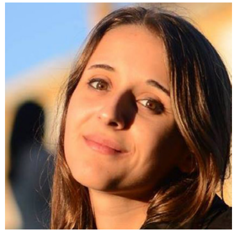
Combining on-site and off-site analysis: towards a new paradigm for cultural heritage surveys

Violette Abergel, Kévin Jacquot, Livio De Luca, Philippe Veron

ISSN 1828-5961



Violette Abergel
Architect, PhD in engineering sciences, Violette Abergel is a post-doctoral researcher at UMR CNRS/MC MAP. Her research activity focuses on digital tools dedicated to cultural heritage survey, analysis and representation. She is member of the digital data WG of the CNRS/Ministry of Culture scientific site for the restoration of Notre-Dame de Paris.

Kevin Jacquot
Architect, PhD in architectural sciences, Kévin Jacquot is researcher at CNRS/MC MAP laboratory, and Associate Professor in Sciences and Techniques for Architecture at the National School of Architecture of Lyon, France. His research activities focus on surveying, geometric modeling and semantic enrichment of digital heritage objects.

Livio De Luca
Architect, Ph.D in Engineering, HDR in Computer Science, he is research director at CNRS and director of CNRS-MAP unit. General Co-chair of DigitalHeritage 2013 and 2015, coordinator and member of national and international actions, his research focus on surveying, geometric modeling and semantic enrichment of digital heritage objects. He is coordinator of the “digital data” working group for the restoration of Notre-Dame de Paris.

Philippe Véron
Philippe VERON is full Professor at Arts et Métiers engineering school in Aix-en-Provence, France and researcher at the Physical and Digital Systems Engineering Laboratory (LISPEN, French Research Ministry unit n°7515), in the field of design science and digital engineering. He is also director of the Carnot ARTS institute and President of the 39 French Carnot institutes Network.

Combining on-site and off-site analysis: towards a new paradigm for cultural heritage surveys

In recent decades, cultural heritage survey practices have significantly evolved due to the increasing use of digitization tools providing quick and easy access to faithful copies of study objects. While these digital data have clear advantages, especially in terms of geometric characterization, they also introduce a paradigm shift by outsourcing ex situ most of the analysis activities. This break between real and virtual working environments now raises new issues, both in terms of data dispersion and knowledge correlation in multidisciplinary teams. Benefiting from the fields of information systems and augmented reality, we proposed an integrated approach allowing the fusion of geometric, visual and semantic features in a single platform. Today, this proof of concept leads to new perspectives for the production of semantically enriched digital data. In this paper, we intend to explore the different possibilities in terms of implementation and their benefits for cultural heritage survey.

Keywords:
Cultural heritage survey; Mixed reality; Information systems; 3D digitization; Semantic description

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

1. INTRODUCTION

The cultural heritage (CH) survey is a matter of observation, analysis and synthesis which consists as much in extracting information from the study object as in conveying it. For an observer in front of his study object, the survey is a special moment of knowledge which serves two main purposes: to memorize the object in its formal components (dimensions, proportions, materials, aspect, ...), but also to select information (techniques, relations between elements, concepts...) and organize them to give the interpretation keys of the object. The progress of the last decades in the fields of computer science and metrology has fostered the development of digital measuring tools. While digital technology has not radically changed measurement methods themselves, improvements in accuracy, automation, and storage capacity have been decisive for many disciplines. In the CH field, the digital technologies allowed a massive collection of 2D and 3D data in response to various needs for study, monitoring, documentation, archiving, or dissemination. It provides valuable new supports for the understanding of study objects, particularly with regard to their morphological characterization.
However, the ability to quickly digitize CH objects through the faithful reproduction of their geometric and visual attributes introduces a major paradigm shift. If in a "classical" survey the observer's analytical engagement is mobilized *in situ* in front of the study object, in a digital survey the identification of significant elements, their interpretation, and their hierarchization take place *ex situ* with a 3D copy. Then, this process depends mainly on the ability to sort, discard and relate virtual data. This break between real and virtual environments introduces a gap that results in a data dispersal effect and a singular difficulty in ensuring knowledge correlation within multidisciplinary teams. From a quantitative and formal point of view, the computational capacity provided by IT tools regarding 2D or 3D digital resources far exceeds that of "conventional" survey processing tools. Conversely, these same tools are far below human cognitive capacities, particularly in knowledge enrichment scenarios. This dichotomy is reflected by recurring contrasts between CH survey methods,

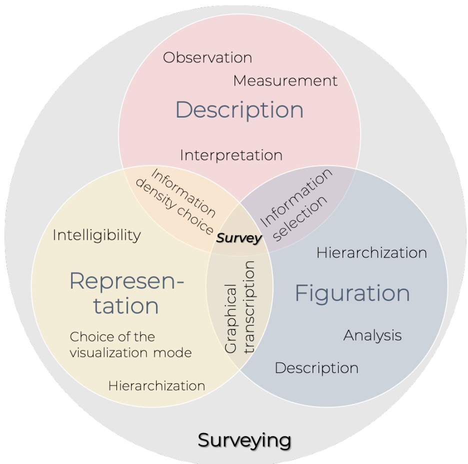


Fig. 1 - The survey interrelated moments: description, figuration, representation.

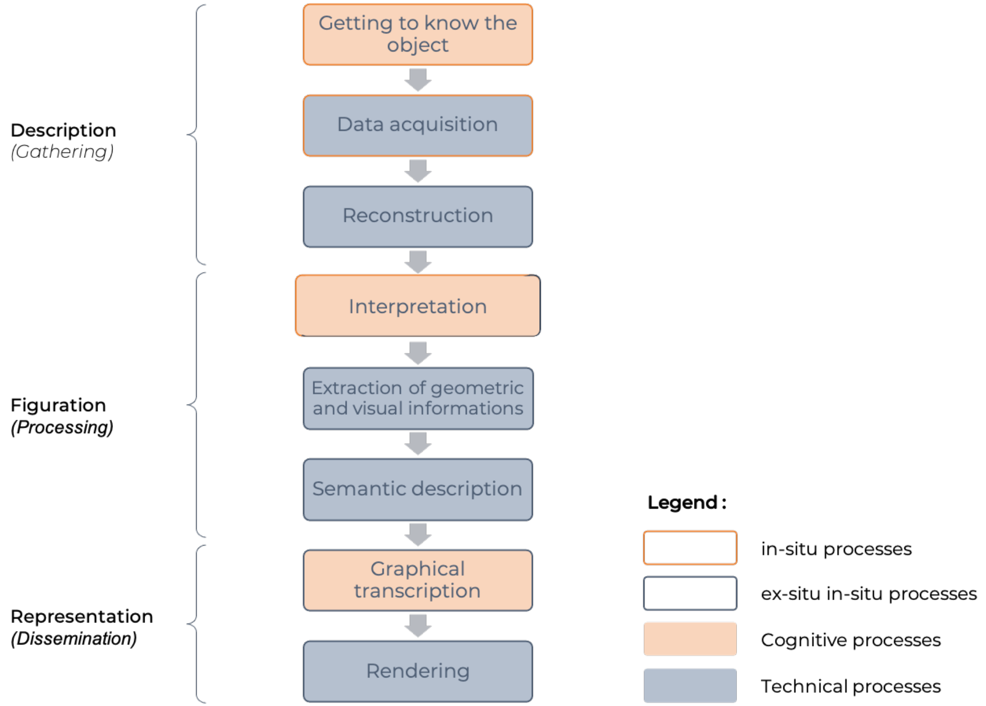


Fig. 2 - The different stages of the digital survey. Contrary to the "classic" survey, this process is sequential.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

which often involve evaluating which one prevails and in which contexts. Our stance is opposite: by considering them as complementary, we seek to bring them together in a single approach in order to benefit from their respective strengths. We sought to define an integrated approach allowing the fusion of geometric, visual and semantic survey aspects in a single multimodal environment. To this end, we stood at the crossroads of two domains: information systems for their ability to federate data and actors, and augmented reality for its ability to contextualize digital data on reality and, in our case, to ensure the continuity of *in situ* and *ex situ* works. From a prototype validating the feasibility of the approach, we focus now on the perspectives that these new interaction modalities are likely to open up for CH survey and semantic annotation.

