# Comparative analysis among photogrammetric 3D models RAW data vs RGB images

Alessio Cardaci[1(0000-0001-7808-517X)], Pietro Azzola[1],
Michele Bianchessi[1], Ruggero Folli[1] e Simone Rapelli[1]

[1] Department of Engineering and Applied Sciences, University of Bergamo
alessio.cardaci@unibg.it; pietro.azzola@unibg.it; m.bianchessi@studenti.unibg.it; r.folli@studenti.unibg.it; s.rapelli@studenti.unibg.it

**Abstract:** The military architecture of the ancient Republic of Venice is a very significant heritage. These are structures with simple volumes in the building and functional in the distribution; they are often little considered because, wrongly, they are considered minor. The cultural capital that must be documented and digitized to preserve, improve and preserve. The case study of the "Torresino da Polvere" in via Beltrami, in the upper town of Bergamo, was used as a pretext for a comparative analysis between the 3D models. Models generated by the direct processing of data from photographic sensors and those obtained from image processing after conventional raw conversion processes. The research, based on rigorous experimentation, proposes a new method of capture and frame management in order to obtain high quality models directly from the raw data. The results allowed some interesting considerations; they highlighted the singular peculiarities of some software in the management of raw data, both for a higher accuracy of the returns, and for an unexpected speed of calculation.

**Keywords:** 3D integrated survey, Torresini da polvere, Bergamo.

## 1 Introduction

The format commonly referred to as raw contains information related to the light intensity measured by the photometric sensor, a 'digital negative' without modification or alteration. Raw data are sequences of bits stored in proprietary files according to specific codes established by the different constructors and, among them, incompatible with each other. Digital manufacturers, for purely business reasons, were less likely to adopt a universal standard, as demonstrated by the limited success of the DNG format [1, 2]. A format developed by Adobe®, patented, open and without loss and may be boycotted by major brands like Canon®, Nikon® and Sony®.
A photograph, even of the same scene and under the same lighting conditions, will therefore be different if captured with different cameras, both because it is dependent on

the different spectral sensitivities of the sensors, and on the different 'preservation' of the information [3]. In addition, the raw data does not represent a structured and visible image thanks to coloured pixels but, instead, a list of values recorded by the photosensitive detectors by means of a filter matrix - reactive to the frequencies of red, green and blue - arranged according to the well-known and usual Bayer Array [4].

The creation of a realistic representation is the result of a conversion performed through raw conversion algorithms which transform the luminosity values of each photoreceptor. Transformation deals sequentially with linear values, white balance, demosaification, colour space conversion and gamma correction in an orderly array of RGB data [5].

The main function of raw-converter software is to 'estimate' the amount of light captured by the chamber and make it visible through colorimetric information; a process strongly conditioned both by the electronics of the chamber and by the choices of the operator [6].

The transformation of the raw data into dng, tif or jpg codes is therefore an action that involves both an alteration of the information (caused by the interpretation of the value) and a loss of part of the contents (as a result of compression and/or reduction from bit depth, generally from 14 to 8, as well as noise reduction and increased contrast and sharpness). A process that in the literature is evaluated as an added value as it allows to improve the quality of photographs, both for the correction of errors made during capture (adjustment of colour shades, increase of contrasts, filtering of background noise, ...), and for the possibility of adding multiple shots with merging and/or HDR elaborations [7].

The proposed research, taking as a pretext the case study of the 'Torresino da Polvere' in the upper city of Bergamo [8], wanted to conduct a comparative study among the 3D photogrammetric models extracted from the direct processing of the photographic sensor data with those, instead, obtained from the processing of RGB images after the traditional raw-converter treatment.

The work is part of a larger search for the digital documentation of military architectures of the former Republic of Venice (Fig. 1); an extensive patrimony of inestimable value made up of structures with simple volumes in construction and functional in distribution and, perhaps for this reason, little considered because poorly considered as minor.

## 2 Laser scanner survey and topography network

The 3D survey of the 'Torresino da Polvere' of via Beltrami, a small architecture with regular geometry and linked to the Venetian Walls of the city, was therefore the necessary study frame for experimentation (Fig. 2).

The strong presence of vegetation on the roof and the partial burial of two of its sides, did not allow the return of the integral model of the building but only part of it; wall sections, however, more than satisfactory for punctual and reliable reflections.

