A Web Platform for the Consultation of Spatialized and Semantically-Enriched Iconographic Sources on Cultural Heritage Buildings

Journal:	*Journal on Computing and Cultural Heritage*
Manuscript ID:	JOCCH-12-0044
Manuscript Type:	Research Paper
Date Submitted by the Author:	14-Dec-2012
Complete List of Authors:	Stefani, Chiara; CNRS , UMR 3495 MAP Busayarat, Chawee; CNRS, UMR 3495 MAP Lombardo, Julie; CNRS, UMR 3495 MAP Véron, Philippe; CNRS, UMR 7296 LSIS De Luca, Livio; CNRS, UMR 3495 MAP
Keywords:	3D digital artefact capture, representation and manipulation, Metadata, classification schema, ontologies and semantic processing for CH multimedia repositories, ICT technologies in support of creating new cultural experiences or digital artefacts , spatio-temporal modeling; 3D-2D semantic annotation; oriented images; architectural heritage; information system

A Web Platform for the Consultation of Spatialized and Semantically-Enriched Iconographic Sources on Cultural Heritage Buildings

CHIARA STEFANI, CHAWEE BUSAYARAT, JULIE LOMBARDO, LIVIO DE LUCA, MAP Laboratory, UMR CNRS/MCC 3495, ENSA Marseille, France
PHILIPPE VÉRON, LSIS Laboratory, UMR CNRS 6168, Arts et Métiers ParisTech, Aix-en-Provence, France

This paper describes an interactive platform for the semantic annotation of oriented iconographic sources based on an accurate 3D model structured according to spatial and temporal features. This platform, called LOCUS IMAGINIS, provides access to cultural information about monuments by collecting personal snapshots taken by visitors. In particular, the platform enables to perform two actions. Firstly, pictures are added into the database: visitors manually align images to the digital model of the monument and than, by means of automatic procedures, images are enriched with 2D semantic annotations that are projected from the 3D model. Secondly, the collection of semantically-enriched data can be queried: when visitors select 2D images of the database, these are oriented in the 3D scene and superimposed onto the 3D model. In this way, this tool provides access to textual and graphic information describing the photographed site (semantic, spatial, temporal features, etc.). Moreover, visitors can interact both with images and the 3D-model to highlight building parts by color. Accessible on-site and on-line, this tool can create awareness about conservation and is adapted to all kinds of audience: students involved in educational workshops, visitors needing to extend their visit, and documentalists interested in the consultation of the iconographic corpus.

Categories and Subject Descriptors: **H.2.8 [Database Management]**: Database Applications—*Image Databases, Spatial databases and GIS*; **H.3.3 [Information Storage and Retrieval]**: Information Search and Retrieval—*Information filtering, Query formulation, Search process, Selection process;* Online Information Services—*Web-based services;* **I.3.6 [Computer Graphics]**: Methodology and Techniques — *Graphic data structures and data types, Interaction techniques*; **I.4.1 [Image Processing and Computer Vision]**: Digitization and Image Capture — *Camera calibration, Imaging geometry*.

Additional Key Words and Phrases: Spatio-temporal modeling; 3D-2D semantic annotation; Oriented images; Architectural heritage; Information system.

ACM Reference Format:

Chiara Stefani, Chawee Busayarat, Julie Lombardo, Philippe Véron and Livio De Luca. 2012. A web Platform for the Consultation of Spatialized and Semantically-Enriched Iconographic Sources on Cultural Heritage Buildings xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

1. INTRODUCTION

Nowadays digital tools and techniques allow producing 2D and 3D data on CH. This huge and often complementary amount of data contributes to describe heritage buildings.

On one side, the spread of digitization technology permits experts to collect accurate 3D data obtained by active and passive systems, such as scanner laser and photogrammetry campaigns. Such data constitutes the basis for further geometric treatments in order to obtain three-dimensional representations of the current state of buildings and to extract 2D representations describing architectural features specific to the analysis goal.

On the other side, the spread of digital camera for general public allows collecting a big amount of snapshots that can reveal useful to describe specific details or even past states of buildings.

This work is supported by the Widget Corporation Grant #312-001.
Author's address: C. Stefani, C. Busayarat, J. Lombardo, L. De Luca, UMR CNRS/MCC 3495, ENSA Marseille, 184 av. de Luminy, 13288 Cedex 09, Marseille, France; email: chiara.stefani@map.archi.fr, cha_v_mek@hotmail.com, julie.lombardo@map.archi.fr, livio.deluca@map.archi.fr; P. Véron, UMR CNRS 7296, Ecole Nationale Supérieure d'Arts et Métiers, 2, cours des Arts et Métiers, 13617 Aix-en-Provence, France.

DOI:http://dx.doi.org/10.1145/0000000.0000000

Finally, digitized historic iconographic sources (such as ancient photographs, drawings, paintings, etc.) are today available to general public and to experts of CH. Such data constitute the basis for obtaining 3D hypothetical restitutions of building previous states, and permits experts to understand and describe buildings throughout their history.