2. CONTEXT: THE CH SURVEY

In 1992, Saint-Aubin defined the survey as the process that "*leads to drawing up, through the use of representation, the dated statement of the actual shape of the building, with the object as a constant reference: its imperfections give an account of the inadequacy of techniques, tools and people; its reshuffles tell its history and uses; its failures as well as its deformations announce its condition*"(Saint-Aubin, 1992). Clearly, the survey is not limited to the memorization of the geometrical features of the object, but involves a selection of significant elements in view of a given area of concern. In other words, the survey can be considered as a cognitive act based on the analysis of the relations between form, geometry, aspect and architectural or archeological meaning. Then, the process involves three interrelated moments: description, figuration, representation.

Description corresponds to the most exhaustive possible capture of the significant data of the object. It is an objective measurement operation oriented towards the satisfaction of a specific request guiding the observation.

Figuration aims to condense the study object into an array of simplified representations and to hierarchize information. This graphical translation operation introduces the need to establish a rendering symbolism aimed at compensating for the loss of information. It is therefore coupled with the qualification of figurative elements: it is a semantic stage, more subjective than the previous one.

Fig. 3 - A methodological challenge: the union of the survey phases in a single moment facing the real object. The survey must be considered as a cycle whose different phases are centered on the study object, possibly through its 3D copy.

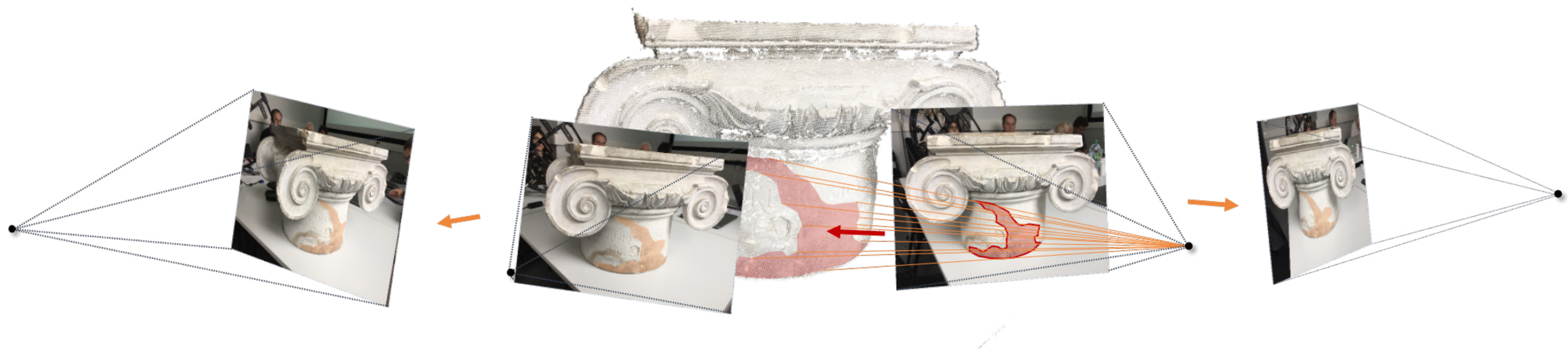


Fig. 4 - Aïoli's approach for 2D/3D propagation of 2D regions, based on the spatialization of project resources.

 DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

Representation aims to produce an intelligible and communicable graphic document. It involves making decisions deeply linked to the previous stages (choice of the scale, information density, visualization modalities, ...). This moment crystallizes a dilemma inherited from the Renaissance, as it implies a difficult trade-off between preserving the vision of the object in its perceptive properties (mainly the restitution of what the eyes see), or in its dimensional and formal properties.

These three stages are deeply interrelated since the choice of the features to be measured is influenced by the understanding of the observed elements themselves and their hierarchization with respect to the initial demand, and consequently, by the anticipation of figuration and representation modalities (Fig. 1). Thus, the survey appears as a process located at the intersection of a cognitive engagement and an objective memorization process.
For the CH survey, digitization presents a major interest: that of being able to express through 3D copy the geometrical complexity of the study objects, reproducing their shapes and appearance. However, the development of 3D acquisition technologies has clearly prioritized the quest for exhaustivity, with the aim of producing digital copies as close as possible to real objects. These exhaustivity and precision expectations tend to sideline the cognitive dimension of surveying practice (Lo Buglio, 2016). In the case of a digital survey, even if the analytical involvement of the observer is still engaged during the choice of acquisition methods and their deployment when he is the operator, most of the selection and interpretation work are done *ex situ*. Thus, where the "classical" survey can be defined as "*a moment when, thanks to the observation of the shape, we restore the features composing the building by an approach that uses the act of 'drawing' as a means to 'understand' the architecture*" (De Luca, 2009), a digital survey appears more like a sequence of distinct successive activities of acquisition, linking and processing of virtual data (Fig. 2).

In this context, surveying today involves the manipulation of heterogeneous data of various natures, origins and levels of detail. This multiplication of study media opens up many opportunities for the research community in terms of information extraction and analysis. However, this richness also presents a paradox, as data often remains under-exploited due to the lack of a versatile tool for manipulation, linking, semantic enrichment and exploitation (Thivet et al., 2019). Beyond interoperability issues, it causes major difficulties in linking up the different viewpoints of the involved actors. The existing approaches to converge multidisciplinary observations within digital representations of a study object (see section 3) are still experimental, but also off-site insofar as they always present semantic enrichment as a final post-processing step. We can therefore notice a clear **break in survey continuity**, between *in situ* data acquisition operations and subsequent *ex situ* processing and interpretation operations. For this reason, it seems essential today to examine how far quantitative (from acquisition procedures and computer processing) and qualitative (from the interpretation of the data acquired by the survey actors) information production opera-

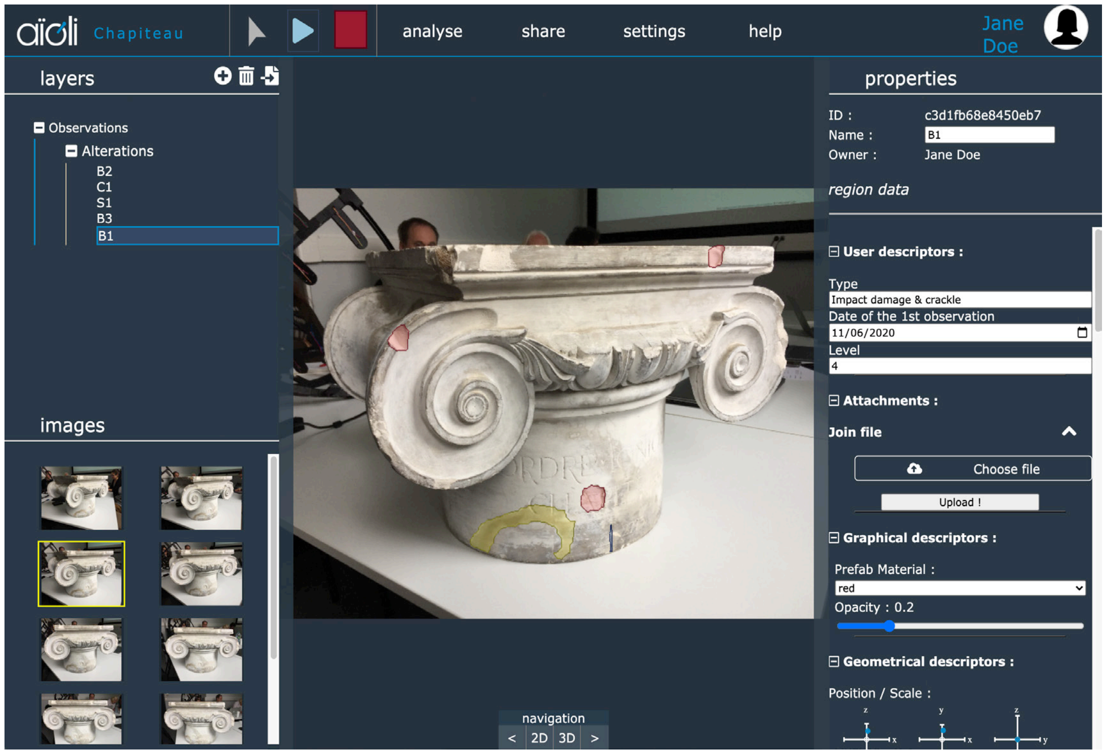


Fig. 5 - Aïoli's interface. Annotation of alteration phenomena observed on a capital.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

tions can converge within a single digital environment allowing both collection and analysis. We sought to define a single integrated approach to federate both survey data and actors. **We consider the survey as a cycle**, a dynamic and iterative process including on-site visits, data acquisition, information extraction, semantic enrichment, and representation steps (Fig. 3). Overall, our ambition raised two issues:

