TOWARDS A KNOWLEDGE MODEL BRIDGING TECHNOLOGIES AND APPLICATIONS IN CULTURAL HERITAGE DOCUMENTATION

F. Boochs[a, *], A. Trémeau[b], O. Murphy[c], M. Gerke[d], J. L. Lerma[e], A. Karmacharya[a], M. Karaszewski[f]

[a] Institut i3mainz, University of Applied Sciences, Mainz, Germany (boochs, ashish)@fh-mainz.de
[b] Laboratoire Hubert Curien, University Jean Monnet, France, alain.tremeau@univ-st-etienne.fr
[c] Digital Arts and Humanities, University College Cork, Ireland, o.murphy@ucc.ie
[d] ITC Faculty, EOS department, University of Twente, Enschede, The Netherlands, m.gerke@utwente.nl
[e] Dept. Cartographic Eng., Geodesy and Photogrammetry, Universitat Politecnica de Valencia, Spain, jllerma@cgf.upv.es
[f] Faculty of Mechatronics, Warsaw University of Technology, Poland, m.karaszewski@mchtr.pw.edu.pl

Commission V

KEY WORDS: Cultural Heritage, Documentation, Optical, Measurement, Knowledge Base

ABSTRACT:

This paper documents the formulation of an international, interdisciplinary study, on a concerted European level, to prepare an innovative, reliable, independent and global knowledge base facilitating the use of today's and future optical measuring techniques for the documentation of cultural heritage. Cultural heritage professionals, color engineers and scientists share similar goals for the documentation, curation, long-term preservation and representation of cultural heritage artifacts. Their focus is on accuracy in the digital capture and remediation of artefacts through a range of temporal, spatial and technical constraints. A shared vocabulary to interrogate these shared concerns will transform mutual understanding and facilitate an agreed movement forward in cultural heritage documentation here proposed in the work of the COST Action Color and Space in Cultural Heritage (COSCH). The goal is a model that captures the shared concerns of professionals for a standards-based solution with an organic Linked Data model. The knowledge representation proposed here invokes a GUI interface for non-expert users of capture technologies, facilitates, and formulates their engagement with key questions for the field.

1. INTRODUCTION

The importance of effective protection and preservation of CH is internationally understood in terms of society, history, identity and memory amongst other concerns - within this context it is paramount to scan, document, analyze, understand, model, virtually reconstruct and visualize/publish CH objects, in particular to

- accurately record artefacts at both micro and nano-scales – to include material properties such as form, color and texture – for today's use and future generations;
- make the resulting e-documentation accessible globally to specialists and the general public;
- monitor the condition of objects for enhanced preventive conservation;
- enhance the knowledge base for art-historical analysis and other scholarly activities;
- support routine applications with specialist know-how and state-of-the-art equipment.

While the level of European technical competence in the precise documentation of spatial or spectral characteristics of surfaces is high, there is no common standard concerning three-dimensional (3D) shape and color existing for precise documentation of CH objects. Despite a general understanding of spatial resolution and accuracy of such documentation, and its potential, within the CH community, there is limited awareness that standards could be improved by direct cooperation within the technical sector. It is, therefore, difficult for CH professionals to use these technologies efficiently or even to define requirements. This paper proposes a knowledge based solution to bridge the gap between the CH community and computer scientists and engineers by fostering information exchange and providing guidelines for using optical technologies for CH documentation.

The paper introduces the COSCH Knowledge Representation (COSCH[KR]) as an optimal framework to overcome those limitations of projects that are usually object-dependent and application-driven, leading to unshared and non-standardized results - providing an interdisciplinary framework for scientists and technicians (developers of measurement systems, software and technologies for a wide range of applications, as well as material scientists, physicists and chemists) and the heritage specialists (art historians, conservators, archaeologists, curators and others) to facilitate the exchange of interests, needs, capabilities, constraints, limits and perspectives.

2. MOTIVATION

Thinking about our tangible cultural heritage, we see a broad field of studies, applications and object categories. As complex as the scope of studies are the **instruments** for non-contact CH documentation used to provide necessary data essentially contributing to the work of human scientists. They may include:

- Digital photography, which provides valuable visible information, but is subjective and cannot be directly used in context of other acquisitions owing to the lack of unique and known scale, unless it is integrated within a 3D workflow, see below.
- Infrared reflectography, which is based on a higher transmission of infrared light, is useful for detecting the underdrawings in paintings.

- Traditional colorimetry and spectrophotometry, which provide accurate information on the optical properties (such as the reflectance) and color appearance (such as color coordinates) of the samples analyzed.
- Imaging systems for specialist analyses, such as computer tomography, which make use of various non-optical parts of the electromagnetic spectrum and may enable examination below the surface and through the object.
- Color, multi- and hyperspectral imaging, which usually relies on many acquisition channels and gives detailed information to the spectral characteristics.
- Structured-light-based techniques, which provide precise spatial models and can be easily combined with color images of the object to give a 3D rendering of real appearance of the object.
- Passive 3D imaging techniques, which are complementary to structured light and use the existing light to collect spatial and visual data which leads to lower quality of the spatial model.
- 3D laser scanning techniques, which scan the object of interest using different strategies and provide generally a vast point cloud allowing accurate and global 3D investigations.
- Integrated multi-imaging systems, which are complex instruments and often a technological compromise. They need to support different technological concepts in order to responds to various wavelengths and to perform dimensional imaging.

