AUTOMATIC ROOF OUTLINES RECONSTRUCTION FROM PHOTOGRAMMETRIC DSM

F. Nex , F. Remondino

3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
<franex, remondino>@fbk.eu, http://3dom.fbk.eu

Working Groups III/4

KEY WORDS: Matching, Automation, Aerial, Extraction, Point cloud

ABSTRACT:

The extraction of geometric and semantic information from image and range data is one of the main research topics. Between the different geomatics products, 3D city models have shown to be a valid instrument for several applications. As a consequence, the interest for automated solutions able to speed up and reduce the costs for 3D model generation is greatly increased. Image matching techniques can nowadays provide for dense and reliable point clouds, practically comparable to LiDAR ones in terms of accuracy and completeness. In this paper a methodology for the geometric reconstruction of roof outlines (eaves, ridges and pitches) from aerial images is presented. The approach keeps in count the fact the usually photogrammetrically derived point clouds and DSMs are more noisy with respect to LiDAR data. A data driven approach is used in order to keep the maximum flexibility and to achieve satisfying reconstructions with different typologies of buildings. Some tests and examples are reported showing the suitability of photogrammetric DSM for this topic and the performances of the developed algorithm in different operative conditions.

1. INTRODUCTION

The extraction of geometric and semantic information from image and range data is one of the main research topics in the geomatics community. Between the different products, 3D city models have shown to be a valid instrument for several applications such as solar radiation potential assessment, urban management and planning, land monitoring, pollutant diffusion, virtual tour, navigation, gaming, etc. As a consequence, the interest for automated solutions able to speed up and reduce the costs for 3D model generation is greatly increased (Haala and Kada, 2011).

In typical mapping and modelling applications, once a point cloud (usually several millions of points) has been extracted, only the first (and shortest) part of the work has been completed. It is afterward required to process them in order to extract metric information (such as shapes, surface normal vectors, dimensions, polylines, etc.) necessary to achieve the final product (3D model, drawing, etc.). In some way, the classification, segmentation, modelling and in general the "understanding" of an unstructured point cloud is the main challenge to be faced nowadays. In literature, several papers dealing with these topics, in particular with the 3D building modelling, have been already presented. Some years ago, only few automated procedures considered the possibility to use images (Paparoditis et al., 2001; Zhang, 2005; Paparoditis et al., 2006; Zebelin et al., 2006) as the information achieved by image matching were considered insufficient to obtain reliable results. Most of the researches were oriented towards LiDAR data (Rottensteiner and Briese, 2002; Habib et al., 2009; Sampath and Shan, 2009; Oude Elberink and Vosselman, 2011). Nowadays a growing number of research works relies on the integration of different data sources (Demir et al., 2009; Vallet et al. 2011) and in particular on the integration of range and image data, exploiting the complementary nature of LiDAR and images (Awrangjeb, et al. 2010; Habib et al., 2010; Nex and Remondino, 2011). But the actual availability of redundant multi-image information with and the improvement of automated image matching methods (Hirshmüller, 2008;

Lemarie, 2008; Haala, 2009; Hiep et al., 2009; Wolff, 2009; Gehrke et al., 2010; Leberl et al., 2010), allow the generation of 3D point clouds and 2.5D raster representations which in the past were only feasible with LiDAR techniques. Several commercial and open source solutions are also available for the production of very satisfactory geometric results (Fig. 1) exploiting the very high radiometric quality, the potentialities of GPU programming and the largely overlapping image blocks.
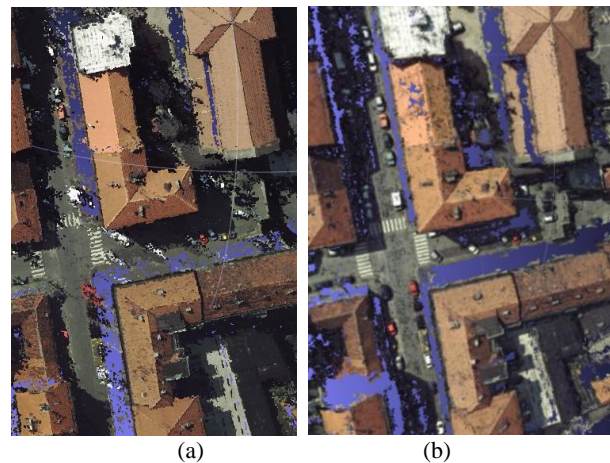

| (a) | (b) |

Figure 1: Examples of image matching results with open-source packages.

