Sampled 3D Models for CH Applications: A Viable and Enabling New Medium or Just a Technological Exercise?

PAOLO CIGNONI and ROBERTO SCOPIGNO
Istituto di Scienza e Tecnologie dell'Informazione (ISTI),
Consiglio Nazionale delle Ricerche (CNR), Pisa, Italy

Any application of three-dimensional computer graphics in the Cultural Heritage (CH) field requires availability of a digital model of the artifact(s) treated. Detailed and accurate digital 3D models can be produced with 3D scanning devices, which allow conversion of reality into digital form in a cost and time-effective manner. We present the capabilities of this technology and the main issues which are preventing its wider use in contemporary applications, highlighting some open problems and a few promising new approaches for 3D model construction. We also briefly review some CH applications which could boost the diffusion and evolution of 3D scanning technology.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/Image Generation—*Digitizing and scanning*; J.5 [**Arts and humanities**]—*Architecture, fine arts*

General Terms: Algorithms

Additional Key Words and Phrases: 3D scanning, scanning pipeline, color data acquisition and mapping, efficient visualization, computer-aided restoration

ACM Reference Format:

Cignoni, P. and Scopigno, R. 2008. Sampled 3D models for CH applications: a viable and enabling new medium or just a technological exercise? ACM J. Comput. Cultur. Heritage 1, 1, Article 2 (June 2008), 23 pages. DOI = 10.1145/ 1367080.1367082 http://doi.acm.org/ 10.1145/1367080.1367082

1. INTRODUCTION

The availability of high-quality digital models is a basic factor in improving our capabilities to study or to communicate our Cultural Heritage (CH). The standard medium currently is two dimensional (images or videos), but in the near future, three-dimensional models will probably be the norm. The availability of digital 3D-models opens a wide spectrum of uses which can astonishingly improve our capabilities to study, analyze, recognize and compare artwork with other heritage items. Three dimensional models

---

We acknowledge the financial support of some active projects: EPOCH (EU Network of Excellence, IST-2002-507382), Fruizione - Tecniche di supporto e modalita innovative (project of the CNR Department on Cultural Heritage) and Media Multidimensionali (project of the CNR Department on ICT). Part of the work presented is also the outcome of several prior projects, such as ViHAP3D (EU IST-2001-32641) and RIS+ (Regione Toscana).

Author's address: Visual Computing Lab, ISTI—CNR, v. Moruzzi 1, 56100 Pisa, Italy; email: r.scopigno@isti.cnr.it, p.cignoni@isti.cnr.it

are the starting point for the design of a large number of applications based on visual encoding and communication, ranging from the passive ones (still images, video, computer animations) to the more interactive and immersive ones (multimedia books, interactive navigation, immersive VR/AR systems, etc.). Unfortunately, very few digital models of 3D artworks exist in the public domain where they can be freely accessed by students, practitioners or art experts. Our first goal is therefore to review the status of the technology supporting the semi-automatic acquisition of digital 3D models and the reasons for the current scarce availability of those 3D resources.

In the case of classical visual media (images, video), we have inexpensive, high-quality acquisition devices produced by consumer-oriented industries. This is unfortunately not the case for 3D acquisition technology. Three-dimensional scanning technology has evolved considerably in the last few years in terms of both hardware (HW) devices [Blais 2003] and algorithms for processing the raw data produced by the scanning devices [Bernardini and Rushmeier 2002]. Three-dimensional scanning devices are usually based on optical technology (laser or structured light) and use either triangulation (ideal for small and medium scale objects) or time-of-flight approach (effective on large scale objects, e.g., architectures). Unfortunately, most of the scanning technologies do not produce the final model in a single shot (as we are used to with the standard 2D media). Conversely, the output is raw data which has to be processed to build up a (hopefully) complete 3D model. A software (SW) process is needed to integrate and compose the raw data, usually requiring substantial user intervention, with long processing times and tedious work.

One could ask if this technology is really the right instrument needed for CH acquisition. What do we miss in transforming automatic 3D acquisition into a robust, low-cost, fast and pervasive technology? Looking at current status, we should say that the cost of the HW devices and the quality of the SW available does not allow for mass use of 3D scanning technology. Even if a considerable number of projects have produced very high-quality digital models of artworks [Levoy et al. 2000; Bernardini et al. 2002; Fontana et al. 2002; Pollefeys et al. 2001; Stumpfel et al. 2003; Bar 2004], it should be noted that most of the projects done so far have been directed by academic groups (very frequently with the aim of testing and assessing new technologies) rather than by CH institutions/experts or even practitioners. This could be considered as an empirical demonstration that 3D acquisition technology is not mature and cannot be considered a consumer technology. One critical point, in our opinion, is to reduce the complexity of the scanning process (make it easier and faster), and thus a basic issue is how to improve the automation of the postprocessing phase (minimizing the human-assisted phases). A second issue is how to reduce the cost of the HW and SW in order to transform a very costly technology into, hopefully, a quite inexpensive resource. We will discuss these issues and potential avenues for research in Sections 2, 3, and 4. Moreover, since scanning is a complex task, both erroneous use of the technology or wrong processing of the data are not so uncommon. We try to list all these potential pitfalls and errors in Section 5.

A second point in our discussion is whether an improved acquisition technology could be the only requirement for transforming a niche market into a mass market. The answer is obviously no. Another important issue is how to manage the complexity of scanned data, that is, how to support visualization and processing of huge 3D models with both extreme efficiency and simple interaction on low-cost platforms. A number of solutions have been proposed recently to manage these issues and are briefly presented in Section 6, together with some open research problems.

A final obstacle which prevents a wider adoption of the 3D medium in the CH domain is that the only application universally accepted is visual communication, that is, using 3D for producing videos, still images, or interactive navigation for products oriented to communication (museum kiosks, multimedia products, broadcasted programs, etc.). A key factor for the complete acceptance of digital 3D models in the CH domain is therefore the development of specific applications, which could transform 3D data into an essential resource for CH management/conservation/restoration; in some cases, application-oriented problems also open interesting research opportunities. We review some of these potential applications in Section 7.

2. ACQUISITION OF DIGITAL 3D MODELS

Automatic 3D reconstruction technologies have evolved significantly in the last decade [Blais 2003]. An overview of 3D scanning technologies is presented in Curless and Seitz [2000]. Among various 3D scanning systems, the more frequently used in CH digitizations are the so-called *active optical* devices. These systems shoot some sort of light over the surface of the artifact and reconstruct its geometry by checking how the light is reflected by the surface. Examples are the many systems based on triangulation (using either laser stripes or more complex light patterns produced with video projectors). These systems, together with the software process needed to transform the raw sampled data into a clean and complete 3D model, are briefly presented in the following section.[1]

Very promising, but still not very common, are the *passive optical* devices, where usually a large number of images of the artifact are taken and a complete model is reconstructed from these images. These approaches, mostly based on consumer digital photography and sophisticated software processing, are presented in Section 4.

3. ACTIVE OPTICAL ACQUISITION

3.1 Hardware Devices

The basic idea of active optical devices is to project a controlled pattern of light on the surface of the artifact. Two different approaches can be followed to reconstruct the space locations of the surface parcels intercepted by the light pattern: *time-of-flight* (TOF) or *triangulation*.

In the first case, the system computes the time elapsed between the emission of a pointwise laser beam and the detection of the return beam reflected by the surface. Knowning the direction of emission, the distance of the reflected surface on this directed line is computed by multiplying the light travel time by the light propagation speed. These TOF systems are generally used to sample large-scale artifacts (such as architectures), are characterized by a very wide working volume (from a few meters to a few hundred meters), medium sampling resolution (usually space is sampled by taking one or a few samples for a squared centimeter), medium accuracy considering the wide operation space (commercial systems guarantee error in the order of 0.5 centimeter), and long acquisition times (sampling speed is on the order of a few thousand samples per second, thus a single high-resolution range map can take up to 20-30 minutes). More recent hybrid systems, which add a phase shift control to the usual TOF approach, support much faster sampling speed (on the order of a hundred thousand samples per seconds) with approximately the same accuracy and are, consequently, getting a major share of the TOF market.

On the other hand, triangulation systems work by using an elementary geometric approach. Given a known light pattern (source and direction of propagation), images of the reflected patterns are taken with an imaging device located at a small distance from the light emitter (thus, the relative positions and orientations of the emitter and the sensor are known). The location in 3D space of each point of reflection is computed by geometric triangulation. Different light patterns are used in existing devices, such as pointwise or a planar sheet of laser light, stripe-based images produced by a digital projector. The performances of this technology are the inverse of those provided by TOF systems: accuracy is

---

[1] General note. The domain of 3D scanning, geometric processing, visualization, and their applications to CH is too large to enable us to provide a complete and exhaustive bibliography; we decided to cite just a few representative references to the literature.

usually very good (on the order of few tens of microns) and fast (usually several hundred thousand samples acquired in 1-2 seconds) sampling resolution is also good (on the order of ten samples for a squared millimeter), but the working volume is usually restricted and depends on the emitter-receiver distance (regions sampled are from 10 × 10 cm up to 100 × 100 cm, and therefore bigger objects have to be sampled by taking a patchwork of scans).

