Automatic coarse registration of range maps

Andrea Fasano, Paolo Pingi, Paolo Cignoni, Claudio Montani, Roberto Scopigno

***Abstract*— Range map registration is still the most time consuming phase in 3D scanning. This is because real scanning set are composed of hundreds of range maps and their registration is still partially manual. We propose a new method which allows to manage complex scan sets acquired by following a regular scanner pose pattern. The method makes proficient use of some initial considerations and auxiliary data structures, which allow to simplify the range map registration problem. It is designed as an iterative solution, where pairs of correspondent vertices are selected through the computation of a regular $n \times n$ kernel which takes into account vertex normals and is defined in the 2D space of the range map (represented in implicit 2D format rather than as a triangle mesh in 3D space). The shape-characterization kernel and the metrics defined give a sufficiently accurate shape matching, which has been proven to fit well the requirements of automatic registration. The solution proposed has been tested on a number of complex scanning set and results are impressively better than previous solutions.**

***Index Terms*— 3D scanning, automatic registration, coarse registration, range maps alignment**

I. INTRODUCTION

THE increasing diffusion of 3D scanning devices and the design of new and efficient algorithms for range data post-processing are at the base of a process where standard CAD tools are going to be replaced by a semi-automatic process based on the direct sampling of real objects' shape. Moreover, automatic acquisition of shape and appearance is no more confined to the classical industrial applications (reverse engineering or quality control), but it is positively affecting new and important fields. In this sense, Cultural Heritage (CH) is probably one of the most indicative and challenging applications. The availability of accurate 3D digital models is becoming a demanding requirement for the knowledge, conservation, restoration, and promotion of CH [13], [1], [10], [25], [12], [19].

Unfortunately, the creation of a digital 3D model from reality is still far enough from being as simple as photography. The user has to manage many complex processing steps (range maps acquisition, registration, fusion, geometry simplification, color attributes recovery). Solutions for a completely automatic scanning system have been proposed, but either these systems are based on the use of complex positioning machinery [13], [4], or adopts silhouette-based approaches which do not guarantee the needed accuracy [27], [29]. An alternative approach is to design new solutions for the classical scanning pipeline which would transform those phases into an unattended process. Many innovative and efficient solutions have been recently proposed for the different phases of the scanning pipeline and, at present, it is possible to assert that the bottleneck of the whole process is represented by the *range maps registration* phase, since this is the only task where a considerable human intervention is still requested.

The accurate acquisition of a real object requires to take many range maps from different locations. If the scanner location and orientation are not tracked, all those range maps are produced in different coordinate spaces (each one depending on the corresponding unknown location and orientation of the scanner). The goal of the range map registration phase thus consists in determining the rigid geometric transformations able to bring back all the coordinates of the acquired data into a unique Cartesian space. Registration is the fundamental precondition to merge all the data into a single and complete digital model. The explicit registration of multiple scans can be simplified or even avoided by adopting different techniques/devices which help in tracking the scanning device (see Section II). However, these solutions usually increase the overall cost of the scanner, support an accuracy which in some cases is much lower than the nominal accuracy of the scanning system, or introduce constraints on weight and working space which make them hard to use in uncontrolled working environments (e.g. acquisitions done in museums or archeological sites). For these reasons, we are interested in avoiding to add gantries or other devices to our standard scanner. Our goal is to design new solutions to make the alignment of range maps a nearly automatic process.

The registration of multiple range scans is implemented by adopting a software approach split in two computational steps. An initial *coarse registration* provides a rough positioning of the range maps, and a

Visual Computing Lab, Istituto di Scienza e Tecnologie dell'Informazione (ISTI), Consiglio Nazionale delle Ricerche (CNR), Via G. Moruzzi, 1, Pisa, Italy. Contact email: fasano@isti.cnr.it , pingi@isti.cnr.it

subsequent *fine registration* brings the scans into tight alignment. In other words, coarse registration is concerned primarily with determining which regions of two scans represent the same portion of the object's surface (the so called "overlapping" region), while fine registration is concerned with minimizing the mismatch between these corresponding overlapping regions.

The solutions proposed in the literature for the *fine local* registration (between two scans [2], [14]) or *fine global* registration (extended to the whole scanning set [20]) are generally based on unattended iteration of the ICP technique and by now they ensure performances and results extremely good. On the contrary, *coarse* registration represents the real bottleneck. Despite the number of solutions proposed (see Section II), most of the commercial software systems available on the market implement the coarse registration adopting an interactive solution; however, this manual process is time-consuming. According to the experience of the authors, a complex scanning set composed of some hundreds of range maps can require some days of hard and boring work to process the manual coarse registration phase.

This paper presents an efficient and innovative automatic solution to the problem of the coarse registration. With respect to the existing feature-based algorithms (i.e. the solutions aiming at the heuristic location of corresponding points in the common parts of the two scans under examination, see Section II) our method presents some innovative aspects which make it particularly robust and efficient for practical 3D scanning campaigns. In Section III we introduce the strategy used to register multiple range maps, while Section IV contains the details about the alignment algorithm given two overlapping range maps. The proposed algorithm originates from the practical need to register hundreds of range maps rather than from theoretical considerations. For this reason, some of the operative choices we adopted can appear disputable from a theoretical point of view but they allowed us to obtain accuracies and performances really good without losing the generality of the algorithm. For example, our method to find corresponding points on the surface of two overlapping scans is not invariant to "heavy"changes of the up direction of the scanning device. This is a limitation in the general case, but it does not represent a problem for standard 3D scanning, where rotating the scanner along his view axe is rather uncommon.
Finally, the results of empirical evaluations of our approach on complex scanning set are presented in Section V. Section VI closes the paper with final remarks and future work.

