123D Catch: Efficiency, Accuracy, Constraints and Limitations in Architectural Heritage Field

by

C. Santagati and L. Inzerillo

123D Catch: Efficiency, Accuracy, Constraints and Limitations in Architectural Heritage Field

C. Santagati and L. Inzerillo

123D Catch: Efficiency, Accuracy, Constraints and Limitations in Architectural Heritage Field

C. Santagati and L. Inzerillo

Abstract

Today, the accurate and detailed reconstruction of geometric models of real objects has become a common process. The diffusion of Image-based 3D modeling techniques, through image-based free, low cost and open source software, have increased drastically in the past few years, especially in the sector of Cultural Heritage (Architecture, Archeology, Urban planning). Nevertheless, web based software (ARC3D, 123D Catch, Hyp3D, my3Dscanner) offer another opportunity respect the desktop systems: they use the power of cloud computing to carry out a semi-automatic data processing. In this way is overcome the considerably slowing-down of the computer of hardware-heavy approaches. Our research investigates the limits and potentialities of 3D models obtained by using Autodesk 123D Catch in Architectural Heritage field, in order to verify the applicability of the method for the practitioner purpose. Our goal is to evaluate Catch efficiency, accuracy, constraints and limitations in order to provide insight into the current state of 3D modeling products. Our methodology forecasts the comparison between 123D Catch models and terrestrial laser scanner 3D acquisition in order to verify metric accuracy. Therefore, we propose a lot of interesting examples applied in architectural heritage field from the detail to the large scale.

Key words: Image-based modeling, 3D reconstruction, Photogrammetry, Computer Vision, Architectural Heritage

1. Introduction

The past few years have seen significant progress toward automatic creation of 3D models. 3D models have become an essential part of many applications ranging from documentation, digital restoration, visualization, inspection, planning, AR/VR, gaming, entertainment, etc.

Carrying out 3D modeling from photographic images is an efficient and intuitive way to create 3D digital models of objects. As Nguyen, H.M., et al. (2012) well described, if compared with conventional geometry-based modeling and hardware-heavy approaches, the image-based modeling method can be employed to extract original texture and illumination directly from images for visual 3D modeling, without the need for complicated processes, such as geometry modeling, shading and ray tracing.

These techniques – named Structure from Motion (SfM)- are usually less accurate, but offer very intuitive and low cost methods for reconstructing 3D scenes and models (Remondino et al., 2012).

There are now a number of software packages that offer the ability to acquire 3D models from a set of images without any a priori information about the scene to be reconstructed. 3D reconstruction from images has undergone a revolution in the last few years.

Thus, there is a growing attention among academics and practitioners due to the great potentialities of these systems.

At present time there are two possibilities: desktop and web based packages.

On the one hand desktop software allow us to customize and verify immediately the automatic reconstruction but have long time processing and hardware-heavy approaches; on the other hand web based packages exploit the power of cloud computing, allowing to carry out other tasks; at the same the time web based software work like a black box and, almost of time the user does not have the possibility to interact with the software in order to improve the outcomes.

Furthermore, there is also another important difference in the output provided by the different software: usually web services provide a textured mesh model on the other hand many of desktop systems provide a point cloud that need to be processed and transformed into a mesh.

In this paper we focus on architectural heritage digitalization by using 123D Catch by Autodesk, one of the more used web based packages. Among all available web based software (ARC3D, 123D Catch, Hyp3D, my3Dscanner) we chose Catch for the easiness of use, the visual quality of the reconstructed scene and the possibility to interact with and develop the results. Furthermore, Catch 3D mesh is suitable for all 3D modeling software.

Our goal is to give an answer on visual and metric accuracy and provide a guide to researchers, academics and, even more, practitioners who want to use this free package.

This way, the practitioner will be able to act in the field of architectural surveying without using expensive technologies and without having an extremely specific expertise. As a matter of fact, tools such as laser scanners and structured lighting systems are often costly, time consuming, not very portable and flexible to use.

We identify the methodology for using the software and we verify and demonstrate its metric reliability.

Several tests have been carried out both on the small object and on the medium and large scale. We tested also on private buildings and we verified the correct methodology for taking a good set of images. Thus, by means of this study the practitioner will be able to understand when/how/where he can use this technology and which is the best methodology to adopt in order to get good results.

Hence, the research is mainly addressed on the detection of limits and criticalities of the algorithm that controls the automatic process. We deeply investigated on the metric reliability of the 123D Catch models comparing them with terrestrial laser scanner acquisitions or reliable Ground Control Points (GCP), on the surfaces reconstruction quality and on the detail quality in relation with the number of images and their resolution.

The main fundamental aspects to investigate are two:

1. the mesh visual accuracy: we obtain a 3D model that reproduces reality;
2. the mesh metric accuracy: we obtain a 3D model that is metrically close to reality.

Therefore, we will deal with two aspects: visual and metric accuracy. Obviously the second is a functional development of the first one. Several tests have been fulfilled. In this paper we present the outcomes carried out on San Rocco church in Acireale, Auteri chapel and San Nicola l'Arena church in Catania whose geometry and complexity allowed us to carry out a good and complete stress-test methodology.

1.1 Related Work

The continuous updating and optimization of algorithms, on the automatic reconstruction process of internal and external camera parameters, make obsolete some works in literature. For instance, the limitations detected by Galizia and Santagati (2012) in the management of high resolution images. The main part of the studies aimed at testing and the exploring of 123D Catch potentialities and

C. Santagati and L. Inzerillo

limits is addressed to the digitalization of little objects (Nguyen et al, 2012), of archaeological finds such as fragments and furnishings (Kersten, Lindstaed, 2012), small archaeological site parts (Lo Brutto, Meli, 2012; Dellepiane et al., 2013) and statues, such as the excellent outcomes obtained by Kestern (Kersten, Lindstaed, 2012) on the Moai statues of Easter Island.

Instead, in architectural field there is a lack of systematic studies, (Kestern, Stallmann, 2012; Manferdini, Galassi, 2013). Hence, rather than getting a comparison between 123D Catch and other image based modeling tools, we consider more useful investigate on 123D Catch architectonic mesh metric and visual accuracy.

Therefore, with this study we started a series of tests in order to fill this gape and provide some methodological directions.

2. Visual Accuracy

2.1 123D Catch

123D Catch web-service, at the present time still available for free in beta version, was developed in Autodesk laboratories and overcomes the previous Photofly project, launched in summer 2010, using technology developed by Realviz (now Acute 3D). Indeed, the used approach underlying 123D Catch technology is well described in Vu, H-H. et al. (2009).

Exploiting the photogrammetric approach and the algorithms of Computer Vision, 123D Catch is able to reconstruct internal parameters of digital camera and the position in space of homologous points finding out a number of correspondences between photographic images sequences, suitably taken.

Indeed, 3D coordinates of all points of the scene are found through the correspondence pixel-pixel and the polygonal model is well reconstructed.

Furthermore, among all the web-services actually available, 123D Catch is the only one that allows to improve the result of the 3D scene reconstruction through the manual stitching of homologous points on triplets of images. At the end of the stitching operation Catch starts again the processing on the cloud.