Even if 3D representations and 2D sources are often complementary, few solutions today allow experts and general public to manage this heterogeneous amount of data. 2D 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.). On the contrary, 3D data are generally difficult to consult because information is usually not structured and common graphic cards do not manage heavy 3D datasets.

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.

1.1 Challenge and aims

Despite the heterogeneity of this corpus of information, both iconography and 3D models relay to buildings. If iconographic sources essentially can represent their morphology and visual aspect (in some cases at various epochs), 3D models have the great advantage of providing a comprehensive and metrically correct representation of the morphology of buildings. For this reason, this research has various challenges.

The first one is to improve early research in conservation and valorization of cultural heritage buildings by providing public with access to information related to cultural sites. The second one is to explore collections of extended iconographic sources (photographs and more ancient sources such as engravings, plans, etc.).

The third one is to provide a joint visualization of 2D and 3D structured and semantically enriched data. 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.

Finally, 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.

It is from this need that the project Locus Imaginis 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. Its aim 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 five sections: section 2 gives an overview on previous approaches for the semantic structuring of 2D-3D data, for their display, and for the spatial referencing of images; sections 3 presents the adopted general approach and section 4 details technical aspects, functionalities and audience of the virtual platform developed in order to consult iconographic documentation; finally the last paragraph evaluates the system, assesses its limits and fixes some research perspectives.

2. PREVIOUS WORKS

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 2D and 3D data to extract information about the building morphology, 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.

Regarding research focusing on data structuring, a first distinction should be done between treatment on images and 3D models. Concerning 3D models, some studies suggest a semantic annotation process based on ontology [Attene et al. 2007], other ones extend this concept to the archaeological field [Manferdini et al. 2008] and to the urban one [Grussenmeyer, Koehl, and Nour El Din 1999]. Relations among several temporal states are usually more difficult to process. In general, various temporal versions of geometric reconstructions are created and stored in databases so to formulate temporal queries [Hetherington, Farrimond, and Presland 2006; R. E. Hetherington and Scott 2004; SanJosé-Alonso et al. 2009], and few studies describe building lifecycles to stress different states and transformations [Stefani et al. 2008; Dudek and Blaise 2008]. Concerning semantic annotation of images, usually semantic concepts are annotated with manual methods [Gong et al. 1994; Shneiderman and Kang 2000] allowing users to associate descriptive keywords to images stored in a database, or with automatic methods allowing semantic scene segmentation, based on content-based image analysis techniques [Flickner et al. 1995] or on recursive neural networks [Socher et al. 2011].

Regarding the spatial referencing of iconographic data, current research lays on manual methods based on the knowledge [Kadobayashi and Tanaka 2005] and on the registration of the position of the cameras [Waldhäusl and Ogleby 1994], on semi-automatic methods based on geometric solutions for the calibration and orientation of the cameras [Tsai 1986], 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 [Lee, Jung, and Nevatia 1992] or homologous points identified on the image [Snavely, Seitz, and Szeliski 2006]. Some browsers have been implemented in order to manage 2D/3D datasets laying on these automatic methods: Photocloud [Brivio et al. 2011] or Photo Tourism [Snavely, Seitz, and Szeliski 2007] manages photographs collections in space, while 4D Cities [Schindler, Dellaert, and Kang 2007] integrate the temporal dimension by means of a constraint satisfaction method, by ordering chronologically the spatialized photos and by automatically applying textures on 3D-models.

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.

3. MAIN APPROACH

This approach organizes and links iconographic collection to a semantically-enriched 3D-model. As defined in [De Luca et al. 2011], the 3D-model is characterized by hierarchically structured elements, classified according to a vocabulary of terms, and having assigned quality attributes and geometric information [De Luca et al. 2010]. Therefore, the iconographic spatialized sources inherit a semantic annotation from the 3D-model.

The approach enables (Figure 1):

- to reference iconography in space and in time, by knowing the geometric models of the cameras, and by associating temporal attributes to sources;
- to structure the 3D-model according to spatial and temporal criteria, by assigning attributes (to this purpose, some tools have been developed to assist the user in the structuring task);
- to annotate iconography from a semantic point of view, using a semantically structured 3D-model and the projective relations existing between 2D/3D data.

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.

