Mixing visual media for Cultural Heritage

Roberto Scopigno [0000-0002-7457-7473]

CNR-ISTI, Pisa, Italy
r.scopigno@isti.cnr.it

Abstract.

The Cultural Heritage (CH) domain is a field where many different visual media are constituent elements of the main activities: study, conservation, dissemination, and presentation to the public (museum visitors, tourists, practitioners). Those media are usually used in isolation, adopting specific visualization tools. This paper aims to present several experiences where multiple visual media have been used in a coordinated manner by fusing or presenting them in the same visualization context. These approaches experimented with new interaction and visualization methodologies to use different media in a synergic way. CH domain is an ideal field of experimentation of the potential of media integration/fusion/cross-analysis. According to our understanding, using multiple media can improve insight capability. We guide the reader in the analysis of some pioneering experiences and approaches and try to deduce, for each of them, the potential improvement granted in terms of data communication or analysis. A final discussion tries to highlight the work needed for a wider acceptance and increased impact of those approaches.

Keywords: Cultural Heritage, visual media, interactive visualization, media integration.

1 Introduction

Visual media play a foundational role in communication concerning art and Cultural Heritage (CH). In our era, visual media are de facto mandatory ingredients in communication and dissemination actions [1,2]. Indeed, the presentation of visual media is not straightforward as text or speech, mostly when more sophisticated media or presentation modes are adopted (requiring user training and specific devices). For example, impaired people deserve proper support in the fruition of visual media. Thus, several methods have been proposed to translate images into other stimuli easier to perceive by impaired people, such as the conversion of paintings into a tactile surface [3]. Anyway, communication is enforced or simplified when it is based on visual stimuli. This work's scope is not to make a historical essay of the use of visual media for CH communication, but conversely to present some flashes on how they are used. The digital domain introduced new visual media modalities that go beyond the usual

picture used as an illustration. These new approaches employ dynamic and interactive fruition (interactive navigation, visual comparison, interactive lighting, etc.).
A CH content producer can choose his preferred media type and visualize it with modern display/interaction tools. But CH practice has also shown that in several applications, a single type of media is not sufficient anymore because an adequate representation and analysis is possible only by merging and integrating multiple media and datasets. Therefore, canonical digitization and visualization technologies require being augmented by software technologies, fusing or creating relations between different visual datasets or media.
This work's primary focus is to discuss a specific opportunity: different visual media can be used in a synergic manner, removing the constraints of accessing a single media at a time. We can experiment with several different ways to integrate, fuse, or cross-compare different media. To enable this, we need technical instruments and methodologies: while single media management is nowadays relatively straightforward from a technical point of view, the integration, fusion, cross-navigation of those media still present some issues and potential for further research (data management, user interfaces, interaction metaphors, new approaches in visual presentation and storytelling). Therefore, this paper aims to present a few previous experiences concerning media integration, showing how to manage those data and which benefits we gather in terms of understanding and communication.
This paper's application context is primarily public fruition and communication of CH to the public and the related and partially intersecting domain of the communication to more advanced and expert communities (such as art historians, conservators, and restorers).
The paper will cover this open subject, endorsing a practical approach. The first step is defining the different media types (Sect. 2). We mention the case of visualization tools and how they could focus on multiple data types (Sect. 3). Then, we review available experiences where visualization technologies allow users to browse and analyze different media in the same shared space, intermixing data content in a single visualization context (Sect. 4). Digital systems supporting conservation and restoration also need similar features (Sect. 5). We conclude with some final remarks (Sect. 6).

2 Data types in CH applications

The technological status is mature if we consider the enabling technologies for building visual assets (either accurate digital samplings of reality or recreation of lost or dismantled artworks/architectures) and presenting them using interactive displays or VR/AR devices [4]. The excellent communication potential of these technologies for museums and the CH community is quite clear.
Digital representations come with different data types and data sources — whether 2D, 3D, 4D [5], or beyond, digitized or modeled. Multiple data types are also used to produce digital representations (e.g., 3D models [6]). We introduce here those different media.

2.1 High-resolution Images

Images are the more common visual media. They have been part of art history and archaeology from the very beginning, initially employing the analogic, printed version and, more recently, digital supports (either digitally native images or scanned from old prints/slides).
While images are fully integrated with the web and HTML since the *Wold Wide Web* birth, few aspects lack a standard solution for archival and visualization purposes. Most of the images produced nowadays are very high-resolution. High-resolution images are now a commodity resource, given the impressive evolution of digital photography (just to mention a single example, recent off-the-shelf smartphones provide 20-80 Mpixel cameras). Moreover, the availability of tools that allow aligning and stitching image patchworks supports users in reaching huge image resolutions.
Visualization on the web of high- or huge-resolution images can be tricky but doable, using specific methodologies (data compression, efficient progressive data transmission). Examples are the multiresolution approaches based on tiling and hierarchical image representations [7,8]. Another important and critical issue could be protecting the data (image watermark technologies [9]).

2.2 Multispectral images

Images depicting the image reflected by a surface while sampling only a specific wavelength are quite common in CH, especially for investigation and conservation purposes. Some examples are *infrared images* (allowing the detection of under-drawings) or *ultra-violet images* (which disclose the presence of patinas or layers of biologic substances, as well as faded or modern paints).

2.3 Reflection transformation Images (RTI)

Re-lightable images (usually named *Reflection Transformation Images - RTI*) are becoming an increasingly used technology to acquire detailed and interactive documentation on quasi-planar objects [10, 11]. This is particularly useful for artifacts characterized by complex light reflection attributes. The advantage of this representation is the possibility of changing the direction of the light incident over the image in real-time (i.e., at visualization time) and using enhanced visualization modes to inspect fine details of the objects' surface. The visual quality and fidelity supported by this media are impressive, in many cases, superior to what we can simulate with 3D models.
RTI images' acquisition is quite simple, requiring a calibrated lighting system and shooting multiple photos for a stationary camera position under variable lighting. Those input images are then processed to produce a single RTI image. The images' lighting information is mathematically synthesized to generate a mathematical surface reflection model for each specific parcel of the surface reflection, enabling users to re-light the RTI image interactively and in real-time. The RTI image encodes, for each pixel image, nor the RGB value but a *function* able to return the surface's color given a specific direction of the incoming light incidence.

