QUALITY FEATURES FOR THE INTEGRATION OF TERRESTRIAL AND UAV IMAGES

E. M. Farella [1], A. Torresani [1], F. Remondino [1]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Web: http://3dom.fbk.eu - Email: <elifarella><atorresani><remondino>@fbk.eu

Commission II

KEY WORDS: photogrammetry, data fusion, terrestrial images, UAV

ABSTRACT:
The paper presents an innovative approach for improving the orientation results when terrestrial and UAV images are jointly processed. With the existing approaches, the processing of images coming from different platforms and sensors leads often to noisy and inaccurate 3D reconstructions, due to the different nature and properties of the acquired images. In this work, a photogrammetric pipeline is proposed to filter and remove bad computed tie points, according to some quality feature indicators. A completely automatic procedure has been developed to filter the sparse point cloud, in order to improve the orientation results before computing the dense point cloud. We report some tests and results on a dataset of about 140 images (Modena cathedral, Italy). The effectiveness of the filtering procedure was verified using some internal quality indicators, external checks (ground truth data) and qualitative visual analyses.

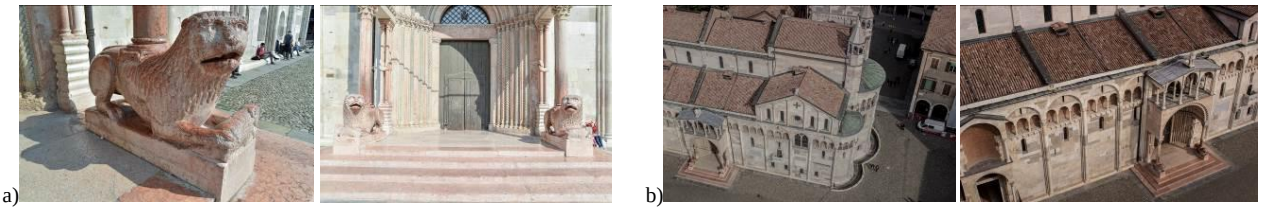


Figure 1: Some examples of the image dataset acquired with a terrestrial camera (a) and UAV platforms (b) for the 3D documentation of the Modena Cathedral (Italy).

1. INTRODUCTION

In the field of Geomatics and Geoinformatics, data integration (or sensor integration) is understood as the combination of data acquired with various sensors in order to derive more accurate information which cannot be deduced from one sensor alone (Zhang, 2010; Gasparovic and Malaric, 2012; Petrasova et al., 2017). The integration of heterogeneous (2D or 3D) data acquired with different sensors and platforms is a relevant and open issue in various fields and applications: 3D modelling of man-made environments, heritage documentation, medicine, autonomous driving, etc. In the cultural heritage field, data integration is often necessary for recording large structures but also complex and small architectural details. Depending on the data sources, several approaches have been explored (Khaleghi et al., 2013; Ramos and Remondino, 2015). However, each developed method is focused on specific issues related to sensors' limitations and data characteristics (point clouds density, accuracies, viewpoints, scales, etc.).

1.1 Aim of the paper

The paper presents a novel approach to integrate – at raw data level – images coming from different platforms (e.g. terrestrial and UAV) and acquired for the documentation of cultural heritage monuments (Figure 1). Such image datasets, if not properly processed by the bundle adjustment method, could lead to errors in the 3D reconstructions due to too large viewpoints, scale and illumination changes, etc. We propose to use a set of quality features (Section 3.1) computed directly on the sparse point cloud derived by the orientation procedure. These features are combined (Section 3.2) and used to remove possible outliers from the 3D tie points before recomputing the orientation parameters of the image dataset and the final dense point cloud (Section 3.3).

2. RELATED WORKS

In the last years many solutions were presented for the integration of geospatial data in various disciplines, such as autonomous guiding of vehicles, robotics, heritage documentation, landslide monitoring, etc. (Luo et al., 2011; Serna et al., 2015; Zieher et al., 2018). All solutions try to exploit advantages of the different data (or sensors) in order to overcome limitations of single data sources.
Data integration can be performed on three levels, depending on the purpose of the data acquisition and the employed sensors (Kochi et al., 2012; Bastonero et al., 2014): (i) raw data level (pixels in case of raster data or 3D points in case of point clouds), (ii) feature level and (iii) decision level. The integration at the raw data level requires sensors to provide similar types of data (i.e. measurements of the same physical entity). Alternatively, data must be integrated based on either extracted features (feature level) or decisions inferred from the single sensors (decision level). The integration of surveying data derived from different sensors or platforms could be accomplished on each of these levels. However, the raw data level should be preferred, as extractable features or inferred decisions may change over time and could not be comparable.
Another way of classifying data integration methods is based on the registration and processing procedures:

- Registration methods based on 3D information: they include ICP-like algorithms optimized for merging data coming from heterogeneous sources or hybrid approaches (Pomerleau, 2015; Mandlburger et al., 2017; Bracci et al., 2018);

Figure 2: An example of the sparse (left) and dense (right) point cloud produced by jointly processing UAV and terrestrial images featuring different scales, viewpoints, etc. The final dense cloud is noisy and with visible misalignments, due to inaccurate orientation results.