II. Previous work

The alignment task is the most time-consuming phase of the entire 3D scanning pipeline, due to the substantial user contribution required by current systems, since range map registration is usually solved by adopting a partially manual process. The standard approach is as follows:
Local pairwise phase

- *Overlapping range maps detection*: for each range map R_i in the scan set $\overline{R}$, detect all R_j in $\overline{R}$ which are partially overlapping with R_i. This pairwise process can be considered as a graph problem: given the nodes (i.e. the range maps), we have to select a subset of arcs such that every node is linked to some others if they are partially overlapping, and thus have to be aligned together (*graph of overlaps*). If the set of range maps is composed by hundreds elements (the scanning of a 2 meters tall statue generally requires from 200 up to 500 range maps, depending on the shape complexity of the object), then the user has a very complex task to perform;
- *Initial Coarse Registration*: provide a first rough registration between each pair of overlapping range maps (R_i, R_j). The initial placement is heavily user-assisted in most of the commercial and academic systems. It usually requires the interactive selection and manipulation of the range maps, either to select a small set of corresponding point pairs or to superimpose the range maps by means of interactive rotations and/or translations;
- *Fine Pairwise Registration*: the scans are finely aligned, usually adopting the Iterative Closest Point process (ICP) [2], [6], [14] which minimizes the alignment error between any pair of range maps;

Global registration

- The pairwise registration produces good results but, since the error minimization takes place sequentially on mesh pairs, the error tends to accumulate and it may result in significant artifacts after a number of pairwise steps. A solution is to perform a global minimization process which distributes the residual error among all pairs in order to spread the error evenly among all range map pairs [20].

The precision of the *coarse registration* phase does not need to be really high because the convergence of the successive *fine registration* is ensured even when the accuracy is low (e.g. a few millimeters of distance between the two maps); however, the manual intervention required for coarse registration is time-consuming and

boring, since it may require some days of work on a complex scanning set.

As already briefly discussed in the introduction, an orthogonal strategy to the one presented here is to add a tracking subsystem to the scanning device, to allow to track its position and orientation in the scanning space. *Tracking solutions* can be based on: magnetic trackers, optical trackers, or accurate mechanical systems (like rotary platforms or robotic arms) for the controlled translation/rotation of the object/scanner. Just to give a few examples: a laser scanner was mounted on a six DOF robotic arm to obtain an automatic scanning and rapid reproduction station [4]; a computer-driven gantry has been designed for the Digital Michelangelo project [13]; optical and magnetic tracking are used in commercial scanning systems [18], [24]. Moreover, physical targets can be introduced into the scene to be scanned or onto the object under examination, to simplify the subsequent software-based registration.

Automatic coarse registration was investigated quite intensively in the last few years, even if a completely automatic and general-purpose solution is not yet found. It is possible to distinguish different approaches to solve the problem. It's not our intention to present here a complete survey of the algorithms proposed in the literature (a good survey paper is [5]); for the sake of conciseness, we will focus here only on some seminal proposals.

The main idea to perform automatic range maps alignment is to *characterize* the *shape* represented by the range maps and to use this characterization to find the correspondences needed for automatic coarse registration. Therefore, some solutions proposed in literature are very general (in some cases, can be conceived as a by product of the *shape similarity* research [28]), while others are more specific. The latter approach is the one we followed, since making profitable use of the peculiar characteristics of the range map registration problem allows to solve it in a more efficient manner.

General solutions define a mapping from the global surface model to some fixed-dimensional vector space, by taking into account the entire mesh (or range map). Lucchese et al. [15] introduced a method operating in the *frequency domain*: to estimate the roto-traslation between a pair of meshes they propose a featureless algorithm which exploits the information about the geometric regularity captured by the Fourier transform.

Delingette et al. ([9], [8]) developed the *Spherical Attribute Image* (SAI) method that defines a direct mapping between an object surface and a spherical surface, thus obtaining a unique representation of any non-convex object (or part of it); the correspondence between two range maps can be obtained from a comparison of the respective SAIs. A similar approach is the *spherical extent function* [30].

Harmonic maps have been adopted as surface descriptors, among others, by Funkhouser et al. [16]. They proposed the *spherical harmonic representation* which provides a rotational-invariant representation of a shape descriptor based on spherical harmonics, enhanced in [17] by factoring out the contribution of anisotropy and geometry.

Other registration techniques adopt the second approach mentioned above (local characterization). One important consideration is that *vertex-based* shape characterizations are more adequate than *object-based* ones in the case of partial and incomplete knowledge (range maps are a very incomplete representation of the overall shape and they overlap only in a small subregion). Most of the methods proposed use some sort of evaluation of the curvature of the mesh (e.g. bitangent curves are used in [31] to characterize mesh vertices).

Stein and Medioni [23] introduced the *splash images*, i.e. small surface patches used to detect local changes in the surface orientation. *Splash images* are then used as primitives to measure the differences between surface normal distributions (this proposal was one of the main inspirations for our work).

Chua and Jarvis [7] proposed the *point signature* , a representation invariant to rotation and translation that encodes the minimum distances of the points on a 3D contour (intersection of the surface with a sphere centered in the point under analysis) to a reference plane. *3D point fingerprints* [26] can be considered an extension of the previous approach. The point fingerprint is a set of 2-D contours that are the projections of geodesic circles onto the tangent plane.

Johnson and Hebert [11] proposed the concept of *spin image*, a more descriptive structure in respect of *splash images* and *point signatures*. A *spin image* is generated using oriented points - 3D points with directions: a *spin image* is a 2D histogram of the surface locations around a point. Matching points that rely on different views of the model have similar *spin images*, so they can be used to find the correct correspondences.

The availability of color or other surface attributes sampled together with the geometry has been used either to improve the accuracy of geometric registration or to perform automatic coarse alignment [21].