The paper presents an automated methodology for the geometric reconstruction of the main roof outlines (eaves, ridges and pitches) from dense point clouds automatically extracted from aerial images. Point clouds generated from image matching can be denser than LiDAR data. In theory, an image block with a GSD (Ground Sampling Distance) of 10 cm would allow the derivation of a point cloud with 100 points/m^2. A typical LiDAR flight for city-modeling applications is in the order of 15-20 points/m^2. The extraction of an higher number of object points allows discontinuities to be better defined and it is directly connected to the Level of Detail (LoD) that can be

achieved in the geometric modeling (Oude Elberink and Vosselman, 2011). However, photogrammetric point clouds and DSM are usually noisier than LiDAR data as they suffers from the radiometric image quality, image overlap, presence of shadows and object texture, as also underlined in (Vallet et al., 2011). The large image overlap can only partly improve the internal accuracy and the reliability of the results but several blunders can be still present in shadowed or almost occluded areas. But the higher number of details that can be detected (i.e. roof tiles) allows to better model surfaces (that are flat in lower resolution point clouds) even if this higher degree of detail is often interpreted as noise during the modelling.

In the literature, many approaches focusing on roof shapes extraction from elevation data have been presented, mainly based on prismatic shapes, point cloud segmentation, feature recognition or DSM simplification (Haala and Kada, 2011). Several commercial software devoted to man-made feature extraction and 3D reconstruction have been developed too. These approaches have been originally implemented based on LiDAR as input data, but several problems arise when photogrammetric DSM are used. Indeed only geometric information can be used in the modelling without any multi-echo pulses or intensity information. Thus the presence of blunders/noise and some gaps in the point cloud can lead to incorrect reconstructions (Fig. 2).
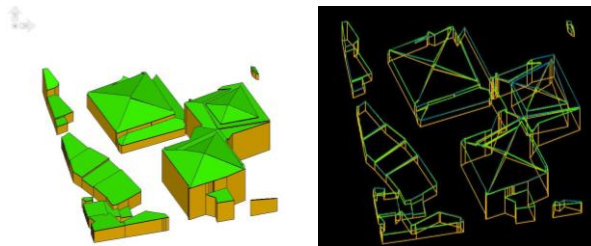

Figure 2: Example of automated 3D roof reconstruction using a commercial software and a photogrammetric point cloud. Some problems are visible on the reconstructed roofs.

The proposed algorithm works on photogrammetric DSMs and it fits in the segmentation methods keeping into consideration the aforementioned problems. The suitability of image-based data for automated building outlines reconstruction is thus reported, considering not only the eave detection but also the ridge, hips and valley extraction. Several iterative steps are performed in order to extract only reliable information for the roof reconstruction. On the other hand, the algorithm maintains the maximum flexibility in order to correctly reconstruct roof building components (Fig. 3) in a big variety of operative conditions. A data driven approach is implemented without any geometrical constraint (parallelism, orthogonality, etc.) in order to define a building model.
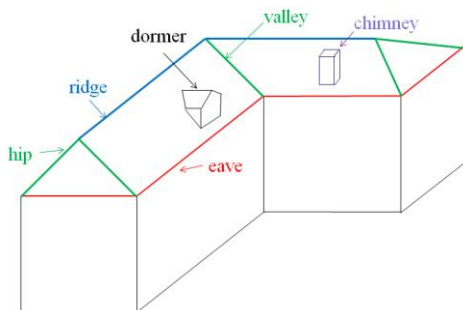

Figure 3: Roof components which should be reconstructed for the highest roof LODs.

In the next sections, the workflow will be described more in detail with examples over different test areas. Finally, conclusions and future developments will be discussed.