Measurements were made using multiple instruments and different acquisition methodologies, a well-established practice in the new integrated investigation procedures [9]. A topographic network has been set up whose vertex coordinates, materialized with

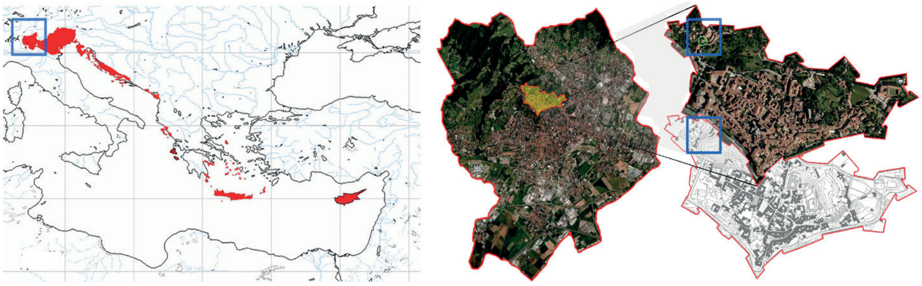


**Fig. 1.** La The fortress of Bergamo in the context of the ancient Republic of Venice of the 16th century and the 'Torresino da Polvere' in the upper city, within the system of the Venetian Walls.

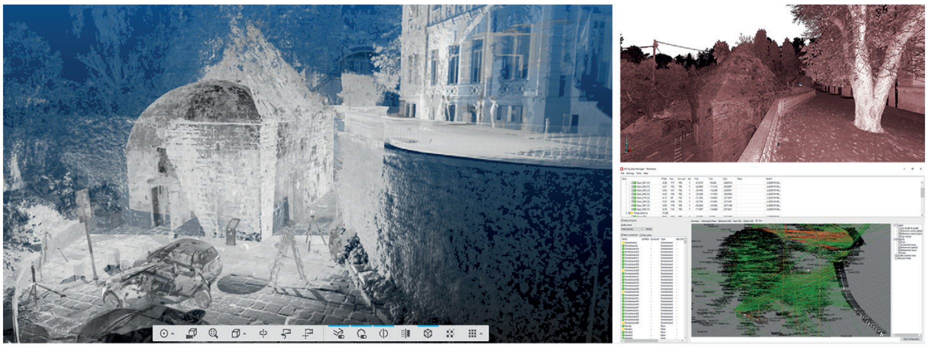
prisms placed on fixed tripods, have been determined with total station and operating a forced centring, a method that allows the interchangeability between signal and instrument has allowed an indeterminacy of less than one millimetre.

The resolution of the hyperdetermined equations of the network (for the surplus of both angular and distance measurements) was carried out through minimum square compensation software.

In this case the well-known commercial product MicroSurvey® Star*Net was used because it was able to combine ease of use and essential functions, with great reliability and high rigor [10].

From the top of the network, through forward intersections and distantly surveys, the coordinates of the centres of paper targets placed on the walls of the building were acquired; they, in large numbers and well-studied in geometry and spatial distribution, have formed the Ground Control Points (GCP) - twelve in total - of the georeferencing system of 3D laser scanner and photogrammetry virtual *maquettes*.

**Fig. 2.** The 'Torresino da Polvere' in via Beltrami in Bergamo; on the left the preview of the 3D model (projection of the photogrammetric model on the 3D laser scanner point cloud); on the right of the photographic images of the architecture after the restoration work of the years 1980.



| Number of scans | Resolution | Sampling time speed | Average Error | Part of point < 4mm | Number of point |
|---|---|---|---|---|---|
| *24* | *7.67 mm/10m* | *122 kpt/s* | *2.528 mm* | *70.1 %* | *728.486k* |

**Fig 3.** The 3D laser scanner model used for comparison of photogrammetric models.

The topographic survey has been integrated with an active sensor survey in order to have a much higher number of cornerstones (twenty-eight in all) at known coordinates; the cloud of points obtained was, in fact, the metric reference for the comparison of the various models [11]. In order to optimize the time, only reflectance information was recorded and not the colorimetric data; the purpose of the experiment was in fact the verification of the accuracy of digital artefacts and not the study and understanding of colour.

The creation of a single global project cloud was carried out with an initial alignment of the individual catches on the targets presented earlier and, therefore, an approximation of the various portions by means of 'shape-recognition' algorithms.

