Quick textured mesh generation in Cultural Heritage digitization
Gabriele Guidi1; Pablo Rodriguez Navarro2; Sara Gonizzi Barsanti1; Laura Loredana Micoli1; Michele Russo3
1Politecnico di Milano, Dept. of Mechanical Engineering, Milan, Italy; 2Universitat Politècnica de València, Department of Architectural Graphic Expression, Valencia, Spain; 3Politecnico di Milano, Dept. Design, Milan, Italy

1. Introduction
The goal of the 3D Icons project, funded under the European Commission’s ICT Policy Support Program which builds on the results of CARARE1 and 3D-COFORM2, is to provide Europeana with 3D models of architectural and archaeological monuments of remarkable cultural importance. The project will end in February 2015, bringing together 16 partners from across Europe (11 countries) with relevant expertise in 3D modeling and digitization. The main purpose of this project is to produce around 4000 accurate 3D models that have also to be generated in simplified form in order to be visualized on low-end personal computers and on the web. To reach this goal a suitable workflow of surveying and modeling has to be outlined.

Politecnico di Milano (POLIMI) is one of the partners of the 3D-ICONS project and its goal within the project is to digitize archaeological items belonging to the Civic Archaeological Museum of Milan (MAM).

Milan has been capital of the Western Roman Empire for more than one century (from 286 to 402 A.D.) and even if the modern city grew covering strata of historical remains, several signs of the past are still visible. The MAM museum has been built on the remains of the Roman circus, dated back to the IV century A.D., and of the city walls, two towers of which are still visible as well as some remains of Roman houses. The Museum holds more than 1000 archaeological objects related to different historical periods (Greek, Etruscan, Roman, Medieval) organized in thematic rooms for each historical period (Fig. 1) and grouped by artifact typology: epigraphs, statues, mosaics, furniture and pottery, with very important pieces as the glass cup called Diatreta (only one intact piece left in the world) and a silver plate called the Patera di Parabiago. The research group selected 527 items to be digitally acquired and modeled.

Fig.1 - One of the rooms of the Museum (Courtesy of MAM)
Starting from the need to produce this huge amount of models in a quite short period, the use of laser scanning was limited due to the following reasons: i) the artifacts material (marble, glass, bronze etc.) resulted less optically cooperative with laser than with digital photography; ii) the highly texturized surfaces of some archaeological objects may generate significant 3D distortions with triangulation laser scanners (i.e. the most suitable for small volumes); iii) the generation of a texturized mesh model has been demonstrated to be far more time consuming with an active device and texturing it with photos, rather than with SFM [Fassi et al., 2013, 73-80].

In this paper the solutions of problems occurred during the survey and modeling of the archaeological objects are described. Three case studies, each based on objects different in shape, material and measures are presented.

2. Data collection and data processing
For 3D digitizing the archaeological objects it was decided to employ the SfM technique. For image acquisition three digital reflex cameras were used: a Canon 5D Mark II, a Canon 60D and a Canon 20D. The Canon 5D features 22 megapixel with a full frame (36 x 24mm) CMOS sensor coupled with a 20 and 50 mm lens. The other two cameras, Canon 60D and 20D, are equipped with an APS-C (22.3 x 14.9mm) CMOS sensor, featuring 18 and 8.2 megapixel respectively. The lens used were a 50 mm on the 60D and a 20 mm on the 20D. The images were acquired at the highest level for each camera (5616 x 3744 pixels for the 5D, 5138 x 3456 for the 60D and 3504 x 2336 for the 20D). The distance from which the images were taken ranges from 0.5 to 3 m due to the objects position within the Museum (e.g. some of them were partially occluded by the walls and it was not possible to move them). The images were taken maintaining around 60% of overlapping between adjacent shots.

