ON THE EVALUATION OF PHOTOGRAMMETRIC METHODS FOR DENSE 3D SURFACE RECONSTRUCTION IN A METROLOGICAL CONTEXT

Isabella Toschi [a, b] *, Alessandro Capra [a], Livio De Luca [c], *, J.-Angelo Beraldin [d], Luc Cournoyer [d]

[a] DIEF – Dept. of Engineering "Enzo Ferrari", University of Modena and Reggio Emilia, Modena, Italy
[b] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
[c] CNRS, UMR 3495 Modèles et simulations pour l'Architecture et le Patrimoine (MAP), Marseille, France
[d] National Research Council, Measurement Science and Standards, Ottawa, Canada, K1A 0R6

Commission V, WG I

KEY WORDS: Metrology, dense stereo, laser scanners, photogrammetry

ABSTRACT:

This paper discusses a methodology to evaluate the accuracy of recently developed image-based 3D modelling techniques. So far, the emergence of these novel methods has not been supported by the definition of an internationally recognized standard which is fundamental for user confidence and market growth. In order to provide an element of reflection and solution to the different communities involved in 3D imaging, a promising approach is presented in this paper for the assessment of both metric quality and limitations of an open-source suite of tools (Apero/MicMac), developed for the extraction of dense 3D point clouds from a set of un-ordered 2D images. The proposed procedural workflow is performed within a metrological context, through inter-comparisons with 'reference' data acquired with two hemispherical laser scanners, one total station, and one laser tracker. The methodology is applied to two case studies, designed in order to analyse the software performances in dealing with both outdoor and environmentally controlled conditions, i.e. the main entrance of Cathédrale de la Major (Marseille, France) and a custom-made scene located at National Research Council of Canada 3D imaging Metrology Laboratory (Ottawa). Comparative data and accuracy evidence produced for both tests allow the study of some key factors affecting 3D model accuracy.

1. INTRODUCTION

1.1 Context

A number of recent publications summarize solutions for the automatic generation of textured dense 3D surface models from 2D images in single/multiple baseline arrangements. Both low cost software packages and open source solutions including web services have become very popular. While providing access to some satisfactory results from a visual point of view, these solutions have the major disadvantage of lacking clear and unambiguous metric results (Remondino et al., 2012). In the field of cultural heritage, the strong geometric and visual consistency of a 3D representation is an essential condition for the documentation and analysis of heritage artefacts. Within the framework of the (Culture 3D Clouds) project, the CNRS MAP laboratory is coordinating the development of image-based acquisition protocols for the accurate and effective 3D reconstruction of cultural objects.

As shown by El-Hakim et al. 2003, there is a need for accuracy evaluation tools for image-based 3D modelling. Key factors and critical configurations affecting 3D model accuracy are also needed. The authors propose a technique that creates simulated data based on the actual project data. The technique gives a valid and realistic measure of the accuracy, and has been applied on a wide variety of data to study the effect of various parameters and configurations. As a result, guidelines for some phases of 3D modelling from images are given. They focus on modelling relatively large structures like monuments and architectures for accurate documentation where knowledge of

uncertainty is important. According to the authors, in practice, it is difficult to achieve optimum network design. Therefore, the goal should be to strive for strong geometric configurations, high redundancy, high image resolutions on natural features and correct calibration. Wenzel et al. 2013 expand this work and the work by Waldhäusl and Ogleby 1994 by proposing a guideline for image data acquisition called "One panorama each step".

1.2 Literature review on approaches for inter-comparison

The European Spatial Data Research Organisation (EuroSDR) project aims at benchmarking image matching approaches for Digital Surface Models (DSM) computation from airborne imagery. A test bed is proposed to software developers, distributors and users of dense matching software in order to evaluate on a continuous basis image-based DSM approaches as the technology is improved. A framework is proposed for the comparison of results based on a common reference surface. Other well-known examples of benchmarks that aim at measuring the performance of different state-of-the-art matching algorithms can be found in (Scharstein and Szeliski, 2002; Seitz et al., 2006). Ahmadabadian et al. 2013 compare four independent dense matching packages for scaled surface reconstruction using stereo camera rigs without measuring any object distances. The tests aim at evaluating both the ability to resolve the scale and to assess the reliability in terms of accuracy. A number of test artefacts are used for the inter-comparison. It is noted by the authors that the laser scanner data is not an absolute reference and therefore the comparisons show relative errors only. Georgantas et al. 2012 present a comparison of an automatic photogrammetric technique based

---

*   isabella.toschi@unimore.it
*   livio.deluca@map.archi.fr

on MICMAC to a terrestrial laser scanning for 3D modelling of a building's stairway with a height of about 12 m. Along with the comparison with laser scanner data, system cost, acquisition and computational time are discussed. The authors conclude that their approach does not reach the geometric quality of a time-of-flight laser scanner, since image-based methods are heavily depended on the presence of texture (reflectance). The authors conclude that, even if the results of the image-based approach may be less accurate than the ones delivered by the range-based method, photogrammetry can be considered an interesting solution thanks to its scalability, low cost and onsite swiftness. Koutsoudis et al. 2013 use a similar methodology i.e. laser scanner data is used as a reference for comparison with dense stereo generated 3D data. Remondino et al., 2012 present a critical insight and a metric evaluation of automated image orientation packages. Different datasets are used in this evaluation. Large and complex scenes with known shapes, precise ground control points (GCP), calibrated cameras and reference scale bars were used in the evaluation of the different software packages. The conclusion is that all the packages evaluated deliver similar results in terms of theoretical precisions of the computed object coordinates and recovered interior camera parameters in the case of a robust image network. Kersten and Lindstaedt, 2012 present also an evaluation of low-cost image-based systems for automatic 3D recording and modelling of archaeological objects. Many scientific and technical communities have understood the importance to benchmark algorithms and methodologies used in image-based 3D systems. Many have also realized the challenges in generating accurate reference data. It is interesting to note that in many publications the metric quality of 3D reference data originate from active 3D image systems either triangulation or time-of-flight-based (Pears et al., 2012; Vosselman and Mass, 2010) and that 3D data is seldom questioned. The current work proposes a reflection on the topic of inter-comparison and shows some elements of solution to clarify the reality of the technologies used for inter-comparison.

