to be less bound to subject domains, as traditionally assumed, but more to classes of user goals. Under this consideration, we want to express the optimism, that the semantic indexing method, which tries to approach human cognitive structures, can be by far more successful in the field, if we change our focus from the "picture telling a story" to the human "telling a story", when he sees a picture.

We propose an approach, to bring an index in better correspondence with relevance for the user, and to organize the immense complexity of the necessary index structures. It is also the base of a respective ongoing implementation based on the Semantic Index System. Moreover, we believe, that the distinct features of the SIS provide a functional advantage over existing systems, especially by its metamodeling and integrative capabilities. In the framework of this paper, many important questions could only be touched, and we are looking forward to a fruitful scientific discussion.

References

[Card93] A.Cardenas,I.T.leong,R.K.Taira,R.Barker,C.M.Breant, "The Knowledge-Based Object-Oriented PICQUERY+ Language," IEEE Transactions on Knowledge and Data Engineering, Vol.5, Mo.4, August 1993.

[Chan92] S.K. Chang, Arding Hsu, "Image Information Systems: Where Do We Go From Here?" IEEE Trans. on Knowledge and Data Engineering, Vol. 4, No. 5, October 1992.

[Chan87] S.K. Chang, Q.Y. Shi and C.W. Yan, "Iconing Indexing by 2-D Strings," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 9, No. 3, May 1987.

[Chri94] M. Christoforaki, Cultural documentation with the CLIO system, Master of Science thesis, Department of Computer Science, University of Crete, February 1994 (in Greek).

[Chri95] M. Christoforaki, P. Constantopoulos and Martin Doerr, "Modeling occurrences in cultural documentation," III International Symposium on Computing in Archaeology, Rome, November 1995.

[Cons91] P. Constantopoulos, J. Drakopoulos and Y. Yeorgaroudakis, "Retrieval of Multimedia Documents by Pictorial Content: A Prototype System," Proc. International Conference on Multimedia Information Systems '91, ACM SIGIR, Singapore, January 1991.

[Cons93] P. Constantopoulos, M. Doerr and Y. Vassiliou, "Repositories for Software Reuse: The Software Information Base," Proc. IFIP WG 8.1 Conference on Information System Development Process, Como, Italy, Sept. 1993.

[Cons94a] P. Constantopoulos, "Cultural Documentation: The CLIO System," Technical Report FORTH-ICS-TR 115-1994, Institute of Computer Science, Foundation of Research and Technology - Hellas, January 1994.

[Cons94b] P. Constantopoulos and M. Doerr, "The Semantic Index System: A brief over-view"

[Cons95] P. Constantopoulos and M. Doerr, "Component Classification in the Software Information Base," in 0. Nierstrasz and D. Tsichritzis, eds., Object-Oriented Software Composition, Prentice-Hall, 1995.

[HiII95] G.Hillebrand,P.Klimathianakis, "Querying the Semantic Index System," 7th ERCIM Database Research Group Workshop on Object Oriented Databases, INESC, Lisbon,Portugal, May 15-16, 1995

[Hjel94] R.Hjelsvold, R.Midtstraum, "Modelling and Querying Video Data," Proc. 20th VLDB Conference Santiago, Chile, 1994.

[Kato9l] T.Kato,T.Kurita,H.Shimogaki,T.Mizutori,K.Fujimura, "A Cognitive Approach to Visual Interaction," Proc. International Conference on Multimedia Information Systems '91, ACM SIGIR, Singapore, January 1991.

[Megh95] C. Meghini, "An Image Retrieval Model Based on Classical Logic", Proc. ACM SIGIR Conference '95.

[Mylo9O] J. Mylopoulos, A. Borgida, M. Jarke, M. Koubarakis, "Telos: Representing Knowledge About Information Systems," ACM Trans. on Information Systems, October 1990.

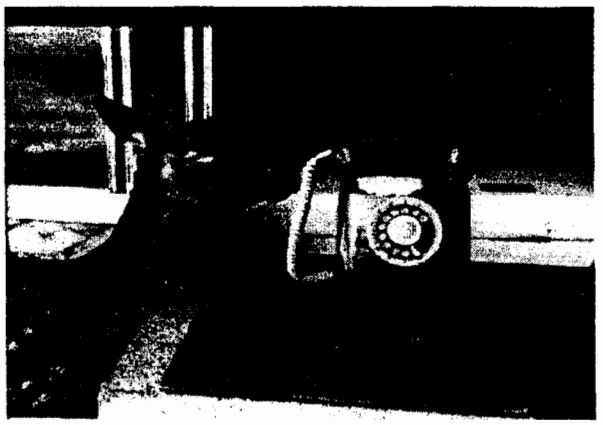
[Oomo93] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, Vol. 5, No. 4, August 1993.

