ASPICO: ADVANCED SCIENTIFIC PORTAL FOR INTERNATIONAL COOPERATION ON DIGITAL CULTURAL CONTENT

Jérôme Godard*, Frédéric Andrès, and Kinji Ono

***Abstract**: In this paper, we present the development of an advanced open source multi-lingual cooperative portal system (ASPICO) dedicated to semantic management, and to cooperative exchange for research and education purpose on digital cultural projects. Advantages of using ASPICO include greater flexibility for digital resource management, generic and systematic ontology-based metadata management, and better semantic access and delivery based on an innovative Information Modeling for Adaptive Management (IMAM).*

***Keywords**: Digital Silk Roads, Semantic Management, Metadata Annotation, Image-Learning Ontology.*

1. Introduction

Following the evolution of cultural heritage archives, new requirements for semantic understanding in a multi-lingual and multi-disciplinary cultural fields such as the historical silk roads have been pointed out in major symposiums [Ono 2001, Ono 2003] related to this field. The *Advanced Scientific Portal for International COoperation* (ASPICO) aims at providing a web portal service in order to enable international and multi-disciplinary researchers and fellows to cooperate on research about cultural projects (e.g. the historical Silk Roads project, the visual cultural topic maps online project). The platform is a java-based open source and available for everyone for research and education purposes. The platform provides multilingual semantic extraction services in order to process digital cultural artefacts from a cross disciplinary point of view, based on cooperative annotation support, metadata extraction and classification. It is based on powerful and industry-leading opensource components such as Linux OS, Enhydra, PostgreSQL, and Dspace. It is an independent platform, allowing internal usage (e.g. intranet), external usage (e.g. extranet) and access for general public via the internet. It provides flexible features, allowing easy customization for individual or community needs. Also it provides in a transparent way a fully multilingual support. So searches on the data can be performed on all the information in any language. The platform is standards compliant based on the usage of XML which allows complex data interaction and analysis to take place both within the server and on the client side. The platform has a distributed architecture so autonomous systems can be located in different institutions in order to be consulted simultaneously and to aggregate final results. In Section 2, we introduce the platform architecture. Then we present the different layers from data collection to semantic management and delivery in Section 3. Section 4 describes a function of ASPICO as a case study, the image learning ontology system including its image content recognition features. We review also the state of the art in the field of active contour as a promising solution for shaping understanding implementation inside the image learning ontology. Finally, Section 4 concludes and gives the direction of the future work.

2. The Platform Architecture

The platform architecture (see Fig. 1) is based on open source components including a storage layer (Dspace[1]), an ontology based metadata management, the query interface and resource entry service, and multi-resolution resource viewing. The system limits the access to data according to users rights to indoor users (Intranet), outdoor users (Extranet) and to the Web users.

* This research is partially supported by a grant (*bourse Lavoisier*) from the French Ministry of Foreign Affairs (Ministère des *Affaires Etrangères*).

[1] MIT's Dspace: http://www.dspace.org

2.1 Ontology-based Metadata Management

A key feature of our system is the multilingual ontology-based metadata support. Our platform follows a promising research approach based on the usage of metadata and ontologies. Metadata is any information which characterizes instance data, and which describes its relationship. Metadata is used to provide an effective use of data, in order to facilitate any data management, any data access, and data analysis [Duval 2002].

An ontology is an explicit specification of the conceptualisation of a domain [Gruber1993]. Ontologies enable domain experts to create an agreed-upon vocabulary and semantic structure for exchanging information about that domain. Ontologies facilitate cataloguing and sharing knowledge, as domain expert are able to contribute to a shared, worldwide, but well-organized knowledge base of technical information. We considered a metadata management architecture and designed multi-layer ontologies to classify and describe resources. It is based on protégé 2000[1]. Each ontology is related to one field such as history, geography, architecture, and art… However, possibility of overlapping exists as different ontologies may have equivalent concepts, or may contain subsets of separate ontologies within themselves.

This problem of ontology integration has been solved by classifying and reorganising ontologies in a logical and semantic sense according to metadata. This points to a need for a formal model for ontology-based metadata management. Ontology is the formal and explicit conceptualization of a particular domain. It includes a set of concepts and their relationships. Based on Protégé 2000, we defined our ontology structure as a 6-tuple:

$$O := \{C,P,A,H^C,prop,att\}$$

where C represents a domain-based set of concepts, P a set of relation identifiers, and A a *set of attribute-value relations. H^C is* a *Hierarchy of Concepts* that are linked together through relations (e.g. specialization, generalization). H^c is a directed transitive relation $H^c \subseteq C x C$ called *concept taxonomy*; function prop: P->CxC relates two concepts non-taxonomically; function att: A->C introduces the relationship between concepts and literal values. As an exemple, let us consider a subset of our ontology structure related to spirituality:

C := {SPIRITUALITY,RELIGION,LANGUAGE,OBJECT,LANGUE,BOUDDHISME}, P := {EXPRESS,CREATE} , and A defines the relations EXPRESS(RELIGION,LANGUAGE) and CREATE(RELIGION,OBJECT).