2. ALGORITHM OVERVIEW

The proposed algorithm processes image-based point clouds in order to extract geometric primitives useful for a more complete and detailed reconstruction of roof buildings. The entire methodology (Fig. 4) is divided in blocks with concatenated processing steps.
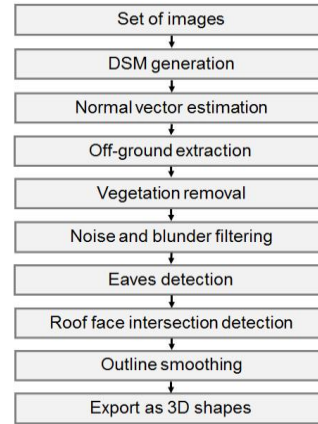

Figure 4. Workflow overview for automated roof shapes extraction from photogrammetric elevation data.

- **DSM generation.** The DSM generation is a fundamental step for the following results as the whole process depends on the quality of the point cloud. The open-source MicMac method (Paparoditis et al., 2006) is applied as, from our experience and after the comparison with other packages, it provides dense and accurate point clouds over urban areas without smoothing effects in proximity of building outlines. A very dense point cloud is strictly recommended (even 1 object point per pixel) in order to have a complete information all over the area (Fig. 5b). Large image overlaps are recommended to reduce occlusions and increase the reliability of the reconstruction while 16-bit resolution images only partially reduce the lack of texture and shadows problems.
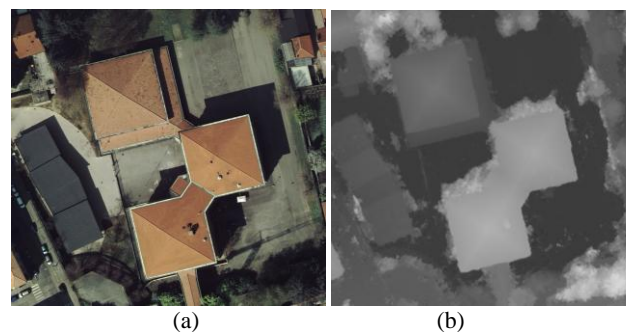

Figure 5: RGB image (a) and corresponding depth map (b).

- **Normal vector estimation.** Man-made objects on urban areas are usually characterized by local flat areas (roofs, roads, etc.) with reduced slopes (maximum 45°). Blunders are usually characterized by chaotic and rough depth variations: this is usually true for the results provided by several matching algorithms. Thus normal vectors on these areas should be almost vertical and regular when matching results are reliable.

On the other hand they suddenly vary in direction and they are almost horizontal on noisy areas. According to this, the normal vector, normally computed using 5x5 pixel patches, can give an indication of the local shape of the DSM. Fig. 6a shows the computed normal vector image on the same area of Fig. 5a: light grey areas indicates almost vertical normal vector directions whereas dark areas describes almost horizontal directions. Horizontal directions are only in correspondence of blunders and rough variations of the buildings (outlines and chimneys).
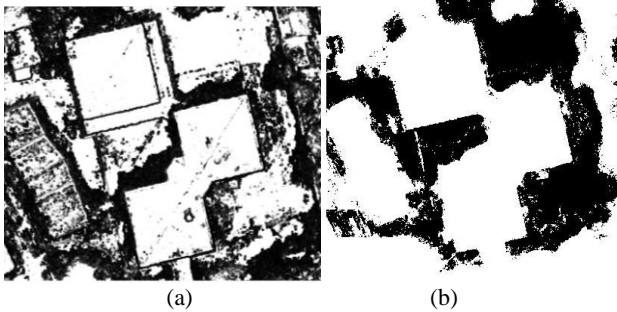

Figure 6: Normal vector image (a) and off-ground area (b).

- **Off-ground extraction.** The automatic off-ground extraction procedure assumes that the height of the ground is lower than the neighbouring non-ground points. The ground filtering is performed with an iterative regular grid filtering. This process considers three different problems: (i) the ground height variations over a big region patch, (ii) the presence of big dimensions buildings that can avoid to determine the correct ground height when too small DSM patches are considered and (iii) the presence of blunders or local noise can influence the determination of ground height value. For these reasons, the ground height is iteratively computed on different DSM patch dimensions, evaluating only almost vertical normal vector areas. In this way, the most representative value of the ground height is determined considering the minimum height value on the different DSM patches. Off-ground points are defined considering points higher than a defined threshold (2-4 m) with respect to the ground height value (Fig. 6b).

- **Vegetation removal.** The off-ground points comprehend buildings, trees and noisy data. The vegetation can be easily removed when NIR images are available, simply using the NDVI value. When only RGB images are available, a combination of height variations and colour information is considered. The achievable results are usually incomplete as some areas are not deleted and thus a morphological filter to remove little areas with irregular shapes is afterwards applied.