1.3 MICMAC and associated tools

We use the suite of tools developed by the French mapping agency (IGN – *Institut Géographique National*) for our tests. This suite includes a number of tools. Among them, we find Apero and MicMac that represent the two main software solutions. The former computes the internal and external orientations of images (Pierrot-Deseilligny and Clery, 2011), whereas the latter performs the surface reconstruction phase by extracting depth maps from oriented images (Pierrot-Deseilligny and Paparoditis, 2006). From an algorithmic point of view, the IGN's suite of tools implements mathematical formulations derived from both the photogrammetric and the computer vision fields, focusing thereby on the accuracy and metric content of the final results as well as the automation of the image-based pipeline. The latter consists of three main consecutive phases. The first one is the tie point extraction step. The tool Tapioca, a Sift[++] implementation of SIFT algorithm (Vedaldi 2010), is provided to the user. Calibration and orientation are then performed through the tool Apero (or its simplified interface, Tapas). As input, the tool can receive inhomogeneous, and possibly redundant, observations, such as: the previously computed homologous points, externally measured GCPs (Ground Control Points) and GPS-measured positions of the camera projection centres. The process makes use of both computer vision techniques, in the initialization phase, and photogrammetric techniques, in the bundle adjustment (Triggs et al., 2000) refinement phase. The two steps

are mixed together, in order to avoid undesirable error accumulation and lead the system to convergence. Both pre-calibration and camera self-calibration are available with a choice of lens models. Finally, the dense image matching phase is carried out with the tool MicMac (simplified interface, Malt). The tool implements a multi-stereo formulation of the NCC (Normalized Cross Correlation) coefficient, specifically adapted to deal with large image datasets. The surface reconstruction is based on a multi-scale, multi-resolution image matching approach, which further reduces the required computational efforts and the possibility of having erroneous matching. In order to reconstruct a geometric surface, MicMac uses a regularization algorithm based on an energetic formulation.

1.4 Paper structure

This paper describes an approach designed to evaluate the metric quality and limitations of a particular image-based 3D modelling technique in a metrological context. The project is presented in Section 2. The first test is carried out in an outdoor environment, whereas the second one is performed within an ISO 1 environmentally controlled laboratory. For both cases, the performance of the algorithms implemented in the IGN's suite of tools is assessed for single views. In Section 3, a description of the origin of the reference data set is given along with the metrological approach. The results summarized in Section 4 are aimed at presenting mainly the metric evaluation of orientation and dense image matching phases. Concluding remarks are presented in Section 5.

2. PROJECT DESCRIPTION

2.1 Experiment in an outdoor environment

2.1.1 Cathédrale de la Major (Marseille, France)
The cathedral of Marseille (Cathédrale de la Major) is a Roman Catholic Church, characterized by a composite "Byzantine-Roman" style. Given the goal of this work, only the main entrance is chosen as test-object (Figure 1) that matches the following requirements:
- Significant depth variations and consecutive depth levels;
- Presence of detailed surfaces;
- Different textures and colours;
- Different materials (stones, marble and wood);
- High availability of open space in front of the scene;
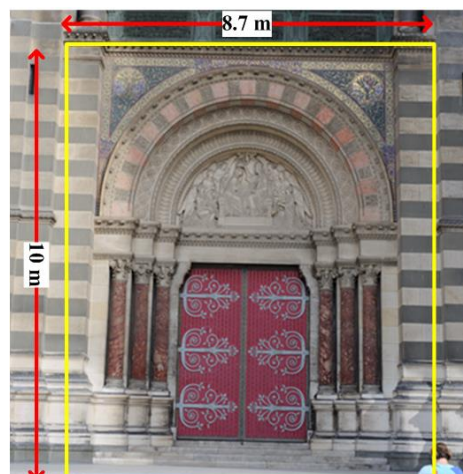- Outdoor conditions.

Figure 1. Test-object and dimensions (Cathédrale de la Major)

2.1.2   Procedural workflow

The IGN's suite includes very parametrical tools that provide the user with the possibility of finely controlling each processing phase through a large amount of attributes and parameters. This flexibility, however, forces the user to deal with the lack of clear rules and good best practices in the software manual. Furthermore, the acquisition protocol doesn't usually follow specific rules, especially in terms of convergence angles and employed focal setting. In order to fill these gaps, the present work aims at analysing the influence of different parametrical choices at each step of the IGN's suite. The effect of the image acquisition protocol employed is studied by examining its influence on the orientation and dense matching phases. For each step of the procedural workflow (Figure 2), a set of most significant parameters is examined, starting from selected acquisition protocols; adequate reference data are always employed. At the end of each analysis, a "best solution" is determined from accuracy results and then the solution is used as input for the subsequent phase.
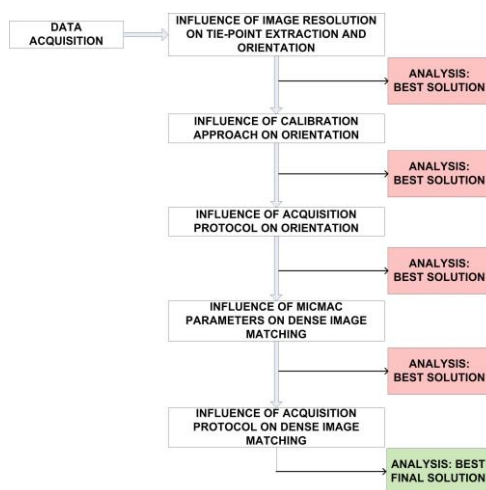

