A DICOM-inspired metadata architecture for managing multimodal acquisitions in Cultural Heritage

Irina-Mihaela Ciortan[1], Ruggero Pintus[2], Giacomo Marchioro[1], Claudia Daffara[1], Enrico Gobbetti[2], and Andrea Giachetti[1]

[1]University of Verona, Department of Computer Science, Italy
[2]CRS4, Visual Computing Group, Italy
irinamihaela.ciortan@univr.it,ruggero@crs4.it,
giacomo.marchioro@univr.it,claudia.daffara@univr.it,
gobbetti@crs4.it,andrea.giachetti@univr.it

Abstract. Quantitative and qualitative analyses of cultural heritage (CH) assets need to interconnect individual pieces of information, including a variety of multimodal acquisitions, to form a holistic compounded view of studied objects. The need for joint acquisition brings with it the requirement for defining a protocol to store, structure and support the interoperability of the multisource data. In our work, we are performing multiple imaging studies in order to analyze the material, to monitor the behavior and to diagnose the status of CH objects. In particular, we employ, in addition to coarse 3D scanning, two high-resolution surface data capture techniques: reflectance transformation imaging and microprofilometry. Given this multivariate input, we have defined a hierarchical data organization, similar to the one used in the medical field by the Digital Imaging and Communications in Medicine (DICOM) protocol, that supports pre-alignment of local patches with respect to a global model. Furthermore, we have developed two supporting tools for multimodal data handling: one for metadata annotation and another one for image registration. In this work, we illustrate our approach and discuss its practical application in a case study on a real CH object – a bronze bas-relief.

Keywords: Metadata, Annotation tools, 3D scanning, Microprofilometry, Reflectance Transformation Imaging

1 Introduction

The importance of exploiting multimodal 3D capture techniques for artwork documentation is widely recognized[1], [2]. Since multiple measurements, often taken with different instruments or at different times, need to be studied together, there is a need for managing all measurements and annotations in a clear and organized structure that could ideally converge to a common standard [3]. This need is fulfilled by specific protocols and associated metadata, which allow one to “describe, identify, facilitate the access, usage and management of (digital) resources” [4]. Metadata are essential to record the full life-cycle of a CH asset [5], as well as any intermediary activities involved in generating a digital model from a physical object. In other words, metadata open the gate to interpretation, providing extra-meaning and a legend on how to read and connect raw scientific datasets. The need for order has convinced end-users and stakeholders of CH to add metadata systems to their projects. This resulted in an inflation of project-tailored metadata schemes [5], [6], [7], that achieve the ad-hoc purpose for which they were created, but make difficult the task of standardization. The basic idea behind our proposed data organization protocol is to follow the same approach used in DICOM (Digital Imaging and COmmunications in Medicine) [9], the standard used to manage medical imaging studies in hospitals. Our proposal only covers the model of the Information Object definitions used in DICOM and the data acquisition management, though the protocol defines all the aspects related to data communication and device interoperability, that is a really crucial aspect in the development of 3D model archives [8].

However, we believe that the development of a complete standard defining all these aspects could be extremely useful also in the CH domain. An attempt to adapt the complete DICOM standard to other domains has been done, for example, in the field of industrial material analysis with the proposal of DICONDE (Digital Imaging and Communication in Nondestructive Evaluation) [10], a standard for handling, sharing, storing and transmitting information between compliant systems.

This paper is structured as follows: in the upcoming section we describe how we adapted some ideas of the DICOM standard for the needs of the Cultural Heritage domain and how we partially included the Aging and Study levels – which are very appropriate to the rapidly changing nature of CH objects and their complexity in need of the multivariate analysis. Afterwards, we describe two of our tools developed to facilitate and catalyze two important steps in handling CH metadata and data: annotation and registration. Then, we go through our protocol step by step to examine a Case Study on a real CH object – a bronze bas-relief, copy of an Italian Renaissance work of art. It is noteworthy for the wide applicability of our proposed method to mention that we put it in practice within an European Horizon 2020 project, Scan4Reco [11], where we deal with a plethora of CH objects, including samples created in laboratory. Finally, our paper concludes with a discussion and ideas for potential future developments.