- **Noise and blunder filtering.** The roof shapes are usually correctly and completely described by the photogrammetric DSM but some noisy points are still visible around it (as shown in Fig. 7a) approximately at the same height of the buildings. Height differences between adjacent points and normal vector variations (black arrows in Fig. 7) of points in correspondence of building boundaries are kept in count in a cost function. In this way blunders close to building outlines are iteratively eroded (red points in Fig. 7b). This process sometimes exceeds in the points removal and an iterative dilation step is then performed. Points located in proximity of building borders are used as seed points (blue points in Fig. 7). For each point, both height variations and local planarity of seed points are considered in a cost function in order to evaluate its suitability to belong to the roof.
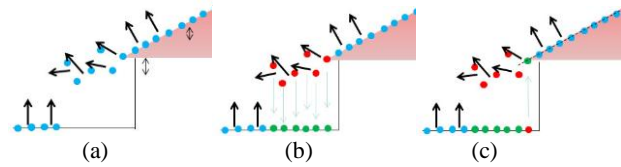

Figure 7: Iterative erosion and dilation process.

At each iteration new points are aggregated on the roof until they cannot be considered suitable (Fig. 7c). The process usually stops to add points in 2-3 iterations. Finally, regions with surface lower than a threshold and with very irregular shape are removed in order to delete not still removed vegetation areas. The filtered DSM achieved by this process is shown in Fig. 8a. Deleted points are replaced by their correspondent ground height, i.e. a mean ground value over a region of few meters.

- **Eaves detection.** Each building can be composed by different roofs with different heights that must be partitioned in order to better reconstruct the entire roof area. The building is thus divided in different volumetric elements (sub-footprint) only when big height variations occurs and completely separated volumes can be defined. Regions of the roof that are connected by a side with other faces (i.e. dormers) are not considered separate sub-footprints. Little separate regions, such as chimneys, are grouped to the roof as they are difficult to be correctly modelled on noisy areas as they could be seen as blunders: for this reason, it was decided to ignore them. Each building's eave is initially stored considering the boundary pixel coordinates of each sub-footprint (Fig. 8b). A first decimation of boundary is performed in order to reduce the number of points and simplify the following smoothing process.
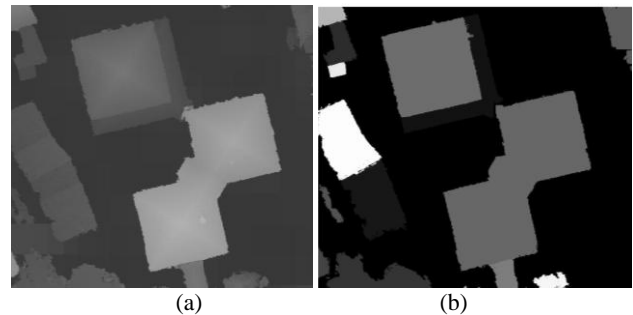

Figure 8: Filtered depth map (a) and corresponding sub-footprints of the buildings (b).

- **Roof face intersection detection.** After the building's eave determination, the ridge, hips and valleys positions has to be defined. The different faces of the roof have to be determined in order to complete the roof shapes reconstruction. Several algorithms already presented (Haala and Kada, 2011) can be adopted: the plane fitting (RANSAC, etc.), region growing algorithms, 3D Hough transform, curvature estimation, etc. In this work a region growing approach was adopted, as it is more robust to noisy data. In particular, this algorithm keeps in count the local maximum gradient on the roof (8 direction are considered), the presence of rough depth variations and the normal vector direction in order to perform a first classification of points. As data are noisy, an over-segmentation of the area is usually determined (Fig. 9a). To solve this problem, the main gradients of the buildings are estimated over each area and the process is repeated. The main gradients of the roof are chosen considering the number of pixels (their area) and their extension on the roof: points with the same orientation spread over the whole roof are usually due to noise and they cannot be

