3D-COFORM – Tools and Expertise for 3D Collection Formation

André Stork, Dieter Fellner

TU Darmstadt, TU Graz, Fraunhofer IGD

Abstract

3D-COFORM has the overall aim to make 3D documentation the standard approach in cultural heritage institutions for collection formation and management. 3D-COFORM is addressing the whole life cycle of digital 3D objects (also called 3D documents) spanning the whole chain from acquisition to processing, and from semantic enrichment to modeling and high-quality presentation – all that on the basis of an integrated repository infrastructure. The paper will give an overview of 3D-COFORM and present its current results and contributions.

1. Introduction

Our heritage – in terms of artifacts – is represented by tens of millions of objects world-wide, all of them being three-dimensional (if we ignore the pure textual content in 3D objects like books etc). 3D-digitization is one (and currently the only) way to warrant long-term preservation of cultural heritage artifacts, making them accessible from almost everywhere and enabling the use of virtual 3D surrogates as a basis for performing (collaborative) research.

Even if digitization, computer graphics and geometric processing technology have made huge steps forward in recent years, cultural heritage items are still challenging in terms of geometric and optical (material) complexity. And, even more challenging, 3D objects scanned with the best technology available resulting in highest quality 3D models are still of moderate use if they cannot be searched and retrieved as conveniently as textual documents.

3D-COFORM (http://www.3d-coform.eu) has the overall aim to make 3D documentation the standard approach in cultural heritage institutions for collection formation and management. 3D-COFORM is currently the only Integrating Project at the European level in this field. 3D-COFORM is addressing the whole life cycle of digital 3D objects (also called 3D documents) spanning the whole chain from acquisition to processing, and from semantic enrichment to modeling and high-quality presentation – all that on the basis of an integrated repository infrastructure.

2. Overview

The main objective of 3D-COFORM is to make the use of 3D digital artifacts a practical and standard way of documenting physical cultural heritage.

For this end an integrated infrastructure is being developed. Research activities in each stage of the documentation chain progress the state-of-the-art in:

- 3D digitization of geometry and material properties
- processing of geometry and material data
- classification and indexing of 3D related and 3D information
- semantic enrichment through annotating
- content-based retrieval of shape and materials (searching)
- modeling of ancient sites through procedural and generic approaches
- presentation of high quality models along with their meta data
- and last but not least: cultural heritage repositories based on CIDOC-CRM