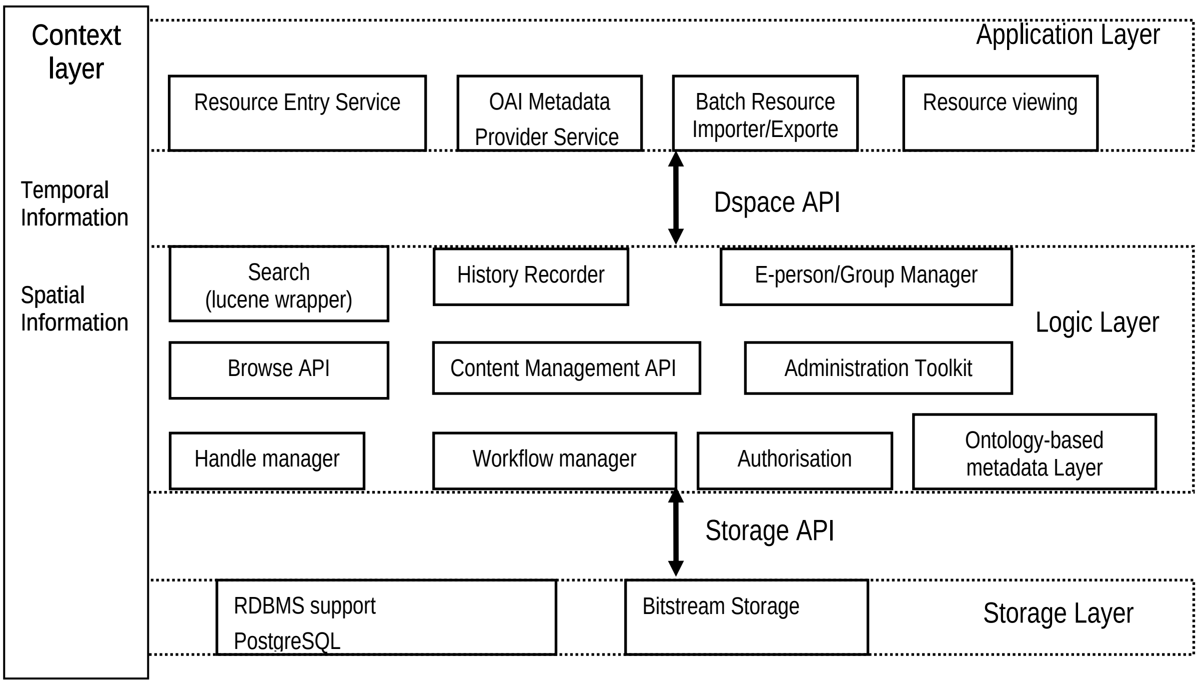

Fig. 1: Architecture of the ASPICO platform

[1] Stanford's Protégé ontology editor and knowledge-base editor: http://protege.stanford.edu/

2.2 Resource Entry

Resource entry is also performed via a web browser interface similar to that used for querying. Users who enter Resource need to log on. Write, modification and suppression rights can be assigned and controlled by the system administrator for each user; some predefined types of user provide community and group management abilities. Information such as the identification profile of the user and date of the entry are automatically filled in by the system. To maintain the integrity of the resource being entered into the system, the controlled lists of relevant vocabulary within the thesaurus are used for each translatable field. When uploading resource via the web interface, users are required to enter some preliminary metadata related to the resource. When a resource is saved into the main database, the metadata is translated into a language independent code representation. The creation of metadata profile is done according to the metadata category such as structural metadata, content metadata and contextual metadata avoiding overlapping between attribute sets.

2.3 Query Interface

Queries are performed via a web browser based interface. Screens for simple or advanced queries can be easily created and the fields to be viewed customized by the system administrator. In addition, date or numeric size fields can be searched by specifying a range of dates or sizes between which searches are performed. Users are able to select the working language and the domain of interest as well as the number of results returned and whether resource results are shown.

The interface is divided in three parts:

1- Historical and material resources related to artifacts;

2- Technical or management information related to photographic resources;

3- Technical or management information related to document resources;

Where applicable, the user can choose technical terms from a list of relevant terms classified alphabetically, or can type something directly in. Ontologies in 21 languages will be able to be consulted on line. Full text searches can be made within each field.