The rooms illumination is based on spotlights pointed directly on the artifacts; the ground floor is illuminated by some skylights and a big glass wall closing the room on one side. The light illuminating the displayed objects is therefore a superposition of natural light entering from the windows and artificial lights pointed over the archaeological artifacts, giving significant changes in the resulting color temperature. Therefore this environment highly influenced the choices during the survey such as for example: i) where to put the photographic set ii) if using or not panels; iii) how to capture objects with faces occluded by the walls.

At least a couple of images with length references where acquired in each photo set in order to scale the 3D measurements to metric values. Most of the time a few coded targets were used on the scene, and all target-to-target distances measured with a metallic tape meter.

The data processing was carried out with the Agisoft Photoscan package, a semi-automatic software in which both the camera orientation and the internal calibration are made automatically, allowing a little interaction to the user. The software implements image orientation and mesh generation through SfM and dense multi-view stereo-matching algorithms. Although the software acts in automatic way, some choices can be done during image orientation, where the operator may set alignment accuracy level and possible control points.
After the image orientation, in the mesh generation other parameters can be set as:

i) the “object type” that can be changed according to the type of object to be modeled. The object type “Arbitrary” defines a 3D free form [Cohen et al., 2012; Pollefeys et al., 2000, 619-626], the “Height Field”, namely a 2.5D surface like a DTM [Doneus et al., 2011, 81-88; Verhoeven et al., 2012, 2060-2070];

ii) the “geometry type” can be set with sharp or smooth function;

iii) the “target quality” specifies the desired mesh quality, higher quality settings can be used to obtain more detailed and accurate geometry, but require longer time for processing;

iv) in the “face count” it’s possible to indicate the maximum number of polygons of the final mesh;

v) the “filter threshold” specifies the maximum face count of small connected components to be removed after surface reconstruction as percentage of the total face count.

3. Case study one - differences between geometry types

The Museum held different kind of objects of different shape, measures and material. Some of them are located close to the walls, others are made of a reflective material, therefore several problems occurred during the survey. Since the final goal of the project is to produce an high number of models in a short period, one of the main factor was to identify the best pipeline in terms of a reasonable ratio between processing time and accuracy. Hence, several test were done with target quality high and medium and selecting different number of polygons. Regarding this last matter, it was seen that Agisoft produces models up to several million points, but in this case the models are too heavy for the final purpose of the project. That’s why it was decided to fix as maximum number 2 million polygons, that we estimated as sufficient to represent with appropriate level of detail all the selected object types.

The first test object is the roman stele of Geminus (A 0.9.6742) (Fig.2a), dated back to the end of the I A.D. Its measures are 174 x 61 x 31 cm., it’s made of limestone from Vicenza and is about the construction of a funeral monument that has been walled-in since 1860 in the arches of Porta Nuova in Milan [Tocchetti Pollini, 1990, AA.VV., 2007]. This object is suitable to verify the different meshing algorithm because on the one hand it has shadowed parts in the niche and the head, and, on the other hand, is has fine sculptured inscription.

The survey was done with 20 images around the object with a Canon 5D Mark II equipped with a 50mm lens, the distance from which the images were taken was 2.50 meter and the consequent GSD is 0.3 mm. The first processing was set up with geometry type sharp, target quality medium, face count 2 million and filter threshold 0.5. Then it was tested the target geometry type smooth with the same other parameters. The processing time between the two settings was quite the same but the results were much different: the sharp geometry type generated a 3D model accurate but full of topologic anomalies and holes that made the post processing of the final model long and complicated (Fig.2). The worst parts were the ones where the object was dark because
shadows make less recognizable the corresponding portion of image, leading to a less robust image matching and, as final result, more erratic 3D point estimates.

Configuring the processing with “geometry type” set to smooth, all these anomalies were hidden because the surface was numerically smoothed and the software interpolated the 3D points trying to reconstruct a complete model.

A comparison between the two models was carried on in Polyworks in order to understand the amount of differences and the presence of possible gaps: the deviation resulted included in the range ± 0.5 mm.

