On-the-fly automatic alignment and global registration of free-path collected 3D scans

Francesco Bonarrigo, Nicola Pezzotti, Alberto Signoroni
Information Engineering Department, University of Brescia, via Branze, 38, I-25123, Brescia, Italy
Email: francesco.bonarrigo@ing.unibs.it, nicola.pezzotti@gmail.com, alberto.signoroni@ing.unibs.it

***Abstract*—We present a complete geometry processing pipeline for multiple 3D scans alignment, composed by a fast, automated feature-based coarse alignment and an efficient global registration, with the aim to enable high-quality and high-throughput cultural heritage digitization. Salient features of our pipeline consist in the capability to provide low-latency, on-the-fly coarse alignment regardless to the number of scans, the capacity to handle interruptions of a continuous acquisition path, as well as an improved implementation of a robust state-of-the-art global alignment technique.**

I. Introduction

Digital representation of humanity's shared heritage, from cultural landscapes to movable artworks, is increasingly taking its place in the toolbox of cultural heritage scientists and scholars and is concurrently becoming an important asset for the virtual fruition and dissemination of culture to the general public. Modern application scenarios require high-fidelity geometric models which in turn are associated to a diffuse need to promote and perform large scale digitization campaigns characterized by easiness, effectiveness and affordability of the 3D objects/scenes acquisition phases. 3D digitization is at the basis of several virtual heritage applications [1], [2], such as digital archives, cataloguing and documentation, interactive virtual museums, archaeological, historical or other specialized studies on works of art including digital preservation, computer-assisted restoration and re-creation (digital art, multimedia historical contextualization), search and retrieval from digital heritage archives. Early works in digitization and modeling as, to name a few, the Digital Michelangelo Project [3], the Michelangelo Deposition Project [4], the Cathedral Saint-Pierre Project [5], the Great Buddha Project [6] not only demonstrated the feasibility of such, even large scale, digitization campaigns, but also sparkled great interest around 3D digital heritage research themes. Basic research on point-based geometry and surface representation, have stimulated heritage-oriented applications and are improving the feasibility and deployment of digitization campaigns. In this paper we focus on still open challenging needs related to the alignment and registration phases of point-based geometry digitization pipelines (examples are [3], [6], [7], [8]) involving high-end 3D range scanning. Such requirements are related to the necessity to grant high throughput, flexibility and fast model reconstruction in the demanding scenarios where a large amount of highly resolved and metrically accurate scans of a given object of interest is generated. This also calls for alignment procedures capable to avoid, or at least minimize, tedious and time-consuming manual intervention. Nevertheless, many existing automatized alignment procedures assume that the acquisition is performed by following a continuous path, implying that each new scan should possess a certain area of overlap with respect to its previous. This is, however, in contrast with the needs of the operator, since he may at some point realize that some area of the object has not been covered during a first pass, so that the need to directly acquire the uncovered area, without any limitation regarding the acquisition path, arises. Therefore there is the need to develop acquisition methodologies that simply remove such path constraints, without resorting to the use of any predetermined (e.g. mechanical or optical) positioning system. The objective of this work is to go some steps further toward such direction by providing a comprehensive point-geometry digitization pipeline that demonstrates to be resilient to interruptions of the acquisition path, and which proved to be accurate and cost-effective even for large, high-resolution datasets. Our achievements also encompass the awareness that, for an effective management and dissemination of cultural heritage models, the quality of the digitization is crucial not only in terms of single objects but also conceived as a general endeavor, in a perspective of enabling and promoting scientifically authenticated repositories [9].

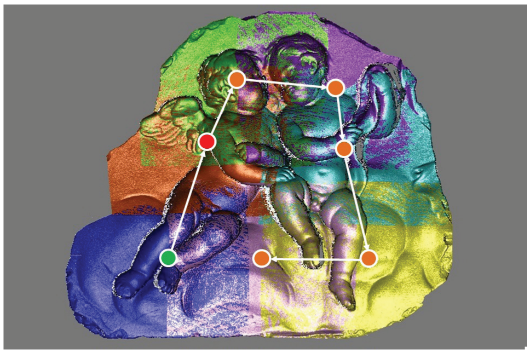

Fig. 1. A sequential, continuous acquisition path.

