INTEGRATED SFM TECHNIQUES USING DATA SET FROM GOOGLE EARTH 3D MODEL AND FROM STREET LEVEL

L. Inzerillo [a]*

[a] DARCH, Department of Architecture, Polytechnic School, 90128 Palermo, Italy - laura.inzerillo@unipa.it

KEY WORDS: SfM, Image Based Modelling, Photogrammetry, Google Earth, 3D model

ABSTRACT:

Structure from motion (SfM) represents a widespread photogrammetric method that uses the photogrammetric rules to carry out a 3D model from a photo data set collection. Some complex ancient buildings, such as Cathedrals, or Theatres, or Castles, etc. need to implement the data set (realized from street level) with the UAV one in order to have the 3D roof reconstruction. Nevertheless, the use of UAV is strong limited from the government rules.
In these last years, Google Earth (GE) has been enriched with the 3D models of the earth sites. For this reason, it seemed convenient to start to test the potentiality offered by GE in order to extract from it a data set that replace the UAV function, to close the aerial building data set, using screen images of high resolution 3D models. Users can take unlimited "aerial photos" of a scene while flying around in GE at any viewing angle and altitude. The challenge is to verify the metric reliability of the SfM model carried out with an integrated data set (the one from street level and the one from GE) aimed at replace the UAV use in urban contest. This model is called integrated GE SfM model (i-GESfM). In this paper will be present a case study: the Cathedral of Palermo.

1. INTRODUCTION

1.1 Introduction

Structure from motion (SfM) represents a widespread photogrammetric method that uses the photogrammetric rules to carry out a 3D model from a photo data set collection (Inzerillo, et al, 2016). Thanks to SfM using, everyone can make with a digital camera a 3D model applied to an object of both Cultural Heritage, and physically Environment, and work arts, etc. (Kersten, et al. 2012; Smith, et al. 2015; Verhoeven, 2011; Westoby, et al. 2012)
However, measurements derived using SfM inherently contain no scale or unit information since source images also lack that information (Falkingham, et al. 2014). Properly scaling SfM measurements requires additional measurement information such as georeferenced camera locations or measured distances between fixed ground control points (GCPs).
The SfM programs semi-automate the SfM process by allowing users to simply upload photos and process them through a software workflow, which shields users from the complex inner workings of SfM. The accessibility of SfM software can be especially advantageous to users in non-technical fields or to those with limited resources.
Nevertheless, some complex ancient buildings, such as Cathedrals, or Theatres, or Castles, etc. need to implement the data set (realized from street level) with the UAV one in order to have the roof reconstruction. Nevertheless, the use of UAV is strong limited from the government rules. (Fig. 1)
In these last years, Google Earth (GE) has been enriched with the 3D models of the earth sites. For this reason, J. Chen and K.C. Clarke (Chen, J, et al. 2016) start to test the potentiality offered by GE. Users can take unlimited "aerial photos" of a scene while flying around in GE at any viewing angle and altitude.

Figure 1. UAV and rules for flight in urban constest

However, the GE models are not accompanied with measures of reliability, so errors and uncertainties in GE models are propagate to the SfM ones. Furthermore, the GE models affected by the images quality and by the operator skill in performing each step of the SfM workflow and this unavoidable affects the GE SfM model. Therefore, the SfM 3D models, with meshing and texture mapping, carried out from GE data set, appear nearly identical to the GE models, with the same imperfections too. In 2016, J. Chen and K.C. Clarke (Chen, J, et al, 2016) demonstrated that the SfM model extract from GE data set, showed a mixed variability in the vertical measurements when compared with the LiDAR data.
Starting from these results, it seems appropriate to study the potentiality of a different use of the Google Earth models, in order to extract from it a data set that replace the UAV function, to close the aerial building data set, using screen images of high resolution 3D models. The challenge is to verify the metric reliability of the SfM model carried out with an integrated data set (the one from street level and the one from GE) aimed at