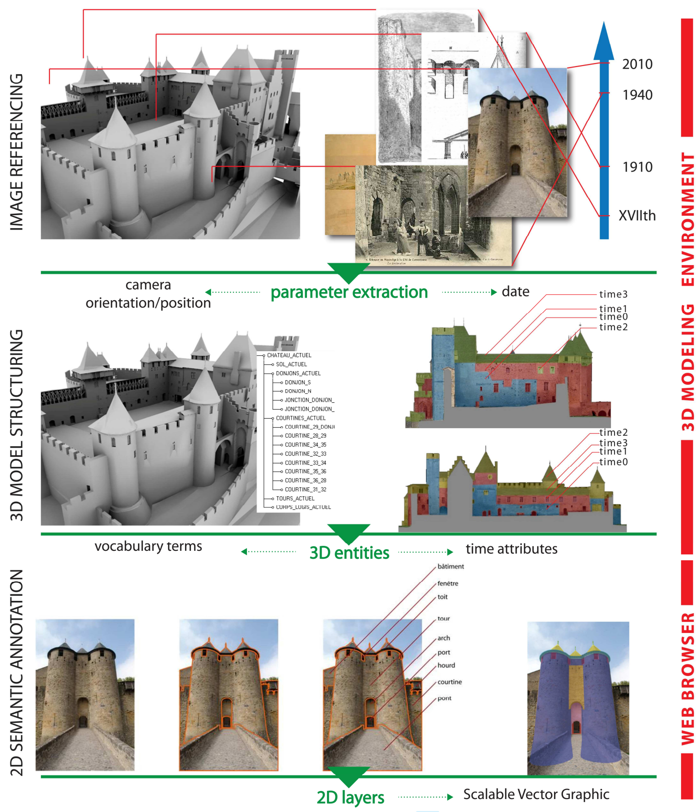

Fig. 1. System overview. Referencing of images (a) and structuring of 3D model (b) are computed in a 3D modeling environment in a specific implemented plug-in. 2D semantic annotation and query of data (c) are possible in a specific web browser.

3.1 Data Input

In this approach, detailed in [De Luca et al. 2010], the input dataset is the 3D model of the actual state. In the case of Comtal Castle, a survey campaign (3D laser scanning + photogrammetry) was conducted in 2004. In total, a dense point cloud of over 30 million points and 380 photographs were collected. This dataset permitted to obtain a restitution of the actual state, by means of the hybrid approach combining photogrammetry principles and parametric operations [De Luca 2006].

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 [De Luca, Veron, and Florenzano 2006]. 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 obtained 3D-model made up the general support for the operations of space-time iconography referencing (section 3.2) like for the semantic structuring step (section 3.3).

3.2 Spatial and Temporal Referencing of Sources

In this first stage, sources need to be referenced in space and time.

3.2.1 *Spatial Referencing.* At a spatial level, sources can be referenced by establishing a projective relation between iconography and the 3D-model of the building current state by means of a spatial resection procedure. As sources can follow different projection rules [Maynard 2005], various geometric models of cameras are handled according to the types of projection of images:

- *Perspective projection.* In the case of photographs and drawings in perspective, the camera geometric model associated to the image [Hartley and Zisserman 2004] is calculated by establishing a set of correspondences between the photograph (2D) and the current state model (3D). This method uses the intrinsic and extrinsic parameters (focal length, distortion, translation and orientation) related to the geometric model of the image. For estimating the spatial resection, the method integrates a versatile camera calibration algorithm [Tsai 1986] needing 11 2D/3D correspondences with the intent of providing intrinsic and extrinsic parameters of the camera. In order to obtain good parameter estimation, characteristic points are accurately selected so to distribute homogeneously the 2D/3D correspondences: they should be chosen on the image maximal extension and depth range of the 3D scene. A detailed description of this approach has already been presented in [De Luca et al. 2010].

- *Other projections.* Sources can follow other projection rules. In the case of plans, elevations, cross sections, to determine the scale of the image and the projection reference plane, in the workspace it is necessary to get 3 correspondences 2D/3D (2 for the scaling, 3 for the projection plane). Once the plane is defined, it can be shifted along the projection line. As for iconography in perspective projection, in order to reduce the correspondence error, it is better to choose 2D/3D coordinates on the image maximal extension. Of course, according to specificities of 3D models and images, a difficulty can consist in choosing both on the 3D model and on the image the 3 coordinates defining an extended plane. In the case of pseudo-perspectives and 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.

3.2.2 *Temporal Referencing.* At a temporal level, a procedure permits to the user to affect temporal attributes to images. Attributes concerns the date of creation of the source (that can be precisely known e.g. for snapshots, otherwise it is unknown), and the period which the represented artifact does not undergo changes in. Period is described by two dates manually inserted by the user (if known).

3.3 Process of Semantic Structuring

This process starts from the actual structured 3D model. It requires images referenced in space and characterized by temporal attributes in order to structure the existing 3D model and to reconstruct missing elements visible in images.

In order to achieve this goal, some modeling tools has been customized: they permit to automatize the task of assigning spatial and temporal attributes while geometry is created, modified or deleted [Stefani et al. 2011].

- *Creating*. 3D modeling of entities existing in the analyzed temporal state but missing in the current state or in other past ones. Geometry is created by inserting geometric primitives.

- *Modifying*. Identifying of the 3D geometric entities already created in another state (by comparison) and splitting, joining or deforming geometric entities according to the visual appearance in the analyzed source.

- *Deleting*. Hiding entities not existing in the analyzed temporal state.

Particularly, in the case of Comtal Castle, the modeling and structuring step according to time was possible by using several techniques depending on the available kind of document and on the kind of architectural element.

