3D Digitization: making it easier and extending it to color

Roberto Scopigno*

Istituto di Scienza e Tecnologie dell'Informazione (ISTI-CNR), Pisa, ITALY

ABSTRACT

The easy construction of detailed and accurate 3D models is becoming a reality by the increasing diffusion of 3D scanning technology. The reduction in cost of the scanning devices and the increasing availability of good processing tools (including emerging open source solutions) makes 3D scanning an enabling technology for the construction of shape models. The talk will present the capabilities of this technology, presenting some recent advances (low cost scanning systems, 3D-from-images technology, improved automation of sampled data processing) and highlighting some open problems. A major focus will be how color or surface reflection characteristics could be sampled and associated with reconstructed 3D shape models. The different approaches proposed will be reviewed, giving more emphasis to the more practical solutions for both acquiring color or surface reflection and mapping those data efficiently on surface meshes. Some examples of the results of current projects, mainly in the Cultural Heritage field, will be shown.

Keywords: 3D scanning, sampled data processing, color acquisition and mapping, cultural heritage applications

Index Terms: I.3.7 [Three-Dimensional Graphics and Realism]: Color, shading, shadowing, and texture;

1 INTRODUCTION

3D technology is nowadays in a consolidate status, since 3D data can be managed on any low-cost computer, thanks to the impressing improvement of technology brought us by the huge 3D computer games market. Any PC comes equipped with everything is needed to manage interactive 3D graphics. New technologies also exist for sampling 3D shapes, usually called 3D scanners. The last ten years have shown an impressive progress of 3D scanning solutions, including both hardware devices (used to sample real objects and to return us sampled 3D point clouds) and graphics software needed to transform those sampled point clouds into good-quality 3D models and to use them in real applications.

Nevertheless, we still miss a significant impact of 3D graphics on Cultural Heritage (CH) applications. Even if we have a series of good practices and some important examples where digital 3D data played an important role, the adoption of those technologies is still far below what we could expect. There are some reasons for that: the 3D graphics field only recently reached a consolidated status; most of the experiences done so far were often driven by academia, rather than being driven directly by CH operators. Some miss concepts or wrong beliefs are also responsible of a very slow diffusion and some skepticism among our CH colleagues. Finally, color acquisition and management on scanned 3D models has been perceived as largely unsatisfactory by art experts, used to the high quality photographic medium. We will try to discuss in this paper some of the more common beliefs, with the aim of demonstrating that some of the perceived cons of this technology are due to problems which have been solved recently.

2 IS 3D SCANNING A TOO EXPENSIVE TECHNOLOGY?

One of the criticism more often raised against the adoption and massive use of 3D scanning technologies is the cost for the deployment of that type of technology. Especially when considering the low budget which characterizes most CH-related activities, 3D scanning cost is often perceived as excessive. Cost issues are raised at several different levels: cost of the hardware required, i.e. of the specific 3D scanning devices; cost of the software needed to process the raw data produced by 3D scanners and processing time (including also the personnel cost, since in some cases a highly skilled operator is still required); and levels of skills required for the operator to ensure proper and successful use of these technologies.

2.1 Cost of hardware

A large number of different 3D digitization devices have been proposed in the last 20 years [1]. A common distinction is between active optical vs. passive optical devices. Active optical systems can be further divided into: triangulation-based systems (using either laser or structured light patterns), and time-of-flight (TOF, also called LIDAR). Passive optical include: silhouette-based systems, stereo and multi-stereo matching solutions (which reconstruct 3D geometry from streams of photographs or videos). Currently, the most diffuse systems are active optical systems (triangulation systems for small/medium scale artifacts and TOF device for large scale artifacts, such as architecture). Unfortunately, the reduction in the cost of these systems was nearly negligible in the last ten years, much slower than the cost reduction experimented in other information technology fields. The price tag of good devices is still in the order of 30-60 KE for triangulation-based systems and 70-100 KE for TOF systems. The slow technical advance and the minor price cut 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 massive R&D effort and large scale production savings. For small/medium scale acquisition, the recent introduction of a low-cost laser-based device sold at 2500 USD is a remarkable news, which should have a giant impact on the domain (it is a triangulation-based system, see at https://www.nextengine.com/). A similar impressive reduction of TOF cost is still a dream, but the good news is that the acquisition of large scale artifacts can now be approached by adopting new passive optical methodologies, in particular the ones that perform 3D reconstruction from a simple sequence of high resolution digital photos of the artifact [9, 12]. These methods are an evolution of the old photogrammetry approach, they have been considerably improved recently and show some interesting potential for a very wide diffusion. They are based on the search of a small set of correspondences between the processed images; these correspondences (usually in the order of tens or one hundred) identify some feature points in the scene as seen from different point 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. An example of result obtained with this technology is shown in Figure 1, where the model presented has been reconstructed by processing