- Registration methods based on automated 2D-2D or 3D-3D correspondence extraction: they focus on the development / improvement of new tie point or 3D feature matching extraction approaches (Tombari and Remondino, 2013; Gerke et al. 2016; Koch et al., 2016; Wu et al., 2018);

3. DEVELOPED METHODOLOGY

Two main approaches are usually employed to process images acquired with different sensors and platforms. In the first approach, image datasets are independently oriented, scaled/georeferenced and different dense 3D point clouds are generated. The obtained point clouds are then aligned and merged with several procedures (i.e. common points among the clouds, matching targets across the datasets, etc.). The final merged dense point cloud often presents visible misalignments as well as color and texture mismatches. In the second approach, all the available images are jointly processed together within the bundle adjustment. However, the derived dense point clouds are often noisy (Figure 2), mainly due to inaccurate orientation results caused by large viewpoints, scale and illumination changes, outliers in the tie points, etc.

The presented work is focused on the joint processing of terrestrial and UAV images and on the implementation of new procedures for the analysis and the improvement of the orientation results. Using some quality features computed on the sparse point cloud produced by the orientation step (bundle adjustment), computed camera parameters are deeply evaluated and refined, before running the dense image matching and generate the final dense point cloud. The entire pipeline is built upon MicMac (Rupnik et al., 2017) as shown in Figure 3. The specific filtering procedure is developed in Python, so it could be coupled to any other photogrammetric processing pipeline.

Contrary to other methods (e.g. Mauro et al., 2014) where features are computed for next best view (NBV) planning and image selection, our method wants to eliminate points that negatively affect the image orientation step and, consequently, the dense point cloud generation.

3.1 Quality features

The evaluation of image orientation's quality within SfM methods can be performed using external checks and internal quality statistics (Roncella et al, 2011). Internal statistics can be used to evaluate the accuracy of the feature extraction and matching techniques whereas external check can control possible block deformations.

Given a set of oriented images, we derive some quality features of the sparse 3D points using in-house developed procedures. We compute the following quality features (Remondino et al., 2017):

- Reprojection error in image space, i.e. the Euclidean distance between a measured image point and the back-projected position of the corresponding 3D point in the same image. It is normally not true that a small reprojection error is a sign of good 3D point. A high reprojection error can negatively affect the quality of the exterior orientation.
- Image redundancy (or multiplicity), i.e. the excess of observations (image points) with respect to the number unknowns (3D object coordinates). This value is related to the number of images where a point has been measured. Therefore, in theory, the higher is the redundancy, the better is the quality of the 3D point – assuming a good intersection angle.
- Intersection angle, i.e. the (maximum) angle between viewing directions generating a 3D point. A high intersection angle increases the quality of a 3D point.
- DoN (Difference of Normals) vector, i.e. the normalized sum of two surface normal vectors derived from different support radii (Ioannou et al., 2012). DoN is a multi-scale filtering operator, used in unorganized point clouds to compute surface discontinuites. High DoN values are expected in areas with substantial surface changes in the given radii range.
- 3D point density, i.e. the distribution of 3D points in object space computed in local neighborhoods. Assuming a good image texture, a low density values can be related to outliers.

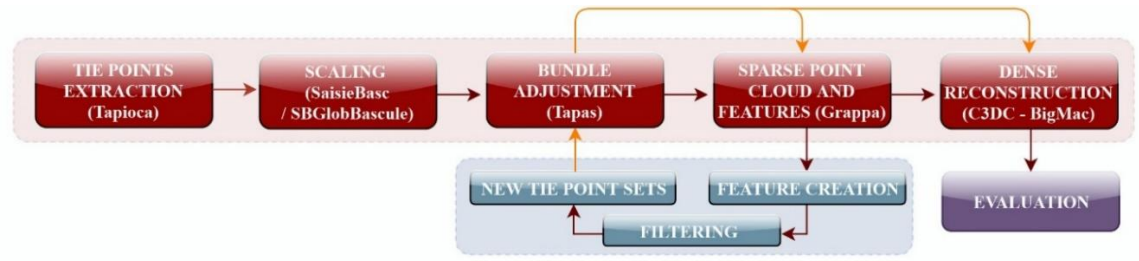


Figure 3: The pipeline flowchart, developed with the MicMac suite. The standard workflow (red blocks) is extended with the filtering procedure developed in Python (blue blocks) for the removal of bad estimated 3D tie points in order to improve the orientation results and the dense 3D reconstruction.

3.2 Features aggregation

All the aforementioned features (Figure 4), directly linked to image orientation results, network geometry and complexity of the surveyed objects, have different ranges and therefore they need to be normalized and rescaled in the same range [0, 1] in order to be aggregated. The normalization is performed using a "feature scaling" method. This method allows to standardize and scale a range of independent variables and features of data. For each input value *x*, the normalized value *y* is given by:

$$y = \frac{(x - min_D) * (max_N - min_N)}{(max_D - min_D)} + min_N \qquad \text{(Eq. 1)}$$

where min_D and max_D are the highest and lowest values in the data and min_N and max_N are referred to the values in the new range.
Using these normalized features, we aggregate them as a linear combination using specific weights for each feature:

$$A_v = (V_{RE} * \alpha) + [1 - (V_M * \beta)] + [1 - (V_A * \gamma)] + [1 - (V_{DoN} * \varphi)] + [1 - (V_D * \mu)] \qquad \text{(Eq. 2)}$$

where:

- A_V is the aggregated value computed for each 3D tie point;
- V_{RE} is the normalized value of the reprojection error and α its weight;
- V_M is the normalized multiplicity value and β its weight;
- V_A is the normalized angle of intersection and γ its weight;
- V_{DoN} is the normalized DoN derived from different support radii and φ its weight;
- V_D is the normalized value of the density computed in a local neighborhood and μ its weight.

