Global refinement of image-to-geometry registration for color projection

Matteo Dellepiane and Roberto Scopigno
Visual Computing Laboratory ISTI - CNR
Via G. Moruzzi 1, 56124 Pisa (PI), Italy
Email: *surname*@isti.cnr.it

***Abstract*—The management, processing and visualization of color information is a critical subject in the context of the acquisition and visualization of real objects. Especially in the context of Cultural Heritage, artifacts are so complex or hard-to-handle that the appearance information has to be extracted from a set of images.**
The images usually have to be registered to the 3D model of the objects, in order to transfer the needed information. Hence, the problem of image-to-geometry registration has been thoroughly studied by the Computer Graphics and Computer Vision community. Several methods have been proposed, but a fully automatic and generic solution is still missing. Moreover, small misalignments often lead to visible artifacts in the final colored 3D models.
In this paper, we propose a method to refine the alignment of a group of images which has been already registered to a 3D model. Taking advantage of the overlapping among the images, and applying a statistical global method based on Mutual Information, the registration error is distributed among all the elements of the dataset. Hence, the quality of color projection is improved, especially when dealing with small details.
The method was tested on a number of heterogeneous Cultural Heritage objects, bringing to a visible improvement in the rendering quality. The method is fully automatic, and it does not need powerful hardware or long processing time. Hence, it represents a valid solution for a wide application on CH artifacts.

I. Introduction

Accurate registration of images on geometry is an important kernel operation among the technologies for Cultural Heritage. This is related to a number of different applications, and is often needed due to the necessity to integrate data coming from different acquisition devices.
The main application is related to the projection and visualization of color information on a 3D model, but registered images are important also in the fields of material properties estimation, referenced images navigation, annotation for monitoring or restoration.
One of the main problems is how to cope with the small misalignments remaining after the registration process. These errors can be due both to the methods used, and to the quality of the initial datasets, and they can result in annoying artifacts.
In this paper we proposed a refinement technique which aims at removing the above cited errors, by applying a global approach where the camera parameters are slightly modified in order to obtain a perfect projection. The method starts from a set of images aligned to a 3D model. It uses Mutual Information to refine each image by taking into account the corresponding projection of all the other images on the model. The camera parameters are refined until convergence, so that the alignment error is distributed among the images, and globally minimized. One of the advantages of the method (which is totally automatic) is that no assumption about the quality of the initial dataset is made: the method will try to obtain the best projection regardless of the quality of both 2D and 3D data. The method was tested on a number of artifacts, showing that the quality of the color projection is greatly improved, especially when dealing with fine decorative details.

II. Related Work

The problem of image-to-geometry registration has been thoroughly studied, and a number of different solutions have been proposed. In the following we tried to divide the methods in groups. We start from the assumption that the images have been taken in a different moment w.r.t. the geometry acquisition, hence we do not take into account methods based on co-located cameras [1], [2], [3].

Semi-automatic methods. A very robust approach is based on setting several 2D-3D correspondences: the correspondences are then used to estimate the camera parameters using a minimization algorithm. Although new procedures to speed up the process were proposed [4], the approach can be very time-consuming, especially when tens or hundreds of images need to be aligned. Automatic planning of the images required could minimize image acquisition and remove the need for registration [5], but this approach can only be used in controlled environments.

Features and Silhouette-based methods. The geometric features of the object can be used to find the registration of images. Features can be points, lines, rectangles [6], edge intensities [7], or the silhouette of the object [8], [9], [10]. However, these methods rely on the presence of these features on the object, hence they could be best applied on peculiar types of artifacts (e.g. architectural scenes, if using clusters of orthogonal lines).

Color-based methods. Another feature that can be used (if present) is the reflectance value (laser intensity) or color information that some 3D scanners acquire. This helps the feature extraction and the establishment of correspondences [11]. Yang et al. [12] made their co-located camera approach robust using this method. Wu et al. [13] exploited color information to align two 3D scenes even from significant viewpoint changes.

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

Statistical methods. Other approaches try to catch the non-linear correlations between the image and the geometric properties of the target surface. A measure which is extensively used in medical imaging (see [14] for a survey) is called Mutual Information (MI). It was pioneered by Viola and Wells [15] and by Maes et al. [16]. Viola and Wells [15] suggested to compare the gradient variations of the image and a rendering of the 3D models showing the surface normals. Corsini et al. [17] extended this algorithm by including other geometric properties, such as ambient occlusion and reflection directions, in the alignment algorithm. Cleju et al. [18] also extended Viola and Wells's work to align more than one image simultaneously. We propose a similar approach to refine the global registration, based on a different optimization framework.