From the point of view of robustness, speed and accuracy, the status of the commercial scanning devices can be considered mature for most applications. The major problem most often raised by practitioners in the CH field is their excessive cost, especially when compared with the low budget which characterizes most CH-related activities. If we consider that the first commercial scanning devices appeared around 20 years ago, we have a technological evolution which is much slower that those we are used in the consumer IT market and, unfortunately, a cost evolution which has been nearly negligible since prices have dropped only approximately 30% in the last ten years. The slow technical advance and the minor price drop is due to the fact that 3D scanning is still a niche market: since the most successful devices sell a few hundred units per year, there are not sufficient revenues for a massive R&D effort and large-scale production savings. The lack of a killer application with a wide market is the key problem. Even if examples of success stories exist (just consider the use of 3D scanners in the movie industry), we still lack a really wide application area.

Finally, most of the existing systems consider just the shape acquisition, while a very important aspect in CH applications is *color sampling*. This is the weakest feature of contemporary technology since those scanners that acquire color information produce low-quality color sampling (with a notable exception of the technology based on multiple laser wavelengths [Blais et al. 2005], unfortunately characterized by a very high price). Moreover it should be noted that existing devices sample only the apparent color of the surface and not its reflectance properties that constitute the characterizing aspect of the surface appearance. There is wide potential for improvement of current technology to cope with these features.

3.2 Processing Scanned Data

Unfortunately, most 3D scanning systems do not produce a final, complete 3D model but rather a large collection of raw data which have to be postprocessed. This is unfortunately the case of all active optical devices. A complete scan of an artifact requires the acquisition of many shots taken from different viewpoints to gather complete information on its shape. Each one produces a so-called *range map*, that is, a single view of the object which encodes the sampled points' geometry. The number of range maps depends on the surface extent of the object and on its shape complexity. Usually we sample from a few tens up to a few hundred range maps. Range maps have to be processed to convert the data encoded into a single, complete, nonredundant, and optimal 3D representation (usually, a triangulated surface). Examples of digital models produced with 3D scanning technology are presented in Figures 1 and 2.

The structure of the postprocessing pipeline is presented in an excellent overview paper by Bernardini and Rushmeier [2002]. Some new algorithms have been proposed since the publication of this review paper, but the overall organization of the pipeline is not changed and therefore the description is still valid. The processing phases (usually supported by commercial tools) are the following.

—Range Maps *Alignment*. By definition, the range map geometry is relative to the current sensor location and has to be transformed into a common coordinate space where all the range maps lie well aligned on their mutual *overlapping regions* (i.e. the sections of the range maps which sample the same portion of the artifact surface).

—Range Maps *Merge* (or *Reconstruction*). A single, nonredundant triangulated mesh is built out of the many partially overlapping range maps. This processing phase reduces the redundancy (after

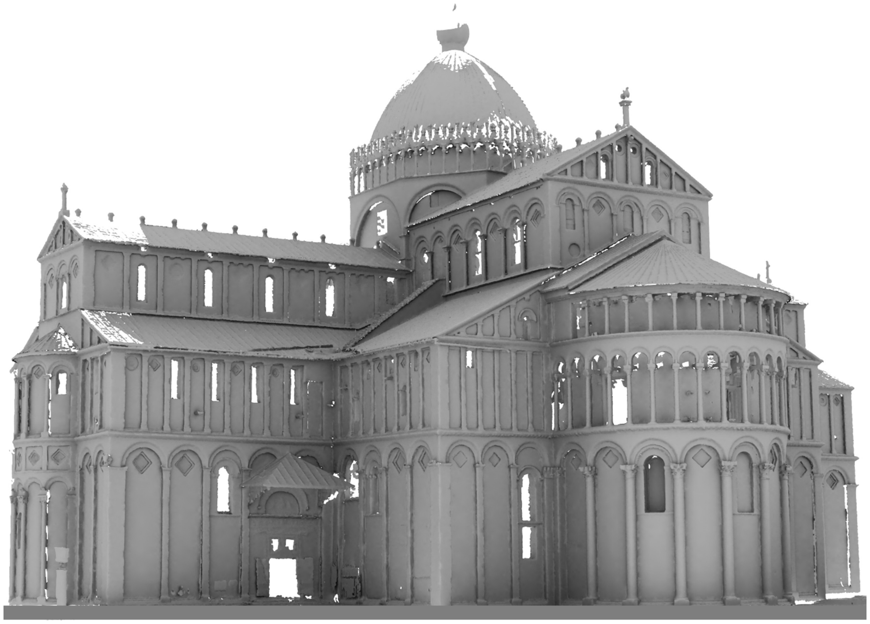

Fig. 1. A digital model produced with time-of-flight 3D scanning (Pisa Cathedral, Italy).

Fig. 2. A 3D scanned model (laser triangulation) of a medieval capitol (S. Matteo Museum, Pisa, Italy). On the left is the digital model rendered as a standard grey surface, while the image on the right is rendered after mapping color detail to the 3D mesh.

merging, each surface parcel of the artifact will be represented by just one geometric element).

—Mesh *Editing*. The goal of this step is to improve (if possible) the quality of the reconstructed mesh, for examples, reducing noisy data or fixing unsampled regions.

—Mesh *Simplification*. The huge complexity of the model obtained usually has to be reduced in a controlled manner by producing discrete a Level-Of-Detail (LOD) or multiresolution representations.

—*Color Mapping*. The information content is enriched by adding color information (an important component of the visual appearance) to the geometry representation.

All these phases are supported either by commercial [INUS Technology 2003; InnovMetrics 2003; Raindrop Geomagic 2003] or academic tools [Curless 2006; Callieri et al. 2002, 2003; Franken et al. 2005].

The alignment task is usually the most time-consuming phase of the entire 3D scanning pipeline due to the substantial user contribution required by current systems and the large number of scans which constitute real scanning results. The initial placement is heavily user-assisted in most of the commercial and academic systems (requiring the interactive selection and manipulation of the range maps). Moreover, this kernel action has to be repeated for all the possible overlapping range map pairs (i.e., on average, 6–8 times the number of range maps). If the set of range maps is composed hundreds of elements (the scanning of a 2-meter tall statue generally requires from 100 up to 500 range maps, depending the on shape complexity and the sampling rate required), the user has a very complex task to perform. For each range map, the user most find which are the partially overlapping ones, given this set of overlapping range maps, determine which one to consider in pairwise alignment (either all of them or a subset); finally, process all those pairwise initial alignments. If not assisted, this becomes the bottleneck of the entire process.

Improved management of a really large set of range maps (from 100 up to 1000) can be obtained both by providing a hierarchical organization of the data (range maps divided into groups with atomic alignment operations applied to an entire group rather than to the single scan) and by using multiresolution representation of the data to make rendering and processing more efficient. Moreover, since the standard approach (user-assisted selection of each overlapping pair and creation of the correspondent alignment arc) becomes impractical on a large set of range maps, tools for the automatic setup of most of the required alignment actions have to be provided (see Section 3.3).

Reconstruction from point samples or range maps has been one of the more active field research on scanning-related subjects in the last few years. A very important feature of a reconstruction code is to perform a *weighted integration* of the range maps (or pointset) and not just join them. Since we usually have a high degree of overlap and sampled data are noisy, a weighted integration can significantly improve the accuracy of the final result by reducing the impact of most of the noisy samples. Another important feature of a reconstruction code is the capability to fill up small holes (i.e., regions not sampled by the scanner, see Section 3.4). Finally, since reconstruction algorithms require a very large memory footprint on a big dataset, they have to be designed to work locally on subsections of the data, loading only the data subset involved in the generation of a single portion of the final results (out-of-core).

If we chose a reconstruction kernel region equal or smaller than the intersampling distance used in scanning, the reconstructed models usually become huge in size (i.e., many millions of faces). Most applications require significant complexity reduction in order to manage these models interactively. Two problems arise when we try to simplify such models: we need solutions working on external memory to cope with these big models [Cignoni et al. 2003]; and simplification has to be accurate [Garland and Heckbert 1997; Hoppe 1999] if we want to obtain high-quality models and accurate interactive visualization (possibly, based on multiresolution data representation schemes). These issues are presented in Section 6.