- The formalization of a digital environment including efficient annotation tools to enable researchers to understand, assimilate, and enrich digital data, i.e. to converge visual, geometric, and semantic aspects within a single work environment.
- The reconciliation of *in situ* and *ex situ* activities through hybridization, to overcome the existing gap between the real and the virtual worlds. Augmented reality (AR) provides a major opportunity, as its purpose is precisely "*to allow a user to carry out sensory-motor and cognitive activities in a new space by associating the real environment and a virtual environment*" (Fuchs et al., 2010)

3. RELATED WORKS

In the CH field, much research focuses on the creation and enrichment of 3D models, or on improving the management of interconnected data. However, reality-based 3D reconstruction and information systems are generally treated as separate topics, and few works consider them as complementary parts of a single framework. Thus, we can distinguish, on the one hand, knowledge-based information systems, such as (Calvanese et al., 2016; Meghini et al., 2017; Niang et al., 2017), designed for the management of documentation activities, but without direct exploitation of the geometric and visual properties of 3D data. On the other hand, approaches with 3D annotation tools such as (Havemann et al., 2009; Serna et al., 2012; Yu et al., 2013) are mainly part of digital collection management frameworks, where semantic enrichment is very efficient for classification and navigation

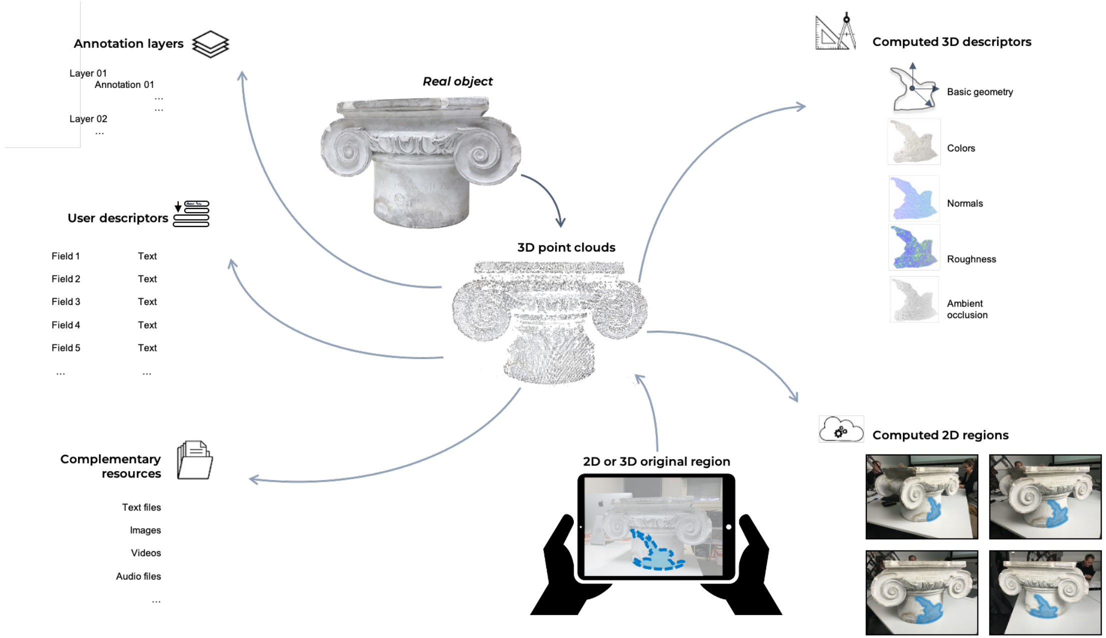


Fig. 6 - Proposal of a multimodal survey approach for architectural heritage, based on the semantic annotation of heterogeneous data and the exploitation of geometric and visual attributes of the 3D copy.

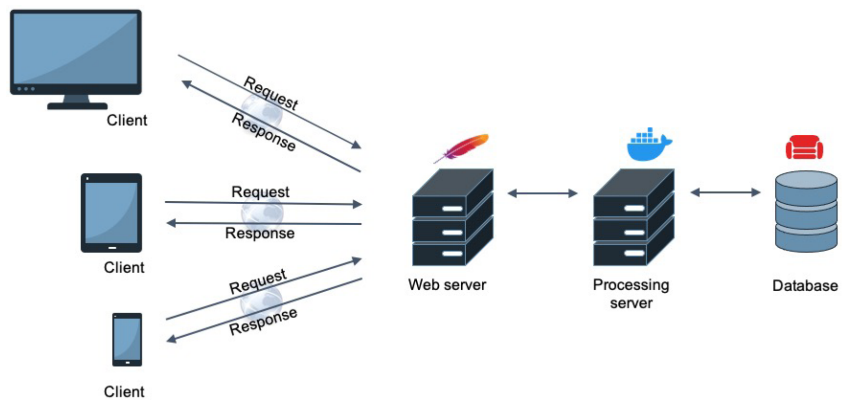


Fig. 7 - Software architecture of the application.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

operations. Information systems adapted to analysis contexts such as (Yu & Hunter, 2013; Shi et al., 2016; Soler et al., 2017; Apollonio et al., 2018) very often require the use of a frozen mesh of the study object and sometimes even its prior segmentation. Except *Cher-Ob*, they focus on a single or isolated representation of the object, missing the great analytical potential of other 2D or 3D resources. Concerning a real-virtual connection, AR applications mainly concern mediation and dissemination (Bekele et al., 2018). The majority target tourism or museum contexts. During this study, we didn't find any AR application dedicated to survey allowing to make on-site semantic annotations.

Put the object in the heart of the survey: the Aïoli platform

(De Luca et al., 2018) propose with Aïoli an integrated approach to merge geometric, visual and semantic aspects to provide an information continuum. It is a collaborative web platform allowing to create projects from simple images. These images are used to produce a 3D point cloud of the study object by photogrammetry. From there, users can annotate any 2D data of the project. Thanks to the projective relations established during the photogrammetric reconstruction, the annotations are automatically propagated to all the 2D and 3D resources on which the region is visible (Manuel et al., 2014) (Fig. 4). These can be freely described by users using descriptors, either as tags, attributes, relations, or by using controlled vocabularies thanks to a link with the OpenTheso platform dedicated to thesaurus management (Fig. 5). This approach seems really interesting as it allows the production of masses of semantic contents rather than masses of uninterpreted raw data. Nevertheless, with regard to the above-mentioned issues, it still has significant limitations. For the survey, the combined use of various 2D and 3D resources is very often essential to allow the observation and understanding of some phenomena or stigmas. However, each annotation can only be made from a single 2D image, which is insufficient in many cases, e.g. if the object is morphologically complex. Moreover, the different representations within the platform remain completely decontextualized: the projective relationship between 2D and 3D resources is well established, 3D data are well reality-based, but the link between the spatialized data and the real object still misses.

4. MAIN APPROACH

Using Aïoli as a working base, our approach is to extend to the real object the projective relation established between the various 2D and 3D resources. We consider that a 3D point cloud in a project is not only a geometrical and visual substitute of the study object, but also a structuring dataset that can be used both to federate the numerous resources involved in a survey, and as an interface between the real object and the digital data itself (Fig. 6). It leads to three distinct requirements:

- An application dedicated to CH survey must not only centralize all the data, information and knowledge related to the study object, but also connect them spatially within a single real-time visualization space allowing their manipulation.
- Virtual data must be accessible in AR to allow users to interact with a hybrid environment comprising both the real object and the related digital information.