* Corresponding author

replace the UAV use in urban contest. This model is called integrated GE SfM model (i-GESfM). (Fig. 2)
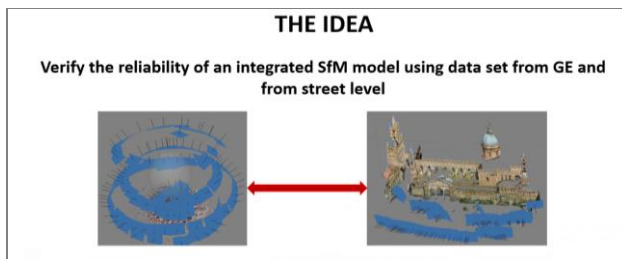

Figure 2. The idea: verify the reliability of the integrated project

In this paper will be present a case study: the Cathedral of Palermo. (Fig. 3)
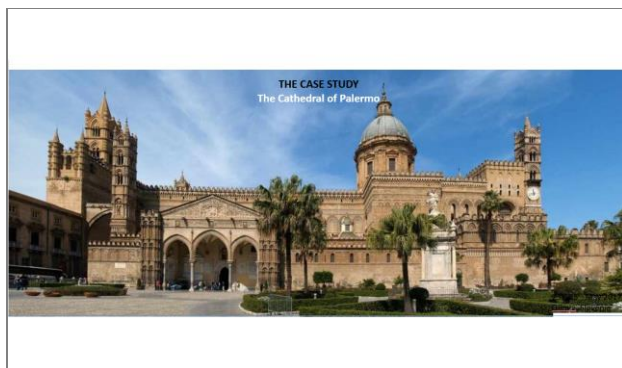

Figure 3. The case study: the Cathedral of Palermo

Simulated flights were done in GE, trying to have more different oblique grades and levels of the ground. The GE video acquisition was performed using the Microsoft Expression 4 Screen Capture Software on a NVIDIA Quadro K1100M monitor, which delivered video images at a range of 3 to 4 magapixel (MP) resolution. This data set was been integrated with the other one made from ground floor with a Nikon 5200 at 24 MP. The SfM elaboration was made using Adobe Photoscan. To use two data set, with so big resolution distance, it is necessary to create two different chunks in Photoscan and align them with markers.

The i-GESfM model was compared with the LIDAR one to verify the metric accuracy.

1.2 State of the art

Nowadays, the only one study conducted on the use of the Google Earth model is the one of Jorge Chen and Keith C. Clarke (Chen, J, et al, 2016) that studied the reliability of the SfM extracted from Google Earth model. They examined three study locations included a residential neighbourhood in Tokyo, Japan. Results of this study support the feasibility of using screen images of Google Earth for SfM modeling. While SfM succeeded in creating models for all three locations that visually resembled Google Earth's own models, quantitative analysis showed that SfM worked best in the built-up areas of Tokyo and UCSB but struggled with the natural environment of Mount Herard. Comparison of sample distances within the SfM models and Google Earth showed planimetric errors of 1% or less and vertical errors of 5% or less for Tokyo and UCSB; however,

absolute errors at Mount Herard -which was compared to LiDAR instead of Google Earth- spanned a range of under 10 m for areas of high relief to values exceeding 100 m for areas with low relief or low texture. The varying qualities of these models reflected not so much limitations of SfM but its reliance on a number of factors that impacted final model quality, such as image quality and operator skill in performing each step of the SfM workflow.

2. METHODOLOGY

2.1 Workflow

The developed methodology involves the realization of a project in Photoscan with two, or more than two, different Chunks. The number of the chunks is closely related to the number of details that you want to include inside the model carried out from GE. (Fig. 4)
You need to know how many Chunks will create at the starting of the project. Despite you can insert a Chunk in any moment, it is better to do it at the starting because you should have some alignment errors.
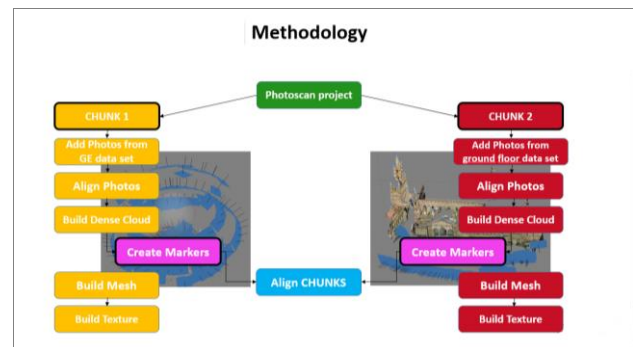