This list might be continued and develops further as technological progress offers new possibilities. Due to the variety of all these instruments, it is also impossible to possess the respective knowledge required to correctly apply and control all these techniques for individual persons, even if they are technicians experienced in the use of non-contact measuring systems. This is certainly still more demanding for users, who in general are primarily interested in data helping their applications without having to know precisely what type of instruments are available.

Overall, complexity increases further through the interaction of measuring techniques and the **object** itself. Looking at the variety of objects, we see another list of characteristics having impact on the choice, use and appropriateness of instruments. Aspects like object size (ranging from small artefacts to large sites), shape (rather flat objects or complete spatial geometries), surface morphology (smooth versus ragged or indented), reflectivity (shiny or diffuse), texture (uniform or varying), spectral / color appearance or material composition play their role and may decide upon the quality of results or the feasibility of techniques. Experienced technicians should know these facts and be able to handle their instruments in the right way, but it is not always possible to overcome restrictions without manipulating the object (like the 3D capture of shiny surfaces): what might be suitable from a technical perspective might be strictly forbidden from the user view. It is therefore essential to know constraints introduced by the user, or the object, respectively.

Another group of variables contributing to the global process of optimal documentation is related to the **environment** or practical conditions for data capture. Constraints might be set through the fact that the object is exposed to the normal outdoor conditions or is in a protected indoor situation allowing preparing and controlling the process of data capture in a way as

best suited for the technology to be used. This concerns questions of susceptibility of the equipment or the object to certain physical influences (like humidity, for example) but also addresses aspects like the ability to prepare or not to prepare the set up in an appropriate manner (control of lighting conditions, for example). Other parameters like the accessibility of an object or a site (measurement in underground caves vs. data capture in a museum) may also play a role. Similarly, the question of whether an object can be observed under geometrical stable circumstances or is moveable and can be transferred to a laboratory, versus a fixed object site to which the equipment has to be moved.

The final important impact is related to the application as such and the needs to be fulfilled by the data. It is obvious that only engineers and technicians (information providers, as pointed out by Letellier, (Letellier, 2007)) knowing about optical measuring techniques are able to define the content and quality of data provided through certain collection processes (Pavlidis, 2007). However, often they also propose optimal **content of data** based on an anticipation of the requirements an information user might have, without clearly understanding the evaluation process realized by conservators, art historians, curators, archaeologist, and other professionals. This might lead in the same way to non-optimal data as the inverse case, when the information user asks for certain input, without understanding the instruments, their characteristics and constraints. It is therefore of real importance to have a dialog between both sides (information users and information providers) to adjust their **respective perspectives**. This addresses the vocabulary (what does accuracy mean?) and the characteristics of the data (scale, resolution, accuracy, composition...) to be provided in order to optimally serve the work of the various groups of information users (archaeologists, architects, conservators, curators, social scientists, art historians and others).

Summarizing all these facts, we are facing a complex scenario when considering optical documentation techniques and their optimal use for many different applications. In particular when not only thinking about simple 3D models but rather about the broad bandwidth of data to be provided for the whole field of questions to be answered for our common tangible cultural heritage.

Many cases exist for single objects and projects, which have been successfully handled in the past and individually answering questions about possible strategies under selected constraints (Boochs, 2008; Böhler, 2005, as examples). However, there are no well-established and commonly accepted standards for precise, non-contact documentation of CH objects that would implement and combine the above-mentioned techniques and relate them to particular applications.

CIPA Heritage Documentation (CIPA, 2014) is the international ICOMOS / ISPRS scientific committee that ensures the right documentation with the existing variety of techniques for preservation, conservation and restoration. A good set of guidelines were undertaken under the RecorDIM initiative (sponsored by The Getty Conservation Institute). CIPA Heritage Documentation encourages and promotes the use of appropriate documentation practice, advises organizations for recording cultural heritage and provides an international forum for exchanging scientific knowledge, ideas and best practices.

The COSCH action (COSCH, 2012) makes a logical attempt to structure this complex scenario and to facilitate the choice of

documentation strategies optimally serving individual applications.

Such an attempt needs to be based on:

- a broad field of competences grasping technical knowledge in the same way as application skills
- a concrete collection of as many relevant factors as possible
- the identification and implementation of a structure expressing dependencies and relations of these facts (see section 4.2)
- the development of an application allowing to make use of this knowledge (see section 4.3)

The first two aspects are directly covered by COSCH. Main focus of this Action is to improve the dialog between technicians and end users from the human sciences. It therefore brings a heterogeneous group of people together covering knowledge ranging from spectral and spatial documentation techniques and related algorithmic processing to various applications including restoration, art historical analysis, and archaeology and conservation science. Consequently, a unique pool of knowledge is available supporting such an attempt.

It still needs a way to structure this knowledge and to make it accessible to everybody. Here it is possible to profit from developments in other areas also handling heterogeneous and extensive information: the Semantic Web (Berners-Lee, 2001). Making appropriate use of techniques developed in this field it is possible to build a knowledge frame grasping instruments, their characteristics, objects and their typical features at the same time as the usage of techniques, the processing of data and the needs decided by a broad field of applications. In the end it needs to build an ontology covering these various domain fields expressing relations and rules in between and thus formalizing what specialist have collected.