2 The proposed protocol for data annotation

Starting from the DICOM protocol, in our proposed metadata architecture, each object of study (artwork or material sample) is treated as a ”patient” and when it undergoes a study, it must be annotated registering its basic information analogously to real patients who get recorded in Radiology Information Systems (RIS). After each object (artwork or sample) is annotated, it is acquired by several imaging studies and relevant metadata is stored. The metadata are stored at the study level of our hierarchical organization and can be retrieved using the standard information model for Query/Retrieve that in DICOM is based on Patient (object), Study and Series – as shown in Fig. 1. In this hierarchy, we inserted an aging level in order to keep track of possibly different natural or artificial aging procedures performed on the objects. Moreover, for each acquisition technique, we have defined a set of specific metadata fields based on tag-value pairs for each of these three acquisition methods: RTI, Microprofilometer and Low-resolution 3D scanning. End-level data stored in our archives are then not necessarily standard images, but typically data (surfaces, clouds, grids) spatially referenced in a global coordinate system.

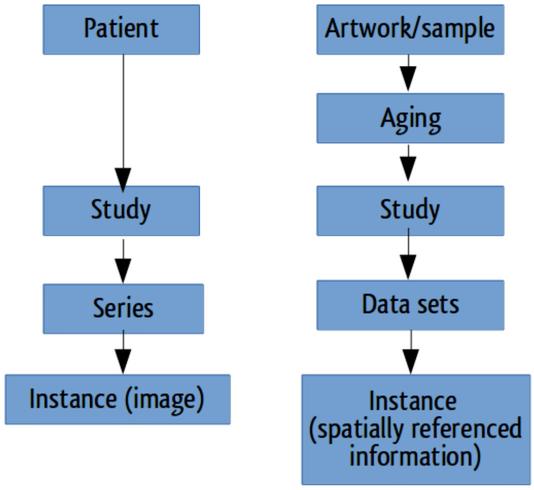


Fig. 1. Query/retrieve Information models for DICOM archives (left) and for our archive (right).

2.1 Object level

The Object metadata file comprises fields that uniquely identify and describe the origin of the CH asset, together with its physical characteristics, as well as the treatments applied (for example protective coatings) and the aging condition of the object at the time of its recording. The full fields and their corresponding descriptions are enumerated in Table 1.

2.2 Aging level

The purpose of the Aging metadata file is to monitor any aging process that reacted on the object. Therefore, it is of interest to record the type of aging

Table 1. Metadata fields for Cultural Heritage object

Field	Description
UID	Represents a unique combination of alphanumeric characters that references to a sample or an artwork in the project's database.
TYPE	The type of artifact: to be chosen between artwork or lab simulated mock-ups (samples).
LOCATION	Geographical place where the object was created.
NAME	The text that describes the name of the sample. It also identifies the sample/artwork, but not necessarily unique.
AUTHOR	The person who created the sample or the artwork.
SOURCE	Where the object comes from: museum collection, research laboratory, cultural institution, etc.
DATE_CREATION	The date when the physical object (sample or artwork) was completed. Format: DD/MM/YY.
DETAILS_CREATION	This might refer on how the sample was created, the history, chronology or steps of creation.
EXTENT	Refers to the physical dimensions of the sample or artwork (height, width, thickness).
SEMANTICS_DESCRIPTION	The description of semantics has to briefly guide through the content of the sample or artwork. Example 1: the artwork is an Icon depicting Virgin Mary.
MATERIALS	The constitutive elements that make up the sample/artwork and their type: metals, pigments, support, etc.
CONDITION	The status of the sample/artwork that might regard the following characteristics: novelty, previous restoration,visible aging effects. Example 1: new, out of the laboratory sample, with no aging effect. Example 2: painting was partially restored, but there is still visible a red pigment discoloration.
TREATMENT	Restoration method that an artwork has undergone or the chemical treatment (such as protective coating) applied to a sample at the moment of creation.

(natural or artificial), the major factors and agents that lead to the aging effects and the attributes of the object that were mainly affected as a result of the aging process, as listed in Table 2.

Table 2. Metadata fields for the aging process

Field	Description
UID	Unique identifier of the aging process, that can be made up of a unique combination of alphanumeric characters, which ought to be representative of the aging method applied and the time frame. Example 1: UV Exposure t1.
TYPE	The category of aging process: artificial (human-provoked) or naturally (due to the passing of time, without human intervention).
DATE	The date when the aging treatment was applied. Format: DD/MM/YY.
LOCATION	Geographical place where the aging process has taken place.
METHOD_NAME	The name of the aging mechanism involved.
SCIENTIST_IN_CHARGE	The person in charge of the conducting or supervising the aging process.
DESCRIPTION	Description of the aging method.
AGENTS	The bio-chemical agents responsible for the aging.
EXPECTED_EFFECTS	The theoretical expected change in the appearance, structure and geometry of the physical object onto which the aging effect was applied.
ATTRIBUTES_AFFECTED	The intrinsic properties of the object that were affected by the aging process.