The aggregated value A_V represents a sort of quality of a 3D point in terms of the combined contribution of the employed features. As expressed in Eq. 2, high values of A_V denote low-quality points.

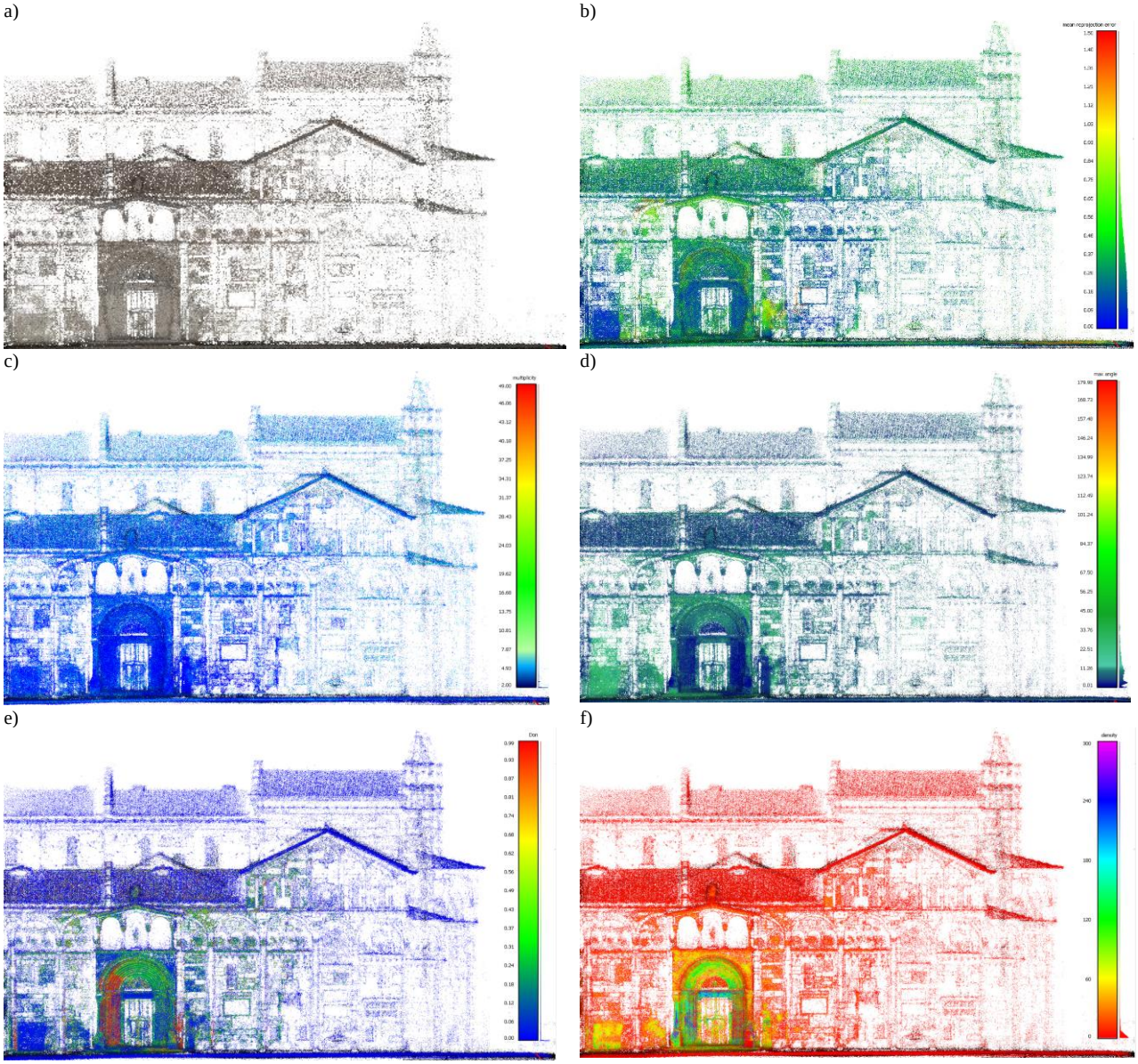


Figure 4: Examples of computed quality features visualized on the sparse point cloud (about 3 million points - a): reprojection error (b), multiplicity (c), intersection angle (d), difference of normals (DoN) (e) and 3D density (f). Please note that the DoN (e) is higher where surface discontinuities are present (i.e. decorations of the main door). The 3D density (f) is higher where multiplicity and intersection angles are also higher.

The optimal threshold value of A_V and the feature weights have been assigned considering both photogrammetric principles and empirical tests (Section 4.4), such as the influence of each feature in the filtering step (Table 2), in terms of deleted 3D tie points and the effective improvement in the orientation results. The choice of assigning appropriate values for the feature weights is essential to avoid filtering out an excessive number of 3D points.

3.3 Filtering and new EO computation

The aggregated value A_V, with the optimal threshold and weights, is then used to discard or keep a 3D point of the sparse point cloud (and the contributing tie points). Low-quality points are automatically filtered out and a new bundle adjustment is re-run in order to compute new orientation parameters and, successively, the final dense point cloud.