II. Free path registration pipeline

Normally, the assessment of an automatic alignment solution is likely to be strongly influenced by the acquisition path chosen by the operator while scanning the artwork. An example of this is depicted in Fig. 1, were in case the alignment of the scan shown by the red dot fails, the views highlighted by yellow dots would fail to align to the first one. Our goal is to remove such constraining dependence on the acquisition path, as well as to replace manual work overburden, in favor of automation and interactivity for a more productive and intuitive scanner usage.

978-1-4799-3170-5/13/$31.00 ©2013 IEEE

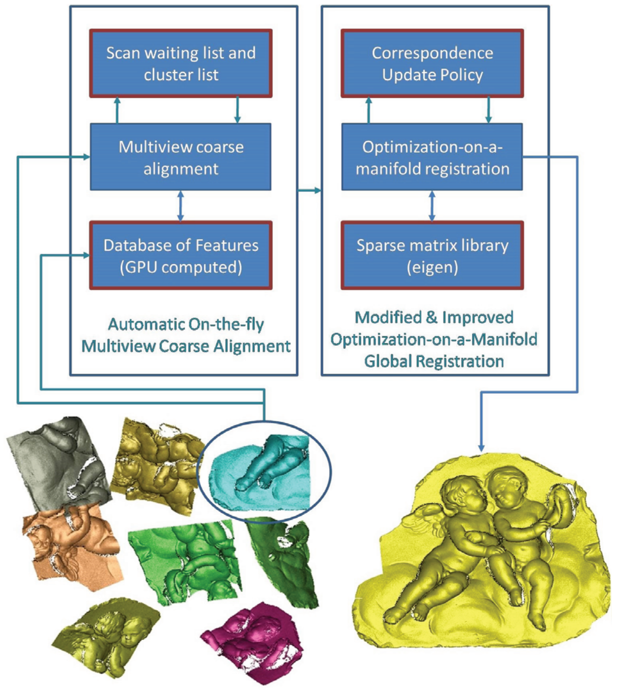

Fig. 2. Functional scheme of the proposed digitization pipeline with new contributions highlighted (red contoured boxes).

A. Proposed digitization pipeline

Basically, we propose an operative reimplementation and concatenation of the tools introduced in [10], [11] and [12] according to the scheme presented in Fig. 2, where new contributions with respect to existing solutions are highlighted in red contoured boxes. Assuming that a way to compute a rigid transform $RT(i, j)$ between two scans i and j exists (Sec.II-B), the proposed solution aims to determine, for each view i, the set of overlapping views with respect to previously acquired scans $O_i = O(i, \{i-1, ..., 1\})$), followed by an estimation of the transform $RT(i, O_i)$ that aligns the new scan to the others. Such estimation is different according to the specific phases, in that the multiview coarse alignment phase (Sec.II-C) computes an approximate rigid transform $RT^{MV}(i, O_i^{MV})$ (where the O_i^{MV} is evaluated on-the-fly during the acquisition), while the subsequent global registration phase (Sec.II-D) estimates the optimal $RT^{GR}(i, O_i^{GR})$ set of transforms on a stabilized view topology. In the following sections, the individual functional components are summarized. We will focus on the improvements and new features introduced in this work, while for a deeper understanding of the methods, the interested reader can refer to the original papers. For the coarse alignment phase, the novel technical aspects are related to the handling of a scan waiting list, as well as the possibility to simultaneously manage several aligned scan clusters, allowing the system to efficiently face interruptions of the acquisition path. Computational performance have also been improved by exploiting the GPU capabilities through CUDA technology for the feature point computation, cluster update and alignment phases. In addition, we also reimplemented a state-of-the-art global registration ([12]), obtaining a quantum leap in terms of computation efficiency by fully exploiting the problem sparseness.