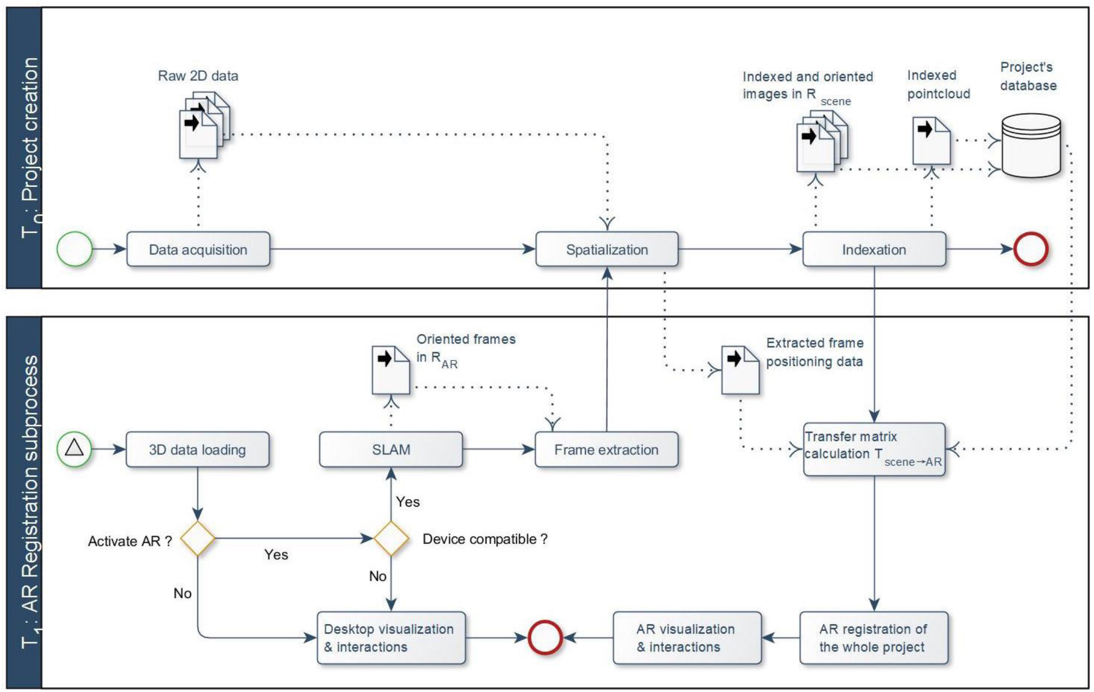


Fig. 8 - AR registration process based on SLAM and on the spatialization of a video frame by spatial resection.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

- The cross exploitation of the geometric, visual, and semantic attributes available must ease analysis and monitoring processes and enrich the understanding of the object. It implies the definition of rendering and semantic annotation modalities (manual, assisted, or automatic) to allow the description of the observed facts.

In this way, the understanding of the object of study would be achieved through the spatial relation of the different resources produced by the actors involved, whatever their nature, and the real-virtual hybridization of contents would essentially rely on our ability to ensure the registration of a single coherent 3D entity. Finally, once the continuity between real and virtual environments is ensured, many possibilities would appear in terms of analysis and interaction.

5. MOBILE AUGMENTED REALITY VISUALIZATION OF SURVEY DATA

The first step of our work consisted in designing an environment allowing to spatialize very heterogeneous resources (visible light photographs, multispectral images, RTI images, normals, curvature or roughness maps, 3D point clouds, 3D annotation paths, ...) in order to make them accessible in a single coherent 3D scene. Its software architecture, derived from Aïoli's, includes four main elements (Fig. 7):

- A database which contains all the application data (users accounts, projets, annotations, etc.);
- A processing server which relies on virtualization technology to provide separate containers specifically dedicated to each calculation process. It hosts in particular the TACO engine, a service allowing to compute the relative poses of 2D images within a photogrammetric project (Pamart, Morlet, et al., 2019);
- A web server, which manage projects data and communicates with the processing server;
- A thin client, running from a browser, which communicates with the web server to access to the services of the application. It integrates a WebGL 3D viewer based on Three.JS and Potree.JS libraries, allowing to model virtual cameras from calibration parameters obtained thanks to TACO, to spatialize them around the global point cloud of the study object, and to ensure the real-time rendering of the whole scene.

These developments are described in more detail in (Pamart, Ponchio, et al., 2019), and include several calculation strategies that can be mobilized according to the nature of the images to be spatialized and their proximity to the project data. Basically, when a user uploads images of any kind, they are stored on the web server and sent to the processing server. TACO then uses various strategies to spatialize them by photogrammetry and compute a dense point cloud if necessary. The images spatialization data - and the new point cloud if any - then transit again through the web server, to be finally displayed in the 3D WebGL viewer on the client side. Some experiments allowed us to confirm our capacity to handle many kinds of images including lasergrammetric data (stations panoramic images), although the fully automatic fusion of photogrammetric and lasergrammetric data still raises some registration issues. In any case, the

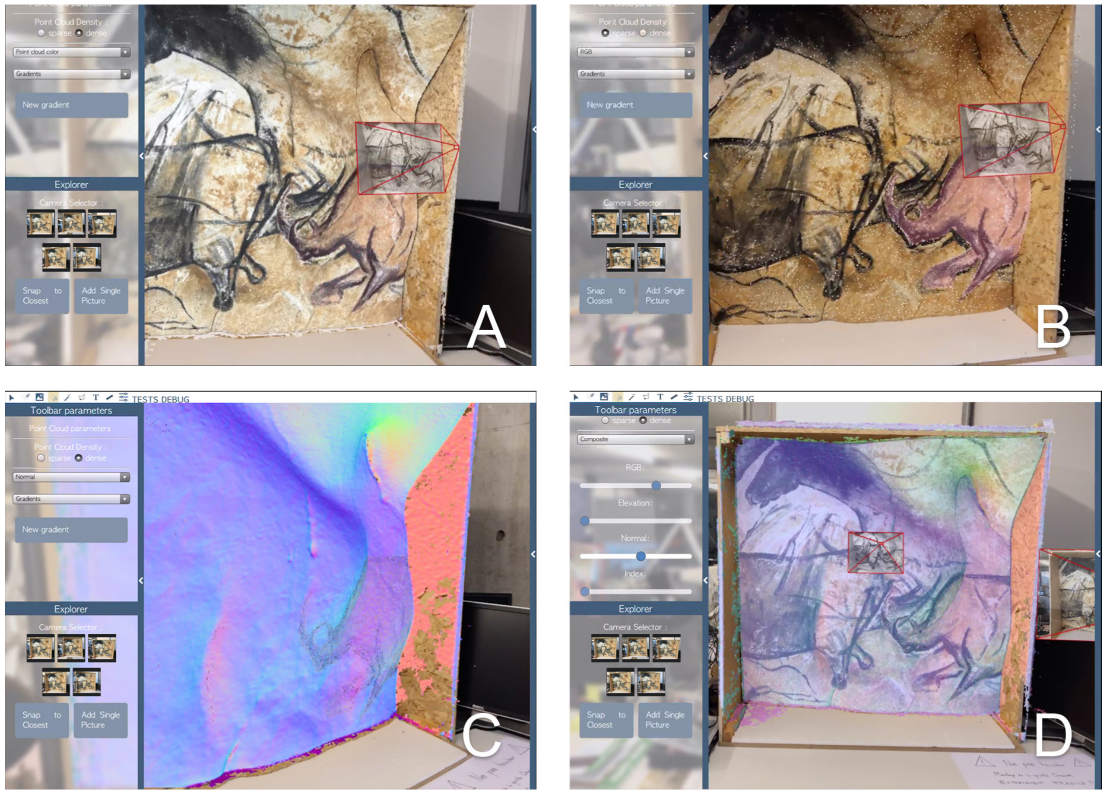


Fig. 9 - Results of our registration approach on the case study of a facsimile of the Chauvet-Pont D'arc cave.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

diversity of the representations gathered in this environment tends on one hand to limit the data dispersion effect, and on the other hand to ease the understanding of the object by taking into account both its geometric and visual complexity. This approach is made possible by the control of the different stages of the spatialization and annotation processes and the storage of computational data, which underlines the interest of an integrated approach.
So far, we were able to focus on the definition of a method to bridge real and virtual environments through AR interaction modalities. We proposed a web compatible markerless registration method, to ensure the coherence of the alignment of the real (study object) and virtual (3D scene of the corresponding Aïoli project) environments. The algorithm, described in (Abergel et al., 2019), is based on the joint use of SLAM tracking, whose poses are recovered using the WebXR API, and a spatial resection calculation. The main idea is that SLAM tracking provides us with a real-time estimation of the mobile camera pose, in an arbitrary local reference system. On the other hand, all Aïoli project data are also represented in another arbitrary local reference system. The spatialization of a single image of the video stream using TACO then allows us to know its pose in both reference systems, and thus to compute the transformation matrix needed for the AR registration (Fig. 8). A prototype allowed us to validate this approach, which offers satisfying results (Fig. 9) despite some technical limitations that remain to be addressed (calculation time, accuracy in low light environments, ...)[1].

