Automatic 3D modeling of archaeological objects

Marco Andreetto Nicola Brusco Guido M. Cortelazzo

Department of Information Engineering

University of Padova

Padova, ITALY

Abstract

A wide-spread use of 3D models in archeology application requires low cost equipment and technically simple modeling procedures. In this context methods for automatic 3D modeling based on fully automatic techniques for 3D views registration will play a central role. This paper proposes a very robust procedure which does not require special equipment or skill in order to make 3D models.

The results of this paper, originally conceived to address the costs issues of heritage's modeling, can be profitably exploited also in other modeling applications.

1. Introduction

3D modeling free-form surfaces is well established in some niche applications such as automobile industry, fiction and cartoon movies, special surgery and industrial mechanics, to name a few, where it is customary to model 3D geometry according to the following pipeline which has range data as input:

1. Pairwise registration of the 3D views.
2. Global alignment.
3. Fusion of 3D data into a single 3D surface.
4. Post processing (holes filling, simplification).

In the scientific community the interest for 3D modeling statues or, in general, cultural heritage's objects such as vases, jewelery, et similar, is gaining interest because of its technical challenges [1]. Major efforts in this area were the "Digital Michelangelo Project" [2] and the "Pietá Rondanini Project" [3], both concerning Michelangelo's artworks, the modeling of the relics of the Museum of Quin Shihuang Terra Cotta Warriors and Horses and of the 13m high Kamakura's Buddha. For a review see [4].

Metrological documentation, whether concerning some industrial application or heritage, is only feasible upon adequate financial resources. Therefore, in that context, automatic modeling and low cost acquisition equipment are somewhat secondary with respect to the data's accuracy. However, low costs and operation simplicity are major issues in order to replace the current iconographic material concerning archeology based on photographs with 3D models, for a better understanding and appreciation of 3D objects. Indeed it has been long recognized that 3D objects cannot be adequately represented by single pictures or collections of pictures, which can only reproduce one viewpoint or a collection of viewpoints, or even by movies, which can reproduce only one visualization trajectory: the one which was followed by the operator during the taking. What characterizes human inspection is the viewer's freedom to choose at any instant his viewing distance and angle in space. The possibility of dynamically visualizing any view of a 3D object upon users' interaction offered by 3D models is the representation tool closest to direct inspection.

With respect to cultural heritage the contribution of this paper concerns a technique for the completely automatic solution of the 3D views registration with great operational simplicity advantages: all one has to do is to make the 3D scans, from them a textured 3D model is automatically obtained. All museum personnel would have to know in order to make 3D models would be to learn how to use a range camera. This is really a minimal and unavoidable requirement in front of the complexity of 3D modeling real life objects.

This paper is organized in 4 sections. Section 2 introduces the techniques suitable to automatize all the phases of the registration process, i.e., the wide-baseline matching, the pairwise registration and the global registration. Section 3 presents a number of study cases and a discussion of the proposed method. Section 4 draws the final remarks.

2. Automatic 3D views registration

The first two steps of the 3D modeling pipeline traditionally concern the registration of all the 3D scans of an object, namely the mutual registration of all the pairs of views, called pairwise registration, and the subsequent registration of all the views, called global registration. The registration of a pair of partially overlapping 3D views is accomplished in two steps: a rough detection of the common region and a fine estimate of the rigid rotation and translation $(\boldsymbol{R}, \boldsymbol{t})$ bringing the two views to the best possible overlap (over the common region). The automatic coarse detection of the common region is an open research area. It is equivalent to the automatic detection of a number of corresponding points between the two views, therefore it is typically referred to as the wide-baseline matching problem and it is considered in Section 2.1. The fine registration of a pair of views is addressed in Section 2.2.