3.3.1 *3D elements belonging to the actual state*. In this case, the temporal structuring was obtained by comparing the 3D model with actual high resolution photographs oriented in space. The structuring was based on the projection of 2D images on the 3D morphological representation (texture extraction), and on the subdivision of the surface in various 3D elements, according to time stratifications visible on the surface of each building. Then, each

3D element has been compared with oriented sources describing past states. In this way, to each 3D element were affected temporal attributes inherited by sources or manually inserted by the user (Figure 2a).

3.3.2 *Missing elements belonging to past states*. In this case, the modeling and structuring process was obtained by comparing the 3D model with old photographs and ancient drawings oriented in space. According to the visual information contained on the image (on the edifice morphology), missing parts have been added or modified by primitive adjustment and face extrusion, laying on coordinates, straights, reference planes defined in the working environment. Finally, temporal attributes were affected, if possible, according to the ones of historic sources (Figure 2b).

At last, as structuring is a complex task, validation of the 3D model is required. This is accomplished by means of a time slider that displays the digital mock up evolution in time.

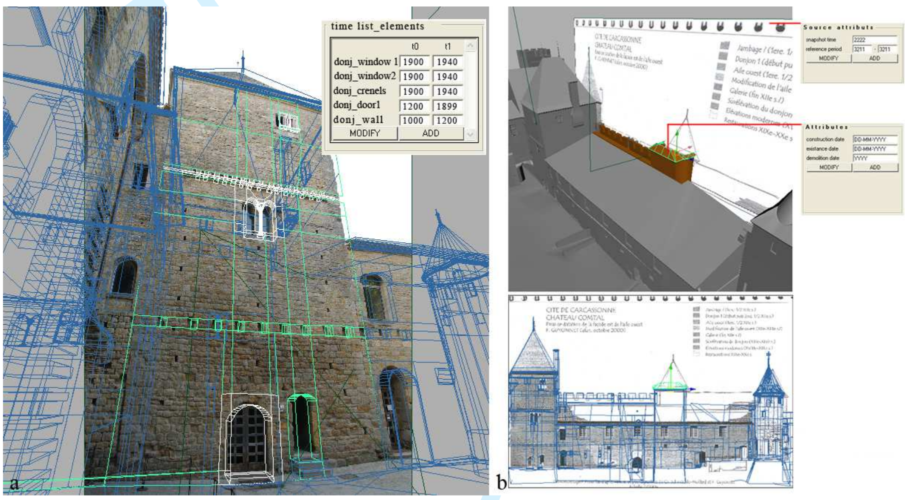

Fig. 2. Temporal stratification on the 3D-model of the Castle, setting time attributes and using high a) resolution photographs and b) other graphic sources.

3.4 Semantic Annotation of Iconography

In contrast with the works in semantic annotation of images (see paragraph 2), our approach does not lean on the direct annotation of semantic concepts on specific areas of each image. 3D-models annotated semantically at the level of the building morphology are used as medium to assign semantic information to images [Busayarat et al. 2008; Busayarat 2010]. The relation between 3D-models and 2D images is created by means of procedures for the spatial referencing of images (see Section 3.2.1). 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 (Figure 3). Each projection (computed through a vector rendering technique detailed in section 4.3) 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 (Figure 4). In the case of spatial referencing of a new image, the building morphological description will be automatically projected on images.

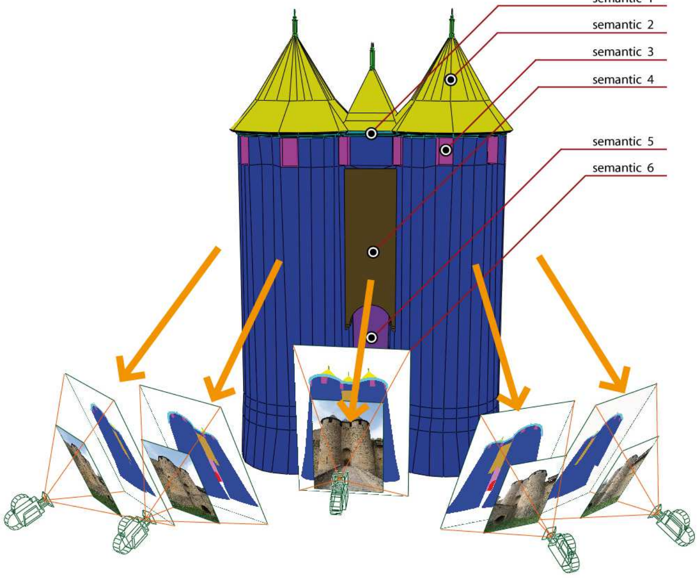

Fig. 3. Semantic annotation of photographs. The building morphology, organized according to a specific description structuring, is projected on the oriented images as semantic layers [Busayarat 2010].

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.