Our method has been devised to make a proficient use of all the specific characteristics of the problem considered – the coarse registration of range maps. We propose a new *vertex-based* shape characterization which

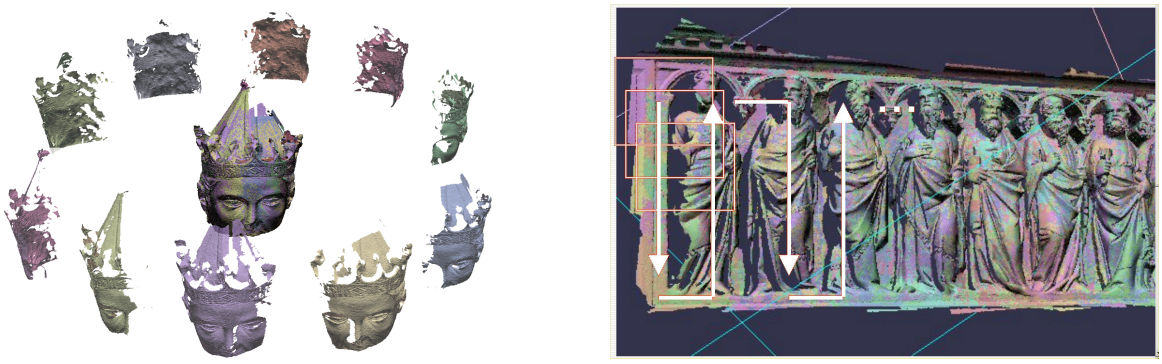

Fig. 1. Range maps are taken in a row-wise order: an example of *circular* stripe around a statue's head (left); an example of *raster-scan* scanning order adopted for the acquisition of a bas-relief (right).

works directly on the range map rather than on the associated triangle mesh.

III. Multiple range maps registration problem

The standard registration task can be simplified by considering some practical aspects. First, the *detection of overlapping range maps* can be reduced to a simpler task: 3D acquisitions are usually done by following a simple selection of the scanner poses. Users usually acquire range maps in *stripes*, following either a *vertical*, *horizontal*, *raster-scan* or *circular* translation of the scanning system (see Figure 1). The different types of stripes 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 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. In practice such coherence is not sufficient to successfully initialize ICP, so approaches like the one in [22] that rely on straightforward application of ICP cannot be adopted.

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 stripes pattern defined above are equivalent: an automatic registration module can processes each couple (R_i, R_{i+1}), to produce in output the roto-translation matrix M_i that aligns R_{i+1} to R_i.

The subset of registration arcs defined above is usually sufficient for a successive ICP application, in order to obtain a fine registration of the range maps. Anyway, depending on the shape of the object and the acquisition pattern, some more arcs are usually needed (interconnecting R_i with all the overlapping range maps). Like other alignment systems, we do not require the explicit creation of all these arcs but we use a *spatial indexing* technique to avoid the user to introduce all these arcs manually. Given a subset of links (where a single arc is known for each node, computed as we propose in the next section), the *graph of overlaps* can be completed in an automatic manner. A *discrete space bucketing* data structure can be easily instantiated, holding for each bucket (a small 3D voxel) the set of range maps passing through that region of space. The initialization of this data structure requires the scan-conversion of every range map in the discrete space. We can easily retrieve groups of overlapping range maps by a simple visit of the bucketing structure, and tell how significant are those overlap extents. Obviously, since we reconstruct the occupancy grid using a discrete resolution which is usually an order of magnitude lower than the spatial sampling used in scanning, the information contained gives an approximate overlapping graph and overlapping factor, but still sufficiently accurate for our purposes. Given the occupancy grid information and once a single alignment arc is provided for each range map, our registration library is able to introduce all needed arcs (in a completely unattended manner, by selecting only those which satisfy a minimum-overlap factor) and to process them using the ICP algorithm. Our tool, *MeshAlign* [3], implements this type of solution to provide automatic fine alignment graph completion.

In conclusion, this approach reduces a $1\ over\ n$ problem into a $1\ over\ 1$ problem (for each range map, find coarse registration matrices for all the overlapping ones in the set of n range maps).

The stripes approach can be seen as an efficient working strategy, in opposition to the more general task to determine a complete adjacency graph.

IV. AUTOMATIC PAIR-WISE RANGE MAP ALIGNMENT

Therefore, we restrict the general problem to the coarse pairwise alignment of two overlapping range maps R_a, R_b.

We assume that the two range maps R_a, R_b, overlap on a reasonable portion of their surface (15-20%). Like other surface matching algorithms, we look for a small set of *feature points* which characterize the first range map R_a: a point-based shape description kernel is proposed in Section IV-A. Then, in a second step, for each of these k points on R_a we search for the potential corresponding points on the second mesh R_b (Section IV-B). Finally, out of those possible k pairs we choose the group of four matching points which gives the best coarse alignment (Section IV-C); if the needed accuracy is not reached, we iterate until convergence (Section IV-D).

A. Starting points selection

A trivial approach would be to choose the starting points $p \in R_a$ random over the range map. Unfortunately, this approach can produce non-representative samples that reside into surface areas having little (or none) geometric features and makes very hard the accurate identification of the corresponding points in the second range map R_b. For this reason, we need a *measure* of how well a given vertex can be representative of the corresponding range map. For the sake of simplicity we consider our input meshes as regularly sampled *2D height fields*. We find this approach effective (as explained in detail in the following), even if this assumption could not hold for some rather infrequent situations (like for example the adoption of a scanner implementing a not regular sampling pattern). The height field assumption gives us a simple local parameterization of the surface that allows us to easily detect for each point p a small and regular kernel of adjacent samples, from which we compute its *variance*. Given a *pivot* vertex $p \in R_a$ and its normal vector N_p, we build a *kernel* K_p by considering the $n \times n$ square of points around p in the range map (represented as a 2D raster). In our implementation, we adopt a kernel size of 13×13. Each element $k_{i,j} \in K_p$ contains the dot product of the pivot's normal vector and the normal vector of the corresponding pivot's neighbor. Then, we calculate the variance of each kernel $K_p \in R_a$:

$$s^2(K^p) = \frac{1}{n^2} \sum_{i,j} (k^p_{i,j} - \mathbf{E}[K^p])^2 \qquad (1)$$

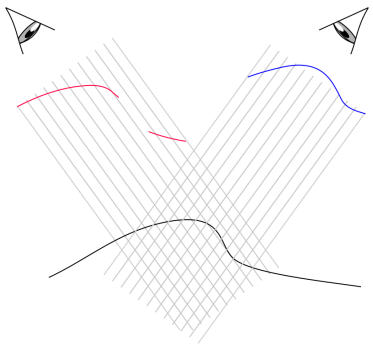

Fig. 2. Example of silhouette edges or self occlusion between two successive view points.