The use of external references, such as spheres or chessboards, allowed for the stricter recording of scans as well as verification of inaccuracies based on differences between targets measured with the laser scanner and the position of the top of the topographic network [12, 13].

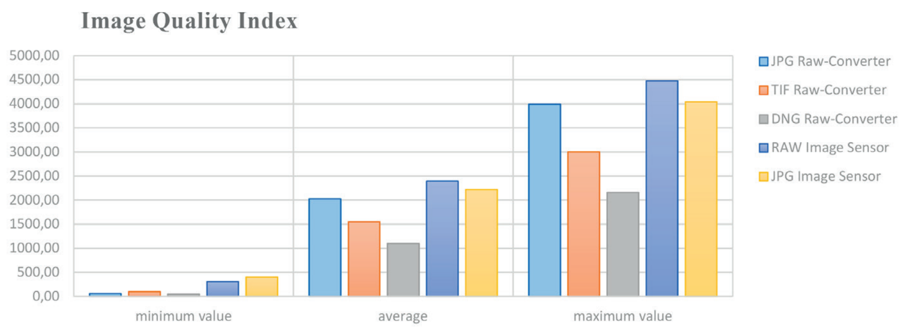
The recording phase was carried out through the software that owns the instrument, the Faro® Scene, and the verification of the accurateness was carried out through the application of the ATS® Quality Manager, able to perform the calculation of voltages, compensations and return the statistics of ground control points (GCP). The excellent overlap of the individual clouds was thus ascertained with an average approximation contained within the instrumental precision (± 1.5 mm).

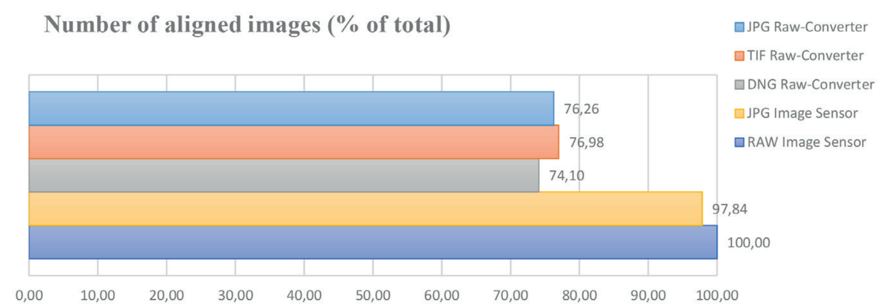
The individual scans, after being subjected to noise reduction and sampling to eliminate the over-posed areas (because they are very close to each other within an ideal sphere of one millimetre of radius), then returned a dense cloud that was used for comparison with the photogrammetric artefacts (Fig. 3).

The deviation between a 'certain' and reliable geometry (that of the 3D laser scanner) and the virtual model recreated on the basis of the different photogrammetric treatments, has therefore made it possible to identify the most accurate archetype.

## 3 Photogrammetry: processing raw data and RGB images

Photogrammetric 3D models were created by processing the same photo capture set. The experiment consisted of the creation of five virtual artefacts: two generated by directly processing the sensor data stored on the camera memory card (RAW image sensor and JPG image sensor), three RGB images created after the raw data processing (JPG Raw-Converter, TIF Raw-Converter and DNG Raw-Converter).
The conversion was carried out through the Adobe@ Camera Raw software by applying to all photographs of the group the same transformation parameters and, therefore, saving them in different formats: DNG, TIF and JPG. Graphical information from the RAW Image Sensor and transformed RGB images - JPG Raw-Converter, TIF