The main steps to use 123D Catch are to:

- Capture a photographic sequence of an object in order that the angle between one shot and the other is about 5-10 degree and the overlapping is about 70%;
- Use the iPhone, iPad, web, or desktop app to upload the photos to the Autodesk cloud where they are converted into realistic 3D models; in this phase the user can decide whether to wait the 3D reconstruction or to be advised by email;

- If the 3D reconstruction is unsatisfactory or some picture are missing, then the user can improve the outcomes through manual stitching of homologous points on triplets of images and submit again the scene to the cloud;
- Create a video, share with others, or even fabricate your project with 123D's 3D printing or laser cutting services.

Furthermore, for a use in cultural heritage visualization field it is necessary to scale and post-process the obtained model in order to define all the imperfections (noise and holes) in mesh quality. The photorealistic mesh model carried out by Catch is exportable in different formats: obj, dwg, fbx, rzi, ipm, las. Thus, you can use whatever post processing software, even open source tools, such as Meshlab (Cignoni, P. et al., 2008).

In the next section we report the right methodology to use Catch to carry out the best visual mesh. For this purpose we applied Catch on an architectural element of the end of XIX century and on two civil buildings.

Figure 1. Datasets used for visual accuracy tests

	Number of images	Stitching
Architectural element	23	0
Civil building (A)	38	9
Civil building (B)	47	7

Table 1. Datasets used for visual accuracy tests
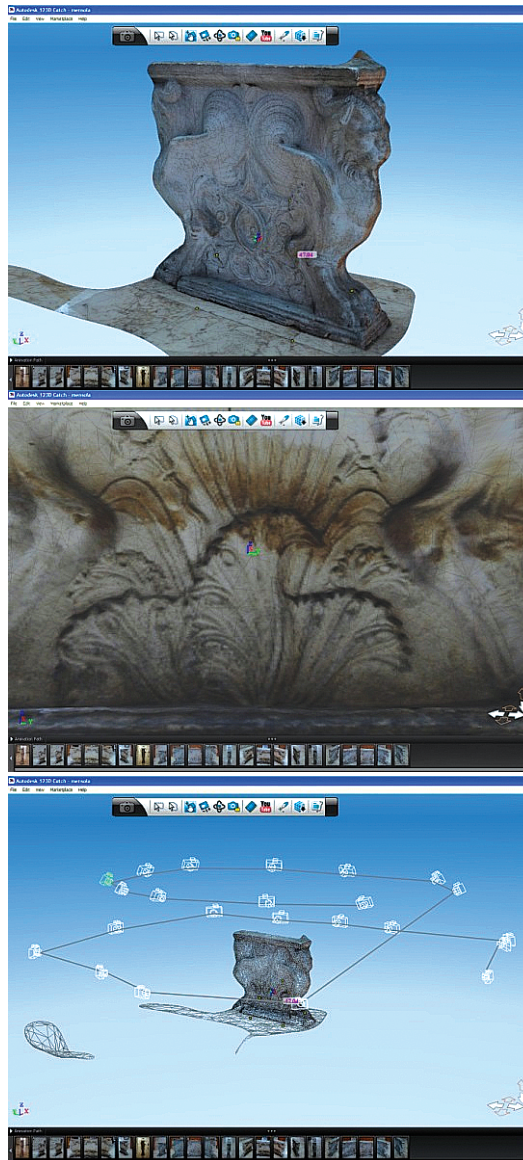
2.2 How Take the Pictures.

Reconstruction process begins by estimating parameters for photos dataset sequence.

1. The sequence is very important to reach satisfactory outcomes. In fact if you change the sequence also the result changes. The pictures must be taken according to a path of continuity around the object and their submission on Catch must be the same. The right sequence to take the pictures is well shown in fig. 2-3.
2. The chosen images should have a large number of

C. Santagati and L. Inzerillo

correspondences. As before said, is necessary to capture a photos dataset sequence in order that the angle between one shot and the other is about 5-10 degree and the overlapping is about 70%. Thus, frame reconstruction can be robustly estimated. This is another strong condition to ensure a good outcome. The best achievement obtained using the correct angle between shots is shown in fig. 2.

3.   The last condition, to carry out an available visual mesh, is to

Figure 2: The photographic sequence needed for the acquisition of an architectural element.

take the picture so as to visualize the entire building. Thus, if

your architecture building is located in a tricky road, it is not possible to create a good cluster of structured pictures and you will never have a mesh. This case is shown in fig. 4. The building must be located in a square or in a large street where you can take photos at the right distance. This is a strong application limit of 123D Catch on architectural buildings. To carry out an architectural mesh of an entire building is necessary that you can make a full turn around it. This the most significant 123D Catch limit.

In case of art objects (small objects) is necessary to make at least two full turns around the object at different heights. For example, for
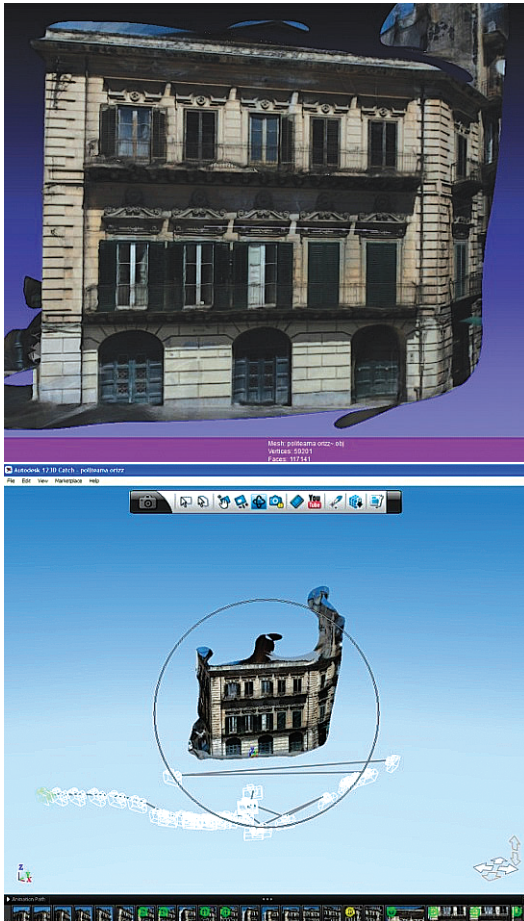
Figure 3: 123D Catch calculation on a civil building (A) in Palermo

the sculpture in fig. 2, have been enough to create a dataset of 23 photos. 123D Catch is very reliable on small objects both on the metric and on the geometric aspect. In figure 2 is shown a detail and you can observe the visual quality achieved.