considered a main representative orientation. On the other hand, frequent orientation on a defined area of the building indicates a representative orientation of one roof face. When the principal orientations are defined, pixel values are constrained to belong to the closer principal orientation in the second iteration (Fig. 9b). This process is still a critical aspect but it provides satisfying results when the roof faces are sufficiently wide to be well modelled by the DSM. The roof face intersections are then defined selecting the boundary of each face. As in the eave detection step, chimneys on the roof are ignored and ridges are constrained to be regular and linear when chimneys are close to the rooftop.
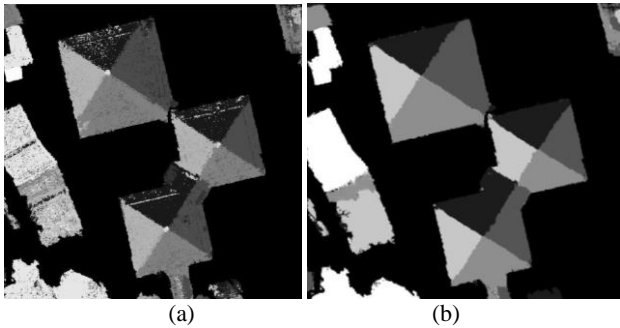

(a)   (b)

Figure 9. Roof faces determination: aggregation procedure.

- **Outline smoothing and footprint generation.** Points provided by image matching algorithms can be randomly noisy. Thus the extracted outlines can also affected by some noise and they cannot be directly used in the roof modelling production. For this reason, a smoothing is needed in order to define a regular shape of the object, easing the roof outlines in set of lines and curves. The great majority of roof outlines can be mainly classified in sets of lines and (more rarely) in second order curves. Therefore, each edge must be split in different basic entities that describe its linear or curved parts separately. Each separate basic entity is then simplified in lines and curves fitting the dominant point information with a robust least square (LS) approach (Nex, 2010). These lines are finally merged together in hierarchical way to reconstruct the geometry of the roof in a whole dataset. The reliability of the fitted lines is defined by residuals of the LS which control the outline displacement to have the extremes coincident. Then, the borders of the hips are moved in order to coincide to the eave corners when the displacement is lower than a threshold.

- **Edge exporting.** The automatically extracted roof edges are exported in CAD as a set of 3D polylines/shapes in order to give a good preliminary idea of the achieved results (Fig. 10).
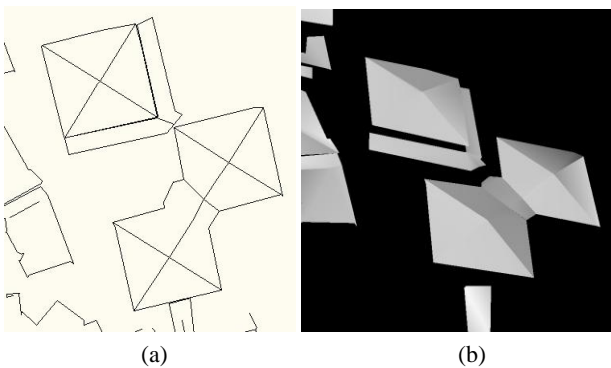

(a)   (b)

Figure 10. Example of exported the 3D shapes seen from above (a) and in an oblique view (b).

3. TEST RESULTS

In the following, some results of the developed methodology are reported. The tests were performed on dense urban areas over the city of Vaihingen (Germany) and Torino (Italy).

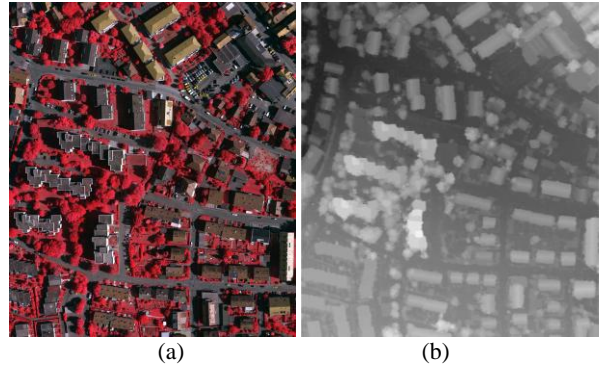## 3.1 Vaihingen

(a)   (b)

Figure 11. One image of the Vaihingen block (a) and the generated DSM shown as depth map (b).

