A database of spatialized and semantically-enriched iconographic sources for the documentation of cultural heritage buildings

Chiara Stefani, Chawee Busayarat, Julie Lombardo, Livio De Luca
Ecole Nationale Supérieure d’Architecture de Marseille
UMR CNRS/MCC 3495 MAP Laboratory
Marseille, France
{chiara.stefani},{livio.deluca} {julie.lombardo}@map.archi.fr; cha_v_mek@hotmail.com

Abstract— Conservation and valorization of cultural heritage buildings can be improved by providing public with access to information related to cultural sites. Particularly, the development of tools adapted to all kinds of audience and suitable to the requirements of the visit can create awareness among the general public about conservation. This article describes LOCUS IMAGINIS, an interactive and collaborative platform providing access to cultural information about monuments through personal snapshots. This platform enables to collect photographs taken by visitors. First of all, pictures taken by visitors are added into the database. Then, they are positioned in the 3D scene on the digital model of the monument and they are semantically annotated. As a consequence, this tool provides access to textual and graphic information describing the photographed site (semantic, spatial, historical entries, etc.). When visitors consult the database and select 2D images, these are superimposed onto a three-dimensional scene. Subsequently, visitors can interact both with images and the 3D-model to highlight building parts by color and to display textual information. Accessible on-site, LOCUS IMAGINIS is a support proposing an innovative approach in the case of educational workshops. Available on-line, visitors can extend their visit by recreating their journey through their own photographs and enrich it with information derived from the database.

Keywords—spatio-temporal modeling; image-based modeling; semantic annotation; architectural heritage; information system

I. INTRODUCTION

One of today solutions for preserving and adding value to cultural heritage buildings is to develop tools to access information relating to them. These tools can reveal particularly useful to the exhibition offer and adapted for all audiences. Digital modeling is the basis of most multimedia devices conceived for virtual visits. So, the new digital tools and techniques can now produce three-dimensional representations of the current state of the building as well as hypothetical restitutions of its previous states. However, beyond three-dimensional data, a large amount of iconographic sources (photographs, drawings, paintings, etc.) describe a building throughout its history. These documents are difficult to consult because databases dedicated to their record and consultation are generally based on criteria of qualitative ranking (such as the name of the building, the type of represented items, the author of the source, the type of support, the representation techniques, etc.). As a consequence, finding the way through a heterogeneous corpus of information, without having a deep knowledge of the morphology of the building, is a real drawback.

Actually, different types of representations can refer to various parts of the building at different scales (volumes, typed elements, elements of detail, decorations, etc.) and also at different temporal states. Despite the heterogeneity of this corpus of information, iconographic sources essentially represent the morphology of buildings. Three-dimensional digitalization of buildings produced by 3D laser scanning has the great advantage of providing a comprehensive and metrically correct representation of its morphology. Therefore, today it is essential to study how three-dimensional representations and iconographic sources can be found within a web browser for the visual retrieval and navigation that uses the morphology of the building as interface to access information.

It is from this need that the project Locus Imaginins is born. Based on a previous work conducted by the MAP laboratory on the actual state of Comtal Castle of Carcassonne in 2004-2005, this project builds on an agreement of scientific cooperation between the “Centre of National Monuments” (CNM) and the MAP laboratory.

The aim of this project is to create a spatialized iconographic database including some interactive features available at any time to the public, both displayed on-line and integrated to the exhibition of the monument on-site. This project, structuring iconic information in consistent way and combining it to an easy and fun consulting interface, seeks to allow the public to access and to contribute to the pictorial history of the monument.

This article has been divided into six sections: section II begins by laying out the theoretical dimensions of research, examining some methods for the structuring of the 3D-model and the spatial referencing of images; sections III and IV
present respectively the adopted general approach and the virtual platform developed in order to consult iconographic documentation; session V defines the user of the platform, and finally the last paragraph evaluates the system, assesses its limits and fixes some research perspectives.

II. RELATED WORKS

Today virtual tools allow us to guide modeling operations according to the goal of the analysis specific to each cultural heritage site. However, the 3D-model by itself does not allow capitalizing and managing the entire knowledge fostering the understanding and analysis of cultural heritage. As a consequence, some research begins, firstly, to structure information about the building morphology from a spatial and temporal point of view, and secondly, to link the documentation produced during the analysis stage, by creating interfaces for handling this data set and linking it to a morphological representation of the building.