Figure 4. The Photoscan project with two chunks

In this case, we have just two Chunks. The PhotoScan process can go ahead with the photo alignment and dense cloud building phases, once uploaded the photos data set.
At the end of the dense cloud, it is necessary to detect the markers in the photos data set collection for each Chunk. Obviously, the markers must have the same insertion sequence. This is a strong obligation to guarantee the chunks alignment.
The order of the sequence must respect the following criterion: first, you create the markers on the detail chunks and then, create the others by the way until you make a scattered mapping on the whole model. There is no study that established how many markers is better to detect sure not less than three. A statistical study should be carried out, in order to determine the optimal markers to detect depending on the size of the façade or the building.
We can align the Chunks and ending the Photoscan process, once detected the markers.
At last, in order to verify the reliability of the ending model, we compare the iGESfM with the Lidar one.
Let's see this work flow applied to the Cathedral of Palermo.

2.2 Build Google Earth Chunk: Chunk 1

The photos data set from GE was done respecting the normal photogrammetry rules (Mikhail, Bethel, & McGlone, 2001). Users can take unlimited "aerial photos" of a scene while

"flying" around in Google Earth at any viewing angle and altitude. The screenshots were optimized using the Microsoft Expression 4 Screen Capture Software on a NVIDIA Quadro K1100M (2560x1440 pixel) monitor, which delivered video images at an effective 3.2 megapixel (MP) resolution (2464x1312 pixel) after accounting for the removal of borders and non-image elements, such as the Google logo, the toolbar, etc.

In our case, for the Cathedral of Palermo, it was made a photos data set with more than two laps. The initial image alignment step, placed some or all images into their correct locations and generated a sparse point cloud of tie points that vaguely resembled the modeled scene.

If you look closely at the model, you notice that in the low area there are some deformations, particularly marked in heights. it will not possible to delete these deformations from the GE model, but it will be possible to correct them. The GE model will reproduce all the errors that the Model in Google Earth presents and, sometimes, will emphasizes them.

180 photos in four laps at different heights compose the data set. (Fig. 5)

Each photo must be adjusted and trimmed in order to erase the google and toolbar details to avoid to have them on the texture of the 3D model. (Fig. 6)

Once completed the Alignment and the Dense Cloud, you proceed to upload the other data set in chunk number 2. In our case, the second Chunk is the one of the southern façade, made from the street level. As said before, you can create all the chunks you need with the data set related to an architectonical detail, or a façade, or a bell tower, etc.

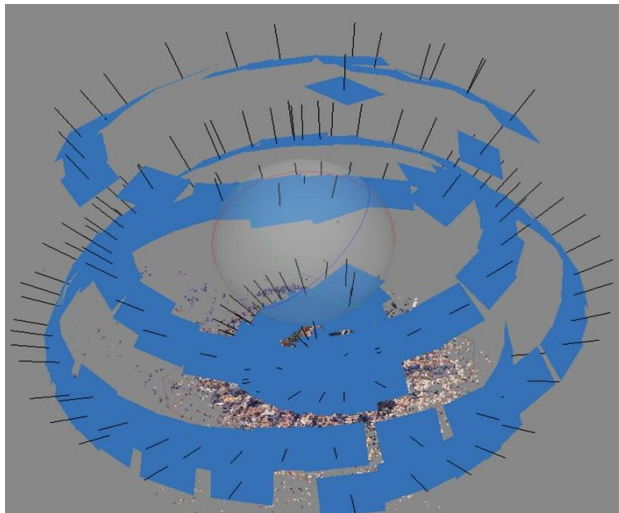In Table 7, there are the SfM reconstruction features. (Tab. 7)

