Archeologia e Calcolatori
Supplement 1, 2007, 163-179

HIGH QUALITY DIGITAL ACQUISITION AND VIRTUAL PRESENTATION OF THREE-DIMENSIONAL MODELS

1. Introduction

Modern 3D graphics technologies allow us to acquire accurate digital models of real objects or complex scenes; these 3D models are the starting point for the design of a large number of applications based on visual presentation, ranging from the passive (video, animations, still images) to the more interactive and immersive ones (multimedia books, interactive navigation, immersive VR/AR systems, etc.). This technology opens great opportunities for a very broad set of applications in the cultural heritage field, but obviously is not limited to this application domain. In fact, 3D scanning technology was developed for more commercial fields, such as movie/animation and industrial design; it is now intensively used in medicine, industrial inspection, urban and terrain management, design, etc.

3D scanning technology has evolved considerably in the last few years, in terms of both hardware devices and algorithms for processing the raw data produced by the scanning devices (Bernardini, Rushmeier 2002). 3D scanning devices are usually based on optical technology (laser or structured light) and use either the triangulation approach (ideal for small and medium scale objects) or the time of flight approach (effective on large scale objects, e.g. architecture). The goal of this paper is not to give a detailed description of the architecture and features of existing scanning hardware; in fact, we will just give a very brief introduction and cite some examples in section 2.

The quality of the contemporary commercial scanning systems is quite good if we take into account the accuracy and speed of the devices; cost is still high, especially for our field of application, which is usually characterized by very low budgets. The latter problem could be mitigated by a wider diffusion of these systems, since a larger number of units sold per year would probably reduce costs significantly. Given the quality of commercial systems, the focus of this paper is on the software, and the discussion involves the issues introduced by the need to efficiently process the huge datasets produced with 3D scanning devices. The quality of the commercial software is still not sufficient to allow a mass-use of this technology. In our research work in the last few years, we have proposed some solutions which are aimed at reducing the complexity of the scanning process (making it easier and faster). We have dealt with two issues in greater detail: how to improve the automation of the post-processing phase (to minimize the human-assisted phases) and how to present complex 3D data with both extreme efficiency and simple interaction.

The scanning of a complex object is performed by taking a [usually large] set of partially overlapping range scans. The classical pipeline which characterizes a 3D scanning session is rather complicated, involving many different operations (introduced in section 3). A few of these processing phases require a substantial user intervention, with long processing times and tedious work; these are the phases where we have focused our research recently, with the aim of designing solutions to improve the automation of those processes and reduce the time required for completion (see subsection 3.1).

Once we have reconstructed a digital 3D model of the scene or of the object of interest, some issues arise from the very dense sampling resolution made possible by modern scanning devices. Being able to sample in the order of ten points per square millimeter or more (in the case of triangulation-based systems) is of paramount value in those 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 can become a problem for many applications. Below we have described the efficient methodologies that allow us to cope with data complexity, i.e. simplification and multiresolution representation.

Finally, the last section is dedicated to a glance at the near future and the presentation of a new technology which could play an important role in 3D acquisition, further reducing the overall cost for the end user.

2. Previous research

Many previous research projects concern the use of 3D technology either to reconstruct digital 3D models of cultural heritage masterpieces or to present those models through digital media (Levoy *et al*. 2000; Bernardini *et al*. 2002; Fontana *et al*. 2002; Polefeys *et al*. 2001; Stumpfel *et al*. 2003; Baracchini *et al*. 2004). An exhaustive description of those projects goes well beyond the scope of the brief overview that we can give in this section. We prefer instead to cite some of the seminal papers on the technologies proposed for 3D scanning and interactive visualization.

Automatic 3D reconstruction technologies have evolved significantly in the last few years. An overview of 3D scanning systems is presented in Curless, Seitz 2000. Among the 3D scanning systems most frequently used in cultural heritage digitizations, are the so called active optical devices. These systems project some sort of light pattern on the surface of the artifact and reconstruct its geometry by registering how the structured pattern is reflected by the surface. Examples are the many systems based on triangulation (using either laser stripes or more complex light patterns produced with a video projector). Very promising, but still not very common, are the passive optical devices; in this case the artifact is usually placed on a rotating platform, a large number of images are taken during rotation and a complete model is reconstructed from these images.

Unfortunately, most 3D scanning systems do not produce a final, complete 3D model but a large collection of raw data (range maps) which have to be post-processed. This is also the case of all the accurate active optical devices. The structure of the post-processing pipeline is presented in the 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 has not changed and therefore the description is still valid.