Figure 2. Procedural workflow (Cathédrale de la Major)

2.1.3   Image acquisition protocols

The image acquisition phase is performed using a Nikon D3X digital camera (6080 × 4044 pixels) and two different lenses: a fixed focal length lens (Nikon AF Micro-Nikkor 60mm f/2.8D) and a zoom-lens, employed at its lowest zoom level (Nikon AF Nikkor 24-85 mm f/2.8 4D IF), i.e. 24 mm. Both lenses are not equipped with optical image stabilizers that would represent a critical factor reducing the camera rigidity. For each lens, images are acquired with three different values of convergence angles α (3°, 5°, 10°), following the crosswise convergent configuration suggested in (Martin-Beaumont et al., 2013). The image acquisition does not use rigidly connected cameras on a stable structure. In order to achieve comparable final results, the camera-object distances are selected using design equations and an interpolation of 1/2 pixel (Blais and Beraldin, 2006) so that the resulting lens performances, in terms of both range uncertainty and lateral resolution, are metrically equivalent. The camera-object distances are 14 m and 26 m for the 24 mm-lens and the 60 mm-lens respectively. All acquisition protocols are performed with the same photographic parameter setup, i.e. focusing fixed at infinity, f8 and ISO(200).

2.2   Experiment in a controlled laboratory

2.2.1   ISO 1 Laboratory (Ottawa, Canada)

The National Research Council (NRC) of Canada Metrological Laboratory, built specifically for 3D imaging metrology work, is an environmentally controlled facility. Controlled air temperature, relative humidity (according to ISO 1), and cleanliness allow accurate and stable measurements (Beraldin et al. 2007). The test-object used, an *ad-hoc* 3D artefact (Figure 3), is characterized by:
- Significant depth and reflectance variations;
- Different textures and materials;
- Presence of detailed reliefs on the surfaces;
- Presence of quasi invisible small structural details.

The 3D scene includes contrast targets, scale bars and spheres. An interferometer-based scale bar is present in the laboratory (not shown in the photograph) and is used for the realization of the SI unit of length for laser trackers (B89.7.5, 2006).
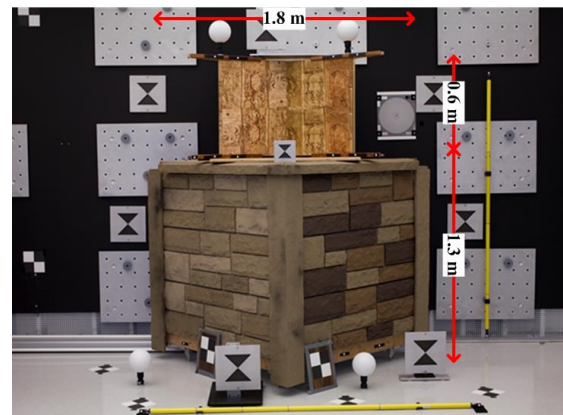

Figure 3. Test-object and dimensions (ISO1 Laboratory)

2.2.2   Procedural workflow

An environmentally controlled laboratory offers an advantageous context for traceable measurements. Uncertainty can be evaluated without worrying about the effect of the environment and a lack of resources (ISO 14253-2). In a way, such environment provides for a procedure to determine the best accuracy achievable with a particular image-based software suite and methodology. The attention is especially focused on the study of different image acquisition protocols and their resulting influence on the algorithm metric performance. Figure 4 summarizes the procedural workflow.
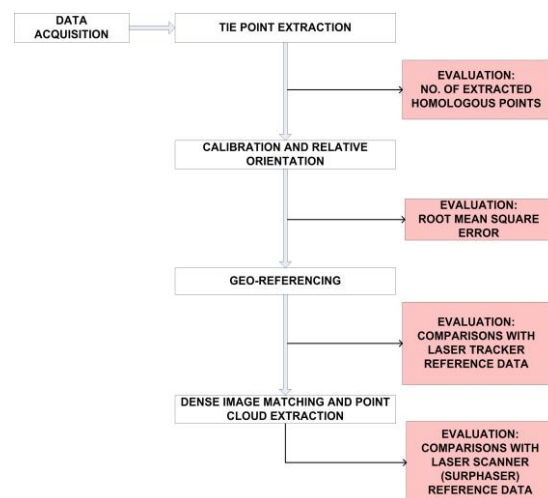

Figure 4. Procedural workflow (ISO 1 Laboratory)

2.2.3   Image acquisition protocols

A Canon EOS 5D digital camera (4368 x 2912 pixels) equipped with a fixed focal length lens (Canon EF 50 mm f2.5 Compact

Macro Lens) is used in the experiments. No automatic optical image stabilization is present. Two different angles of convergent images are tested (5° and 10°), following a crosswise convergent configuration without the use of a stable mounting structure. The 3D artefact is acquired at a focusing distance of 4.75 m, after having adequately glued the lens so that its focus setting doesn't change. F-stop and ISO sensibility are kept fixed at f-8 and 100 respectively. A diffused and controlled ambient light (fluorescent) provides an illumination without cast shadows.