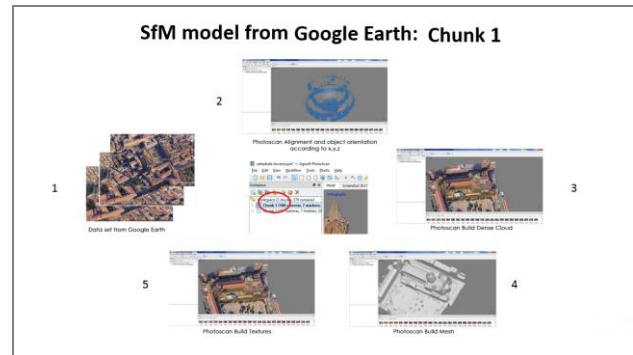Figure 6. SfM process for GE model: chunk 1

GE SfM Model	
camera	/
# photos	180
Sensor Width (Sw)	/
Focal lent	/
Image size in pixel	~2464 x 1312
Dataset time processing	345435
Photoscan time processing	345435

Table 7. GE SfM Model features

2.3 Build street level Chunk: chunk 2

A very rich southern face, that is over 100 meters long, characterizes the cathedral of Palermo. This façade was chosen for the second Chunk. The data set was made at nine in the morning to avoid both the sun on the façade and a big number of tourist in the later hours. Furthermore, there are a lot of natural and architectonic hurdles, hedge, etc. However, the photogrammetry rules are always the same. (Falkingham, P. L., et. al 2014). (Fig. 8)

It was used a Nikon D5200, 18 /55 mm and was made 191 photos with 6000 x 4000 image size in pixel.

The same procedure used for the GE model, is used in this case, to carry out the Level Street model. (Fig. 9)

From the Table 10, it is apparent that the data provided by a SfM made with a camera are much more complete than that achieved through screenshots. In this case, it is possible to carry out all the parameters for the reliability and the metric control.
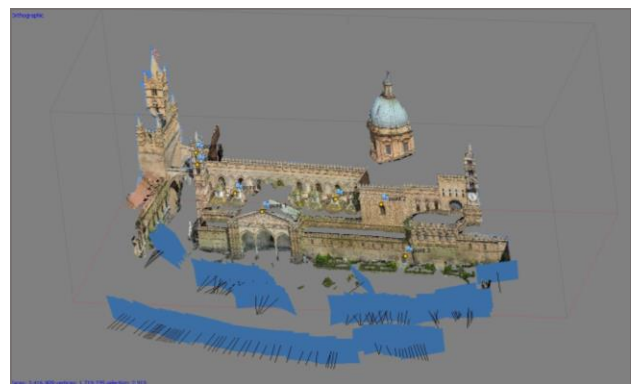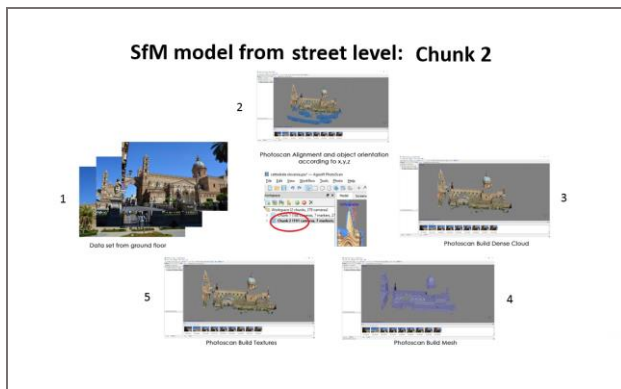

Figure 5. Data set for GE model: chunk 1

Figure 8. Data set for Street Level model: chunk 2

Figure 9. SfM model from street level: chunk 2

Street Level SfM Model	
camera	Nikon D 5200
# photos	191
Sensor Width (Sw)	23.5 x 15.6 mm CMOS sensor
Focal lent	18 mm
Image size in pixel	6000 x 4000
Dataset time processing	3234
Photoscan time processing	23423

Table 10. Street Level SfM Model features

2.4 Detect markers in both chunks

To detect the markers is necessary to study both data set for each chunk. In fact, the markers are not randomly selected: each marker must be present in at least three photos for each chunk.
The point chosen for the marker must be clearly identifiable in each of the three photos for each chunk. (Fig. 11, 12)
Besides, you must enter the markers starting from the details chunk and then switch to the others. This condition is very restrictive: the marker are detected with a number progression. So, if you start from marker number 1 on a point that does not exist in all chunks, during the alignment phase, the software would link different points to the model but with the same numbering.