2.4 Video

Video is another medium that is extremely easy to produce (a mobile phone is enough for many application cases) and present to the public (the commercial platform YouTube is an immediate example). But we need more professional grabbing devices and a solid knowledge of the cinematographic language if the goal is to produce high-quality videos.
Moreover, videos are not just grabbed but also synthetically produced with computer animation tools. This latter source of data is becoming common in CH.

2.5 Panoramic/360 images and videos

Panoramic or 360 degrees images[1] or videos are a quite new media which adds an *interactive* opportunity, allowing users to navigate and interact with these visual assets. Acquisition easiness and speed, together with the richness of details granted by the high-resolution supported, makes panoramic images an ideal medium for CH[2].
The acquisition of panoramic videos is straightforward, either from multiple poses taken by a camera positioned over a tripod or using specific devices (based on multiple cameras).

2.6 Terrains models

Terrain models are commonly termed as 2.5-dimensional data. They are quite common in geographic or land representations and are often used to represent the context of CH discoveries visually. These data are managed with GIS approaches or as standard 3D data.

2.7 3D models

3D representations also become quite common in CH. Two classes of models are used:

- *Sampled models*, usually produced with active 3D scanning (laser-based systems or systems using structured light) or with photogrammetry approaches (production of 3D models from set of 2D images);
- *Modelled representations,* produced using the user-driven modeling systems designed for 3D modeling and computer animation applications (e.g., Blender, Maya, etc.).

Sampled models give much more control on the accuracy of the representation than modeled representations. Conversely, the latter are more common for broad public applications (e.g., to produce videos or virtual reconstructions).

---

[1] https://en.wikipedia.org/wiki/360_photography
[2] For an example of panoramic images adopted to enable the virtual visit to museum see: http://www.youvisit.com/tour/louvremuseum (accessed on 26 November 2020).

There is a pressing need for platforms supporting easy and free publication on the web of 3D models. SketchFab [12] is a recent commercial solution, supporting automatic web publishing and a nice and easy to use interface.

Animated/deformable 3D models are used in CH to present and add an interactive behavior to the artifact. They are complex models which allow encoding both the shape and the functioning of an artifact. Therefore, not just how it looks but also how one could operate/manipulate/act.

Depth images (or *depth maps*) are another type of visual data with a predominant 3D interpretation. These are usually the raw result of 3D sampling performed with an active scanning device: an image where each pixel encodes a point sampled in the 3D space. We can render a depth map as an image (thus, rendering in false color the distance from the observer), or we can convert it into a cloud of points.

2.8 Beyond visual – Sound

Sound is another essential component for the simulation or representation of virtual spaces. A silent 3D scene is not realistic nor sufficiently immersive in a virtual reality context. Thus, the sound should be taken into account while producing sophisticated visual products, as computer games that pair visual and sound contents. Anyway, the scope of this work is limited, and we cannot cover here also the case of sound and its use in CH applications.

3 Data integration at the level of the generalist visualization tool

The primary task with visual media is to provide proper visualization tools. Visual media are usually complex data and thus, we need efficient visualization processes, taking care of several technical issues (data simplification, multi-resolution encoding, compression, progressive transmission, GPU-enabled rendering). All those themes have been the subject of intense research. Nowadays, we dispose of an arsenal of efficient technical solutions for the implementation of effective visualization tools.
Visualization tools are available for each different media (e.g., RTI images [13], 3D models [14, 15, 121, 16], videos [17]). An issue perceived by the CH community is that each tool has a different interface and its specific design; moreover, many of them have been designed for technical communities and expose to the user a very complex set of features and parameters. Not all those features are of interest for the average CH user, and more are the features more complex is usually to master the tool. Moreover, since those tools have different interfaces, the synchronous analysis of different media becomes a rather cumbersome process.

Workflows requiring the contemporary analysis of multiple media are common in the CH domain. Therefore, a comprehensive approach has been endorsed by Yale University while designing the CHER-Ob system [18]. Their starting point was to provide support for the many different media types and formats needed in CH through a single tool. The goal is to offer a single and congruent GUI (an example is the manipulation modality, implemented frequently in graphics systems with different interpretations of the trackball concept), unified terminology, and the usual basic visual analysis functionalities. CHER-Ob is an open-source software tool, having a single intuitive interface consistent with CH needs. The system handles various types of 2D and 3D data and preserves user-generated metadata and annotations (see Figure 1).

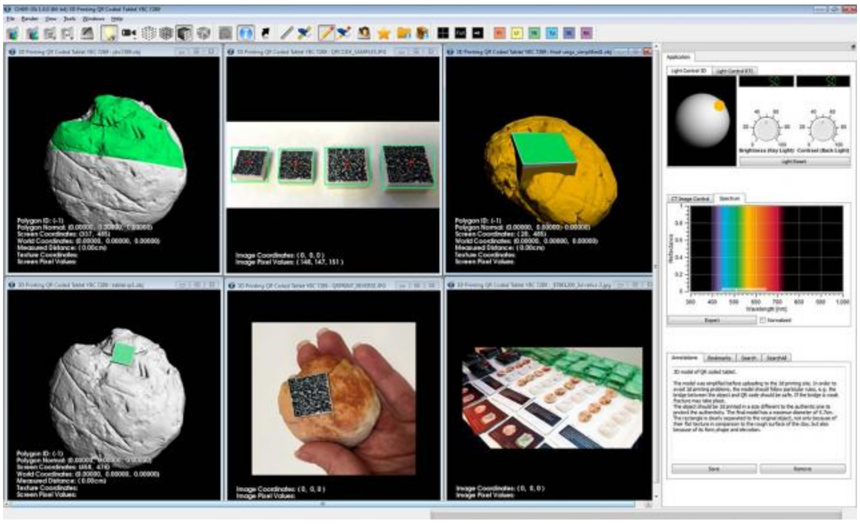

Fig. 1. A snapshot from the CHER-O system (YALE University, https://graphics.cs.yale.edu/software-packages/cher-ob-open-source-platform-shared-analysis-and-video-dissemination-cultural)