where $\mathbf{E}$ is the average of each kernel:

$$\mathbf{E}[K^p] = \sum_{i,j} \frac{k^p_{i,j}}{n^2} \qquad (2)$$

The variance is used to cluster all range map points in buckets characterized by a similar surface curvature. Low values of $s^2(K^p)$ are relative to flat areas where the normal vectors are relatively uniform. On the other hand, high values of $s^2(K^p)$ correspond to zones having high curvature. Note that if a mesh has open boundaries, then vertices on the proximity of these boundaries produce a high variance value, due to the absence of information on kernel points lying outside the surface.
We have chosen to discard all the points having either high or low variance, using two opportune threshold values selected according to empirical experience. There are several motivations for this choice. First, we obviously discard *flat areas* because they cannot give sufficient information to detect the correspondent point in the other mesh. We discard also the vertices with *high variance*, even if this could seem less intuitive. First, high variance vertex could belong to open boundaries (and thus to an incomplete local sampling of the surface); they have to be discarded since the matching range map might have a more complete sampling of the same zone, that will hardly match with the previous one. Second, a high variance vertex could be generated in the proximity of a silhouette vertex (where we have a false step due to a self occlusion of the mesh, see Figure 2). This portions of the mesh are dangerous because these steps do not exist in reality, but depend on the scanner location at the time of the take of range map R_a; the same portion of the surface, seen from a different point of view, could have a very different shape descriptor. Figure 2 shows how silhouette edges or self occlusion can create a false high-variance ridge that does not really exist on the surface object.

Once the bad vertices are discarded, a small set of candidate starting points are chosen randomly among the remaining points (the ones with medium variance). The number of these points is usually small, around 20 in our empirical tests (see Section V).

B. Finding a matching point

Then, given a selected vertex $p \in R_a$, we have to find the best matching vertex $q \in R_b$ (if it exists). The matching algorithm's quality is tightly bounded to the metric used for the comparison. A perfect metric should adopt mesh attributes which are invariant with respect to the roto-translation that occurs between the range maps. Our method builds up on the same kernel defined in the previous subsection. In particular, we compute the kernel for every vertex $q \in R_b$. Given $p \in R_a$ and its kernel K_p, the metric consists in finding the more similar kernel K_q relative to the point $q \in R_b$. So, for each K_q, we compute the squared difference with K_p:

$$d^2(K_p) \quad = \quad \frac{1}{n^2} \sum_{i,j} (k^p_{i,j} - k^q_{i,j})^2 \qquad \forall q \in R_b \quad (3)$$

and we choose as best potential matching point the one having minimum distance $d^2(K_p)$.
This kind of metric is invariant 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 view axe is rather uncommon (the scanner is usually connected to a tripod, which makes impossible to apply a substantial rotation along the view axe).
Obviously, this metric does not ensure the convergence to the correct matching: the selected corresponding point could be incorrect, since multiple vertices can present a shape signature similar to the one of the vertex considered (or because a point on a non-overlapping region of R_b can be very similar to the selected point p of R_a). Therefore, we need to validate the matching.
Other approaches exist that try to avoid as much as possible any false match by defining more complex shape signatures. We have chosen to follow a strategy which couples a computationally efficient shape descriptor with a further validation phase which proves the correctness of (or purges) the selected matching points.

C. Matching points validation

Using the approach described in the previous subsection we get a set of t corresponding points pairs (p, q) with $p \in R_a, q \in R_b$ that can include some (or eventually many) false matches. The naive approach – use all the t couples in order to determine the matching matrix M according with [2] – can easily fail if there are many wrong matching pairs. Therefore, we adopted a different approach. For each combination of 4 different pairs extracted by the set of t pairs ($Q_j = [(p,q)_1 \dots (p,q)_4]$) we compute a matrix M_j; in this manner we obtain $\binom{t}{4}$ different matching matrices M_j. Then, we choose the matrix that provides the smaller alignment error ϵ_j computed only on the 4 pairs which defines Q_j:

$$\epsilon_j = \frac{1}{4} \sum_{(p,q) \in Q_j} (\|p - M_j q\|^2) \quad (4)$$

If the alignment error ϵ_j is equal or smaller than the user-selected threshold value $coarse\_err$ [1], we have found a sufficiently correct alignment. Otherwise, M_j should not be considered as a correct alignment matrix.

The computational cost of our approach depends on the value of t and the resolution of the range maps. Larger is t and larger is the subset of all possible roto-translation between R_a and R_b which are considered and checked. On the other hand, we should set t to small values to maintain the computational cost affordable. In the next paragraph, we'll show how to find an alignment matrix using a very low value for t: $20 \approx 40$ vertices (remember that a standard range maps contains around 300K vertices).

D. Iterative matrix computation

The best matrix found yields an error greater than the given threshold $\epsilon_{best} > coarse\_err$ if the set of t pairs contains mostly incorrect matches, and it is not possible to compute a correct alignment matrix from these t pairs. Instead of setting a bigger value for the t parameter, we adopt an iterative approach: all the previous steps (choice of a new set of t vertex pairs, evaluation of the alignment error) are iterated until a proper alignment matrix is found. To improve accuracy and speed up convergence of the method, we reuse on each iteration the best results obtained in the previous cycle.

At iteration i, we select the local set of t matching vertices by computing the respective kernels. Then, to find the best roto-translation matrix, we add to the local t pairs the 4 pairs that generated the best solution found in the previous iteration. In this manner, the space of possible solutions is augmented and this helps the algorithm to converge faster to a consistent solution.

[1] The threshold $coarse\_err$ is usually given in metric units, i.e. in millimeters in the case of our laser-scanner range maps.

The main drawback of this method is that we might be trapped in a *local minimum*, i.e. when the algorithm is not capable to find a better solution than the previous one. We detect a local minimum stall by checking if in subsequent iterations the alignment error ϵ_{best} doesn't improve. To remove the stall, we perform a perturbation of the best current solution by discarding one of the 4 couples that generated the current solution. According to the results of our empirical experiments (see Section V), this simple heuristic detected and recovered all local minima stalls. Moreover, in order to detect the occurrence of an impossible matching (e.g. when the overlap $R_a \bigcap R_b \approx \emptyset$), we set a threshold value max_{iter} for the maximum number of iterations allowed.