The display or the output format of the results (e.g. HTML, XML, plain text, formatted tabular, list of images, graphical, statistical analyses etc) is independent of the storage structure in order to optimize the delivery process. It typically follows a methodology based on context-dependent cultural resource accesses [Godard 2003].

2.4 Multi-Resolution Resource Viewing

Another key component of the resource management system is the capability to remotely view multi-resolution resources including high resolution images of both 2D paintings and 3D objects. Each image resource is stored as both a JPEG thumbnail for rapid previewing and in tiled pyramidal TIFF format for high-resolution viewing. A java applet permits multi-resolution viewing in conjunction with the storage layer. This viewing system is based on the Internet Imaging Protocol. The viewer works by requesting only the tiles at the appropriate resolution required for viewing a particular part of the image. The requested tiles are then dynamically JPEG encoded by the server and sent to the applet. In this way, images of any size can be viewed quickly across the internet.

2.5 Multilingual Ontology-based Metadata

Multilingual support is becoming very critical in the cultural domain. This can be accounted by: (1) the increasing share of cultural contents accessed over internet, (2) efforts to develop standards for cultural data from diverse fields for the purpose of digital archiving and research sharing, and (3) the increase in use of tools to extract semantic from cultural digital data. Furthermore, multi-lingual Ontology-based metadata approach enables searching by semantic and by contextual content as it relies on multi-lingual annotated documents and features extraction processes. Each set of ontologies is based on an object-identifier bridge and mono-lingual Unicode (UTF-8) encoding ontologies.

Controlled lists of technical terms (e.g. Art and Architecture Thesaurus AAT, Library of Congress Authorities) from each ontology as well as the free text information fields (such as the titles) have been translated with the support of domain experts.

3. From Data Collection to Semantic Management

We describe our solution of collecting information using ontology-based metadata management applied to cultural contents. Metadata are often categorized in three sets: (1) Production-related descriptors, (2) Physical descriptors, and (3) Administrative descriptors. Let us describe the whole acquisition process from the field data collection to the semantic management.

3.1 Data Collection and Contextual Metadata Acquisition

In this section, we describe the semantic model for ontology-based metadata management used to support metadata associated to a set of data collection. The semantic model consists of: (1) the resource layer, (2) the metadata layer, (3) the mapper layer, and (4) the context layer.

A resource r is the base physical representation of data in the metadata management framework. At the resource layer, resources are stored in the form {O,P,T} where O is the object representation of the resource r, P is the parent object of O such as archive or collection, and T is the type or schema of the resource r. data resources are represented as a tree-structure in terms of hierarchical parent-child relations between resources.

As metadata is often considered as synonym of instance of ontologies, our approach follows the formal definition of metadata structure [Guarino 1998]. The metadata structure is defined as a 6-tuple {*O,I,L,inst,instr,instl*} that consists of an ontology *O*, *I* a set of instance identifiers, *L* a set of literal values, a function *inst* called concept instantiation $C\text{->}2^{I}$, a function *instr* defined by $P\text{->}2^{IxI}$ called relation instantiation, and a function *instl*, called attribute instantiation defined as $P\text{->}2^{IxL}$.

As examples from our ontology-based metadata, we provide the following subset of literal values:

{INDIA,FRENCH,MANDARA,CROWN}

inst can then be applied as follows:

inst(INDIA)=COUNTRY, *inst*(FRENCH)=LANGUAGE, *inst*(MANDARA)=OBJECT, *inst*(CROWN)=OBJECT

Furthermore, we can express relations between the instances and an attribute for the OBJECT instances. We give the following examples:

CREATE(BOUDDHISME,MANDARA), CREATE(BOUDDHISME,CROWN)

Mappings store how resources and metadata concepts are linked as well as how metadata layer concepts are linked between different vocabularies. There are two categories of mapping links namely: (1) resources-to-concepts and (2) concept-to-concept. For the resources-to-concepts mapping, we define MappingR2C={type, ER/AR,EC/AC} where type defines the typing of mapping, ER is the set of equivalent resources, AR is a set of resources to be aggregated. EC is the set of equivalent concepts and AR is the set of concepts to be aggregated.

The context layer stores the information related to the environment of the end-users and also the rules to determine if a mapping is applicable. A context rule is defined formally as CR={E,C,A} where E is the set of events, C is the set of conditions, and A is the set of actions to be performed if the conditions evaluate to true.