4. EXPERIMENTS AND DISCUSSION

4.1 Test case

Although various case studies were performed to evaluate the developed method, hereafter only the Modena dataset (Figure 1) is reported. It is composed of 138 images of the Modena Cathedral in Italy (82 terrestrial and 56 from an UAV platform). The terrestrial images were acquired with a Nikon D750 mounting a 28 mm lens (pixel size of 5.98 µm). The UAV was equipped with a Canon EOS 600D (focal length of 28mm, pixel size of 4.4 µm).

4.2 Processing environments for features extraction

The open-source suite MicMac was chosen for performing the image processing and 3D reconstruction. MicMac allows to implement and share new tools and algorithms to enrich the photogrammetric workflow and to deepen the qualitative evaluation of the results. It already provides the possibility to interact with the processing pipeline and to access some qualitative indicators (e.g. bundle block adjustment residuals, covariance matrices, etc.) for the evaluation of the results.
We have enriched the MicMac suite with a new tool, named "Grappa", developed to extract the following quality features (Figure 4 a-b-c):

- mean re-projection error of the points in the sparse cloud;
- multiplicity values (the number of images where each point is visible and measured);
- max intersection angles among image observations generating a 3D point.

Grappa is run after the bundle adjustment step and it extracts the quality features by processing orientation results, 3D tie points coordinates and RGB values.
On the other hand, the difference of Normals (DoN) feature is computed with PCL library (version 1.8). The DoN (Ioannou et al., 2012) is a multi-scale approach used to process unorganized point clouds. For each 3D point p, two normal vectors are estimated with different support radii (r_1, r_2). The estimation of the point normal vector is generally based on support radii or considering a fixed number of neighbours. Depending on the chosen radius and the underlying geometry of the surface, the normal vector is clearly different. The DoN $\Delta\hat{n}$ is the normalized vector obtained as difference between these point normals estimated with different radii:

$$\Delta\hat{n}\,(\mathrm{p}, r_1, r_2) = \frac{\hat{n}\,(p, r_1) - \hat{n}\,(p, r_2)}{2} \qquad \text{(Eq. 3)}$$

where $r_1 < r_2$ and $\hat{n}\,(\mathrm{p}, r)$ is the surface normal estimated for the point p, considering the support radius r. Several pairs of support radii values were tested (Figure 5): small support radii highlight small changes in the surface structure of the scene whereas bigger radii values are most effective in highlighting surface changes in large scale structures. In our experiments, considering the normal surface estimation on small and complex surface details, the radii 0.03-0.3 were chosen as good values for the successive filtering step.
Finally, the point cloud density is computed scripting a CloudCompare plugin. The sparse point cloud density is used to identify points with low density values in their neighborhood. Such points could be low-quality 3D points. Density values are computed considering the number of neighbors in a sphere of a chosen radius r or determining the distance to the nearest neighbours. In our experiments a sphere of radius 0.05 m (corresponding to the lowest spatial resolution of the sparse cloud reconstruction) helped to recognize low quality points and reduce noise in the data, particularly in areas like roof, windows, etc.

4.3 Tests on joint and independent orientation procedures

Preliminary photogrammetric tests were performed to compare the two orientation strategies discussed in Section 3 (i.e. independent and joint) in terms of number of 3D tie point extracted, reprojection error, multiplicity values, points in the dense clouds and RMS Error (RMSE) on Check Points (CPs) collimated in the images (Table 1). Results show that both independent (followed by a merging of the dense point clouds) and joint orientation of terrestrial and UAV images lead to similar results, although a visual inspection of the obtained 3D result shows less noisy dense reconstructions adopting the joint processing procedure.

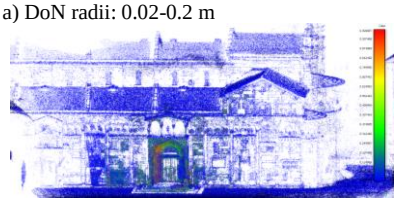
4.4 Thresholds, weights and point filtering

The aggregated value A_v in Eq. 2 is calculated for each 3D tie point and it is a sum of the normalized and weighted quality feature values. Therefore, choosing the appropriate weight for each feature and the threshold of A_v is a crucial step in the filtering procedure, which can greatly affect the aggregated value calculation.

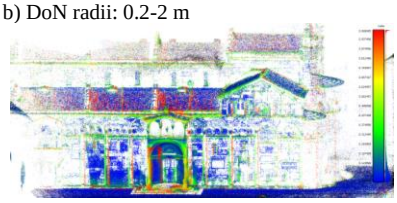
	Orientation of only terrestrial images	Orientation of only UAV images	Independent orientation and combined clouds	Joint orientation of terrestrial and UAV images
Computed 3D tie points	696K	1 Mil	1. 7 Mil	1.6 Mil
# pts in the dense cloud	13.3 Mil	10.5 Mil	18.4 Mil	18.9 Mil
Reprojection error (pix)	0.791	0.672	0.742	0.718
# pts in 2 images	~ 63%	~52%	~ 56%	~ 57%
# pts in 3 images	~ 17%	~19%	~ 19%	~ 18%
# pts in 4 or more images	~ 7%	~ 10%	~ 9%	~ 8%
RMSE on CPs (cm)	1.8	2.2	1.8	1.7

Table 1: Results of the orientation tests on the Cathedral images run with the independent and joint datasets. Number of reconstructed 3D points (sparse and dense clouds), reprojection error and multiplicity values with two different approaches: jointly-orientated and independently-oriented-and-combined.