Such a structured knowledge base will serve many aims. It

- will cover technical and application views at the same time
- will simplify the selection of the most appropriate documentation techniques
- will help the humanities to utilize technical knowledge
- will help the humanities to make more informed decisions for documentation purposes
- will make all COSCH collected knowledge accessible for the scientific community
- will provide a flexible base which can constantly grow

3. STATE OF THE ART

In any rational case, information generates knowledge and the generated knowledge shapes the attitude that leads to the behavior (Kollmuss, 2002). Knowledge is an essential factor in every domain of activities to hold things together. The key factor is the use of existing knowledge in more than one situation or for more than one individual. This is the base purpose of any knowledge management system and its underlying models. Knowledge Management as a concept has existed for decades and is no longer a new research area (Davenport, 1998). Formally, "Knowledge Management is a discipline that promotes an integrated approach to identifying, capturing, evaluating, retrieving, and sharing all of an enterprise's information assets which may include documents, databases, policies, expertise and experiences (captured/uncaptured) in the individuals" (Duhon, 2008).

Computational Information Science has taken long strides in last few decades. Now we consider the semantics of the piece of information rather than information itself. Semantics are captured through conceptual models that structure the information sets. These conceptual models support organizing information along generic abstractions through primitive concepts as entity, activity, agent, and goal (Mylopoulos, 1998). These concepts are popularly known as "Ontologies" and play key role in formalizing knowledge. Ontology is the study of existence and essence and a recognized term in humanities. The term in computing refers to sets of controlled vocabularies; and such studies on defining formal ontologies that formalize very general concepts that hold true across domains and disciplines are popular in both pre-computational and current computational eras. These ontologies form Knowledge Bases defining formalized representations of facts, rules, and heuristics that could be used for inferring new knowledge on objects and events. Formal ontologies like Cyc (http://www.cyc.com), DOLCE (a Descriptive Ontology for Linguistic and Cognitive Engineering, http://www.loa-cnr.it/DOLCE.html), PROTON (PROTo Ontology, http://proton.semanticweb.org) are already being used in linguistic and semantic indexing areas. A comprehensive survey on existing formal ontologies is presented in (Mascardi, 2007). Likewise, ontologies are developed for specific areas. These ontologies are define formal descriptions of the concepts in specific application areas and are better known as Domain Ontologies (Musen, 1998).

CIDOC-CRM (CIDOC-CRM, 2013) is currently the knowledge model for documentation and information sharing within the Cultural Heritage domain (Boeuf, 2013). This ISO 21127:2006 conceptual relationship model constitutes a scalable ontology representing concepts within cultural heritage and museum documentation. Though it is a domain ontology specifically designed for the domain of CH, it constitutes components of a formal ontology and classes defining events and objects in time and space. CIDOC-CRM is based on the objective of the integration of differing, large numbers of information resources and offers a platform to make information compatible to CRM in order to benefit from semantic interoperability (Boeuf, 2013). CIDOC-CRM constitutes different concepts and their relationships against each other for CH documentation. Though an expressive model, it is not designed for knowledge discovery per se. In relation to optical measurement perspectives, in this case, it lacks the definitions of facts, rules and heuristics that could be used infer the knowledge discovery. Projects like ResearchSpace (RS) Project (Alexiev, 2013), use the model for inferring knowledge. The Europeana Data Model (Europeana, 2014) is a knowledge model based on existing standards, thesauri and knowledge models like CIDOC-CRM to accommodate data models from different data sources within different countries of Europe (Charles, 2013). Data integration and interoperability have always been the major objectives behind the use of knowledge technology in CH domain. MultimediaN E-Culture uses semantic technologies to bridge the information gaps between several cultural institutions (Rijksmuseum, Louvre, Tropenmuseum, and so on) through diffused thesauri and bring different online repositories of the cultural heritage (Schreiber, 2006). Likewise project "The Museum Finland" uses seven different domain ontologies (artifacts, materials, actors, situations, locations, times, collections) that have been defined through analysis and

incorporate national and local thesauri to integrate different CH databases of the country (Mantegari, 2009).

Knowledge models have been traditionally used for content management and retrieval in cultural heritage. The "Centre de Recherche et de Restauration des Musées de France" (C2RMF) manages its content through its customized Content Management System (CMS) EROS CMS. The metadata schema of EROS database system is mapped to CIDOC-CRM. The intention is to provide a suitable online platform through open-source interaction model mSpace (http://www.mspace.fm) to explore the C2RMF's content semantically (Pillay, 2007). The 4TB of high resolution and high dynamic range imaging which include X-Ray, Infrared, visible light among them are to be shared through the platform. Limited knowledge models has also been researched on managing techniques and technologies in cultural heritage. (De Luca, 2013) presents a synthesis on existing research works on structuring heterogeneous data and their semantically enriched 3D models of cultural heritage from different contexts. Today, ontology based knowledge models are used to bridge gaps between different data providers working strictly in cultural heritage domains or to add structured semantics to data, in order to improve or facilitate its use. Existing knowledge models like CIDOC-CRM or Europeana promote data interoperability. However comprehensive, they do not address techniques and technologies that are vital in any conservation and restoration activities. A need for a generic knowledge model, which addresses issues on these techniques and technologies, is being constituted as part of the COSCH action. The model takes both aspects of expertise under its semantic umbrella. This paper describes the initial work towards achieving such a model.