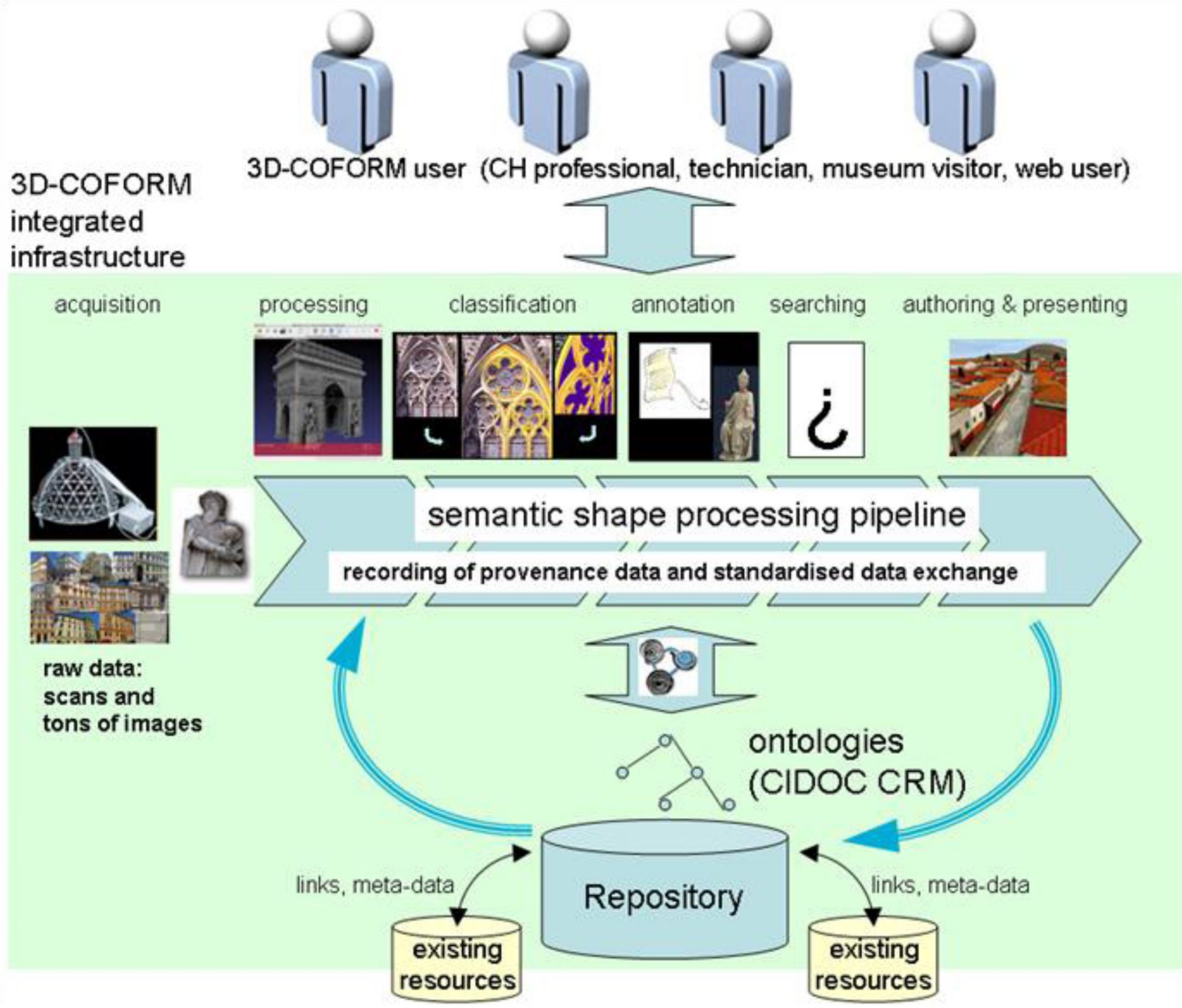

Figure 1: 3D-COFORM integrated infrastructure

The 3D-COFORM research program is accompanied with training and testing activities involving real users with real test cases. The enterprise activities are rounded-off with the analysis of business models for 3D digitization and documentation campaigns. Sustainability of 3D-COFORM results shall be achieved by establishing a Virtual Competence Center on 3D (VCC-3D).

In the remainder of the paper we highlight some of the current results of the 3D-COFORM project.

3. 3D acquisition techniques

3D-COFORM is developing a range of 3D acquisition technology to better and faster digitize shape and material of CH artifacts with the following characteristics:

- Immoveable Heritage
- Movable, regular objects: In-hand scanning
- Movable, regular objects: Dome-based acquisition
- Movable, optically complicated objects
- Reflectance Acquisition

Among these technologies are photometric stereo devices (Mini-Dome), domes to capture the geometry along with the BTF (Multiview Dome) of an artifact and scanners to measure the meso-structure of quasi-planar objects that will be introduced in the following subsections.

<u>Mini-Dome</u>

The Mini-Dome [TvG11] is a compact portable system (see Figure 2) that has only one camera and uses a whole collection of light sources on a sphere to calculate surface normals. Using a multigrid solver, these surface normals can be integrated giving a 3D reconstruction following a photometric stereo approach. Due to its simple mechanical structure, it easily packs into a suitcase, and can be brought to the object rather than the other way around. This is in contrast with the Multiview Dome, which consists of as many camera viewpoints as light sources and is described in the following section. Therefore the latter can rely on multi-view stereo reconstruction techniques, which is challenging nonetheless.

Figure 2: Mini-Dome setup

Generally, photometric stereo generates a normal field. Integrating over the normals, results in quasi 3D-models with somewhat limited accuracy. Within 3D-COFORM, advances in the precision of photometric stereo have been achieved. With the help of the mini-dome the following objects have been acquired. Figure 3 shows an oracle bone during a session at ShangDong, China. Objects varied from 5 to 20cm.

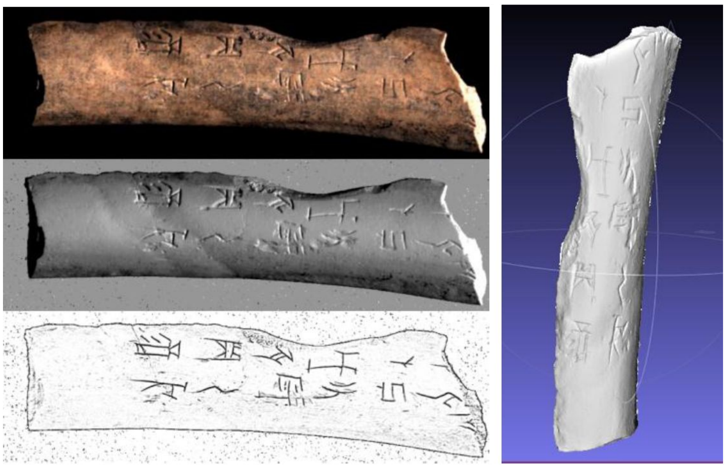

Figure 3: Jia Gu: oracle bone, ShangDong China.
Left: Picture, normal-map and line-view. Right: 3D model.

<u>Multiview Dome</u>

In contrast to the mini-dome, the multiview dome [WSRK11] takes pictures of a 3D artifact from 151 different viewing orientations. As with the mini-dome, the light directions vary for each shot.