Currently some research focuses on structuring the 3D-model in space and time. If some studies suggest a semantic annotation process based on ontology [1], others extend this concept to the archaeological field [2] and to the urban one [3]. However, very few studies point to describe relations among several temporal states: generally this research structures the 3D-model and isolates geometric reconstructions stored in databases so to formulate temporal queries [4-6], and some of them describe building lifecycles to stress different states and transformations [7], [8].

Regarding the linking of the model with other graphical documents, today several studies explore the spatial referencing of iconographic data (such as photographs, drawings, paintings, etc.). They can lay on manual methods based on the knowledge [9] and on the registration of the position of the cameras [10], on semi-automatic methods based on geometric solutions for the calibration and orientation of the cameras [11], or still on automatic methods based on analysis and image processing. These last methods estimate the position and orientation of the cameras, starting from the automatic extraction of vanishing points [12] or homologous points identified on the image [13], [14].

Finally, little research has developed interfaces for the management of the iconography collection in space and in time: some of them [15], [16] allow navigation in the 3D space reconstructed from the spatialized high-resolution photos. However, other studies integrate the temporal dimension based on a constraint satisfaction method, by ordering chronologically the spatialized photos and by automatically applying textures on 3D-models [17].

Clearly, these works have shown that the retrieval of the body of knowledge fostering the comprehension and analysis of heritage buildings can be significantly improved, on one hand, by connecting the iconographic collection to the building, and on the other hand, by semantically annotating the morphology of buildings non only as a whole, but also in terms of their parts and subparts, their attributes and relations. However, one major problem remains open. Today the semantic relation between the 3D-model and the collection of spatialized images is lacking. The use of semantics could become a support for displaying measurements made on the accurate 3D-model, information collected about analytical data, or the conservation state of the building.

III. MAIN APPROACH

The main idea of this work is to organize the iconographic documentation and to link it to a semantically structured 3D-model, characterized by elements defined by geometric information hierarchically structured, classified according to a vocabulary of terms, and having assigned quality attributes, as defined in [18]. Therefore, the iconographic spatialized sources inherit a semantic annotation from the 3D-model.

The approach, already detailed in [19], permits:

- referencing iconography in space and in time. This is based on the knowledge of the geometric models of the cameras, and on the association of temporal attributes depending on the period of the sources.
- structuring the 3D-model according to spatial and temporal criteria. This is possible by assigning attributes, and is based on tools developed to assist the user in the structuring task.
- annotating iconography from a semantic point of view. This is possible by means of a semantically structured 3D-model and the projective relations existing between the 3D-model and the 2D sources.

The adopted approach is based on the idea that the 3D-model should be structured from a semantic point of view. The semantic description consists in three main levels described in detail in [20]. The first level concerns the decomposition, which is the morphological organization of the building into elements and sub-elements (in which geometry is associated to each unitary concept). The second one is the structural level, and defines the relations between such parts (creating hierarchical relations between entities). Finally, the last one is that of representation, that associates a representation to each element according to the various types of data processing. The three following sections detail the three aspects of the process: the referencing of sources, the semantically structured modeling and the semantic annotation of images.

A. Spatial and Temporal Referencing of Sources

This method consists of two mains steps. The first one is to establish a projective relation between iconography and the 3D-model of the building current state by means of a spatial resection procedure. Sources can follow different projection rules [21]. Various geometric models of cameras are handled according to the types of images:

- Perspective projection. In the case of photographs and drawings in perspective, a set of correspondences established between the photograph (2D) and the current state model (3D) are necessary to determine the camera geometric model associated to the photograph [22]. It uses the focal length and distortion (intrinsic parameters) and the translation and orientation (extrinsic parameters) related to the geometrical model of the image. For the estimation of the spatial resection
we use a versatile camera calibration algorithm [11] needing at least 11 2D/3D correspondences with the intent of providing intrinsic and extrinsic parameters of the camera. A detailed description of this approach has already been presented in [19].

- Other projections. For drawings such as technical ones (plans, elevations, cross sections), that are based on cylindrical projection, 3 correspondences 2D/3D are sufficient to determine the scale of the iconographic source in the 3D work space and to determine the projection reference plane. Once the plane is defined, it can be shifted along the projection line. In the case of pseudo-perspectives/pseudo-axonometries, a manual method is adopted. It consists in manually referencing the image with the support of existing elements. Of course for this kind of drawings a correct 2D/3D correspondence is not possible.