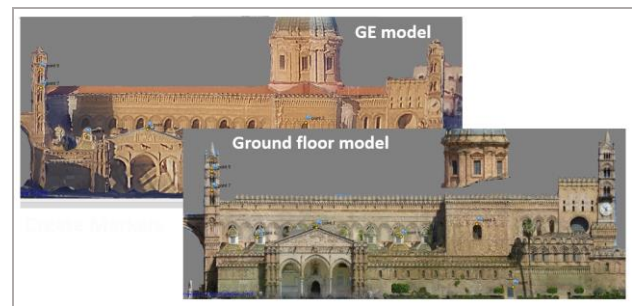Figure 11. Detect Markers in Chunk 1 and Chunk 2

Figure 12. Detect Markers in Chunk 1 and Chunk 2

In this case study, with 2 chunks, 12 markers were chosen on the south façade.

2.5 Align chunks and complete the PhotoScan process

Once you detect markers, you need to align the chunks and complete each Photoscan processing for each chunk.
Once aligned, the model is scaled and both views of the two models fit perfectly. This means that the alignment has been successful and the markers will perfectly coincide in the aligned model. (Fig. 13)

Figure 13. The aligned model

3. GROUND SAMPLING DISTANCE

The Ground Sampling Distance (GSD) is the distance between two consecutive pixel centres, measured on the ground. The bigger the value of the image GSD, the lower the spatial resolution of the image and the less visible details. The GSD is related to the flight height: the higher the altitude of the flight, the bigger the GSD value.
For example:
•A GSD of 5 cm means that one pixel in the image represents linearly 5 cm on the ground (5*5 = 25 square centimeters).
•A GSD of 10 m means that one pixel in the image represents linearly 10 m on the ground (10*10 = 100 square meters).
Even when flying at a constant height, the images of a project may not have the same GSD. This is due to terrain elevation differences and changes in the angle of the camera while shooting. Since the orthomosaic is created using the 3D point cloud and the camera positions, an average GSD will be computed and used. (Fig. 14)
It is usually recommended to process images captured at the same flight height, as they have the same Ground Sampling Distance (GSD). It means that all images will have the same level of details. This facilitates the matching of keypoints between images and therefore, helps the reconstruction.
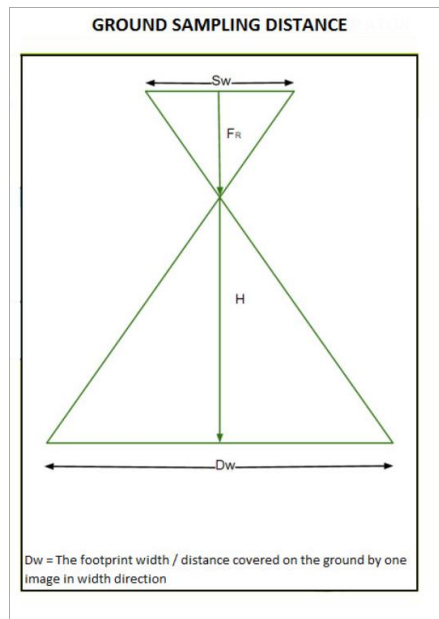
Figure 14. Ground Sampling Distance

However, Pix4Dmapper can also process images with different GSD as long as the highest GSD is smaller than twice the lowest GSD. Considering that the flight height and the GSD have a linear relationship, for the same project, captured with the same camera, the highest flight height at which images are taken should not exceed two times the lowest flight height:

$$GSD1 \leq 2 * GSD2$$

$$(Sw * H1 * 100) / (Fr * Iw) \leq 2 * (Sw * H2 * 100) / (Fr * Iw)$$

$$H1 \leq 2 * H2$$

Where:
GSD = Ground Sampling Distance [cm/pixel].
Sw = sensor width [mm].
H = flight height [m].
Fr = real focal length [mm].
Iw = image width [pixel].
In the case study of the Cathedral of Palermo we have a relevant slope along the southern front. Moreover, as seen in the figure 5, the data set from GE was made at different heights and there isn't the value of Sw because the data set was made by screenshots.
As said before, with Pix4D is possible to determine the GSD of a data set at different heights. Nevertheless, in our case study there are chunks that have data set made with screenshot and data set made with Nikon camera.