V. Experimental Results

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). The laser scanner used in all our acquisitions is a Minolta Vivid 910, which returns range maps of resolution 640*480 (around 300K samples). After the automatic selection of an initial coarse registration matrix with the proposed algorithm, all datasets were finely aligned (pairwise local and global registration) using our *MeshAlign* tool [3]. All performance figures have been measured on a Pentium IV 2.4GHz PC with 1GB RAM.

A. Single range map pairs alignment

Figure 3 shows a set of four matching points selected on two range maps sampling a portion of the Arrigo VII's head (the results are shown over the depth maps, rendered as they are with no roto-traslations added at visualization time). In this case each range map contains about 150k vertices (here large portions in the 2D maps contain null data). The solution was found in 30.8 seconds, setting $t = 20$ and $coarse\_err = 1.0mm$. The algorithm iterates four times to converge to a coarse alignment satisfying $coarse\_err$. The error obtained in the four iterations is shown in the graph presented in Figure 3. For each iteration we show the best alignment error ϵ_j obtained. The result of this coarse alignment is presented in Figure 4 (this figure requires to be seen in color since the two range maps are rendered with a different color shade). It's possible to notice that the alignment's quality is rather good and thus the ICP algorithm can be applied successfully to the meshes.

Another interesting example is the one shown in Figure 5, which shows the matching points selected on two different range maps (they represents the back

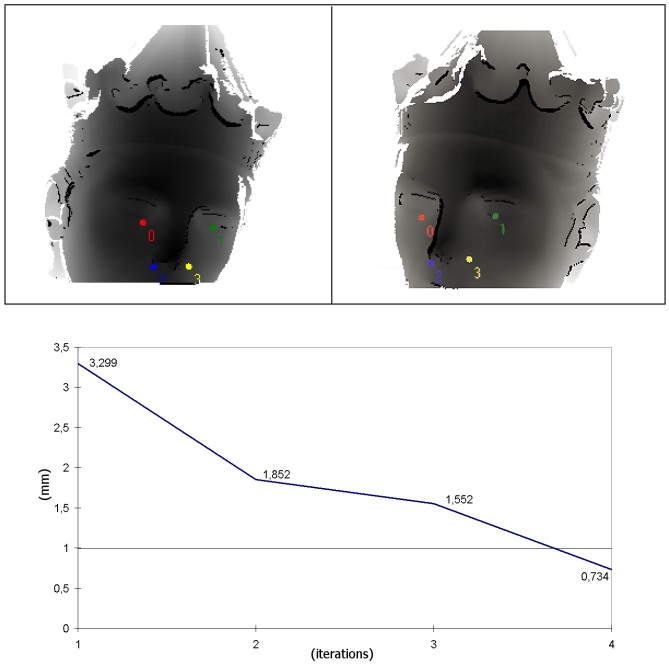

Fig. 3. The four matching point pairs selected by the algorithm on two range maps. The coarse alignment for the two range maps required four iterations (error in millimeters).

Fig. 4. The result of the coarse alignment.

section of the Arrigo's head). In this case the running time was longer, 1min:28sec, since the algorithm was temporarily trapped in local minima and thus required more iterations than the previous example (9 in total in this case, 3 were local minima as shown in of Figure 5). After the stall's detection (at the 5^{th} iteration), the current solution was perturbed; we can notice that, at the sunbsequent iteration, the solution found was worse than the previous one. Anyway, at the 7^{th} iteration the alignment error improved. The technique adopted to exit from a local minimum stall, as described in Section IV-D, influences the alignment error curve in the way previously described: after an initial worsening, usually error is improved; we observed the same behavior in all the tests we carried out, whenever a local minimum was

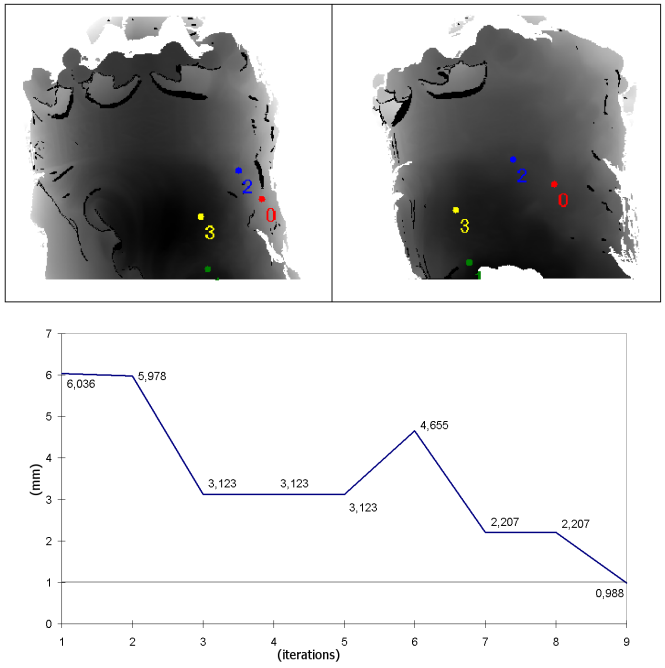

Fig. 5. The matching point pairs selected on the back of the Arrigo's head; the graph shows the alignment iterations and the two local minima successfully managed.

detected.

B. Stripes alignment

We present now some examples of the tests performed on bigger datasets. As explained in Section III, in a real scan campaign we usually acquire the range maps in a row or column-wise order; this simplifies the planning process and also the successive alignment phase. For this reason, at the end of a standard scan campaign, we have the scan set subdivided into a few stripes, where each stripe contains many consecutive and adjacent range maps. A few ad hoc scans are also usually produced at the very end of the scanning session, to cover object regions which have not been sampled by the previous stripes.

Figure 6 shows the result of the stripe-based alignment of a spiral column with artistic carvings (real size is about 85 cm tall, with a diameter of 25 cm.). Five *circular stripes* were scanned, for a total of 62 range maps (about 11M vertices). The automatic coarse alignment completed processing in 1h:13min, using $t = 20$ and $coarse\_err = 1.0mm$. The graph on the bottom of Figure 6 shows how many range maps of the column were aligned by the algorithm in a certain amount of iterations: most of the range map pairs required few iterations ($\leq$ 10), while only 2 meshes required about thirty iterations.