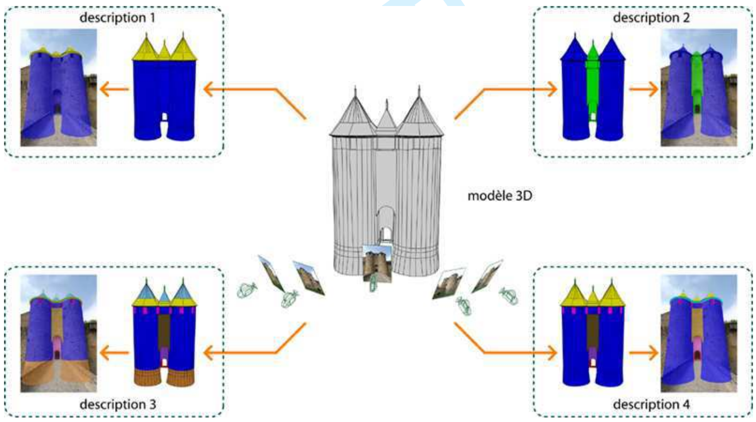

Fig. 4. Change of semantic projection on image according to the modification of the spatio-temporal structuring.

4. LOCUS IMAGINIS: AN INTERACTIVE AND COLLABORATIVE WEB BROWSER

In order to query the system, an interactive web browser has been developed. It gathers and queries semantically structured data (sections 4.1 and 4.2). 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 perspective and parallel projections.

- 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.

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/.

4.1 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 geometric 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.

4.2 Spatial Referencing of Sources

Iconographic sources are related with the 3D model by means of a 3D interface for the real-time spatial referencing. 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 defined in the paragraph 3.2. These two parameters (types of representation and projection) determine the approximation levels of spatial referencing.

In the modeling environment referencing is done by the Tsai algorithm, whilst in the web browser referencing is done manually: firstly the user configures the projection type, and then interactively manipulates the 2D plane of the iconographic source in the 3D space. The user can change the position and orientation of the camera view of the selected image in order to make the 3D view coincide with the camera view of the selected image (Figure 5).

At the moment, 192 photographs (belonging to the funds of the Center of National Monuments) and 130 photographs (acquired in situ during the survey) have been inserted and spatially aligned to the Castle. The first set of photographs has been inserted by manual referencing, while for the second one the semi-automatic process was used (section 3.2.1). Today, CNM teams (Direction of cultural development and of publics, cultural and educational sector of Comtal Castle of Carcassonne) are contributing to the extension of the database by manual insertion.

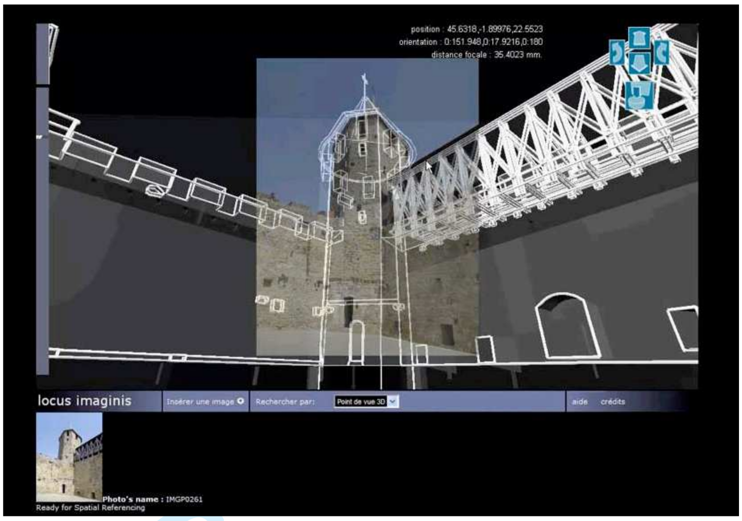

Fig. 5. Manual spatial referencing of images. By means of the keyboard and the buttons of the interface, the user can displace and rotate the navigation camera on which is projected the 2D plane of the image to be positioned in space.

4.3 Semantization of Images

Compared to other methods [Snavely, Seitz, and Szeliski 2007], the most important difference consists in the use of a 3D geometric polygonal model instead of point clouds to distribute and spread semantic annotations. This allows us to manage projected regions that result more complex, accurate and readable. In particular, the task of semantic projection is based on the communication between the MySQL database and the 3D modeling software Maya. The process embeds mechanisms to collects camera parameters (associated with the aligned images) from the database and to generate the geometric models of the cameras in Maya (which the polygonal representation of the 3D scene is loaded in).