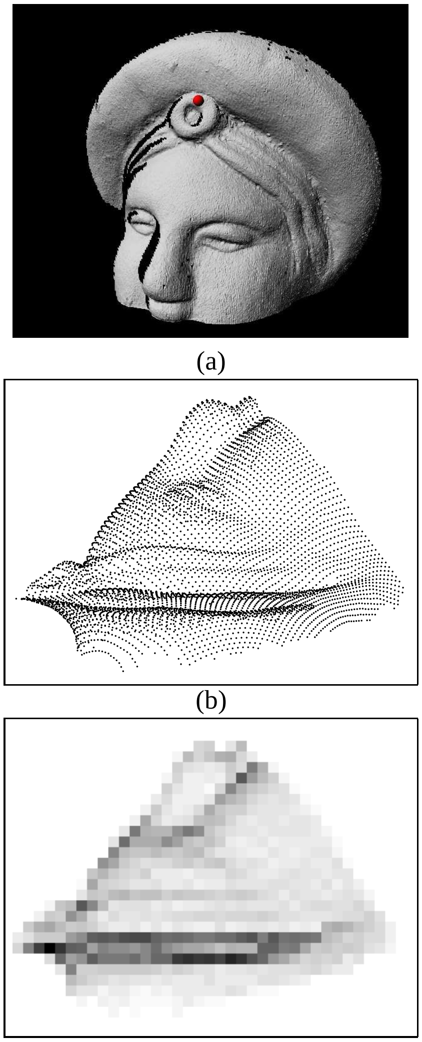
Before entering the details of pairwise registration it is appropriate to remind that even in front of very accurate results for each pair of 3D views, the pairwise registration procedure alone cannot model full objects because of error accumulation. The errors may come from the pairwise registration procedure itself as well as from measurement noise and discretization effects on the acquired surface. There are several methods for redistributing the registration errors upon the results produced by a pass of pairwise 3D views registration, for a recent literature review see [5].

2.1 Automatic detection of the common region

There exist two fundamental approaches to the automatic detection of a rough common region between a pair of partially overlapping 3D views. One rests on the use of invariant statistics for each point of the mesh such as [6] [7], the other on the use of special features based either on geometry such as [8], or on texture [9]. We reference only the above two works because there is no room in this paper for a larger reference list.

This section reports about the results we obtained with the method of the spin-images introduced by Johnson and Hebert in [7], where the reader is referred for a detailed presentation.

Fig. 1a) shows a view of a 3D model of a Pompeian statuette, hereafter called “Lady”, see Table 1. As Fig. 1 exemplifies, spin-images depend only from the intrinsic surface characteristics (or statistical properties) and not from the surface’s spatial position and orientation. In other words the spin-image associated with a vertex point is invariant with respect to rigid rototranslations.



(a)

(b)

(c)

Figure 1: Lady: (a) A single 3D view (b) Spin-map (c) Spin-image

It is also very important to observe that the spin-images of neighbour points are very similar because the features of the local surface area near the point have a greater impact in the spin-image formation. Therefore, by way of spin-images, one may associate a collection of images to a 3D surface mesh, as every point of the surface can generate a spin-image. A pair of surfaces representing the same object from different view-points will be associated to a pair of sets of different spin-images: corresponding points in the common region between two 3D views will have similar spin-images because, as already pointed out, spin-images strongly depend on local shape’s characteristics. The spin-images of corresponding points of two partially overlapping 3D views will not be identical because of surface discretization effects and because the two patches share only a portion of their surface. However, if the overlap is substantial (no less than 30% of regions characterized by an adequate