Finally, we have to support the reconstruction of textured meshes from a sampling of the object's surface reflection properties. The most common approach is the acquisition and mapping of the *apparent color* (reflected color, illumination-dependent) using digital photo cameras. This is the easiest and more practical approach, since setting up a controlled lighting for a more sophisticated acquisition of the reflection properties of the object's surface (BRDF acquisition [Lensch et al. 2003]) is often impossible or impractical for CH artifacts, especially when acquisition has to be done in museum conditions. The issues involved in color acquisition and management are discussed in Section 3.5.

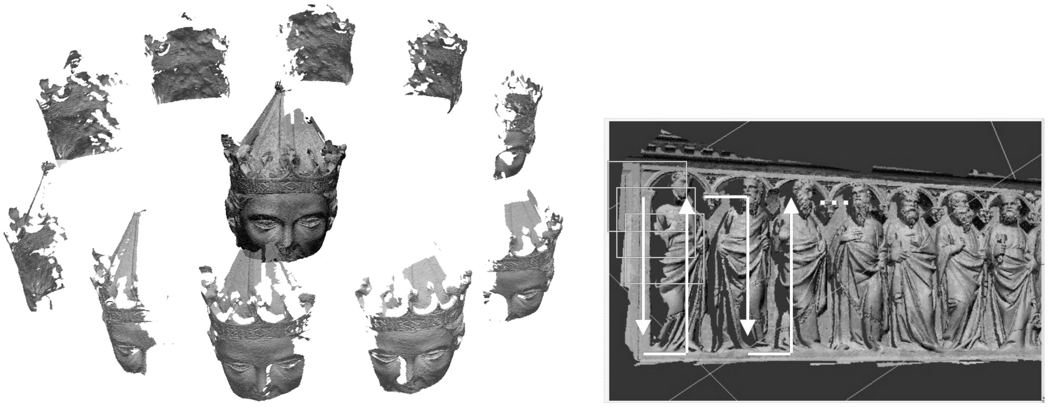

Fig. 3. Range maps are taken in a rowwise order: an example of circular stripe around a statue's head (left); an example of raster-scan scanning order adopted for the acquisition of a bas-relief (right).

3.3 Making Alignment an Automatic Process

Solutions for completely automatic scanning systems have been proposed, but either those systems are based on the use of complex and costly positioning/tracking machinery (e.g., see Levoy et al. [2000]) or they adopt passive silhouette-based approaches which do not extend well to medium or large-scale artifacts. An alternative approach is to design new solutions for the classical scanning pipeline which would transform it into a mostly unattended process. In particular, the range maps alignment phase is the only task where a considerable human intervention is still required. Several papers have proposed methods for automatic alignment usually based on some form of shape analysis and characterization (see Campbell and Flynn [2001] for a survey paper).

The general alignment problem can be made more tractable by considering some assumptions which usually hold in practical use. While designing a new solution [Fasano et al. 2005], we started from a few initial conditions directly gathered by our experience in 3D scanning. First, the detection of the pairs of overlapping range maps can be reduced to a simpler task, once we notice that 3D acquisition is usually done by following simple scanning pose paths. Users usually acquire range maps in *sequences*, following either a *vertical*, *horizontal*, *raster-scan* or *circular* translation of the scanning system (see Figure 3). The different types of sequences share a common property: they contain an ordered set of n range maps, such that range map R_i holds a significant overlapping with at least R_{i-1} and R_{i+1}. Vertical, horizontal or raster-scan stripes are often produced when acquiring objects like bas-reliefs, walls, or planar-like items. Circular stripes are more useful when acquiring objects like statues, columns, or cylindrical-shaped objects.

If we can assume that the acquisition has been performed using one of these stripe-based patterns, then we may search for overlapping and coarse registration on each pair of consecutive range maps (R_i, R_{i+1}). An automatic registration module can processes each couple (R_i, R_{i+1}) to produce in output the rototranslation matrix M_i that aligns R_{i+1} to R_i. Matrix M_i can be computed by applying some basic geometric processing to the two range maps: feature points can be detected by evaluating a shape descriptor on the two meshes (or point set); potential corresponding feature points pairs can be detected by adopting RANSAC-like approaches. From these possible point pairs (Figure 4), we can select the one which produces, after ICP alignment, the matrix M_i which preforms the best alignment (see Fasano et al. [2005] as an example of this type of solution). As an alternative to geometry-based solutions, other

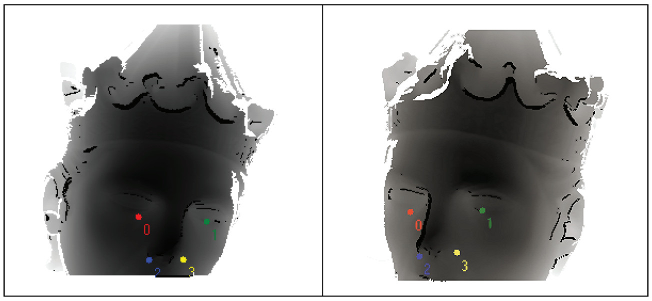

Fig. 4. The four matching point pairs selected by the automatic alignment algorithm on two consecutive range maps.

researchers have proposed an image-based approach for the selection of the correspondent point pairs [Bendels et al. 2004] by looking to image features in the RGB channel associated to the XYZ samples (under the assumption that the scanner produces self-registered XYZ RGB shape and color samples).

The subset of registration arcs is not complete when we restrict the search to consecutive pairs in linear sequences (since we usually have many other potential overlaps between range maps), but it is sufficient for the application of an intelligent ICP-based solution. A smart alignment system can be designed to be able to complete the needed arcs (interconnecting R_i with all the overlapping range maps not just R_{i-1} and R_{i+1}) in an automatic manner. Once we have an initial alignment (even if defined on a very sparse set of arcs), a simple process based on a *spatial indexing* data structure allows us to complete it. This spatial index encodes for each 3D cell the set of range maps passing through that subvolume, allowing for an easy automatic detection of the potential overlaps and for running ICP-based alignment on those corresponding range map pairs. Given the occupancy grid information, and once a single alignment arc is provided for each range map, the alignment system may introduce all needed arcs (in a completely unattended manner) by selecting and processing only those which satisfy a minimum-overlap factor.

The automatic registration approach just sketched [Fasano et al. 2005] has been tested on many complex scanning campaigns (where each range map is usually affected by noise, artifacts, and holes). An example concerning a bas-relief in Figure 5 is shown, whose approximate length is 2.5 meter. In this case, two raster-scan (snake-like) stripes were acquired and processed, for a total of 117 meshes (about 45.5M vertices). The overall automatic alignment requires usually much less than the raw scanning time, and it is therefore sufficiently fast to run in the background during the acquisition, processing all the scans in sequence as soon as they are produced.

Solutions similar to those presented in this section are unfortunately still not provided by commercial software. They would significantly speed/up the processing of scanned data, reducing the manpower required and the overall costs.

3.4 Coping with Incomplete Surface Samplings

According to the experience of people who performed 3D scanning of complex artifacts, obtaining a complete sampling of the artifact surface is often impossible [Levoy et al. 2000]. The reasons why we usually end up with an incomplete scan are various: presence of self-obstructing surfaces; presence of small cavities or folds; sections of the surface which are not cooperative with regard to the optical scanning technology adopted (highly reflective materials like polished metals, transparent components such as glass or gems, surfaces which do not reflect the emitted light probe, etc).

In all these cases, we have to decide if the model has to be completed or if we have to leave it incomplete. In the CH domain, we are usually asked to produce models which contain only sampled

Fig. 5. The coarse alignment of the bas-relief (top) and the final model (middle); alignment has been performed in a nonattended mode.

geometry, that is, the use of software solutions which fill up the gaps is not allowed. On the other hand, incomplete digital models perform very poorly in visualization since the holes usually attract the observer's attention much more that the other clean parts. For visualization application, a clean sampled surface obtained by closing all the gaps with plausible surface patches is therefore needed. A very nice extension to available data formats would be an attribute that could allow us to differentiate between sampled and interpolated geometry, making it possible to make visually evident those two different data components. This would be an interesting addition to provenance data encoding.

Gap filling can be obtained by two orthogonal approaches: *volumetric* and *surface-oriented*. In the first case, the holes can be filled at reconstruction time, for example, by enhancing the volumetric reconstruction approaches based on discrete volume distance with a diffusion process which extends the distance field in regions not covered by scanned samples [Davis et al. 2002] or after the reconstruction by putting the model in a volumetric grid and devising solutions capable of producing a water-tight surface [Ju 2004; Podolak and Rusinkiewicz 2005]. Unfortunately, these approaches make it very hard to differentiate sampled and interpolated geometry. Moreover, methods based on volumetric diffusion are very complicated to use because steering the diffusion process to obtain a plausible completion surface is not easy; a time-consuming trial-and-error process is usually needed to find the parameters that best fit a given dataset.