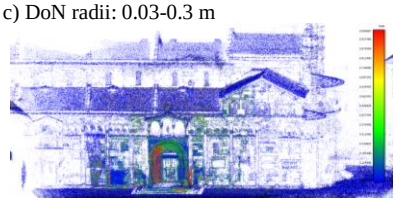
a) DoN radii: 0.02-0.2 m



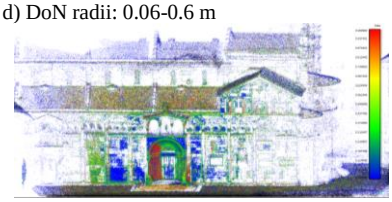
b) DoN radii: 0.2-2 m



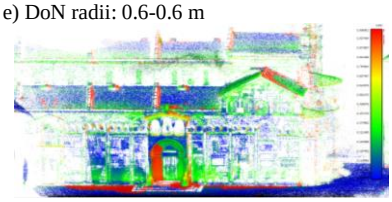
c) DoN radii: 0.03-0.3 m



d) DoN radii: 0.06-0.6 m



e) DoN radii: 0.6-0.6 m



f) DoN radii: 2-20 m

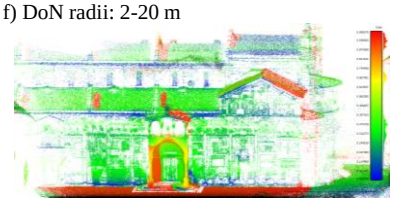
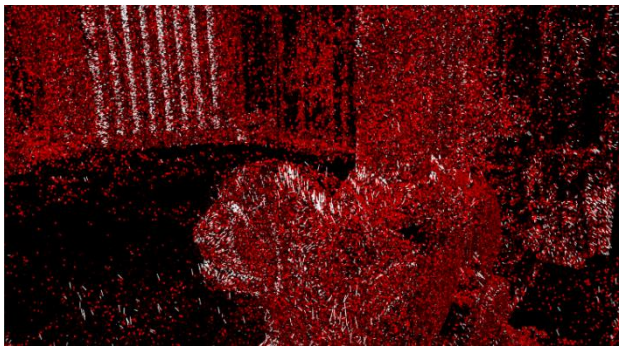


Figure 5: Vector magnitude of the Difference of Normals (DoN) operator using several combinations of radii. The DoN operator produces high response in areas with significative surface changes in the given radii range. For small radii values (i.e. case –*b*), greater variations are expected in fine-detailed areas. Increasing the radii values (i.e. case -*f*), DoN detects higher changes in the large architectonical parts (roof, façade,..).



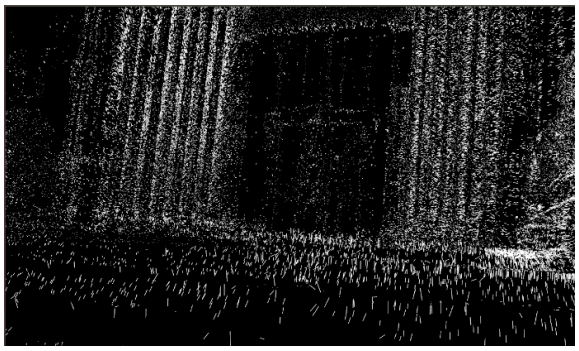
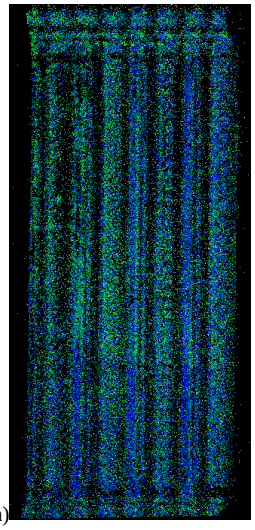
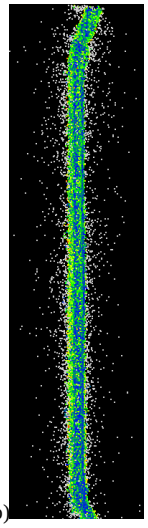


Figure 6: Two different visualizations of the DoN vectors (radii 0.03-0.3 m) computed in the area of the main entrance of the Cathedral: fine-detailed parts (architectural ornaments – left and steps - right) present vectors (white segments) with strong magnitude. On the left, the visualization of the sparse point cloud (red points) and the DoN vectors (white segments). On the right, only the DoN vectors are showed.

Feature	Threshold	Deleted 3D points	Variation of features' mean values		
			Reproj. error	Multiplicity	Intersec. angle
Reprojection error (px)	1	109K (~ 4%)	**~ -10%**	~ -1%	~ -2%
Multiplicity	3	1.6 Mil (~ 62%)	~ -34%	**~ 61%**	~ 57%
Max intersection angle (degree)	10	1.2 Mil (~ 47%)	~ -16%	~ 27%	**~ 58%**
DoN (normalized degrees)	0.001	930K (~ 35%)	~ -2%	~ -4%	~ 17%
Density (numb. of neighbors)	2	320K (~ 12%)	~ -1%	~ 1%	~7%

Table 2: Variation of the employed quality features when filtering is performed using a single feature. For each feature, the filtering threshold, the number of removed 3D tie points and the variation of the quality features after the points removal are shown. Negative variations represent an improvement only in the reprojection error case.

a) 

b) 

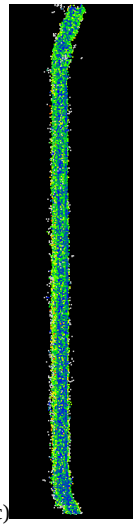
c) 

Figure 7: Filtering results with the DoN feature on an area full of architectural ornaments (a). Sideview before (b) and after (c) filtering.