GSD for Ge model is not possible to calculate due the no value of Sw

GSD model for Level Street model: 1.30 m

4. RESULTS

We have to compare the carried out model with Lidar one in order to verify the reliability of the ending model. Considering the results obtained by Chen (Chen, J, et al, 2016), verifying the

stability of the heights of the model obtained is of paramount importance.
The comparison was made between the orthophoto carried out from the LiDAR and the SfM models
The first comparison is the one between the GE model and LiDAR. The second one is between the aligned model and LiDAR.
Three points to compare have been chosen on the two orthophoto: two points on the bell tower, and one on the highest point of the portico. Furthermore, In addition, a line was drawn at the height of the central merlot.
From the figure 15, you can see that some points correspond, others are completely misaligned. Furthermore, there is no coherent relationship: some are lower, others more right, others more left. Others are perfectly matching. (Fig. 15)
The point on the bell tower on the left side is lower than the right position. The point on the bell tower on the right side, is lower too but the bell tower is inclined on the left. The point on the porticato is lower and on the left.
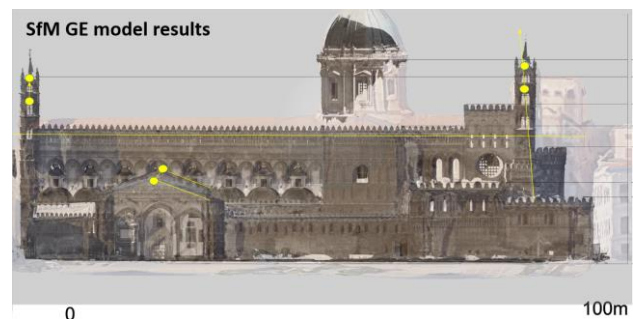

Figure 15. The aligned model

The result is disappointing, the intentions of using models from GE is to be abandoned.
This result would be expected, in fact it confirms the results found by Chen.
Go ahead and make the comparison between the LiDAR and aligned model. (Fig. 16)
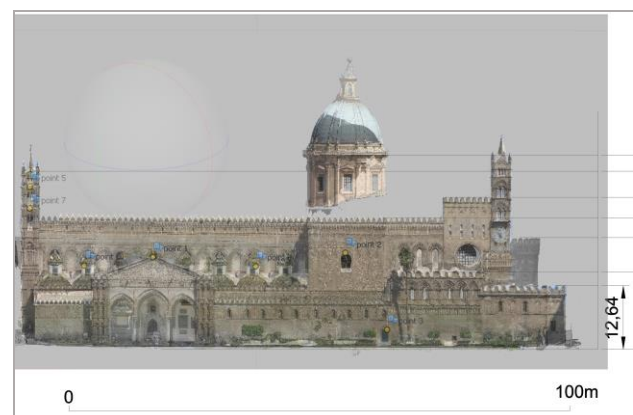

Figure 16. The aligned model

In order to verify the reliability of the aligned model, all the markers were considered for the comparison.
You can see that all points match perfectly.

5. CONCLUSIONS

The results obtained are astonishing and exceed the expectations of the goals. The GE model was corrected in the alignment phase: it has adapted to the model carried out from the street level, and its deformations have been corrected.