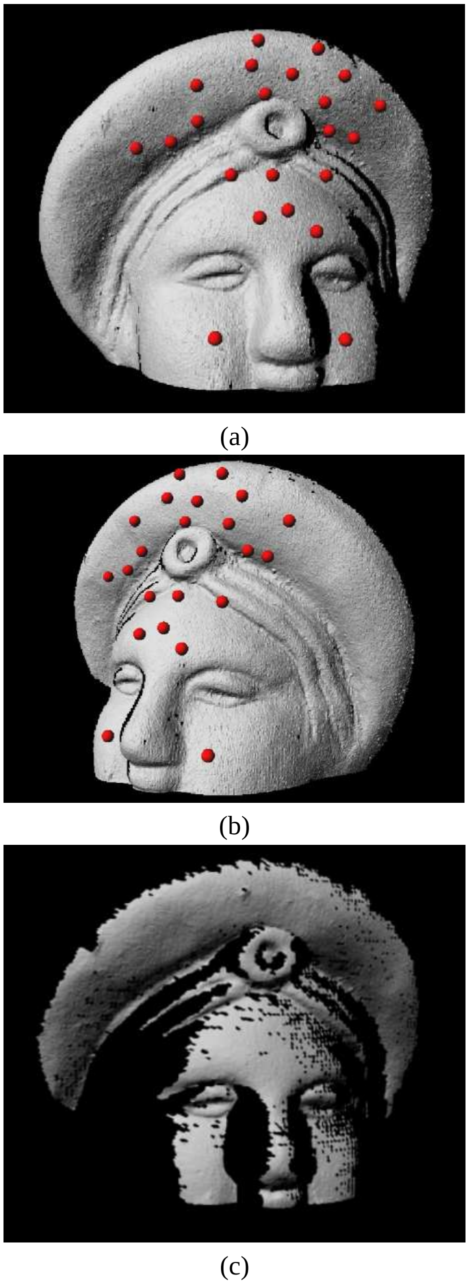

(a)
(b)
(c)

Figure 2: (a)-(b) Example of putative correspondences between two partially overlapping meshes (c) The estimated common region

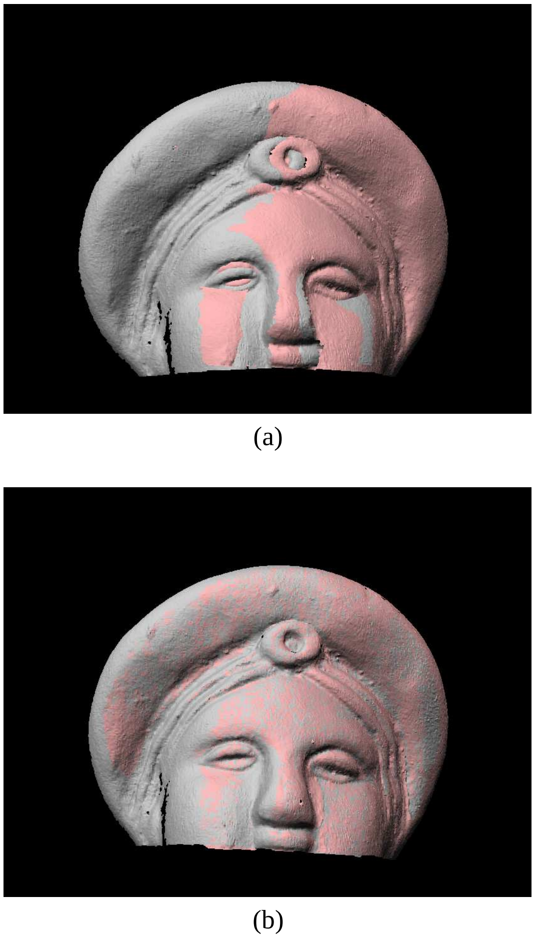

(a)
(b)

Figure 3: (a) Automatic prealignment by the spin-images method; (b) Final pairwise registration by the frequency domain method followed by the ICP

presence of geometrical features) corresponding points of the two 3D views located in the common region will have similar spin-images. The detection of the common region between two partially overlapping 3D views in this way can be turned into the recognition of the most similar images of two sets of (spin) images, a problem for which a number of techniques are available.

In order to detect corresponding points a set of putative point correspondences is determined upon the similarity of the spin-images, as the example of Fig. 2 shows.

The correspondences are organized within geometrical consistent groups and for each group one determines the rototranslations $(\boldsymbol{R}, \boldsymbol{t})$ moving the points of the first 3D view as close as possible to their corresponding points in the second 3D view, via Horn's algorithm. The rototranslation $(\boldsymbol{R}, \boldsymbol{t})$ giving the widest overlap between the two views is selected. The common region between the two views is taken to be this overlapping area.

Fig. 3a) shows the the rough alignment of the two 3D views of of Fig. 2 obtained by bringing to coincidence the corresponding points found by the spin-images method. The geometrical artifacts due to the poor estimate are clearly noticeable in the central part of the face; see, for instance, the doubling of the Lady's little disk (in the middle of the front) and nose.