The high resolution meshes produced with 3D scanning are in general very hard to render with interactive frame rates due to their excessive complexity. This problem gave rise to intense research on: simplification and multiresolution management of huge surface meshes (GARLAND, HECKBERT 1997; HOPPE 1999; CIGNONI *et al*. 2003); and interactive visualization, where both mesh-based (CIGNONI *et al*. 2004) and point-based solutions (BOTSCH *et al*. 2002; RUSINKIEWICZ, LEVOY 2004) have been investigated.

3. PROCESSING SCANNED DATA

Since most of the current scanning systems acquire only a portion of the given artifact in a single shot, a complete 3D scanning requires the acquisition of many shots of the artifact taken from different viewpoints, to gather complete information on its shape (digital acquisition of the geometry and topology). Therefore, to perform a complete acquisition a first phase is to acquire many shots, producing the so called range maps, i.e. single views of the object which encode the sampled points' geometry; the number of range maps requested depends on the extent of the surface of the object and the complexity of its shape. Usually, we sample from a few dozen up to several hundred range maps. This set of range maps has to be processed to convert the data encoded into a single, complete, non-redundant and optimal 3D representation (usually, a triangulated surface). An example of digital model produced with 3D scanning technology is presented in Fig. 1. The processing phases (usually supported by standard scanning software tools) are:

– Range map alignment, since by definition 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 region (i.e. the sections of the range maps which correspond to the same portion of the artifact surface).
– Range map merger (or fusion), to build a single, non redundant triangulated mesh out of the many, partially overlapping range maps.
– Mesh editing, to improve (if possible) the quality of the reconstructed mesh.

Fig. 1 – A digital model produced with 3D scanning from 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.

– Mesh simplification, to accurately reduce the huge complexity of the model obtained, producing different high quality Level Of Details (LOD) or multiresolution representations.
– Color mapping, to enrich the information content by adding color information (an important component of the visual appearance) to the geometry representation, producing in output textured meshes.

The Visual Computing Lab of ISTI-CNR has designed and implemented a set of scanning tools: *MeshAlign, MeshMerge, MeshSimplify* (Callieri *et al.* 2003), *Weaver* (Callieri *et al.* 2002) and *TexAlign* (Franken *et al.* 2005) which support all the post-processing phases described above. The second generation of our tools has been produced in the framework of the EU IST "ViHAP3D" project (2002-2005). Our intention is not to give a comprehensive description of these tools here, but simply to give first a very brief overview and then a characterization of the problems and the bottlenecks in 3D scanned data processing.

MeshAlign makes it possible to register multiple range maps; it adopts a classical approach based on first, a pairwise local and then a global alignment (Pulli 1999). This canonical approach has been implemented with a number of innovations to reduce the user contribution, to improve efficiency and ease of use, and finally to support the management of a large number of range maps (in fact, we were able to process range datasets containing up to six hundred range maps).

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. 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. 6-8 times the number of sampled range maps). If the set of range maps is composed by hundreds of elements (the scanning of a statue 2 meters tall generally requires from 200 to 500 range maps, depending on the complexity of the shape of the statue), then the user has a very complex task to perform: for each range map, he must find which are the partially-overlapping ones; given this set of overlapping range maps, he must determine which ones to consider in pair-wise alignment (either all of them or a subset); finally, he must process all those pair-wise initial alignments. Our goals in the design of *MeshAlign* were:

– To support the management of very large sets of range maps (from 100 up to 1000); this can be achieved by providing both a hierarchical organization of the data (range maps divided into groups) and by using multiresolution representation of the data to make rendering and processing more efficient.
– Since the standard approach (user-assisted selection and initialization of all the overlapping pairs and the creation of the correspondent alignment arc) becomes impractical on large set of range maps, we planned to provide instruments for the automatic setup of most of the required alignment actions (see next subsection).
– Finally, provide visual/numerical presentation of the intermediate status of the alignment process and of the accuracy achieved.

MeshMerge (Callieri *et al.* 2003), our volumetric reconstruction tool, is based on a variant of the volumetric approach (Curless, Levoy 1996). *MeshMerge* can manage large range map sets (many million sample points) on low-cost PC platforms with an excellent level of efficiency.

A very important feature of a reconstruction code is the performance of a weighted integration of the range maps and not just joining them. Since we usually have a high degree of overlap (and considering that sampled data contain some noise), a weighted integration can significantly improve the accuracy of the final result, reducing the impact of the possible noisy samples which are located in proximity of other more accurate samples taken with overlapping range maps. Another important feature of a reconstruction code is the ability to fill up small holes (i.e. region not sampled by the scanner); this is an optional feature of *MeshMerge*.

Since the adoption of a volumetric approach requires a very large memory footprint on big dataset, *MeshMerge* provides a split-reconstruction feature: to process huge datasets it works on sub-sections of the data (out-of-core), loading only the range maps involved in the generation of that single section of the voxel set. The various parts of the final model are joined after the split-reconstruction process with a small time overhead; the boundary of the sub-blocks are guaranteed to be identical so the joining of resulting sub-meshes is trivial.