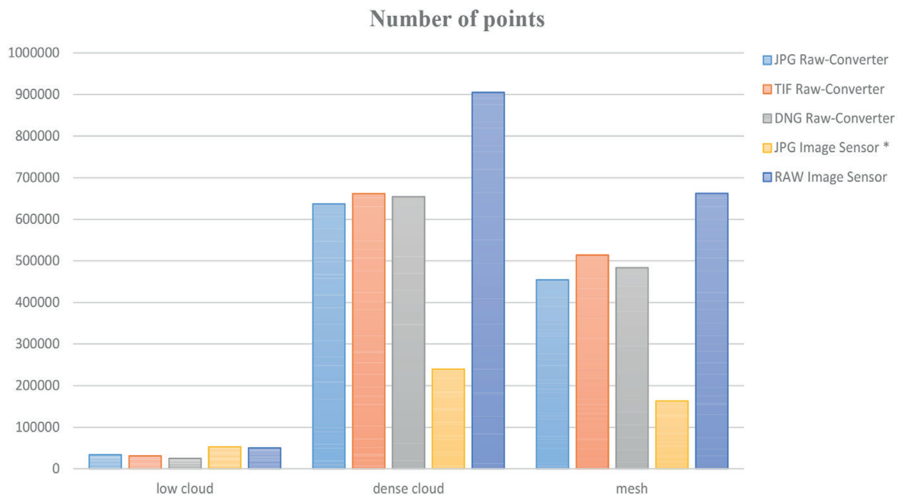



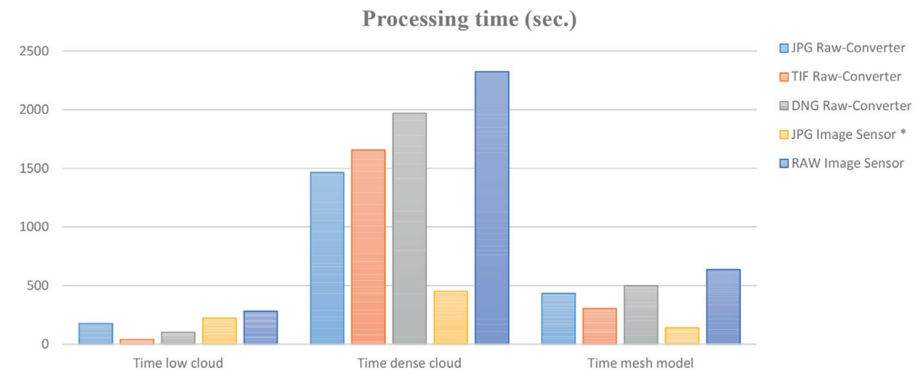

| Name | Minimum Quality | Average Quality | Maximum Quality | Image Aligned |
|---|---|---|---|---|
| *JPG Raw-Converter* | *60,00* | *2027,50* | *3995,00* | *76,26 %* |
| *TIF Raw-Converter* | *98,00* | *1550,00* | *3002,00* | *76,98 %* |
| *DNG Raw-Converter* | *48,00* | *1103,00* | *2158,00* | *74,10 %* |
| *JPG Image Sensor* | *402,00* | *2221,00* | *4040,00* | *97,84 %* |
| *RAW Image Sensor* | *309,00* | *2393,50* | *4478,00* | *100,00 %* |

**Tab 1.** The image quality index: comparative tables and graphs between the various formats.

Raw-Converter and DNG Raw-Converter - was captured/stored at the same resolution as about 21 Mega-Pixels (5616x3744), while JPG Image Sensor was slightly lower at 5.2 Mega-Pixel (2784x1856). It is because of the settings of the device used, a mark II EOS 5D Canon@ reflex with fixed lens Canon@ 24mm/1.4, put to get small files.
In all, a hundred shots were taken with an overlap between the images between 60% and 75%. The digital development of raw sensor information is recommended practice - sometimes hobbled because the raw data is not importable in many commercial applications - in the photographic field, both because it can correct any errors during capture (exposure, white balance, histogram ...), and because it allows to obtain more realistic 'personal' RGB images thanks to the manipulation of the colour temperature, light and shadow regulation, noise reduction and sharpness improvement [14].
The initial checks immediately showed that the processing of the raw converter seems to bring only an improvement in the visual quality, perceptible from the eye of the photographer but not from the 3D reconstruction software.
More accurate tests aimed at quality assessment have in fact shown that in the processed images the blurring and areas with uniform colouring textures has significantly increased, as shown by the tables and graphs reported (Tab. 1).
The quality index value was returned by specific multi-slate analysis processing algorithms with frequency decomposition [15], a more reliable solution than the simplest commonly used edge detection applications. A rigorous method aimed at knowing the overall characteristics of each individual image [16]. When returned to the table, the minimum value, average value, and maximum value of the entire set of 140 images were reported.
The best results, as you can guess, were those relating to direct raw catches; it surprised the comparison between the two groups in jpg with very similar parameters, as if to demonstrate that the raw-converter treatment - for this format - is independent of the conversion algorithm used. The deterioration of the quality index was also marked by the first steps with 3D Image Based Reconstruction software, in which a reduction of about 30% of the images used for the external positioning of the cameras was evident.
The greater number of non-aligned photographs, from careful observation of the models, was caused both by the smaller number of keypoints recognized by the software, and by the greater inaccuracy in the positioning of target centres; this feature, as will be emphasized below, will result in less accuracy of the model. In fact, the deviation between the coordinates of the references attributed by the software (for the lower sharpness and greater blurring of the edges) and those forced by the direct assignment of the values measured with the topographic network, is the cause of a greater error of RMS reprojection.
The experiment was based on the use of the latest version of the 3DFlow® Zephyr; in fact, it is one of the few commercial applications for photogrammetry that can import raw data. The tests were able to take advantage of the possibilities offered by the new Windows Imaging Component (WIC) Raw Format engine that can support most of the raw formats of large photographic equipment manufacturers.
Specifically, the software utilizes the potential of the codec pack developed by Microsoft® as part of the LibRaw project. It is a free and open source software library [17] that can process the raw files of many manufacturers while maintaining metadata,