Multi-view methods. Almost all the above mentioned methods are based on the alignment of a single image on the geometry. A more recent group of works rely on the fact that if a group of images has to be aligned on a model, it is possible to take advantage also on the relations between images. Several works exploit Structure From Motion (SFM) during 2D/3D registration process, like those of Zhao et al. [19], Stamos et al. [20] (which is an extension of the work of Liu et al. [21]), Zheng et al. [22], Pintus et al. [23], and Corsini [24].

Refinement of existing registrations. Our method aims at refining an initial registration of a group of images. Hence the results of any of the above methods could be a starting point for the refinement. Other methods aimed at improving the color projection by modifying the initial image set, using for example Optical Flow methods [25], [26]. These methods can obtain extremely accurate results, but they are limited by the resolution of the images, or by the need of advanced hardware and very long processing time. Our method aims at improving the camera parameters without modifying the rest of the dataset. We make use of Mutual Information, but on a more global and interconnected way w.r.t. the usual statistical methods.

III. Global refinement using Mutual Information

The input of our method is composed by:

- A 3D model
- A set of images
- A set of camera parameters associated to each image of the dataset. In our case, each camera is defined by seven parameters: three for camera position, three for camera orientation, and one for the focal length (see next Section for details).

The assumption is that an initial registration of the images on the 3D model is already provided. The registration can be obtained using any of the above mentioned methods, or it could come from external systems (like multi-view stereo reconstruction tools).
The goal of the method is to modify the camera parameters of all the images so that the global Mutual Information will be maximized. This means that each of the images will have the maximum Mutual Information Value w.r.t. all the other images projected on the 3D model. In order to do this, we will treat the system of registered images as a graph, and we will try to distribute the alignment error in the graph by improving all its nodes.
This approach has several points in common with the one by Pulli [27], which aimed at the global refinement of groups of range maps. It obtained this by treating the system of range maps as a graph, and improving the alignment of a range map at a time, trying to distribute the alignment error in a balanced way.

A. *Using Mutual Information to align a single image*

Mutual Information (MI) measures the information shared by two random variables A and B. The Mutual Information $\mathcal{MI}$ between two images I_A and I_B can be defined as:

$$\mathcal{MI}(I_A, I_B) = \sum_{(a,b)} p(a,b) \log \left(\frac{p(a,b)}{p(a)p(b)} \right) \quad (1)$$

where $p(a)$ $(p(b))$ is the probability that the value of the pixel I_A (I_B) is a (b) and $p(a,b)$ is the joint probability of the event (a,b). The joint probability distribution can be easily estimated by evaluating the joint histogram of the two images and then dividing the number of occurrences of each entry by the total number of pixels. A joint histogram is a bi-dimensional histogram made up of $N \times N$ bins; the occurrence (a,b) is associated with the bin (i,j) where $i = \lfloor a/m \rfloor$ and $j = \lfloor b/m \rfloor$ and m is the width of the bin. This value can be seen as an expression of the *nonlinear correlation* between the variables A and B.

The image-to-geometry registration problem in this case can be defined as the estimation of the camera parameters that produce a rendering I_B of the 3D model that maximizes MI with respect to the image to align (I_A). The generation of the rendering is the main issue to be solved, since generally there's a lack of knowledge of not only the color and materials of the object but also of the lighting conditions. At the same time, since the Mutual Information expresses a correlation between the images, the photorealism is not a requisite: it is important that the rendering contains a high amount of information "in common" with (I_A). Corsini et al. [17] proposed several rendering types following this aim. In particular, they showed that ambient occlusion correlates well since the occluded parts of the geometry often correspond with the dark parts in the real image due to the poor illumination arriving at these points, and normal maps are strongly correlated with more directional illumination.
In this context the registration can be formalized as an optimization problem in a 7D space:

$$\begin{aligned} \mathcal{C}^* &= \arg\max_{\mathcal{C}\in\mathbb{R}^7} \mathcal{MI}(I_A, I_B(\mathcal{C})) \quad (2) \\ \mathcal{C} &= (t_x, t_y, t_z, \theta_x, \theta_y, \theta_z, f) \end{aligned}$$

where f is the focal length, (t_x, t_y, t_z) and $(\theta_x, \theta_y, \theta_z)$ define the position and orientation of the camera, I_A is the image to align and I_B is the rendering of the 3D model. Obviously, I_B depends on the camera parameters $(\mathcal{C})$. The Equation (2) can be solved by a non-linear optimization algorithm such as NEWUOA [28].

B. *Extension to groups of images*

The proposed method aims at extending the single image registration problem to a group of images. Instead of using a pure rendering of the 3D model, the goal is to take advantage of the overlaps among the projections of the images on the surface of the model. In this way, the best alignment will be reached when all the images will project the same color details on the same part of the surface.

Fig. 1. An example of the images used to calculate the value of an arc. Top, the original image. Bottom: a rendering of another image projected on the top image plane. The parts which are not covered by the image are shown using a "combined" (normal map + ambient occlusion) rendering.