According to our experience, this is the first complex open-source tool focusing on data visualization of CH digital assets rather than on generic geometry processing tasks. It provides tools for the integrated visualization of various types of visual data (2D images, hyperspectral images, volume data encoded using the DICOM standard and triangulated 3D models, with or w/out textures). But in this system each data resides in its specific visualization window, i.e., data are not fused or integrated.

4 Towards data integration in interactive visualization

A step beyond is providing methodologies and tools for integrating and fusing different data types in the framework of a single visual presentation context. The following subsections present some related experiences. In each of these subsections we present a representative example(s) and describe which research questions could be more easily answered, the lessons learned, or the potential impact of this approach for the CH community.

4.1 2D images and 3D models

The study of artworks, the visual or narrative presentation provided to a generalist user (museum visitor, tourist), often requires both 2D and 3D media. In our vision, there is not a predominant medium, since each one has specific advantages (visual quality, resolution, completeness, easy of manipulation or analysis, etc.). Thus, we should offer access to both media, possibly in a coordinated manner.

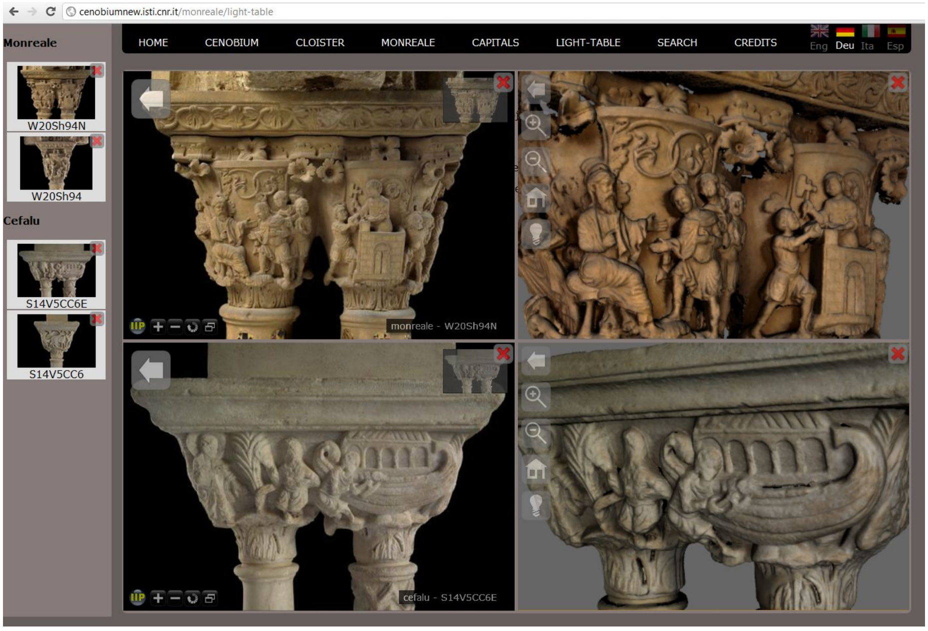

Fig. 2. The light table component of the Cenobium system, supporting side-by-side analysis of a selected group of 2D / 3D representations.

This requirement was at the basis of the design of the Cenobium[3] system on 2006 [19]. We designed this system to support art scholars and students in studying medieval sculptures (in this specific case, decorated capitols) through multiple visual representations. From the very beginning, the Cenobium system provided what we called the *light table*, i.e., a facility to support the side-by-side visual analysis of different media. The Cenobium system provides several high-resolution images and a 3D model for each capitol; the user can select any image or 3D model (belonging to the same or different capitols) and selectively activate their visualization in a common interactive space (see Figure 2).

[3] Cenobium project: http://cenobium.isti.cnr.it/

A much more sophisticated approach for the joint management of 2D images and 3D models was introduced by the seminal paper PhotoTourism [20]. This paper presents a single working space where the user can navigate a 3D scene and, in a synchronous manner, a large set of geo-located 2D images. The availability of geo-location information for each 2D image (position and orientation of the camera at shooting time) allows locating them precisely in the 3D space and navigating and visualizing all data with a unified interface.
This kind of joint dataset (a high-resolution 3D model plus an extensive collection of pictures), describing a real-world object or a scene, can be generated by multiple sources and is becoming increasingly common due to the advent of photogrammetric methodologies for 3D acquisition [21, 22].

This approach has been later on refined for CH use by the PhotoCloud[4] system [23]. It extended the PhotoTourism approach by: redesigning the GUI with a more flexible management of the image thumbnail-bar [24]; adding increased 3D data flexibility (allowing the use of high-res triangulated 3D models and not just pointsets); proposing a methodology for computing image cameras over unregistered image dataset; and improved visualization and navigation features that fully exploit the underlying 3D dataset.

The potential of this new visualization paradigm is of high interest for CH applications. There are several CH conservation or monitoring activities where the capability of managing in a unified context a 3D representation and multiple sets of images (taken at different times and, maybe, also by other people) could be a crucial asset. Another example is the added value of presenting historical photos of an urban context to tourists, by immersing the old photographs in the 3D scene and enabling an easy navigation and selection of the images.
One major innovation introduced by this approach is the new methodology introduced for browsing a large set of images, solved with the adoption of a *space navigation* approach: while I navigate the space, the system presents in real-time the subset of 2D images nearer to my current view. That seems to be a major innovation with respect to previous image archives supporting a query-based interface.
Thus, one could wonder why we have not seen a real impact in CH applications of this innovative approach. It might be due to the increased implementation complexity of a system such as PhotoCloud compared with a standard image browsing system or a 3D viz tool. Since CH is an under-funded domain, unfortunately, complex approaches (either to implement or to use) often do not pay.