[Pate94] M.J.Patel, "Situation Assessment and Adaptive Learning: Theoretical and Experimental issues," in proceedings of the Second International Round Table on Abstract Intelligent Agent, Rome: ENEA Headquarters. And as ERCIM Research Report ERCIM-03/94-RO30 CNR, Rocquencourt, France (1994b)

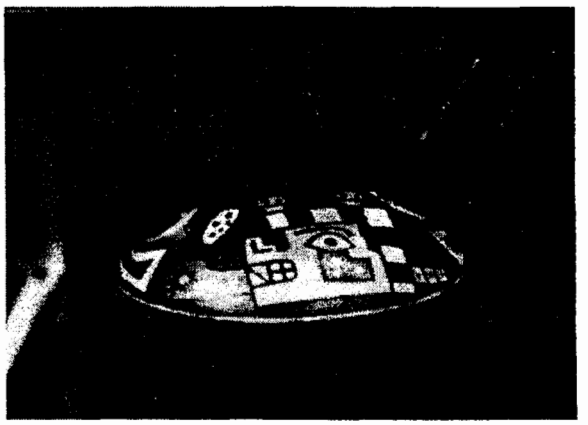
[Pras94] A.Prasad Sistla, C.Yu, R.Harad, "Reasoning about Spatial Relationships in Picture Retrieval Systems," Proc. 20th VLDB Conference Santiago, Chile, 1994.

[Salz94] B.Salzberg, "On Indexing Spatial and Temporal Data," Information Systems Vol. 19, No 6, pp 447-465, 1994.

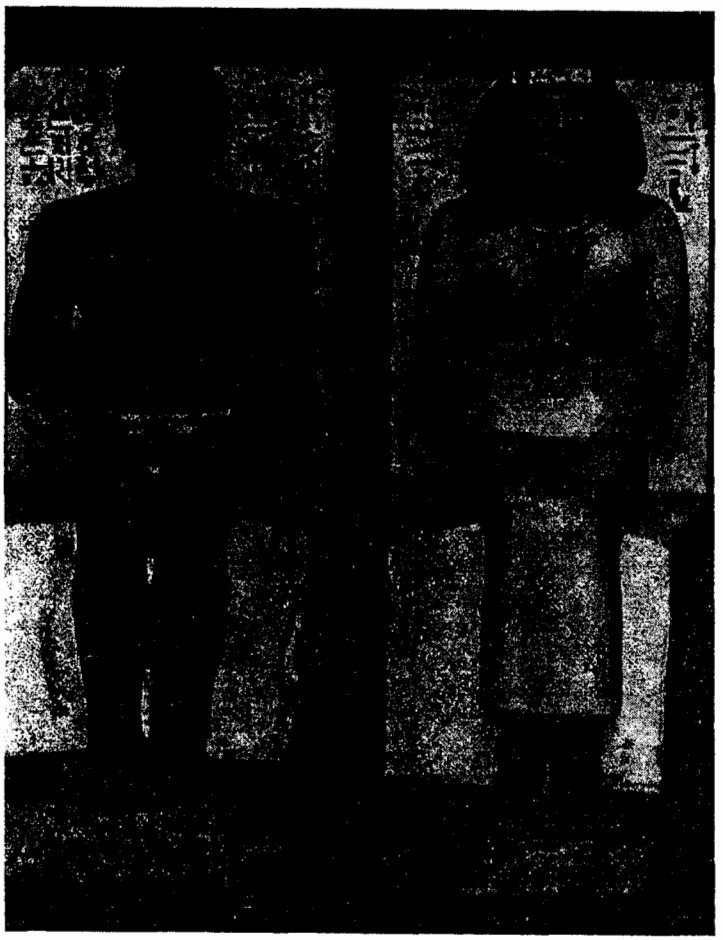
[Sven89] E. Svenonius, *Design of Controlled Vocabularies*, Encyclopedia of Library and Information Science, Marcel Dekker, New York, 1989