1) Representing the images as a graph: In order to handle the connections among all the elements of the registration project, it's necessary to encode them in a structure. Hence, the registration project is represented as a graph. The nodes of the graph are represented by each image of the dataset. The nodes are connected through arcs, and each arc is associated to a weight. The value of the weight between an image I_1 and another image I_2, indicated with $w(I_1, I_2)$, corresponds to

$$w(I_1, I_2) = \mathcal{MI}(I_1, \text{proj}(I_2, I_1))OV(I_2, I_1) \qquad (3)$$

where the first term is the MI calculated between the image I_1 and the projection of the image I_2 on the image plane of I_1. The projection is generated by projecting I_2 on the 3D model, and then generating a rendering from the point of view of I_1. The parts of the 3D model which are not covered by I_2 are represented using the *combined rendering* (ambient occlusion + normals map) proposed by Corsini et al. [17]. Figure 1 shows an example of a couple of images used for the calculation of the arc weight.
The value of the arc is also weighted by OV, which represents the amount of overlap between the images. This is the ratio between the pixels on I_1 image plane which is covered by I_2, and the total number of pixels covered by the 3D model. This term aims at giving a more important role to the images which share a bigger common projection surface. At the same time, the arc between two nodes is created only if the value of OV is bigger than 0.2: this simplifies the graph structure and prevents artifacts coming from images which share very small common projection surfaces.
According to this definition, the graph related to each dataset analyzes each image, and creates an arc for each couple of images where there is enough overlap. The result of the building phase is a weighted directed graph, since $(w(I_1, I_2)$ is usually different from $w(I_2, I_1))$. It is interesting to note that the 3D model is not directly represented in the graph, but it plays the role of a "medium" due to the projections involved.

2) Graph-based registration refinement: The refinement of the graph is obtained following a similar procedure to the one used by Pulli [27] for the alignment of range maps. In that method, the refinement is reached by considering one node at a time, and using Haussdorf distance as the value to minimize. In our case, we used Mutual Information as a value to be maximized.
The refinement loop follows these steps:

- **Node selection**: among the nodes which were not already refined, the one with more connections with already refined nodes is chosen. When more than one node has the same number of connections with refined nodes, the one with the biggest largest number of entering arcs is chosen.
- **Node refinement**: the refinement is obtained by maximizing the MI (Equation 2) between the image associated to the node and a rendering of the 3D model where all the images associated with connected nodes are projected on the geometry. Figure 2 shows an example of a rendering used for the refinement. Since several images can project onto the same portion of geometry, the color assigned to the pixels is a combination of the contributions of all the images. The contribution is weighted by the value of the arc connecting the node image and each image which projects on the model. This approach aims at having the other images of the dataset as a guide for the chosen node to converge to a common alignment. If portions of the geometry are not covered by any other image in the set, *combined rendering* is used.
- **Node update**: when the maximization procedure ends, the node is labeled as *refined*, and the graph is updated (all the weights of the arcs involving the node are recalculated). The procedure goes back to step 1, until all the nodes are refined.

When all the nodes have been refined, it is possible to start the refinement again to further improve the global alignment. The procedure should be able to converge until all the camera parameters associated to the images are not modified anymore. See Figure 3 for two examples of the improvements of the registration of images.
Comparing two cameras is not trivial: one of the most reliable methods is to compare the projection on the image planes

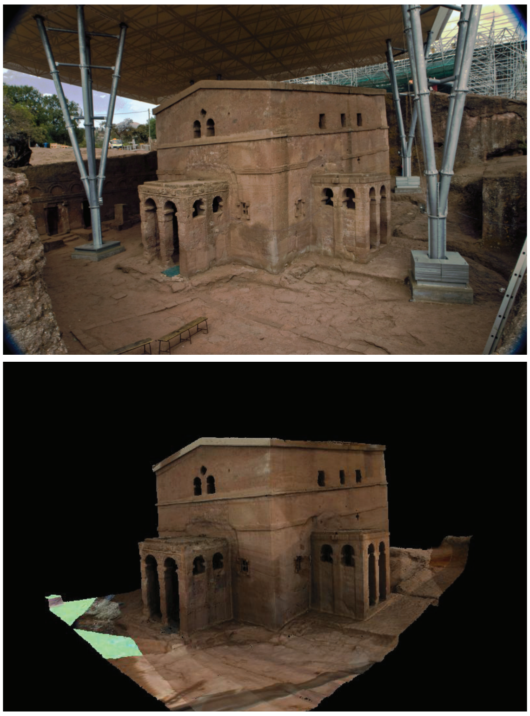

Fig. 2. An example of the renderings used for the refinement of camera parameters. Top: the image associated to the node to refine. Bottom: a rendering obtained by projecting all the other images on the 3D model. The small portions not covered by images are rendered using the ”combined” rendering.