3. METROLOGICAL APPROACH

3.1 Context

Many publications have looked at accuracy tests results for image-based methods. Most of the comparisons are performed using either 'reference' data acquired with laser scanners or geometric artefacts with known form and size (spheres, flat planes, and gauge blocks). Incidentally, some authors use the expression "ground truth" to signify a reference data set. The VIM3 offers explanatory notes on the use of the word "true" and thereafter we use "reference data set". Although this general approach may be seen as reasonable, the latter approach has a better metrological traceability than the former. For instance, roundness measurement on a reference artefact is a well-established technology with a traceability chain. On the other hand, laser scanners measurements are performed on a regular basis but internationally recognized standards are nowhere to be found let alone a traceability chain. It is true that manufacturers of laser scanners or 3D imaging systems provide to their customers a datasheet or a calibration certificate with some numerical values and sometimes a terminology that is much closer to VIM3 and GUM 2008. Nevertheless, it is important to understand that these documents are not generated according to an international standard but instead they are linked to internal company guidelines that may or may not be available to the customer. Fortunately, efforts are being made in international organizations like ASTM (E57), ISO/TC 213 (ISO 10360) and ISO/TC 172/SC 6 (ISO 17123) and by the German national body VDI-VDE (2634). Current methods rely on a comparison between a given measurement dataset from one "instrument" with another measurement dataset from a reference "instrument" which has a smaller uncertainty (at least four times smaller as per ISO 14253). In reality, the user has to be satisfied with an unmediated comparison via dissimilar instrument technologies or in the absence of laser scanner data, use a reference data set derived from the different data sets under study where some statistical method has been applied to the whole data set. The metrological approach we use in this paper attempts to fill the void by reviewing the metrological aspects of the problem and by proposing an avenue for solution with dissimilar technologies.

3.2 Fundamental questions in metrology

There are three fundamental questions that one must answer when working in metrology. The first question is "how do you know what you are measuring?"; this is the "measurand" issue in metrology. According to the VIM3, a measurand is the quantity intended to be measured. The measurand cannot be specified by a value but only by a description of a quantity. The set of quantity values being attributed to a given measurand together with any other available relevant information are the measurement results or results of measurement. Incidentally,

according to VIM3, accuracy is a *qualitative* term. Uncertainty should be used to express the accuracy of a measurement (the symbol u is typically used and in the case of an expanded uncertainty, U). The second question is "how do you know you can trust the measurement?" This is the calibration issue (VIM3). The third question is "how do you know measurements are equivalent?". This is the traceability issue (VIM3). Here both an uncertainty evaluation of a measurement result and an explicit connection to the metre in the case of dimensional metrology are required. If one wants to perform an inter-comparison or an evaluation of a particular 3D imaging system in a metrological context, all three questions must be answered. Furthermore, one shall strive to consider these three questions in metric surveys and develop an adequate understanding of the uncertainty components in measurement (as per ISO 14253-2).

3.3 Performing data inter-comparison

3.3.1 Methodology
Our accuracy tests are performed using 'reference' data acquired with two hemispherical laser scanners (LS1 and LS2), one total station (TS), one laser tracker (LT), and some contrast targets. Only the TS and LT have a clear measurement traceability route; the laser scanners may have one but no information is present in the data sheets. The two experiments are conducted according to the protocols presented in Section 2. Small volumes are measured in order to restrict the study to estimating the accuracy over a limited number of instrument stations and single 3D points clouds. PolyWorks v12.1.18 IMAlign™ software package from InnovMetric Software Inc. provides the main image alignment and 3D point cloud comparison techniques. The alignment is based on an iterative algorithm that computes an optimal alignment by minimizing the 3D distances between surface overlaps in a set of 3D images acquired from unknown viewpoints. Here, a measurand around that distance can be defined. After each iteration, the algorithm applies to each 3D image a transformation matrix corrected by an incremental transformation matrix that best improves the image alignment with respect to the other 3D images. This incremental matrix is computed using a linear least-squares technique, and results from averaging the best alignment parameters of each image point. When the alignment process is set to perform no iteration, we get a comparison between 3D point clouds. This feature is useful when the two 3D point clouds being compared are already registered in the same coordinate system. A histogram or an error map of image alignment errors bounded by a maximum acceptable distance error between an image point and another 3D image can be generated.

3.3.2 Measurements uncertainty and instrumentation
A recently calibrated Leica FlexLine™ TS06plus Total Station is used to survey the contrast targets in the main entrance of the Cathedral in Marseille. The survey is performed from a single station in the morning when the wind was weak, the temperature was about 10 °C ± 2 °C, relative humidity was about 60% ± 10% and the barometric pressure was fairly stable at 1019 hPa ± 2 hPa. The specification sheet quotes a distance measurement accuracy of 2 mm + 2 ppm (standard deviation as per ISO-17123-4 and without reflector) and angle measurement (Hz, V) accuracy between 2" and 7" (ISO-17123-3). The maximum distance within the scene is about 16.6 m. Using this information and the ISO-17123-4, the uncertainty budget calculation shows that the main uncertainty comes from the range distance estimation. The elevation and azimuth angular uncertainties at 16.6 m are about one quarter of the range

uncertainty. The combined uncertainty for the range is about $u_R(TS)=2.18$ mm (1 σ) and expanded uncertainty $U_R(TS)=4.36$ mm (k=2). A Hemispherical FARO® Laser scanner Focus[3D]model 120 is used on that same entrance. The alignment of two 3D point clouds acquired from two distinct distance positions in front of the main entrance, gave instead a 1 σ value of about 1 mm which is selected as our local measurement uncertainty $u_R(LS1)=1$ mm. This last value is a more realistic representation of the noise level present in the 3D point clouds. Already, one can see that the TS may become a limiting factor in the determination of the accuracy of any techniques in the present context. In a separate laboratory test, the laser scanner lateral resolution was evaluated between 5 m and 20 m. With a contrast target characterized by a pattern with a spatial frequency of 5 lp/mm, the lateral resolution was shown to be in the 1 mm to 2 mm range between 5 m and 10 m.