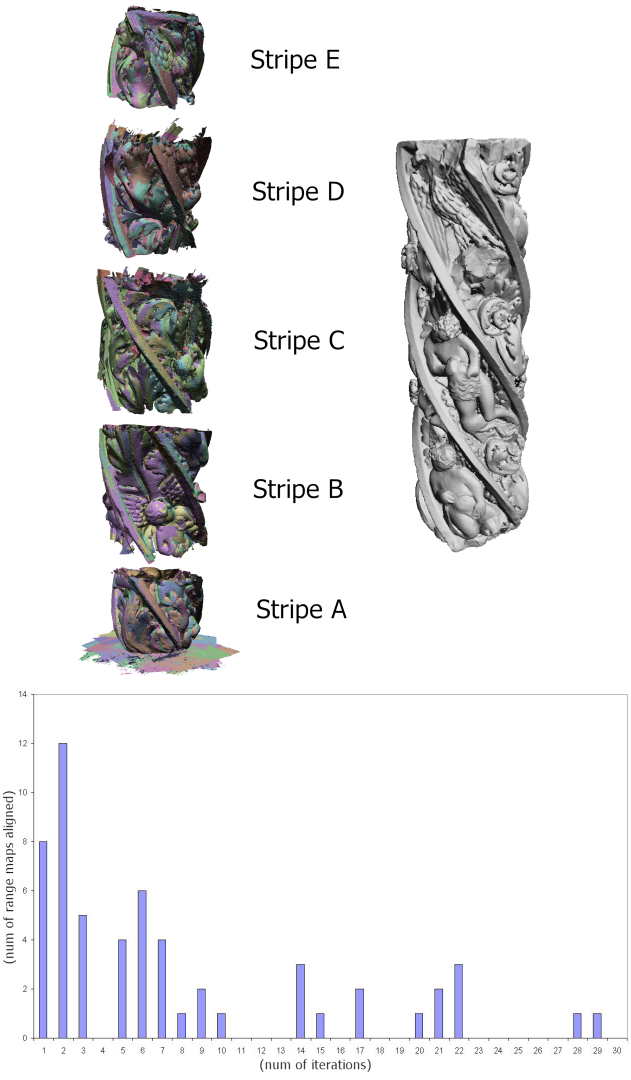

Fig. 6. The result of the coarse alignment applied to the five circular stripes *(left)* and the final model obtained *(right)*. The histogram shows how many range maps were aligned with that number of iterations; most of the alignments required less than 10 iterations.

An example concerning a bas-relief is shown in Figure 7, whose 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 algorithm performed the overall alignment in 1h:50min. As shown in the graph presented in Figure 7, almost all pair-wise alignments were done with a single iteration, so the entire process required a shorter time to complete than the previous column example. This run produced an overall alignment totally satisfactory for the subsequent application of ICP and global alignment. Again, the set of arcs has been completed automatically by the *MeshAlign* system.

Figure 8 shows an important aspect of the *local vs. global* alignment: Figure 8.a shows the result of the local

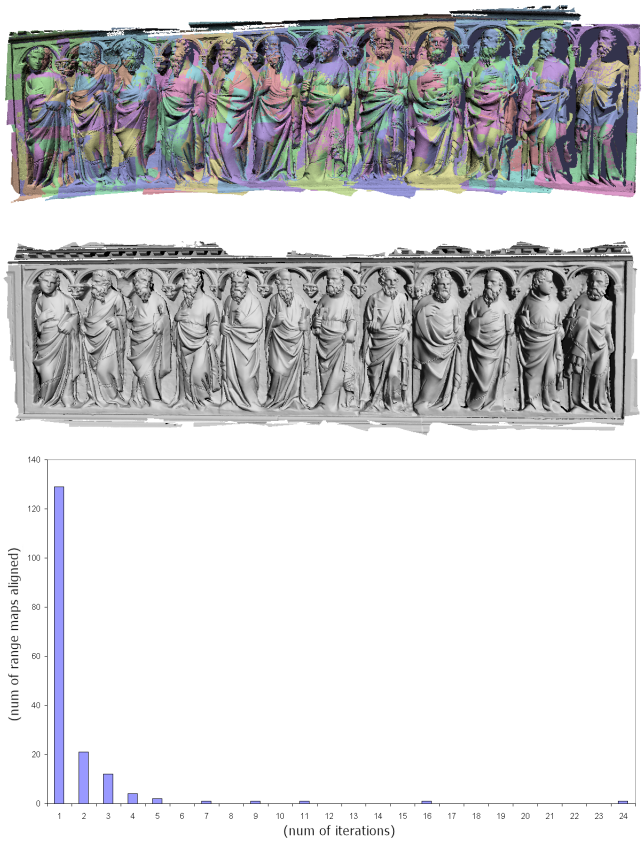

Fig. 7. The coarse alignment of the bas-relief *(top)* and the final model *(middle)*; almost all of the alignments required just 1 iteration.

coarse alignment of the first four meshes out of a circular stripe representing the head of the Minerva of Arezzo (a bronze statue, 1.55 cm tall); the final result of the local coarse alignment is shown in Figure 8.b. The substantial accumulated error between the first (green) and the last range map (violet) in the circular scan is easily visible (see for example the major miss-match in the nose region). This is due the well-known alignment error accumulation along the circular stripe, common to all local registration algorithms. The subsequent application of a global registration phase [20], [3] produces a correct fine alignment, as shown in Figure 8.c.

An overall presentation of numerical figures relative to all the scans set presented is given in Table I. The last two scan set mentioned in this table are presented in Figures 9 and 10.