This dataset belongs to the test area project of ISPRS WG III/4 "Urban Classification and 3D Building Reconstruction" (http://www.commission3.isprs.org/wg4/).
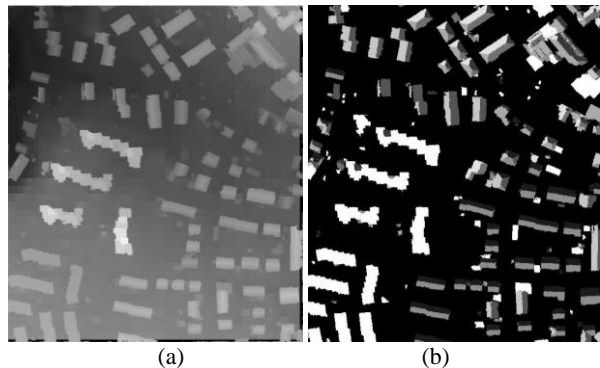

(a)   (b)

Figure 12. DSM after the filtering process (a) and the detected roof shapes and faces (b).

Images, provided as pan-sharpened colour infrared images, were acquired with a DMC camera over the city of Vaihingen (Germany) with 80/70% overlap. They feature 8 cm GSD and a radiometric resolution of 11 bits.

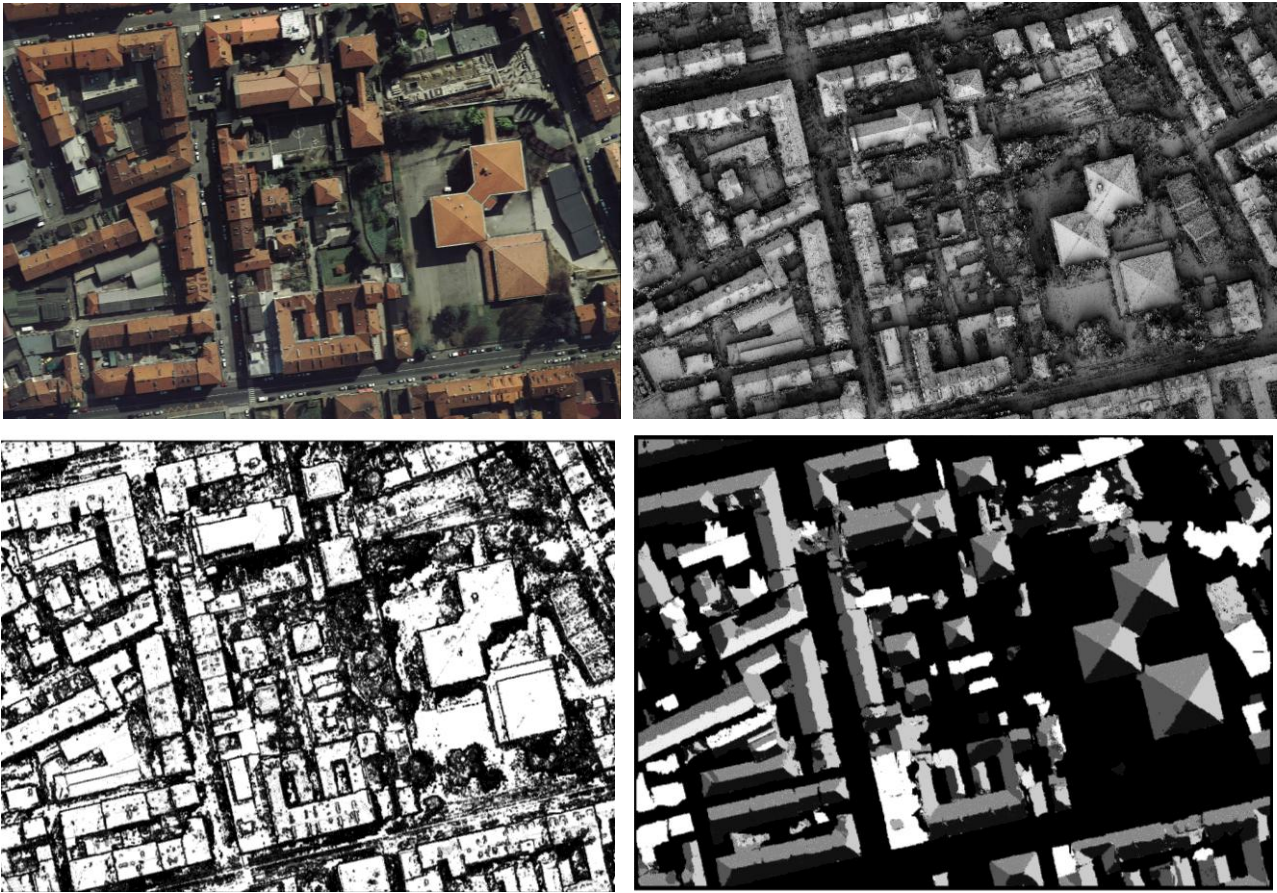Figure 13. 3D shapes of the test area exported in CAD.