2.3 Study level

Whereas the aging level offers a preliminary versioning of the Cultural Heritage object, the Study level represents the supporting data for each point on the time axis that the object is passing through. The study level can be split into as many acquisition techniques as are implemented (in our case, we used three techniques). Even though each acquisition has its characteristic method, we created groups of metadata fields as a way of coping with the immediate divergence of the techniques. Therefore, the Acquisition metadata fields in our model are further grouped into several wrappers: Study identification, Setup specifications, Hardware specifications, Software specifications, Output files and Spatial reference.

For brevity, in this text we only include the full set of metadata fields for the 3D low resolution imaging system (Table 3). The tables corresponding to the other two techniques (RTI and Microprofilometry) can be fully viewed online on our project's website, Scan4Reco [11], in the document on the Metadata and Joint acquisition protocol published as a deliverable of the project [12].

Similarly to the 3D low-resolution study described in Table 3, RTI groups are populated by Identification (Acquisition ID, Date, Type, Location Name, Operator Name), Setup specifications (Type, Description), Hardware specifications (the optical characterization of the Digital Camera and other optical accessories: lens, filter, lights, etc), Software specifications (Camera settings for capture: ISO, aperture, focus, the use of tethering tools), Output files (raw data, derived data and corresponding formats) and Spatial referencing (Camera axis direction and center position; intrinsic parameters). We have defined fields for the microprofilometric acquisition analogously [12].

Therefore, the metadata files for all the acquisition methods manage to both preserve the specificity of each technique and, at the same time, to maintain inter-connectivity between the various methods by exploring a standardized, grouped way to cover all the essential information in similar groups.

Table 3. Metadata fields for Low-resolution 3D scanning

Field	Description
ACQUISITION_ID	Unique ID of the acquisition
ACQUISITION_DATE	Date and time when the acquisition was performed.
ACQUISITION_TYPE	The type of study applied to the object, such as: RTI, Microprofilometry, Low-res scanning
LOCATION_NAME	Geographical location of where the acquisition was performed. Example: Verona, Sardinia
OPERATOR_ID	The unique identifier of the operator who performed the acquisition.
OPERATOR_NAME	The name of the person who conducted the acquisition.
SETUP_DESCRIPTION	Description of the particularities of the setup.
SCANNING_DEVICE	The scanning device used for capturing the object and the cloud of points corresponding to the object's geometry.
RESOLUTION	The resolution of the scanning device.
ACCURACY_VALUE	A numeric value that indicates the accuracy of the 3D scanning procedure.
TOOL_VERSION	A numeric value that represents the tool's version used for scanning.
OUTPUT_FILE	3D scan file name
OUTPUT_FILE_FORMAT	Mesh/Point cloud format
ORIGIN	XYZ coordinates of the origin in the reference space
ORIENTATION	Unit vectors of the acquisition space in the reference space

3 Metadata recording tools

When storing the metadata and using the actual data, there are two main challenges to address. In the first case, the metadata fields need to be annotated in an error-proof manner; in the latter, the multimodal data need to be fetched into the same coordinate system. We are addressing these challenges by implementing helper tools for both annotation and registration processes.

3.1 Metadata annotation interface

In order to facilitate the annotation of the metadata fields, a form-based GUI for the generation of the metadata files has been developed using the Qt framework (Figure 2). The tool lists the fields that need to be filled-in by the data operator. Moreover, to reduce the possibility of error, where possible we are including predefined answers into drop-down boxes, check boxes or radio buttons. Further, the fields include tooltips that provide additional clarification on their meaning. In addition, in case the mandatory fields are not filled-in, the metadata files cannot be saved and the mandatory fields are flagged with a change in the background color. The tool generates the metadata as simple text files and allows both export and import (in case the user wants to verify or modify an existing metadata file). Although new tabs dedicated to generate the metadata of the acquisition level can be added to the graphical user interface in a similar fashion, this tool has been developed for the specific purpose of annotating the Artifact and Aging metadata, following that the Study level can be supported by the acquisition software.