The reconstructed models (when produced using a voxel size equal or smaller than the inter-sampling distance used in scanning) are usually huge in size (i.e. many millions 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 a solution working on external memory to cope with these big models; simplification has to be accurate if we want to obtain high-quality models and accurate visualization.

Our *MeshSimplify* tool (CIGNONI *et al.* 2003) has no limits in terms of maximal size of the triangle mesh in input, since it adopts an external-memory approach; at the same time, it ensures high-quality results, since it is based on edge collapse and takes into account both accuracy of geometry and shape curvature (GARLAND, HECKBERT 1997; HOPPE 1999).

Finally, the *Weaver* tool (CALLIERI *et al.* 2002) supports the reconstruction of textured meshes from a sampling of the object appearance. We usually perform the acquisition of the apparent color (reflected color, illumination-dependent) using digital photo cameras. This is the easiest and most practical approach, since setting up a controlled lighting for a more sophisticated acquisition of the reflection properties of the object's surface (BRDF acquisition, see LENSCH *et al.* 2003) is often impossible or impractical for cultural heritage artifacts. To map color data on the 3D model we first compute the inverse projection and intrinsic parameters for each photo (from the image to the 3D mesh) using our *TexAlign* system (FRANKEN *et al.* 2005). Then, the *Weaver* tool computes an optimal coverage of the 3D mesh with sections of the original images, packs all the used portions in a new texture map and stores UV parameterization in the triangle mesh. Finally, it reduces color (hue/intensity) disparity on boundaries between overlapping photo parcels.

3.1 *Making alignment an automatic process*

Solutions for a completely automatic scanning system have been proposed, but either these systems are based on the use of complex positioning machinery, or adopt passive silhouette-based approaches which do not guarantee the required accuracy. 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 map registration phase is the only task where considerable human intervention is still requested. Several papers have proposed methods for automatic alignment, usually based on some form of shape analysis (see CAMPBELL, FLYNN 2001 for a survey paper).

In designing a new solution (FASANO *et al.* 2005), we started from a few initial conditions directly gathered from our experience in 3D scanning. First, the detection of the pairs of overlapping range maps can be made much simpler, 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

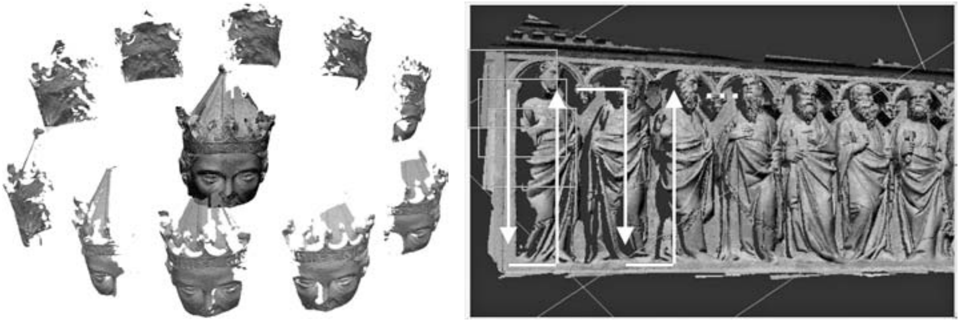


Fig. 2 – Range maps are taken in a row-wise order: an example of circular stripe around a statue's head (left); an example of raster-scan scanning order adopted for the acquisition of a bas-relief (right).

vertical, horizontal, raster-scan or circular translation of the scanning system (Fig. 2). The different types of sequences share some common properties: they contain an ordered set of *n* range maps, such that range map R_i holds a significant overlapping with at least and 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 indeed 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}. From the point of view of the registration algorithm, all the stripe patterns defined above are equivalent: an automatic registration module can processes each couple R_i, R_{i+1}, in order to produce in output the roto-translation matrix M_i which aligns R_i to R_{i+1}.

The subset of registration arcs defined above is not complete (since we usually have many other potential overlaps between range maps), but sufficient for the application of an intelligent ICP-based solution. Our MeshAlign system is able to complete the required arcs (interconnecting R_i with all the overlapping range maps, not just R_{i-1} and R_{i+1}) in an automatic manner. MeshAlign adopts a spatial indexing technique, which for each 3D grid cell stores the set of range maps passing through that region of space, to detect possible overlap and to run on the corresponding range map pair the required ICP-based alignment. Given the occupancy grid information and once a single alignment arc is provided for each range map, our registration system is able to introduce all the arcs needed (in a completely unattended manner), by selecting and processing only those which satisfy a minimum-overlap factor.