2.3 Camera Parameter Estimation

Figure 4: 123D Catch calculation on a civil building (B) in Palermo
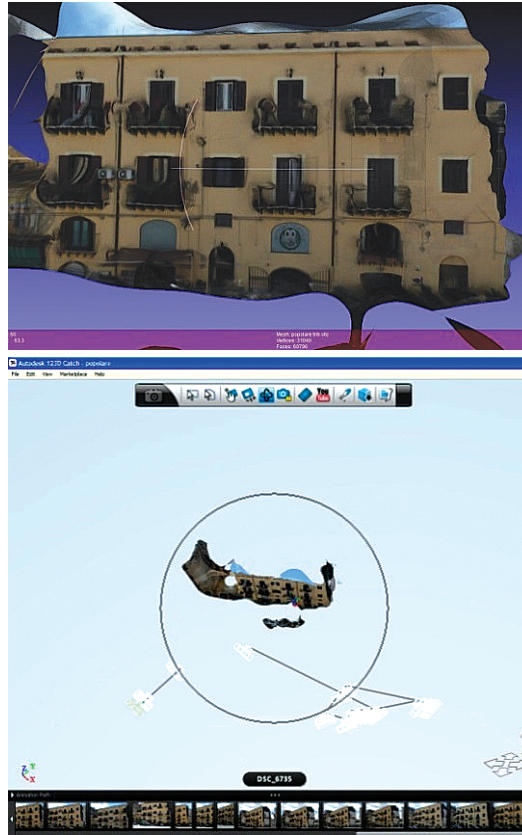
Given a set of matching images, the goal of this stage is to recover simultaneously the geometry of the scene and the Structure for Motion (SfM) (Furukawa and Ponce, 2007; Remondino, et al., 2012; Snavely, 2008; Wu, 2011). SfM includes the extrinsic (position, orientation) and intrinsic parameters of the camera for the captured images.

123D Catch processes only photos taken by a single camera. In case of change of lens or camera gear, or use of wide-angle, or use images downloaded from the web, Catch's algorithm is not able to process the mesh because it does not recognize the homologous points. Photos dataset have to follow the parameters shown in the previous section.

The amount of pictures to take is relative to the object to be processed and the amount of detail to provide (view section 3.3.2).

2.4 Mesh Quality

Catch allows to choose an output quality for mesh. There are three choices:

1. the mobile one which is fast, suitable for viewing on mobile devices;
2. the standard one -which is the recommended one from Catch- with high resolution textured mesh and it is the best for the visualization on the desktop;
3. the maximum one which is a very high density mesh, suitable for manipulating in external applications.

However, the resolution quality carried out from the standard output quality mesh is very approximate if compared with the maximum one (fig. 5).
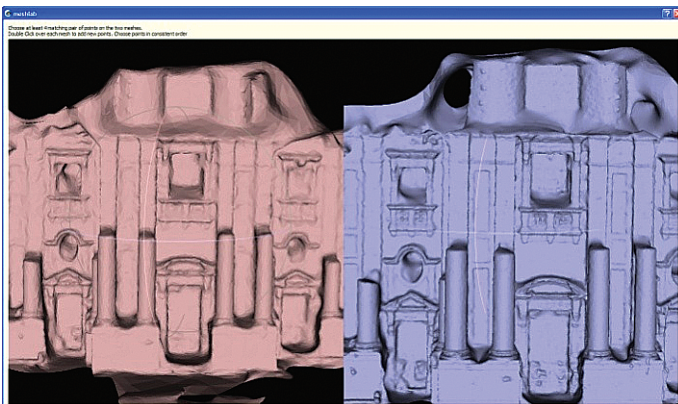
Figure 5: On the left side there is the standard output quality mesh and on the right side the maximum one

3. Metric Accuracy

As stated in the introduction, our methodological approach is focused on testing 123D Catch on different architectural size objects to verify its reliability and give to practitioners some milestone on which to build.

Thus, our chosen case studies span from architectural detail to large scale architecture. For testing we chose some architectural details of the church of San Rocco in Acireale, the façade of San Rocco (mid-scale), the façade of San Nicola l'Arena in Catania (large-scale) and Auteri Chapel in Catania.

As it is well visible in fig. 6, all chosen objects have both a strong three-dimensionality and a complex geometry; thus, allowing us to conduct a good test methodology.

We did not carry out any test on primitive geometries (i.e. pyramid, cylinder, parallelepiped) thanks to:

C. Santagati and L. Inzerillo

- geometric features of case studies;
- comparison between surfaces and not between point clouds of the models.

Architectural details

Entire architecture
(Auteri Chapel, Catania)

Mid-scale architecture
(San Rocco, Acireale)

Large-scale architecture
(San Nicola, Catania)

For metric comparison we used point clouds carried out by TOF (Time of Flight) laser scanner 3000 HDS by Leica Geosystem of Laboratory of Architectural Photogrammetry and Survey "Luigi Andreozzi" (University of Catania).

3D model made in 123D Catch has been exported in obj format.

Metric comparison has been carried out in Meshlab, an open source software able to scale, align and process both point clouds and meshes.

The alignment has been carried out scaling and glueing Catch mesh on laser mesh using ICP algorithm. Alignment outcomes between two meshes have been verified applying Hausdorff distance filter and visualized through vertex quality filter. The achievements are shown in a red-green-blue scale, where red means good and blue means bad.

Furthermore, we carried out a series of vertical and horizontal cross-sections on the two aligned meshes in JRC Reconstructor environment to quantify/visualize better the gaps between the two meshes.

3.1 Architectural Details/Elements

Regarding architectural details we chose a part of the base and a capital of San Rocco church in Acireale.

3.1.1 Part of the Base

Table 1 summarizes all the information about dimension, resolution, images number and information on the mesh quality in terms of number of vertices and triangles (faces). The shots have been carried out by a Coolpix E8800. The base 3D model of the includes 14 images and its visual appearance is very good.
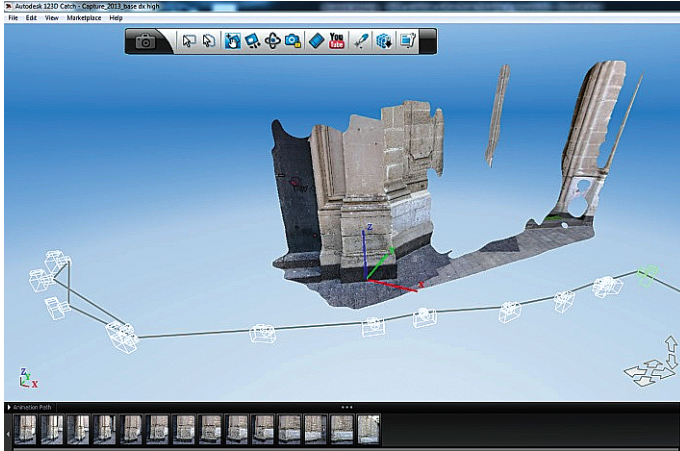

Figure 7: Part of the base 123D Catch model

We observe that 123D Catch model is more detailed than point cloud one. The outcomes of the comparison denote an excellent overlap. The average error is contained between 0.0025 and 0.001m.

Architectural element	Part of the base
Dimension of the object	1.50x1.50 m
Number of images	14
Resolution	8.5Mpixel
123D Catch mesh	693,736 triangles 368,357 vertices
Laser scan point cloud	154,901 vertices
Average error	0.0025-0.001m

Table 1. Part of the Base dataset
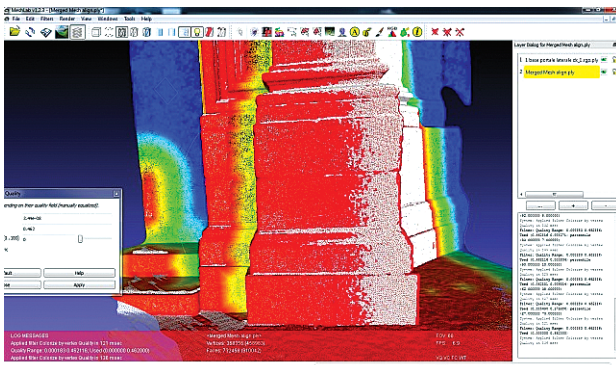
Figure 8: Part of the base mesh alignment error visual evaluation: red (good) blue (bad)

C. Santagati and L. Inzerillo

Furthermore, both the horizontal and vertical cross sections carried out with a step of 0.05 m reveal a very good quality of 123D Catch model (figure 10).