On the other hand, *geometric processing* solutions can be devised to detect and fill unsampled regions. Surface-oriented approaches try to detect and close holes preserving triangle shape and curvature [Liepa 2003; Tekumalla and Cohen 2004]. The problem is not simple because we need geometrically robust solutions that are able to close any gap with a surface patch which should share curvature continuity with regard to the adjacent regions. Some open issues are the necessity for dealing with self-intersections and isles, and the algorithm speed and robustness, and the difficulty in creating fully automatic but reliable methods.

Moreover, in some cases, a basic concept of curvature is not enough since a missing region can contain surface features (such as a given texture or a carved detail) that we may want to reproduce in a way conforming with adjacent regions. So-called *surface inpainting* methods have been proposed to support an intelligent cut-and-paste of surface detail from one completely sampled region to a partially sampled region [Sharf et al. 2004; Bendels et al. 2005].

3.5 Mapping Complex Photographic Detail on 3D Models

CH is an obvious example of an application domain which requires not just shape sampling but also an accurate acquisition and management of color data. Accurate approaches for sampling the surface reflection characteristics of an artifact have been proposed (e.g., BRFD sampling [Lensch et al. 2003]) but are still too complicated to be massively applied to the CH field where we do not usually work in controlled lab conditions but rather in crowded museums.

For most practical cases, a simpler approach is still widely used: the so-called apparent color is acquired and mapped to the 3D model. A series of pictures can be taken with a digital camera, trying to avoid shadows and highlights by taking them under a favorable lighting setup; these photographs are then stitched onto the surface of the object. However, even in this simpler case, the processing needed in order to build a plausible texture is not straightforward [Callieri et al. 2002]. Naive mapping of apparent color on the mesh can produce severe discontinuities that are due to the varying illumination over the surface sampled by the photos. Many different approaches have been proposed to reduce the aliasing and to produce seamless color mapping. We cite here only some representative papers: using the range intensities produced by some active optical scanning devices to correct the color information [Umeda et al. 2005], detecting and removing cast shadows [Troccoli and Allen 2005], which are usually a major problem in color acquisition in outdoor scenes, devising methods for computing the inverse illumination, that is, recovering approximate surface reflectance from a sampling of the real surface under known illumination conditions [Rushmeier and Bernardini 1999; Debevec et al. 2004], and finally, applying a multivariate blending approach, which allows for the inclusion of all possible heuristics for image content characterization to produce, for each point on the surface, an optimal weighted average of the available pixel samples [Callieri et al. 2008].

A basic problem in managing color information is how to register the images with the geometric data in a time-efficient way. Once intrinsic (lens focal length and distortion) and extrinsinc parameters (view specs) have been computed for each image by registering it onto geometry, many approaches exist to map the color infomation on the geometry, based on mesh parameterization or color-per-vertex encoding. The bottleneck in the color pipeline is the image registration step, a complicated time-consuming phase which requires substantial intervention of a human operator since available approaches are based on the selection of several corresponding point pairs which link each 2D image to the 3D mesh [Tsai 1987]. Analogous to the range map alignment problem, this phase should be solved automatically as much as possible and assisted. No fully automatic and robust approach has been proposed yet for the general problem (i.e., a large and complex object where each image covers only a subset of its overall extent). Automatically finding the correspondences between 2D images and a 3D mesh is not an easy task, because sampled geometry presents less features than sampled color. A possible research direction could be to move from the usual search for image-to-geometry correspondences to a context were we use both image-to-geometry and image-to-image correspondences. As shown in Bendels et al. [2004] and other recent works, finding correspondences in images is simpler than detecting image-to-geometry correspondences. Since a large number of image-to-image correspondence pairs can be detected in an automatic manner, we can deploy that information to speed up the overall image registration process

Fig. 6. The David model is shown with color mapping. On the left is the prerestoration status (61 images mapped), while the postrestoration status is shown on the right (another set of 68 images). The two-colored David models are rendered in real time with the Virtual Inspector system.

or to solve those cases where a single image covers a region where the surface has insufficient shape features to allow an accurate selection of image-to-geometry correspondences.

In recent research, we designed a new tool to support image registration, TexAlign [Franken et al. 2005], which builds a graph of correspondences, where the 3D model and all the images are represented as nodes and a link is created for any correspondence defined between two nodes (implementing either an image-to-geometry or an image-to-image correspondence). This graph of correspondences is used (a) to keep track of the work done by the user, (b) to infer automatically new correspondences from the instantiated ones, and (c) to find the shortest path in terms of the number of correspondences that must be provided by the user to complete the registration of all the images. The goal is to assist the user in the management of large set of images. This system has been used recently to map a complex photographic sampling (61 + 68 images mapped on the David model, see Figure 6). Adding a filter for automatic image-to-image correspondences detection to *TexAlign* could make image registration a nearly automatic process.

Another open issue concerning color mapping is how to manage the mapping of very high resolution color sampling on 3D meshes. Very high resolution samplings of surface reflectance can be obtained with either a high-end scanning system [Blais et al. 2005] or by adopting standard photographic devices. Currently, the resolution of prosumer digital cameras is at least an order of magnitude higher than the resolution of triangulation-based scanning devices. A medium-size object (e.g., a statue) can be easily sampled with 50-100 images, which means some hundred million pixels. The need to manage such texture resolution in real time opens several issues concerning optimal texture mapping representation, multiresolution representation of geometry and texture data, and efficient visualization.

Obviously, we envision future systems able to encode not only the reflected radiation (apparent color), but able to sample the reflection properties of the surface (i.e., its BRDF). To make those solutions practical, we need to improve current methods in terms of ease of operation, reduced dependency and highly controlled lighting environment, and reduced processing complexity.

Other promising approaches support sophisticated sampling of the reflection properties by adopting image-based rendering approaches. We have a spectrum of possibilities from the one-view and multiple-lighting approach of Polynomial Texture Maps (PTM) [Malzbender et al. 2001] to the more demanding multiview approach granted by Light Field rendering [Levoy and Hanrahan 1996]. In all these cases (see Levoy [2006] for an overview), we get rid of the 3D model and adopt approaches which sample the artifact with a multitude of images. To produce any view requested at visualization time, we process/interpolate the knowledge granted by the sample image set. In most cases, this implies a huge quantity of image data to be acquired (thus long acquisition times are needed), stored, and accessed in real time, for example, making Web-based visualization difficult.

4. PASSIVE OPTICAL ACQUISITION OF 3D MODELS

One of the common issues in the common practice of CH 3D scanning is the high cost of the HW devices, nonsustainable for low-budgeted projects since the price of high-quality laser-based technologies spans in the range of 64,000–140,000 Euro.[2] Passive optical acquisition methods, proposed since the early stages of 3D scanning research, are a cheaper but lower-quality alternative, performing 3D reconstruction from a simple sequence of high resolution digital photos of the artifact. Some of them have been considerably improved recently and show some interesting potential for a wider diffusion.

4.1 Carving from Silhouette

Both geometry and color of an artifact can be reconstructed by taking photos of the artifact while it performs a calibrated rotation in front of the photocamera. Early methods were based on the extraction of silhouettes and carved the 3D space by intersecting the perspective conoids bounded by those silhouettes. The main limitation of this approach is that it reconstructs the *visual hull* rather than the real shape of the artifact since concave regions are not sampled by any silhouette. This limitation has been recovered by recent solutions which adopt either stereo matching or other approaches to reproduce faithful geometric detail on concave regions of the artifact surface (e.g., see Esteban and Schmitt [2003]). Moreover, these methods are able to collect apparent color data and to automatically encode that information in the mesh produced in output.

Among the positive characteristics of this technology is that it is based on digital photographic technology. HW is therefore extremely cheap. The impressive increase in resolution and optical quality of digital cameras has an immediate impact on the accuracy of the 3D models that can be produced. Unfortunately, only a few systems were on the market at the end of the 90s, without a remarkable success. This could be due both to the early stage of the technology (they produced visual hulls, carving of concave regions was not supported) and to the inherent low flexibility of the silhouette-based method: it can acquire only small-scale artifacts since it has to put them on a rotary platform and frame the entire artifact in each photo.

4.2 3D from Photo Streams

The recovery of three-dimensional structure out of a sequence of photos is a well-studied field in the computer vision literature, but, until recently, it was difficult to really harness the results of the many

---

[2] The recent introduction of a low-cost laser device sold for \$ 2500 is remarkable news which could have a giant impact on the domain (it is a triangulation-based system, https://www.nextengine.com/).

Fig. 7. A 3D model reconstructed from just a sequence of 100 high resolution photos. The reconstruction was performed using the tools developed within the EU IST NoE *Epoch*.

presented algorithms in a single working framework [Pollefeys et al. 2001; Vergauwen 2006]. These approaches are based on the search of a small set of correspondences between the processed images; these correspondences (usually on the order of tens or one hundred) identify some feature points in the scene as seen from different points-of-view. Depending on how these corresponding image points are located in the different pictures, the 3D position of these feature points and the orientation of the camera are recovered. Starting from these few sparse points, a dense depth-range map can be reconstructed from each image by interpolating these recovered points and applying stereo-matching techniques on the pixel in the in-between regions.