---

[4] PhotoCloud system: http://vcg.isti.cnr.it/photocloud/ and https://www.youtube.com/watch?v=-Bb9k2Gy2Yg

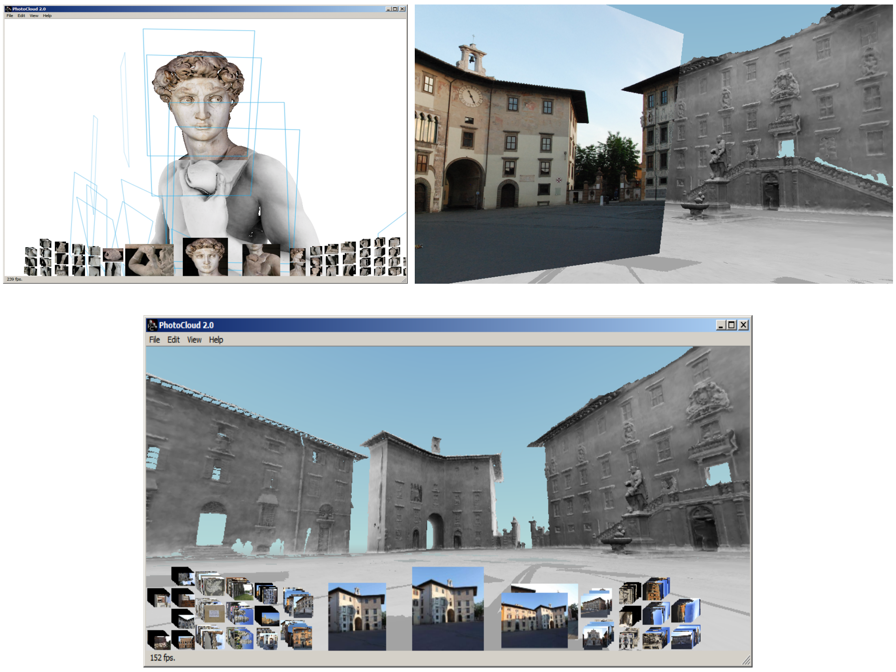

Fig. 3. Two examples of use of PhotoCloud: over a statue dataset and in an urban context (one of the 2D images is projected on the 3D scene); the third image (on the bottom) presents the thumbnail-based interface provided to browse the image set.

Research questions / Lesson learned / Impact:
Presenting 2D and 3D data in a unified rendering context allows:

- To simplify the search of images in an extensive database (replacing search queries based on text or codes with virtual navigation);
- To enable easy comparison of images taken at different times and covering the same physical space (e.g., monitoring of changes in urban contexts);
- To use images to increase the visual quality of 3D models (e.g., mapping high-resolution 2D content over a low-resolution 3D model; adding effects depending on illumination, weather condition, time of the day);
- Requires sophisticated solutions able to geo-locate precisely the input images over the 3D scene (now possible in an automatic manner with AI-based solutions).

4.2 Textual data and 3D models

The integration of textual descriptions and visual model is another potential subject of work in the fusion of different media. One early attempt was the inclusion of 3D models in pdf files pioneered by Adobe[5], thus evolving the usual static figures into interactive 3D illustrations.

An ideal testbed to experiment a tighter integration of 3D models and text was a project concerning the interactive presentation of a peculiar artwork, very complex in size and decorative apparatus. The Ruthwell Cross is a big (5 meters tall) medieval stone cross whose surface contains carved figures, symbols, and runic inscriptions related to an ancient poetic text. It is an artwork that cannot be presented by just a visual representation, as well as we cannot describe it using only text. We verified that those media do a much better work jointly.

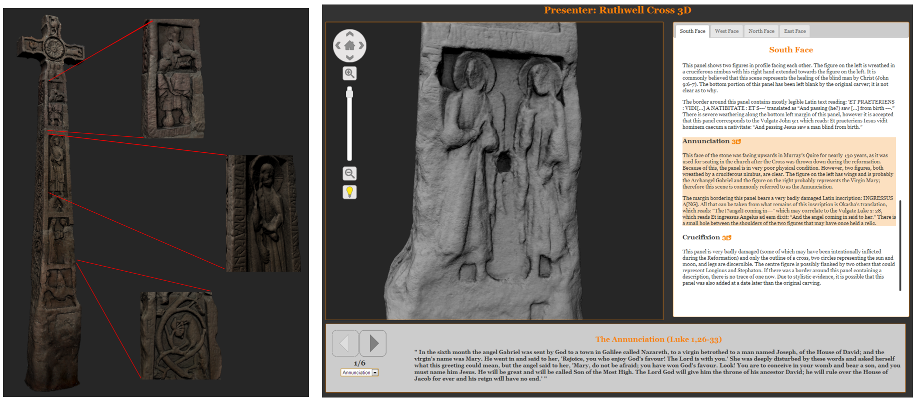

Fig. 4. The Ruthwell cross (left); (right) the snapshot presents the system used in Lecture mode, the text description has induced the system to open a view over the 3D model which poses in foreground the related carved panel.