4. COSCH^{KR}: KNOWLEDGE REPRESENTATION MODEL FOR COSCH

4.1 General Description

Cultural Heritage is the domain where arguably one of the largest numbers of interdisciplinary activities concentrates in a common goal. From discovery to restoration and analysis of objects, various activities from completely different disciplines are involved. These disciplines need to communicate with each other through a reliable platform for achieving the appropriate result. COSCH^{KR} (COSCH Knowledge Representation) intends to provide such platform. The purpose is twofold: 1) to bridge between technical expertise in documentation and the expertise in restoration and analysis and 2) to reuse already existing knowledge within individual and cross domains for the obvious benefits. In this sense, the COSCH^{KR} should present a platform to store and represent knowledge of individual domain with interrelations to other expert domains.

COSCH^{KR} benefits from the recent developments in the Semantic Web framework and its underlying technologies. The knowledge model is expressed through the Web Ontology Language (OWL), which is W3C recommendation to define ontology since 2004 (Horrocks, 2007).

4.2 Knowledge Representation (How?)

COSCH^{KR} is a knowledge model with enriched facts, rules and heuristics binding different expert domains and thus is different from CIDOC-CRM. Figure 1 illustrates the top-level ontology of the model as defined for the first version.
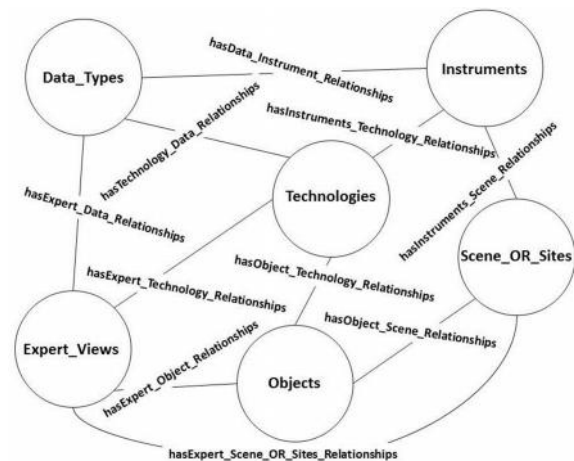

Figure 1: Top-level ontology of COSCH^{KR}

The top-level ontology constitutes of six major top-level classes. Each class encapsulates the expert knowledge from every domain within CH. For example, the class Technologies has sub-classes consisting of *Acquisition_Technology*, *Documentation_Technology*, *Measuring_Technology* and *Usage_Of_Technology*. Within every specialized class, the expert knowledge of each technical domain is presented. Let us consider for example the specialized class *Acquisition_Technology* (within class *Technologies*) where the knowledge of data acquisition is stored and represented. This class further specializes into spectral, spatial, and other related domains where each holds the knowledge of data acquisition techniques of respective specific domain.
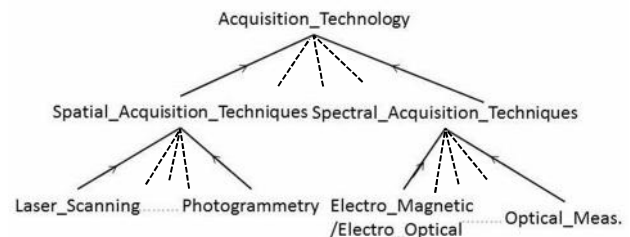

Figure 2. Taxonomical hierarchy of incorporating data acquisition techniques of different technical domains

These two acquisition techniques differ in terms of their usages and requirements while documenting objects in CH and are thus represented within the knowledge model. *Spatial_Acquisition_Techniques* can only be considered while documenting the geometries of the objects whereas *Spectral_Acquisition_Techniques* can only be considered while documenting colors at different spectral channels. They are presented through relevant rules within the model. For instance, *Spatial_Acquisition_Techniques* is suitable to the objects with geometry and is presented through given rule within the model:

Spatial Acquisition Techniques (is Suitable technology for) Physical Objects (having) Geometry   *(1)*

The rules will be expressed as eq. 1 in this paper for simplicity and easy to understand. However, there are Description Logic notations (Baader, 2001) to define them.

Similarly, the rule that defines *Spectral Acquisition Techniques* should consider color properties and other physical properties (e.g. albedo, roughness and surface normal, like (Chen, 2012))

related to visual appearance rather than geometries of the object. In some applications, such as pigment analysis, the data obtained with optical measurement systems can be completed with the ones collected by non-optical measurement systems (Rebollo San Miguel, 2011). In some study cases, spectral acquisition techniques not only refer to colors, but also a whole lot of other properties recorded by non-optical sensors (e.g. LIBS, XRF, Raman and FORS spectroscopies).

Relationships represented through the arcs in Figure 1 provide the necessary bridges between different knowledge domains. For instance, in a real world scenario, the purpose of documenting CH objects triggers the technology to be chosen which in turn affects the choice of instruments. The selection of instruments - on the other hand - depends on the nature of the site or the budget available and so on. All these aspects are bound through respective interdependent relationships. The idea is to tap the underlying knowledge within those relationships and model them as a real world scenario in the COSCH[KR] knowledge model. The knowledge model then can be inferred by different domain experts for their sought after answers.