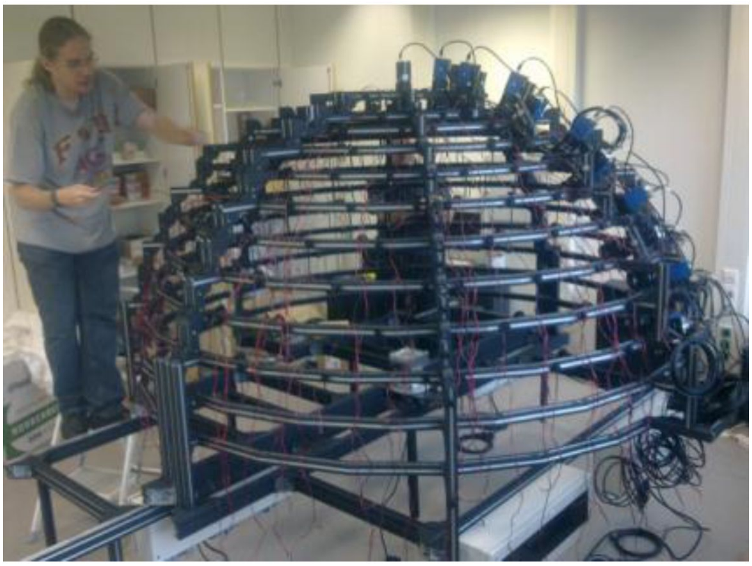

Figure 4: Multiview Dome at Bonn University

In combination with structured light fringe projection the multiview dome allows for accurate 3D reconstruction of shape and material. Given this accurate geometry, also great improvements of the quality of the material property estimation have been achieved (see Figure 5). This new acquisition pipeline for geometry and spatially varying surface materials [SWRK11] was tested on several different and challenging objects with remarkable results.

Figure 5: Achieved quality of an object acquired with the Multiview Dome object acquisition pipeline [SWRK11], employing a 3D reconstruction obtained with structured light [WSRK11] together with BTF reflectance data.

Figure 6: Test of faithfulness of the Multiview Dome acquisition. Left: Picture taken with calibrated camera and light-source. Right: Rendering of digitized object under the same view and light conditions.

<u>Meso-structure scanner</u>

Nowadays digital consumer cameras have 10 million pixels and more, whereas projectors used for fringe projection for 3D reconstruction from structured light are typically in the range of half a million to 2 million (absolute maximum).

This observation motivated the development of a new approach to improve depth resolution: an enhancement of the projector with an extra, high-precision optical lens shifter. Note, the lens shifter is added as a device in front of the projector's standard lens.

Using low-cost off-the-shelf hardware components – in this case a head attenuation motor from a hard disk – allows for shifting the stripe pattern with sub-pixel accuracy, actually it allows for 2000 steps per pixel. We measured accuracy and repeatability with a laser measurement device.

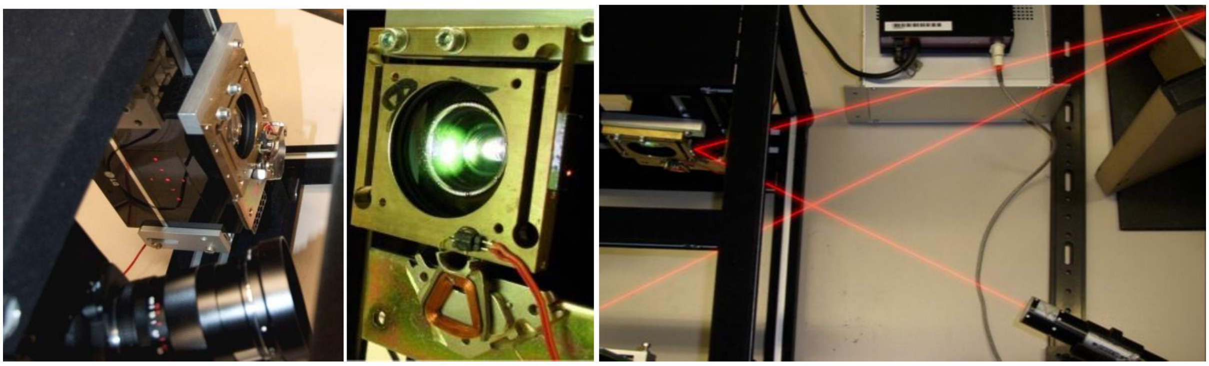

Figure 7: Meso-structure scanner

Analyzing depth resolution by means of two ‘measurement devices’, namely a ‘paper staircase’ and a precision-welded aluminum bar, showed a depth resolution of 20 microns which is 3 to 4 times higher than without using the lens shifting device. For further details see [RSGS10].

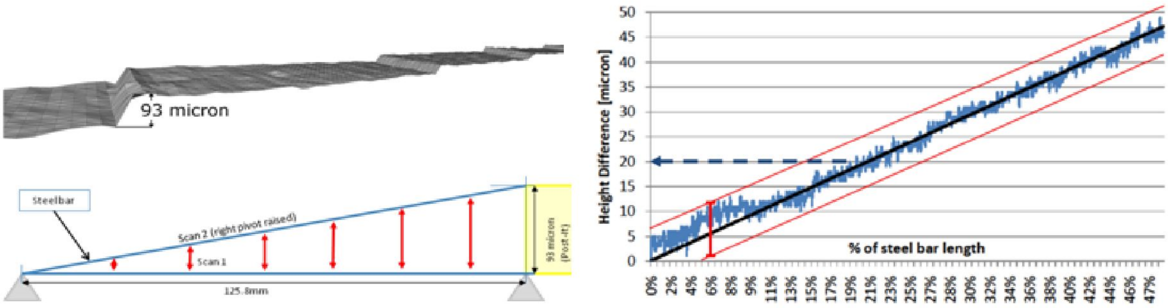

Figure 8: Accuracy measurement of meso-scanner

The following figure shows a number of 3D scans of ancient coins provided by the Victoria and Albert Museum, London.