Firtsly, the influence of quality features' values was investigated when a single feature filtering was applied (Table 2). Qualitative evaluations for DoN and density values variation (respectively in terms of surface changing variation response and noise reduction) were also performed (Figure 6-7). This allowed to quantify the effect of each feature in the filtering process and to find the optimal value of A_v. After various empirical tests, the following features' thresholds were chosen: the reprojection error should be smaller than 1 px, the multiplicity should be at least 3 (images) and the intersection angle should be greater than 10°, the DoN operator should be 0.001 and the number of neighbors for the cloud's density should be at least 2.

Considering the effect of the single feature's filtering, in terms of removed tie points and general improvements of the internal quality values, we realized that the filtering could not rely only on the single features as highly correlated among them. Therefore other tests were performed using, instead, the aggregated filtering approach (Eq. 2) with ad-hoc weights.

Table 3 reports the variation of three quality features when the filtering procedure considers aggregated and weighted values. Three different combination of feature weights were used (Figure

Case	Weight					Threshold A_v	Deleted points	Variation of features' mean values		
	Reproj. error	Multipl.	Inters. angle	DoN	Density			Reproj. error	Multiplicity	Inters. angle
1	10%	40%	30%	10%	10%	0.941	53K (2%)	~ -4%	~1%	~1%
2	10%	40%	30%	10%	10%	0.847 (90%)	1.8 Mil. (68%)	~ -7%	~45%	~66%
3	**10%**	**40%**	**30%**	**10%**	**10%**	**0.894 (95%)**	**809K (30%)**	**~ -10%**	**~13%**	**~26%**

Table 3: Weights, aggregated filtering thresholds, deleted 3D tie points and internal quality values improvements for the best weights' combination (highlighted box in Figure 8) applied to filter the sparse point cloud. The first case considers the A_v threshold derived by photogrammetric principles and empirical experiments (Section 4.4). Case 2 and 3 investigate a lower A_v threshold (90% and 95%). Case 3 had the most balanced ratio between number of deleted points and improvements of the considered quality features (lower reprojection error and higher values of multiplicity and maximum angle).

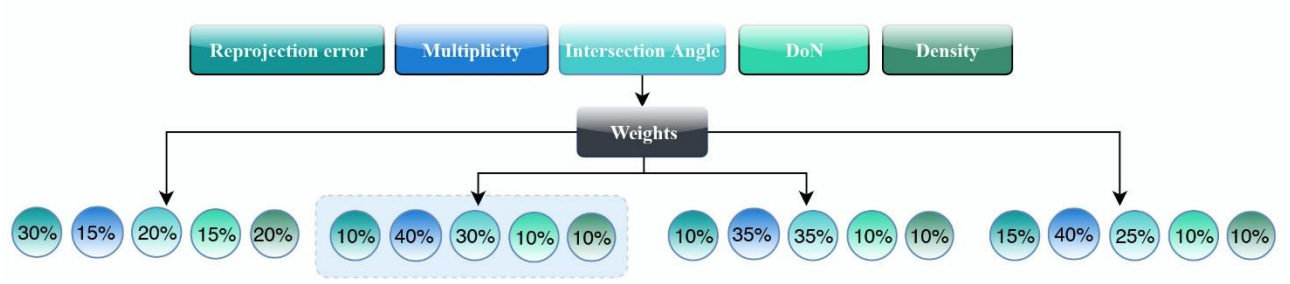


Figure 8: Graphical representation of the tested weights combinations. The highlighted box shows the combination yielding the best results in terms of quality feature improvements (see also Table 3).

8). For each combination, three different thresholds of A_v are used: the optimal value (computed using the before-mentioned photogrammetric principles) and two lower thresholds (respectively 90% and the 95%). In particular, Case 3 highlights the best weights combination, in terms of acceptable deleted points and general improvements of the considered quality values.

4.5 The filtering step

Given the optimal thresholds and weights, the developed automatic Python filtering tool handles the removal of the 3D tie points - and the corresponding 2D tie points - not fulfilling the given conditions. 3D tie points with an aggregated value above the optimal threshold are automatically removed. A new MicMac workspace is then generated for the computation of a new set of orientation parameters and a new dense 3D point cloud.

4.6 Results evaluation using external checks

Beside validating the proposed methodology using internal quality values (Table 3), external checks were also used to evaluate the quality of the newly derived dense point clouds. The ground truth is a point cloud of the Cathedral acquired with a Leica HDS7000. A seven parameters Helmert transformation was used to co-register the photogrammetric and ranging clouds. The evaluation was performed considering:

- RMS error (RMSE) on Check Points (CPs) (Table 4);
- RMSE on plane fitting (Figure 9 and Table 5);
- mean error and standard deviation values with a cloud-to-cloud distance on selected areas (Table 6).

RMSE on Check Points (cm)	
Original dense cloud	3.5
Filtered dense cloud	2.6
Average variation	**~ 22%**

Table 4: RMSE on some natural CPs measured on the dense point clouds and variation (improvement) after the filtering procedure.