COSCH[KR] knowledge model is developed to fill in gaps between different domain experts especially those of users working on documenting CH objects (archaeologists, rest orators, museum experts, and so on) and that of technical experts (engineers). Views and expectations on single particular terms or events can be interpreted completely differently by these two different expert communities. One of the major objectives of "the Semantic Web" framework is to plug these differences through semantically mapping them together. COSCH[KR] follows the same in conceptually mapping both previously perceived as differing datasets within the one graph. The top-level class *Expert_Views* (see Fig.1) intends to provide the framework where experts from different domains (especially from those two expert communities) define different aspects from their point of view. The model maps them semantically together to give satisfactory answers for both communities.

4.3 COSCH[KR] Application

COSCH[KR] Knowledge model encapsulates the expert knowledge from different domains of CH, which will be utilized through an interactive frontend tool. COSCH[KR] Application is intended to provide a common platform to experts and CH users alike to put forward their queries and get answers without worrying about the complexity of the backend model. The application should allow seeking answers in varying nature: simple to complex and should invoke the knowledge model to infer underlying facts and heuristics.

5. EXAMPLES

This section presents examples on how COSCH[KR] bridges the gaps between technical experts in spatial and spectral domains and experts working in CH to suggest appropriate solutions to their requirements.

5.1 Color and Spectral Image Acquisition

The intention of the following example is to provide a view of relationships between the needs of users and the knowledge of experts in the selection of a color images acquisition device. It is difficult for a user, even for an expert, to have a good (comprehensive) view of all parameters (factors) that are decisive in the selection of an acquisition device. Meanwhile, one expert could argue that the most decisive factor is the

accuracy of the color measurement and that in this context the best instrument is a 2D grey level camera with 32 spectral filters (LCTF). Another expert could also argue that the most decisive factors are: (1) the cost; (2) the speed; (3) the portability; (4) the weight and (5) the usability of the system and that in this context the best instrument is a 2D RGB camera with color filters. In this case, the objective of the Knowledge Representation Model is to help users (by implicit reasoning) to identify themselves *what* are the main factors useful for each case study and to identify *what* are the other factors that cannot be satisfied.

As example, let us consider that a user (*case study 1*) wants to digitize a small painting (e.g. The "Pot of Geraniums" painted by Matisse, height 41.3 cm, width 33.3 cm) in order to measure the spectral properties of its pigments. For this specific purpose, (Zhao, 2008), (Tamplin, 2005) and (Berns, 2003) recommend using a 2D grey level camera with 32 spectral filters (LCTF) to perform such acquisition. Now, let us consider that another user (*case study 2*) wants to digitize another small painting (e.g. The "Fish" mentioned in (Chen, 2012), height 20 cm, width 32 cm) in order to measure the visual appearance of this painting, i.e. to characterize both its color appearance and its physical properties (e.g. diffuse albedo, specular albedo, specular roughness, and surface normal). For this specific purpose, (Chen, 2012) recommends using a 2D RGB camera with color filters to perform such acquisition.

From these two examples (a 2D surface), it appears that the choice of a given technology depends also of the intent (expectation) of the user (i.e. *Usage_Of_Technology*) and of the relationships between this intent and the rules and factors listed below (see eq. 2) . Let us note that for these two study cases only three of the six major top-level classes have an impact on the decision:

Users [(has Intention on) some Usage] and [(has Impressions on) Physical Objects (on) Reflectivity and Roughness] and [(has Expectation on) Spectral Techniques (with Accuracy in) Color and Spectral Channel] (2)

The list of rules defining expectations of users can go on and cover other knowledge classes in the top-level ontology.

The number of useful factors varies from one object type to another one. Meanwhile some factors are implicit (e.g. the *Instruments_Characters*) others have to be guided by implicit reasoning from *Expert_Views Knowledge* (e.g. Physical properties). Meanwhile in some study cases (e.g. *case study* 1) some factors are at top level in some cases study (e.g. spectral accuracy), others are less decisive (e.g. lighting conditions) (see eq. 3). In other study cases (case study 2) the opposite happens, e.g. lighting conditions are at the top level and spectral accuracy is less decisive (see eq. 4).

Physical Objects (is Situated inside) Condition (having) Any Lighting AND Spectral Technology (has Accuracy) High Spectral Accuracy (3)

Physical Objects (is Situated inside) Condition (having) Good Lighting AND Spectral Technology (has Accuracy) Any Spectral Accuracy (4)

The relationships mentioned above can be encapsulated by heuristics based on knowledge provided by the state of the art. The COSCH[KR] Model can also help the used to optimize the number of factors to set at top level (to relax the number of

constraints) when no system satisfies his/her expectations. Thus, for the case study 2, according to the current *Expert_View* there is no low cost and fast instrument, which can measure the visual appearance of a painting. Even if these two factors seem to be decisive for the user, the model can suggest to the user to set the accuracy of the measurement (in particular the *Physical_Properties*) as major factor rather than the cost or the speed (i.e. *Instruments_Characters*). Moreover, even if the *Lighting_Instrument* (*hasObject_Instrument_Relationships)* and the *Calibration_Instrument* (*hasExpert_Objects_Relationships)* sub-classes seem to be secondary for the user, the model can raise awareness the user that these factors can have a decisive impact on the result (see (Chen, 2012)).