In the laboratory, an absolute distance meter (ADM)-based laser tracker model FARO® X is used to locate the contrast targets of the GCPs and Check Points (CPs) instead of a TS. Typically, specifications for laser trackers are given by an ASME B89.4.19-2006 assessment where a maximum permissible error (MPE) is specified with a traceability route. The equation below gives the error in range L when distance is measured from the centre of the instrument (0.1 m to 35 m):

$$E_{L,MPE} = \pm(20\,\mu m + L \bullet 0.8\,\mu m) \tag{1}$$

Metrologists interpret these extreme values as a 2σ. The transverse capability of the laser tracker (elevation and azimuth angles) is given by

$$E_{T,MPE} = \pm(36\,\mu m + L \bullet 6\,\mu m) \tag{2}$$

For a working distance of 4.75 m, the radial expanded (k=2) uncertainty is approximately $U_R(LT)=23.8$ μm and the transverse expanded (k=2) uncertainty is approximately $U_T(LT)=64.5$ μm. The radial measuring capability of laser trackers in terms of uncertainty is much better than its angular measuring capability. A hemispherical 3D scanner model Surphaser® 25HSX is used in the laboratory. Tests results show that the 1σ noise after the alignment of two 3D point clouds acquired from two distinct distance positions in front of the 3D artefact gave a value of about 0.3 mm at a range of 5 m which is selected as our local measurement uncertainty uR(LS2)=0.3 mm. This last value is again a more realistic representation of the noise level present in the 3D point clouds. Lateral resolution was evaluated using a star pattern and the lateral resolution is found to be in the ½ mm range at a range of 5 m.

3.4 Uncertainty and the importance of an error budget

From the calculation shown above for a single point for the TS, an expanded uncertainty (k=2) of 4.36 mm for radial measurements and 1.09 mm for angular measurements are expected on the contrast targets. A local single 3D point expanded uncertainty for the FARO® laser scanner of about 2 mm is anticipated on the entrance. In retrospect, a TS with a lower measurement uncertainty should have been used at the Marseille site. The LT provides coordinates measurements with an uncertainty well below those achievable with both the Surphaser® laser scanner and the image-based technique being evaluated. From calculations, an expanded uncertainty (k=2) of 0.094 mm for a 3D coordinates are expected on the NRC contrast targets. A local single 3D point expanded uncertainty for the Surphaser® laser scanner of about 1 mm is anticipated on the surfaces of the 3D artefact.

4. RESULTS

4.1 Results achieved in an outdoor environment

4.1.1 Metrological assessment of image orientation
Both image correspondences (Tapioca) and measured GCPs are employed in order to compute camera poses and orientations. In particular, the registration of image-based results within the Total Station (TS) reference frame is achieved by using seven well-distributed planar targets. Internal orientation parameters, pre-computed within a calibration procedure performed with an *ad-hoc* image dataset, are refined in the process by adopting a Fraser-derived formulation (Fraser, 2001). Once orientations are defined, each pixel visible in at least two images can be back-projected into the 3D absolute space, defining its 3D position as intersecting rays. The 3D coordinates of nine CPs are thereby computed and then compared to the ones measured with TS. This accuracy assessment is performed starting from all available acquisition protocols, after having oriented the corresponding dataset with the procedure mentioned earlier. Final standard deviations, σ, (Table 1) show that results achieved by the 60mm-lens are generally better than the corresponding ones gathered by the 24mm-lens. This is due in part by the better range and lateral accuracy provided by longer focal length, according to triangulation equations. Also, the combined datasets (highlighted in yellow) show the best metric performance.

24 mm-lens			
α	# Images	σ x (mm)	σ y (mm)	σ z (mm)
3°	5	11.9	15.4	12.6
5°	3	20.6	28.6	14.3
10°	3	18.9	27.7	24.2
3°+5°+10°	9	9.0	13.2	11.3
60 mm-lens			
α	# Images	σ x (mm)	σ y (mm)	σ z (mm)
3°	5	9.2	13.5	5.8
5°	3	18.7	16.5	20.8
10°	3	14.5	15.9	14.3
3°+5°+10°	9	4.8	2.2	4.6
24+60 mm-lenses			
α	# Images	σ x (mm)	σ y (mm)	σ z (mm)
3°+5°+10°	29	2.9	2.1	3.1

Table 1. Accuracy assessment of image orientation phase, $u_R(TS)=2.18$ mm (Cathédrale de la Major)

4.1.2 Metrological assessment of dense image matching
Depth maps are extracted, after having selected the best parametrical setup in accordance with the scene characteristics. In particular, many tests are carried out in order to deepen our understanding of three chief parameters, i.e. the regularization factor, the Z-quantification factor and the final Z-resolution (see MicMac Documentation). A metric evaluation of the dense image matching accuracy is then performed through two different studies. Point clouds are first analysed with the software PolyWorks: best-fit geometrical primitives are thereby extracted from significant portions of the acquired 3D scene, such as planar surfaces (e.g. pillar) and cylinders (e.g. columns). Resulting standard deviations and RMSE values are then evaluated. Secondly, the metrological assessment is completed by performing comparisons with reference data, i.e. the point cloud acquired with the FARO® Laser scanner Focus[3D] (LS1)

at a mean instrument-object distance equal to 5 m. The tests are carried out within the software PolyWorks, after having registered the LS1 data with the image-based point cloud (registration to the TS reference frame is possible only with the image-based point cloud). Three significant portions of the symmetrical test-object are selected for the comparisons. For each subset, standard deviations (σ) of the distances between the compared point clouds and the corresponding histograms are computed and analysed. Results are listed in Table 2.