A result of the application of this technology is shown in Figure 7. The model shown in Figure 7 was reconstructed by around one hundred 6M pixel photos of the *Arc du Triomphe* (Paris, France), shot all around the monument. This particular reconstruction was performed with the ARC 3D Web service (http://www.arc3d.be/) developed within the EC IST Network of Excellence *Epoch* [Epoch 2006]. Users registered to this Web service can simply upload their photo sequences and a remote server automatically converts the photos into a sequence of aligned range maps (one for each photo) which can be downloaded and processed by the user.

The advantages of this new approach are quite evident. The only hardware required is a simple good quality digital photographic camera, and the scanning process requires just taking a reasonably large number of photos all around the object. On the other hand, this approach still exhibits a lower geometric precision than the well assessed laser-based 3D scanning technologies. Moreover, since the reconstruction process is based on the detection of corresponding features on consecutive photos, it encounters difficulties in the reconstruction of artifacts with large flat and uniformly colored parts that does not exhibit evident features to be recovered.

5. 3D ACQUISITION PITFALLS AND DEFICIENCIES

While 3D scanning hardware and software technologies have considerably evolved in the last years and their use in the Cultural Heritage field has gained acceptance, these technologies are far from being easy to use and error-free. Many problems can arise during a scanning session; most of these issues are related to various deficiencies and inadequacies in both the technologies and the background of the operators. Some of these problems have already been briefly mentioned in the article; we mention and discuss them in detail in this section.

5.1 Technological Deficiencies

First of all, 3D scanning technologies are not able to capture many kinds of materials that occur in the CH field: transparent surfaces (like glass and precious stones), mirrored, polished, and very shiny objects (like metals), fluffy and fuzzy substances like feathers and furs. All these materials are quite difficult to sample with the current off-the-shelf optical technologies. For some of the previous cases, experimental research projects have shown the feasibility of technological solutions that overcome these limitations but their practical applicability to real projects has still to be assessed.

Moreover, as already noted, most of current scanning technologies focus only on shape, ignoring the issues related to the acquisition of the optical appearance of the scanned surface. Even when the adopted technologies are able to capture the color (e.g., by mapping photos over the surface), in almost all the cases, this leads only to the acquisition of the apparent reflected color without considering nontrivial reflectance properties of the surface (e.g., its shininess). The acquisition of a correct representation of the reflectance properties of a surface together with its 3D shape is still the domain of research experiments and it is not an off-the-shelf technology.

Another issue of 3D scanning is the cost of the hardware equipment: usually it is very high and it can become prohibitive for many low-budgeted cultural heritage institutions.

On the positive side, we have to note that the research in this field is very active so we can easily hope that, in the near future, the capabilities of 3D scanning hardware will improve gradually eliminating these shortcomings.

5.2 Misuse and Pitfalls of the Technicians

The people involved in a scanning Session can be roughly partitioned into two sets: the technical staff, who actually perform the scanning task with a not-so-strong CH background, and the scanning purchaser, who is usually more strictly related to the CH world but often lacks a deep knowledge of the technology side. Obviously there are some notable exceptions, typically in most of the successful 3D scanning projects.

5.2.1 *Wrong Choice and Use of the Hardware.* As we have sketched in the previous sections, there are many possible technologies available, each one with its own pros and cons. External constraints (e.g., the availability of a given device or the previous knowledge of a specific technology) might affect the choice of the preferred hardware leading to the selection of a nonoptimal technology.

5.2.2 *Wrong Data Postprocessing.* The desire to provide clean and professional results to the scanning purchaser often causes the undocumented (and often unnecessary) editing of the acquired data. 3D scanning technologies are often unable to completely recover the whole surface, leaving holes and unsampled regions in the final model. Any editing action over the originally acquired data should be documented in the most evident way in order to make it possible to distinguish between ground-truth sampled data and the parts that have been added or heavily edited by the subsequent geometric processing phases. Excessive data smoothing or inaccurate simplification are clear examples of actions which can destroy data accuracy.

5.3 Misuse and Pitfalls of the Scanning Purchaser

On the other hand, lack or incomplete technical knowledge of 3D scanning can cause various inconvenience in the overall process of creating trusted digital reproduction of works of art.

5.3.1 *Evaluating and Assessing.* Being able to evaluate in an objective way the final quality of a 3D scanned model is a basic resource for the wide adoption of this technology. Conversely, to more traditional media like photography where established protocols and procedures exists and the quality can be determined by an analysis of the (digital) photo and its provenance data, the assessment of 3D media is much more complex. Quality depends on a large number of factors: scanning HW characteristics, modality of use of the scanning device, pipeline and parameters of the SW postprocessing which transforms the raw data into the final 3D model. Therefore, the whole pipeline including HW and SW used to produce the digital object has to be considered when assessing the quality of the digital model. Provenance data are therefore more complex than the ones characterizing photography. While common practices for the use of general 3D visualization in the research and communication of Cultural Heritage have already been discussed in the London Charter [Lon 2006], there is still an absence of clear protocols for the management of 3D scanning technologies and data. Some attempts should be made to define the concept of provenance data for 3D models in a sufficiently general and standardized way.

The availability of free and standard tools supporting the quality assessment (ideally, by considering both geometry and reflectance data) would be of paramount value, especially for the usually nontechnical people who order a scanning session. This is an area that desperately needs research. Even a well-informed consumer at this point does not have the tools and standards available to make a purchasing decision and a good assessment of the results that are produced by a (usually expensive) project.

5.3.2 *Presenting and Preserving.* An interesting point of discussion is what are the most fruitful uses of the acquired data, but this will be the subject of a deeper analysis in the following Sections 6 and 7. Another consideration that is often neglected is the long-term preservation of these data. While for traditional media like written material, photography, and films, the issues about the preservation, classification, and future access to the acquired data is a rather well studied subject, for this kind of 3D digital data there are no established practices. Without going into detail, let us just mention here the issues related to the format of the delivered 3D data: often the results of a 3D scanning session are delivered only as some files in a closed format, accessible trough proprietary software whose longevity is not guaranteed. In these cases, a serious risk is that the data produced are buried in a hidden format without the possibility of being utilized in the near future.

6. VISUAL PRESENTATION OF HUGE MODELS IN LOCAL AND WEB-BASED CONTEXTS

Some issues arise from the very dense sampling resolution generated by modern scanning devices. Being able to sample on the order of ten points per-squared-millimeter or more (in the case of triangulation-based systems) is of paramount value in many applications which need a very accurate and dense digital description. On the other hand, this information is not easy to process, render, and transfer, therefore, excessive data density may become a problem for many applications. This created much intense research on simplification and multiresolution management of huge surface meshes [Garland and Heckbert 1997; Hoppe 1999; Cignoni et al. 2003] and interactive visualization methods with both mesh-based [Cignoni et al. 2004, 2005] and point-based solutions [Rusinkiewicz and Levoy 2000; Botsch et al. 2002] proposed in the literature.

6.1 Simplification and Multiresolution Management of Huge Models

Data complexity can be managed by adopting a *data simplification* approach to reduce the data resolution at the expenses of a loss of geometric accuracy. Many solutions have been proposed for the accurate simplification of 3D triangulated surfaces, usually based on an iterative elimination of selected vertices or faces driven by an error-minimization cost function. This approach allows the construction of any level of resolution we need often with a rather expensive computation (from a few seconds to a few hours, depending on the solution used and the complexity of the initial surface). Simplification is very handy to produce models which fit a specific application requirements (e.g., a simple small model for a Web presentation which should be downloadable in a given short time or a model used for a very accurate rapid reproduction by a 3D printer).

Another approach is to store not just the final simplified model but all the intermediate results reached during the iterative simplification. All these intermediate results have to be encoded in an efficient data structure (*multiresolution* encoding) that will allow our interactive application to extract different resolution on the fly within very short time compatible with real time applications (e.g., visualization). A view-dependent variable resolution representation can be produced for each frame from these schemes according to the current view specification (e.g., higher resolution for the portions in the foreground, progressively lower resolution for data in the background) and the requested visualization accuracy.

Recent research on multiresolution schemes has produced a number of solutions based on higher granularity than previous methods (i.e., instead of focusing on single triangles, patches of triangles become the elementary entity). These solutions provide for the management of huge 3D scanned models at an interactive frame rate on consumer PC's [Cignoni et al. 2004, 2005]. *Virtual Inspector* is an example of a visualization system which adopt this approach for supporting the inspection of large, complex 3D models in real time [Callieri et al. 2008].