The result is even more surprising when you consider that the Google Earth starting model brought many inaccuracies. Furthermore,

If you implement the project with the other chunks, it generates a model with multiple insights. (Fig, 17, 18)

You might also experiment with creating an hypertext where on the Ge model you can edit the links from which the SfM model opens with architectural details.
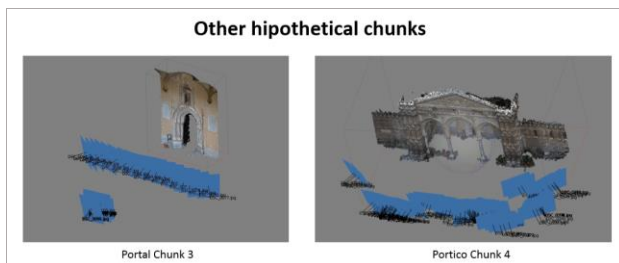

Figure 17. Data set for Portal and Portico

In this way it will be possible to create a 3D model of a complex monument like this, without sacrificing the details of rich architectural portions of decorations and with complex geometries.
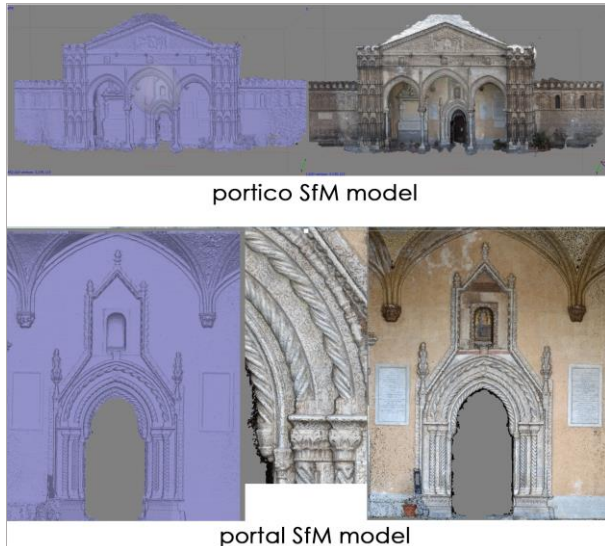

Figure 18. SfM model of the Portico and Portal

Nevertheless, there are very significant limitations concerning the impossibility of determining:

- the GSD value for the entire model;
- the quality standardization of the model obtained;
- the GE model may not be up to date with any restoration or collapses.

As regards the GSD value, we said before that it is not possible to calculate due the absence of Sw parameter of the GE model.

The model obtained is the result of the Google Earth model of the starting. Some are very accurate and well-defined, others are not accurate and often deformed if you look them closely. This means that there is no quality uniformity in the starting model and, so, there will not a quality uniformity of the aligned model. For example if we consider the Google Earth model of Basilica of Saint Peter in Rome and the Cathedal of Palermo, you can see the difference of the 3D model quality. (Fig. 19)

At last, it may happened that the Google Earth models aren't up to date. In fact, the building may be restored or collapsed and the model on Google Earth is not yet up to date.

We can conclude that this experiment opens up new horizons of research and pushes further insights including the opportunity to insert some Augmented Reality objects (Inzerillo, 2013) to complete a high accuracy model.
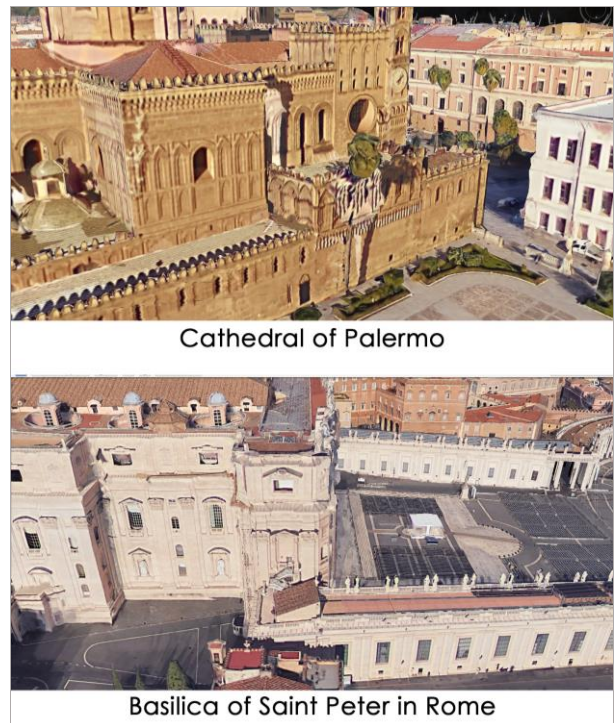