of several samples. For this reason, in order to measure the variation in the camera system after a refinement step, a group of N 3D sample points $X = x_1, x_2, ..., x_N$ is extracted from the 3D model. The average variation of cameras is calculated as:

$$Var = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} |pro(x_i, C_j^{After}) - pro(x_i, C_j^{Before})|}{MN} \quad (4)$$

where $pro(x_i, C_j^{After})$ is the projection of the point x_i on the image plane of C_j^{After}, the camera associated to image I_j. The average of the variation in pixel of the projection of the sample points on the image planes before and after the refinement gives a reasonable estimation of the amount of perturbation applied.
Hence, at the end of the refinement step, the *Var* value is calculated. If it is below a pre-defined threshold, or the maximum number of iteration has been reached, the refinement stops. In the examples shown in the Results Section, the number of samples used was 5000, the *Var* threshold was set to 1.2 pixels, and the maximum number of iterations was set at 5.

3) Selective refinement of nodes: One of the limitations of the statistical approach is that the Mutual Information is a pure number. Hence, it is not possible to compare its value between couples of images. Hence, one of the limitations of the proposed system is that an image which exhibits a misalignment could ”guide” all the others to a sharp, but uncorrect color projection (see Figure 4).

Fig. 3. Two examples of the refinement of single images after the application of our method, showing an image with the 3D model in transparency. In the first one (first and second row) small details are better aligned. In the second one (third and fourth row) the silhouette is better matched.

In order to cope with this potential problem, the proposed system gives the possibility to the user to indicate some *anchor nodes*, which are associated to images which already have an accurate registration. In this case, the nodes will always be considered as *refined*, and their role will be to ”guide” the other images to a more accurate registration.
This modality gives the possibility to handle the datasets where only a few images are misaligned: the time needed for the refinement process will be much smaller, and the user will have control on the registration procedure.

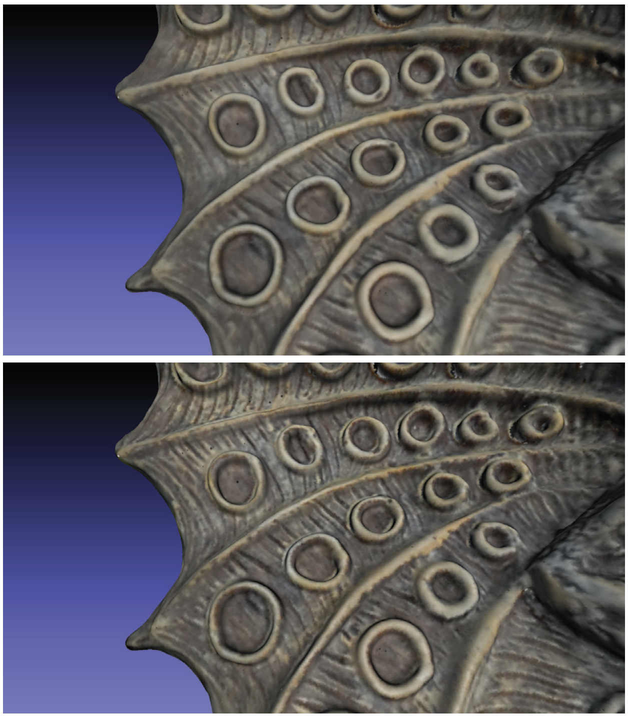

Fig. 4. An example of a possible limitation of the approach. Top: a portion of the model before refinement. Bottom: the same portion after the refinement. Although the color is sharper, the alignment w.r.t. the geometric features (e.g. the circles) is less accurate. This is because a misaligned image "guided" the others to a misaligned common position.

IV. Results

The proposed system was tested on a number of real cases covering a wide range of possible objects, from small artifacts to architectures, and a variety of datasets, from a few to tens of images, with varying quality of both 2D and 3D data. All the objects are examples of Cultural Heritage, and they were acquired using 3D Scanners. Photographic datasets were mostly acquired with the purpose of color projection.
The initial registration of the images was obtained using a single image registration method, the one based on MI, proposed by Corsini [17] and implemented in MeshLab [29]. In the case of *Formella* and *Maryam Church* datasets, the registration was obtained using an evolution of Corsini et al, *Mutual Correspondences* [30], which gives the possibility to the user to guide the registration process with a simple interaction. All the processing was performed on an Intel Core i7 CPU, with 24 GB RAM and an NVidia GeForce GTX 560 Ti.
The datasets are described in Table 5. The Table describes the physical size of the object, the complexity of the 3D model, the amount and resolution of the images, the processing time needed for the convergence of global refinement. From a general point of view, the method proved to be applicable regardless of the physical size of the object and of the complexity of the dataset. The time needed for the refinement is partially dependent on the size of the 3D model, and on the number and resolution of images, but the initial misalignment and the amount of overlap among the images play a critical role. Anyway, the processing time is reasonable even in the case of very big datasets (hundreds of images), also considering that the refinement step is an *una tantum* operation, needed just before the color projection.