We mentioned simplification and multiresolution technologies in this section not because we consider them as open field for research (in our opinion, these are nearly mature technologies) but to stress with the CH community the point that these solutions are ready and can be proficiently used. One of the more common negative concern raised by practitioners against 3D scanning is the size of the models obtained, which according to this common believe makes them unusable for CH applications. This is not true since the availability of simplification and multiresolution technologies can successfully cope with the data complexity of 3D scanned data.

6.2 Usability of Virtual Heritage Worlds

Ease of use of visual CH tools oriented to ordinary people (still not very competent with 3D graphics and computer games) is an important factor for their success. One of the most complicated actions to perform is directing navigation in the virtual space. Therefore, free navigation should be requested only in those cases where this action really adds something to the learning experience. The risk is to have the visitor lose orientation (e.g., discovering himself lost in sidereal space, maybe just because he turned his back to the scene) and abandoning the system.

Other important characteristics of a visualization system are its flexibility, completeness, and configurability. To fulfill these objectives developers could design complicated systems characterized by a very complete set of functionalities and involute GUI (e.g., consider commercial modeling systems). Conversely, while designing our *Virtual Inspector* tool [Callieri et al. 2008] as a system oriented to non-expert users (e.g., museum visitors), our approach was to define a restricted set of functionalities and to provide the tool with an easy installation interface for the selection of the subset of these functionalities that the designer of the specific installation wants to use (e.g., in a museum kiosk).

Le braccia della statua, oggi ridotte a moncherini, dovevano essere disposte in modo da attenuare l'andamento diagonale delle pieghe sul petto mentre le mani dovevano recare le "insignia imperalia", verosimilmente in metallo e dunque probabile causa delle prime mutilazioni.
Gli oggetti-simbolo del potere quali lo scettro e il globo, accompagnati alla corona, caratterizzano infatti l'iconografia regale e imperiale.

La testa dell'imperatore è spezzata rispetto al resto del corpo e risulta male impostata sul collo sottostante. E' stata notata una certa sproporzione rispetto al resto della figura ed una distanza stilistica tra la fattura minuta del volto in confronto al resto del corpo e alle immagini dei "consiglieri"; si peraltro è ipotizzato per la testa l'uso di un materiale non omogeneo, fino a mettere in dubbio la pertinenza alla statua o che le singole parti siano da attribuire a mani diverse. Non sembra tuttavia probabile che la testa fosse destinata ad altra figura perché al livello della parte superiore del petto e sulle spalle il marmo è stato scalpellato e dunque l'andamento irregolare della frattura non offre una prova sufficiente. Anche la presunta differenza di materiale richiederebbe di essere verificata attraverso l'analisi mineralogico-petrografica: infatti la venatura grigia è generalmente discontinua, più o meno diffusa in zone diverse dello stesso blocco.

L'affinità con la testa di Arrigo giacente in Cattedrale -seppure non unanimamente riconosciuta- sembra confermare l'appartenenza della statua in trono al monumento funebre. Il problema attende necessari chiarimenti: proprio sulla testa coronata si basa di fatto l'identificazione di tutto il gruppo con l'imperatore e il suo seguito, nonché la tradizionale attribuzione alla tomba di Arrigo.

Fig. 8. Virtual Inspector. The Arrigo VII statue rendered with active hot spots (top); a short pop-up panel with a short description of the missing hand appears when the mouse passes over the hotspot located on the hand (left), or a more complex page associated with the hotspot on the neck (right).

6.3 Not Just 3D Data: Adding Other Knowledge

Visualization tools usually focus on just the visual presentation of the shape characteristics of the artifact. This is not enough if we want to provide a comprehensive presentation of the global knowledge available on the artifact. On the other hand, the 3D shape can become some sort of visual 3D map which allows us to integrate, link, and present all the available information in an intuitive and visually pleasing way. The goal is, therefore, to transform standard 3D browsers into tools able to connect the shape with all the available multimedia (MM) data linked with the artifact. This opens up the discussion to the wider topic of how to support and implement data annotation on 3D models.

Hot spots can be a very handy resource to associate multimedia data to any point or region of a 3D model. This resource allows for the design of interactive presentations where the 3D model becomes a natural visual index to historical/artistic information, for example, using standard HTML format and browsers (see Figure 8). But hot spots are only a first step that implement a pointwise association metaphor. We should devise more flexible and powerful instruments to associate information to 3D meshes, for instance supporting links between subregions of the surface to MM documents.

Moreover, visualization instruments should be extended to provide tools to integrate and enrich the data associated with the 3D digital model, supporting a democratic and cooperative approach such as that at the base of the Wikipedia effort.

6.4 Distributing 3D Data on the Web

In the modern world, we cannot avoid considering the issues related to the Web-based access and distribution of 3D data. The peculiar aspect of scanned data with regard to. other 3D models is the average size of the models. To make these data usable on the Web, we should either deploy efficient geometric compression technology [Rossignac 2004] or adopt remote rendering approaches [Koller et al. 2004]. A Web-based application which could boost the usage of 3D scanning considerably is mentioned in Section 7.5.

7. WHICH USES OF 3D MODELS?

The question one can ask at this point is does if the reason for not having a more wide spread use of 3D graphics, in CH applications depend just on quality and the cost of the acquisition technologies. Technological issues play a role, but in our opinion, this is not the only (or major) reason for the scarce use of 3D graphics in CH applications. Another key factor is the lack of some killer applications in the realm of the daily needs of our CH colleagues. Visualization is the most diffuse utilization of digital 3D models, but unfortunately, in many cases virtual presentation is still considered by CH practitioners as a medium yet to reach the public rather than a basic instrument for CH research and conservation. In our opinion, there are many possibilities for the development of useful tools oriented to the CH domain. This is mostly an applied research or research transfer activity (designing tools according to specific needs), but, in some cases, specific needs can also open interesting issues concerning basic research. We introduce some of these possible applications in the following sections.

7.1 Computer-Aided Restoration

Restoration of CH artifacts can be positively affected by the use of accurate 3D models. Today restoration is a very complex task in which multidisciplinary skills and knowledge are required. A complex set of investigations usually precedes the restoration of a valuable artwork: visual inspection, chemical analysis, different type of image-based analysis (RGB or colorimetric, UV light reflection, X-Ray, etc.), structural analysis, historical/archival search, etc. These analysis might also be repeated over time to monitor the status of the artwork and the effects of the restoration actions. An emerging quest is how to manage all the resulting multimedia data (text/annotations, historical documents, 2D/3D images, vectorial reliefs, numeric data coming from the analysis, etc.) in an integrated framework, making all information accessible to the restoration staff (and possibly to experts and ordinary people as well). The final goal is to guide the restorer in the choice of the proper restoration procedure by the evidence of the analysis performed and to assess in an objective manner the results of the restoration (to compare the prerestoration and postrestoration status of the artwork, to document the restoration process). Since most of the information is directly related to spatial locations on the artwork surface, 3D Three-dimensional models can be valuable media to index, store, cross-correlate and obviously visualize all this information. 3D models can also be a valuable instrument in the final assessment phase, supporting the interactive inspection of the multiple digital models (pre and postrestoration status) to check possible shape and color variations.

Moreover, a number of investigations can be performed directly on the digital 3D model by adopting computer-based simulations or computations. This has been done in the past to assess the static and structural status of buildings or sculptures in order to detect risky conditions due to an exaggerated stress of the materials. Deterioration is another effect that can be simulated to give a preview of the future conditions of the artworks subject to corrosion or deterioration (e.g., the erosion of sculpted stone decorations in our polluted historical towns). Very few works focused on have this subject, which involves an accurate simulation of both shape and reflection properties and the modification/evolution of the inspected surface.

A similar task, but with a different goal, is the virtual presentation of the forecasted effects of a restoration action in order to allow the restorers to show a plausible model of the expected results to decision makers and to the public before the execution of the restoration.

Therefore, a future goal for computer-aided restoration technologies would be the possibility of simulating the geometric and appearance effects of degradation of the different materials. This is a highly challenging task since we need to couple high-quality geometry acquisition techniques with accurate models of the physical and chemical properties of materials. Being able to simulate how a given metallic artifact would oxidate or a marble stone would degrade/erode under the attack of pollution, rains, and other effects would be a valuable instrument for CH management. Here the ideal goals are both to guess future damage or to bring back an endangered artwork to a plausible original status.

7.2 Reconstruction of Fragmented or Disassembled Artifacts

Another field of application is the digital reconstruction of disassembled or fragmented artworks. Physical reassembly is a process done manually by archeologists. The adoption of a computer-aided approach can be justified either in the case of the extreme fragility of the artifact or of complicated manipulation (e.g., the fragments are either too heavy or too many to be manipulated easily by an archeologist). Early methods have been proposed for special cases, such as the reconstruction of sherds of ancient pottery where some hypothesis on regularity and symmetry of shape can simplify the reassembly task [Kampel and Sablatnig 2003]. A recent result has shown that the generic process can be solved in a robust manner [Huang et al. 2006] by also taking into account the nonprecise and eroded fractures of archeological remains. The joint improvement of 3D scanning and automatic reassembling methods can open new insights into very complex reconstruction problems.