A fair comparison with existing solutions is not easy since no source code is available on public domain and timings are often not presented. Moreover, most of the approaches presented in literature takes into account (and present results on) only the simple case of a pair of range maps, even of very small size ([7], [11], [26]). The only previous paper presenting results obtained on

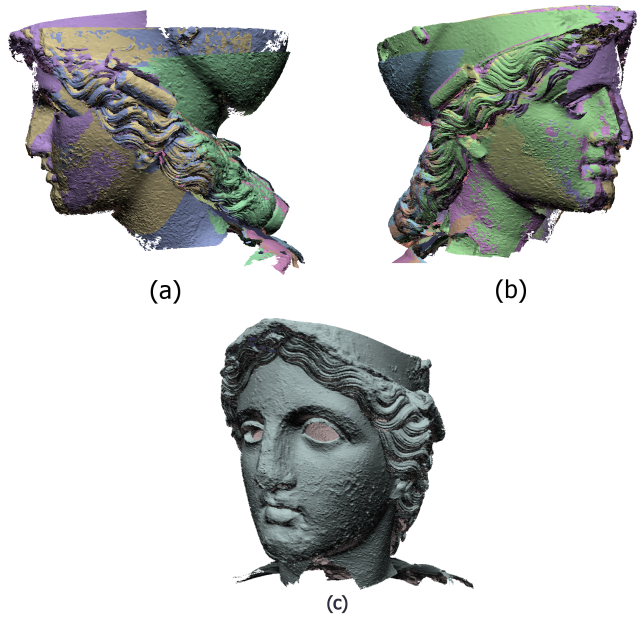

Fig. 8. Local coarse alignment of the first four range maps of the Minerva's head *(a)*; a substantial accumulated error is visible between the first (green) and the last range map (violet) in the circular stripe *(b)*; the final result after the application of global alignment *(c)*.

Fig. 9. The coarse alignment of a portion of the archway of S. Ranieri's door, Pisa Cathedral *(up)*. The final model after ICP and global alignment *(down)*.

Fig. 10. The coarse alignment of the Arrigo VII's sepolcro.

Dataset	Range maps	Total vertices	Alignment time
Minerva's head	10	1.7M	7min
Spyral column (Arrigo VII)	62	11M	1h 13min
Bass-relief (Archway)	110	30M	52min
Bass-relief (Arrigo VII)	117	45.5M	1h 50min
Sepolcro (Arrigo VII)	136	28M	41min

TABLE I

NUMERICAL FIGURES ON THE DATASET CONSIDERED IN THE EMPIRICAL EXPERIMENTS.

rather complex scan set is [15], but unfortunately authors do not offer information on the running times of their solution.

VI. CONCLUSIONS

We have presented a new automatic registration method which has demonstrated very good performances while converging to valid solutions. The method is based on some simplifying assumptions, which allow us to face the automatic range map alignment problem with an iterative solution based on the automatic detection of corresponding point pairs. The solution presented is based on a new shape characterization kernel that focuses on surface vertices; it works in the 2D space of the range maps and characterize 3D geometry by processing surface normals evaluated on a kernel of regularly sampled adjacent points. The method has demonstrated to work well on real complex scan set with good performances even using a standard PC (as shown in table I). We are planning to include this solution as a background process in our scanning front end (which drives the Minolta scanner). Due to the reduced computational complexity, automatic alignment can be run in background during acquisition, to have all range maps transformed on the fly in a common space. This will allow the user to monitor the completion status of the acquisition and to detect well in advance not sampled surface regions.

ACKNOWLEDGMENTS

This work was supported by the EU NOE "Aim@Shape" project (EU 506776) and "MACROGeo" project financed by MIUR - Action FIRB. We gratefully acknowledge the cooperation with the Tuscan Archaeologic Superintendency (Florence, Italy) – concerning the acquisition of the Minerva statue – and the Opera Primaziale Pisana and Museum of the Cathedral (Pisa, Italy) – concerning the acquisition of the Arrigo VII statues.

REFERENCES

[1] F. Bernardini, H. E. Rushmeier, I.M. Martin, J. Mittleman, and G. Taubin. Building a Digital Model of Michelangelo's Florentine Pieta'. *IEEE Computer Graphics & Applications*, 22(1):59–67, Jan-Febr. 2002.

[2] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and machine Intelligence*, 14(2):239–258, February 1992.

[3] M. Callieri, P. Cignoni, F. Ganovelli, C. Montani, P. Pingi, and R. Scopigno. Vclab's tools for 3d range data processing. In A. Chalmers D. Arnold and F. Niccolucci, editors, *VAST 2003 and EG Symp. on Graphics and Cultural Heritage*, page (in press), Bighton, UK, Nov. 5-7 2003. Eurographics.

[4] M. Callieri, A. Fasano, G. Impoco, P. Cignoni, R. Scopigno, G. Parrini, and G. Biagini. Roboscan: An automatic system for accurate and unattended 3d scanning. In *3DPVT'04: Second Int. Symp. on 3D Data Processing, Visualization and Transmission*, page (in press). IEEE Comp. Soc., Sept. 6-9 2004.

[5] Richard J. Campbell and Patrick J. Flynn. A Survey Of Free-Form Object Representation and Recognition Techniques. *Computer Vision and Image Understanding*, 81(2):166–210, 2001.

[6] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *International Journal of Image and Vision Computing*, 10(3):145–155, April 1992.

[7] C. S. Chua and R. Jarvis. Point signatures: A new representation for 3d object recognition. *Int. J. Comput. Vision*, 25:63–85, 1997.

[8] H. Delingette, M. Hebert, , and K. Ikeuchi. A spherical representation for the recognition of curved objects. In *Proc. IEEE Int. Conf. On Computer Vision*, pages 103–112, 1993.

[9] H. Delingette, M. Hebert, and K. Ikeuchi. Shape representation and image segmentation using deformable surfaces. *Image and vision computing*, 10:132–144, 1992.

[10] R. Fontana, M. Greco, M. Materazzi, E. Pampaloni, L. Pezzati, C. Rocchini, and R. Scopigno. Three-dimensional modelling of statues: the minerva of arezzo. *Journal of Cultural Heritage*, 3(4):325–331, 2002.

[11] A.E. Johnson and M. Hebert. Surface registration by matching oriented points. In *3DIM97*, pages 4 – View Registration, 1997.