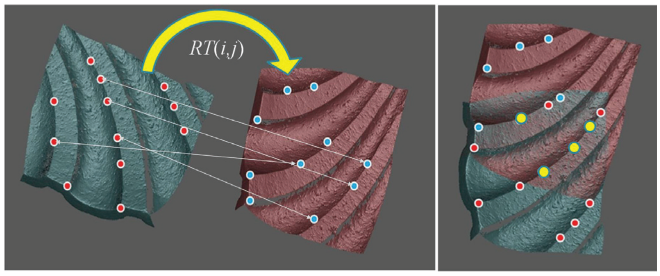

Fig. 3. Feature-based pairwise alignment principle

B. Feature-based alignment core

Both pairwise and one-to-many automatic alignment are based on the robust feature-based approach described in [10]. In its most basic form, it allows to compute $RT(i, j)$ out of correspondences detected among distinctive keypoint descriptors found to be similar between the views i and j, as depicted in Fig. 3. Technical solutions regarding multiscale keypoint detection (geometric extension of SIFT detectors), keypoint description (saliency and normal information of keypoint neighborhood, encoded in a sector-based grid), correspondence matching (based upon descriptors similarity), correspondence selection (ordering and skimming) and final roto-translation estimation are described in [10] along with justifications and comparisons with respect to other state-of-the-art feature-based techniques and RANSAC-based correspondence skimming. The multiview alignment presented in [10] remains critical when the number of views increases due to possible alignment time peaks caused by the feature database growth. This raises a scalability problem when the number of views to align grows up to a critical point. Another important factor impacting on the computational performance is the time required to compute the keypoints, which is reduced here by means of a GPU implementation of the keypoint detector.

C. On-the-fly multiview coarse to fine alignment

In [11] effective multiview automatic alignment (Fig. 4.a) is reached by an effective two-step correspondence matching phase acting on a feature database F_{db}, where features are organized in clusters of limited size through an iterative and adaptive variant of the k-means algorithm. At first an inter-cluster search is performed by matching each feature of the current view F_n with respect to the cluster centroids, this is followed by an intra-cluster search to detect the most similar feature (the correspondence we are looking for) within that cluster (Fig. 5). More precisely, while during the final intra-cluster search a given feature is matched with respect to the ones populating the cluster by means of a circular correlation to maximize feature discrimination, during inter-cluster search such correlation can be replaced by estimating a principal direction for each feature descriptor. This way, direct matching comparison (further quickened by feature subsampling) can be performed: this loss of matching resolution is fully justified at inter-cluster stage, since the centroids are enough different from each other. This way near-constant time correspondence search irrespective to the number of scans to align can be guaranteed. In addition, the clustering update phase is kept lightweight by terminating it once sufficiently close to convergence. Eventually, one-to-many ICP fine registration is

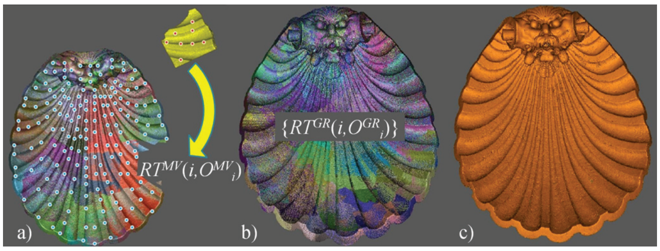

Fig. 4. a) Addition of a new image in a multiview coarse alignment, b) Global registration phase, c) Final model

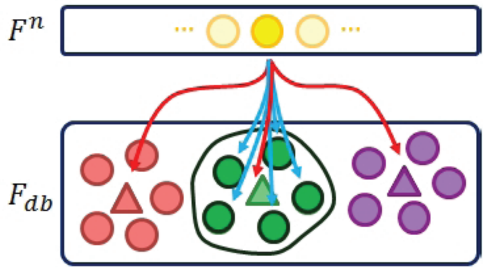

Fig. 5. Two-step correspondence search: red arrows identify inter-cluster matching, sky ones indicate intra-cluster matching.

performed at the end of each successful view alignment.
What [11] lacked to elaborate was the possibility to implement a smart on-the-fly alignment scheme (the one we realized here) capable to handle the case in which incoming scans do not find immediate successful alignment. For such scans, our approach foresees the implementation of a waiting list. When the succession of incoming scans keeps failing to align (which is a symptom that the acquisition path may have been broken), the latest scan in this list is set as seed to form a new cluster of scans, initializing a different feature database. Features extracted from each new scan will now be matched against all existing scan clusters (we determined a maximum number of 5 concurrent instances, which we found adequate for most of the cases). At the end of the acquisition process, each scan still populating the waiting list is re-matched. At last, scan clusters are aligned among each other, with the rationale that the composition of more scans is more likely to lead to successful alignment.