Examples of contextual values are the geographical location (e.g. *PARIS, GPS location*), time (e.g. *Oct 14th 1992 9:00pm*), and rules (e.g. *Tibet and 14th Century and Opaque watercolors).*

Each resource in our archive has a relationship with a *Resource Categorization Tree* (RCT) where each node label has been defined from a specific classification standard (e.g. ATT for the English culture). Each monolingual RCT is fully dependent on the cultural classification of the related language. In addition, the RCT provides a metadata form describing each resource. For each entity instance, the attributes lists and values (denoted *Resource Description* and *profile* for the other three entities) are then manipulated by database-like operators (creation, index, comparison...) [Godard 2004a], and provide the foundations of an Information Modeling for Adaptive Management (IMAM).

3.2 Resource Provider and Metadata Acquisition

Expert and researchers play a role of resource provider. Their role is to register their datasets using a metadata form. They only input from the research field mainly administrative descriptors (e.g. location address, administrative description) part of the metadata form. Also they can annotate resources according to their point of views and they can share their comments according to cross-disciplinary and multi cultural backgrounds. This registration process generates an update version of the metadata form. Each resource and its related metadata form are the knowledge base of the digital archive.

3.3 Semantic Management

Semantic Extraction

The semantic extraction (shown in Fig. 2) relies heavily on automatic process based on ontologies that provide shared conceptualizations of specific domains and on metadata defined according these ontologies enabling comprehensive and transportable machine understanding. The global process is based on four steps: (1) OCR batch recognition, (2) XML cleaning, (3) XML MPEG-7 metadata descriptors generation and (4) Topic Map Generation.

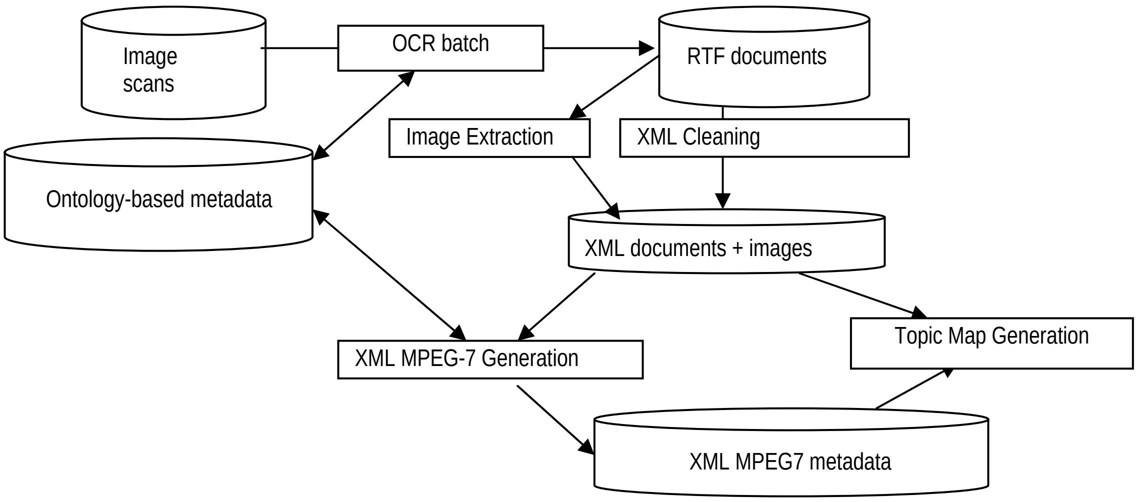

Fig 2: Semantic Extraction Example from Visual and Textual Resources

Semantic Storage, update and versioning

Several research projects such as the arXiv e-print archive[1], the Networked Computer Science Technical Reference Library (NCSTRL)[2] or the Kepler project[3] in the field of digital libraries or digital research archives, tried to solve issues of sharing and semantic storage of research information. They generally provide a common interface to the technical report collections based on the OAI (Open Archives Initiative) infrastructure[4]. This mechanism enables interoperability among large scale distributed digital archives. In many cases, the network environment services include automated registration service, tracking of connected clients, and harvesting service of clients' metadata. Query service enables accesses to resources and to its related metadata.

The Open Archives Initiative (OAI) has created a protocol (Open Archives Initiative Protocol for Metadata Harvesting, OAI-PMH) based on the standard technologies HTTP and XML as well as the Dublin Core metadata scheme[5]. OAI presently supports the multipurpose resource description standard Dublin Core which is simple to use and versatile.