Plane fitting RMSE (cm)			
Plane #	Original dense cloud	Filtered dense cloud	Variation
1	1.44	1.40	~ 4%
2	0.60	0.56	~ 11%
3	1.50	0.70	~ 52%
4	1.60	1.40	~ 16%
5	0.40	0.41	~ -2.00%
Average	**1.10**	**0.90**	**~ 17%**

Table 5: Improvement of the average plane fitting RMSE on five selected planar areas, considering the original and the filtered dense point clouds.

Figure 9: The areas where the plane fitting analyses were performed.

Cloud to cloud distance on sub-areas (cm) - Average values			
	Mean	St.deviation	Variation
Original dense cloud	5,4	5,8	**~ 12%**
Filtered dense cloud	5,0	5,3	

Table 6: Average cloud-to-cloud distance variation.

A visual inspection of the newly computed dense point cloud (Figure 10) shows also a general improvement of the results with respect to the dense cloud generated with the orientation parameters before the filtering procedure.

5. CONCLUSIONS AND FUTURE WORKS

The paper presented a new workflow for improving the joint processing of UAV and terrestrial images. An extended photogrammetric pipeline, involving a filtering step, is used to remove bad 3D tie points in the sparse point cloud before computing the dense reconstruction. The filtering is based on quality features derived from the sparse point cloud. The method was evaluated analyzing internal quality checks, external reference data and qualitative visual analysis. Obtained results showed a general improvement of all the evaluated parameters on the tested dataset.
Further tests are planned to investigate the effectiveness and the reliability of the proposed pipeline on different datasets. Moreover, additional quality features (DoN combined with curvatures, flatness and roughness values), other normalization functions and different weights combinations will be investigated to verify the robustness of the filtering equation.

ACKNOWLEDGEMENTS

Authors are thankful to Michaël Gaudin (IGN France) for supporting the realization of the implemented methodology and SAL Engeneering (https://www.salengineering.it/) for acquiring the UAV images of the Cathedral.

REFERENCES

Bastonero, P., Donadio, E., Chiabrando, F., Spanò, A., 2014. Fusion of 3D models derived from TLS and image-based techniques for CH enhanced documentation. *ISPRS Annals of Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. II-5, pp. 73-80.

Bracci, F., Drauschke M., Kühne S., Márton Z.-C., 2018. Challenges in fusion of heterogeneous point clouds. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 42(2), pp. 155-162.

Gasparovic, M. and Malaric, I., 2012. Increase of readability and accuracy of 3D models using fusion of close-range photogrammetry and laser scanning. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 34(B5).

Gerke, M., Nex, F. C., and Jende, P. L. H., 2016. Co-registration of terrestrial and UAV-based images: experimental results. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 40(3/W4), pp. 11-18.

Ioannou Y., Taati B., Harrap R., Greenspan N., 2012. Difference of Normals as a multi-scale operator in unorganized point clouds. *Proc. 2nd Int. Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission* (3DIMPVT).

Khaleghi, B., Khamis, A., Karray, F., O., Razavi, S. N., 2013: Multisensor data fusion: A review of the state-of-the-art. *Journal on Information Fusion*, Vol.14(1), pp. 28-44.

Koch, T., Zhuo, X., Reinartz, P. and Fraundorfer, F., 2016. A new paradigm for matching UAV and aerial images. *ISPRS Annals of Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. III-3, pp. 83-90.

Kochi, N., Kitamura, K., Sasaki, T., Kaneko, S., 2012. 3D Modeling of Architecture by Edge-Matching and Integrating the Point Clouds of Laser Scanner and Those of Digital Camera. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 39(B5).

Luo, R., Chang, C. C., Lai, C. C., 2011. Multisensor fusion and integration: theories, applications, and its perspectives. *Sensors Journal*, Vol. 11(12), pp. 3122-3138.

Mandlburger, G., Wenzel, K., Spitzer, A., Haala, N., Glira, P., Pfeifer, N., 2017. Improved topographic models via concurrent airborne LiDAR and dense image matching. *ISPRS Annals of Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. IV-2/W4, pp. 259-266.

Mauro, M., Riemenschneider, H., Signoroni, A., Leonardi, R., Van Gool, L., 2014. A unified framework for content-aware view selection and planning through view importance. *Proc. BMVC.*

Petrasova, A., Mitasova, H., Petras, V. and Jeziorska, J., 2017. Fusion of high-resolution DEMs for water flow modeling. *Open Geospatial Data, Software and Standards*, Vol. 2(1), pp. 6.

Pomerleau, F., Colas, F. and Siegwart, R., 2015. A Review of Point cloud registration algorithms for mobile robotics. *Foundations and Trends in Robotics*, Vol. 4(1), pp. 1-104.

Ramos, M., M. and Remondino, F., 2015. Data fusion in Cultural Heritage – A Review. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 40(5/W7), pp. 359-363.

Remondino F., Nocerino E., Toschi I., Menna F., 2017. A critical review of automated photogrammetric processing of large datasets. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 42(2/W5), pp. 591-599.

Roncella, R., Re, C. and Forlani, G., 2011. Performance evaluation of a structure and motion strategy in architecture and cultural heritage. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol.38(5/W16), pp. 285-292.

Rupnik, E., Daakir, M., Deseilligny, M. P., 2017. MicMac–a free, open-source solution for photogrammetry. *Open Geospatial Data, Software and Standards*, Vol. 2(1), 14.