5.2 Influence of the goal of digitization on the process

The following example shows the differences in creating documentation of cultural heritage objects caused by different requirements to the final model. Let us assume that we want to create a computer model of a moderately large object, for example a 2 meter high sandstone vase, placed in the garden of King Jan III's Palace Museum at Warsaw. This vase was made by J. A. Karinger and J. A. Siegwitz in the 18th century. It is ornamented with sculptures of mythological characters. The natural color of sandstone has been changed locally by atmospheric conditions, so the information about it is also desirable to climatologists, geographers and planners outside the CH domain.

In this example, two digitization processes are described. The first one is aimed at obtaining very accurate and dense 3D model with good, detailed color representation in the form of 3D point cloud to be used by art conservators for various analyses and as a true copy for professional documentation. For this purpose as the object's material is sandstone, some experts proposed spatial resolution of about 2500 points / mm^2 (Bunsch, 2011), the same for geometry and color. The second digitization process' purpose is to create a model for visualization in the Internet. In this case, the accuracy and density of the model is not as crucial. In addition, color information should be present, but it is not required to be of such high density as in the previous case. Moreover, the output data type may be different (sparse 3D point clouds are not well suited for visualization purposes) (Meyer, 2007) and even 2D images can be sufficient. Finally, from a cost perspective, the system for creating Internet content should be low-cost, easy and fast to operate and possibly portable to allow for digitization of wide range of objects, including this one outdoors.

For those two examples, the choice of digitization technology depends on the end-user expectations related to some general rules and factors. The significant influences on the final decision have all top-level classes of COSCH^{KR} ontology. The following (eq. 5) illustrates the first case where the user requires very high resolution result with detailed representation of the object.

User (has Intention of) Analysis AND Analysis (has Requirements) Technology (which Produces) 3D model (with) Dense and high Detailed and high Resolution (And) Color (with) High Accuracy   *(5)*

Inferring inside the model; *Users (needs to use) Technology (which Produces) 3D model (with) Dense and high Detailed and high Resolution (has Color) Color (with) High Accuracy*
(6)

The second case is to produce result that can be used to disseminate through Internet (eq. 7 - 9).

User (has Intention of) Internet Publication   *(7)*

Internet Publication [(has Requirements) Technology (which Cost) Cheap in Price AND Technology (Producing) 2D images (with) Good Resolution (is) Cheap in Price]   *(8)*

Inferring inside the model: *Users (needs to use) Technology (which Produces) 2D images (with) any Density and Good Resolution (And) Color (with) Good Accuracy*   *(9)*

Note, the model infers to recommend a completely different technology.

5.3 Geometric Camera calibration knowledge schema

Based on the top-level ontology of COSCH^{KR}, the geometric camera calibration issue has strong interrelations among the six classes: *Technologies, Instruments, Data_Types, Expert_Views, Objects, Scene_OR_Sites*. The Technologies class is clearly covered by the Optical measurement method_Photogrammetry. Knowledge about the Objects is mandatory, in particular (Size of the artefacts, Surface, Outer Geometry and Texture) and its interdependence with *Scene_OR_Sites*. It is understood that the devices used in the survey fulfil the specifications as regards Instruments_Characters (Complexity, Cost, Speed and Mobility). The features of the instruments used in the survey have to be clarified as regards *Supporting_Instruments* (Camera, Lens) and related metadata, for instance, number of images, camera set up, existence of *Calibration pattern* (as part of Calibration_Instruments) on the image (*Data_types*, 2D, Images).

A tentative list of 2D Camera Calibration Algorithms-is covered in photogrammetric and computer vision textbooks and papers such as Brown (1971), Fraser (1997) and Zhang (2000). Nevertheless, the right selection of additional parameters is even more important than the method, as well as the flexibility of the solution in order to cope with local and/or global parameters, especially when dealing with brand new zoom autofocus compact digital cameras.

The case study 1 (The "Pot of Geraniums" painted by Matisse) presented in Section 5.1 for color and spectral image acquisition is presented next to clarify the performance of the COSCH^{KR}. The purpose of the solution is to correct the distortion of a single image acquired with a zoom digital camera for documentation of paintings based on straight lines (Lerma, 2007). It can be a typical scenario in many museums, art galleries. The following eq. 10 presents the intention of the users while those presented through eq. 11 presents the existing scenarios that would be taken into consideration by the knowledge model.