6. ENHANCEMENT OF GEOMETRICAL AND VISUAL PROPERTIES AND NEW ANNOTATION TOOLS

Although the prototype reached by the IT implementation of the stated approach is highly perfectible, what really interests us are the perspectives that this demonstration opens up in the field of CH studies. Considering that it is now possible to register projects in AR and therefore that informa-

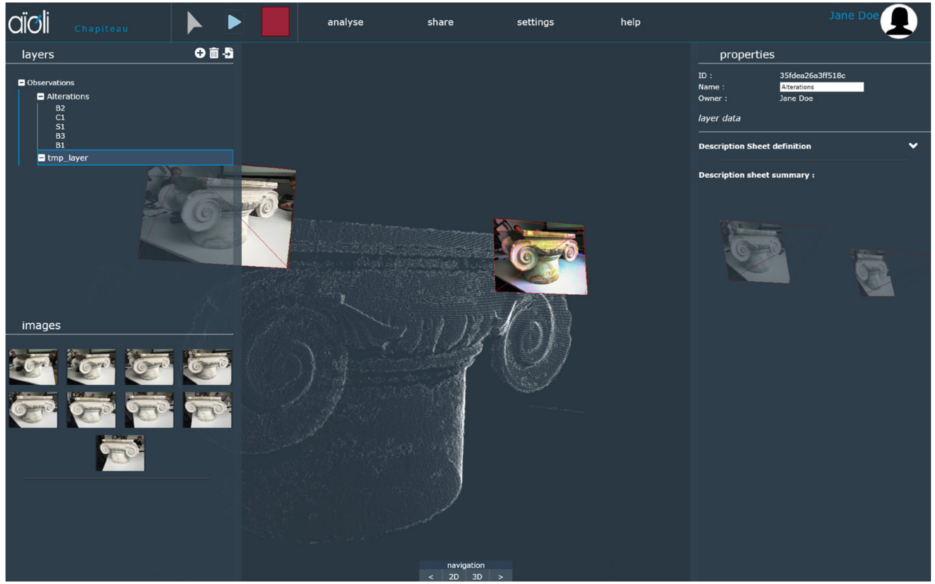


Fig. 10 - The simplest and most obvious approach is to directly process images likely to support annotations. Since the processing does not interfere with their pose, the propagation process is unchanged.

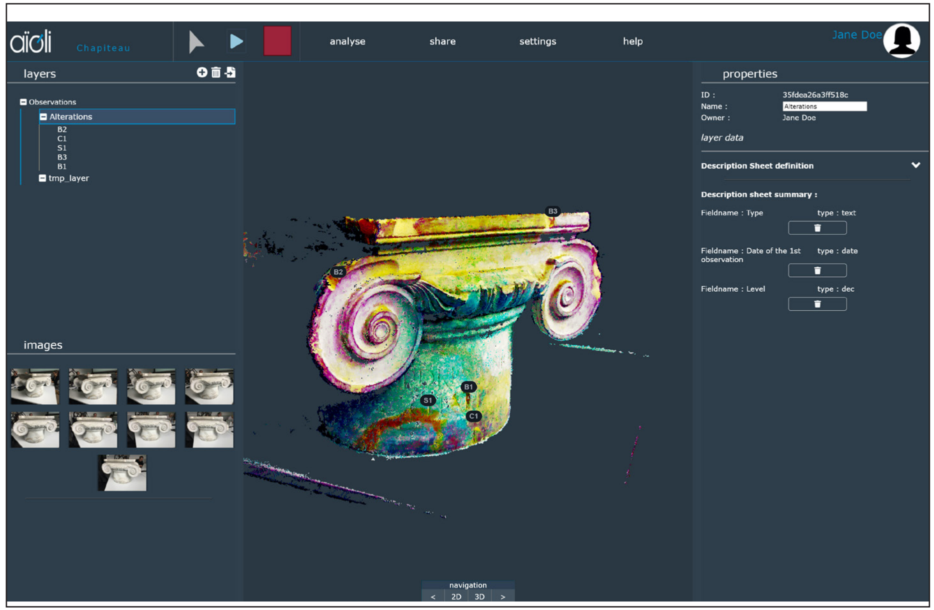


Fig. 11 - Point cloud processing permits the enhancement of geometrical or visual attributes while benefiting from the advantages of 3D support, particularly valuable in the case of annotations of morphologically complex phenomena.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

tional continuity is guaranteed between real and virtual environments, we can now consider how to reconcile the perceptual capacities of observers with the computational capacities of IT. Here the aim is to assist the observer in reducing the object complexity to the benefit of the intelligibility of the concepts it supports. This involves two aspects: on one hand, visualization, i.e. how can we best benefit from the attributes of existing data? And on the other hand, interaction: since phenomena are observable, how can they be effectively annotated?

6.1 EASE OBSERVATION

First of all, such a survey environment offers many analysis opportunities, primarily through the possibility of processing the geometrical and visual attributes of the manipulated data to help the observation of certain phenomena difficult to perceive with the naked eye (engravings, fine reliefs, pictorial traces, ...). For this purpose, three main strategies emerge according to the nature of data, which we will illustrate in this section by taking the example of decorrelation stretching, an algorithm particularly prized for the study of decorated objects because of its ability to help the observation of paintings (Le Quellec et al., 2015).
For *ex situ* analysis, images are a privileged annotation support. From this point of view, a first strategy simply consists in proposing different processes applicable to 2D project data. The user has to select one of the project images and a processing method according to a list of algorithms that would be implemented on the calculation server, and then save the output as a new image. In the case of the decorrelation stretching, the idea is to consider that the colors components of the various elements appearing on an image are correlated with each other, (i.e., that each element is associated with a specific proportion of red, green, and blue in the case of the RGB color space), and we can highlight these combinations by replacing the colors of the original image by others, more

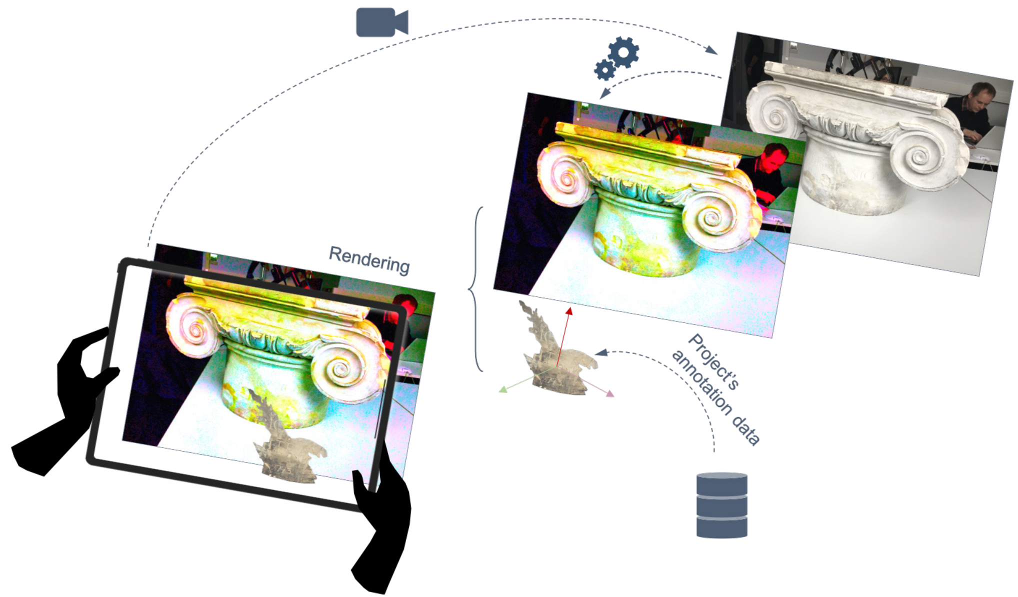


Fig. 12 - Schema of real-time video processing. The processing of the video stream by algorithms such as decorrelation stretching or edge detection could ease the observation of imperceptible phenomena when working *in situ*.