The second step is to use a procedure that allows the user to attach temporal attributes to the image. These attributes include, according to the kind of image, the creation date and the state described by the image. Creation date can be precisely known e.g. for snapshots, otherwise it is manually inserted. The state, representing a period which artifact does not undergo changes in, is characterized by starting and ending dates, manually inserted by the user (if known).

B. Structured Modeling Based on Iconography

The structuring of the 3D-model according to spatial and temporal criteria relies on the approach detailed in [19]. This is an approach of structuring and reconstruction from images referenced in space and characterized by temporal attributes. In the case of Comtal Castle, the collected data concerned the ones of 3D laser scanning and photogrammetry. The survey campaign was realized in 2004 and was constituted of 25 stations, including 153 scans and the acquisition of 380 photographs. Laying on these data, a restitution of the actual state has been made by means of the hybrid approach combining photogrammetry principles and parametric operations detailed in [23].

The 3D-model made up the general support for the operations of space-time iconography referencing (described in III.A) like for the semantic structuring step. Particularly, the structuring step according to time (Fig. 1) was possible by using several techniques depending on the available kind of document.

- In the case of old photographs and ancient drawings, the 3D representation has been compared to the 2D spatialized image and, according to the visual information contained on the image (specifically on the morphology of each building of the castle) some parts have been modeled.

- In the case of actual high resolution photographs, the 3D representation has been compared to the 2D spatialized image and structured according to time stratifications readable on the surface of each building.

Figure 1. Temporal stratification on the 3D-model of the Castle, using high resolution photographs.

This process was facilitated by the customization of modeling tools permitting to automatize the task of assigning spatial and temporal attributes while geometry is created, modified or deleted [24].

C. Semantic Annotation of Iconography

Most of the current research in semantic annotation uses manual methods [25], [26], allowing users to associate descriptive keywords to images stored in a database, or automatic methods [27], based on content-based image analysis techniques.

In contrast with the works in this area, our approach on semantic annotation of images does not lean on the direct annotation of semantic concepts on specific areas of each image [28], [29]. 3D-models annotated semantically at the level of the building morphology are used as medium to assign semantic information to images. The relation between 3D-models and 2D images is created by means of procedures for the spatial referencing of images (see Section III.A). By aligning images to the 3D-model, the 3D-model contour is projected on the 2D image with the goal of superposing a semantic layer on the original image. This layer is produced by projecting the spatial extension of the 3D-model which is associated to morphological entities of semantic description (Fig. 2). Each projection (computed through a vector rendering technique detailed in section IV.C) is processed as a 2D polygon associated with the identifier of the related 3D entity.
The projective relation between 3D representations and the spatialized iconographic sources provides automatic procedures for adding and updating semantic annotations:

- In the case of addition or change of the 3D morphological description, the new structuring description is re-projected on images.
- In the case of spatial referencing of a new image, the building morphological description will be automatically projected on images.

As described in the next session, semantic annotation of images by projection generates structured data well-adapted for searching iconographic sources relating to the architectural elements composing buildings.

IV. THE INTERACTIVE AND COLLABORATIVE DEVICE

In order to query the system, an interactive device has been developed. It gathers and queries semantically structured data (sections IV.A and IV.B). The system architecture is composed of three main elements:

- A database containing the iconographic sources inserted by the institutions, the photographs inserted by the general public, and the 3D representations of the building related to several temporal states resulted from 3D scanning and photogrammetry acquisition.
- An interface for the consultation of the database (by themes, attributes, types) and for the administration of the insertion and referencing of iconographic sources (historical photos, posters, postcards, drawings, prints, paintings, etc.) by taking into account different types of projections (perspective, axonometry, plans, elevations, sections, etc.).
- A 3D real-time browser interface for the visual retrieval of iconographic sources according to spatial and temporal criteria.

The web application is built laying on a MySQL database (containing the three-dimensional models, the bi-dimensional images accompanied by coordinates on their spatial referencing, as well as descriptive notes concerning the architectural elements of the building and the images). A graphic bar permits to display, scroll and manipulate images and to display them into a real time 3D scene based on the 3DVIA plug-in. The platform is compatible with some of the common browsers (Internet Explorer and Safari) on Windows and Mac OS X platforms.

A. Database

The database organizes four types of information: semantic data, historical information, iconographic data, and 3D representations.