24 mm-lens			
α	# Images	σ (mm)	Histogram
3°	5	7.65	
5°	3	6.66	
10°	3	6.22	
3°+5°+10°	9	4.83	
60 mm-lens			
α	# Images	σ (mm)	Histogram
3°	5	8.38	
5°	3	8.34	
10°	3	7.14	
3°+5°+10°	9	5.07	
24+ 60 mm-lenses			
α	# Images	σ (mm)	Histogram
3°+5°+10°	29	3.73	

Table 2. Statistics delivered by inter-comparisons performed on the half portal 3D data, u_R(LS1)=1 mm (Cathédrale de la Major)

The best accuracy level (highlighted in yellow) is delivered by the dataset including images acquired by both lenses. In this case, an accuracy level below half a centimetre is achieved. The color-coded map associated with this acquisition protocol is shown in Figure 5, together with a view of the compared 3D scene. The colour scale ranges from -10 mm (blue) to 10 mm (red).
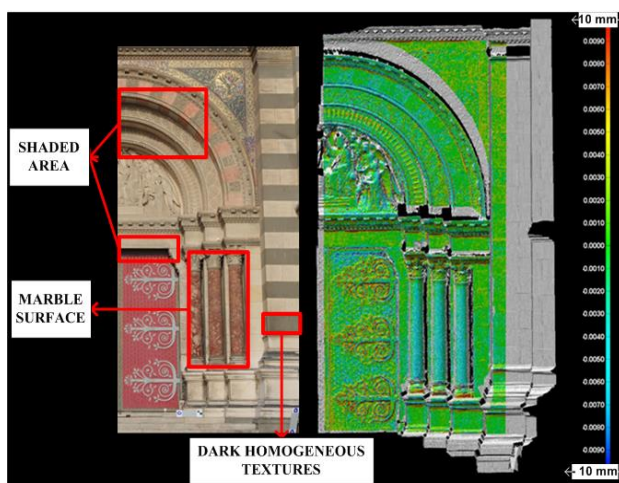

Figure 5. Colour-coded error map delivered by inter-comparison performed on the portal (right) and the resulting problematic areas shown on the colour image (left).

The error distribution shows that the areas delivering the largest deviations are mainly the followings ones:

- Areas subject to cast shadow, that deliver large differences also when airborne images are employed (Haala, 2013);
- Area characterized by dark and homogeneous textures, such as the dark pattern of the pillar;
- Marble columns, where the computed differences are negative. This problem is connected to the performance of time-of-flight laser scanners in the presence of translucent surfaces. Here, an apparent depth penetration 5-6 mm is observed (El-Hakim et al., 2008).

4.2 Results in an environmentally controlled laboratory

4.2.1 Metrological assessment of image orientation
Once image correspondences and relative orientations are computed (Tapioca and Tapas), the datum ambiguity is solved in this experiment by using four well-distributed ground control points. In particular, GCPs are selected within the NRC-targets measured with the LT: the same data are employed in order to register also the LS2-acquired point clouds in the same LT reference frame. Camera calibration parameters are computed by performing a self-calibration within the bundle adjustment procedure. In order to evaluate the metric accuracy of the photogrammetric orientation results, the remaining 10 targets (contrast and NRC-targets) are then assumed as independent check points and matched in at least three images. Their 3D coordinates, thereby computed as intersections of homologous rays, are finally compared to the ones measured with the LT and LS2, delivering the standard deviations, σ, listed in Table 3.

50 mm-lens				
α	# Images	σ_x (mm)	σ_y (mm)	σ_z (mm)
5°	15	0.42	0.60	0.36
10°	15	0.49	0.80	0.32
5°+10°	27	0.47	1.34	0.40

Table 3. Accuracy assessment of the orientation phase (ISO 1 Laboratory)

Results show that the orientation algorithm achieves a considerably good accuracy level: all standard deviations are below 1 mm, with the only exception of one value. The results show that they are consistent with the expanded measurement uncertainty of both LT and LS2. By analysing the individual residuals, the higher deviations correspond to a few targets lying on the floor, whose position requires an unfavourable acquisition direction for all the instruments employed.

4.2.2 Metrological assessment of dense image matching
Depth maps are then extracted, by adopting the best parametrical setup. Starting from previously computed information, i.e. orientations (both internal and external), depth values, origin and steps of depth quantification, point clouds are finally delivered in PLY file format. Figure 6 shows the result achieved with the 5°-dataset.
The raw image-based point clouds are finally compared with measured reference data, i.e. the 3D point cloud acquired with the LS2. The tests are performed within the software PolyWorks. No alignment was necessary, just a straight comparison is performed in the software (iteration set to 0). The standard deviations (σ) of the distances between the compared entities and corresponding histograms are listed in Table 4.
The dense image matching algorithm is able to reach the same accuracy level previously pointed out by the orientation metrological assessment: all tests, in fact, deliver sub-millimetre standard deviations and comparable results.
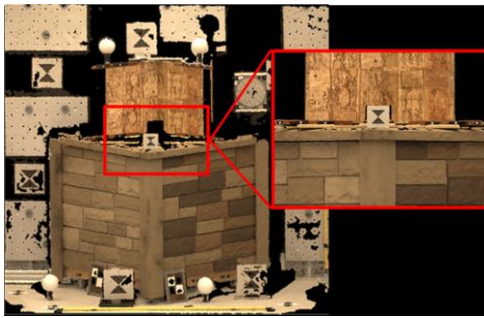
Figure 6. The point cloud extracted with the 5°-dataset.

50 mm-lens			
α	# Images	σ (mm)	Histogram
5°	15	0.88	
10°	15	0.67	
5°+10°	27	0.67	

Table 4. Results obtained in the ISO 1 Laboratory.