After this comparison it was stated that the deviation was not so significant even more so the target of the project regards only the visualization of the model on the internet. The sharp model is clearly better in terms of geometry and it’s much more clear-cut but the smooth model does not have neither holes nor topological anomalies, that simplify a lot the post processing, and the texture balances the geometrical differences between the two types.

4. Case study two - camera and lens differences

The object of the second case study is a sculpted group composed by Aphrodite and Eros with dolphin (A 0.9.1181-1182) dated half of the I-II century A.D. Both statues are in Greek marble, with anthropomorphic real size (96 and 44 cm height), characterized by a medium/high shape complexity [Camporini E., 1979, AA.VV., 2008]. The two statues are fixed, at a distance of about 15 cm between one and the other, on the same basement, positioned between a wall, 2 pilaster and other 2 unmovable statues (Figs. 3a, 3b). In an ideal condition the photo shoot scheme of an all-round object expected to rotate around the subject at different heights, taking images realizing parallel orbits, maintaining an uniform focus distance. In the shown situation it was impossible to use the ideal scheme, therefore, the position of the artifacts represent the most significant constrain of the survey.
At the first step we used a Canon 60D, with a APS-C sensor (14.8x22.2 mm) equipped with a 50 mm lens, equivalent to an 80 mm lens on a full-frame camera. The captured area at 1 meter is of 28x18,5 cm with an average GSD of 0.125 mm. Due to this configuration it was necessary to take 70 images to cover all the surface of the Aphrodite’s statue, keeping out Eros. Regarding the back, all the images were focused only on a portion of the object, so there aren’t images of the entire statue from these points of view Fig.4a). In the data processing phase, this has caused the failure of the alignment process due to the impossibility to find and adequate number of homologous points on the different images. For the second survey campaign we used a Canon 20D, characterized by the same type of sensor, equipped with a 20 mm lens, equivalent to an 32 mm lens on a full-frame camera and the captured area at 1 meter distance is about 70x46 cm, the average GSD is 0.25 mm. In this case it was possible to locate the camera around the object realizing images that focused the entire Aphrodite statue and, in same case, also the Eros one. The shooting schema it’s quite similar to the ideal one with the difference that in the back side the distance from the object is smaller than in the front side and to keep the whole surface we used a larger tilt angle (fig 4b). The processing of the acquired data in this case gave good results both in the alignment and in the meshing phase and we obtained the 3D textured models of both the statues with 2 million of polygons.

The negative feature we have to highlight is related to the illumination con-
dition. Since in this case it was impossible to move the object, so we had to shoot with ambient natural light coming only from the front and upside. The images of the side part of the Aphrodite model show the different illumination effects both in the textured (Fig.4c) and in the smoothness of the obtained surface (Fig.4d). The 3D model produced is acceptable for visual if shown with its texture on, while it is not satisfactory if showed just as shaded mesh, because all the asperity of the surface, due to the bad matching of poorly lit areas, are visible.

5. Case study three: materials and photographic layout differences
The objects of this case study are two sculpture of man’s portrait [Camporini E., 1979, Aa.Va., 2008]. the first one (A 0.9.1160) is a sculpture of the I century A.D., 27 cm high, realized in clear opaque marble, with fine decoration of hair and beard. The second one (A 0.9.1157) is a precious exemplar of III century A.D., 42,5 cm high, realized in bronze casted with lost wax and finishes by burin, chisel and polishing.

Fig.5 - Man’s sculpted portrait: a) in marble (ref. A 0.9.1160); b) in bronze (ref. A 0.9.1157)
Both the sculpture are movable and it was possible to place them in a free space, with the possibility to turn around and shoot exploiting the ideal schema (fig.6a).