Figure 6 shows the results of the registration refinement. The color was transferred from the images to the 3D Model using the Masked Photo Blending approach [31], which is a method which blends the contribution of all the images during color projection. This method proved to be very robust, being able to deal also with very complex datasets. Nevertheless, one of the limitations is that small misalignments could introduce *aliasing* (or *ghosting*) effects on the colored model. This is due to the fact that color details are projected in slightly different positions on the surface of the 3D model.
Figure 6 shows a rendering of the colored 3D models obtained with a color projection applied before and after the global refinement. We can observe that, in general, the quality of the color information is clearly improved. In the case of small and medium size objects (*Gargoyle* and *Formella*) the method is able to recover the fine details of the decoration of the objects.
In the case of bigger objects, the method recovers also quite big starting misalignments, regardless of the number of images taken into account. In the case of the *Maryam Church*, given the low number of images and the simple shape of the object, the color was encoded in a Texture (all the other objects are represented using color-per-vertex). Also in this case, there was an improvement of the fine color details. In the case of the *Cathedral*, the method proved to be able to handle a massive amount of images, that usually produce blurry color, due to the accumulation of misalignments. Also in this case, most of the color detail was recovered, although part of the finer elements was lost. This was due both to the quality of the images (average resolution) and to the accuracy of the 3D model, which was not high. Additional snapshots of the results are shown in Figure 7

The proposed method has some limitations. The main one is shared with all the statistic-based registration approaches: if the quality of the dataset is low, or the elements do not share enough information, the method will not be able to converge, and it could lead to the degeneration of the camera parameters estimation. This limitation is generally shared also with the other registration approaches, except some of the ones needing a strong user intervention.
As already mentioned in Section III-B3, the other limit comes from the fact that it is difficult to compare the quality of the registration of the single images. This can bring to a convergence of the refinement obtained by "following" an image which had a lower quality registration. This can be partially solved by applying a selective refinement.
Nevertheless, the proposed method is simple and completely automatic, and it does not need complex hardware and long processing time. It reduces the time that the user needs to spend in order to obtain very accurate image registration. Finally, it helps overcoming the limitations of most of the state-of-the-art color projection tools.

$Object$	Size (cm)	3D Model (MTri)	N. Images (Resolution)	Processing Time (sec.)
$Gargoyle$	12	3	11 (3872x2592)	191
$Formella$	65	5	10 (3872x2592)	127
$Neptune$	580	10	44 (1728x1152)	740
$MaryamChurch$	930	3 (with texture)	8 (3872x2592)	112
$Abside$	3500	9	310 (1936x1296)	3250

Fig. 5. Table of data for the five test cases.

V. Conclusion

In this paper, we presented a method for the refinement of image-to-geometry registration. The goal is to improve the quality of an already registered set of images, in order to solve eventual misalignments and improve color projection.
In order to achieve this, the set of images is treated as a graph, and the estimation of the camera parameters is calculated taking into account the projection of all the other images on each image plane. In order to refine the registration, a statistical method based on Mutual Information was implemented. The graph representing the images is refined node-by-node until convergence. The 3D model acts as a simple medium for color projection, because the final goal is to have a registration where all the images project the same color details on the same part of the geometry, regardless of its quality.
The method proved to be robust and reasonably fast. It was tested on a number of Cultural Heritage objects, covering different physical sizes and dataset complexities. All the tests showed an improvement in the color quality. This method proves to be extremely useful especially in the Cultural Heritage field, where most of the times the only way to obtain a basic color information of an object is to transfer it from a set of uncalibrated images.
The future improvements of the method include: the study of mechanisms to prevent the worsening of color quality, in the case of low quality datasets, and the implementation of simple interaction procedures to give the possibility to the user to guide the refinement process.

Acknowledgment

This research work was partly funded by the EU Community FP7 ICT *Harvest4D* project (Grant Agreement 323567). We'd like to thank the people from Zamani Project (http://www.zamaniproject.org/) for the Maryam Church data.

References

[1] K. Pulli, H. Abi-Rached, T. Duchamp, L. G. Shapiro, and W. Stuetzle, "Acquisition and visualization of colored 3D objects," in *Proc. of the 14th International Conference on Pattern Recognition (ICPR'98)*, vol. 1. IEEE Computer Society, 1998, pp. 11–15.

[2] V. Sequeira and J. G. Goncalves, "3D reality modelling: Photo-realistic 3d models of real world scenes," *3D Data Processing Visualization and Transmission, International Symposium on*, pp. 776–783, 2002.