Figure 19. Model from Google Earth of the Cathedral of Palermo and Basilica of Saint Peter in Rome

REFERENCES

Chen, J., Clarke, K. C. 2016, Rapid 3D Modeling Using Photogrammetry Applied to Google Earth. In: Proceedings, AutoCarto2016. *The 19th International Research Symposium on Computer-based Cartography Albuquerque*, New Mexico, USA. September 14-16, 2016. Edited by S.M. Freundschuh. Online Proceedings published by the Cartography and Geographic Information Society.

Falkingham, P. L., Bates, K. T., & Farlow, J. O. 2014. Historical Photogrammetry: Bird's Paluxy River Dinosaur Chase Sequence Digitally Reconstructed as It Was prior to Excavation 70 Years Ago. PLOS ONE, 9(4), e93247. http://doi.org/10.1371/journal.pone.0093247

Kersten, T.P., Lindstaedt, M. 2012. Potential of Automatic 3D object reconstruction from multiple Images for applications. In

Architecture, Cultural Heritage and Archaeology. International Journal of Heritage in the Digital Era1.3. pp. 399–420.

Kyriakaki, G., A.Doulamis,N. Doulamis,M. Ioannides, K. Makantasis, E.Protopapadakis, A.Hadjiprocopis, K. Wenzel, D.Fritsch, M.Klein, and G.Weinlinger. 2014. 4D Reconstruction of Tangible Cultural Heritage Objects from Web-Retrieved Images. *International Journal of Heritage in the Digital Era3.2* (2014):431–452. doi:10.4399/97888548993917

Inzerillo, L. 2013. Augmented reality. Past, present, future. In Proceedings of *SPIE- The International Society for Optical Engineering*, 8649, 86490 E

Inzerillo, L., Santagati, C., 2016. Crowdsourcing cultural heritage: From 3D modeling to the engagement of young generations. In: *6th International Euro-Mediterranean Conference on Digital Heritage, EuroMed 2016*; Part I. Lecture Notes In Computer Science, vol. 10058, p. 869-879, Cham: Springer International Publishing AG, ISBN: 978-3-319-48496-9, ISSN: 0302-9743, doi: 10.1007/978-3-319-48496-9_70 Nicosia, Cyprus, Greece, 31 October 2016 through 5 November

Mikhail, E. M., Bethel, J. S., & McGlone, J. C. 2001. Introduction to modern photogrammetry. New York: Chichester: Wiley.

Moussa, W.,M. Abdel-Wahab, and D. Fritsch. 2012. An Automatic Procedure for Combining Digital Images and Laser Scanner Data. In The *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. vol. XXXIXpartB5 (2012):229–234.

Smith, M. W., Carrivick, J. L., & Quincey, D. J. 2015. Structure from motion photogrammetry in physical geography. Progress in *Physical Geography*, 309133315615805.

Stathopoilou, E.K.,Georgopoulos, A., Panagiotopoulis, G., Kaliampakos, D., 2015. Crowdsourcing lost cultural heritage. In *ISPR Annals of Photogrammetry Remote Sens. Spat. Inf. Sci*. II-5/W3, 295-300.

Verhoeven, G. 2011. Taking computer vision aloft archaeological three-dimensional reconstructions from aerial photographs with photoscan. Archaeological Prospection, 18(1), 67–73. http://doi.org/10.1002/arp.399

Vincent, M. L., Coughenour,C., Remondino, F., Gutierrez, M.F., Bendicho, L., Menchero, Manuel, V., Frtisch, V. 2015. Crowd-sourcing the 3D digital reconstructions of lost cultural heritage. In: *Digital Heritage*. 171–172.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. 2012. "Structure-from-Motion" photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300–314. http://doi.org/10.1016/j.geomorph.2012.08.021