bit-data depth (up to 32 bits per channel) and the high dynamic range of the sensor. The choice of a single application penalises the generalization of the problem as the results can be conditioned by the structure of the programme; aware of this limitation, the authors are conducting the same tests with multiple software (in order to be able to make general hypotheses not dependent on the applications used) but this, at the moment, is not related to the treatment of the essay because the findings are still unripe and deserve further study.
In principle, however, they appear to confirm the hypotheses advanced so far. The creation of the photogrammetric models followed the usual workflow distinguished, first of all, by a single choice of parameters for the various processing steps for all reconstructions. Standard values have been set of the variables that regulated the creation of the artefacts, avoiding extremes and particular situations that could have altered the generality of the results favouring one specific format over another (Fig. 4). The first phase saw the recognition by the software of particular points in the photographs and therefore the subsequent combination of them between the homologous frames; based on these 3DF® Zephyr indications continued by operating the spatial positioning of the chambers and the consequent generation of the scattered clouds.
The next steps were to build the dense cloud, convert the discontinuous point model to the continuous mesh model, and finally, reprojection on the faces of individual RGB image triangles (after accurate blending to ensure uniformity of exposure and colour). The georeferencing of the model and its 'scale' was carried out on the basis of the spatial coordinates of the Ground Control Points (GCP) which also fulfilled the role of Quality Control Points (QCP); as in the procedures concerning the data acquired by passive sensors, in fact, some cornerstones were used for the comparison of differences and differences between the coordinates of the models and those of the topographic network (Tab. 2, Chart 1).
In order to achieve a better result instead of proceeding with the processing of the entire set of images within a single environment - more burdensome in terms of time and hardware resources - partial processing of the images was chosen and then the models






| Name | Point Low Cloud | Time Low Cloud | Points dense cloud | Time Low Cloud | Mesh | Time Mesh |
|---|---|---|---|---|---|---|
| *JPG Raw-Converter* | *33483* | *177 sec* | *637266* | *1465 sec* | *454357* | *435 sec* |
| *TIF Raw-Converter* | *30898* | *41 sec* | *662138* | *1656 sec* | *514134* | *304 sec* |
| *DNG Raw-Converter* | *24840* | *101 sec* | *654083* | *1970 sec* | *483776* | *498 sec* |
| *JPG Image Sensor* | *53118* | *221 sec* | *239585* | *453 sec* | *163371* | *139 sec* |
| *RAW Image Sensor* | *50405* | *281 sec* | *904689* | *2324 sec* | *662435* | *637 sec* |

**Tab 2.** Quantitative comparison tables between the various photogrammetric models.

were brought together. In particular, the two façades were rebuilt separately, keeping the areas close to the edge and roof as common joining elements.
This artifice has also made it possible to obtain more points (both scattered and dense cloud) as well as smaller and more 'interpretative' meshes of the real geometry of the building, thanks to the double processing of common frames (Fig. 5).
The comparison between the models highlights, with absolute clarity, how the best reconstructions are those obtained from the raw data, both in terms of accuracy and in the percentage of parts returned. The reconstruction of the cover, from shots captured in poor conditions with the inclined and more distant chamber, is guaranteed by the