Moreover, creating assemblies with 3D CG or VR technologies could help to better understand the past. Three-dimensional scanned or modeled artifacts can be used as building blocks in interactive CG/VR applications (or to produce passive animations) to understand how different materials or components to worked in the past in building architectural structures [Levy and Dawson 2006] or complex instruments (ranging from the simple compound tool to the big industrial machine).

7.3 Virtual Repainting

The availability of accurate 3D models opens interesting capabilities for dissemination of the original aspect of ancient sculptures or architectures. In many cases, we have statues which have either completely lost their original painting (this is the case with many archeological masterpieces) or present severe deterioration. Using cheap video projectors, we can virtually restore color to the surface of these artworks by repainting the digital 3D model and projecting the model back on the surface of the original artifact [Raskar et al. 2001]. The same approach can also be used to present different reconstruction hypothesis of the original painting or to digitally paint solid copies produced with rapid reproduction technologies.

An open issue in this type of applications is how to register the original (or the physical copy) with the projected digital image. Manual registration is a slow and complicated process, and the same action could be transformed into a semi-automatic process by coupling the video projector with a video acquisition channel and adopting image-based solutions which could iteratively improve the mapping of the rendered image on the original.

7.4 3D Catalogs

CH cataloging efforts are still based on textual data enriched by a few, often B/W, photographs. It is very easy to forecast a future where catalogs will adopt digital 3D models as the main representation resource for characterizing the shape of the artifacts. Unfortunately, whether this will happen in the future depends more on the availability of funds and improved technological skills of the CH governing bodies than on technological issues.

The wide availability of 3D catalogs is the prerequisite for devising interesting applications concerning 3D search and retrieval methodologies, which are able to find shape-similar artifacts in large repositories given a sample object. A considerable research effort is now devoted to this topic; current solutions are in our opinion too naive to work well in a field so demanding such as CH, but technology might improve fast. Moreover, the availability of 3D models of similar artworks can be very valuable to support further study such as, for example, understanding how representation of God changed as religion moved to different parts of the world [Ikeuchi et al. 2003] or supporting the assessment of attribution hypothesis [Dellepiane et al. 2007].

While we consider the current 3D scanning technology mature for cataloging efforts (even if the data-encoding issue is still not completely solved with so many alternative 3D encoding solutions), some technical issues are still open if we consider the problem of data preservation in long time frames. The usual issues concerning the deterioration of hardware support are made more complex in the case of 3D graphics by the very fast renovation cycle of the software tools. This problem is less critical for 3D repositories than for 3D interactive applications. In the first case, the adoption of publicly known and documented data formats make data obsolescence a more tractable problem. But this issue has to be considered and solutions have to be planned well in advance.

7.5 Geographic Web Browsers Deploying 3D Models

Web-based systems such as Google Earth or Microsoft Local Live could boost considerably the usage of 3D reconstruction and the availability of 3D models depicting CH artifacts. These systems allow for the addition of small resolution 3D models on top of terrain models and easy linking of high resolution 3D models, which could represent CH resources (both architectures and artworks which characterize a given territory or urban context). We could easily forecast that, in the near future, local authorities will be active sponsors for massive data gathering to populate those systems, with the goal of increasing the tourism and providing a wide dissemination of local CH masterpieces. This could give an impressive boost to the use of both image-based approaches for the reconstruction of low-poly models and 3D scanning for higher accuracy models.

8. CONCLUSIONS

We have tried to demonstrate in this article that, 3D scanning can be considered a nearly mature technology. The research performed in the last few years has produced significant results, but some issues still remain open, mostly in the way this technology can be productively and effectively used, and as discussed in Section 5.3.1, how the quality of the obtained results can be objectively assessed. We have presented some recent results and some open research fields, aimed at improving ease of use and efficiency of the scanning technology. We also summarized some innovative applications which could be a stimulus for a wider adoption of 3D scanning in the CH field.

Since the focus of this article is broad, it is very hard to cite all interesting papers that have appeared so far. Rather than presenting a complete list of citations (which is impossible under the given article length constrain), we would like to refer the interested reader to some events specifically focused on the topics mentioned here: the VAST series of events (proceedings available in the Eurographics Digital Library at http://www.eg.org/), the CIPA symposia, the bi-annual 3DIM events (http://www.3dimconference.org/), and the 3DPVT conference (http://3dpvt.org/), together with the standard general conferences (ACM SIGGRAPH and Eurographics, for the CG field, and ICCV and CVPR in the Computer Vision domain) and journals.

ACKNOWLEDGMENTS

The David model (geometry only) is from the Stanford Digital Michelangelo project, provided by the Stanford repository, and the Pisa Cathedral is from the Digital Cathedral Project.

REFERENCES

BARACCHINI, C., BROGI, A., CALLIERI, M., CAPITANI, L., CIGNONI, P., FASANO, A., MONTANI, C., NENCI, C., NOVELLO, R. P., PINGI, P., PONCHIO, F., AND SCOPIGNO, R. 2004. Digital reconstruction of the Arrigo VII funerary complex. In *the 5th International Symposium on Virtual Reality Arhcaelogy and Cultural Heritage (VAST'04)*, N. S. Y. Chrysanthou, K. Cain and F. Niccolucci, Eds. Eurographics, 145–154.

BENDELS, G. H., DEGENER, P., WAHL, R., KÖRTGEN, M., AND KLEIN, R. 2004. Image-based registration of 3d-range data using feature surface elements. In *The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'04)*, Y. Chrysanthou, K. Cain, N. Silberman, and F. Niccolucci, Eds. Eurographics, 115–124.

BENDELS, G. H., SCHNABEL, R., AND KLEIN, R. 2005. Detail-preserving surface inpainting. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'06)*. M. Mudge, N. Ryan, and R. Scopigno, Eds. Eurographics Association, Pisa, Italy, 41–48.

BERNARDINI, F. AND RUSHMEIER, H. E. 2002. The 3D model acquisition pipeline. *Comput. Graph. For. 21,* 2, 149–172.

BERNARDINI, F., RUSHMEIER, H. E., MARTIN, I., MITTLEMAN, J., AND TAUBIN, G. 2002. Building a digital model of Michelangelo's Florentine Pieta. *IEEE Comput. Graph. Appl. 22,* 1, 59–67.

BLAIS, F. 2003. A review of 20 years of range sensor development. In *Proceedings of SPIEIS&T Electronic Imaging*. Vol. 5013. 62–76.

BLAIS, F., TAYLOR, J., COURNOYER, L., PICARD, M., BORGEAT, L., DICAIRE, L.-G., RIOUX, M., BERALDIN, J.-A., GODIN, G., LAHNANIER, C., AND AITKEN, G. 2005. Ultra-high resolution imaging at 50m using a portable xyz-rgb color laser scanner. In *International Workshop on Recording, Modeling and Visualization of Cultural Heritage*.

BOTSCH, M., WIRATANAYA, A., AND KOBBELT, L. 2002. Efficient high quality rendering of point sampled geometry. In *Proceedings of the 13th Eurographics Workshop on Rendering* , S. Gibson and P. Debevec, Eds. Eurographics, 53–64.

CALLIERI, M., CIGNONI, P., CORSINI, M., AND SCOPIGNO, R. 2008. Masked photo blending: Mapping dense photographic dataset on high-resolution sampled 3d models. *Comput. Graph. 32.* (To appear)

CALLIERI, M., CIGNONI, P., GANOVELLI, F., MONTANI, C., PINGI, P., AND SCOPIGNO, R. 2003. VCLab's tools for 3D range data processing. In *The 4th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'03)*. A. C. D. Arnold and F. Niccolucci, Eds. Eurographics, Bighton, UK, 13–22.

CALLIERI, M., CIGNONI, P., AND SCOPIGNO, R. 2002. Reconstructing textured meshes from multiple range RGB maps. In *7th International Fall Workshop on Vision, Modeling, and Visualization*. IOS Press, 419–426.

CALLIERI, M., PONCHIO, F., CIGNONI, P., AND SCOPIGNO, R. 2008. Virtual inspector: A flexible visualizer for dense 3d scanned models. *IEEE Comput. Graph. Appl. 28,* 1.

CAMPBELL, R. J. AND FLYNN, P. J. 2001. A survey of free-form object representation and recognition techniques. *Comput. Vision Image Understand. 81,* 2, 166–210.