D. Fast, accurate and robust global registration

Even in its multiview and improved form, the proposed feature based coarse alignment basically performs a locally incremental (intra-cluster) alignment of the scans, so that it cannot ensure the reaching of the minimum misalignment error on the entirety of the scans. As anticipated, optimal global registration (see Fig. 4.b) can be estimated upon the views topology determined by the coarse alignment phase (after possible manual addition/correction of any residual unaligned scan), in order to allow the highest accuracy point based model to be built (Fig. 4.c). In [12], an improved version of the iterative optimization-on-a-manifold global registration method has been introduced, which is able to bring down both methodological and computational limitations associated to the original approach [13]. In particular, it is demonstrated that, while performing a correspondence update process at each iteration may seem to downgrade the original quadratic convergence rate to linear, this is however to the benefit of reaching a smaller residual error figure. Reported computational improvements are an effective correspondence update process and the modification of the minimization procedure from a pure Gauss method to a less demanding damped Gauss-Newton. We improve upon such method by fully exploiting the sparsity of the problem to solve, switching from the *newmat* matrix library [14] (which cannot handle sparse matrices computation) to the *eigen* library [15]: this allowed us to obtain a quantum leap in terms of computational performance.

III. Experimental results

For an experimental assessment of the versatility, robustness, accuracy and efficiency of the proposed digitization pipeline we considered a set of artifacts, provided by the authors of [10], [11] and [12] and shown in Fig. 6, which are representative of different subjects (movable artworks, statues, building decorations) and styles, different number of views and critical issues (repeating or convoluted geometric patterns). Object dimensions span from 20 cm of the Rose to 1.5 m for the Neptune dataset. The scans were collected with a state-of-art structured-light 3D scanner, which theoretical resolution is around 1,3 million of points per scan. Each scan has approximately 3/4 valid points. All developed software is written in C++, and runs on a off-the-shelf PC (Intel I5 3450 quad core 3.1Ghz, 8 GB RAM, with NVIDIA GeForce GTX 660 GPU). Table I summarize all quantitative statistics generated by our tests. In particular, to stress the flexibility and robustness of the proposed scheme, we setup two operating scenarios: a) a 'Free-Path', FP configuration, i.e. freely chosen path during the on-the-fly digitization, which foresees a limited number of path interruptions (analogously to a typical acquisition scenario), and b) a 'Path-Free', PF scan, i.e. the extreme case where a random path (obtained by scrambling the scan order) is provided. MV statistics are referred to the multi-view coarse alignment, while GR ones are related to the global registration stage. Measured distances are computed as the average of the scan distances in overlapping areas.
At first we observe that the extreme case (PF) of random path scan arrival does not determine any kind of problem to our pipeline, actually a slightly better result is reached for the Decoration dataset, likely due to the fact that the 'flowers' area, containing lots of repetitive features, gets separated into various feature datasets. Similarly, MV time increase for the PF configuration is justified by the increased use of the waiting list, even if no appreciable variations in the final scan clusters number can be identified. The MV distance results are remarkable for a 'coarse' alignment method, indicating that the feature-based multiview alignment, our waiting-list and scan cluster solutions produce reliable and accurate results. The GR time results testify that 1) a dramatic improvement with respect to the original (IOM [12]) method has been obtained trough sparse matrices exploitation, which is proportionally more influential for bigger datasets (reaching two order of magnitude for the Neptune); 2) the reached minimum is virtually the same both in FP and in PF scenarios, reaching an average figure of the same order of magnitude of the scanner metric accuracy, demonstrating the optimality and robustness of the approach.

TABLE I. ON-THE-FLY MULTIVIEW ALIGNMENT AND GLOBAL REGISTRATION PIPELINE PERFORMANCE

Dataset	**#RIs**	**MV rate [%]**				**MV time [sec]**		**MV avg dst[μm]**		**GR avg dst[μm]**		**GR time [sec]**		
		FP	clust.	PF	clust.	FP tot (view)	PF tot (view)	Free-Path	Path-Free	Free-Path	Path-Free	IOM [12]	FP	PF
Rose	24	100	1	100	1	13 (0.54)	13 (0.54)	14	14	13	13	49	12	12
Hurricane	32	100	1	100	1	47 (1.47)	53 (1.66)	60	65	35	34	29	22	25
Decoration	47	93.6	4	100	1	85 (1.81)	105 (2.23)	27	26	23	23	100	24	22
Venus	61	100	1	100	1	92 (1.51)	163 (2.67)	21	28	18	19	113	118	102
Shell	98	99.0	2	98.0	3	241 (2.46)	326 (3.32)	91	86	48	49	2824	167	153
Neptune	169	98.8	3	98.2	4	535 (3.16)	693 (4.10)	67	74	40	39	13830	273	295

Fig. 6. Considered artworks: a) Venus, b) Rose, c) Hurricane, d) Decoration, e) Neptune and f) Shell.