Shortcomings of such research projects generally include a too general metadata attributes schema for fine-grained information (e.g. cultural domains) and the non-support of community building so the semantic management and storage are weak points. However, OAI-PMH itself has been created to provide an XML-wrapper. As part of our project, it has been extended in our project to support multi-disciplinary metadata schemas such as Object ID for historical buildings, CDWA[6] for historical artefacts, or VRA[7] for Visual Resources so the semantic storage has been enhanced. The platform provides service for data handling, registration for

---

[1] Cornell's arXiv.org e-Print archive: http://arxiv.org

[2] Networked Computer Science Technical Reference Library: http://www.ncstrl.org/

[3] Kepler – A Digital Library For Building Communities: http://kepler.cs.odu.edu/

[4] Open Archives Initiative: http://www.openarchives.org

[5] Dublin Core Metadata Initiative: http://dublincore.org/

[6] Categories for the Description of Works of Art: http://www.getty.edu/research/conducting_research/standards/cdwa/

[7] The Visual Resources Association: http://www.vraweb.org/

identification; and semantic handling based on cross-disciplinary metadata schemas to create OAI-compliant metadata and semantic management. All manipulated and exchanged information are in XML. As it has been pointed out in [Westermann 2003], XML database solution is the most appropriate approach for semantic storage and management according to the current state of the art but still issues still remain such as versioning and update propagation. A native semantic database engine will be part of the evolution of the platform.

Semantic Access and Delivery

Let us present how this architecture takes advantage of the semantics in order to provide adaptive processes for data access and delivery. In the case of collaborative environments, we are convinced that the available and well-structured knowledge, which is available through the semantic management, can allow us to provide relevant personalized services. Service-oriented architectures definitely require strong knowledge management frameworks in order to perform automated adaptive tasks. But fully automated service composition still remains illusive [McDermott 2003]. However, by operating within known and defined domains, it is nowadays possible to prepare mappings or wrappers in advance so the automated processes can be efficient. The semantic management layer described previously is the basis to enable ASPICO delivering powerful services dedicated to users and to improve automated data management for communities. Using ontology-based metadata management, ASPICO enhances the ability of personalization by dealing with multi-domain semantic control and a strong algebraic metadata model. We defined a set of four entities (resources, users, communities, devices), which allows us to structure all the available knowledge that can be used in order to process adaptive services. IMAM Services aim at offering personalized data access and delivery for users who are involved in communities related to multi-domain interests. We introduced two main services [Godard 2004b]:

- *Viewpoint*, which is a query optimizer. It selects relevant resources from the query answer depending on environment profile (combination of user and device profiles).
- *Placement*, which dispatches, in an authoritarian way, new resources on devices depending on the correlations between the resource and community or user, and on devices capacities.

In the next section, an image-learning ontology is described as a typical application for our semantic management framework.

4. Case Study: The Image-learning Ontology System

Content-based image retrieval is a challenging and active research area [Jain 1998] with the potential to provide powerful tools for image searching and semantic understanding. Although many techniques have been described in the research literature, the capabilities of current content matching systems are still basic general purpose approaches. Our research focuses on the combination of ontology-based metadata and image content understanding, calling this approach *Image-learning Ontology.* We have been applying this innovative approach to the architecture domain. As it is shown in Fig 3, the image-learning ontology system is based on two phases:

- a training phase where the system learns semantic links between technical terms and image objects.
- a semantic discovery phase where the system proposes to the experts new semantic links between technical terms and image objects.

Large set of digital image archives on architecture at NII allows us to sample and group together images with similar characteristics and categories thereby providing the reference material for testing the behavior of image content recognition. Furthermore cooperation with domain experts provides a multi-lingual support on these categories.

In the remainder of the paper, we detail the image content recognition processing according to semi-automatic and automatic properties. General techniques based on such features as color distribution, texture, outline shape, and spatial color distribution have been quite popular in the research literature and in content based retrieval systems.

4.1 Semi-Automatic Content Recognition of Images

Regarding this aspect, requirements have been determined according to the resource type as part of the resource metadata. If the resource is a document, the recognition process is done for each image included inside the document. Iconography characterization is one the most complex issue in this field so semi-automatic content recognition process based a learning phase and sketching is the most suitable method. Some shape recognition methods work well such as portrait, landscape, buildings, or themes such as crucifixion, or virgin and child. The experts confirm through sketching [Sciascio 1999] the contents of the images.