[3] C. Früh and A. Zakhor, "Constructing 3D city models by merging aerial and ground views," *IEEE Computer Graphics and Applications*, vol. 23, pp. 52–61, 2003.

[4] T. Franken, M. Dellepiane, F. Ganovelli, P. Cignoni, C. Montani, and R. Scopigno, "Minimizing user intervention in registering 2D images to 3D models," *The Visual Computer*, vol. 21, no. 8-10, pp. 619–628, sep 2005.

[5] K. Matsushita and T. Kaneko, "Efficient and handy texture mapping on 3D surfaces," *Computer Graphics Forum*, vol. 18, no. 3, pp. 349–358, 1999.

[6] L. Liu and I. Stamos, "Automatic 3D to 2D registration for the photorealistic rendering of urban scenes," in *CVPR*, vol. 2. IEEE Computer Society, 2005, pp. 137–143.

[7] P. J. Neugebauer and K. Klein, "Texturing 3D models of real world objects from multiple unregistered photographic views," *Computer Graphics Forum*, vol. 18, no. 3, pp. 245–256, 1999.

[8] D. G. Lowe, "Fitting parameterized three-dimensional models to images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, pp. 441–450, May 1991.

[9] L. Brunie, S. Lavallée, and R. Szeliski, "Using force fields derived from 3D distance maps for inferring the attitude of a 3D rigid object," in *Proc. of the Second European Conference on Computer Vision (ECCV'92)*. Springer-Verlag, 1992, pp. 670–675.

[10] H. P. A. Lensch, W. Heidrich, and H.-P. Seidel, "Automated texture registration and stitching for real world models," in *PG '00: Proceedings of the 8th Pacific Conference on Computer Graphics and Applications*. IEEE Computer Society, 2000, pp. 317–326.

[11] K. Ikeuchi, A. Nakazawa, K. Hasegawa, and T. Ohishi, "The great buddha project: Modeling cultural heritage for vr systems through observation," in *Proc. of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR'03)*. IEEE Computer Society, 2003, pp. 7–16.

[12] G. Y. Gehua Yang, J. Becker, and C. V. Stewart, "Estimating the location of a camera with respect to a 3d model," in *Proc. of the Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM2007)*. IEEE Computer Society, 2007, pp. 159–166.

[13] C. Wu, B. Clipp, X. Li, J.-M. Frahm, and M. Pollefeys, "3d model matching with viewpoint-invariant patches (vip)," in *CVPR'08*. IEEE Computer Society, 2008, pp. 1–8.

[14] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[15] P. Viola and I. William M. Wells, "Alignment by maximization of mutual information," *Int. J. Computer Vision*, vol. 24, no. 2, pp. 137–154, 1997.

[16] F. Maes, A. Collignon, D. Vandeermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions in Medical Imaging*, vol. 16, pp. 187–198, 1997.

[17] M. Corsini, M. Dellepiane, F. Ponchio, and R. Scopigno, "Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties," *Computer Graphics Forum*, vol. 28, no. 7, pp. 1755–1764, 2009.

[18] I. Cleju and D. Saupe, "Stochastic optimization of multiple texture registration using mutual information," in *Proceedings of the 29th DAGM conference on Pattern recognition*. Springer-Verlag, 2007, pp. 517–526.

[19] W. Zhao, D. Nister, and S. Hsu, "Alignment of continuous video onto 3d point clouds," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 8, pp. 1305–1318, Aug 2005.

[20] I. Stamos, L. Liu, C. Chen, G. Wolberg, G. Yu, and S. Zokai, "Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes," *Int. J. Comput. Vision*, vol. 78, pp. 237–260, July 2008.

[21] L. Liu, I. Stamos, G. Yu, G. Wolberg, and S. Zokai, "Multiview geometry for texture mapping 2D images onto 3D range data," in *Proc. of the 2006 IEEE Computer Society Conference on Computer Vision*