CIGNONI, P., GANOVELLI, F., GOBBETTI, E., MARTON, F., PONCHIO, F., AND SCOPIGNO, R. 2004. Adaptive tetrapuzzles: Efficient out-of-core construction and visualization of gigantic multiresolution polygonal models. *ACM Trans. on Graph. 23,* 3, 796–803.

CIGNONI, P., GANOVELLI, F., GOBBETTI, E., MARTON, F., PONCHIO, F., AND SCOPIGNO, R. 2005. Batched multi triangulation. *IEEE Visualization Conference (VIS'05)*. 27–35.

CIGNONI, P., MONTANI, C., ROCCHINI, C., AND SCOPIGNO, R. 2003. External memory management and simplification of huge meshes. *IEEE Trans. Visualiz. Comput. Graph. 9,* 4, 525–537.

CURLESS, B. 2006. Vrippack users guide. Tech. rep. TR-UUCS-04-019, University of Washington. November 25.

CURLESS, B. AND SEITZ, S. 2000. 3D photography. *ACM SIGGRAPH Course Notes*, Course 19.

Davis, J., Marshner, S., Garr, M., and Levoy, M. 2002. Filling holes in complex surfaces using volumetric diffusion. In *1st International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'02)*. IEEE Computer Society, 428–438.

Debevec, P., Tchou, C., Gardner, A., Hawkins, T., Poullis, C., Stumpfel, J., Jones, A., Yun, N., Einarsson, P., Lundgren, T., Fajardo, M., and Martinez, P. 2004. Digitizing the Parthenon: Estimating surface reflectance properties of a complex scene under captured natural illumination. In *Proceedings of the Visim, Modeling and Visulization Conference (VMV04)*.

Dellepiane, M., Callieri, M., Fondersmith, M., Cignoni, P., and Scopigno, R. 2007. Using 3D scanning to analyze a proposal for the attribution of a bronze horse to Leonardo da Vinci. In *The 8th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'07)*. Eurographics 117–124.

Epoch. 2006. The European Network of Excellence on ICT Applications to Cultural Heritage (IST-2002-507382). `http://www.epoch-net.org/`.

Esteban, C. H. and Schmitt, F. 2003. Silhouette and stereo fusion for 3D object modeling. In *The 4th International Conference on 3D Digital Imaging and Modeling (3DIM'03)*. IEEE Computer Society, 46–53.

Fasano, A., Pingi, P., Cignoni, P., Montani, C., and Scopigno, R. 2005. Automatic registration of range maps. *Comput. Graph. For. 24,* 3.

Fontana, R., Greco, M., Materazzi, M., Pampaloni, E., Pezzati, L., Rocchini, C., and Scopigno, R. 2002. Three-dimensional modelling of statues: the minerva of arezzo. *J. Cultur. Herit. 3,* 4, 325–331.

Franken, T., Dellepiane, M., Ganovelli, F., Cignoni, P., Montani, C., and Scopigno, R. 2005. Minimizing user intervention in registering 2D images to 3D models. *Visual Comput. 21,* 8-10, 619–628.

Garland, M. and Heckbert, P. 1997. Surface simplification using quadric error metrics. *SIGGRAPH Computer Graphics Conference Proceedings*. (Annual Conference Series). IEEE, Addison Wesley, 209–216.

Hoppe, H. 1999. New quadric metric for simplifying meshes with appearance attributes. In *Proceedings of the IEEE Conference on Visualization (VIS'99)*. IEEE, 59–66.

Huang, Q.-X., Flöry, S., Gelfand, N., Hofer, M., and Pottmann, H. 2006. Reassembling fractured objects by geometric matching. *ACM Trans. Graphics 25,* 3, 569–578.

Ikeuchi, K., Nakazawa, A., Hasegawa, K., and Ohishi, T. 2003. The great buddha project: Modeling cultural heritage for VR systems through observation. In *The 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 7.

InnovMetrics. 2003. PolyWrks^{TM}inspection and modeling solutions. http://www.innovmetric.com/

INUS Technology. 2003. Rapidform^{TM} 2004. http://www.rapidform.com/index.htm.

Ju, T. 2004. Robust repair of polygonal models. *ACM Trans. Graph. 23,* 3, 888–895.

Kampel, M. and Sablatnig, R. 2003. Virtual reconstruction of broken and unbroken pottery. In *The 4th International Conference on 3D Digital Imaging and Modeling (3DIM 2003)*. IEEE Computer Society, 318–325.

Koller, D., Turitzin, M., Levoy, M., Tarini, M., Croccia, G., Cignoni, P., and Scopigno, R. 2004. Protected interactive 3D graphics via remote rendering. *ACM Trans. Graph 23,* 3, 695–703.

Lensch, H. P. A., Kautz, J., Goesele, M., Heidrich, W., and Seidel, H.-P. 2003. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph. 22,* 2, 234–257.

Levoy, M. 2006. Light fields and computational imaging. *IEEE Comput. 39,* 8, 46–55.

Levoy, M. and Hanrahan, P. 1996. Light Field Rendering. In *SIGGRAPH Computer Graphics Conference Proceedings* (Annual Conference Series). Addison Wesley, 31–42.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., and Fulk, D. 2000. The Digital Michelangelo Project: 3D scanning of large statues. In *SIGGRAPH Computer Graphics Proceedings*. (Annual Conference Series). Addison Wesley, 131–144.

Levy, R. and Dawson, P. 2006. Reconstructing a thule whalebone house using 3d imaging. *IEEE MultiMedia 13,* 2, 78–83.

Liepa, P. 2003. Filling holes in meshes. In *EG/ACM Symposium on Geometry Processing (SGP'03)*. Eurographics, 200–206.

(The) London Charter. 2006. The London Charter: For the use of 3D visualization in the research and communication of cultural heritage. More info on: `http://www.londoncharter.org/`.

Malzbender, T., Gelb, D., and Wolters, H. 2001. Polynomial texture maps. In *SIGGRAPH '01: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'01)*. ACM Press, New York, NY, USA, 519–528.

Podolak, J. and Rusinkiewicz, S. 2005. Atomic volumes for mesh completion. In *EG/ACM Symposium on Geometry Processing (SGP'05)*. Eurographics, 33–41.

POLLEFEYS, M., GOOL, L. J. V., VERGAUWEN, M., VERBIEST, F., AND TOPS, J. 2001. Image-based 3d acquisition of archeological heritage and applications. In *International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'01)*, D. Arnold, A. Chalmers, and D. Fellner, Eds. 255–261.

REINDROP GEOMAGIC. 2003. Geomagic^{TM} studio: Revere engineering and custom design software. `http://www.geomagic.com/products/`.

RASKAR, R., WELCH, G., LIM LOW, K., AND BANDYOPADHYAY, D. 2001. Shader lamps: Animating real objects with image-based illumination. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*. S. J. Gortler and K. Myszkowski, Eds. Eurographics, 89–102.

ROSSIGNAC. 2004. Surface simplification and 3D geometry compression. In *Handbook of Discrete and Computational Geometry*, Vol. 2 J. E. Goodman and J. O'Rourke, Eds. CRC Press.

RUSHMEIER, H. AND BERNARDINI, F. 1999. Computing consistent normals and colors from photometric data. In *Proceedings of the 2nd International Conference on 3D Digital Imaging and Modeling*. 99–108.

RUSINKIEWICZ, S. AND LEVOY, M. 2000. QSplat: A multiresolution point rendering system for large meshes. In *Computer Graphics Conference Proceedings (SIGGRAPH 00)*. ACM Press, 343–352.

SHARF, A., ALEXA, M., AND COHEN-OR, D. 2004. Context-based surface completion. *ACM Trans. Graph. 23,* 3, 878–887.

STUMPFEL, J., TCHOU, C., HAWKINS, T., DEBEVEC, P., COHEN, J., JONES, A., AND EMERSON, B. 2003. Assembling the sculptures of the Parthenon. In *The 4th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST'03)*. A. C. D. Arnold and F. Niccolucci, Eds. Eurographics, 41–50.

TEKUMALLA, L. S. AND COHEN, E. 2004. A hole filling algorithm for triangular meshes. Tech. rep. TR-UUCS-04-019, University of Utah.

TROCCOLI, A. AND ALLEN, P. K. 2005. Relighting acquired models of outdoor scenes. In *The 5th International Conference on 3-D Digital Imaging and Modeling (3DIM'05)*. 245–252.

TSAI, R. 1987. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE J. Robotics Autom. 3,* 4.

UMEDA, K., SHINOZAKI, M., GODIN, G., AND RIOUX, M. 2005. Correction of color information of a 3d model using a range intensity image. In *The 5th International Conference on 3-D Digital Imaging and Modeling (3DIM'05)*. 229–236.

VERGAUWEN, M. 2006. Harnessing structure-and-motion. Ph.D. thesis, Katholieke Universiteit Leuven, Belgium.

Received December 2006; revised March 2007, August 2007; accepted November 2007