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

[13] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The Digital Michelangelo Project: 3D scanning of large statues. In *SIGGRAPH 2000, Computer Graphics Proceedings*, Annual Conference Series, pages 131–144. Addison Wesley, July 24-28 2000.

[14] Marc Levoy and Szymon Rusinkiewicz. Efficient variants of the ICP algorithm. In *Third Int. Conf. on 3D Digital Imaging and Modeling (3DIM 2001)*, pages 145–152. IEEE Comp. Soc., May 28th - June 1st 2001.

[15] Gianfranco Doretto Luca Lucchese and Guido Maria Cortelazzo. A frequency domain technique for range data registration.

IEEE Trans. on Pattern Analysis and Machine Intelligence, 24(11):1468–1484, 2002.

[16] T. Funkhouser M. Kazhdan and S. Rusinkiewicz. Rotation invariant spherical harmonic representation of 3d shape descriptors. In *Eurographics Symposium on Geometry Processing*, pages 156–164, 2003.

[17] T. Funkhouser M. Kazhdan and S. Rusinkiewicz. Shape matching and anisotropy. *ACM Trans. on Graphics (SIGGRAPH 2004)*, 23(3):(in press), Aug. 2004.

[18] Polhemus. The FastSCAN Cobra scanning system. More info on:`http://www.polhemus.com/fastscan.htm`, 2003.

[19] M. Pollefeys, L. J. Van Gool, M. Vergauwen, F. Verbiest, and J. Tops. Image-based 3d acquisition of archeological heritage and applications. In D. Arnold, A. Chalmers, and D. Fellner, editors, *VAST 2001 Conference Proc.*, pages 255–261, Athens, Greece, Nov. 28-30 2001. ACM Siggraph.

[20] K. Pulli. Multiview registration for large datasets. In *Proc 2nd Int.l Conf. on 3D Digital Imaging and Modeling*, pages 160–168. IEEE, 1999.

[21] G. Roth. Registering two overlapping range images. In *3DIM'99: Second Int. Conf. on 3D Digital Imaging and Modelling*, pages 191–200, October 1999.

[22] S. Rusinkiewicz, O. Hall-Holt, and M. Levoy. Real-time 3d model acquisition. In *Comp. Graph. Proc., Annual Conf. Series (SIGGRAPH 02)*, pages ??–?? ACM Press, July 22-26 2002.

[23] F. Stein and G. Medioni. Structural indexing: Efficient 3-d object recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14:125–145, 1992.

[24] Steinbichler Optotechnik. COMET series. `http://www.steinbichler.de/`, 2004.

[25] J. Stumpfel, C. Tchou, T. Hawkins, P. Debevec, J. Cohen, A. Jones, and B. Emerson. Assembling the sculptures of the parthenon. In A. Chalmers D. Arnold and F. Niccolucci, editors, *VAST 2003 and EG Symp. on Graphics and Cultural Heritage*, pages 41–50, Bighton, UK, Nov. 5-7 2003. Eurographics.

[26] Y. Sun, J. Paik, A. Koschan, D.L. Page, and M.A. Abidi. Point fingerprint: A new 3-d object representation scheme. *IEEE Transactions on Systems, Man and Cybernetics*, 33(4):712–717, August 2003.

[27] Richard Szeliski. Rapid octree construction from image sequences. *CVGIP: Image Understanding*, 58(1):23–32, July 1993.

[28] J.W. Tangelder and R.C. Weltkamp. A survey of content based 3D shape retrieval methods. In *Int. Conf. on Shape Modeling and Applications (SMI 2004)*, page (in press), Genova, Italy, 2004. IEEE Comp. Society.

[29] S. Tosovic and R. Sablatnig. 3D modeling of archaeological vessels using shape from silhouette. In *Proc. of 3rd International Conference on 3D Digital Imaging and Modeling (3DIM)*, pages 51–58, Qubec City, Canada, May 28 - June 1 2001. IEEE Comp. Soc.

[30] D. Vranic and D. Saupe. 3D model retrieval with spherical harmonics and moments. In *Proc. of the DAGM*, pages 392–397, 2001.

[31] J. Vanden Wyngaerd and L. Van Gool. Automatic crude patch registration: Toward automatic 3d model building. *Computer Vision and Image Understanding*, 86(2):8–26, 2002.

Andrea Fasano is a research collaborator at the Istituto di Scienza e Tecnologie dell'Informazione of the National Research Council in Pisa. His research interests include 3D scanning, automatic registration, ray tracing and applications of Computer Graphics. Fasano received an advanced degree in Computer Science in 2003 from the University of Pisa.

Paolo Pingi is a PhD student at the Istituto di Scienza e Tecnologie dell'Informazione of the National Research Council in Pisa. His research interests include 3D scanning, automatic registration, sensor tracking and applications of Computer Graphics. Pingi received an advanced degree in Computer Science in 1999 from the University of Pisa.

Paolo Cignoni is research scientist at the Istituto di Scienza e Tecnologie dell'Informazione (ISTI) of the National Research Council (CNR) in Pisa, Italy. His research interests include computational geometry and its interaction with computer graphics, scientific visualization, volume rendering, simplification and multiresolution. Cignoni received in 1992 an advanced degree (Laurea) and in 1998 a PhD in Computer Science from the University of Pisa.

Claudio Montani is a research director with the Istituto di Scienza e Tecnologie dell'Informazione (ISTI) of the National Research Council (CNR) in Pisa, Italy. His research interests include data structures and algoritms for volume visualization and rendering of regular or scattered datasets. Montani received an advanced degree (Laurea) in Computer Science from the University of Pisa in 1977. He is member of IEEE and AICA.

Roberto Scopigno is senior research scientist at the Istituto di Scienza e Tecnologie dell'Informazione (ISTI) of the National Research Council (CNR) in Pisa, Italy. He is currently engaged in research projects concerned with scientific visualization, volume rendering, multiresolution data modeling and rendering, 3D scanning and applications of 3D computer graphics to Cultural Heritage. Scopigno received an advanced degree (Laurea) in Computer Science from the University of Pisa in 1984. He is member of IEEE and Eurographics. He is elected member of the Executive Committee of the Eurographics association and, since 2003, Vice Chair of the association.