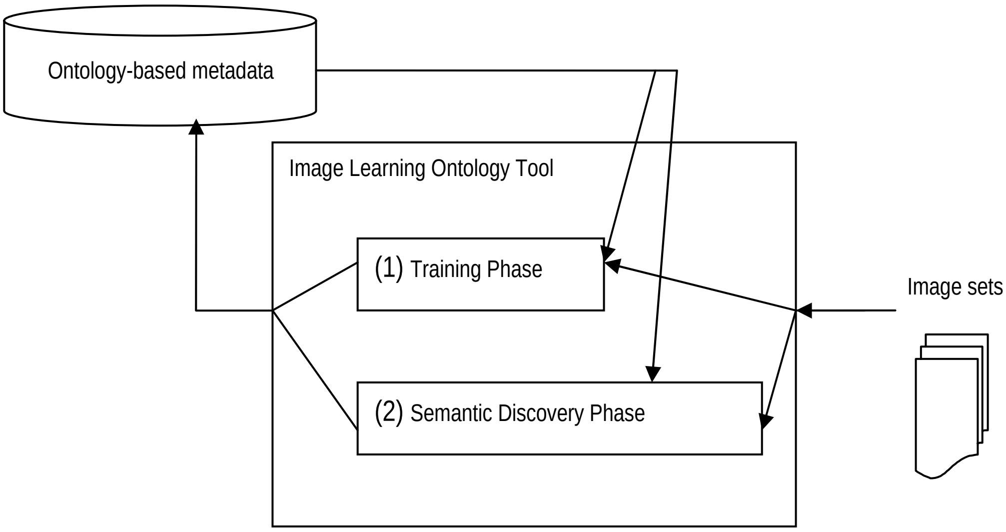

Fig 3: Architecture of the Image-learning Ontology System

4.2 Automatic Shape Identification

One user-requirement for the DSR project is the identification of shapes for painting and buildings in order to provide richer statistics for searches. This is useful for restricting areas of interest and avoiding backgrounds. This is carried out using recognition of deformable models. Research activities concerning deformable models can be partitioned in two types in [Jain 1998]:

- Free-form model, also called active contours, which allows representing any shape by using a minimizing energy algorithm.
- Parametric model allows defining and encoding specific geometric properties of the shape (moments, angles…)

Let us review the two classes of model.

Free-form model

An initial contour, or snake, C is defined by the coordinates $\{x(s), y(s)\}$, 0<s<1. The method was initially introduced by [Kass 1988] and involves the energy-minimization contour C by controlling the three forces:

- The internal forces E_{int}, which define the constraints concerning the shape of the model (more or less smooth).
- The images forces E_{image} which distort the contour according to the variations of pixels values (grey level or colour values).
- The external forces E_{con}.

The willed contour is thus obtained by minimizing the energy given by:

$$E_{snake} = \int_0^1 \left(E_{\text{int}}\left(x(s), y(s)\right) \quad + \quad E_{image}\left(x(s), y(s)\right) + E_{con}\left(x(s), y(s)\right)\right) ds$$

The external forces E_{con} will be not used in what follows.

The internal forces are mainly defined by the coordinates of the snake C:

$$E_{\text{int}}(s) = \tfrac{1}{2}\left(\alpha(s)\left(x_s(s)^2 + y_s(s)^2\right) \quad + \quad \beta(s)\left(x_{ss}(s)^2 + y_{ss}(s)^2\right)\right)$$

Where the subscripts on x and y define derivative form. The coefficients α and β indicate the strength of the elasticity and of the rigidity. In practice, for the digital images applications the problem must be discretized [Davison 2000]. Energies must be sampled at N equally spaced knots v_i around the edge C:

$$E_{\text{int}} = \frac{1}{2h}\sum_{i=0}^{N-1}\alpha_i\left|v_i - v_{i-1}\right|^2 + \frac{1}{2h^3}\sum_{i=0}^{N-1}\beta_i\left|v_{i-1} - 2v_i + v_{i+1}\right|^2$$

In general, the first curve is initialized by B-splines, widely described by [Blake 1998] and used, for instance, by [Stammberger 1999] for a magnetic resonance imaging application.

The image energy E_I depends on the variations of grey level $g(x_i,y_i)$.

$$E_I = k_I\sum_{i=1}^{N} g(x_i, y_i)$$

$$E_{image} = -\nabla E_I$$

The energy-minimization is usually realized iteratively by a gradient descent algorithm until a minimum. Intuitively, a major drawback appears. Indeed, the contour C which depends on the initial position can be attracted by a local minimum, far-off the shape desired. A control of the final contour must be thus checked. Many approaches have been proposed to erase these problems in [Jain 1998], and in [Tsechpenakis 2004]. Moreover the method fails sometimes for very complex shapes. Nevertheless, the method remains very powerful for image segmentation and its implementation is very fast.

Lastly the use of colour information allows improving the performance of active contours. [Ngoi 1999] proposed thus a new active contour model for shape extraction of images acquired in outdoor conditions.

Parametric models

The active contours are based on an energy minimizing calculated from the coordinates of the pixels belonging to the contour; the basis of the parametric models is the study of the shape deformation by using geometrical parameters. The model needs now more specific a-priori knowledge of the shape. We can differentiate two parametric models [Jain 1998]:

- Analytical deformable models which are defined by analytical curves.
- Prototype-based deformable templates defined an “average” shape of a class of objects.