Figure 9: Coins scanned with our meso-structure scanner

4. Data processing

Basically, all acquired raw data needs some processing. Examples for data processing are:

- 3D reconstruction from sets of digital 2D images
- Texture registration and mapping to 3D geometry
- Assembly of digitized fragmented artifacts
- etc.

Multi-view stereo

Multi-view stereo (MVS) is a technique that calculated 3D shape information out of a set of pictures of an object taken from different orientations. MVS is a well-known approach but has made large steps forward in the last couple of years, esp. stimulated by research on scale-invariant features (SIFT).

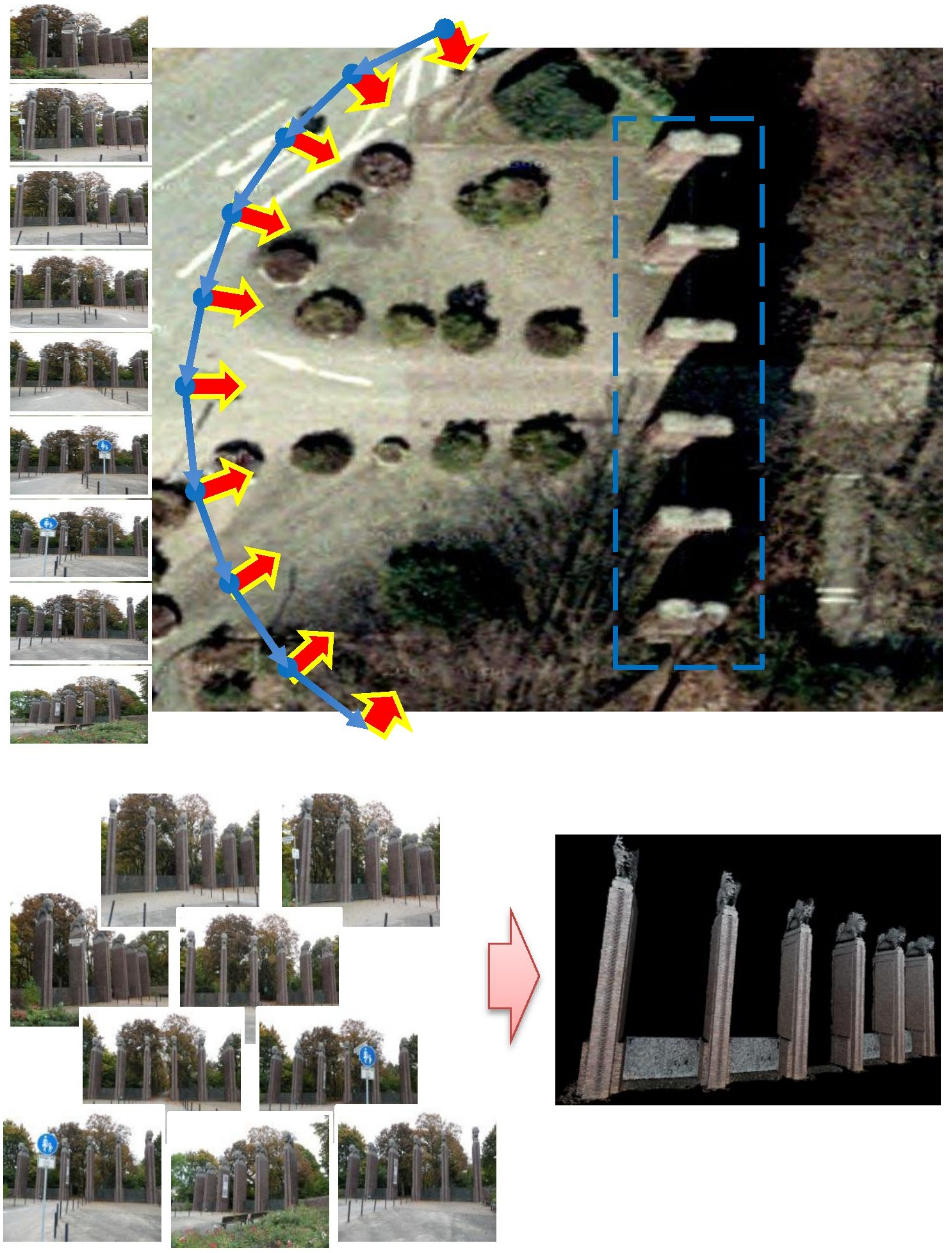

Figure 10: Multi-view stereo approach applied to the Löwentor, Darmstadt

Current research in this filed is addressing robustness of the reconstruction algorithms towards the input images with respect to lighting conditions, scale, obstacles in images, etc. and creating dense(r) 3D point clouds by incorporating photometric stereo (PS) [ARSG10].