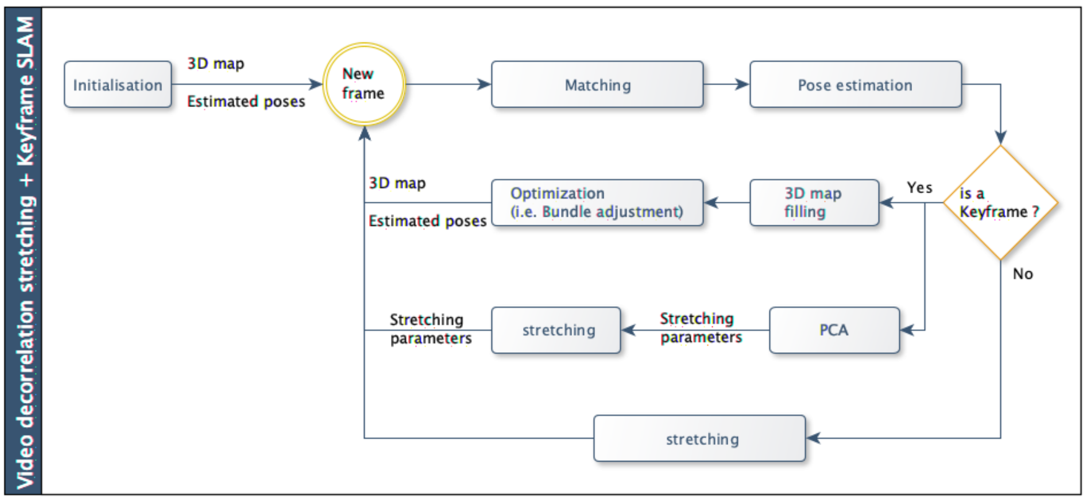


Fig. 13 - Possible approach for the optimization of the frame rate.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

distant from each other. The implementation of this algorithm to apply it to the project 2D data is quite straightforward, it consists in a linear transformation applied to each pixel of the input image, whose parameters are determined by statistical values extracted from a Principal Component Analysis (PCA), and in the writing of a new image with the output values (Alley, 1999). This transformation is given by the following equation:

$$b = s_t \ R^t s_c \ R\,(a-m) + m_{target}$$

Where:
a and *b* are vectors containing the RGB values of the pixel respectively on the input and output image;
s_t is the stretching matrix, whose diagonal elements are the made of the desired standard deviation;
R is a matrix formed by the eigen vectors from the PCA;
s_c is the scaling matrix, whose diagonal elements are made of the inverse of the standard deviation values in the new eigenspace;
m is a vector containing the mean of each band in the input image;
m_{target} is a vector containing the desired mean for each band of the output image.

Since the processed image shares the same pose as the original one already spatialized, it remains fully compatible with Aïoli's annotation process (Fig. 10). In the case of more morphologically complex elements, a second strategy consists in processing 3D point clouds to generate new textures that better account for their geometric or visual properties. Many approaches are already documented in literature (see non-photorealistic rendering or expressive rendering) and are really valuable for the understanding of CH objects. An AR use of these processes can provide information about relief variations, light behavior, or even possible links between surface aspect and morphology, while allowing to contextualize the observed phenomena in their real environment in order to put them in relation to each other. Moreover, as we will see later, point clouds can also support annotations, which can be propagated to all the project resources. Keeping the example of decorrelation stretching, this algorithm can be exploited to process not the pixels of an image but the colorimetric attributes of 3D points (Fig. 11). It can help, for example, the annotation of cave paintings that are often made on winding walls. This is only a slight variation of the previous algorithm, since it involves reading and writing a PLY file rather than an image, with no consequence for the linear transformation.
Finally, a third strategy consists in exploiting the video stream of the AR display, to process it in real time and thus directly modify the user's perception of his environment. It involves processing each video frame on the server, before sending it back to the client for display, still superimposing the project data on the output video stream (Fig. 12). While this operation is theoretically feasible, it raises significant latency issues that hamper the efficiency of the visualization. We can limit this effect by reducing the frames resolution, or by limiting the number of calculations. For example, in the case of decorrelation stretching, we could choose to apply PCA only to the tracking keyframes and to keep the stretching parameters until the next one (Fig. 13). The real-time use of this kind of processing would then constitute a valuable exploration tool, particularly to help users detect changes in materials or pigments, and thus direct their subsequent analyses towards specific areas of interest. However, in the case of a web environment such as ours, even with optimized calculations, the latency due to the number of communications between server and client seems for now to be incompatible with a truly comfortable AR use.

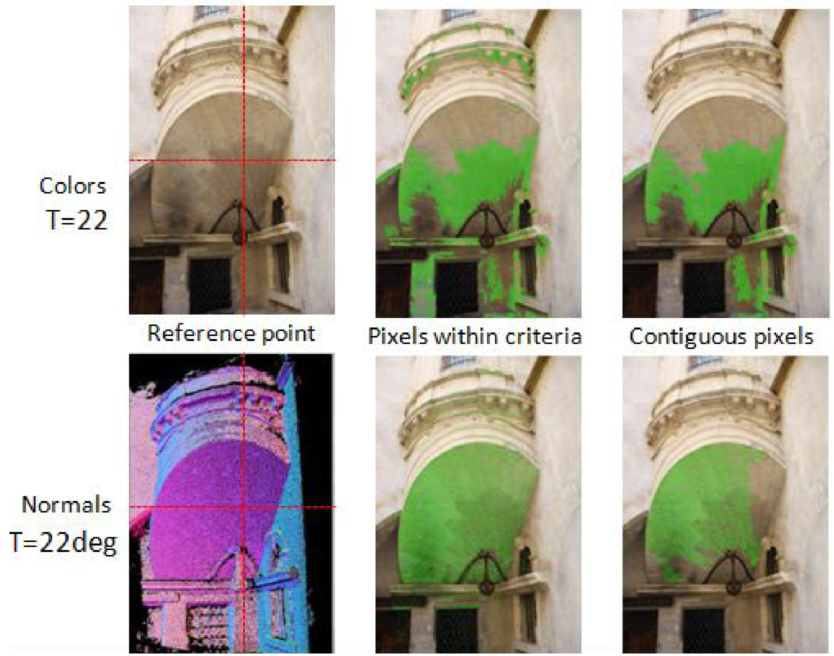


Fig. 14 - The transcription of the 3D magic wand tool can be very useful for the annotation of architectural elements, whose paths often follow simple geometric primitives. Here, the normals allow to quickly select planes, and to reproject them in 2D on the project images.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

6.2 ANNOTATE THE OBSERVED PHENOMENA

Assuming that the possibility of *in situ* work associated with enhancement techniques does indeed promote the understanding of the study objects in the scope of a survey, it remains to allow the transcription of the observations by the semantic annotation of the various project data. To do so, four ways can be considered:

Manual annotation of 2D and 3D data: First, we can extend Aïoli's annotation principle to a wider variety of media, as our approach allows us to spatially relate all the object representations in the same reference system and thus, to rely on projective geometry to directly transfer annotations from one resource to another. Thus, in addition to 2D annotation, we can propose 3D annotation modalities by allowing the user to draw the regions boundaries directly on a point cloud using ray casting. Once a path has been validated, the 3D-2D projection of the region allows to find the boundaries of corresponding areas on each 2D image of the study object.
We can also assist in the annotation drawing with selection tools. For example, we have implemented a tool inspired by the famous "magic wand" in use in many image editing softwares, which consists in manually selecting a reference pixel and then automatically testing all the other ones to check if their values are similar to the reference ± a threshold. This operation can be extended to 3D data, as a ray cast allows to quickly retrieve all the attributes of a selected 3D point, and thus to turn them into reference values, whether you are interested in colors, normals or any other geometrical attribute. Then, the attributes of the other points are tested through the shader's color function. Thus, by using normals, a single click allows a user to select all the points belonging to a same plane, or even to a same curved surface (Fig. 14).