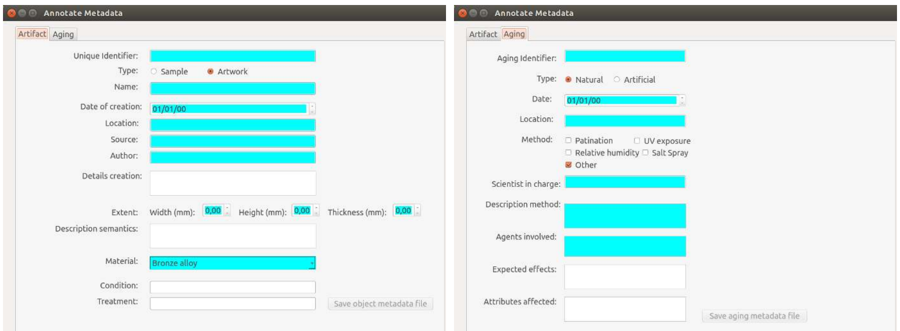


Fig. 2. The interface of the metadata annotation tool. The tool has two tabs: Artifact and Aging. The mandatory fields are emphasized with a cyan background and the save buttons are inactive, since the mandatory fields are not yet inserted.

3.2 Multimodal Data Registration tool

The RTI and the micro-profilometry data are acquired separately and in distinct times. RTI outputs an image stack, where each 2D image corresponds to a different light position, captured with a fixed camera viewpoint and fixed object scene. On the other hand, the microprofilometer provides 3D information in the form of a height map, by measuring the distance from the probe to the studied surface across an XY grid. So, we need to fuse those multi-modal signals into one common reference frame and keep the fusion information as manageable metadata for further processing. In this way, we can embed together not only geometry data from optical high-res profilometry, but also other geometric information (e.g., normal map) and appearance data (e.g., albedo) from the RTI image stack.

To exploit the fused signals, they need first to be registered. The goal of the procedure that registers the RTI and micro-profilometry data is to compute the mapping (position and orientation) between the 2D domain of the RTI camera sensor and the 3D surface (reference depth map) from the micro-profilometer. The input information for the registration is the micro-profilometry depth map and some data/metadata from RTI – e.g., intrinsic parameters of the camera sensor and/or only one image with the same view point as the RTI data. Since the RTI image stack is captured from a fixed viewpoint and varying lighting conditions, it is sufficient to align just one image to the 3D geometry in order to register all the information in the stack onto the 3D geometry. The image used for the registration could be one of the images in the stack, or, more likely, an image computed by an image stack processing routine (e.g., the normal map field).

A graphical user interface has been developed to help the user with the registration (see Sec. 4 and Fig. 5). It shows the user with two images: one from the RTI (Fig. 5, normal map on the right) and the other from micro-profilometry data (Fig. 5, geometry on the left). Since the intrinsic camera parameters have been already computed for the RTI data in a previous calibration step, the user only needs to select a small set of 2D-3D correspondences (at least three, but likely more for a better, more robust first estimation). Although out of the scope of this paper, an automatic refinement step could be added at the end of the registration pipeline, which might be based on an ICP-like (Iterative Closest Point) optimization or on the possibly available signals that allow for Mutual Information computation.

Finally, the exported registration metadata includes the list of correspondences (i.e., pairs of 2D and 3D positions), the pointers to the geometry and RTI files (i.e., local or remote location of the files), the extrinsic camera parameters (i.e., the view matrix), and the intrinsic camera parameters (i.e., the projective matrix, the distortion coefficients, and the image resolution).

Of course, if the sampling resolution is similar to that of the RTI, the same procedure and metadata can be used to perform and document the registration between RTI data and the geometry acquired with a low resolution 3D scanner.

This example shows how, in addition to providing ways to record registered study information, our proposed metadata architecture is flexible enough to also store information related to the registration process. For instance, our approach allows one to track errors or to improve registration starting from the same data by repeating processing steps using improved/alternative software on the same registration information.

4 Case study: bronze bas relief

To better illustrate our proposed protocol and the supporting tools, we have chosen a real Cultural Heritage example as a case study. The object consists of a bronze (Cu90-Sn10 alloy) bas-relief, created in 2004 for educational purposes and that was loaned for educational purposes by the Opificio delle Pietre Dure in Florence, Italy. The bas-relief – shown in Figure 3 – is a copy of one of the bronze panels of the Paradise Door made by Lorenzo Ghiberti for the Baptistery of Florence. The dimensions of the bas-relief are 39 x 12 x 2 cm (width x height x thickness). The surface underwent an artificial patination through the application of iron(III) chloride, giving the surface a brownish appearance. In 2007 it was coated with a protective product and exposed outdoors in an urban environment (central Florence) until 2016. In the following subsections, we will show how we generate the metadata files from this object and how we obtain images referenced in the same coordinate system from multimodal image data.