---

*E-mail: r.scopigno@isti.cnr.it

Figure 1: Image from a 3D model obtained with passive reconstruction from a set of digital images (by ARC 3D and MeshLab tools).

some photos, shot all around the statue, using the ARC 3D Webservice (http://www.arc3d.be/) developed within the EC IST Network of Excellence “Epoc” (http://www.epoch-net.org). The raw data returned by the ARC 3D system have been processed with the MeshLab tool (http://meshlab.sourceforge.net) [5] .

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 geometric precision that is much less predictable than the well assessed laser-based 3D scanning technologies: since the reconstruction process is based on the detection of corresponding features on consecutive photos, these approaches encounter difficulties in the reconstruction of artifacts with large flat and uniformly colored parts that do not exhibit evident features to be recovered (e.g. uniformly painted walls) and have even more significant problems with non-diffuse surfaces.

2.2 Cost of software

Unfortunately, 3D scanning systems do not produce a final, complete 3D model but rather a large collection of raw data, which have to be post-processed. A complete scan of an artifact requires the acquisition of many shots taken from different viewpoints to gather complete information on its shape. Each shot produces a range map, that is a single partial view of the object. The number of range maps required to sample an artifact 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, non redundant, and optimal digital 3D representation (usually, encoded by a triangulated surface). The processing phases (usually supported by commercial tools) are:

- 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 two adjacent range maps which sample the same portion of the artifact surface).

- Range Maps Merge (or Reconstruction). A single, non-redundant triangulated mesh is built out of the many partially overlapping range maps. This processing phase reduces the redundancy (after 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 un-sampled regions (generating surface patches to close small holes).

- Mesh Simplification. The huge complexity of the model obtained usually has to be reduced in a controlled manner to transform the usually huge master model (millions of samples and triangles) into a model of size appropriate to the specific application. A huge mesh can be either simplified or converted into a discrete a Level-Of-Detail (LOD) model or a multiresolution representation.

- 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 (e.g. INUS Technology, InnovMetrics, Raindrop Geomagic) or academic tools [10, 3, 5] . Unfortunately, commercial software is still very costly (around 10K-20K Euro for each installed workstation). The diffusion of open source solutions could be an important resource for fostering an increased diffusion of this technology; some academic labs are following this policy [5].

2.3 Time required to process the raw data

The results of the last ten years of research on sampled 3D data processing had a profound impact on the time and effort requested to the user to transform the raw, point-based sampled data into a good quality 3D model. Processing large sampling was a nightmare until recently. Taking into example the case of a single statue, processing time has been reduced from several weeks to a few days (1-3), thanks to a progressive automation of the 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 a multiresolution representation of the raw data, to make rendering and processing more efficient. Moreover, since the standard approach (user-assisted selection of each overlapping pair and selection of the correspondent alignment pairs) becomes impractical on a large set of range maps, some solutions for a completely automatic range map alignment have been proposed. These methods are based on the characterization of a few feature points contained in each range map and subsequent search for matching points in the adjacent maps. These solutions have been demonstrated to work well (90-95% reduction of the processing time) but are still available only in academic solutions [7, 3]. Scanning systems that automatically track the scanner location and therefore produce aligned range maps also exist (based on magnetic or optical tracking), but usually cost twice than standard high quality devices and are thus of limited diffusion.