Semi-automatic or automatic annotation of 2D and 3D data: Another way is to offer semi-automatic or automatic annotation tools based on a prior segmentation of the input data. This process consists in letting the user select an input resource (2D or 3D) as well as a segmentation algorithm from a list. The parameterization of the algorithm must then be performed by the user on the client side, before it is executed on the server. Finally, the segmentation results are returned to client for possible editing and validation of the paths (Fig. 15). To assess the viability of this approach, we tested using a point cloud plane segmentation algorithm. The aim of the experiment, described in (Manuel et al., 2018), was to generate an annotation layer automatically. From a mathematical point of view, a plane is a geometrical element quite simple to identify [2]. Here, the segmentation method was based on a split-and-merge approach. The results of this operation are regions made of coplanar 3D points from the input cloud. These regions can be reprojected on each project image (Fig. 16), so we can automatically generate annotation layers. Of course, regarding semantic enrichment, the responsibility always lies with the user to describe the identified regions.

AR annotation: A last possibility is to consider the fact that the analysis of the study object can be done by assigning elements that would not have been subject to a 2D or 3D acquisition. Current digitization approaches are mainly interested in the production of representations of the surface features of the study object. However, it would be very restrictive to limit a CH object to its physical shell, particularly in the case of architecture for which spatiality

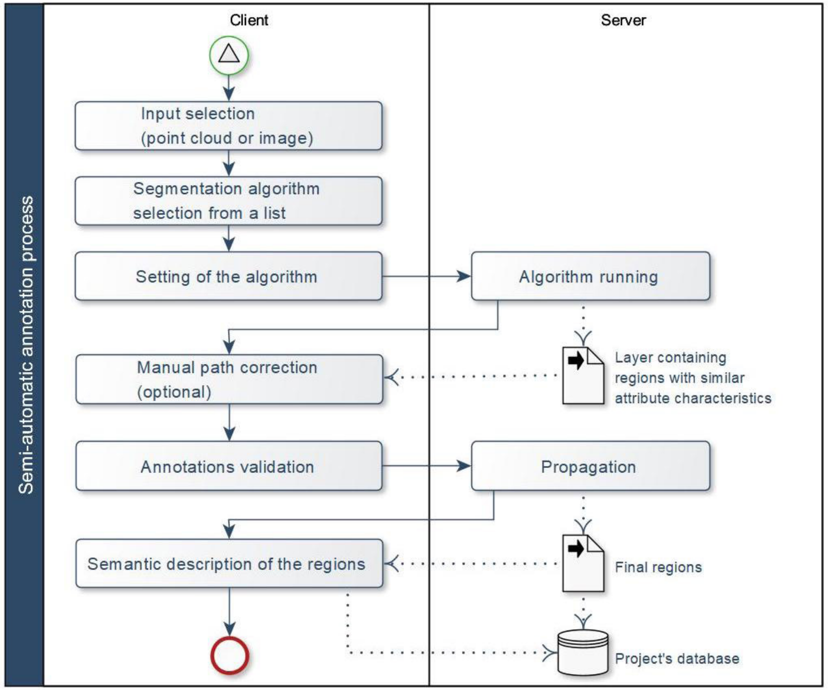


Fig. 15 - Semi-automatic annotation process based on segmentation algorithms.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

is a major matter that requires thinking about volumes. Therefore, it may be useful to take advantage of AR to propose annotation tools allowing the creation of freely positioned geometries in the environment, linked to semantic descriptions.

A simple approach would be to create a spline according to the successive positions of the mobile (Fig. 17), taking advantage of the movement freedom allowed by this hybrid visualization mode. However, it could lack precision and prove restrictive in terms of user experience: in many archaeological sites, wandering is restricted for conservation reasons. Another approach would be to rely on AR tracking anchors to create and place solid geometries. The free placement of virtual objects in the real world is a common AR use case, usually handled with a hit-test. It consists in a ray cast starting from the center of the camera and with a direction calculated from the screen coordinates of the point touched by the user on his mobile screen. If it intersects one or more tracking elements, the function returns an array containing the relevant anchors and their 3D positions. Overall, this approach considers real-world geometry as a magnetic support constraining the drawing, as in common 3D drawing or modeling softwares. For the survey, it allows us to consider a large formal variety for the analysis: simple solid geometries with symbolic aim (cube, sphere, ...) for punctual observations, polylines or polyhedrons for more complex elements. This annotation modality also offers a certain freedom regarding the nature of the descriptions: in addition to the usual text-based semantic models (tags, attributes, ontologies), we could take a deeper interest in the relational model, e.g., by giving the user the possibility to leave an audio comment located in space, or by exploiting other capture formats.

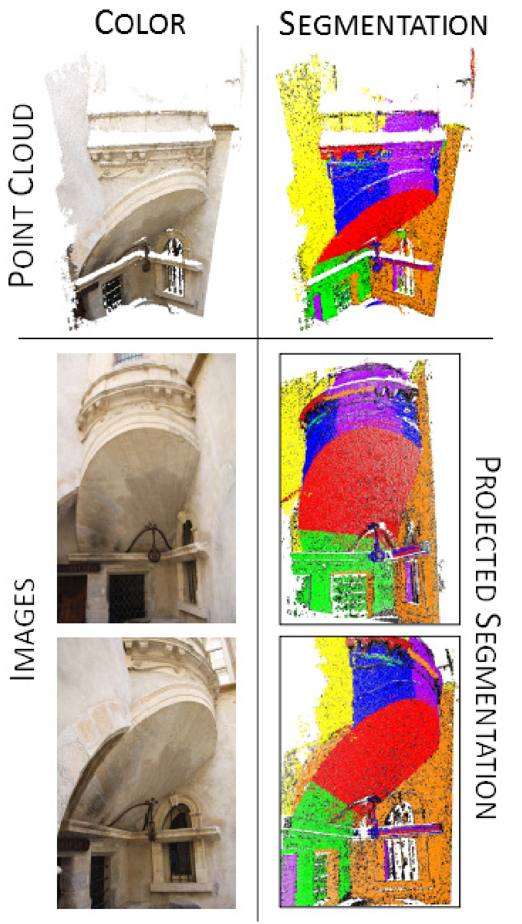


Fig. 16 - Example of segmentation result, and its reprojection on the project images.

7. TRACEABILITY: A STEP TOWARDS UNDERSTANDING THE OBSERVER/OBJECT RELATIONSHIP

The collaborative nature of the platform and its use contexts require a special care to be given to traceability issues, whether of data or observations. It refers to underlying matters such as the processes reproducibility, legal aspects concerning the content ownership, metadata production, or simply the ability to preserve the history of an element in order to follow its evolution (creation, editing, descriptions, ...). The convergence of the survey moments in a unique environment can be valuable, as we can save in the database as much information as necessary concerning the calculation processes performed or the users' actions on the data (author of an annotation, authors of a semantic description, original path, media and tool used, ...). It is no longer just a matter of ensuring contents and actions documentation, but rather of considering traceability as a key to the understanding of the relationship that links an observer to his study object. By memorizing the supports primarily exploited for the annotations carrying a given semantics, we can hope to understand which geometric or visual attributes best represent a concept. Could a link be established between the disciplines involved in an interdisciplinary survey and the privileged media to annotate? What similarities or differences can be noted in the way actors from different profiles look at the same object? From this point of view, the multimodal survey considerably broadens the perspectives. The successive poses of the camera on the mobile not only inform us about the user trajectories during his visit, but also about his areas of interest. We could even get more complete information with oculometry functions, using the mobile front camera to know at each instant which point of the rendered image is looked by the user and its 3D position. From these data, we can wonder: how does the context of the visit influence the way the object of study is perceived? In the case where different actors show an interest in the same area, do these overlaps result at the annotation level in an increase in the correlation between the regions of the different layers? These perspectives will require accumulating and crossing a significant amount of quantitative and qualitative data, which can only be done in the longer term, so our experiments are only a very first step.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

ISSN 1828-5961

8. CONCLUSION

It seems that the major survey difficulty lies in the delicate balance between the preservation of the object in its perceptive or in its geometrical properties, in an exercise aiming at the transcription of the observed facts. In this respect, the multimodal survey could constitute an interesting approach to reconsider the relevance of the usual planimetric representation paradigms. This practice is also an answer to Moravec's paradox: far from opposing each other, human capacities for high-level tasks and computer capacities for low-level tasks can converge and complete each other. Our approach consists in offering the survey actors a simplified and immediate access to the state of knowledge regarding their object of study and the possibility to contribute to it in a collaborative way. We think that this method favors scenarios in which it will be possible to identify potential correlations between different observations, different disciplines, or even to reveal under-studied elements that would benefit from a more in-depth analysis. Of course, we can also wonder whether, conversely, it doesn't run counter to the idea of a discovery process: is the access to observations and pre-established descriptions by other actors likely to influence the observer and orient his analysis to the point of causing interpretation biases? While a prototype has allowed us to validate the technical feasibility of the approach, its potential in terms of uses and acceptability remains yet to be explored. Thus, in addition to improving the reliability of the prototype and its experimental tools, further works will focus on the response to specific analysis scenarios and the field evaluation of our proposal with its future end-users.