Figure 14: Original DMC image over Torino (top-left), derived DSM shown in shaded mode (top-right), normal vector (bottom-left) and roof shapes detection results (bottom-right).

Unfortunately the images were acquired in the early morning (8 a.m.) and very long shadows influenced the quality of the extracted DSM (Fig. 11b). Noisy areas in correspondence of shadows had to be removed with some filtering, followed by the off-ground area identification (based on the NDVI index) and the detection of the correct building outlines (Fig. 12a). Finally the buildings were classified in sub-footprints and the ridges extracted as shown in Fig. 12b where each colour represents a different face orientation, whereas white areas refers to flat roofs.

The extracted outlines were finally smoothed in set of polylines and exported as 3D shapes in CAD format for their final visualization (Fig. 13). The results are quite satisfactory also due to the availability of the NIR band which allows the use of NDVI index to filter out the vegetation.

3.2 Torino

The dataset contains 6 aerial RGB images (DMC camera, GSD=12 cm) over a urban area of Torino (Italy). The test area (ca 0.5x0.5 km) is characterized by several high buildings, trees and variation of the ground height. Some blunders are present in the generated DSM (Fig. 14) around several buildings, but the developed methodology could face these problems.

After the filtering process (normal vectors are shown in Fig. 14 bottom-left), the DSM was improved by deleting the residual noise and the roof faces were correctly detected: the achieved result is shown in Fig. 14 bottom-right.

Anyway, some buildings regions are missing: this problem is usually concentrated on noisy areas, especially on buildings lower than surrounding ones. The roof face were correctly detected and the building outline extraction allowed to evaluate more in detail the reliability of the results. Big and higher building roofs were correctly reconstructed, while smaller man-made entities were partially missing.

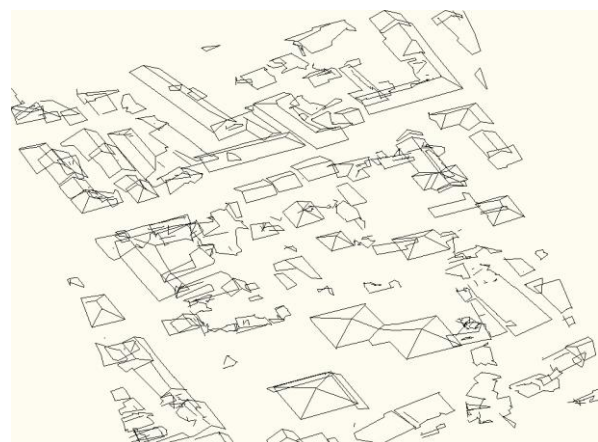Finally, the outlines and the ridges were smoothed and exported in CAD format as shown in Fig. 15.

Figure 15. 3D shapes of the test area exported in CAD. Small polylines associated to small man-made structures are also present.

4. CONCLUSIONS AND FUTURE DEVELOPMENTS

In this paper an algorithm for the automated extraction of roof building shapes from photogrammetric DSM was presented. The presented algorithm keeps in count the difference between LiDAR and photogrammetric DSM, showing the suitability of the latter data for roof shapes extraction too. Photogrammetric DSM could became a valid alternative to LiDAR data in the building extraction, thanks to their higher density.

The proposed method is data driven: the approach is able to process different typologies of buildings, but the results are strongly influenced by DSM quality which directly depends on the image quality. In general, very high overlap image blocks with reduced shadows are strictly recommended to extract very accurate and dense point clouds. The performed tests have demonstrated how different can be the quality of the achieved results as a function of the image quality and available bands. The algorithm produces the correct reconstruction in most of the cases, when the input data is smooth and complete. Anyway, some problems can be found when very complicated building geometries have to be reconstructed (i.e. industrial sheds or old city centres). Moreover the noise and blunder filtering step can delete wrong points, in particular when they are limited to small areas. On the other hand, some holes in the DSM can be generated when very noisy areas are analysed.