We designed a web-based presentation platform [25] to provide full integration of 3D content with heterogeneous multimedia data; it allows to easily encode and visualize interlinked modules, covering different types of data (i.e., text, 3D models, and digital editions of the poetic text). Users can select the navigation over the 3D model, discovering the minutes geometric and iconographic details (automatically, he will also receive indication by the system on the specific text section(s) which describe the region under visual focus). Conversely, the user can read the descriptive text (in this case, the system will highlight the region over the 3D model associated with the specific text section and automatically pose the user in the ideal view to visually analyze that region). The Ruthwell cross is an ideal testbed for this approach since the artwork is very dense of figurative carvings and carved inscriptions; decoding and understanding those elements is quite complicated and requires a dense, descriptive apparatus.
Since the Ruthwell cross project was directed towards the CH community (specifically, to provide didactical support for teaching), great attention was given to the design of

[5] 3D pdf: https://helpx.adobe.com/it/acrobat/using/displaying-3d-models-pdfs.html (accessed on 26 November 2020).

an intuitive and easy to use GUI[6]. Navigation over the cross was implemented to facilitate interaction (an example is the support for constrained view selection).

Research questions / Lesson learned / Impact:

- Complex artworks require sophisticated descriptions (text) and long interactive visualization sessions. The availability of a common data presentation context (3D & text) allows to structure the access to those data channels in a coordinated manner;
- Testing this approach with students produced very encouraging results;
- A system which allows both *free navigation/inspection* and *guided navigation* is also convenient. Guided navigation is provided through a sequence of predefined "lectures" (these are similar to the slides of a Power Point presentation, they allow to disclose knowledge on the artwork by presenting to the student several starting points to both the text and the 3D model);
- Structuring the content is a time-consuming process (writing text, structuring it in sections, creating the links with the visual content, designing the "lectures"); it requires a motivated tutor, but it could also be an ideal cooperative task for a course project, involving the teacher and some students.

4.3 Panoramic images and 3D models

Panoramic or 360 images are an excellent type of media. Even if they are based on 2D content, they provide a visual experience that emulates 3D navigation, offering some advantages (the visual quality offered by 360 images is often much better than the one of 3D models, while data encoding complexity is smaller) and disadvantage (they are usually sampled sparsely and limit the interaction of the user to just a rotation of the view direction).

Integration with a 3D representation could be an excellent solution to solve the inherent sparsity of 360 images. This approach has been endorsed and experimented by the Zamani Project[7] [26]. The Zamani Project is an excellent example of a wide-scope digitization project taking place in entire Africa. It undertakes data collection and analysis, heritage communication, training, and capacity building for experts and the public. It provides access to high-quality spatial heritage data (most of them related to endangered African heritage sites in complex-to-reach locations). It aims at fostering public capability to learn from, preserve, and protect cultural heritage.

The Zamani staff proposed an excellent example of media integration: the navigation of 3D scenes enriched by panoramic images, as it is exemplified in one of their videos[8].

---

[6] A video of the system presenting the Ruthwell cross is at:
https://www.youtube.com/watch?v=Wov-2ik4ibY (accessed on 26 November 2020).

[7] Zamani Project: https://zamaniproject.org/

[8] Zamani – Use of 3D & Panoramic images in the video "3D Heritage Documentation of African Heritage Sites" (accessed on 26 November 2020):
https://www.youtube.com/watch?v=8jTlKjUAzn8&list=PLl-WdPQN2XnmR4lj19bRRCkT0FEQW1E9b8

Here, the approach is to immerse several 360 images inside a 3D scene, marking their position in the scene with a *glass sphere* depicting the related 360 image's visual content. The idea of using the *glass sphere metaphor* to present visually the presence of a panoramic image in a specific location is space is brilliant: very easy to understand and to interact with. Once the user is navigating the 3D scene, as soon as he passes in the proximity of a 360 image, he is trapped by the sphere. The visualization context switches from the 3D navigation to the 360-image navigation. The user can then escape from the sphere/panoramic image and continue navigation over the 3D scene. This is shown at the suggested link by a pre-recorded video (see Figure 5), but can be implemented in interactive browsers as well.

Fig. 5. A clip from a video of the Zamani Project, showing the use of 3D navigation enriched by the presence of interactive 360 images (represented with *glass spheres* immersed in the 3D scene).

A similar approach is also endorsed in [27]. The authors propose the integration of high-resolution spherical panoramas, a variety of maps, GNSS, sound, video, and text information for representation of numerous cultural heritage objects. The focus is to include other media (for example maps) inside the navigation of the panoramic images; thus, maps and other data are rendered inside each single panoramic image.

Research questions / Lesson learned / Impact:

- It is an ideal way to solve the limitations of the two media: the scene contains only a few spherical images, but the visual quality provided by each of them is very high; the 3D model is by definition continuous and complete, covering all the visible scene, but we know it is not easy to sample at full resolution and high-fidelity (color and surface reflectance) large scenes. Joining the two media allows summing the respective advantages.

- It is relatively easy to integrate spherical images and 3D into an interactive visualizer. Switching among media is also quite intuitive.

5 Data integration while supporting conservation and restoration

CH conservation and restoration are important professional domains, where media integration is also required.

5.1 Integration of different media for conservation condition analysis

Multiple scientific investigation technologies are commonly used to assess the conservation conditions of artworks. In many cases, the diagnostic results produced are visual data (RGB or hyperspectral images, 3D models, discrete volumetric representations, etc.). The clear comprehension of a phenomenon often requires to analyze all those investigation data, not in isolation but jointly.
A very recent example is the study of a painting (the Ecce Homo by Antonello da Messina, oil on wood). This very fragile artwork has been subject to a complex scientific investigation project [27]. One of the conservator's research queries was to get insight into the relations linking several areas on danger visible over the surface (cracks and lifting of the painted surface) to possible woodworm galleries existing in the underlying wood substrate.
The investigations executed to gather data on this phenomenon were: the high-resolution 3D scanning of the surface, aimed at producing a very detailed sampling of the status of the painted surface; the CT scan, aimed at sampling the conditions of the wooden table in its internal volume. The data produced can be analyzed, using standard 3D visualization tools (for the 3D scanned geometry) and tools for the segmentation and visualization of volumetric data (for the CT scanning data). Nevertheless, the conservator's analysis would have been simplified by a system able to render both datasets in a joint context. Therefore, one of the tasks in the Ecce Homo conservation project was to segment the CT data, detect all the worm galleries, produce the related 3D geometries, and design a visualization and analysis tool able to render both geometries (panted surface and worm galleries). The goal was to enable the interactive search of correspondences of color detachments and underlying worm galleries, to assess potential critical areas over the painted surface.
The resulting system, implemented as a web application on top of 3DHOP [15] allows using transparency and cut-through/sections to discover these potential pairs and to compute metric distances between locations over the painted surface and the underneath wood cavity. An example is presented in Figure 6 or, better, in a video[9].