To solve the rough registration of range map R_i over R_{i+1}, we have developed an efficient shape characterization kernel which works directly on the

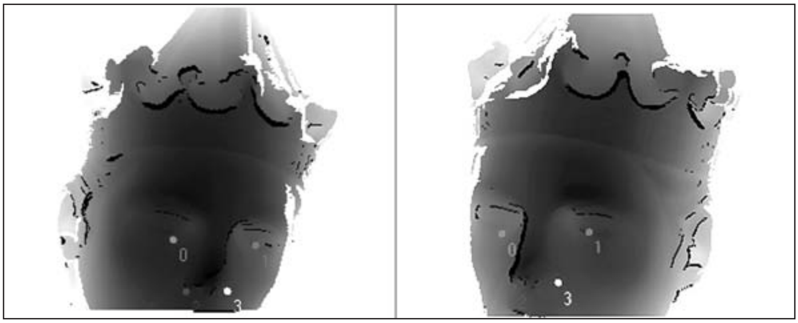


Fig. 3 – The four matching point pairs selected by the algorithm on two range maps.

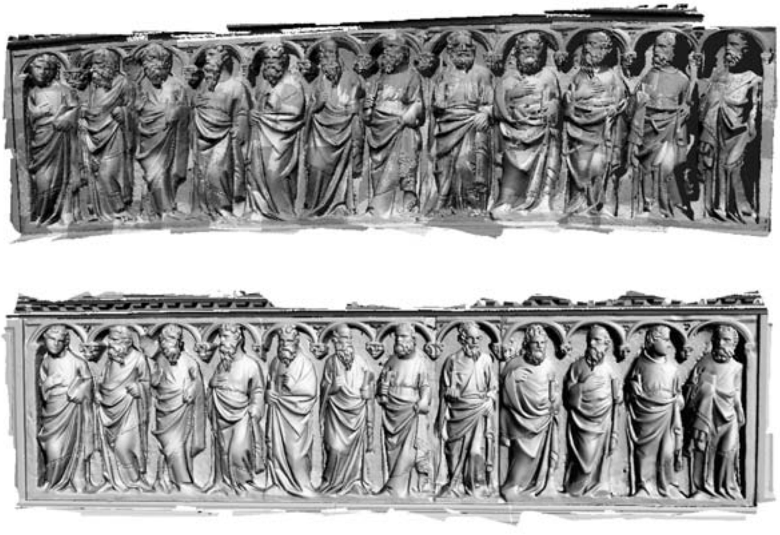


Fig. 4 – The coarse alignment of the bas-relief (top) and the final model (middle); almost all of the alignments required just 1 iteration.

discrete range map space. Like other surface matching algorithms, we look for a small set of feature points which characterize the first range map. For the sake of simplicity we consider our input meshes as regularly sampled 2D height fields. In a second step, for each of these *k* points on R_i we search for the potential corresponding points on the second mesh R_{i+1}. Finally, out of those possible *k* pairs we choose the group of four matching points which gives the best coarse alignment (see Fig. 3); if the required accuracy has not been achieved, we iterate until convergence, by selecting and checking several different points.

The metric defined to select matching points is invarianble with respect to the usual transformations (translations and rotations) that occur to the meshes belonging to a strip. This metric is not invariant to consistent rotations over the view direction of the scanning device. However in standard 3D scanning rotating the scanner along his viewing axis is rather uncommon (the scanner is usually connected to a tripod, which makes impossible to apply a substantial rotation along the viewing axes).

The proposed registration algorithm was tested on many large datasets coming from real scanning campaigns (each range map contains therefore real raw data, usually affected by noise, artifacts and holes).

An example concerning a bas-relief is shown in Fig. 4, with an approximate length is 2.5 meters; in this case two raster-scan (snake-like) stripes were acquired, for a total of 117 meshes (about 45.5M vertices). The overall alignment required 1h:50min (on a Pentium IV 2.4GHz), i.e. much less than the raw scanning time (approximately 4 hours in the case of the basrelief). The solution presented is sufficiently fast to run in the background during the acquisition, processing all the scans in sequence as soon as they are produced.

3.2 *Mapping complex photographic detail on 3D models*

Many applications require sampling not just the geometry, but also the color information. Cultural heritage is obviously a good example of an application field were an accurate management of color data is required. Accurate approaches for sampling the reflection characteristics of an artifact surface have been proposed (e.g. BRFD sampling), but are still too complicated to be massively applied to cultural heritage applications were, usually, we do not work in controlled lab conditions but in crowded museums.

For most practical cases a simpler approach is adopted: a series of pictures taken by a digital camera are stitched onto the surface of the object, trying to avoid shadows and highlights and taking pictures under favorable light conditions. However, even in this simpler case, the pictures need to be processed in order to build a plausible texture for the object (CALLIERI *et al*. 2002).