Furthermore, the colour-coded error maps show that the largest errors, in terms of deviation from the reference data, are mainly located at sharp surface gradients, such as the ones corresponding to the edges between the vertical walls of the corners and to the small grooves among the bricks. As evidence, the error distribution associated to results achieved with the 10°-dataset is shown in Figure 7: the colour scale ranges from -5 mm (violet) to +5 mm (red). These sharp edges are problematic for active laser scanners when the spot diameter is large compared to the structural (lateral) resolution (VDI/VDE 2617) being analysed. In the present situation, there may be a mismatch between the structural resolution of the LS and the image-based 3D point clouds. Finally, the figure shows that the image-based reconstruction can resolve small creases on the glued surfaces.
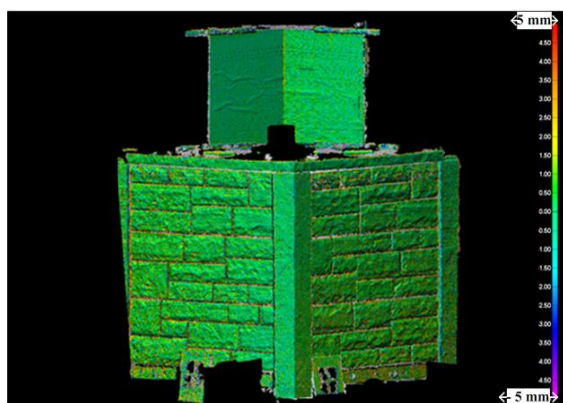

Figure 7. Colour-coded map delivered by inter-comparison and starting from the 10°-dataset

5. CONCLUSION

An increasingly number of software solutions for the automatic generation of textured dense 3D point clouds from a set of un-oriented images has recently appeared in the market. So far this development has not been supported by the definition of clear standards and guidelines that are necessary to evaluate the resolution, repeatability, reproducibility and measurement uncertainty. It is within this context that this paper strives to provide an approach to assess both the metric quality and limitations of the tested image-based 3D modelling technique and that within a metrological context. The metrological approach adopted here is based on reference instruments, whose uncertainty is estimated through both specifications and experimental tests. This approach is applied to the metric evaluation of the algorithmic solutions implemented in the IGN's suite of tools (Apero/MicMac); in particular, the influence of the image acquisition protocol used is studied in some depth, by examining its effects on the orientation and dense matching phases. Two experiments are carried out, in order to analyse the software performance in dealing with both outdoor and environmentally controlled conditions for the case of a single 3D image. The results achieved within these two contexts are not general since they are of course influenced by the specific operative conditions affecting each case study, i.e., datasets, hardware/software means, ambient and operators. Nevertheless, these studies offer a possible reference procedural workflow that can be further applied to different case studies in order to set specifically-adoptable best practices. Aspects of metrological nature should be developed further for the good of the different communities and for market growth. These aspects include defining what the measurand is in a given comparison, how calibration is performed and how often, what the traceability route is that links a measurement value to the SI unit, and, finally how the measurement accuracy of a system is quoted. Many of these aspects offer great opportunities for research and development for academia, national measurement institutes (NMI) and industry.

REFERENCES

References from Books:

Fraser, C.S., 2001. Photogrammetric Camera Component Calibration: A Review of Analytical Techniques. In: *Calibration and Orientation of cameras in Computer Vision*, Grün and Huang (Eds.), Springer Series in Information Sciences 34, ISBN 3-540-65283-3, pp. 95-121.

Pears, N., Liu, Y., Bunting, P., 2012. *3D Imaging, Analysis and Applications*. Springer-Verlag (London, UK), ISBN 978-1-4471-4062-7, 499 pp.

Vosselman, G., Mass, H.-G., 2010. *Airborne and terrestrial laser scanning*. Whittles Publishers (Scotland, UK), ISBN 978-1904445-87-6, 336 pp.

References from Journals:

Ahmadabadian A. H., Robson, S., Boehm, J., Shortis, M., Wenzel, K., Fritsch, D., 2013. A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 78, pp. 157-167.

Blais, F., Beraldin, J.-A., 2006. Recent Developments in 3D Multi-modal Laser Imaging Applied to Cultural Heritage. In: *Machine Vision and Applications*, pp. 395-409.

Gallo, A., Muzzupappa, M., Bruno, F., 2013. 3D reconstruction of small sized objects from a sequence of multi-focused images. In: *Journal of Cultural Heritage*, Vol. 15, Issue 2, pp. 173-182.

Koutsoudis, A., Vidmar, B., Ioannakis, G., Arnaoutoglou F., Pavlidis, G., Chamzas, C., 2013. Multi-image 3D reconstruction data evaluation. In: *Journal of Cultural Heritage*, Vol. 15, Issue 1, pp. 73–79.

Pierrot-Deseilligny, M., De Luca, L., Remondino, F., 2011. Automated image-based procedures for accurate artifacts 3D modeling and orthoimage generation, In: *Geoinformatics FCE CTU Journal*, Vol. 6, pp. 291-299.

Scharstein, D., Szeliski, R., 2002. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. In: *International Journal of Computer Vision*, 47(1-3), pp. 7-42.

References from Other Literature:

Beraldin, J-A., 2004. Integration of Laser Scanning and Close-Range Photogrammetry – The Last Decade and Beyond. ISPRS, Istanbul, Turkey, Vol. XXXV, Part. B5, pp. 1031-1042.

Beraldin, J.-Angelo, Rioux, M., Cournoyer, L., Blais, F., Picard, M., Pekelsky, J., 2007. Traceable 3D imaging metrology. In: *Electronic Imaging 2007*. International Society for Optics and Photonics, pp. 64910B-64910B-11.