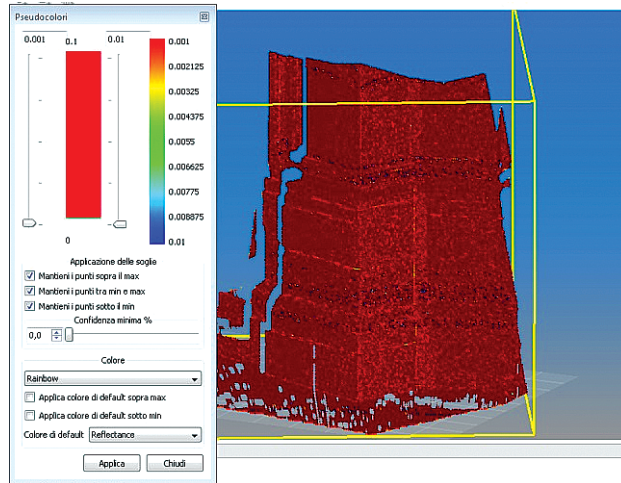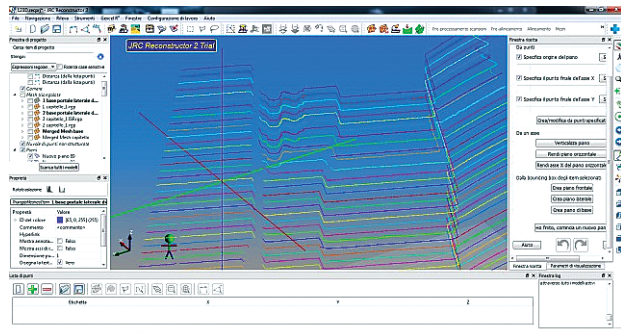Figure 9: Part of the base mesh alignment evaluation histogram: error red (good) blue (bad)

Figure 10: Part of the base mesh alignment evaluation error through horizontal cross sections

3.1.2 Capital

The 3D model of the capital includes 10 images and, as well shown in figure 11, model's visual appearance is good although the shots are taken from the ground.

All the information regarding this second dataset are shown in Table 2, that reports also the information on mesh quality regarding number of vertices and triangles (faces). Thus, once again 123D Catch model is more detailed than points cloud one.

The comparison outcomes declare a good overlap. The average error is contained between 0.01m and 0.001m.
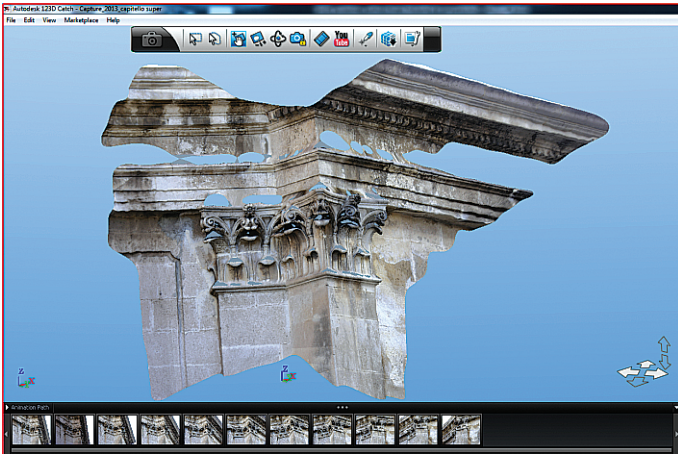
Architectural element	Capital
Dimension of the object	1.50x1.50 m
Number of images	11
Resolution	8.5Mpixel
123D catch mesh	872,000 triangles 438,845 vertices
Laser scan point cloud	179,784 vertices
Average error	0.01-0.001m

Table 2. Capital dataset

Furthermore, both horizontal and vertical cross sections carried out with a step of 2.5 cm show a very good quality of 123D Catch model.
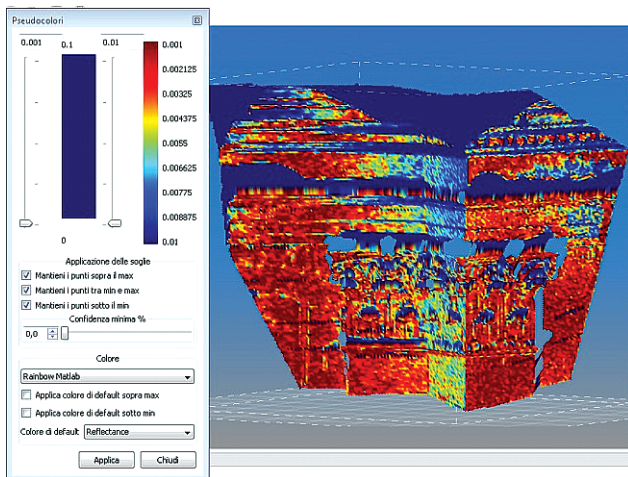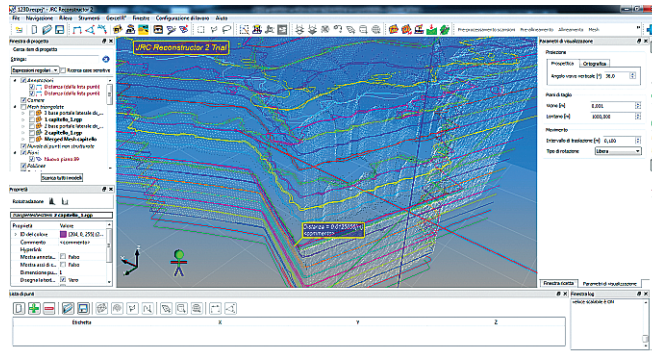

Figure 12: Capital mesh alignment evaluation histogram: error red (good) blue (bad)

Considering achieved results, we assert that on architectural details both visual and metric accuracy is reliable as much as TOF 3D laser scanner accuracy (0.006m).

Figure 13: Capital mesh alignment error evaluation through horizontal cross sections


3.2 Medium Scale Architecture

The dataset regarding the church of San Rocco in Acireale, mainly covers the first order of the façade, because the church is located in a street that does not allow to capture the entire façade.

Therefore, we used a dataset of 45 images covering the first order of the façade surface about of 17x8m.

Figure 14: San Rocco façade 123D Catch model
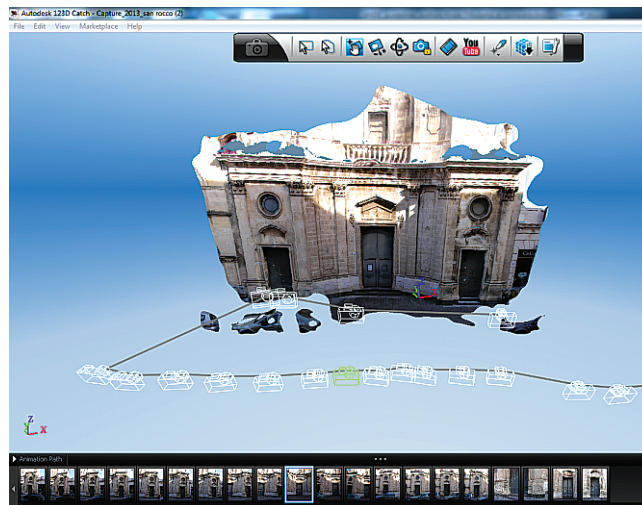
In Table 3 you get all information you need about San Rocco dataset: mesh quality, number of vertices and triangles (faces).