**Fig 4.** Selected parameters within the image set process.

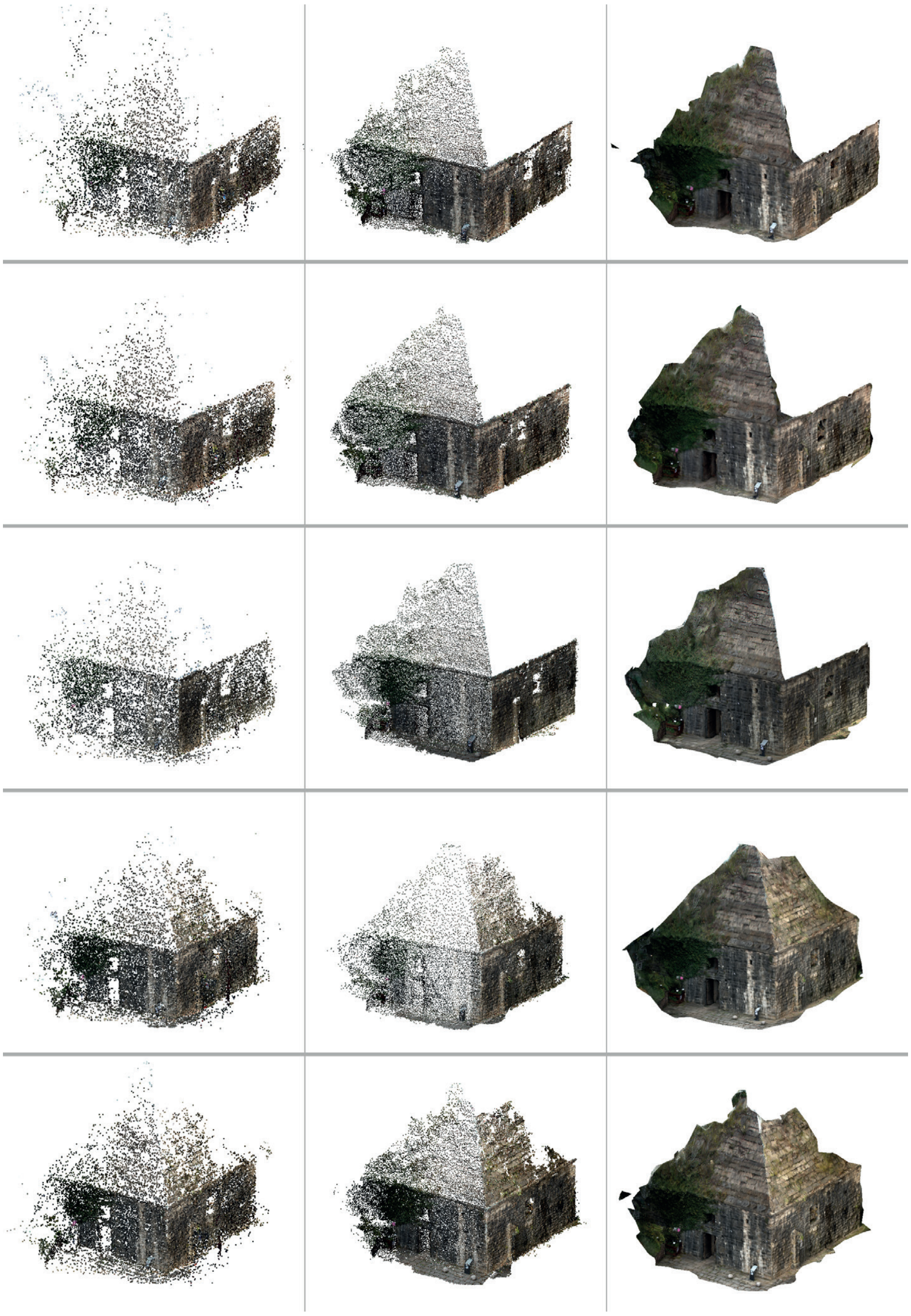


**Fig 5.** Model comparison: scattered point cloud, dense point cloud and mesh model (from left to right); retashing from the JPG Raw-Converter, JPG Raw-Converter, DNG Raw-Converter, JPG Image Sensor, and RAW Image Sensor image set (up to down).