Texture registration and mapping to 3D geometry

The projection of a photographic dataset on a 3D model is a robust and widely applicable way to acquire appearance information of an object. The first step of this procedure is the alignment of the images on the 3D model. While any reconstruction pipeline aims at avoiding misregistration by improving camera calibrations and geometry, in practice a perfect alignment cannot always be reached. Depending on the way multiple camera images are fused on the object surface, remaining misregistrations show up either as ghosting or as discontinuities at transitions from one camera view to another.

Based on the computation of Optical Flow between overlapping images, to correct the local misalignment by determining the necessary displacement, the symptoms of misregistration, instead of searching for a globally consistent mapping, which might not exist, can be corrected. The method scales up well with the size of the dataset (both photographic and geometric) and is quite independent of the characteristics of the 3D model (topology cleanliness, parameterization, density) [DMCCS11].

The method is robust and can handle real world cases that have different characteristics: low level geometric details and images that lack enough features for global optimization or manual methods. It can be applied to different mapping strategies, such as texture or per-vertex attribute encoding.

Figure 11: Example of color projection

Fragment re-assembler

Many heritage items only exist anymore in pieces and sometimes it is neither clear nor easy to assemble the puzzle of fragments together. Thus, automatic processes tend to fail. A combination of manual input and geometric processing, fitting and optimization algorithms is a more promising approach where the human cognitive capabilities are combined to steer fitting algorithms to put fragments together. The process is as follows:

- the user loads two fragments that she thinks once where one object (or part of one object)
- she specifies three corresponding points by graphical interaction

- the algorithm uses this information as starting point for an optimization process that tries to find a best fit taking the geometry (shape) of the faces the corresponding points belongs to into account

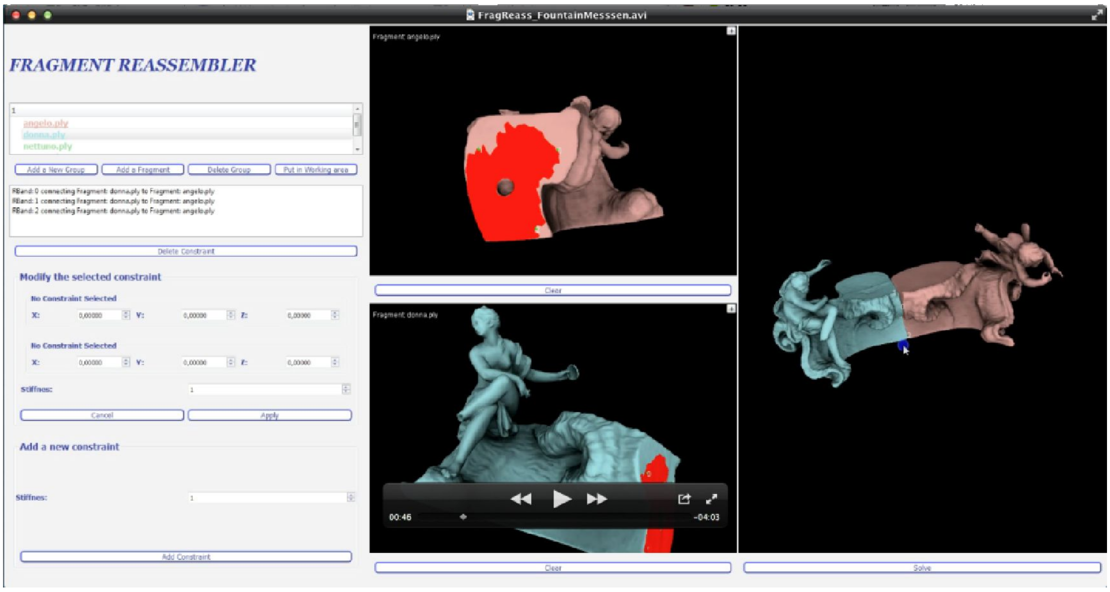

Figure 12: Fragment re-assembler

5. The Repository Infrastructure (RI)

Spending the effort to create visually pleasing digital 3D models of CH items is of limited benefit if they cannot be stored, retrieved, re-used, enriched, etc.

The 3D documentation process places specific requirements to a repository supporting these uses [PBHTGF10]:

- distributed architecture
- managing metadata and mass multi-media data in an integrated manner
- being based on standard CH ontologies, namely CIDOC-CRM
- support complex, content-based query mechanisms integrating metadata-based with shape-based retrieval
- support the continuous semantic enrichment (through annotations) of the knowledge represented in the repository

The structure of the 3D-COFORM repository is sketched in the following figure. Clients access the repository via Web Service interfaces. There is a central Metadata Repository (MR) and a centralized inventory of all ingested multimedia datasets, which themselves can be stored on distributed local resources (locations).

Queries can be issued using the SPARQL query language to the MR. Based on the results clients can download and display digital 3D models alongside with their metadata. Additional indexing services support shape-based and material-based retrieval of objects.