2.2 Automatic pairwise registration of 3D views

The final estimate of the rotation and translation $(\boldsymbol{R}, \boldsymbol{t})$ between two 3D views is typically accomplished by the ICP algorithm [10] or some of its many variants. The rough prealignment obtained by spin-images is often adequate to bring the ICP to convergence. However this was not found to be always the case, consistently with the fact that the ICP algorithm is known to be able to produce very precise estimates of $\boldsymbol{R}$ and $\boldsymbol{t}$ when properly started, as it may otherwise be trapped by local minima. This problem may be overcome by initiating the ICP from different starting points, strategies for their selection are proposed in the literature. In practice this issue is commonly solved by manually selecting a few corresponding points from which an estimate of $(\boldsymbol{R}, \boldsymbol{t})$ is determined by Horn's algorithm.

In order to have a fully automatic procedure we use instead the frequency domain method of [11] as a robust device for determining a first estimate of the $(\boldsymbol{R}, \boldsymbol{t})$ parameters to use as starting point for the ICP. The method of [11] and the ICP operate on very different principles and their combination turns out remarkably robust. It may be worth quickly recalling the rationale of both methods in order to motivate their synergy.

As well known, the general idea behind the ICP algorithm is to find a set of matching points on the overlap of the two 3D views and to minimize the distance between these correspondences, in the Euclidean sense or according to similar criteria. It can be proved that this procedure suitably iterated converges to a local minimum [10].

The method of [11] rests upon the fact that if textured surface $l_2(\boldsymbol{x})$ is a version of $l_1(\boldsymbol{x})$ rigidly rotated and translated by $(\boldsymbol{R}, \boldsymbol{t})$, i.e,

$$l_2(\boldsymbol{x}) = l_1(\boldsymbol{R}^{-1}\boldsymbol{x} - \boldsymbol{t}), \tag{1}$$

the Fourier transform of $l_1(\boldsymbol{x})$ and $l_2(\boldsymbol{x})$, respectively denoted as $L_1(\boldsymbol{k})$ and $L_2(\boldsymbol{k})$ are related as

$$L_2(\boldsymbol{k}) = L_1(\boldsymbol{R}^{-1}\boldsymbol{k})e^{-j2\pi\boldsymbol{k}^T R\boldsymbol{t}}. \tag{2}$$

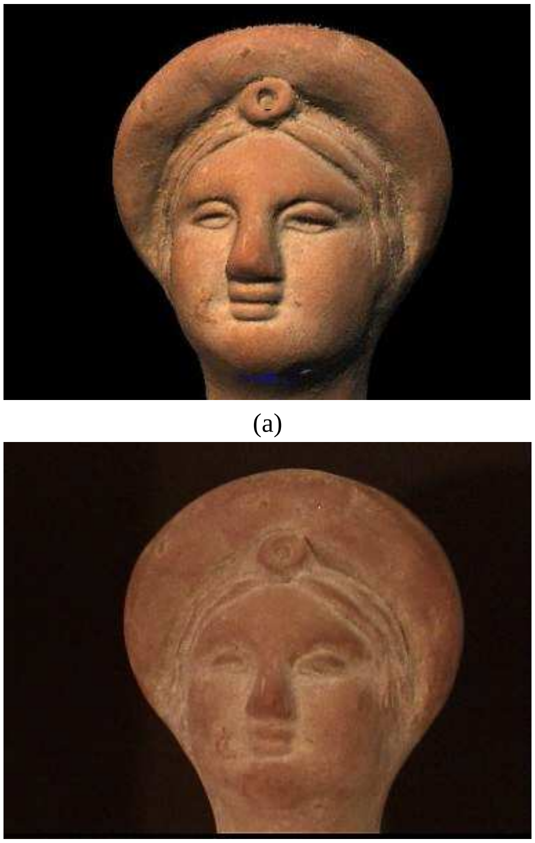
In terms of magnitudes Equation (2) simplifies in

$$|L_2(\boldsymbol{k})| = |L_1(\boldsymbol{R}^{-1}\boldsymbol{k})|. \tag{3}$$

It can be proved [11] that the versor of the rotation axis, is the vector $\boldsymbol{k}$ which equates to 0 the difference function

$$\Delta(\boldsymbol{k}) = \left| \frac{|L_1(\boldsymbol{k})|}{L_1(\boldsymbol{0})} - \frac{|L_2(\boldsymbol{k})|}{L_2(\boldsymbol{0})} \right|. \tag{4}$$

An effective technique to solve this minimization problem is given in [11], together with a procedure for estimating the rotation angle upon the rotation axis, and the translation vector $\boldsymbol{t}$ upon knowledge of $\boldsymbol{R}$.