Fig. 3. The CH object used for our case study: a bronze bas-relief depicting a female figure.

4.1 Filling in the Metadata fields

To generate the metadata files corresponding to the bas-relief and the details on its aging we used the metadata annotation tool presented in Section 3. The user-friendly interface of the tool after correct data entry is shown in Figure 4.

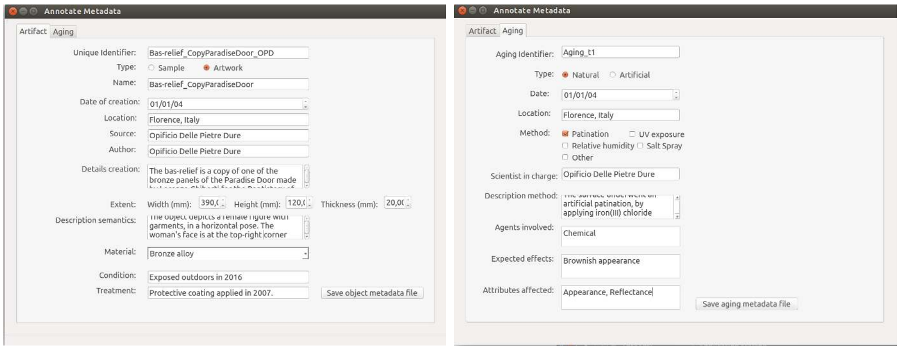


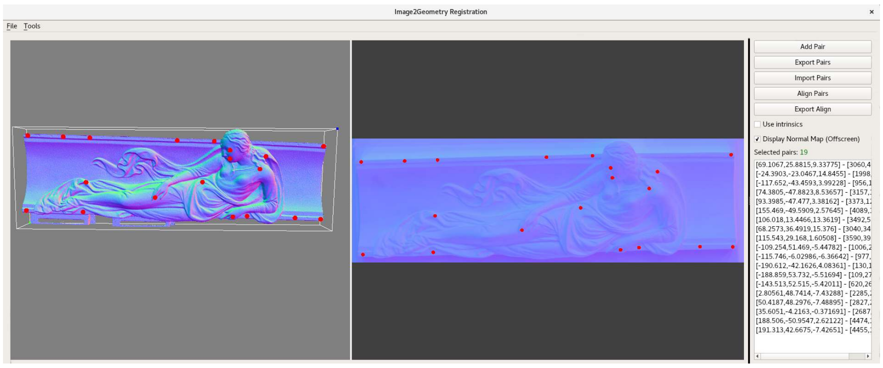
Fig. 4. The interface of the metadata annotation tool, with all the (mandatory) fields filled in. The save buttons are now active.

4.2 Demonstration of multimodal data registration

Another type of metadata are those related to multi-modal registration, which are automatically computed and exported by an alignment routine provided with minimal user input. Figure 5(a) shows an example of how to use the registration tool to select correspondences between the 2D domain of the high resolution normal map, computed from the RTI image stack, and the 3D geometry. In this case, the user has selected 19 pairs and, although the intrinsic camera parameters are typically available from a pre-calibration step, here we show how the tool can calibrate both the camera matrix and the distortion coefficients together with the extrinsic parameters. Figure 5(c) and 5(b) respectively show the final registered geometry from micro-profilometer and the normal map from RTI. The user performed a similar procedure to register the low resolution geometry from the 3D scanner (Fig. 5(d)) to the high resolution micro-profilometry data; in this case an open source software (Meshlab [13]) was used to register the two surfaces. These three figures demonstrate how the three multi-modal signals perfectly overlap.

5 Discussion

In this paper, we have proposed a DICOM-inspired metadata protocol designed for Cultural Heritage objects. This protocol features inclusion of aging processes and is suitable for annotating multi-modal data. We showed how metadata annotation is crucial for further data reaching a meaningful and holistic analysis of distinct studies that are performed on a CH object. The peculiar aspect of the CH acquisitions is that they are typically done both in 2D and 3D on complete objects or small regions. It is therefore mandatory to record all the information required to align the captured information with a reference 3D frame in order to allow for jointly processing the data. Acquisition devices should store all the metadata useful for the alignment and specific tools for information mapping should be coupled with the them. We implemented this approach in the



(a) Registration tool

(b) Aligned normal map

(c) Aligned geometry (Micro-profilometer)

(d) Aligned geometry (Low-res Scanner)

Fig. 5. Image-to-geometry registration tool. The registration between 2D images and 3D geometries is assisted by a GUI (a) where the user can select correspondences between points in space and pixels in the image. In this case we use this tool to register the RTI image stack, by using the resulting normal map (b), and the geometry from the micro-profilometer (c); in addition we use Meshlab [13] to align high resolution data (c) with the surface acquired by the low resolution 3D scanner (d).