A basic problem in managing color information is how to register the images with geometric data. In most cases, the set of images is taken after the scanning, using a consumer digital camera. This registration step is again a complicated, time-consuming phase which requires substantial intervention of a human operator. Unfortunately, no fully automatic powerful approach has been proposed for the general problem (i.e. a large and complex object, where each image covers only a subset of its overall extent). The user is usually required to provide correspondences, or hints on the correspondences, which link the 2D images and 3D geometry.

In a recent research project we designed a new tool to support image-to-geometry alignment, *TexAlign* (FRANKEN *et al*. 2005), in which the main goals were: to reduce user intervention in the process of registering a set of images with a 3D model; to improve the power of the process by giving the user the possibility of selecting correspondences which link either 2D points to 3D geometry (image-to-geometry correspondences) or 2D points to 2D points (image-to-image correspondences). The latter can help a lot in all those cases where a single image covers a region where the surface does not have sufficient shape features to allow an accurate selection of image-to-geometry correspondences. The *TexAlign* tool tries to solve the problem by setting up 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. This graph of correspondences is used to keep track of the work done by the user, to infer automatically new correspondences from the one represented and 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.

In all those cases where the operator has a large number of images to align and map to the 3D shape, *TexAlign* makes it possible to reduce the time needed to perform the alignment and to improve the overall accuracy of the process. Some results are reported in (FRANKEN *et al*. 2005). This system has been recently used to map a complex photographic sampling (more than 61 pre-restoration and 68 post-restoration images to be mapped on the David model, see Plate IX).

4. INTERACTIVE VISUAL PRESENTATION OF VERY LARGE MODELS

Some issues arise from the impressive increase in data complexity (and richness) provided by the evolution of 3D scanning technology: how to manage/visualize those data on commodity computers; how to improve the ease of use of the visualization tools (as potential users are often not expert with interactive graphics); how to support the presentation of other multimedia information together with the visualization of complex 3D geometry. Our Virtual Inspector browser (CALLIERI *et al*. 2007) has been designed to give a solution to these issues.

4.1 *Simplification and multiresolution management of very large models*

One of the major issues is how to cope with the complexity of the data. A first approach is to adopt a data simplification approach, i.e. to reduce the data resolution at the expense of a loss of geometric accuracy. Many solutions have been proposed for the simplification of 3D triangulated surfaces, usually based on the iterative elimination of selected vertices or faces. The basic idea is: to select at each iteration the local change (e.g. collapse a given triangular face into a vertex) which minimizes the loss of accuracy; and to repeat these local changes (each one reducing the size of the mesh of a few triangles) until either the required model size ("no more than 100K faces in the final model") or degradation of accuracy ("error should not be greater than 0.5 mm") is met. This approach allows the construction of any level of resolution we want, usually 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 the specific application (e.g. a model for a web presentation which should be downloadable in a given short time, or a model to be used for rapid reproduction by a 3D printer).

Another approach is to store not just the final simplified model, but all the intermediate results obtained during simplification. The latter have to be encoded in an efficient data structure (multiresolution encoding) that will allow our application to extract in real time models at a resolution optimized to the requirements of each single frame of the application (e.g. a visualization browser).

Virtual Inspector is a visualization system that allows non-expert users to inspect a large complex 3D model at interactive frame rates on standard PC's. To support the efficient manipulation of massive models, Virtual Inspector adopts a new multiresolution approach where view-dependent variable resolution representations can be extracted on the fly (Cignoni *et al.* 2004). For each frame, the best-fit variable resolution LOD is selected according to the current view specification (higher resolution for the portions in foreground, progressively lower resolution for data in the background) and the required visualization accuracy.

4.2 *Usability of virtual heritage worlds*

The ease of use of a tool intended to present cultural heritage material to ordinary people (still not very used to 3D graphics and computer games) is an important factor for the success of an application. One of the most complicated actions that has to be performed is to drive navigation in virtual space. Therefore, free navigation should be required only in those cases where this action really adds something to the learning experience. The risk is to have the visitor become disoriented (e.g. discover himself lost in sidereal space, maybe just because he turned his back to the scene) and abandon the system.

Virtual Inspector is mainly intended for the visualization of single works of art (sculptures, pottery, architectures, etc.), and adopts a very intuitive approach to guide the virtual manipulation and inspection of the digital replica, based on a straightforward metaphor: we provide a dummy representation of the current inspected model on a side of the screen, which can be rotated on its axis; to select any given view the user needs only to point with the mouse to the corresponding point on the dummy (Plate IX).