[9] The EcceHomo project is presented at: http://vcg.isti.cnr.it/activities/eccehomo/ and a video is at: https://www.youtube.com/embed/_cG0uR_h8VM (accessed on 26 November 2020).

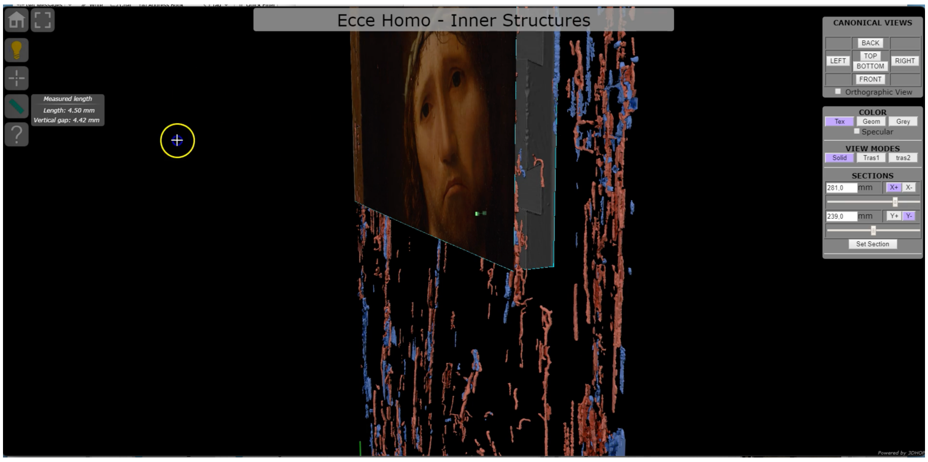

Fig. 6. Analysis of the status of the Ecce Homo painting: the painted table is tilted, only part of its panted surface is rendered using interactive sections, the underlying worm's cavities are rendered. The depth of the nearer worm gallery to the point selected by the user (green point, on the right-bottom of the Ecce Homo mouth) is computed and shown in the left panel (4.42 mm).

Research questions / Lesson learned / Impact:

- We can visualize and interrogate data on the painted surface (optical 3D scan) and the interior (radiographic CT scan). This allows easily to create correspondences and to compute measures (e.g., distances) to give numeric evidence to the conservator's queries.
- The same approach can be extended to multi-spectral images to add other data to be cross-related with the 3D data.

5.2 Integration of different media in restoration documentation

A major focus in CH restoration and the related digital documentation is: (a) to allow storing data and documents by means of the geo-referenced interconnection of those sources of informations over the digital 3D representation and (b) to annotate the 3D model to encode information or insight gathered. An example is the documentation system developed for the Nettuno Fountain restoration [29]. This system enriches a high-quality 3D model with hot-spot links to other media (pdf text, images, drawings) and with annotated regions, defined over the 3D surface and associated with textual descriptions. It was implemented as a web-based system using 3DHOP technology [15]. In this information system (based on an underlying relational data base) the 3D model acts as a spatial index to the information contained and referred (i.e., to several other types of data) and as a canvas where the restores can draw and select spatially-defined annotations.

The AIOLI[10] restoration documentation system developed by CNRS adopts also a similar approach. It manages 3D representations (based on pointset), RGB images and RTI images in the same visualization space [30]. This system also supports creating and visualizing annotations, defined on a single media (in this case, 2D images) and projected automatically on all other media depicting the same spatial area. Automatically inheriting annotations on any media supported is an essential plus of this system.

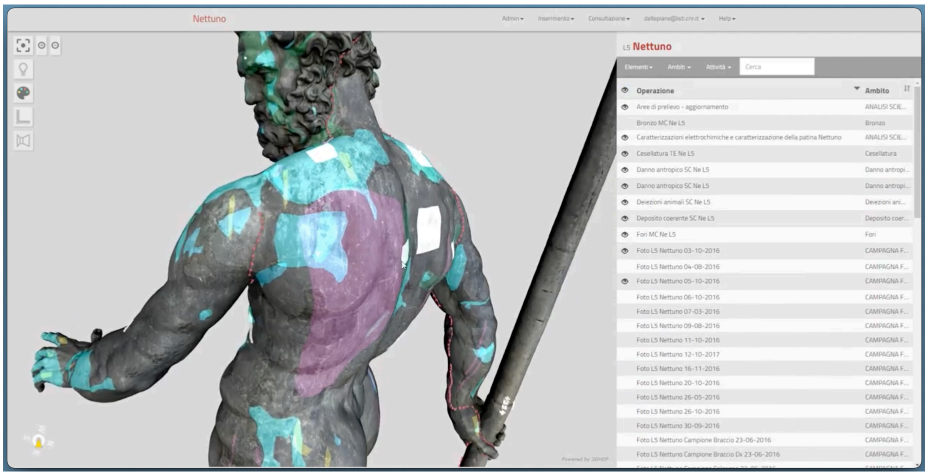

Fig. 7. The interface of the Neptune information system: a portion of the monument is rendered in forefront, some annotated regions and the related subset of data linked to this statue are presented with the list on the right and can be interactively accessed with a mouse click.

Research questions / Lesson learned / Impact:

- Restoration documentation is an application domain where multiple media types have to be stored, interconnected, and consulted.
- The 3D media could be the ideal index to available information, either augmenting the 3D model with geo-referenced links to other media, or mapping information on its surface through spatial annotations.
- The restoration of Neptune fountain demonstrated that this approach is doable, can be quickly learned by restorers and curators, and it can support cooperative and on-site work.