Serna, C. G., Pillay, R., Tremeau, A., 2015. Data fusion of objects using techniques such as laser scanning, structured light and photogrammetry for Cultural Heritage applications. *Computational Color Imaging*, pp. 208-224.

Tombari, F., Remondino, F., 2013. Feature-based automatic 3D registration for cultural heritage applications. Proc. IEEE Conference "Digital Heritage 2013", Vol. 1, pp. 55-62.

Wu, B., Xie, L., Hu, H., Zhu, Q., Yau, E., 2018. Integration of aerial oblique imagery and terrestrial imagery for optimized 3D modeling in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 139, pp. 119-132.

Zhang, J., 2010. Multi-source remote sensing data fusion: status and trends. *Int. Journal of Image and Data Fusion*, Vol. 1(1), pp. 5-24.

Zieher, T., Toschi, I., Remondino, F., Rutzinger, M., Kofler, Ch., Mejia-Aguilar, A., Schlögel, R., 2018. Sensor- and scene-guided integration of TLS and photogrammetric point clouds for landslide monitoring. *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. 42(2), pp. 1243-1250.

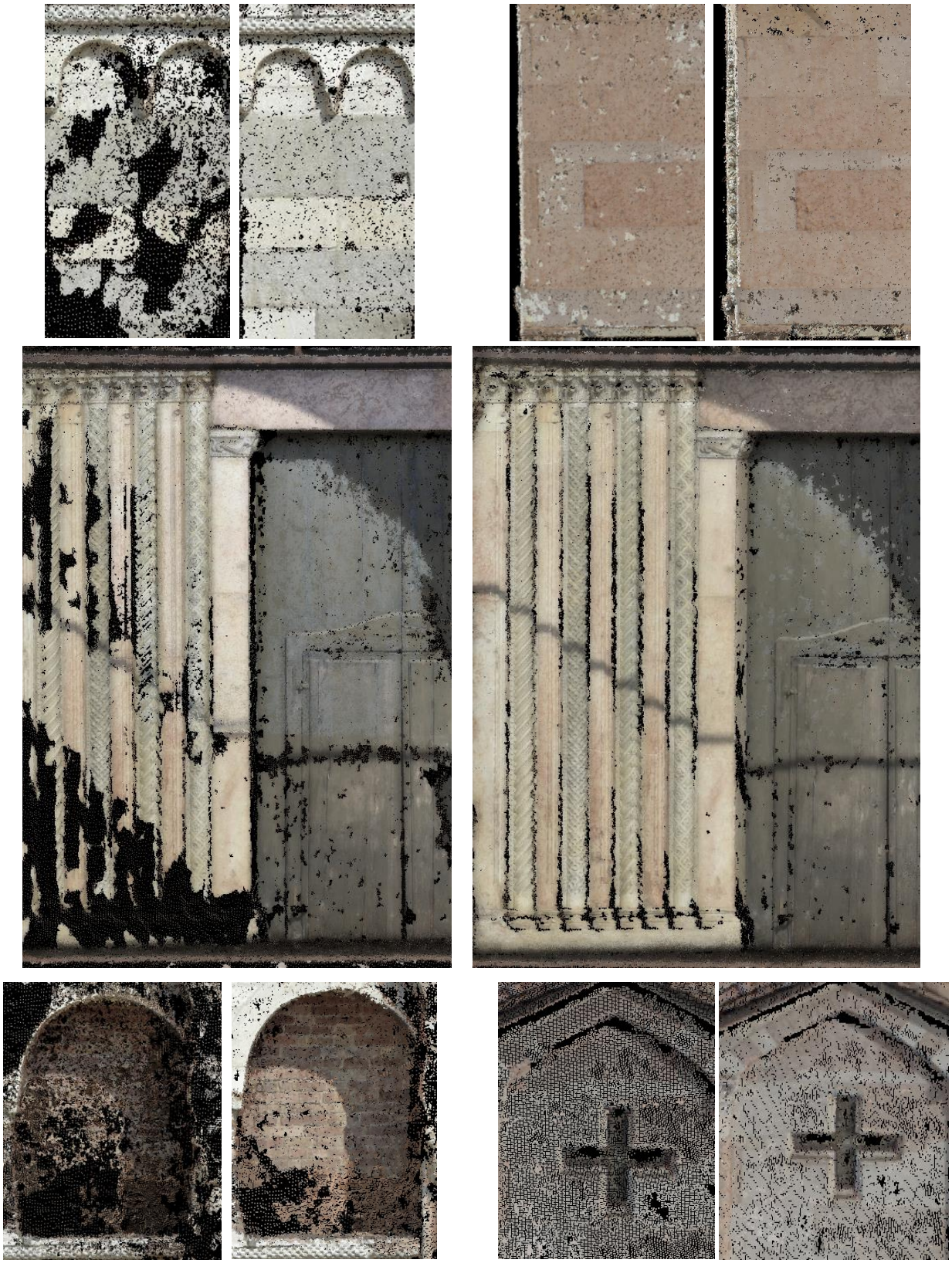


Figure 10: Qualitative evaluation of the dense point clouds derived with the standard photogrammetric pipeline (left images for each box) and the clouds derived after the proposed filtering method (right images). A visual comparison highlights clear improvements in the dense reconstruction (less holes, higher density, etc.).