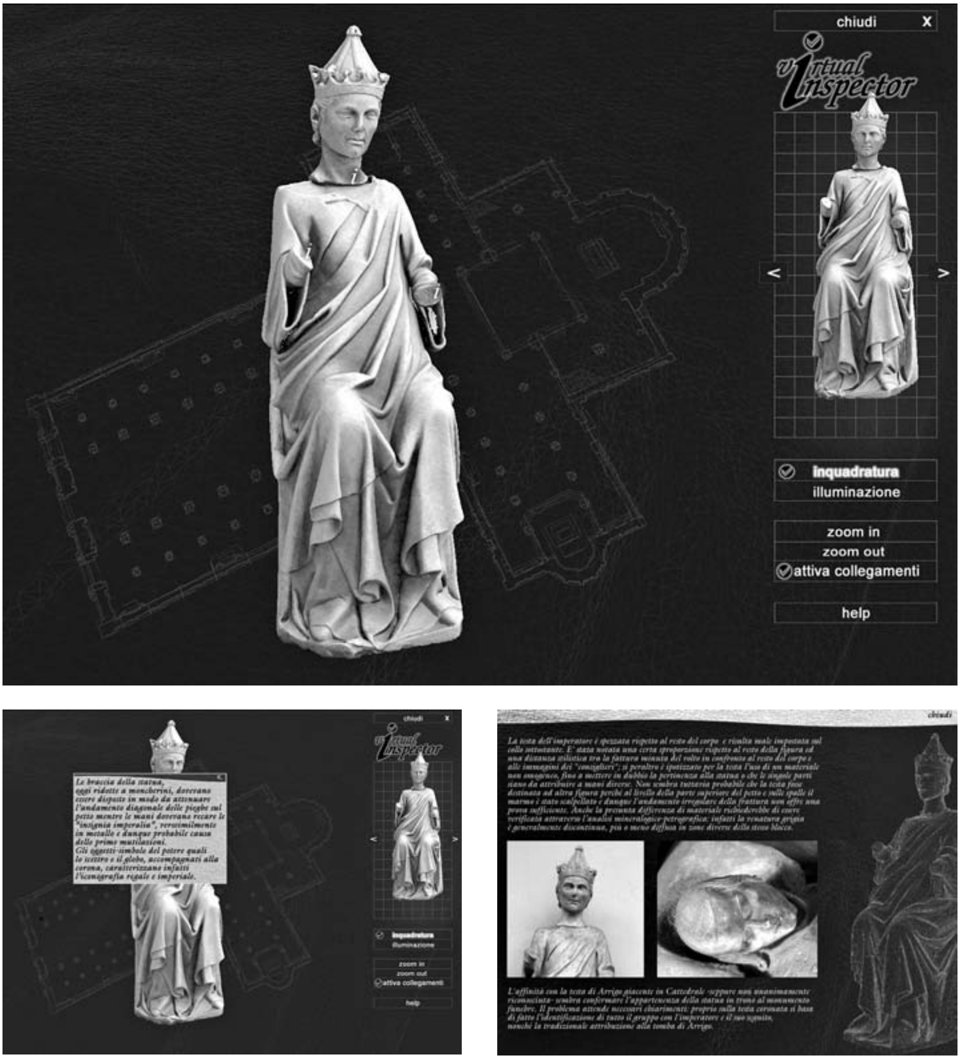
Another important characteristic of a visualization system is its flexibility and configurability. To fulfill this objective developers would be forced to design very complicated systems characterized by a very complex set of functionalities (e.g. consider scientific visualization tools). Conversely, while designing Virtual Inspector as a system oriented to non-expert users (i.e. museum visitors), our approach was to define a restricted set of functionalities and to provide the system with an easy installation interface for the selection of the subset of these functionalities that the designer of the specific installation want to include in the installation (i.e. a museum stand).

All main parameters of a Virtual Inspector installation can be easily specified via XML tags contained in a initialization file, such as: which 3D models are to be rendered (a single mesh or multiple ones), the system layout characteristics (i.e. how the different models will be presented on the screen, where GUi buttons are located), the rendering modes (e.g. standard Phong-shaded per-vertex colors or BRDF rendering) and the interaction mode (e.g. model manipulation via the standard virtual trackball, the dummy-based "point and click" interaction, or both). Therefore, the design of the graphic layout can be done easily by a professional graphic designer, since the layout of the application, all icons and background graphics elements can be completely redesigned with respect to previous incarnations of the Virtual Inspector system. This can be done by the easy specification of the new images and location on the screen of all icons and elements of the GUI in the XML initialization file and does not require either programming nor recompilations of Virtual Inspector. It is a task that can be easily assigned to an operator with very limited IT competence.

4.3 *Not just 3D data: adding other knowledge*

Hot spots are a very handy resource for associating multimedia data (e.g. html pages) to any point or region of a 3D model. This allows us to design interactive presentations where the 3D model is also a natural visual index to historical/artistic information, presented using standard HTML format and browsers (Fig. 5).