(a)

(b)

Figure 4: (a) A view of the final textured model (b) a picture of the actual statue

We want to point out that the frequency domain algorithm of [11] does not operate on a point-to-point correspondence logic as the ICP, since the Fourier transform makes a synthesis of all the available spatial information. Furthermore, since the Fourier transform doesn't only take into account the 3D shape information, but also the texture information, the method of [11] is able to register textured symmetrical objects (such as painted vases, that could not be aligned on the basis of local geometrical information only). This possibility plays a rather valuable service for practical modeling of heritage's objects.

Since the frequency domain method does not operate directly on the 3D surface data, but it has to turn them into small volumes, it is not as precise as the ICP.

3. Experimental results and discussion

As first example, we consider the automatic modeling of "Lady" introduced in Section 2. Some data about this piece are reported in Table 1. The spin-images method was able to automatically find the common regions between all the pairs of partially overlapping 3D scans and to preliminarily align them by bringing to coincidence the common regions as shown in Fig. 3a) for a pair of upper scans. Table 1 reports the average execution times on a Pentium III 1.8 GHz, which is 27.40 s.

The displacement error of a single point is the distance between its position after the prealignment by way of spin-images and its final position after global registration. Some spread in the displacement errors due to the error accumulation of the pairwise alignment procedure is expected.

The values of the average registration error between corresponding points for each pair of overlapping 3D views, prealigned by the spin-images method turn out rather uniform. This indicates a regular performance of the spin-images method with data with uniform characteristics. The comparison of the average registration error after the prealignment by way of spin-images with the average registration error after global registration, which is reported in Table 1, indicates that they are about three orders of magnitude greater.

The proposed procedure can be applied to objects acquired by any number of scans provided that the overlap between the scans is around 30% and that the common region be adequately characterized. This means that one must exercise some care during the taking in order to scan according to the above requirements.

As Fig. 4 shows, the model is visually coherent both geometrically and colorimetrically with the original. The differences between the actual picture of the original object shown in Fig. 4b) and the view of the 3D model of Fig. 4a), are mostly due to the flash light used in the actual taking of the picture (which produces a kind of uniform strong illumination not easy to emulate with synthetic lights) and to a much lesser extent to the approximate replica of the picture's viewpoint.

As a second study case we consider the model of an Etruscan skull, found in a tomb. This object was acquired without texture since we were told that only its geometry is of anthropological archaeology's interest. The basic data about this object are given by Table 2.

The average execution time is about twice the average execution time of "Lady", since the scans have a much greater number of object's points. The average displacement error of the skull is comparable with that of "Lady" in spite of a couple of views concerning the back of the skull with high displacement error peaks, due to their specific characteristic.

4. Conclusions

A specific points of interests of the proposed method for the automatic realization of 3D models is the use of the frequency domain technique of [11] as a robust device for the automatic determination of effective starting points of the ICP algorithm upon the crude registration parameters obtained by the method of the spin-images.

(a)

(b)

Figure 5: Etruscan skull: (a) Some 3D scans (b) A view of the model

The proposed technique, as shown in Section 3, is suited to the automatic modeling of archaeological objects, which is considered one of the most challenging applications of 3D modeling.

In the introduction we explained that low costs and limited technical skills are essential for the widespread practice of 3D modeling in archeology and in this connection automatic 3D modeling procedures are essential. Automatic procedures may also play valuable services in many other areas: we are currently testing variations of the proposed method in the 3D modeling of palatal calcs and of anatomical surfaces which are applications of odontosthomatological and plastic surgery's interest respectively.