The comparison outcomes reveal a good overlap. The average error is contained between 0.005 and 0.02m. Considering that the accuracy of the used terrestrial laser scanner is 0.006 m we still assume that, on mid-scale architecture, you get good achievements.

Furthermore, both horizontal and vertical cross sections carried out with a step of 0.25 m denote a very good quality of 123D Catch model.

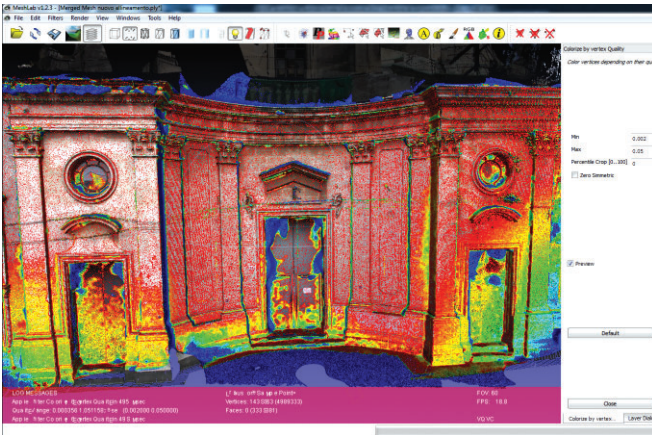| Medium Scale Architecture	San Rocco
Dimension of the object	17x8 m
Number of images	45
Resolution	8 Mpixel
123D catch mesh	2,656,863 triangles 1,362,712 vertices
Laser scan point cloud	1,593,930 vertices
Average error	0.005-0.02m

Table 3. San Rocco dataset

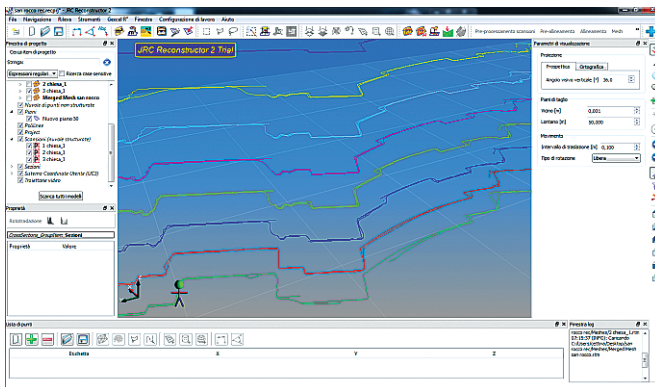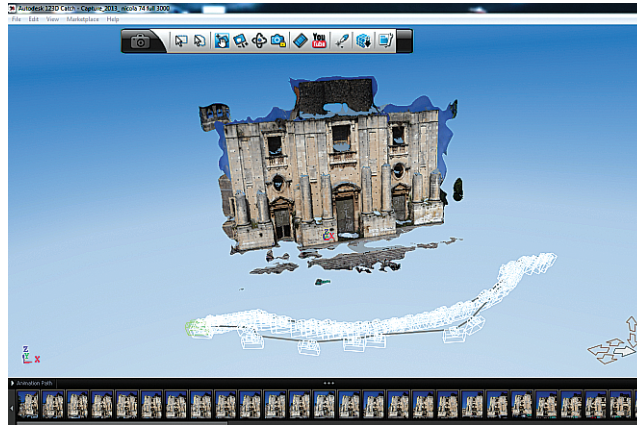Figure 15: Mesh alignment error evaluation: red (good) blue (bad)

Figure 16: Mesh alignment error evaluation, through horizontal cross sections, on left side of the façade

Figure 17: Mesh alignment error evaluation, through horizontal cross sections, on right side of the façade

## 3.3 Large Scale Architecture

To get the large scale architectural building test, we chose the monumental façade of the church of San Nicola l'Arena in Catania of 47x25m. The shots have been carried out by a Coolpix L22. Table 3 denotes all the information regarding dataset dimension, resolution and images number.

Dealing with a large scale architecture, we verified how and if the images number affected both visual accuracy and metric accuracy of 3D reconstruction. Hence, we got two datasets with 74 and 136 photos.

3.3.1 San Nicola 136 dataset

Table 4 reveals all resolution, dimension of the object, images number information. During Catch processing it was necessary to stitch manually 50 images to sketch in 86 automatically processed.

Figure 18: San Nicola façade 123D Catch model
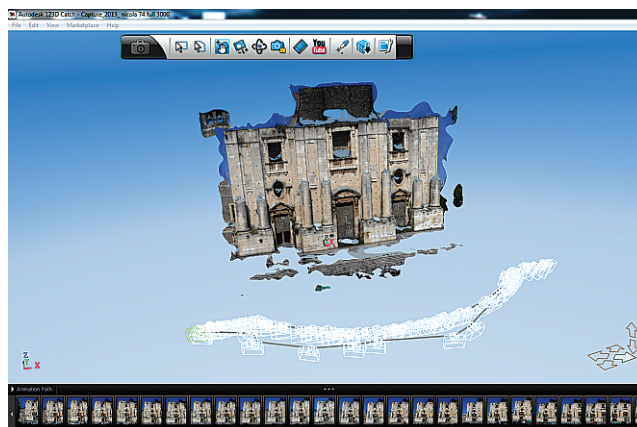
Because of terrestrial laser scanner mesh heaviness, we carried out a light mesh in JRC Reconstructor preserving all sharpen edges. Then we got meshes alignment. Whereas, the comparison has been carried out with original point cloud mesh.

Large Scale Architecture	San Nicola
Dimension of the object	47x25 m
Number of images	136
Resolution	8.5 Mpixel
123D catch mesh	2,481,252 triangles 1,261,619 vertices
Laser scan point cloud	2,801,344 vertices
Average error	0.005-0.02m

Table 4. San Nicola 136 photos dataset

The comparison outcomes reveal a very good overlap. The average error is contained between 0.02 and 0.005m. Hence, if we consider that used terrestrial laser scanner has an accuracy of 0.006, we still assert that on large-scale architecture we get good outcomes.