The specification of hot spots is extremely easy in Virtual Inspector; modifications to the 3D models are not required. We provide a simple 3D browser to the person in charge of the implementation of the multimedia presentation, which makes it possible to query the 3D coordinates of any point on the surface of the artifact (by simply clicking with the mouse on the corresponding point). Then, a new hot spot is specified by introducing a new XML tag in the Virtual Inspector specification file. The hot spot XML tag specifies basically the 3D location and the action that has to be triggered when



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 imperialia", 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 e 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 scalpellinato 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. 5 – Virtual Inspector: the "Arrigo VII" statue rendered with active hot spots (top); a short popup panel with a short info, describing the missing hand, appears when the mouse passes over the hotspot located on the hand (left), or a more complex page associated to the hotspot on the neck (right).

clicking on the hot spot (e.g. the name of the html file, if we want to open a multimedia page). After activation, the control passes to the html browser, while Virtual Inspector remains sleeping in the background and automatically regains control of the interaction whenever the html browser is closed.

Fig. 6 – A 3D model reconstructed from just a sequence of approximatively one hundred high resolution photos. The reconstruction was performed using the tools developed within the EU IST NoE "Epoch".

5. A GLANCE AT NEAR-FUTURE TECHNOLOGIES

One of the most frequent issues in the common practice of cultural heritage 3D scanning is the high cost of the technologies involved, which often become unsustainable for low-budget projects. Classical high quality laser-based technologies, like the approaches based on time-of-flight or triangulation cited in the introduction, employ high end hardware whose price spans in the range of 40-100 thousands Euro. Luckily enough, a cheaper and lower quality alternative approach is emerging, performing 3D reconstruction from a simple sequence of high resolution digital photos of the artifact.

The recovery of three dimensional structure out of a sequence of photos is a well studied field in computer vision literature, but, until recently, it was difficult to really exploit the results of the many algorithms presented within a single framework (Pollefeys *et al.* 2001; Vergauwen 2006).

A result of the application of this technology is shown in Fig. 6. The model shown in Fig. 6 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 experimental tools and web service that have been developed within the European IST Network of Excellence Epoch (http://www.epoch-net.org/). Users registered on this web service can simply upload their photo sequences on a remote server that automatically converts the photos into a sequence of aligned range maps (one for each photo) that can be downloaded and processed by the user. In exchange the user has to provide public, non commercial access to the reconstructed 3D data.

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 simply taking a reasonably large number of photos (in the order of many dozens or about one hundred) all around the object. On the other hand, this approach still exhibits less 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 some difficulties in the reconstruction of artifacts with flat and uniformly colored parts.

6. Conclusions

3D scanning can be considered as a nearly mature technology. The research performed in the last few years has produced significant results, but some issues still remain open. We have presented some recent results on two different sides: how to increase the automation of the scanning process (which, unfortunately, is still user-assisted if we want to produce a good-quality model); and how to manage efficient rendering of very large models, supporting also the integration of multi-media data to the 3D mesh with the classical hyperlink approach.

Roberto Scopigno, Paolo Cignoni, Claudio Montani
Visual Computing Lab
Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo"
CNR – Pisa

Acknowledgements

We wish to acknowledge the financial support of some active projects: the EU Network of Excellence IST-2002-507382 "EPOCH", projects "Fruizione – Tecniche di supporto e modalità innovative" of the CNR Patrimonio Culturale Dept. and "Media Multidimensionali" of the CNR ICT Dept. However, the work presented is also the outcome of several prior projects, such as projects EU IST-2001-32641 "ViHAP3D" and Regione Toscana "RIS+". The David model (geometry only) is a result of the Stanford's Digital Michelangelo project, provided by the Stanford repository.

REFERENCES

Baracchini C., Brogi A., Callieri M., Capitani L., Cignoni P., Fasano A., Montani C., Nenci C., Novello R.P., Pingi P., Ponchio F., Scopigno R. 2004, *Digital reconstruction of the Arrigo VII funerary complex*, in N.S.Y. Chrysanthou, K. Cain, F. Niccolucci (eds.), *VAST 2004 (Bruxelles, BE, 2004)*, Eurographics, 145-154.

Bernardini F., Rushmeier H.E. 2002, *The 3D model acquisition pipeline*, «Computer Graphics Forum», 21, 2, 149-172.

Bernardini F., Rushmeier H.E., Martin I., Mittleman J., Taubin G. 2002, *Building a digital model of Michelangelo's Florentine Pietà*, «IEEE Computer Graphics and Applications», 22, 1, 59-67.

Botsch M., Wiratanaya A., Kobbelt L. 2002, *Efficient high quality rendering of point sampled geometry*, in S. Gibson, P. Debevec (eds.), *Proceedings of the 13th Eurographics Workshop on Rendering (RENDERING TECHNIQUES '02) (Aire-la-Ville, CH, 2002)*, Eurographics, 53-64.