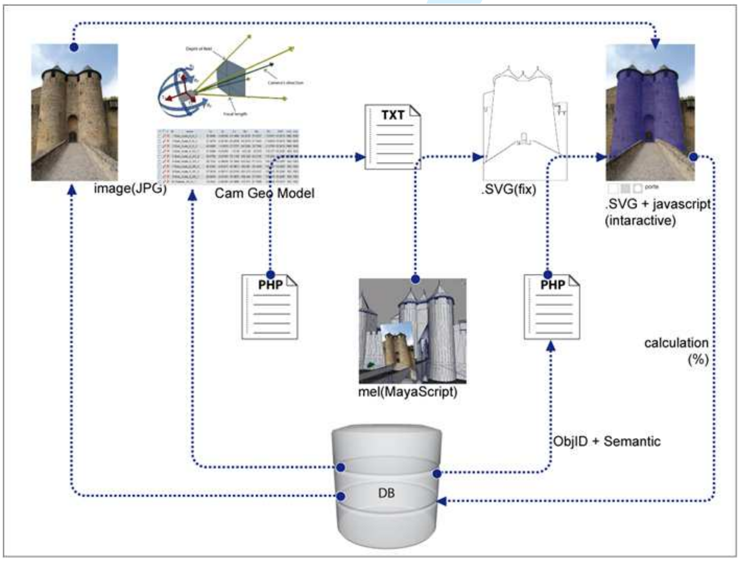

Fig. 6. Scheme of the semantic annotation of images. The process consists of four steps: the compilation of data into the database, the reconstruction of image viewpoints, the projection of semantic layers, the adding of additional attributes and interactive behaviors in the semantically annotated images.

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.

In the projected vector image, each segment of the polygon perfectly matches to the shape of 3D objects of the scene (according to viewpoints of images). Moreover, each 2D polygon combines all the semantic attributes qualifying 3D elements. This completely automatic process can be applied on images spatialized both manually (in the web interface) and semi automatically (in the modeling environment). Of course, the precision level varies according to the adopted calibration method: in the case of manual method, there can be 2D-3D misalignments due to the user ability.

4.4 Locus Imaginis Functionalities

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, it enables various actions specific to the level of display on the 3D scene:

- *color coding:* this feature enables to display the 3D-model of the monument according to its historical states and the current one by different color parameters by affecting time attributes to the parts of the 3D-model (Figure 7).
- *graduated visualization:* the 2D-3D visualization can be modified by means of the variation parameters concerning images (transparency), volumes and edges of the polygonal 3D model.

Particularly, wireframe, colors and transparency features permit, on one side, to better perceive the 3D shape of past states behind images, and on the other one, to appreciate and to easily interpret temporal stratifications visible on images.

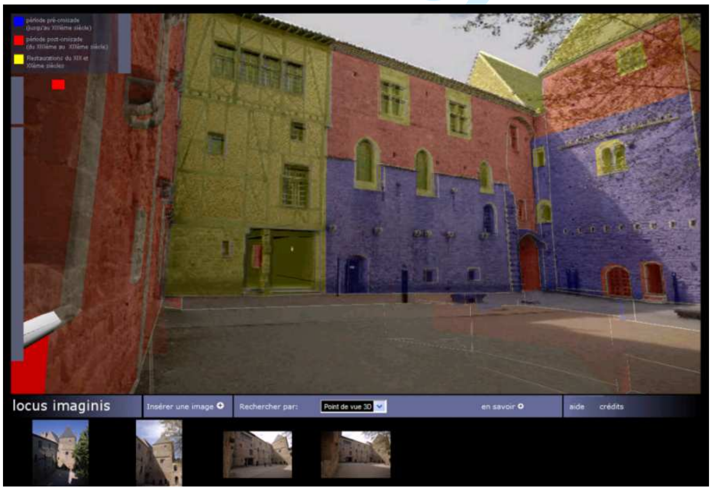

Fig. 7. Display of the 3D-model by colors according to the different temporal states.

4.5 Visual Retrieval and Navigation in Locus Imaginis

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:

Fig. 8a. Search by 3D viewpoint

Fig. 8b. Search by source

4.5.1 *Search by 3D viewpoint.* This type of search permits 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. 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 a 2D-3D joint interaction by means of a double dialog between a 2D thumbnail bar and a 3D scene.

- *3D to 2D synchronization:* when the user navigates in the 3D scene, images of the 2D bar are displayed and adapted to the ones whose viewpoint is included in the visual piramid (frustum) of the navigation camera of the 3D scene. Images displayed in the middle of the 2D bar match with the ones whose cameras are the nearest to the camera navigation frustum (Figure 8a).
- *2D to 3D synchronization:* when a thumbnail is selected, the navigation camera of the visual browser moves in real time in the scene (positioning in the viewpoint of iconographic sources on the 3D scene), and displays 2D iconography overlaid to the 3D representation of the building.

In order to manage image detection, 3 parameters can be modified: the field of view of the navigation camera (FoV), the distance between the navigation camera and the image views (D), and the orientation difference between the navigation camera and the image views (O). The default settings of the system were defined as follows:

FoV = 90.6 degrees,
D = 15 m,
O = 90 degrees.

4.5.2 *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. According to the kind of source (photograph, engraving, drawing), at the selection of an image, the 3D-model is displayed in conic or cylindrical projection (Figure 8b).