6 Discussion and concluding remarks

The paper has presented some examples of interactive systems where a single interactive visualization context manages different media. The aim has been to show that we should not just select a single medium and restrict our analysis capabilities. Still, we could use as many media are needed/useful for a given communication or study task.

10 The AIOLI system is presented at http://www.aioli.cloud/ (accessed on 26 November 2020).

We tried to show that using multiple media in a single application framework may increase the information we can communicate, grasp, and learn. The examples shown demonstrate that several innovative approaches have been presented and evaluated; the required enabling technology is available (even if not yet at the level of ready-to-use commercial tools).

Concerning the potential *impact* of digital solutions for CH data, we often miss proper assessment of the approaches presented in literature since only a minority of the works report the results of structured user tests involving substantial user communities. In effect, it is not so common to have formal user tests run in the framework of applications directed to the CH community (or, often, the user tests presented in scientific papers are done on tiny samples, reducing the analysis's effectiveness and level of trust). This issue might slow-down acceptance by the community.
We should say that the apparent impact of the approaches presented in this paper has been a bit disappointing. Moving to a broader discussion level and considering also other similar works, it seems that several nice methods presented in the literature by academic projects did not have a substantial impact on the day-by-day work of CH professionals or that this impact is very slow to emerge. We could draft several hypotheses.

The first one could be the relatively *slow* process of *technology endorsement* in the CH domain. Let me consider the example of 3D digitization: the basic technology was introduced in 1995-2000 and has improved since then. But the effective adoption and use of 3D digitization technology by CH professionals became a concrete reality only very recently, at least 15 years after technology appearance. It could hold the same for sophisticated visualization systems, such as those presented here and proposed in the literature in 2005-2020. Should we patiently wait the year 2030?

Let us move to another debate, concerning *complex/sophisticated methods vs. easy/straightforward approaches*. We could consider just a subset of the CH&ICT related literature: the papers authored by influential users (e.g., museum curators who implemented digital content for their expositions, teachers developing digital resources for their courses, art historians or archaeologists reporting use of digital media in their work) but NOT co-authored by academics. Those papers could be considered a report of what the CH community is doing autonomously and independently. Unfortunately, this literature subset mostly describes the adoption of simple or basic visualization approaches. An example of the tools reported is Sketchfab, a success story due to the simple approach provided for publishing 3D models on the web and making them available to colleagues or the public. Many of the latter papers report the selection of Sketchfab as the chosen viz platform. Let me move to another example concerning a more flexible visualization platform, 3DHOP. 3DHOP was designed to support: (a) easy and standardized viz tasks with a basic viewer, analogous to the Sketchfab approach; and (b) more complex communication and presentation functions, enabled by its configurable and extendable design. The analysis of what external users have done with 3DHOP [31] shows that many have just used the basic viewer. CH users very often ignore the potential of the configurability and the extended features.

These examples and our experience with CH professionals lead us to think that the technology adoption process is very slow and that *easy and basic solutions win*. Moreover, one could believe that most of the more sophisticated approaches proposed in the literature are just the result of technology geeks. An open query is if the real CH user community does require that level of sophistication. Thus, concerning the issue of who is driving innovation, we are back to the usual technology- or user-driven debate.

We still need a strong effort in *disseminating* innovative technologies and sophisticated solutions to the community of users. Since the CH domain is still characterized mainly by personnel having a human-science background, we need intense training and consulting. A related problem is the perduring *hiring policy* in museums and CH institutions. Since digital communication cannot be implemented on the shoulders of just art historians and curators, we need real multidisciplinary teams established in CH institutions.

Finally, the role of *private companies* is critical in any digital domain. In many CH experiences, open-source resources and academic prototypes are common solutions, while in other application domains, this role is taken by commercial software tools. This is a weakness for the CH domain, justified by the limited financial resources dedicated to digitization and digital communication, which is not attracting enough the private sector. The CH domain's economic growth could be beneficial for attracting companies who could deliver more sophisticated and easier to use software solutions.

To summarizing, we envision some *actions* which might have an impact on some of the issues mentioned:

- CH institutions need financial resources to enable a concrete and tangible jump in the digital era, but these resources should be bound for a significant share to *permanent hiring* of staff with ICT background, and only the rest should finance projects;
- The availability of funds dedicated to digitization and creation of digital content (for example, as part of the EU Recovery Plan) should include the CH domain and could be crucial for implementing the point above and for attracting private companies towards the CH domain;
- Academia should still pair research efforts with dissemination and training actions directed to CH professionals.

7 References

1. Scopigno R., Callieri M., Cignoni P., Corsini M., Dellepiane M., Ponchio F., Ranzuglia G., 3D models for Cultural Heritage: beyond plain visualization, *IEEE Computer*, Volume 44, Number 7, pp. 48-55 (2011). http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=5958706
2. Adamopoulos E., Rinaudo F., 3D interpretation and fusion of multidisciplinary data for heritage science: a review, Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W15, pp.17-24 (2019).
3. Reichinger A., Maierhoffer S., Purgathofer W., High-quality tactile paintings, Journal on Computing and Cultural Heritage (JOCCH) 4, 2 (2011).