Let us comment these two models:

- Analytical deformable templates

In those methods, templates are defined by parametric models such as ellipses, or circle parametric function. The model, which possesses only few degrees of freedom, fit the desired shape by energy minimizing applied to the model parameters. The most popular example of such a method is the eye template of [Yuille 1992]. In this model, parameters for which the variations are carried out are the centre and the radius of the circle and the coefficients of the parabola. In this model, we distinguish also two kinds of energy, the internal energy which is defined by a parametric function characterizing a shape and the external energy which represents the features of the image. Minimizing energy algorithm is then used. Because of the parametric function is chosen previously, the analytical deformable templates are required for segment objects with a known shape.

- Prototype-based deformable templates

For the prototype-based deformable templates, a particular model is previously built according to the shape we want to extract (for example a model of a sculpture or a model of a building). The performance of the prototype template depends, obviously, on the description of the shape. Recent research works have adopted learning method from a set of samples. So as to do it, [Cootes 1994] and [Cootes 1995] have thus used this kind of methods. From those samples templates, a mean shape is calculated and used as the generic model and the variations are determined by eigenvectors of the covariance matrix. Other deformable templates based on a prototype have been also described in [Jain 1998].

Digital Silk Roads archives contain high resolution colour images acquired in outdoor (high luminosity, reflections, shadows). If the natural conditions of imaging and the variability of the conditions complicate the segmentation, the images to be segmented present objects with particularly simple shapes (doors, roofs, mosaics, arch…), making the use of deformable models easier. The main advantages of active contours are the speed and the flexibility. Specific a-priori knowledge is not required. On the other hand, it seems that the deformable templates are very adapted for locating specific structures in the images but this method needs a more specific knowledge about the shape we want to extract. Another difficulty is to define accurately objects. The segmentation tool must be precise for segment small objects such as frieze; tiles, lock and writings without over segment the images by defining objects without any architecture signification. A quite "supervised" process is thus preferable.

5. Conclusion

This paper introduced the ASPICO platform as a semantic advanced archive management system. It described the knowledge structure and the interface based on multilingual ontology-based metadata management. Introducing the image-learning ontology system, we investigated the possible solutions for automatic shape identification and then motivated the appropriate strategy to be adopted by cultural digital resource archive in order to perform indexing and efficient retrieval on digital cultural content through ASPICO. In a next step, we are planning to add innovative functions related to real time semantic acquisition based on remote control data acquisition.

Bibliography

[Blake 1999] Blake A. and Isard M. (1999). *Active contours*. Springer. 352 p.

[Cootes 1994] Cootes T.F., Hill A., Taylor C.J., Haslam J. (1994). *Use of active shape models for locating structures in medical.* In: Image Vision Computing 12(6), p.355-366.

[Cootes 1995] Cootes T.F., Taylor C.J., Cooper D.H., Graham J. (1995). *Active shape model – their training and application.* In: Computers Vision Image understanding 61(1). p.38-59.

[Davison 2000] Davison N.E., Eviatar H., Somorjai R.L. (2000). *Snakes simplified.* In: Pattern Recognition 33. p.1651-1664.

[Duval 2002] Duval E., Hodgins W., Sutton S., Weibel S. (2002). *Metadata Principles and Practicalities.* D-Lib Magazine April 2002 Volume 8 Number 4 ISSN 1082-9873.

[Godard 2003] Godard J., Andres F., Grosky W., Ono K. (2003). *Management of Geomedia Content: Context-dependent Data Access*. NII Journal No. 7, p.9-17.

[Godard 2004a] Godard J., Andrès F., Grosky W. (2004). *Knowledge Management Framework for the Collaborative Distribution of Information*. DataX (EDBT Workshop), Heraklion, Crete, Greece, p.2-16.

[Godard 2004b] Godard J., Andrès F., Andaroodi E., Maruyama K. (2004). *Towards a Service-oriented Architecture for Collaborative Management of Heterogeneous Cultural Resources*, Digital Library Architectures, Sixth Thematic Workshop of the DELOS EU Network of Excellence, S. Margherita di Pula, Cagliari, Italy, p.183-194.

[Guarino 1998] Guarino N. (1998). *Formal ontology and information systems*. In N. Guarino, Ed., Proceedings of FOIS `98, Trento, Italy, June. IOS Press, Amsterdam, 1998, p.3-15.

[Gruber1993] Gruber T.R. (1993). *Towards Principles for the Design of Ontologies Used for Knowledge Sharing*. In Roberto Poli Nicola Guarino, editor, International Workshop on Formal Ontology, Padova, Italy.