Furthermore, both horizontal and vertical cross sections, carried out with a step of 0.50 m, denote 123D Catch model quality.
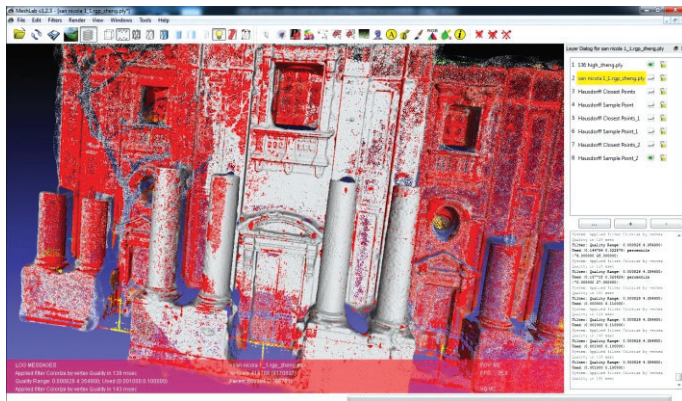

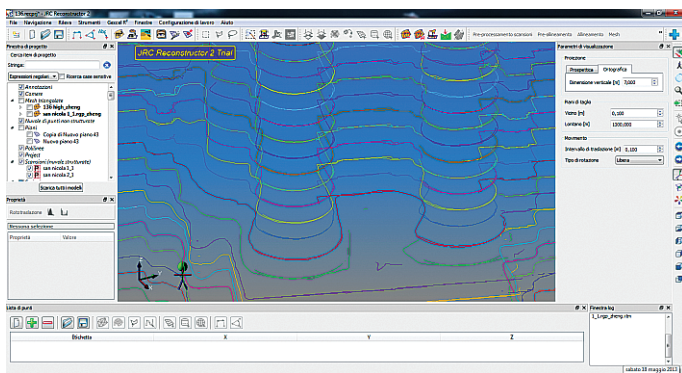Figure 19: Mesh alignment error evaluation: red (good) blue (bad)

Figure 20: Mesh alignment error evaluation through horizontal cross sections

C. Santagati and L. Inzerillo

3.3.2 San Nicola 74 Dataset

We applied the same investigation methodology to verify metric accuracy of 74 images dataset 3D model.

Table 5 reports all information regarding 74 images dataset. The comparison outcomes reveal a misalignment error of about 0.10 m. Dealing with an architectural monument, even if a large scale one, this value is not acceptable.

Table 5. San Nicola 74 photos dataset

Large Scale Architecture	San Nicola
Dimension of the object	47x25 m
Number of images	74
Resolution	8.5 Mpixel
123D catch mesh	982,324 triangles 1,955,322 vertices
Laser scan point cloud	2,801,344 vertices
Average error	0.05-0.15 m

The average error is contained between 0. 15 and 0.05 m. Furthermore, both horizontal and vertical cross sections carried out with a step of 0.50 m highlight 123D Catch model poor quality.

Hence, we assert that low accuracy in terms of clearness and sharpness of the edges causes a major uncertainty in the geometrical identification of the model so as to generate very serious errors.

Figure 21: Mesh alignment error evaluation: red (good) blue (bad)
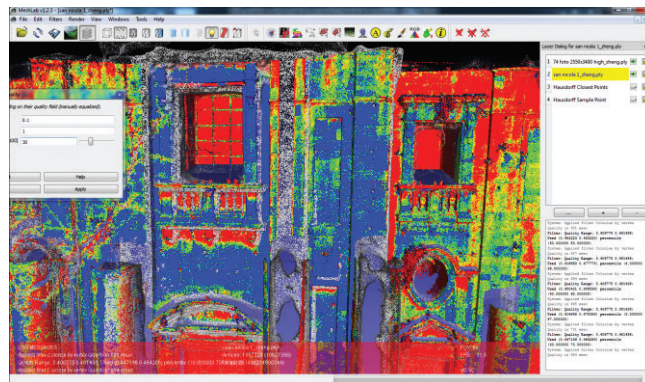
3.3.3 Input Image Resolution

The last test we carried out deals with the incidence of image resolution on metric accuracy.

We worked on San Nicola datasets considering four different datasets accordingly with images number and resolution as shown in Table 6.
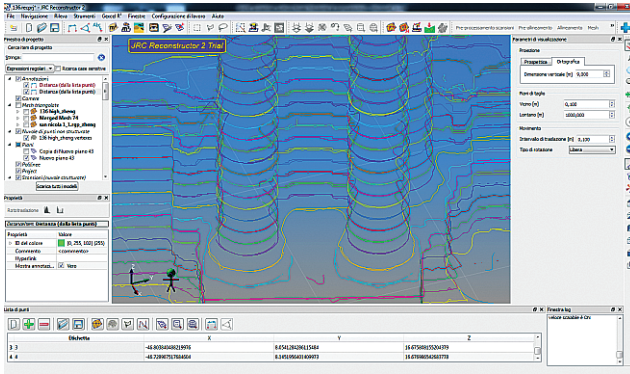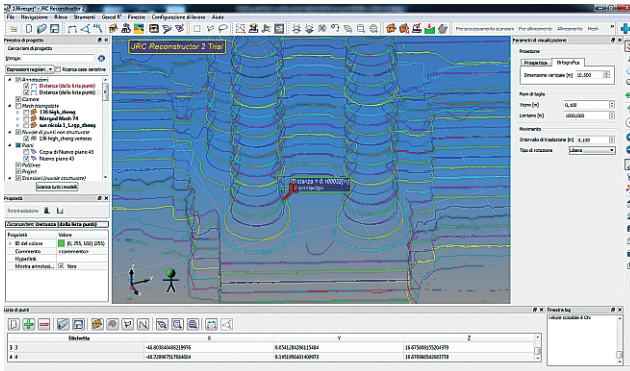
Figure 22: Mesh alignment error evaluation through horizontal cross sections

San Nicola	74 photos	136 photos
1050x1400 pixel	340082 faces 170897 vertices	502,389 faces 251,915 vertices
2550x3400 pixel	982,324 faces 1,955,322 vertices	2,577,777 faces 1,294,529 vertices

Table 6. Datasets for image resolution/ images number tests

Since in the previous section we demonstrated that quality of mesh affects the accuracy of metric model, it seemed sufficient to perform just only a visual comparative analysis.

The comparisons were made in Meshlab visualizing the meshes both in textured and in smooth mode.

The presence of holes has been considered of secondary importance compared to the quality of mesh's graphic detail.

In figures 23 and 24 are clearly visible some common 3D reconstruction deficiencies: holes and lack of sharpness.

The latter is more important to carry out a high visual and metric mesh accuracy.

123D Catch manages leaner, faster and reliable a photos dataset not so much wide. On the other side a wide range of photos dataset guarantees more resolution, sharpness and geometry accuracy, especially in details.

C. Santagati and L. Inzerillo

123D Catch tends to close automatically mesh holes without considering the actual geometry of the object. The reliability of the holes closed automatically by Catch is not acceptable.

We can conclude that it is preferable to give a photo dataset with high resolution images.