4. Bekele M., Pierdicca R., Frontoni E., Malinverni E.S., Gain J.E., A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage, ACM Jour. on Computing and Cultural Heritage, March 2018, Article No.: 7 (2018) https://doi.org/10.1145/3145534
5. Rodríguez-Gonzálvez, P. and Muñoz-Nieto, A. L. and del Pozo, S. and Sanchez-Aparicio, L. J. and Gonzalez-Aguilera, D. and Micoli, L. and Gonizzi Barsanti, S. and Guidi, G. and Mills, J. and Fieber, K. and Haynes, I. and Hejmanowska, B., 4D reconstructionnand visualization of Cultural Heritage: analyzing our legacy through time, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. Vol. XLII-2/W3, pp.609--616 (2017).
6. Remondino F., Rizzi A., Reality-based 3D documentation of natural and culturalheritage sites—techniques, problems, and examples, Appl Geomat, 2:85–100, (2010).
7. Hussain A.J., Al-Fayadh A., Radi N., Image compression techniques: A survey in lossless and lossy algorithms, Neurocomputing, Volume 300, Pages 44-69 (2018). ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2018.02.094.
8. Google Maps (http://maps.google.com) and Google Earth (https://www.google.it/intl/it/earth/)
9. Potdar V. M., Han S. and Chang E., "A survey of digital image watermarking techniques," *INDIN '05. 2005 3rd IEEE International Conference on Industrial Informatics,* Perth, WA, Australia, pp. 709-716 (2005). doi: 10.1109/INDIN.2005.1560462.
10. Malzbender T., Gelb D., Wolters H., Polynomial texture maps. In: Proc. of the 28th ACM SIGGRAPH Conference, pp. 519–528. ACM (2001).
11. Palma G., Corsini M., Cignoni P., Scopigno R., Mudge M., Dynamic Shading Enhancement for Reflectance Transformation Imaging, ACM Journ. on Computers and Cultural heritag, Volume 3, Number 2 (2010).
12. Sketchfab, a 3D web viewer (https://sketchfab.com/)
13. Palma G., Siotto E., Proesmans M., Baldassarri M., Baracchini C., Batino S., Scopigno R., Telling the story of ancient coins by means of interactive RTI images visualization, CAA 2012 Conference Proceeding, page 177—185 (2012).
14. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., Ranzuglia, G.: MeshLab: an open-source mesh processing tool. In: Sixth Eurographics Italian Chapter Conference, pp. 129–136, Eurographics (2008).
15. Potenziani, M., Callieri, M., Dellepiane, M., Ponchio, F., Scopigno, R.: 3DHOP: 3D Heritage Online Presenter. Computer & Graphics 51(Nov. 2015), 129-141 (2015).
16. Ponchio F., Potenziani M., Dellepiane M., Callieri M., Scopigno R., ARIADNE Visual Media Service: Easy Web Publishing of Advanced Visual Media, CAA2015, Volume 1, page 433-442 (2016) (tool available at: https://visual.ariadne-infrastructure.eu/)
17. YouTube, video sharing platform (https://www.youtube.com/)
18. Wang, Zeyu et al. CHER-Ob: A Tool for Shared Analysis and Video Dissemination." ACM Journal on Computing and Cultural Heritage, vol 11, issue 4 (2018).
19. Dercks U., Ponchio F., Scopigno R., CENOBIUM 10 years after: an evolving platform for Digital Humanities, Archeologia e Calcolatori, Volume 2018, Number 10, page 123-141 (2018).
20. Snavely N., Seitz S.M., and Szeliski R., Photo tourism: exploring photo collections in 3D. ACM Trans. Graph., vol. 25, pp. 835–846 (2006).
21. Remondino F., Spera M.G., Nocerino E., Menna F., Nex F., State of the art in high density image matching, Photogramm Rec., 29 (146), pp. 144-166 (2014).
22. Aicardi I., Chiabrando F., Lingua A.M., Noardo F., Recent trends in cultural heritage 3D survey: the photogrammetric computer vision approach, Journal of Cultural Heritage, Vol. 32, Pages 257-266 (2018), ISSN 1296-2074.

23. Brivio P., Benedetti L., Tarini M., Ponchio F., Cignoni P., Scopigno R., PhotoCloud: interactive remote exploration of large 2D-3D datasets. IEEE Computer Graphics and Applications, Vol. 33(2), pp. 86-96 (2013).
24. Brivio P., Tarini M., Cignoni P., Browsing Large Image Datasets through Voronoi Diagrams, IEEE Transaction on Visualization and Computer Graphics (Proceedings of Visualization 2010), Volume 16, Number 10, page 1261-1270 (2010).
25. Leoni C., Callieri M., Dellepiane M.,O Donnell D., Rosselli Del Turco R., Scopigno R., The Dream and the Cross: A 3D Scanning Project to Bring 3D Content in a Digital Edition, ACM Journ. on Computing and Cultural heritage, Volume 8, Number 1, art.5 (2015).
26. Rüther, Heinz, et al., From point cloud to textured model, the Zamani laser scanning pipeline in Heritage documentation, South African Journal of Geomatics 1.1: 44-59 (2012).
27. Koeva, M.; Luleva, M.; Maldjanski, P. Integrating Spherical Panoramas and Maps for Visualization of Cultural Heritage Objects Using Virtual Reality Technology. *Sensors* **2017**, *17*, 829.
28. Albertin F., Ruberto C., Cucci C., Callieri M., Potenziani M., Siotto E., Pingi P., Scopigno R., Bettuzzi M., Brancaccio R., Morigi M.P., Castelli L., Taccetti F., Picollo M., Stefani L., de Vita F., Beyond the analyses of the "Ecce Homo" by Antonello da Messina: from non-invasive investigations to dissemination, submitted paper, 2020.
29. Apollonio F.I., Basilissi V., Callieri M., Dellepiane M., Gaiani M., Ponchio F., Rizzo F., Rubino A., Scopigno R., Sobrà G., A 3D-centered information system for the documentation of a complex restoration intervention, Journal of Cultural Heritage, Volume 29, page 89-99 (2017).
30. Messaudi T., Veron P., Halin G., De Luca L., An ontological model for the reality-based 3D annotation of heritage building conservation state - Journal of Cultural Heritage - Vol. 29, p.100-112 (2018).
31. Potenziani M., Callieri M., Scopigno R., Developing and maintaining a web 3D viewer for the CH community: an evaluation of the 3DHOP framework, 16th Eurographics Workshop on Graphics and Cultural Heritage, page 169-178 (2018).