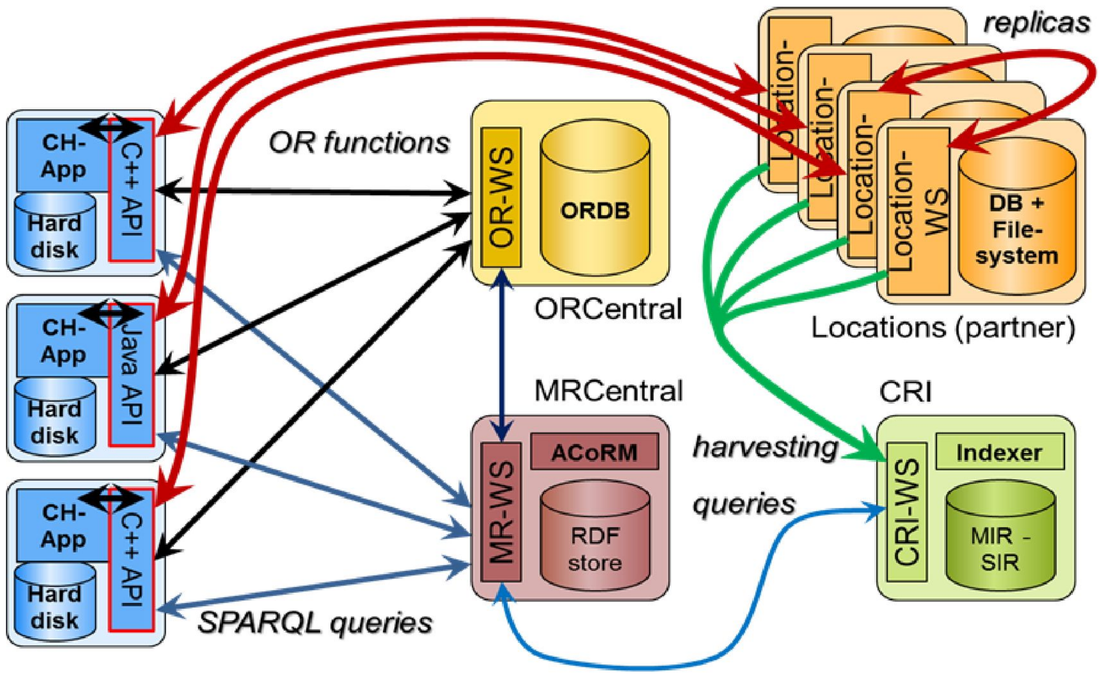

Figure 13: The RI architecture

Clients cannot only issue queries and download data but also ingest new data and thus enrich the content of the RI in two ways:

- add multi-media data (3D models, pictures, text files, etc.)
- add metadata and enrich the knowledge represented in the semantic net of the CIDOC-CRM ontology

6. Integrated Viewer Browser (IVB)

The purpose of the IVB is to search, browse, view and annotate collection items in the RI.

The search interface enables the user to specify combined queries. Combined queries can incorporate different concepts from the CIDOC ontology as well as shape and material information.

Different browsing modes allow the user to get an overview over the returned result set together with the metadata of each object found in the repository. Information displayed in the browsing part of the IVB can be used to refine the query. The interplay of search, browse and search again supports iterative query development.

Figure 14: The 3D-COFORM Integrated Viewer Browser
(left: search and browse; right: 3D viewing and annotating)

In the viewing component(s) of the IVB different media types can be displayed ranging from text over 2D digital pictures to 3D models.

A generic concept to define ***areas*** on any type of media, have been defined as a basis for generating ***annotations*** on objects. We distinguish between two types of annotation unary and n-ary ones. Unary annotations are used for commenting one object. N-ary annotations create links (relation-ships) between objects (which still can be of different type). Annotations are ingested into the RI thus enriching the knowledge represented in the repository. Annotations can be searched for, using the search component of the IVB. The semantic net represented in the RI, containing these annotations can be browsed from the browsing and viewing component. For creating/representing annotations we adopt semantic web standards, in this case TRiG-files [PSDTGCS11].

From within the IVB, objects can be segmented as well. Segments, e.g. an arm of a statue, can be used to drive a shape-based search. Segments are objects by their own right and are also stored in the RI together with metadata generated by the segmentation event.

7. Modeling CH artifacts and sites

Not all our heritage still existing in physical form, either completely or partially. In many cases only ruins or drawings exist that can serve as a basis for virtual reconstructions using 3D modeling tools and a manual process. 3D modeling is known to be a tedious process, thus research is being per-formed in order to decrease modeling time. In 3D-COFORM two approaches are being pursued:

- procedural / generative modeling (GML compositor)
- sketch-based modeling

<u>GML compositor</u>

The idea of generative modeling is generally to combine parametric building blocks, thus generating an increasingly complex model. The building boxes can be procedurally combined to generate repeti-tive elements that follow a certain structure which for man-made objects and architecture is often the case [HUF12].

In GML compositor the user has an easy-to-use interface to define, parameterize and arrange the building boxes and specify the rules they follow to build up a larger structure such as a building or parts of a building. The following figure is exemplifying this process. In the upper left corner, the UI of GML compositor is depicted, side by side with the final rendering result. In the bottom row, three snapshots of the generative process are shown. Starting with a plain façade, the arcades are added successively and decoration to the arcades finishes the model.