- Iconographic sources inserted by cultural institutions (mainly the Centre of the National Monuments). If images are already contained into an online database, a URL-type link allows the connection.
- Digital photographs inserted by the general public. These photographs are automatically filled in with regard to the EXIF metadata relative to the digital camera, and manually filled in by web users with regard to the information that links the photograph to its author.
- Three-dimensional representation of the building is organized according to different temporal states. Its spatial distribution bases on a common reference used also as geometrical reference for the spatial alignment of iconographic sources.
- The values of spatial referencing of iconographic sources calculated using the web interface (see the next paragraph), consisting in a set of coordinates on the camera position, and parameters of distortion and focal length.

B. Spatial Referencing of Iconographic Sources

The 3D interface for the real-time spatial referencing relates the iconographic sources with the 3D representation of the building. Spatial referencing is based on the identification of the types of representation (current photographs, old photographs, drawings of survey, sketches, etc.) and the types of projection (perspective, axonometry, plan, elevation, section, etc.) defined in the paragraph III.A. The types of representation and projection determine the approximation levels of spatial referencing. In the browser, the referencing is done manually: after having configured the projection type, the user interactively manipulates the 2D plane of the iconographic source in the 3D space. The user can change the position and orientation of the viewpoint relating to the iconographic source according to its spatial superposition on the 3D-model of the building (Fig. 3).
In the early developments of the interface, a first set of images has been inserted into the database and spatially aligned to 3D-model of Comtal Castle: 322 photographs including 192 belonging to the funds of the Center of National Monuments, and 130 photographs acquired in situ during the survey stage of the building. To date, more images have been inserted and spatially positioned in the database thanks to the contribution of the CNM teams (Direction of cultural development and of publics, cultural and educational sector of Comtal Castle of Carcassonne).

C. Procedure of Semantization of Images

This process consists of four steps: the compilation of data into the database, the reconstruction of image viewpoints, the projection of semantic layers, the adding of additional attributes and interactive behaviors in the semantically annotated images (Fig. 2). Concerning the technical development, the system of semantic projection is based on the communication between the MySQL database and the 3D modeling software Maya. This communication permits to collect the necessary information from the database and to generate the geometric models of the cameras associated with the aligned images in Maya (which the polygonal representation of the 3D scene is loaded in). In order to create a vector image of the 3D scene, a library of rendering functions called MayaVector has been used. The developed MEL procedure launches an automatic rendering of the 3D scene starting from each reconstructed virtual camera (related to the viewpoints of spatially referenced images). This library enables to easily project the 3D element silhouette on the entire image and to store it in the database. This produces a vector image in SVG format, whose viewpoint corresponds to that of the original image. In this image, to each projected region (SVG polygon) are associated the semantic attributes stored in the database. The rendering procedure is programmed so to calculate only the contours of the external edges of the object in the scene (including the edges of intersection among the various objects) and to ignore other polygonal edges.

D. Browser Interface for Visual Retrieval and Navigation

The browser interface for the visual retrieval/navigation in the iconographic corpus relies on a web architecture that integrates three parts: the database, an interactive 3D scene, and a set of dynamic pages.

This system permits to navigate in the 3D scene in real time, to locate the viewpoint of iconographic sources on the 3D scene, and to display 2D iconography overlaid to 3D representations of the building. Moreover, specifically to the level of display on the 3D scene, various actions are enabled by this system:

- to display the 3D-model of the monument according to its historical states and the current one, by color coding. This is possible by associating time attributes to the parts of the 3D-model and displaying them by different color parameters (Fig. 4).
- to graduate the 3D visualization of the model, by means of the variation parameters concerning images, volumes and edges.

At the functional level, this application allows positioning the images of the castle (photographs, engravings, drawings, paintings, etc.) on the 3D building representation and looking for iconographic sources inserted into the database according to three main criteria:

- **Search by 3D viewpoint.** This type of search allows the user to navigate in real time through the 3D virtual environment of the monument and to detect the images whose frustum corresponds to the one of a virtual observation camera (Fig. 5). This search uses information contained in the geometric representation of the monument and in the geometric models of the cameras associated to each image (determined in the phase of spatial referencing). Based on the continuous comparison between the projection parameters of the observation point in the scene and those of iconographic sources, the browser interface enables to switch from the image of the building (2D source) to its morphology (3D representation).
- **Search by source type.** This type of search helps identifying the images stored in the database depending on the type of iconographic source (photograph, drawing, painting, and engraving). It depends on the attributes specified by the user at the time of inserting the image in the database (Fig. 6).
- **Search by term.** This type of request enables to find all the images in which an architectural element, specified by a vocabulary term, is represented (Fig. 7). This search uses the semantic decomposition of the 3D representation of the building. This decomposition is then projected on all the 2D images (spatially referenced) by adding over them a semantic layer containing the identifier and the semantic attributes associated to the represented elements (Fig. 8-9).
Figure 4. Display of the 3D-model by colors according to the different temporal states.