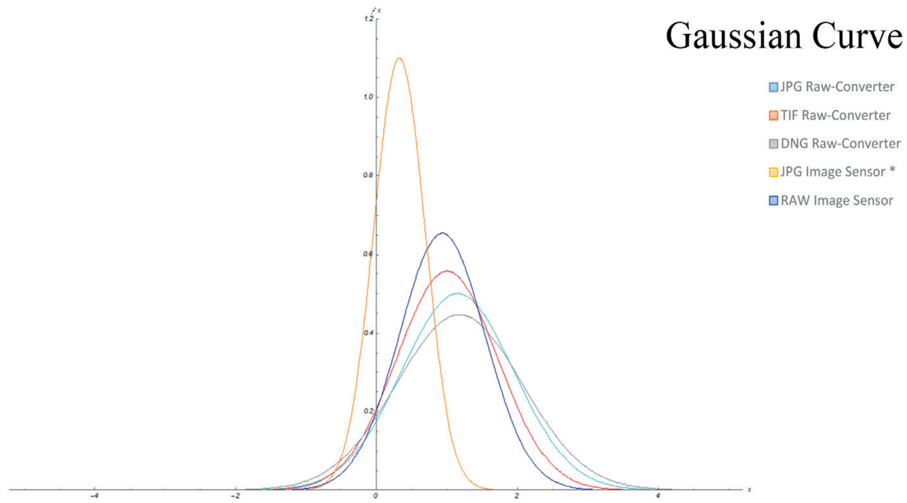



**Chart 1.** Comparing the accuracy between models of the RMS index.

sensor information (raw and jpg) and not possible with RGB images after raw-converter treatments. It is important to highlight the results obtained by the jpg image sensor format because it is the result of pictures with much lower resolution and high compression ratio.
This is actually the basis of shorter processing times and less points and triangles making up clouds and meshes. A model that, although less accurate (it must not mislead the graph because it is compared with groups of photographs of different resolutions) has nevertheless reconstructed the model in its entirety, including the roof (Fig. 6).

# 4 Conclusions

Image Based Reconstruction 3D photogrammetry is based on the use of digital photos but many software does not allow the direct processing of raw files but only that of RGB images.
The first consideration is the convenience of using raw data for the creation of virtual artefacts based on photographic captures, because they can combine high metric accuracy with processing times that are still contained. The lack of digital development of the shots, while allowing to reduce 'time and work', is a constraint for greater attention during the capture of the scene; in fact, it is not possible to correct any 'grip' errors because, to date, it is not allowed to make changes to raw files within the software. The external transformation of the raw data (not carried out by the chamber processor and/or with proprietary software) has instead shown a general 'loss' of information quality; the same raw jpg data saved directly from the camera allows the creation of models much more accurate than those obtained as a result of raw-converter transformations and saved in dng, tif and jpg formats.
This manifested itself with particular evidence in the very 'glimpsed' photographic

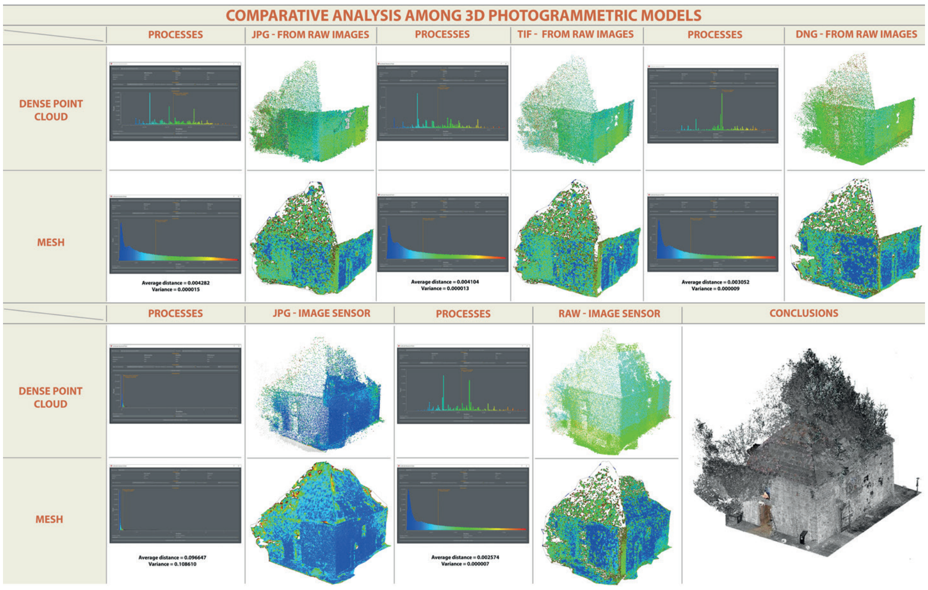



**Fig. 6.** Comparative analysis between photogrammetric models and the 3D laser scanner cloud.

sockets (such as the roof to the left of the roof photographed from below and up close by the presence of a wall) where the number of recognized keypoints was even insufficient to reconstruct the geometry.