3 IS 3D SCANNING LIMITED TO GEOMETRY SAMPLING?

Most 3D scanning systems consider just the geometric 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 usually produce low-quality color sampling (with a notable exception of the technology based on multiple laser wavelengths, unfortunately characterized by a very high price). Moreover it should be noted that existing devices sample only the apparent color of the surface (reflected color) and not its reflectance properties, which constitute the characterizing aspect of the surface appearance. The availability of a digital model encoding how a given surface reflects the light is of extreme importance if we would be able to see the digital 3D replica under different lighting conditions. Let me introduce here just a few

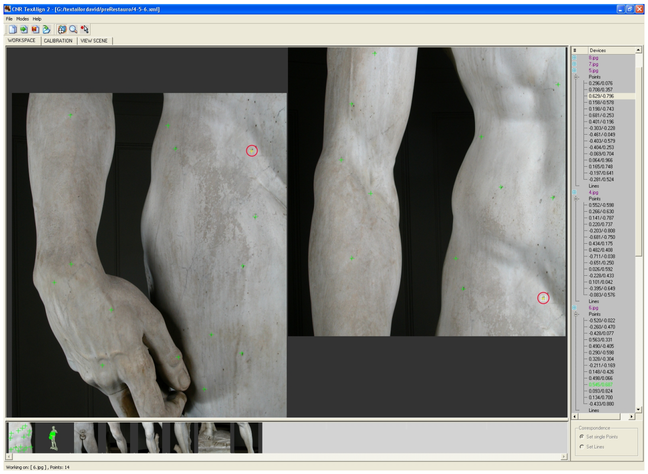

Figure 2: Screenshot of our Image Alignment tool with an example of partially overlapping RGB images (red circles indicate an image-to-image correspondence).

Figure 3: Two set of around 60 images each (depicting the pre- and post-restoration status) have been mapped onto the digital model of Michelangelos David and rendered in real time using the Virtual Inspecor system. Digital model courtesy of Stanford University (Digital Michelangelo Project) and Museo Gallerie dellAccademia, Florence.

examples to justify the need of dynamic illumination capabilities: being able to move a light source interactively around the digital replica and to synthesize accurately how the object is lighted, e.g. simulating razing light; reproducing different daylight conditions on an architecture; or simulating the visual appearance of an artifact or an architecture when it is lighted with different illuminants (electric light, different type of flames, solar light, etc).

Several accurate approaches for sampling the surface reflection characteristics have been proposed; a majorexample is the methodology devised by MPII to acquire the Bidirectional Reflection Function Distribution (BRFD) [11]. Unfortunately, most of these solutions are still too complicated to be massively applied to the CH field, where it is very hard to setup the controlled lab conditions needed to estimate the light reflection and diffusion. Moreover, since these methods makes use of controlled lighting conditions (usually in lab conditions) to sample the reflection function, they are nearly impossible to use on architectures.

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 [4]. 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. Some approaches have been proposed to reduce the aliasing and to produce seamless color mapping. A new flexible solution has been proposed in [2], where a multivariate blending function weights all the available pixel data with respect to geometric, topological and colorimetric criteria. The blending approach is efficient, since it mostly works independently on each image, and can be easily extended to include other image quality estimators. The resulting weighted pixels are then selectively mapped on the geometry, preferably by adopting a multiresolution per-vertex encoding to make profitable use of all the data available and to avoid the texture size bottleneck.

A basic problem in managing color information is how to register the images with the 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 and robust 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 (see Figure 2).

In a recent research we designed a new tool to support *image-to-geometry alignment*, *TexAlign* [8], whose main goals were: to reduce the user intervention in the process of registering a set of images with a 3D model; to improve the robustness 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 were a single image covers a region where the surface has not sufficient shape feature 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 instantiated 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* allows to reduce the time needed to perform the alignment and to improve the overall accuracy of the process. Some results are reported in [8]. This system has been recently used to map a complex photographic sampling (more than 70 images to be mapped on the David model, see Figure 3) and [6].