Figure 5. Search by 3D viewpoint. At the selection of an image on the bottom bar, the camera of the visual browser moves on the viewpoint of the selected image and displays on the bar images detected in the visual pyramid of the browser camera.

Figure 6. Search by type of source. At the selection of a cross section, the 3D-model is displayed in cylindrical projection.

Figure 7. Search by term. At the selection of an annotated image, the semantic annotation can be highlighted on the 3D scene at the mouse rolling over.

Figure 8. Search by term. At the selection of a specific term, the 3D-model displays by coding colors 3D elements belonging to the selected category.

V. DEFINITION OF USES

The policy and development of this database has been chosen according to three main uses, defined with the CNM. This 3D database addresses to different user profiles for different purposes:

- **On-line and on-site use, as a complement to the tour.** Visitors (general public) can use the database on-line to prepare the visit of the monument and to extend it, by setting their own photographs taken during a visit or on-site, as a complementary cultural offer for the visitors to the Castle. In this way, the public can also contribute to the enrichment of the iconography of the building.

- **In-situ use in the context of educational workshops** Young students, within the framework of classroom activities, can explore new modalities for understanding the morphology and its relation with the history of a monument. The exercise of finding the
viewpoints of the photographs taken during the visit within the 3D-model allows the construction of spatial reference points essential to the understanding of the monument. Particularly, the project Locus Imaginis has been integrated into the partnership agreement between the CNM and the Jules Fil high school of Carcassonne (pilot school for new technologies), as innovative educational program for the cultural heritage approach tested by students.

- **On-line use for the collection and storage of iconographic sources.** Documentalists and researchers (historians, architects, archaeologists, etc.) interested in the consultation of the iconographic corpus related to heritage buildings can use the database to deepen their knowledge on the building morphology, having the opportunity of accessing iconographical sources starting from the exploration of 3D representations relating to different historical periods.

The browser interface developed for this project is available on-line to general public at the following web page: http://vinci.gamsau.archi.fr/htdocs_locus2. Moreover, a video explaining the process and guiding the user is available at http://vinci.gamsau.archi.fr/htdocs_locus2/video/.

### VI. CONCLUSIONS

This work described a system for storage of various kinds of images. It lays on the spatial referencing of iconographic sources on the 3D space and their segmentation into semantic layers. This project introduces an innovative modality of visual search, based both on the positioning of a virtual observation camera in the 3D scene and on content (architectural features, semantic attributes, temporal attributes, etc.).

Despite the results obtained with this project, some issues need to be resolved and some reflections should prompt further research.

Firstly, a technical limit concerning the focal length. The observation of images of close-ups makes it difficult to understand the context, since the focal length of visualization of the model corresponds to that of the image. It would be wise to obtain a simultaneous scaling both of the focal image and of the 3D-model. Conversely, when an image presents a distant point of view of the castle, the 3D-model appears excessively far. In this case, it would be better to “lock” the focal length of the image on the 3D-model so as to modify the distance of the observation point without disconnecting the two elements image / 3D-model.

Secondly, concerning the technological aspects to develop, the introduction of PC-tablet (such as iPad 2) integrating digital cameras, could make it possible to access the features of Locus Imaginis (both for research of iconographic sources and for the insertion of new photographs) directly on-site. Therefore, using the application Locus Imaginis on PC-tablet will explain better the informational continuum that this project has tried to establish between the monument and its actual digital representation.

From a conceptual point of view, in future developments, we plan to treat also a research on the 2D/3D projective relation. In our work we carried on this relation only in one direction: 2D images are described starting from a semantically enriched 3D-model. However, in view of recent results in the field of photogrammetry and computer vision concerning 2D/3D matching, automatic calibration and orientation of great sets of photos, this relation could be operated in the opposite direction: the semantic characterization (segmentation and annotation) could be led on the photos and be projected on the 3D-model. A first research has been already carried on the annotation of surface degradations so to create degradation mappings projected and queried by users on the 3D scene [30]. In future works, the semantic characterization of photos (segmentation and annotation) could be done directly on images and be projected on the 3D-model in order to structure image-based models such as point clouds, polygon meshes, etc.

### REFERENCES


Figure 9. Search by terms. Semantic annotation is projected on the 2D images and the user can display by terms the semantic decomposition on each image containing the selected term.