Users (has Intention to) Distortion Correctness and (has Data Input) Image (with Present Number) 1   *(10)*

Image [(Recovered Format) EXIF] AND [(Taken With) Camera (with) Lens (that is) Zoomable (having) High Distortion] AND [(has Calibration Pattern) None] AND [(has Frame) Right] AND [(contains) Geometry (that is) Straight Line]   *(11)*

Inferring inside the model; *Users (should Use) Straight Line Calibration (is A) Calibration Algorithm (is Best for) Distortion Correctness (is Suitable with) Images equal to 1 (has*

Calibration Pattern) None and (has Frame) Right and (contains) Geometry (that is) Straight Lines (12)

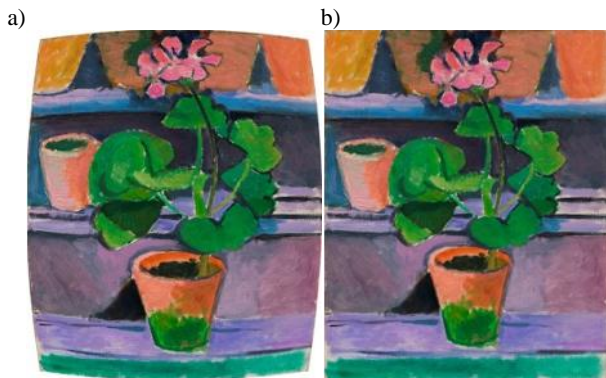The result of the camera calibration is presented in Fig. 3.

a)   b)

Figure 3. Result of the camera calibration: a) input image; b) output images corrected for lens distortion

5.4 3D documentation by Structure-from-motion

This last example combines elements from the two former ones. It is about the spatial/geometric documentation of a quite large object, which is permanently installed in a museum exhibition. In contrast to the vase example in 5.2 we are not aiming at such a high resolution; for the virtual museum and archeological documentation we are satisfied with an absolute point accuracy of less than 1 cm (maximum error) for signalized points and a point density of 1 per cm². Another condition is that the time available for the data acquisition is limited to minimize impact on museum visitors.

The COSCH[KR] system will come to the conclusion that a classical close range photogrammetry or structure from motion technique (sfm) in combination with a dense matching approach might be best suitable for this kind of object. The user of the COSCH[KR] also will be informed that a camera pre-calibration is an integral part of the photogrammetric workflow, so no special attention needs to be put on this, given that the accuracy requirements permit an on-the-job-calibration. An example is given below: The Lamassu is a human-headed winged bull dated to Ashurnasirpal II's (883-859 B.C). It is made of Alabastrine Limestone, has a size of approx. 1 x 4.5 x 4.5 m and is currently exhibited in the Iraq national museum in Baghdad, see Fig. 4.

Figure 4. Lamassue: Excavation activities, location in the museum and 3D point cloud model

In a research case study, the mean point error from a modern close range photogrammetric workflow was confirmed to be less than 5mm for signalized points and up to 1.5cm in poorly textured areas (Alsadik, 2014).

6. CONCLUSIONS

In this paper, we present initial developments towards establishing a structured knowledge base to allow linking of the complex and different worlds of technologies for non-contact optical object documentation on one side, and of applications and interests of users related to cultural heritage objects on the other side. The developments have their motivation in the increasing functionality and power of optical documentation techniques raising questions of appropriateness and suitability of possible strategies for an individual uses or application cases. Without correct identification and evaluation of manifold criteria, a user is unable to select the best way of documenting a case and resulting in oversized data, useless data or costs exceeding a reasonable level.

On the other hand, today's techniques for handling and structuring knowledge are well suited to developing a framework for solving the problem of this vast amount of factors characterizing optical measuring techniques. One important precondition for success is a good overview and access to representative knowledge in the science fields having an interest in these questions. Here the group of people engaged in the Cost Action COSCH acts as a good base, as it integrates technicians (spectral, spatial acquisition, algorithms & processing) and users (curators, conservators, art historians, archaeologists...) at the same time. All other people interested in supporting this idea and processes are cordially invited to join the Action and to contribute with their skills and experience.

7. References

Alexiev, V., Manov, D., Parvanova, J., Petrov, S., Kiryakov, A., 2013. Large-scale Reasoning with a Complex Heritage Ontology (CIDOC CRM). In: *Proceedings of workshop Practical Experiences with CIDOC CRM and its Extensions (CRMEX 2013)*, Malta.

Alsadik, B., Gerke, M., Vosselman, G., Daham, A., Jasim, L., 2014: Minimal camera networks for 3D image based modeling of cultural heritage objects. *Sensors*, 14(4), pp. 5785-5804.

Baader, F., Sattler, U., 2001. An overview of tableau algorithms for description logic, *Studia Logica,* 69(1), pp. 5 - 40.

Berners-Lee, T., Hendler, J., Lassila, O., 2001. The Semantic Web. *Scientific American*, pp. 34-43.

Berns, R.S., Taplin, L.A., Imai, F.H., Day, E.A., Day, D.C., 2003. Spectral Imaging of Matisse's Pot of Geraniums: a case study. In: *Proceedings of IS&T 11th Color Imaging Conference*, pp. 149-153.

Böhler, W., 2005. Comparison of 3D laser scanning and other 3D measurement techniques. *Recording, Modelling and Visualisation of Cultural Heritage*. Taylor & Francis Group, London, pp. 89-99.

Boeuf, P. L., Doerr, M., Ore, C. E., Stead, S., 2013. *Definition of the CIDOC Conceptual Reference Model*. ICOM/CIDOC CRM Special Interest Group. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.1-draft-2013May.pdf.

Boochs, F., Huxhagen, U., Kraus, K., 2008. Potential of high-precision measuring techniques for the monitoring of surfaces from heritage objects. In: *In-situ Monitoring of Monumental Surfaces, Intl. Workshop, SWM08,* Florenz, Italy.