Fig.6 - a) Alignment of the camera for the bronze portrait; b) The final texturized model of the marble portrait
Furthermore it was possible to use background panel to isolate the object from the contest. This mean that in the alignment phase we cannot use external references besides the subject, but if the subject has the adequate characterization of the surface this is acceptable and some advantages in the following steps of the process can be obtained. The benefit of this set up is that we can mask images, both during the alignment or the meshing phase, using automatic selection of the uniform color’s image portion (black or white) behind the main subject. The masked images are faster to be processed by the software, the obtained 3D model has fewer topological anomalies and the user’s work to refine the model is lower.

To scale the obtained models to the real size we equipped the survey set with few targets supplied by the Agisoft Photoscan software. Thanks to this, after the meshing, the automatic identification of the target allowed to impose the real distance measured during the survey.

For the marble head, it was used a black panel in order to have a good contrast between the color of the object and the background. Thanks to the material of the object and the possibility to move it from its place, the final model was really good (Fig.6b). For the bronze head it was decided to do the opposite. Since the color of the object was very dark and shadowy the background chosen was white.

![Fig.7 - Differences between the two models obtained by matching images taken with: a) white background panel, amesh distortion in correspondence of the nose is evident; b) black background panel, the mesh configuration is far more coherent with the true one; c) The final texturized model of the bronze head obtained with the use, during the survey, of a black panel](image)

However the resulting 3D model was non to satisfactory and the texture resulted very gloomy. Due to the high contrast between the dark head and the white background, the image exposure, being evaluated on the average of the framed area, tended to result too low. As a result the darker image areas containing the head (i.e. the most informative portion of the whole images) were not enough exposed to give a sufficient amount of data for the following image matching phase. For this reason some topological errors occurred on the model, for example on the nose (Fig.7a). So the survey was carried on placing a black panel behind the object. The problems of excessive light background reflection was disappeared and the color of the texture was more balanced.
and bright (Fig.7c). Both the survey were done using a Canon 5D Mark II with a 50 mm lens, mounted on a tripod, having the diaphragm fixed, a manual focusing and an automatic exposure with diaphragm priority, obtaining a GSD of 0.18 mm.

6. Summary of the experience and conclusions
As a project for the implementation of Europeana with 3D models, the 3D Icons project is permitting experiment a quick 3D acquisition and modeling approach based on the SfM technique, on different objects, situations and materials. Having the necessity to produce a high number of models in three years, it was essential to organize the work in a strict workflow that allowed to avoid time consuming operations. From this need derived the choice of the specific SfM implementation made in the Agisoft Photoscan software.

It seems to be a very good tool for generating good quality meshes from images in a semi-automatic way, giving the possibility to avoid manual selection of homologous points, as the traditional photogrammetry process require, but permitting an acceptable interaction with the user. Thanks to the presence of several different artifacts in the Museumit was possible to face different problems.

After all the tests done, some best practice were highlighted. Regarding the survey, the best choice was to move, where possible, the objects in a place where the lighting conditions were good enough to avoid flickering or huge contrasts among the different part of the objects, especially those with an articulated geometry or parts with deep decoration. In this case, the use of reflective panels helped in having light conditions homogeneous and as much as possible correct. Similar, the use of panels to avoid the background in the images, helping the following masking process, was really useful in clipping the processing time. During the matching phase, the choice of the parameters played an important role in avoiding time consuming situations. Defining the geometry type as smooth, all the post processing necessary to clean the model and to fill the mesh lacks were avoid, maintaining a good accuracy in the final model. It was also decided to limit the creation of polygons to 2 million, considered a good compromise between the time in the processing and the precision of the model, bearing in mind also the purpose of the project itself.

Acknowledgment
The authors desire to thank all the staff of the Civico Museo Archeologico di Milano, and in special way the director, Dr. Donatella Caporusso.

Notes
1 http://www.carare.eu
2 http://www.3d-coform.eu/

References
Doneus M., Verhoeven G., Fera M., Briese Ch., Kucera M., Neubauer W., 2011, From Deposit to Point Cloud - A Study Of Low-Cost Computer Vision approaches for the straightforward documentation of Archaeological Excavations, Geoinformatics CTU FCE, 81-88.