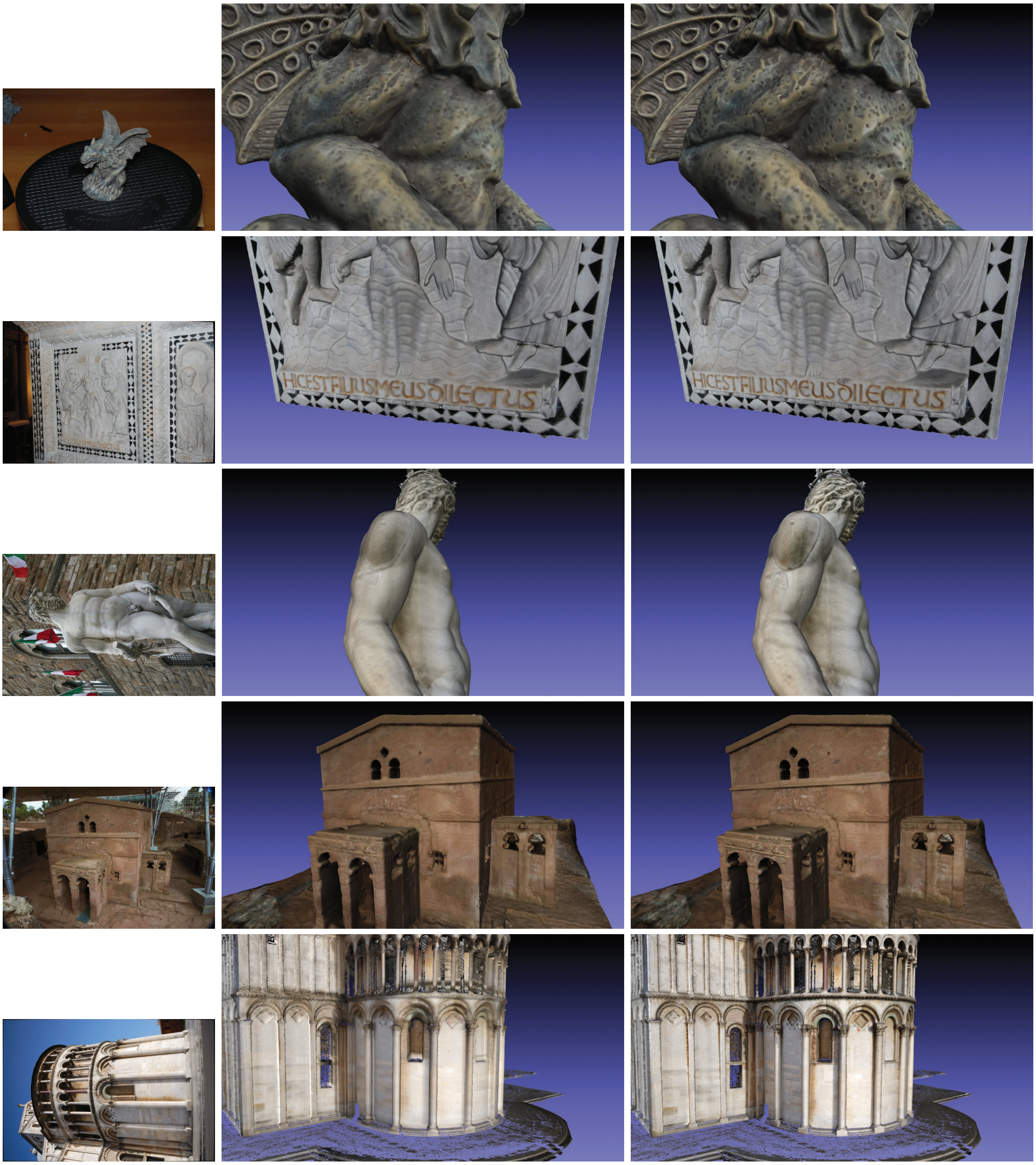

Fig. 6. The test cases. First column: a representative image of the ones used for color projection. Second column: color projection before refinement. Third column: color projection after refinement.