4.5.3 *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. 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 (Figure

9b). Specifically, at the selection of an annotated image, the semantic annotation can be highlighted on the 3D scene at the mouse rolling over (Figure 9a). Then, at the selection of a specific term, the 3D-model displays by coding colors 3D elements belonging to the selected category (Figure 9b). From a technical point of view, vector layers (SVG), stored in the database, are associated to images, so creating temporary image files. These files are edited on the fly, and vector regions of images corresponding to the selected entities are colored. This feature permits users to know the position of elements in the retrieved images. Results of the query are ranked according to the area extension of the region identified on the image.

Fig. 9a. Specific term of the 3D-model displayed by coding colors

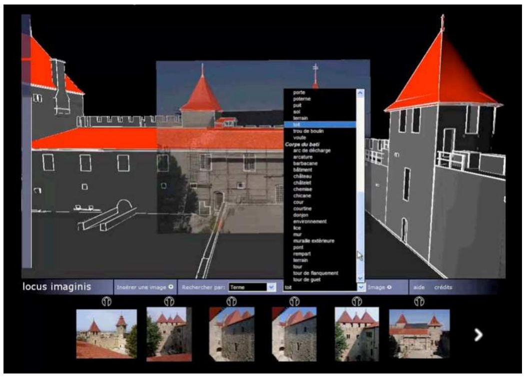

Fig. 9b. Display of 2D and 3D elements belonging to the category

4.6 Target Users

This 3D database addresses to different user profiles, defined with the CNM, for different purposes:

- Complement to the tour. Visitors can use the database on-line to prepare the visit of the monument and on-site, as a complementary cultural offer for the visitors to the Castle.
- 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. Particularly, the project Locus Imaginis has been integrated into the partnership agreement between the CNM and the Jules Fil high school of Carcassonne and will be tested by students.
- Collection and storage of iconographic sources. Documentalists and researchers (historians, architects, archaeologists, etc.) can use the database to deepen their knowledge on the building morphology, by accessing on-line iconographic sources starting from the exploration of 3D representations related to different historical periods.

5. CONCLUSIONS

This work described a system for the storage and the spatial referencing of iconographic sources on the 3D space and for 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.). Furthermore, for the adopted process of semantic structuring, this project is sustainable: firstly, 2D-3D data can be easily updated, and secondly, semantic annotations can be easily upgraded.

Overall, some issues need to be resolved. Specifically, in the case of close-up and background images, the comprehension of the context is difficult, since 3D-model appears excessively near or far. For close-up images, it would be better to obtain a simultaneous scaling of the focal length and of the 3D-model; on the contrary, in the case of background images, the focal length should be locked on the 3D-model, so as to modify the distance of the observation point.

Finally, some reflections should prompt further research. In our work, we explored the 2D to 3D projective relation, as 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, this relation could be operated in the 3D to 2D direction: the semantic characterization (segmentation and annotation) could be led on the photos and be projected in the 3D scene. 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 [Stefani et al. 2012]. In future works, the semantic characterization of photos (segmentation and annotation) could be done directly on oriented images and be projected on the 3D-model in order to structure image-based models such as point clouds, polygon meshes, etc.

REFERENCES

Attene, M., F. Robbiano, M. Spagnuolo, and B. Falcidieno. 2007. "Semantic Annotation of 3D Surface Meshes based on Feature Characterization." In *Lecture Notes in Computer Science*, 4816:126–139.

Brivio, P., L. Benedetti, M. Tarini, F. Ponchio, P. Cignoni, and R. Scopigno. 2011. "PhotoCloud: Real-time Web-based Interactive Exploration of Large Mixed 2D-3D Datasets." *IEEE Computer Graphics and Applications*, IEEE Press edition.

Busayarat, C. 2010. "La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales : définition d'une approche pour la sémantisation de sources iconographiques par référencement spatial." Aix-en-Provence, France: Arts et Métiers ParisTech.

Busayarat, C., L. De Luca, P. Véron, and M. Florenzano. 2008. "An On-line System to Upload and Retrieve Architectural Documents Based on Spatial Referencing." In *IDMME - Virtual Concept 2008 Conference*. Beijing, China.

Dudek, I., and J.-Y. Blaise. 2008. "Visual assessment of heritage architecture life cycles." In *Journal Of Universal Computer Science*. Graz, Autriche.

Flickner, M., H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, et al. 1995. "Query by Image and Video Content: The QBIC System." *Computer* 28 (9): 23–32. doi:10.1109/2.410146.

Gong, Y., H. Zhang, H. C. Chuan, and M. Sakauchi. 1994. "An image database system with content capturing and fast image indexing abilities." In *Proceedings of the International Conference on*, 121–130. IEEE Comput. Soc. Press. doi:10.1109/MMCS.1994.292444.

Grussenmeyer, P., M. Koehl, and M. Nour El Din. 1999. "3D Geometric and Semantic Modelling in Historic Sites." In *XVII CIPA International Symposium*. Olinda, Brazil.

Hartley, R.I., and A. Zisserman. 2004. *Multiple View Geometry in Computer Vision*. Second. Cambridge University Press.