Considering the various technologies and methodologies used for 3D digitization, the subset of techniques for surface reflection acquisition and mapping on digital 3D models is the topic where greater is the potential for improvement to cope with the pressing requirements of CH applications.

Some results of high-quality mapping of color data on 3D meshs are presented in Figures 3 and 4.

Figure 4: An example of colored 3D digital model of one of the terrecotte that decorated the front of the Luni temple. On the left, the current color and on the right, a preliminary hypothesis of the original painted status.

4 CONCLUSIONS

As briefly presented in the previous sections, we think that the evolution and improvement of 3D scanning technologies makes this approach highly effective for applications in the CH domain. This technology is now affordable and satisfies the data accuracy and density required by many applications. We forecast a wide adoption in the near future. What still remains, in parallel with the further improvement of the technology (3D sampling HW, SW for geometric post-processing), is the required management of metadata and provenance data, which should be archived and managed through the entire workflow of geometric data.

ACKNOWLEDGEMENTS

We acknowledge the financial support of project *"START: Scienze e Tecnologie per il Patrimonio Artistico, Architettonico e Archeologico Toscano"* financed by Regione Toscana, Italy.

REFERENCES

[1] F. Blais. A review of 20 years of range sensor development. In *Videometrics VII, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol. 5013*, pages 62–76, 2003.

[2] M. Callieri, P. Cignoni, M. Corsini, and R. Scopigno. Masked photo blending: mapping dense photographic dataset on high-resolution sampled 3d models. *Computers and Graphics*, 32:464–473, 2008.

[3] M. Callieri, P. Cignoni, F. Ganovelli, C. Montani, P. Pingi, and R. Scopigno. VCLab's tools for 3D range data processing. In A. C. D. Arnold and F. Niccolucci, editors, *VAST 2003*, pages 13–22, Bighton, UK, Nov. 5-7 2003. Eurographics.

[4] M. Callieri, P. Cignoni, and R. Scopigno. Reconstructing textured meshes from multiple range rgb maps. In *7th Int.l Fall Workshop on Vision, Modeling, and Visualization 2002*, pages 419–426, Erlangen (D), Nov. 20 - 22 2002. IOS Press.

[5] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia. Meshlab: an open-source mesh processing tool. In *Sixth Eurographics Italian Chapter Conference*, pages 129–136. Eurographics, 2008.

[6] M. Dellepiane, M. Callieri, F. Ponchio, and R. Scopigno. Mapping highly detailed color information on extremely dense 3D models: the case of Davids restoration. *Computer Graphics Forum*, 27(8):(in press), 2008.

[7] A. Fasano, P. Pingi, P. Cignoni, C. Montani, and R. Scopigno. Automatic registration of range maps. *Computer Graphics Forum (Proc. of Eurographics '05)*, 24(3):(in press), 2005.

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

[9] R. Koch, M. Pollefeys, and L. V. Gool. Realistic surface reconstruction of 3D scenes from uncalibrated image sequences. In F.-E. Wolter and N. M. Patrikalakis, editors, *The Journal of Visualization and Computer Animation*, volume 11(3), pages 115–127. John Wiley & Sons, Ltd., 2000.

[10] S. C. G. Laboratory. Vrippack volumetric range image processing package. More info on: `http://www-graphics.stanford.edu/software/`, 2003.

[11] H. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transaction on Graphics*, 22(2):234–257, Apr. 2003.

[12] M. Pollefeys, R. Koch, M. Vergauwen, A. A. Deknuydt, and L. J. V. Gool. Three-dimensional scene reconstruction from images. In B. D. Corner and H. Nurre, Joseph, editors, *Proceedings of the Conference on Three-Dimensional Image Capture and Applications II (3d-image-capture-appl-00)*, pages 215–226, Bellingham, Washington, Jan. 24–25 2000. SPIE.