Brown, D. C., 1971. Close-range camera calibration. *Photogrammetric Engineering*, 37(8), pp. 855-866.

Bunsch, E., Sitnik, R., Micho ski, J., 2011. Art documentation quality in function of 3D scanning resolution and precision. In: *Proceedings of SPIE 7869, Computer Vision and Image Analysis of Art II*, Vol. 7869, San Francisco, USA.

Charles, V., 2013. Introduction to the Europeana data model. In: *Proceedings of the 17th Int'l Conference on Theory and Practice of Digital Libraries*. Malta.

Chen T., Berns, R. S., 2012. Measuring the total appearance of paintings using a linear source, studio strobes, and a dual-RGB camera. Report for Andrew W. Mellon Foundation, Rochester, USA.

CIDOC-CRM, 2013. CIDOC-CRM, http://www.cidoc-crm.org/ (accessed July 26, 2013).

CIPA, 2014. CIPA Heritage Documentation, http://cipa.icomos.org/, date of access: 04.05.2014

COSCH 2012, COST Action Color and Space in Cultural Heritage, http://www.cost.eu/domains_actions/mpns/Actions/TD1201, date of access: 04.03.2014

Davenport, T. H., Prusak, L., 1998. *Working Knowledge - How Organizations Manage What They Know.* Boston, Massachusetts: Harvard Business School Press.

De Luca, L., 2013. 3D Modelling and Semantic Enrichment in Cultural Heritage. In: *Keynote at 54th Photogrammetric Week*. Stuttgart.

Duhon, B., 2008. It's all in Our Heads. *Inform 12(8)*, pp. 9 - 13.

Europeana, 2014, http://pro.europeana.eu/home, date of access: 05.03 2014.

Fraser, C. S., 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4), pp. 149-159.

Horrocks, I., Patel-Schneider, P. F., McGuinness, D. L., Welty, C.A., 2007. OWL: A Description Logic Based Ontology Language for the Semantic Web. In: *The Description Logic Handbook: Theory, Implementation, and Applications (2nd Edition)*, chapter 14. Cambridge University Press. pp. 1-32.

Kollmuss, A., Agyema, J., 2002. Mind the Gap. *Environmental Education Researc*h, 8(3), pp. 239 - 260.

Lerma, J. L., Cabrelles, M., 2007 A Review and Analyses of Plumb-Line Calibration. *The Photogrammetric Record*, 22(118), pp. 35-150.

Letellier, R., Schmid, W., LeBlanc, F., 2007. *Recording, Documentation, and Information Management for the Conservation of Heritage Places: Guiding Principles*. The Getty Conservation Institute, Los Angeles, USA.

Mantegari, G., 2009. Cultural Heritage on the Semantic Web: From Representation to Fruition. Ph.D. Dissertation, Università degli Studi di Milano-Bicocca.

Mascardi, V., Cordi, V., Rosso, P., 2007. A Comparison of Upper Ontologies, Technical Report DISI-TR-06-21. WOA 2007, pp. 55 - 64.

Meyer, E., Grussenmeyer, P., Perrin, J.P., Durand, A., Drap, P., 2007. A web information system for the management and the dissemination of Cultural Heritage data, *Journal of Cultural Heritage,*, pp. 396-411

Musen, M.A., 1998. Domain Ontologies in Software Engineering: Use of Protégé with EON architecture. *Methods Inf Med* 37(4 – 5), pp. 540 - 550.

Mylopoulos, J., Jurisica, I., Yu, E., 1998. Computational mechanisms for knowledge organization. In: *Proceedings of the 5th International Conference of the International Society of Knowledge Organization*, Lille, France. Vol 6, pp. 125-132.

Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., Chamzas, C., 2007. Methods for 3D digitization of Cultural Heritage, *Journal of Cultural Heritage*, 8(1), pp. 93-98.

Pillay, R., Aitken, G., Pitzalis, D., Lahanier, C., 2007. Archive Visualization and Exploration at the C2RMF. In: *Proceedings of International Cultural Heritage Informatics Meeting (ICHIM07)*. Toronto, Canada.

Rebollo San Miguel E. P., 2011, Applications of imaging spectroscopy to the chemistry of cultural heritage field. Ph. D. Dissertation, Università degli Studi di Padova, Dipartimento di Scienze Chimiche.

Schreiber, G., Amin, A., de Boer, V., Hardman, L., Hildebrand, M., Hollink, L., Huang, Z., van Kersen, J., de Niet, M., Omelayenko, B., van Ossenbruggen, J., Siebes, R., Taekema, J., Wielemaker, J., Wielinga, B., 2006. MultimediaN e-Culture demonstrator. In: *Proceedings of the 5th International Semantic Web Conference*: *Lecture Notes in Computer Science*, Athens, USA, pp. 951 - 958.

Tamplin, L. A.; Berns, R.S., 2005. Practical spectral capture systems for museum imaging. In: *Proceedings of 10th Congress of the International Color Association*, pp. 1287- 1290.

Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 22(11), pp. 1330-1334.

Zhao, Y.; Berns, R. S.; Taplin, L.A.; Coddington, J., 2008. An investigation of multispectral imaging for the mapping of pigments in paintings. In: *Proceedings of SPIE, 6810*, Vol. *681007(9).* San Jose, USA.

7. ACKNOWLEDGEMENTS