Figure 23: San Nicola 74 dataset – on the left HR photos 3D model, on the right LR photos 3D model
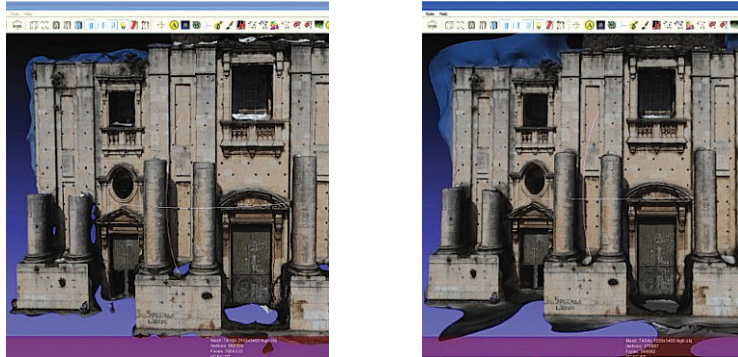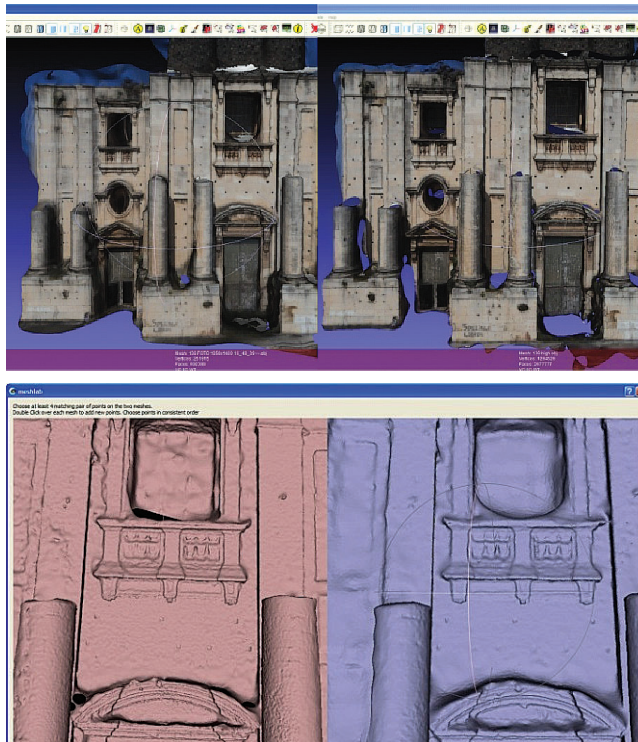

Figure 24: San Nicola 136 dataset – on the left HR photos 3D model, on the right LR photos 3D model

3.4 Entire Architectural Object

We tested 123D Catch visual and metric accuracy on a façade of a building or church and verified its excellent outcomes.

In this section, we will get 123D Catch 3D model of an entire architectural object.

The third strong condition, just treated on section 2.2, has reduced our choice on architectural building. It was really difficult to find a building isolated from the context and proportionate to the road so that it could be photographed entirely. It becomes more difficult, if you think architectonical typologies and urban planning of Latin culture with high buildings in narrow streets.

We tested Catch on a little chapel in Catania: Auteri Chapel. We chose this architectural building for its simple cylindrical shape.

Entire architectural object	Auteri Chapel
Dimension of the object	4,20x5,14 m
Number of images	40
Resolution	4 Mpixel
123D catch mesh 306,001 vertices	590,869 triangles

In Table 7 you get all information you need about Auteri Chapel dataset: dimension of the object, mesh quality, number of vertices and triangles (faces).

123D Catch 3D model visual accuracy is very good apart from the covering (this problem is because we couldn't take picture of it). Figure 26 shows the excellent sharpness of cylindrical surface both of the building and of niches and columns.
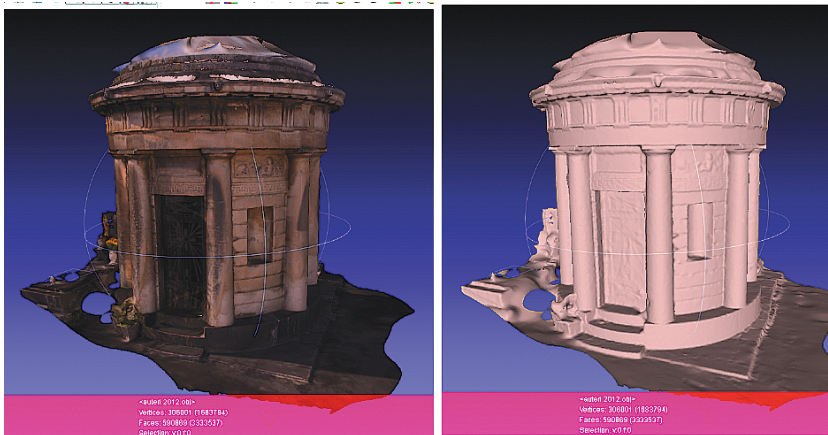

Figure 26: Auteri Chapel 123D Catch model, on the right smooth visualization

123D Catch 3D reconstruction encountered some problems for the lacking of continuities of images in dataset, due obstacles.

For metric accuracy test we used horizontal and vertical cross sections compared with Ground Control Points (GCP) carried out by a reliable direct survey .

C. Santagati and L. Inzerillo

Therefore, 123D Catch model was scaled and referred for overlapping to direct survey drawings. Figure 28 shows metric accuracy evaluation considering two significant horizontal cross-sections: one at 0.75 and the other 1.50 m. Except for the areas where 3D reconstruction is not geometrically exact, the gaps are about of 0.01m, thus confirming what previously tested.

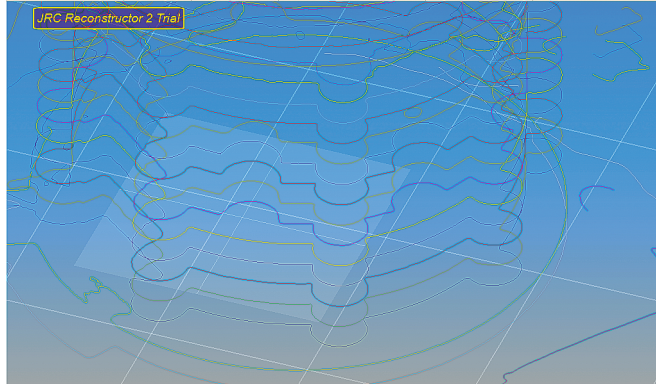Figure 27: Auteri Chapel horizontal cross sections

Figure 28: Auteri Chapel reconstruction error evaluation on horizontal profiles: yellow 123D Catch model, red GCP survey

4. Conclusions and Future Works

We described the overall design of image-based reconstruction algorithms, and evaluated a number of 3D reconstruction models. We can conclude that 123D Catch is an excellent tool for Image Based Modeling. However, to give a reliable overview, we have noted Catch tips, advantages and disadvantages.

Advices:

1. Metric accuracy was significantly affected by mesh quality. Therefore, we must apply the two parameters that control mesh quality, namely: resolution of dataset and number of images. The number of pictures must be appropriately

selected depending on size and level of detail and according to parameters that regulate photogrammetry. For large scale Architectures is better produce your photo datasets as large as possible to ensure a metric accuracy of a few centimeters.

Points of strengthen:

1. Low processing times;
2. Processing on cloud;
3. Metric accuracy for statues, work of arts, archeological and architectural details/elements applications is of the order of mm.
4. Metric accuracy for medium scale architectural buildings applications is of the order of 1-2 cm
5. Metric accuracy for large scale architectural buildings applications is of the order of 1-2 cm
6. In order to obtain processing reliable according to metric accuracy, it is necessary to use cameras with a resolution between 6-12 Mpixel. Therefore, it is possible to use non professional cameras without specific lens.

All those advantages cut down the costs not only in terms of required equipment but also in terms of hours/man, hours/machine you have to consider when you start a 3D digitalization project. As a matter of fact, when you manage to plan a laser scanner project, all those ratios affect very weightily on the intervention.

Points of weakness:

1. Photos dataset must be structured.
2. The building or the object to capture should be shot in its entireness. 123D Catch is not able to manage the overlapping between two frames in height.

This latter condition strongly limits the use of this tools in several architectural applications. Almost of time, you are not in the optimal condition to capture a good photo dataset.