[Kass 1988] Kass M., Witkin A., Terzopoulos D. (1988). *Snakes: active contours model*. In: International Journal of computers vision 1 (4), p.321-331.

[Jain 1988] Jain A.K., Zhong Y., Dubuisson-Jolly M-P. (1998). *Deformable template models: A review*. In: Signal Processing 71, p.109-129.

[McDermott 2003] McDermott D. V. (2003). *PDDL2.1 - The Art of the Possible? Commentary on Fox and Long*. J. Artif. Intell. Res. (JAIR) 20: p.145-148.

[Ngoi 1999] Ngoi K.P., Jia J.C. (1999). *An active contour model for color region extraction in natural scenes*. In: Image and vision computing 17, pp.955-966.

[Ono 2002] Ono K. (2002). Proceedings of the Tokyo Symposium for Digital Silk Roads, National Institute of Informatics, Tokyo, Japan, ISBN 4-86049-007-X, 321p.

[Ono 2004] Ono K. (2004). Proceedings of the Nara Symposium for Digital Silk Roads, Nara, Japan, ISBN 4-86049-024-X, 510p.

[Sciascio 1999] Sciascio E. D., Mongiello. M. (2003). *Query by sketch and relevance feedback for contentbased image retrieval over the web*. Journal of Visual Languages and Computing, 10(6) p.565-584.

[Stammberger 1999] Stammberger T., Eckstein F., Michaelis M., Englmeier K-H., Reiser M. (1999). *Interobserver reproducibility of quantitative cartilage measurements: comparison of B-spline snake and manual segmentation.* In: Magnetic Resonance Imaging, 17(7). p.1033-1042.

[Tsechpenakis 2004] Tsechpenakis G., Rapantzikos K., Tsapatsoulis N., Kollias S. (2004). *A snake model for object tracking in natural sequences.* In: Signal processing: Image communication 19, p.219-238.

[Westermann 2003] Westermann U., Klas W. (2003). *An analysis of XML database solutions for the management of MPEG-7 media descriptions*. ACM Computing Surveys, Volume 35 (4), p.331-373.

[Yuille 1992] Yuille A., Hallinan P., Cohen D. (1992). *Feature extraction from faces using deformable templates.* In: International Journal of Computer Vision, 8(2), p.99-111.

Authors' Information

Jérôme Godard – National Institute of Informatics (NII), Graduate University of Advanced Studies (Sokendai) Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: jerome@grad.nii.ac.jp

Frédéric Andrès – National Institute of Informatics (NII), Graduate University of Advanced Studies (Sokendai) Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: andres@nii.ac.jp

Kinji Ono – National Institute of Informatics (NII), Waseda University Hitotsubashi 2-1-2 Chiyoda-ku, Tokyo 101-8430, Japan; e-mail: ono@nii.ac.jp

MPEG-7 BASED IMAGE RETRIEVAL ON THE WORLD WIDE WEB

Rajeev Agrawal, Farshad Fotouhi, Peter Stanchev, and Ming Dong

Abstract: Due to the rapid growth of the number of digital media elements like image, video, audio, graphics on Internet, there is an increasing demand for effective search and retrieval techniques. Recently, many search engines have made image search as an option like Google, AlltheWeb, AltaVista, Freenet. In addition to this, Ditto, Picsearch, can search only the images on Internet. There are also other domain specific search engines available for graphics and clip art, audio, video, educational images, artwork, stock photos, science and nature [www.faganfinder.com/img]. These entire search engines are directory based. They crawls the entire Internet and index all the images in certain categories. They do not display the images in any particular order with respect to the time and context. With the availability of MPEG-7, a standard for describing multimedia content, it is now possible to store the images with its metadata in a structured format. This helps in searching and retrieving the images. The MPEG-7 standard uses XML to describe the content of multimedia information objects. These objects will have metadata information in the form of MPEG-7 or any other similar format associated with them. It can be used in different ways to search the objects. In this paper we propose a system, which can do content based image retrieval on the World Wide Web. It displays the result in user-defined order.

Keywords: XML, MPEG-7, Metadata, Multimedia, Content Based Image Retrieval (CBIR)

1. Introduction

The CBIR has been a very active research area in the last decade. Conventional content-based image retrieval systems [1, 2, 3] use low-level features such as color, texture, shape, automatically extracted from the images. Another focus of this research is on improving the low level features. The modifying the similarity measures make the retrieval as better as possible. It is argued in [4] that unconstrained object recognition is still beyond of current technology. The content based systems can at best capture only pre-attentive similarity, not semantic similarity. So far there has not been a single system, which can perform this task automatically without human intervention due to the nature of this problem.