Further investigations will be performed in order to increase the performances of the algorithm in these areas too. Then, symmetries, parallel walls and regular shapes in general are very difficult to be directly achieved from data: several constraints should be imposed in order to achieve a more effective solution. In particular the possibility to integrate image information to the DSM information will be investigated in order to improve the completeness and reliability of the approach. Several tests over new areas will be performed in order to evaluate the performances of the developed methodology and evaluate the geometrical accuracy of this technique. For sure the potentiality of photogrammetrically derived DSMs is enormous and such data are getting a reliable alternative to LiDAR point clouds.

REFERENCES

Awrangjeb, M., Ravanbakhsh, M., Fraser, C.S., 2010: Automatic detection of residential buildings using LiDAR data and multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 65(5), pp. 457-467.

Demir, N., Poli, D., Baltsavias, E., 2009: Detection of buildings at airport sites using images & LiDAR data and a combination of various methods. *IAPRS&SIS*, Vol. 38(3/W4), pp. 71-77.

Gehrke, S., Morin, K., Downey, M., Boehrer, N., Fuchs, T., 2010: Semi-global matching: an alternative to LiDAR for DSM generation?, *IAPRS&SIS*, Vol. 38(1).

Haala, N., Kada, M., 2010: An update on automatic 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, pp. 570-580.

Haala, N., 2009: Comeback of digital image matching. Proc. *PHOWO* 2009, pp. 289-301.

Habib, A.F., Zhai, R., Kim, C., 2010: Generation of complex polyhedral building models by integrating stereo aerial imagery and LiDAR data. *PE&RS*,, Vol. 76(5), pp. 609-623.

Hiep, V.,H., Keriven R., Labatut, P., Pons J-P., 2009. Towards high resolution multi-view stereo. Proc. *Computer Vision and Pattern Recognition*, pp. 1430-1437.

Hirschmüller, H., 2008: Stereo processing by semi-global matching and mutual information. *IEEE Trans. on PAMI*, Vol. 30(2), pp. 328-341.

Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S., Wiechert, A., 2010: Point clouds: Lidar vs 3D vision. *PE&RS*, 76(10), pp. 1123-1134.

Lemaire, C., 2008: Aspects of the DSM production with high resolution images. *IAPRS&SIS*, Vol. 37(B4), pp. 1143-1146.

Nex, F., 2010: Multi-image matching and LiDAR data new integration approach, PhD Thesis, Politecnico di Torino, Torino.

Oude Elberink, S., Vosselmann, G., 2011: Quality analysis on 3D building models reconstructed from airborne laser scanning data. ISPRS *Journal of Photogrammetry and Remote Sensing*, Vol. 66(2), pp. 157-165.

Paparoditis, N., Souchon, J-P. Martinoty, G., Pierrot-Deseilligny, M., 2006: High-end aerial digital cameras and their impact on the automation and quality of the production workflow. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 60(6), pp. 400-412.

Rottensteiner, F., Briese, F., 2002: A new method for building extraction in urban areas from high-resolution LiDAR data. *IAPRS&SIS*, Vol.34(B3-A), Graz, Austria, pp. 295-301.

Sampath, A., J. Shan, 2008: Building Reconstruction from Airborne LiDAR Data Based on Clustering Analysis. *IAPRS&SIS*, Vol. 37(B3a), Beijing, China.

Vallet, B., Pierrot-Deseilligny, M., Boldo, D., Brédif, M., 2011: Building footprint database improvement for 3D reconstruction: A split and merge approach and its evaluation. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 66(5), pp. 732-742.

Zebedin, L., Klaus, A., Gruber-Geymayer, B., Karner, K., 2006: Towards 3D map generation from digital aerial images. *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 60(6), pp. 413-427.

Wolff, K., 2009: DGPF project: evaluation of digital photogrammetric aerial based imaging systems- generation of digital surface models. *IAPRS&SIS*, Vol. 38(1-4-7/W5), Hannover, Germany (CD-Rom).

ACKNOWLEDGEMENTS