Automatic 3D modeling can be improved in many aspects. The wide baseline matching problem is a central issue. Methods capable to deal with symmetric objects, to efficiently take into account structure and texture informa-

“Lady”	
Height	187.9 mm
Total no. of scans	24
Memory occupancy of a single scan (PIF format)	1.659 Mb
Memory occupancy of the whole VRML model	57,7 Mb
Texture type	RGB
Average displacement error	$4.32\ mm$
Average pair-wise registration error	$0.402\ mm$
Average execution time	$27.40\ s$
Final registration error	$3.405 \cdot 10^{-4}\ mm$
Average triangle side	$0.217\ mm$
Total number of triangles	2084308
Total surface area	12996.08 mm^2

Table 1: Data about “Lady”

“Etruscan Skull”	
Height	154.28 mm
Total no. of scans	17
Memory occupancy of a single scan (PIF format)	6.32 Mb
Memory occupancy of the whole VRML model	58.6 Mb
Texture type	-
Average displacement error	$5.82\ mm$
Average pair-wise registration error	$0.813\ mm$
Average execution time	$46.81\ s$
Final registration error	$1.536 \cdot 10^{-3}\ mm$
Average triangle side	$0.548\ mm$
Total number of triangles	741358
Total surface area	76690 mm^2

Table 2: Data about the Etruscan Skull

tion, to be more accurate than the existing ones are currently investigated. Among the various approaches the one of [8] looks very promising. We are currently experimenting with extensions of the concept of spin-images which also include texture and not only shape information.

The use of inexpensive acquisition equipment is also very important for cost containment, in this connection it would be worth exploring the practical performance in cultural heritage's applications of the acquisition methods of [12] [13] [14].

Acknowledgments

The examples shown in this paper come from a 3D modeling project between the Archaeological Museum of ..., the ... consortium of ... and our group. We would like to acknowledge ... , director of the Archaeological Museum of ... for their assistance and guidance. This work has been partially supported by ... .

References

[1] G. Godin et al., “Active Optical 3D Imaging for Heritage Applications,” *IEEE Computer Graphics and Applications*, pp. 24–36, September-October 2002.

[2] M. Levoy et al., “The digital Michelangelo project: 3D scanning of large statues,” in *Proceedings of SIG-GRAPH Computer Graphics Proceedings, Annual Conference Series*, 2000, pp. 131–144.

[3] F. Bernardini, I. Martin, J. Mittleman, H. Rushmeier, and G. Taubin, “Building a Digital Model of Michelangelo's Florentine Pietà,” *IEEE Computer Graphicd and Application*, vol. 22, no. 1, pp. 59–67, Jan-Feb 2002.

[4] F. Bernardini, I. Martin, and H. Rushmeier, “High-quality texture synthesis from multiple scans,” Tech. Rep., February 2000.

[5] H. Rushmeier F. Bernardini, “The 3D Model Acquisition Pipeline,” *Computer Graphics Forum*, vol. 21, no. 2, pp. 149–172, 2002.

[6] K. Higuchi, M. Hebert, and K. Ikeuchi, “Building 3D models from unregistered range images,” *Graphical Models and Imag. Proc.*, vol. 57, no. 4, pp. 315–333, 1995.

[7] A. E. Johnson and M. Hebert, “Using Spin-Images for Efficient Multiple Model Recognition in Cluttered 3-D Scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.

[8] J. Vanden Wyngaerd and L. Van Gool, “Coarse registration of surface patches with local symmetries,” in *Proc. European Conference on Computer Vision (ECCV'02)*, May 2002, vol. 2351, pp. 572–586.

[9] G. Roth, “Registering two overlapping range images,” in *Proceeding of the Second Intl. Conf. on 3D Digital Imaging and Modeling*, Ottawa, Canada, October 1999, pp. 191–200.

[10] P. J. Besl and N. D. McKay, “A method for registration of 3D shapes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–259, November 1992.

[11] L. Lucchese, G. Doretto, and G. M. Cortelazzo, “A frequency domain technique for 3-D view registration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1468–1484, November 2002.

[12] R. Cipolla and P. Giblin, *Visual Motion of Curves and Surfaces*, Cambridge Univesity Press, Cambridge, United Kingdom, 2000.

[13] J. Bouguet and P. Perona, “3d photography using shadows in dual-space geometry,” *International Journal of Computer Vision*, vol. 35, no. 2, pp. 129–149, November-December 1999.

[14] S. Savarese, H. Rushmeier, F. Bernardini, and P. Perona, “Implementation of a shadow carving system for shape capture,” in *Proc. of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT2002)*, Padova, Italy, 2002, pp. 12–23, IEEE Press.