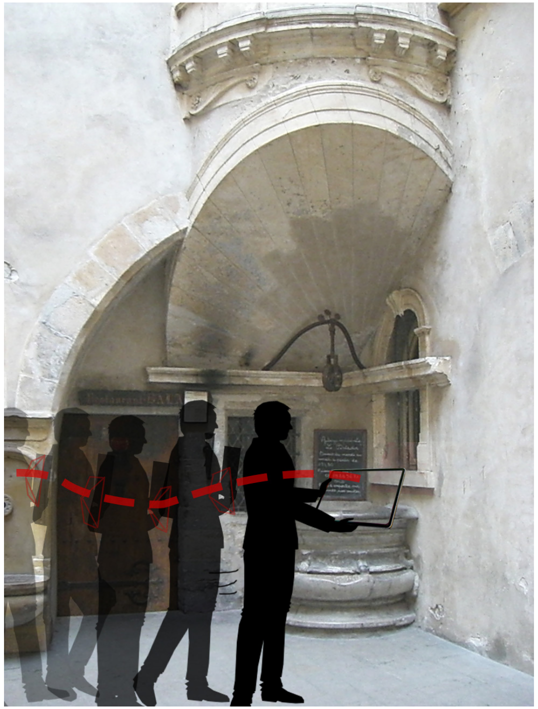


Fig. 17 - Illustration of an annotation tool relying on the user's movements.

NOTE

[1] The main technical difficulty for the registration process lies in the fact that this is a Web environment, which is still poorly adapted to AR uses despite the gradual emergence of standards, which has required "tailored" developments.

[2] In this experiment, the refinement criteria were chosen to detect coplanar elements, but their modification could allow the detection of other geometric primitives.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13

REFERENCES

Abergel, V., Jacquot, K., De Luca, L., & Veron, P. (2019). Towards a SLAM-based augmented reality application for the 3D annotation of rock art. *International Journal of Virtual Reality*, 19(2), 17-26.

Alley, R. (1999). *Algorithm theoretical basis document for : Decorrelation stretch*, Jet propulsion Laboratory, V.2.2, Pasadena, CA, USA.

Apollonio, F. I., Basilissi, V., Callieri, M., Dellepiane, M., Gaiani, M., Ponchio, F., Rizzo, F., Rubino, A. R., Scopigno, R., & Sobra', G. (2018). A 3D-centered information system for the documentation of a complex restoration intervention. *Journal of Cultural Heritage*, 29, 89-99.

Bekele, M. K., Pierdicca, R., Frontoni, E., Malinverni, E. S., & Gain, J. (2018). A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. *Journal on Computing and Cultural Heritage*, 11(2), 1-36.

Calvanese, D., Liuzzo, P., Mosca, A., Remesal, J., Rezk, M., & Rull, G. (2016). Ontology-based data integration in EPNet : Production and distribution of food during the Roman Empire. *Engineering Applications of Artificial Intelligence*, 51, 212-229.

De Luca, L. (2009). *La photomodélisation architecturale : Relevé, modélisation, représentation d'édifices à partir de photographies*. Eyrolles.

De Luca, L., Manuel, A., Alaoui, A., Abergel, V., Pamart, A., & Cao, I. (2018). *Aïoli – A reality-based 3D annotation platform for the collaborative documentation of heritage artefacts*. Retrieved January 11, 2018 from http://www.aioli.cloud/

Fuchs, P., Hugues, O., & Nannipieri, O. (2010, décembre). Proposition d'une Taxonomie Fonctionnelle des Environnements de Réalité Augmentée. A*FRV2010. Cinquième Journées de l'Association Française de Réalité Virtuelle et de l'Interaction 3D.*

Havemann, S., Settgast, V., Berndt, R., Eide, Ø., & Fellner, D. W. (2009). The arrigo showcase reloaded—Towards a sustainable link between 3D and semantics. J*ournal on Computing and Cultural Heritage*, 2(1), 1-13.

Le Quellec, J.-L., Duquesnoy, F., & Defrasne, C. (2015). Digital image enhancement with DStretch ® : Is complexity always necessary for efficiency? *Digital Applications in Archaeology and Cultural Heritage*, 2(2-3), 55-67.

Lo Buglio, D. (2016). *Caractérisation de formes architecturales, Une approche expérimentale intégrant complexité et intelligibilité des représentations numériques* (doctoral dissertation) Université Libre de Bruxelles, Belgium.

Manuel, A., De Luca, L., & Véron, P. (2014). A Hybrid Approach for the Semantic Annotation of Spatially Oriented Images. *International Journal of Heritage in the Digital Era*, 3(2), 305-320.

Manuel, A., M'Darhri, A. A., Abergel, V., Rozar, F., & Luca, L. D. (2018). A semi-automatic 2D/3D annotation framework for the geometric analysis of heritage artefacts. *2018 3rd Digital Heritage International Congres*s, 498-505.

Meghini, C., Niccolucci, F., Felicetti, A., Ronzino, P., Nurra, F., Papatheodorou, C., Gavrilis, D., Theodoridou, M., Doerr, M., Tudhope, D., Binding, C., Scopigno, R., Vlachidis, A., Richards, J., Wright, H., Geser, G., Cuy, S., Fihn, J., Fanini, B., & Hollander, H. (2017). ARIADNE : A research infrastructure for archaeology. *Journal on Computing and Cultural Heritage*, 10, 1-27.

Niang, C., Marinica, C., Markhoff, B., Leboucher, E., Malavergne, O., Bouiller, L., Darrieumerlou, C., & Laissus, F. (2017). Supporting Semantic Interoperability in Conservation-Restoration Domain : The PARCOURS Project. *Journal on Computing and Cultural Heritage*, 10(3), 1-20.

Pamart, A., Morlet, F., & de Luca, L. (2019). A fully automated incremental photogrammetric processing dedicated to collaborative remote-computing workflow. *ISPRS, XLII-2/W9*, 565-571.

Pamart, A., Ponchio, F., Abergel, V., Alaoui M'Darhri, A., Corsini, M., Dellepiane, M., Morlet, F., Scopigno, R., & De Luca, L. (2019). A complete framework operating spatially-oriented RTI in a 3D/2D cultural heritage documentation and analysis tool. *ISPRS, XLII-2/W9*, 573-580.

Saint-Aubin, J.-P. (1992). *Le relevé et la représentation de l'architecture.* Inventaire général des monuments et des richesses artistiques de la France.

Serna, S. P., Schmedt, H., Ritz, M., & Stork, A. (2012). Interactive Semantic Enrichment of 3D Cultural Heritage Collections. *VAST: International Symposium on Virtual Reality*, 33-40.

Shi, W., Kotoula, E., Akoglu, K., Yang, Y., & Rushmeier, H. (2016). CHER-Ob : A Tool for Shared Analysis in Cultural Heritage. *Eurographics Workshop on Graphics and Cultural Heritage*, 187-190.

Soler, F., Melero, F. J., & Luzón, M. V. (2017). A complete 3D information system for cultural heritage documentation. *Journal of Cultural Heritage*, 23, 49-57.

Thivet, M., Verriez, Q., & Vurpillot, D. (2019). Aspectus : Développement d'une plateforme collaborative en ligne d'analyse des données 3D pour l'archéologie et le patrimoine. *In Situ*, *39*.

Yu, C.-H., Groza, T., & Hunter, J. (2013). Reasoning on Crowd-Sourced Semantic Annotations to Facilitate Cataloguing of 3D Artefacts in the Cultural Heritage Domain. *International Semantic Web Conference*, 8219, 228-243.

Yu, C.-H., & Hunter, J. (2013). Documenting and sharing comparative analyses of 3D digital museum artifacts through semantic web annotations. *Journal on Computing and Cultural Heritage*, 6(4), 1-20.

DOI: https://doi.org/10.20365/disegnarecon.26.2021.13