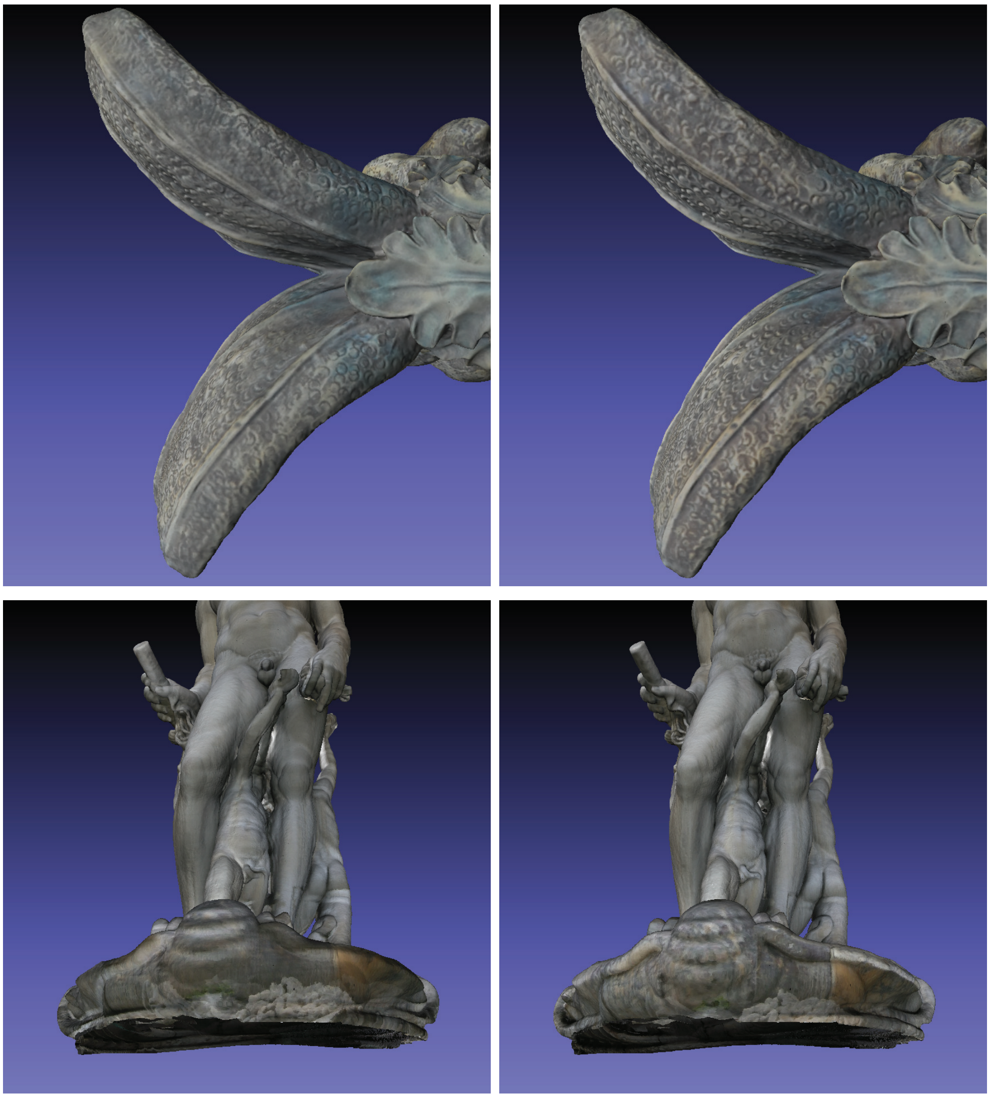

Fig. 7. Snapshots of the color projection before and after refinement. First two rows: a detail of the Gargoyle, before and after refinement. Last two rows: the bottom part of the Neptune, before and after refinement.

and Pattern Recognition (CVPR'06), vol. 2. IEEE Computer Society, 2006, pp. 2293–2300.

[22] H. Zheng, I. Cleju, and D. Saupe, "Highly-automatic mi based multiple 2D/3D image registration using self-initialized geodesic feature correspondences," in *ACCV (3)*, 2009, pp. 426–435.

[23] R. Pintus, E. Gobbetti, and R. Combet, "Fast and robust semi-automatic registration of photographs to 3d geometry," in *Proceedings of the 12th International conference on Virtual Reality, Archaeology and Cultural Heritage*, ser. VAST'11. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2011, pp. 9–16. [Online]. Available: http://dx.doi.org/10.2312/VAST/VAST11/009-016

[24] M. Corsini, M. Dellepiane, F. Ganovelli, R. Gherardi, A. Fusiello, and R. Scopigno, "Fully automatic registration of image sets on approximate geometry," *International Journal of Computer Vision*, vol. 102, no. 1-3, pp. 91–111, 2013. [Online]. Available: http://dx.doi.org/10.1007/s11263-012-0552-5

[25] M. Eisemann, B. De Decker, M. Magnor, P. Bekaert, E. de Aguiar, N. Ahmed, C. Theobalt, and A. Sellent, "Floating textures," *Computer Graphics Forum (Proc. of Eurographics)*, vol. 27, no. 2, pp. 409–418, Apr 2008.

[26] M. Dellepiane, R. Marroquim, M. Callieri, P. Cignoni, and R. Scopigno, "Flow-based local optimization for image-to-geometry projection," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 18, no. 3, pp. 463 –474, March 2012.

[27] K. Pulli, "Multiview registration for large data sets," in *Proc. of the 2nd international Conference on 3-D digital imaging and modeling (3DIM'99)*. Washington, DC, USA: IEEE Computer Society, 1999, pp. 160–168.

[28] M. J. D. Powell, "Developments of NEWUOA for minimization without derivatives," *IMA Journal of Numerical Analysis*, vol. 28, no. 4, pp. 649–664, Oct. 2008.

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

[30] M. Sottile, M. Dellepiane, P. Cignoni, and R. Scopigno, "Mutual correspondences: an hybrid method for image-to-geometry registration," in *Eurographics Italian Chapter Conference 2010*. EG, Nov 2010, pp. 81–88.

[31] M. Callieri, P. Cignoni, M. Corsini, and R. Scopigno, "Masked photo blending: mapping dense photographic dataset on high-resolution 3D models," *Computer & Graphics*, vol. 32, no. 4, pp. 464–473, Aug 2008.