Scan4Reco project [11] performing on the same objects and at different aging stages the acquisition of high resolution 3D scans, low resolution 3D scans and multi-light image acquisition using custom setups and processing tools.

The case study shown in this paper – the analysis of a bronze bas-relief – recreated the complete path of our pipeline from object annotation to multimodal annotation and, finally, end-data registration.

Our current work can be further improved by adding support for other acquisition methods, integrating them in our protocol and the supporting tools. Another useful future development, that would however require a richer database of studied CH objects, would consist in creating an ontology that would simplify and allow the partial automation of the metadata annotation process.

Acknowledgments. This work was partially supported by the Scan4Reco project funded by European Union's Horizon 2020 Framework Programme for Research and Innovation under grant agreement no 665091.

References

1. Remondino, Fabio. "Heritage recording and 3D modeling with photogrammetry and 3D scanning." Remote Sensing 3, no. 6 (2011): 1104-1138.
2. Pamart, Anthony, Odile Guillon, J-M. Vallet, and L. De Luca. "Toward a multimodal photogrammetric acquisition and processing methodology for monitoring conservation and restoration studies." In Proceedings of the 14th Eurographics Workshop on Graphics and Cultural Heritage, pp. 207-210. Eurographics Association, 2016.
3. Ronzino, P., S. Hermon, and F. Niccolucci. "A metadata schema for cultural heritage documentation." V., Capellini (ed.), Electronic Imaging & the Visual Arts: EVA (2012): 36-41.
4. Stasinopoulou, Thomais, Lina Bountouri, Constantia Kakali, Irene Lourdi, Christos Papatheodorou, Martin Doerr, and Manolis Gergatsoulis. "Ontology-based metadata integration in the cultural heritage domain." In ICADL, pp. 165-175. 2007.
5. D'Andrea, Andrea, and Kate Fernie. "CARARE 2.0: a metadata schema for 3D Cultural Objects." In Digital Heritage International Congress (DigitalHeritage), 2013, vol. 2, pp. 137-143. IEEE, 2013.
6. Doerr, Martin, and Maria Theodoridou. "CRMdig: A Generic Digital Provenance Model for Scientific Observation." In TaPP. 2011.
7. Doerr, Martin, Stefan Gradmann, Steffen Hennicke, Antoine Isaac, Carlo Meghini, and Herbert van de Sompel. "The europeana data model (edm)." In World Library and Information Congress: 76th IFLA general conference and assembly, pp. 10-15. 2010.
8. David Koller, Bernard Frischer, and Greg Humphreys. "Research challenges for digital archives of 3D cultural heritage models" Journal on Computing and Cultural Heritage (JOCCH):2 3 (2009), ACM.
9. Digital Imaging and Communication in Medicine. [Online]. Available: http://dicom.nema.org/. [Accessed: 10- March- 2017]
10. ASTM E2339-15, Standard Practice for Digital Imaging and Communication in Nondestructive Evaluation (DICONDE). [Online]. Available: https://www.astm.org/Standards/E2339.htm. [Accessed: 10- September- 2017]

11. Scan4Reco project - Multimodal Scanning of Cultural Heritage Assets for their multilayered digitization and preventive conservation via spatiotemporal 4D Reconstruction and 3D Printing. [Online]. Available: http://scan4reco.eu/scan4reco/. [Accessed: 12- September- 2017]
12. Scan4Reco project deliverable on metadata: Protocol for joint acquisition of Low-res 3D, Microprofilometry and RTI. [Online]. Available: http://scan4reco.eu/scan4reco/sites/default/files/documents/scan4reco_d5.2_final.pdf. [Accessed: 12- September- 2017]
13. Cignoni, Paolo, Marco Callieri, Massimiliano Corsini, Matteo Dellepiane, Fabio Ganovelli, and Guido Ranzuglia. "Meshlab: an open-source mesh processing tool." In Eurographics Italian Chapter Conference, vol. 2008, pp. 129-136. 2008.