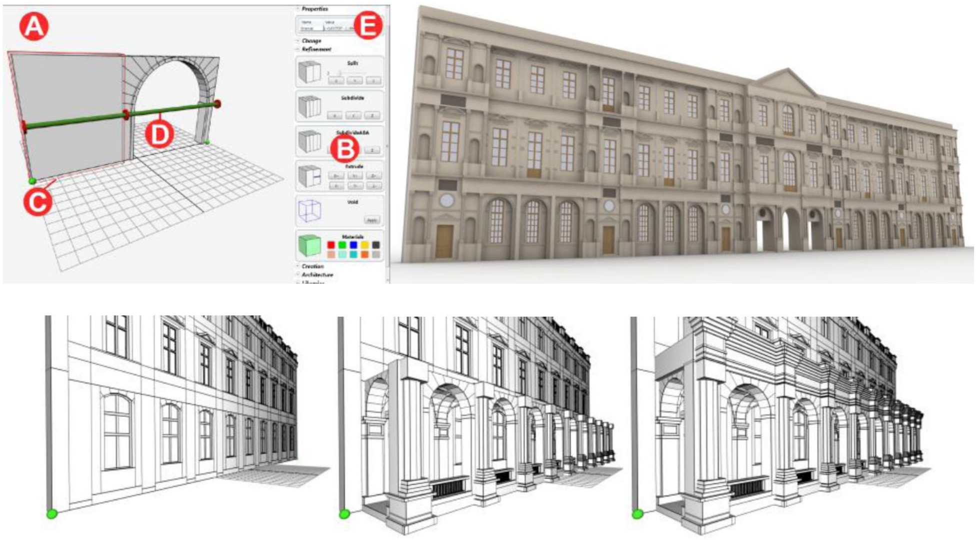

Figure 15: GML compositing (upper left: UI of tool; upper right: final result; second row: stages of facade modeling)

<u>Sketch-based modeling</u>

The purpose of sketch-based modeling focuses on different kinds of objects, namely more organically shaped ones and to create 3D hypothesis for none existing parts. With sketch-based modeling the user can create geometry on the basis of scanned parts to enrich them by possible/plausible add-ons and compare different alternatives rapidly. The sketch-based interface uses stroke based interaction and create smooth sub-division surfaces that can be represented compactly [BHSF09].

The following figure shows from left to right: a scan of a vase with missing handles; next: first hypothesis for handles; next: alternative hypothesis; right: a vase similar to the left (handles maintained)

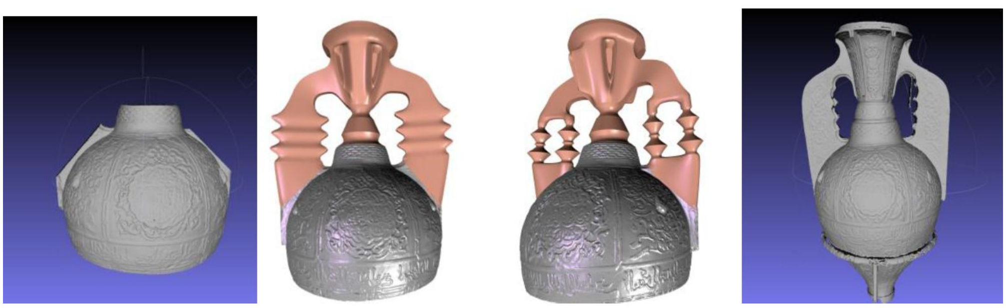

Figure 16: Sketch-based modeling

8. Presenting History

Presenting history to CH professionals as well as to the wider public requires realistic rendering and appropriate interaction techniques. With sophisticated rendering we close the loop to the corresponding material acquisition techniques presented in the first section of this paper.

Figure 17: Left: virtual model of an optically complicated object; right: Community Presenter

Presenting history to a wider audience, e.g. in a museum, specific dedicated user interfaces are required that take the variety of visitors (age, experience, …) into account. Experience tells us that virtual exhibitions in museums must be

- easy-to-use
- stable / robust
- simple with respect to interaction devices
- restricted as needed to convey the story with respect to navigation techniques

Community Presenter (depicted in fig. 17 on the right hand side) is a highly configurable tool to create dedicated presentations of virtual 3D artifacts for large audiences.

Media from the repository can be exported to Community Presented as the basis for the story that shall be told in the presentation. A designing stage takes place to define the narrative of the story in Community Presenter (CP). Scripting capabilities of CP allow easily realizing the implementation of the story-telling in CP and the parameterization of the user interface.

To give an example on how simplistic a user interface for such a presentation can get, we refer to a community presentation of antique coins which have been acquired in the course of 3D-COFORM as digital 3D objects. To keep things simple, the presentation was restricted to 2D-interaction to relieve users from fiddling with 3D navigation, i.e. instead of rotating a coin about an arbitrary axis, CP – in this case – just provided the option to flip the coin on user request. This example unveils how unfamiliar common museum visitors are with 3D interaction. Similar experience exist with even more immersive virtual exhibitions which have shown that in many cases limiting interaction to 2D for museum visitors can be beneficial.