IV. CONCLUSION

Aware of the requirements of modern digitization of cultural heritage, we presented an entirely automated multiview alignment pipeline which allows free-path acquisitions of the artifacts with immediate user feedback. The system comprises a multiview feature-based coarse alignment technique (based on incremental and adaptive feature clustering) followed by an improved global registration technique (based on an optimization-on-a-manifold approach). These solutions are conceived to work together in a scheme which enables flexible on-the-fly path decision during the acquisition, as well as superior computational performance. A series of representative and diversified artworks has been acquired with a professional structured light scanner, whereas the multiview coarse alignment was capable to guarantee near real-time feedback independently of the number of views, while the global registration provided high accuracy, robustness to possible misalignments and at the same time competitive alignment time (compatible with practical modeling applications) even for large amounts of high-resolution scans. In synergy with high-end 3D scanners and properly designed interactive acquisition software interfaces, the proposed pipeline can be used for high-quality, highly usable and cost-effective (high throughput) digitization campaigns, as required in modern cultural heritage applications.

ACKNOWLEDGMENT

This work was partially supported by the collaborative research project *ABC3D* funded by MIUR (Italian Ministry of University and Research) and Regione Lombardia (Italy).

REFERENCES

[1] F. Stanco, S. Battiato, and G. Gallo, Eds., *Digital Imaging for Cultural Heritage Preservation: Analysis, Restoration, and Reconstruction of Ancient Artworks*. CRC PressI Llc, 2011.

[2] R. Scopigno, M. Callieri, P. Cignoni, M. Corsini, M. Dellepiane, F. Ponchio, and G. Ranzuglia, "3D models for cultural heritage: Beyond plain visualization," *IEEE Computer*, vol. 44, no. 7, pp. 48–55, 2011.

[3] 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*, 2000, pp. 131–144.

[4] F. Bernardini, H. Rushmeier, I. Martin, J. Mittleman, and G. Taubin, "Building a digital model of michelangelo's florentine pieta," *Computer Graphics and Applications, IEEE*, vol. 22, no. 1, pp. 59–67, 2002.

[5] P. Allen, A. Troccoli, B. Smith, S. Murray, I. Stamos, and M. Leordeanu, "New methods for digital modeling of historic sites," *Computer Graphics and Applications, IEEE*, vol. 23, no. 6, pp. 32–41, 2003.

[6] K. Ikeuchi, A. Nakazawa, K. Hasegawa, and T. Ohishi, "The great buddha project: modeling cultural heritage for vr systems through observation," in *IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003*, 2003, pp. 7–16.

[7] F. Bernardini and H. Rushmeier, "The 3D model acquisition pipeline," *Computer Graphics Forum*, vol. 21, no. 2, pp. 149 – 172, 2002.

[8] A. Vrubel, O. R. P. Bellon, and L. Silva, "A 3D reconstruction pipeline for digital preservation," in *CVPR 2009*, 2009, pp. 2687–2694.

[9] D. Koller, B. Frischer, and G. Humphreys, "Research challenges for digital archives of 3D cultural heritage models," *J. Comput. Cult. Herit.*, vol. 2, no. 3, pp. 7:1–7:17, Jan. 2010.

[10] F. Bonarrigo, A. Signoroni, and R. Leonardi, "Multi-view alignment with database of features for an improved usage of high-end 3D scanners," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 148, 2012.

[11] N. Pezzotti, F. Bonarrigo, and A. Signoroni, "Boosting the computational performance of feature-based multiple 3D scan alignment by iat-k-means clustering," in *3DIMPVT 2012*, oct. 2012, pp. 89 –96.

[12] F. Bonarrigo and A. Signoroni, "An enhanced 'optimization-on-a-manifold' framework for global registration of 3D range data," in *3DIMPVT 2011*, 2011, pp. 350 – 357.

[13] S. Krishnan, P. Y. Lee, J. B. Moore, and S. Venkatasubramanian, "Optimisation-on-a-manifold for global registration of multiple 3D point sets," *International Journal of Intelligent Systems Technologies and Applications*, vol. 3, pp. 319 – 340, 2007.

[14] http://www.robertnz.net/nm_intro.htm.

[15] http://eigen.tuxfamily.org.