El-Hakim, S., Beraldin, J.-A., Picard, M., Cournoyer, L., 2008. Surface reconstruction of large complex structures from mixed range data – The Erechtheion experience. ISPRS, Beijing, China, Vol. XXXVII, Part. B5, pp. 1077-1082.

Cléry, I., Pierrot-Desseilligny, M, 2011. An ergonomic interface to compute 3D models using photogrammetry. *Proc. XXIII CIPA Symposium*, Prague, Czech Republic,12-16 Sept. 2011.

El-Hakim, S., Beraldin, J.-A., Blais, F., 2003. Critical Factors and Configurations for Practical 3D Image-Based Modeling. *6th Conf. on 3D Measurement Techniques*, Zurich, Switzerland, Vol. II, pp. 159-167.

Georgantas, A., Brédif, M., Pierrot-Deseilligny, M., 2012. An accuracy assessment of automated photogrammetric techniques for 3D modelling of complex interiors. ISPRS, Melbourne, Australia, Vol. XXXIX, Part B3, pp. 23-28.

Haala, N., 2013. The Landscape of Dense Image Matching Algorithms. *Proc. Photogrammetric Week 2013*. Dieter Fritsch (Ed.), Stuttgart, pp. 271-284.

Kersten, T.P., Lindstaedt, M., 2012. Image-Based Low-Cost Systems for Automatic 3D Recording and Modelling of Archaeological Finds and Objects. *Proc. EuroMed 2012 Conference*, M. Ioannides et al. (Eds.), LNCS 7616, Springer Berlin Heidelberg, pp.1-10.

Martin-Beaumont, N., Nony, N., Deshayes, B., Pierrot-Deseilligny, M., De Luca, L., 2013. Photographer-friendly work-flows for image-based modelling of heritage artefacts. ISPRS, XXIV International CIPA Symposium, Strasbourg, France, Vol. XL, Part. 5/W2, pp. 421-424.

Nony, N., De Luca, L., Godet, A., Pierrot-Deseilligny, M., Remondino, F., Van Dongen, A., Vincitore, M., 2012. Protocols and Assisted Tools for Effective Image-Based Modeling of Architectural Elements. *Proc. EuroMed 2012 Conference*, M. Ioannides et al. (Eds.), LNCS 7616, Springer Berlin Heidelberg, pp. 432-439.

Phillips, S., Krystek, M., Shakarji, C., Summerhays, K., 2009. Dimensional measurement traceability of 3D imaging data. *Proc. SPIE 7239*, Three-Dimensional Imaging Metrology, 72390E (January 19, 2009).

Pierrot-Deseilligny, M., Clery, I., 2011. APERO, an open-source bundle adjustment software for automatic calibration and orientation of set of images. ISPRS Int. Workshop 3D-ARCH 2011, Trento, Italy, Vol. XXXVIII, Part. 5/W16, pp. 269-277.

Pierrot-Deseilligny, M., Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach: an application to surface reconstruction from SPOT5-HRS stereo imagery. ISPRS Workshop on Topographic Mapping from Space, Ankara, Turkey, Vol. XXXVI, Part. 1/W41, pp. 73-77.

Remondino, F., Del Pizzo, S., Kersten, T.P., Troisi, S., 2012. Low-cost and open-source solutions for automated image orientation – A critical overview. *Proc. EuroMed 2012 Conference*, M. Ioannides et al. (Eds.), LNCS 7616, Springer, Heidelberg, pp. 40-54.

Seitz, Steven M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *Computer vision and pattern recognition, 2006 IEEE Computer Society Conference on*. IEEE, pp. 519-528.

Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment – a modern synthesis. In: *Vision Algorithms: Theory and Practice*, Lecture Notes in Computer Science 1883, Springer-Verlag, Berlin, Germany, pp. 298-.372.

Waldhäusl P., Ogleby C., 1994. 3 x 3 Rules for simple photogrammetric documentation of architecture. ISPRS, Vol. XXX, Part5, pp 426- 429.

Wenzel, K., Rothermel, M., Fritsch, D., Haala N., 2013. Image acquisition and model selection for multi-view stereo. ISPRS, Volume XL-5/W1, 3D-ARCH 2013, 25–26 Feb., Trento, Italy 251-258.

References from Websites (Last accessed on 2014, Mar. 12)**:**

ASTM (E57), http://www.astm.org/COMMITTEE/E57.htm

B89.4.19, 2006. Performance Evaluation of Laser-Based Spherical Coordinate Measurement Systems, http://asme.org

B89.7.5, 2006. Metrological Traceability of Dimensional Measurements to the SI Unit of Length, http://asme.org

Culture 3D Clouds, http://c3dc.fr/

EuroSDR, http://www.eurosdr.net/

GUM 2008. Guide to the expression of uncertainty in measurement, http://www.bipm.org

ISO 3650:1998. Geometrical Product Specifications (GPS) - Length standards - Gauge blocks, http://www.iso.org/

ISO 1:2002. Standard reference temperature for geometrical product specification and verification, http://www.iso.org/

ISO/TC 172/SC 6 (ISO 17123), http://www.iso.org/

ISO/TC 213 (ISO 10360), http://www.iso.org/

MicMac Documentation, http://logiciels.ign.fr

PolyWorks, by InnovMetric, http://www.innovmetric.com

VDI/VDE (2617), 2005. VDI/VDE 2617 Blatt 6.2 Accuracy of coordinate measuring machines with optical distance sensors, http://www.vde.com/

VDI-VDE (2634), 2012. VDI/VDE 2634 Blatt 2 Optical 3-D measuring systems - Optical systems based on area scanning, http://www.vde.com/

Vedaldi, 2010. http://www.robots.ox.ac.uk/

VIM3, JCGM 200:2012. International vocabulary of metrology, http://www.bipm.org/en/publications/guides/vim.html