9. Conclusions and outlook

This paper gave an overlook about the Integrating Project 3D-COFORM funded by the European Commission. The showcases in the paper have been selected to exemplify how 3D-COFORM is addressing all steps of digital 3D documentation in the lifecycle of virtual CH artifacts.

Beside these technical developments the 3D-COFORM project is addressing the practicalities of 3D documentation process including also research on business models which on the longer run will enable museum to find re-financing possibilities for the investments involved in 3D digitization.

The project so far has considerably pushed the leading edge in digital documentation in the CH field and produced more than 100 research papers. However, with each research challenge answered two new ones pop-up. Thus, the following challenges will remain interesting research topics in the future, among others:

- complicated objects
- reasoning about shape
- ‚industrialization‘ of 3D acquisition
- shortening time from 3D models to new user experiences
- embedding into existing information structures
- adaptation in practical workflows

We are convinced that digital 3D documentation in Cultural Heritage is the way to go, and current trends like 3D Internet will largely contribute to a break-through of 3D in this field, making it an everyday choice for museums, CH professionals and increasing attractiveness for museum visitors.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement n° 231809; 2008-2012. The authors like to thank their contributors to this paper, especially the following institutions and names:

- Leuven University and ETH Zürich: Luc van Gool and his researchers
- Bonn University: Reinhard Klein and his researchers
- Graz University: Sven Havemann and his research team
- CNR ISTI: Roberto Scopigno and his research team
- FORTH: Martin Dörr and his team

3D-COFORM is bringing together more than 100 researchers all over Europe, many cannot be mentioned here explicitly – but we appreciate all their valuable support!

References

[TvG11] Tingdahl David, Van Gool Luc « A Public System for Image Based 3D Model Generation » Computer Vision/Computer Graphics Collaboration Techniques, 5th International Conference, MIRAGE 2011, Rocquencourt, France, October 10-11, 2011. Proceedings, Lecture Notes in Computer Science, Volume 6930, 2011, DOI: 10.1007/978-3-642-24136-9, pages 262-273

[WSRK11] Michael Weinmann, Christopher Schwartz, Roland Ruiters, Reinhard Klein « A Multi-Camera, Multi-Projector Super-Resolution Framework for Structured Light » In Proceedings of 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), pp. 397-404, May 2011

[SWRK11] Christopher Schwartz, Michael Weinmann, Roland Ruiters, Reinhard Klein « Integrated High-Quality Acquisition of Geometry and Appearance for Cultural Heritage » In proceedings of the 12th International Symposium on Virtual Reality, Archeology and Cultural Heritage VAST 2011, pp. 25-32, October 2011

[RSGS10] Ritz, Martin; Scholz, Manuel; Goesele, Michael; Stork, André: High Resolution Acquisition of Detailed Surfaces with Lens-Shifted Structured Light. In: Artusi, Alessandro (Ed.) u.a.; European Association for Computer Graphics (Eurographics): VAST 2010. : Eurographics Symposium Proceedings. Goslar : Eurographics Association, 2010, pp. 1-8; 179 (Color plate)

[ARSG10] Ackermann, Jens; Ritz, Martin; Stork, André; Goesele, Michael: Removing the Example from Example-based Photometric Stereo. In: You, S. (Organizer) u.a.: Reconstruction and Modeling of Large-Scale 3D Virtual Environments : ECCV 2010 Workshop W10. 2010, 14 p.

[DMCCS11] Matteo Dellepiane, Ricardo Marroquim, Marco Callieri, Paolo Cignoni, and Roberto Scopigno: Flow-based local optimization for image-to-geometry projection. In: IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS: IEEE VisWeek 2011, IEEE 2011.

[PBHTGF10] Pan, X., Beckmann, P., Havemann, S., Tzompanaki, K., Doerr, M. and Fellner, D.W. (2010). A Distributed Object Repository for Cultural Heritage. Proc. The 11th Intern. Symp. on Virtual Reality, Archaeology and Cultural Heritage (VAST 2010), pp. 105–114, Eurographics Press

[PSDTGCS11] Pena Serna S., Scopigno R., Doerr M., Theodoridou M., Georgis C., Ponchio F., Stork A. (2011) 3D-centred media linking and semantic enrichment through integrated searching, browsing, viewing and annotating. In VAST11: The 12th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (Prato, Italy, 2011). Note: this paper has been selected as a Best Paper at the conference and will be invited for publication of an extended version on ACM J. on Computing and Cultural heritage.

[HUF12] Havemann S., Ullrich T. & Fellner D.W., *The Meaning of Shape and some Techniques to Extract It*, Multimedia Information Extraction, to appear.

[BHSF09] Bein, Matthias; Havemann, Sven; Stork, André; Fellner, Dieter W.: "Sketching Subdivision Surfaces". ACM SIGGRAPH / Eurographics Symposium on Sketch-Based Interfaces and Modelling. August 1-2, 2009, New Orleans, USA.