Hetherington, R. E., and J. P. Scott. 2004. "Adding a Fourth Dimension to Three Dimensional Virtual Spaces." In *Proceedings of the Ninth International Conference on 3D Web Technology*, 163–172. Monterey, California: ACM.

Hetherington, R., B. Farrimond, and S. Presland. 2006. "Information Rich Temporal Virtual Models Using X3D." *Computers & Graphics* 30 (2): 287–298.

Kadobayashi, R., and K. Tanaka. 2005. "3D viewpoint-based photo search and information browsing." In *SIGIR*, 621–622. Salvador, Brasil.

Lee, S. C., S. K. Jung, and R. Nevatia. 1992. "Automatic Pose Estimation of Complex 3D Building Models." *IEEE Workshop on Application of Computer Vision*.

De Luca, L. 2006. "Relevé et multi-représentation du patrimoine architectural : définition d'une approche de reconstruction 3D d'édifices". Aix-en-Provence: Ecole Nationale Supérieure d'Arts et Métiers (ENSAM).

De Luca, L., C. Busayarat, C. Stefani, N. Renaudin, P. Véron, and M. Florenzano. 2010. "An Iconography-Based modeling Approach for the Spatio-Temporal Analysis of Architectural Heritage." In *2010 Shape Modeling International Conference*, 78–89. Aix-en-Provence, France. doi:http://doi.ieeecomputersociety.org/10.1109/SMI.2010.28.

De Luca, L., C. Busayarat, C. Stefani, P. Véron, and M. Florenzano. 2011. "A semantic-based platform for the digital analysis of architectural heritage." *Computers & Graphics, Pergamon Press, Inc. Elmsford, NY, USA* 35 (2). Elsevier: 227–241. doi:10.1016/j.cag.2010.11.009.

De Luca, L., P. Veron, and M. Florenzano. 2006. "Reverse Engineering of Architectural Buildings Based on a Hybrid Modeling Approach." *Computers & Graphics* 30 (2): 160–176. doi:10.1016/j.cag.2006.01.020.

Manferdini, A.M., F. Remondino, S. Baldissini, F. Gaiani, and B. Benedetti. 2008. "3D Modeling and Semantic Classification of Archaeological Finds for Management and Visualization in 3D Archaeological Databases." In Limassol, Cyprus.

Maynard, P. 2005. *Drawing distinctions: the varieties of graphic expression*. Cornell University Press.

SanJosé-Alonso, J.I., J. Finat, J.D. Pérez-Moneo, and J.J. Fernández-Martín. 2009. "Information and knowledge systems for integrated models in Cultural Heritage." In *Proceedings of the 3rd ISPRS International Workshop 3D-ARCH 2009*. Vol. XXXVIII-5/W1. Trento, Italy: F. Remondino, S. El-Hakim, L. Gonzo.

Schindler, G., F. Dellaert, and S.B. Kang. 2007. "Inferring Temporal Order of Images From 3D Structure." In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*.

Shneiderman, B., and H. Kang. 2000. "Direct annotation: a drag-and-drop strategy for labeling photos." In *Information Visualization, 2000. Proceedings. IEEE International Conference on*, 88 – 95. London , UK. doi:10.1109/IV.2000.859742.

Snavely, N., S.M. Seitz, and R. Szeliski. 2006. "Photo Tourism: Exploring Image Collections in 3D." In *SIGGRAPH*. Boston, Massachusetts USA.

Snavely, N., S.M. Seitz, and R. Szeliski. 2007. "Modeling the World from Internet Photo Collections." *International Journal of Computer Vision*. doi:10.1007/s11263-007-0107-3.

Socher, Richard, Cliff Chiung-yu Lin, Andrew Y. Ng, and Christopher D. Manning. 2011. "Parsing Natural Scenes and Natural Language with Recursive Neural Networks." In *ICML 2011*.

Stefani, C., X. Brunetaud, S. Janvier-Badosa, K. Beck, L. De Luca, and M. Al-Mukhtar. 2012. "Developing a Toolkit for Mapping and Displaying Stone Alteration on a Web-based Documentation Platform." *Journal of Cultural Heritage*.

Stefani, C., L. De Luca, P. Veron, and M. Florenzano. 2011. "A Tool for the 3D Spatio-Temporal Structuring of Historic Building Reconstructions." In *Digital Media and its applications in Cultural Heritage*, 153–168. Amman, Jordan: J. Al-Qawasmi, Y.Alshawabkeh, F. Remondino.

Stefani, C., L. De Luca, P. Véron, and M. Florenzano. 2008. "Reasoning about space-time changes: an approach for modeling the temporal dimension in architectural heritage." In *Proceedings of the IADIS International Conference*. Amsterdam, Netherlands.

Tsai, R.Y. 1986. "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision." In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 364–374. Miami Beach, FL.

Waldhäusl, P., and C.L. Ogleby. 1994. "3 x 3 Rules for simple photogrammetric documentation of architecture." *J.G.Fryer* XXX. International Archives of Photogrammetry and Remote Sensing: 426 – 429.