PIC. 1

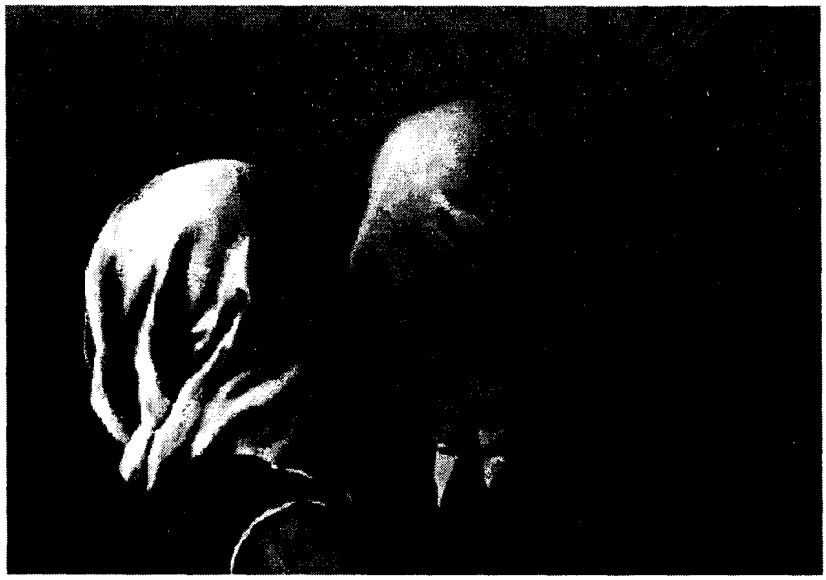


PIC. 2



PIC. 3

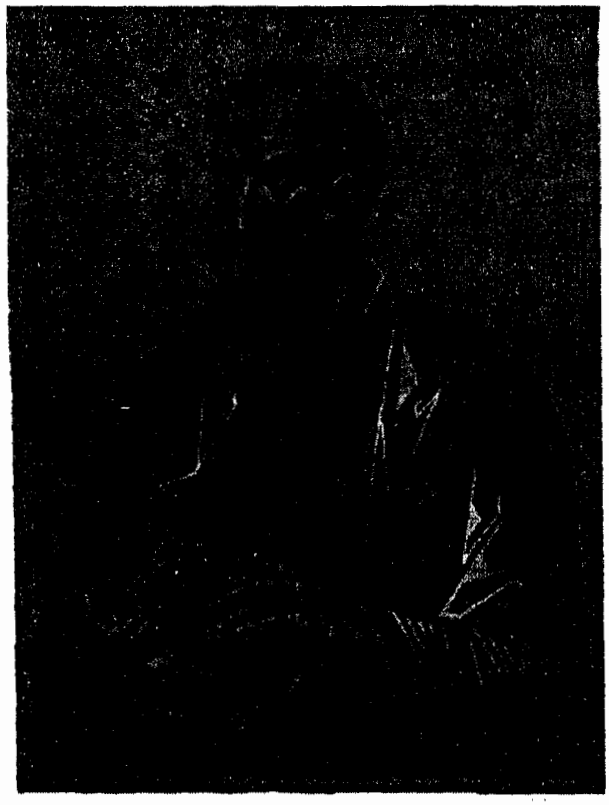
PIC. 4



PIC. 5

PIC. 6

PIC. 7



PIC. 8

PIC.1 Tame deer with red telephone, Nara Park with its traditional tame deers, Nara, Japan, 1983 (Photo: Martin Doerr).

Comment: Documentary snapshot. It does however not illustrate the historical park and its deers. The documentary value seems restricted to the creator. May be suitable for advertizing or comics.

PIC.2 Vase from early Maya period. Museum Rietberg, Zurich, Switzerland, 1978 (Photo: Martin Doerr).

Comment: Documentary image on isolated object. Good prototypicality. Technical quality sufficient for object studies, but not for publication. The decoration is an image in its own right, with standardized meaningful patterns. The interpretation needs a specialist.

PIC.3 Prince Rahotep and his wife Nofret, Medum, Egypt, early 4th dynasty, Kairo, Museum 223 (Taken from: Kurt Lange,"Aegypten,"Hirmer Verlag, Munich 1978).

Comment: Documentary image on isolated object. Publication quality, prototypical. Kurt Lange comments, that the innatural equality of height expresses equality of status.

PIC.4 Luxor, Amun-Mut-Chons-Temple. Ramses II(1290-1224 B.C.) and his wife, queen Nefertari. (Taken from: Kurt Lange,"Aegypten",Hirmer Verlag, Munich 1978).

Comment: Documentary image on an object in its environment. The wife of Ramses appears here in innatural inequality of size with respect to her husband (the small person right behind his knee). A comparison with pic.3 could be subject of a scientific study about a potential change of the social roles of men and women in ancient Egypt.

PIC.5 Rene Magritte, Les Amants (The lovers) Zeisler Collection, New York. (Taken from: D.Larkin,"Magritte,"Pan/Ballantine, London 1972)

Comment: Imaginary Scene. Object features probably irrelevant for retrieval. Work of art. Besides iconographic descriptions, better keywords than "masked persons," may be "psychology of love," "psychology of communication," "human relations."

PIC.6 Maria Loizou wears the traditional wedding dress of E.Tsambazh (Taken from: E.Tsenoglou, "The women dresses of Kastellorizo",Ethnographika, Peloponnesian Folklore Foundation,Vol.4-5,Navplio, Greece, 1983)

Comment: Documentary image on isolated object. Multiple aspects: The actual person, the actual dress, and the prototypicality of the dress for the island Kastellorizo, the latter being predominant.

PIC.7 The "Valide Tzami," Platia Komarou, Heraklion, Crete, Greece, 19th century. Copper or steel engraving, artist unknown. (Property of Martin Doerr)

Comment: Documentary image on a full scene. Many aspects: Daily life, dresses, coexistence of Greek and Turkish, architectural styles, estate of the actual buildings, some of which are lost now, the trees. Many details may become relevant in future studies.

PIC.8 Saint Paulus, by Euphrosynos, 1542CE, Mount Athos, Greece (Taken from: Nikos Psilakis, "Monastiria kai Erimitiria tis Kritis". Vol.1, Stamou Broth., Athens 1992, ISBN 960-220-310-2).

Comment: Byzantine Icon from Crete. Even though it shows a historical person, a byzantine icon intends to render the spiritual world in a standardized (linguistic) way, and has neither documentary intentions nor values. Many details are predescribed, and need not be annotated at the work of art, but at the class as a whole. Besides the rich icongraphic aspects, these paintings play an important religious and social role, which is an aspect of usage not shared with religious paintings in general.