Important reflection should be done for jpg images provided directly from the camera and produced with proprietary algorithms from different manufacturers. They are of far better quality than those generated by external software (ex: Abobe @CameraRaw); however, they have the disadvantage of preventing the photographer from intervening on the conversion parameters and, therefore, adjusting white balance, brightness, ..., because everything is decided by the machine.

## Acknowledgements

The study was conducted within a collaboration between the University of Bergamo and the Municipality of Bergamo, aimed at the knowledge, conservation and enhancement of the complexes of the Venetian Walls of Bergamo. The authors thank the 3DFlow team for valuable information and advice regarding the structure and optimal use of Zephyr software.

## Riferimenti bibliografici

1. Triantaphillidou S., Allen E.: Digital image file formats. In: Allen E., Triantaphillidou S. (eds): The manual of photography and digital imaging, pp. 315-328. Routledge, New York (2010).

2. Schewe J.: Il negativo digitale,: file RAW con Photoshop, Lightroom e Camera RAW. Apogeo, Milano (2014).
3. Nguyen R., Prasad D. K., Brown M. S.: Raw-to-raw: Mapping between image sensor color responses. In: Proceedings on IEEE Conference on Computer Vision and Pattern Recognition, pp. 3398-3405. Doi: 10.1109/CVPR.2014.434 (2014).
4. Fraser B., Understanding Digital Raw Capture. Adobe Systems Incorporated, San Jose (2004). Homepage: https://www.adobe.com/digitalimag/pdfs/understanding_digitalrawcapture.pdf
5. Rojtberg, P.: Processing RAW images in Python (2017). LNCS Homepage, https://www.researchgate.net/publication/314239357_Processing_RAW_images_in_Python
6. McHugh S. T.: Understanding Photography: master your digital camera and capture that perfect photo. No Starch Press, San Francisco (2018).
7. Verhoeven, G. J. J.: It's all about the format–unleashing the power of RAW aerial photography. International Journal of Remote Sensing, 31(8), 2009-2042, (2010).
8. Cappellini P.: Le polveriere venete. Tipografia Cesare Ferrari, Bergamo (1987).
9. Cardaci A., Versaci A.: Rilievo e restauro: un binomio imprescindibile: approcci metodologici e applicazioni operative finalizzate alla conoscenza e alla conservazione del patrimonio culturale. Aracne, Roma (2018).
10. Bonfanti C.: Topographic and photogrammetric approaches for the multiscale documentation. Parthica: incontri di culture nel mondo antico (17), 57-62, (2015).
11 Shan J., Toth C. K.: Topographic laser ranging and scanning: principles and processing. Routledge, New York (2018).
12. Liscio E., Hayden A., Moody J.: A comparison of the terrestrial laser scanner & total station for scene documentation. In Journal Association Crime Scene Reconstruction, (20), 1-8, (2016).
13. Cox R.: Real-world comparisons between target-based and targetless point-cloud registration in Faro Scene, Trimble RealWorks and Autodesk Recap. (2015). LNCS Homepage, https://eprints.usq.edu.au/29195/
14. Carpiceci M.: Fotografia digitale e architettura, elaborazioni con le odierne attrezzature fotografiche e informatiche. Aracne, Roma (2012).
15. Pirotti F., Vettore A., Guarnieri A.: Algoritmi di classificazione contextual con GRASS. Geomatics Workbooks, (6), 1-12, (2016).
16. Alfio V. S., Costantino D., Pepe M.: Influence of image TIFF format and JPEG compression level in the accuracy of the 3D model and quality of the orthophoto in UAV photogrammetry. Journal of Imaging, 6(5), 30, (2020).
17. AAVV.: LibRaw raw images decoder library: project Goals and Objectives. LNCS Homepage, http://www.libraw.org