Practitioner needs to have as much as possible information on the entire architectural building, to carry out horizontal and vertical cross-sections, elevations, etc, for its professional activity.

Since our study is addressed to verify 123D Catch, giving a guide to practitioner to use such a powerful low cost tool, we can assume that: even though visual and metric accuracy are excellent, nevertheless, the need to capture the building in its entireness, considerably reduces the possible case studies, and then, a full use by practitioner.

Otherwise, it's an excellent tools for other application fields such as:

1. Researcher investigations;
2. Archeological survey;
3. museums visual art collections survey;
4. architectural elements survey.

123D Catch visual and metric accuracy and reliability testing on both the small and on a large scale was a critical step, so far lacking in literature.

The achieved outcomes are promising. Nevertheless, among some issues, that still remain open, we suggest:

1. Comparison on the same datasets with other available SfM tools both on line and desktop (i.e. Photoscan by Agisoft and the open source package VisualSfM developed by Wu and its plugin CMVS (PMVS2) by Yasutaka Furukawa);
2. Data mining of all professional-technical representation such as orthophotos, profiles, etc;
3. Possibility of use this tool for integrating lacking in laser scan point clouds without losing metric accuracy.

5. References

Agarwal, S. (et al.), 2010. Reconstructing Rome. Computer, 43 (6), pp. 40-47

Bandiera A. (et al.), 2011. Nascita ed utilizzo delle tecniche digitali di 3D imaging, modellazione e visualizzazione per l'architettura e i beni culturali. Ikhnos, Lombardi editore, pp. 81-134.

Cignoni, P. (et al.), 2008. MeshLab: an Open-Source Mesh Processing Tool, Sixth Eurographics Italian Chapter Conference, pp. 129-136.

Dellepiane (et al.), 2013. Archeological excavation monitoring using dense stereo matching techniques. Journal of Cultural Heritage, Volume 14 Number 3, page 201-210 – May-June 2013.

Ferdani, D., 2012. Dense stereo matching based systems in archaeology: applications, benefits and limitations. In: Proceedings of MIMOS, Rome, Italy, 9-11 October 2012, pp. 217-224.

Filippucci, M., 2010. Nuvole di pixel. La fotomodellazione con software liberi per il rilievo d'architettura. Disegnarecon, 3, pp. 150–163.

Fratus de Balestrini, E., Guerra, F., 2011. New instruments for survey: on line softwares for 3d recontruction from images. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII-5/W16, pp. 545- 552.

Furukawa, Y., 2010. PMVS2, Department of Computer Science and Engineering, University of Washington, Seattle, Washington.
http://grail.cs.washington.edu/software/pmvs/ (accessed 15 May 2013)

Furukawa, Y., Ponce, J., 2007. Accurate, Dense, and Robust Multi-View Stereopsis, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, July 2007.

Furukawa, Y. (et al.), 2010. Towards Internet-scale Multi-view Stereo, IEEE Computer

Society Conference on Computer Vision and Pattern Recognition, July 2010.

Galizia M., Santagati C., 2012. Low-cost image-based modeling techniques for archaeological heritage digitalization: more than just a good tool for 3D visualization?, Proceedings of ARQUEOLÓGICA 2.0, Seville, Spain, 20-22 June 2012.

Hiep V.H. (et al.), 2009. Towards high-resolution large-scale multi-view stereo. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, July 2009.

Kersten, Th. (et al.), 2012. Automatische 3D-Objektrekonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie. Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation, 21, pp. 137-148.

Kersten, Th., Stallmann, D., 2012. Automatic texture mapping of architectural and archaeological 3d models. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXIX-B5, 273-278.

Kersten, Th., Lindstaedt, M., 2012. Image-Based Low-Cost Systems For Automatic 3d Recording And Modelling Of Archaeological Finds And Objects, M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, R. Caffo (Eds): EuroMed 2012, Progress In Cultural Heritage Preservation, LNCS 7616, Springer-Verlag Berlin Heidelberg 2012, Pp 1-10

Lo Brutto M., Meli, P., 2012. Computer vision tools for 3d modelling in archaeology, International Journal of Heritage in the Digital Era, Volume 1, Supplement , pp 1-6

Manferdini AM, Galassi M., 2013. Assessments for 3d reconstructions of cultural heritage using digital technologies. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W1, 2013, pp 167-174

Nguyen, H.M., (et al.), 2012. 3D Models from the Black Box: Investigating the Current State of Image-Based Modeling. Proceedings of the 20th International Conference on Computer Graphics, Visualization and Computer Vision (WSCG 2012), Pilsen, Czech Republic, June 25-28, 2012

Pierrot-Deseilligny, M. (et al.), 2011. Automated image-based procedures for accurate artifacts 3D modeling and orthoimage generation, Proc. CIPA, 2011

Remondino, F., 2011. Accurate and Detailed Image-Based 3D Documentation of Large Sites and Complex Objects. F. Stanco, S. Battiato and G. Gallo, (eds) Digital imaging for cultural heritage preservation: analysis, restoration, and reconstruction of ancient artworks, CRC Press, pp.127-158

Remondino, F. (et al.), 2012. Low-Cost and Open-Source Solutions for Automated Image Orientation–A Critical Overview, M. Ioannides, D. Fritsch, J. Leissner, R. Davies, F. Remondino, R. Caffo (Ed.): EuroMed 2012, Progress In Cultural Heritage Preservation, LNCS 7616, Springer-Verlag Berlin Heidelberg 2012, 40-54

Santagati, C. 2006. "L'azzurro del cielo". Un polo museale tra arte, architettura, natura nel cimitero di Catania, Documenti DAU n. 31, Edizioni Caracol, Palermo.

Snavely, N. (et al) 2006. Photo tourism: Exploring photo collections in 3D. ACM Transactions on Graphics, 25(3):835–846, 2006.

Snavely, N., Seitz, S.M., Szeliski R., 2008. Modeling the World from Internet Photo Collections. International Journal of Computer Vision, 80(2), pp. 189-210

Vergauwen, M., Van Gool, L., 2006. Web-based 3D Reconstruction Service.

C. Santagati and L. Inzerillo

Machine Vision and Applications 17, pp. 411–426.

Vu, H-H. (et al.) 2009. Towards high-resolution large-scale multi-view stereo. Conference on Computer Vision and Pattern Recognition (CVPR), Miami, Jun 2009.

Wang Y F. A Comparison Study of Five 3D Modelling Systems Based on the SfM Principles. Technical Report, Visualsize Inc. TR 2011-01, Sept 8: 1-30

Wu, C., VisualSFM : A Visual Structure from Motion System http://homes.cs.washington.edu/~ccwu/vsfm/ (2011)

Wu C., Agarwal, S., Curless, B., Seitz, S. M., 2011. Multicore Bundle Adjustment. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2011.

Wu, C., SiftGPU: A GPU Implementation of Scale Invariant Feature Transform (SIFT) http://www.cs.unc.edu/~ccwu/siftgpu/ 10/04/2012

6. Acknowledgements

C. Santagati [a], L. Inzerillo [b]

[a]DARCH, Civil Engineering Faculty, 95125, Catania, Italy
cettina.santagati@dau.unict.it

[b]DARCH, Civil Engineering Faculty, 90100 Palermo, Italy
laura.inzerillo@unipa.it