Callieri M., Cignoni P., Ganovelli F., Montani C., Pingi P., Scopigno R. 2003, *VCLab's tools for 3D range data processing*, in A.C.D. Arnold, A. Chalmers, F. Niccolucci (eds.), *VAST 2003 (Brighton, UK, 2003)*, Eurographics, 13-22.

Callieri M., Cignoni P., Scopigno R. 2002, *Reconstructing textured meshes from multiple range rgb maps*, in G. Greiner (ed.), *7th International Fall Workshop on Vision, Modeling, and Visualization 2002 (Erlangen, D, 2002)*, IOS Press, 419-426.

Callieri M., Ponchio F., Cignoni P., Scopigno R. 2007, *Virtual Inspector: A flexible visualizer for dense 3D scanned models*, «IEEE Computer Graphics and Applications», in print.

Campbell R.J., Flynn P.J. 2001, *A survey of free-form object representation and recognition techniques*, «Computer Vision and Image Understanding», 81, 2, 166-210.

Cignoni P., Ganovelli F., Gobbetti E., Marton F., Ponchio F., Scopigno R. 2004, *Adaptive tetrapuzzles: Efficient out-of-core construction and visualization of gigantic multiresolution polygonal models*, in *SIGGRAPH 2004*, «ACM Transaction on Graphics», 23, 3, 796-803.

Cignoni P., Montani C., Rocchini C., Scopigno R. 2003, *External memory management and simplification of huge meshes*, «IEEE Transactions on Visualization and Computer Graphics», 9, 4, 525-537.

Curless B., Levoy M. 1996, *A volumetric method for building complex models from range images*, in *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, ACM Press, 303-312.

Curless B., Seitz S. 2000, *3D Photography*, ACM SIGGRAPH 2000 Course Notes, Course 19.

Fasano A., Pingi P., Cignoni P., Montani C., Scopigno R. 2005, *Exploiting the scanning sequence for automatic registration of large sets of range maps*, «Computer Graphics Forum», 24, 3, 517-526.

Fontana R., Greco M., Materazzi M., Pampaloni E., Pezzati L., Rocchini C., Scopigno R. 2002, *Three-dimensional modelling of statues: The Minerva of Arezzo*, «Journal of Cultural Heritage», 3, 4, 325-331.

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

Garland M., Heckbert P. 1997, *Surface simplification using quadric error metrics*, in *SIGGRAPH 97 Conference Proceedings*, Annual Conference Series, Addison Wesley, 209-216.

Hoppe H. 1999, *New quadric metric for simplifying meshes with appearance attributes*, in *Proceedings of the 10th IEEE Conference on Visualization (VIS '99) (San Francisco, USA, 1999)*, IEEE Computer Society, 59-66.

Lensch H.P.A., Kautz J., Goesele M., Heidrich W., Seidel H.-P. 2003, *Image-based reconstruction of spatial appearance and geometric detail*, «ACM Transaction on Graphics», 22, 2, 234-257.

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

Pollefeys M., Gool L.J.V., Vergauwen M., Verbiest F., Tops J. 2001, *Image-based 3D acquisition of archeological heritage and applications*, in D. Arnold, A. Chalmers, D. Fellner (eds.), *VAST 2001 Conference Proceedings*, Athens, ACM SIGGRAPH, 255-261.

Pulli K. 1999, *Multiview registration for large datasets*, in *Proceedings of the 2nd International Conference on 3D Digital Imaging and Modeling (Ottawa, CA, 1999)*, IEEE, 160-168.

Rusinkiewicz S., Levoy M. 2000, *QSplat: A multiresolution point rendering system for large meshes*, in *SIGGRAPH 2000 Conference Proceedings*, Annual Conference Series, ACM Press, 343-352.

Stumpfel J., Tchou C., Hawkins T., Debevec P., Cohen J., Jones A., Emerson B. 2003, *Assembling the sculptures of the Parthenon*, in A.C.D. Arnold, A. Chalmers, F. Niccolucci (eds.), *VAST 2003 (Brighton, UK, 2003)*, Eurographics, 41-50.

Vergauwen M. 2006, *Harnessing structure-and-motion*, Ph.D. dissertation, Katholieke Universiteit Leuven, Belgium.

ABSTRACT

Detailed and accurate digital 3D models can be produced with 3D scanning devices, which allow to convert reality in digital form in a cost- and time-effective manner. The capabilities of this technology and the global methodology are presented here in a synthetic manner. Moreover, we focus on the main issues which are preventing its wider use in contemporary applications, such as: the considerable user intervention required, the usually incomplete sampling of the artifact surface and the complexity of the models produced. Another emerging issue is how to support the visual presentation of the models (local or remote) with guaranteed interactive rendering rates. Some practical examples from the results of current projects in the cultural heritage field will be shown.