

## Segmentation of UAV-based images incorporating 3D point cloud information

A.Vetrivel\*, M.Gerke, N.Kerle, G.Vosselman

University of Twente, Faculty of Geo-Information Science and Earth Observation (ITC), P.O. Box 217, 7500 AE, Enschede,  
The Netherlands  
[a.vetrivel@utwente.nl](mailto:a.vetrivel@utwente.nl)

**KEY WORDS:** Image segmentation, Building detection, UAV images, 3D point cloud, Texture, Region growing

### ABSTRACT:

Numerous applications related to urban scene analysis demand automatic recognition of buildings and distinct sub-elements. For example, if LiDAR data is available, only 3D information could be leveraged for the segmentation. However, this poses several risks, for instance, the in-plane objects cannot be distinguished from their surroundings. On the other hand, if only image based segmentation is performed, the geometric features (e.g., normal orientation, planarity) are not readily available. This renders the task of detecting the distinct sub-elements of the building with similar radiometric characteristic infeasible. In this paper the individual sub-elements of buildings are recognized through sub-segmentation of the building using geometric and radiometric characteristics jointly. 3D points generated from Unmanned Aerial Vehicle (UAV) images are used for inferring the geometric characteristics of roofs and facades of the building. However, the image-based 3D points are noisy, error prone and often contain gaps. Hence the segmentation in 3D space is not appropriate. Therefore, we propose to perform segmentation in image space using geometric features from the 3D point cloud along with the radiometric features. The initial detection of buildings in 3D point cloud is followed by the segmentation in image space using the region growing approach by utilizing various radiometric and 3D point cloud features. The developed method was tested using two data sets obtained with UAV images with a ground resolution of around 1-2 cm. The developed method accurately segmented most of the building elements when compared to the plane-based segmentation using 3D point cloud alone.

### 1. INTRODUCTION AND RELATED WORK

Automatic detection of individual building and recognition of its distinct sub-elements from remote sensing data are crucial for many applications including 3D building modelling, building level damage assessment and other urban related studies (Dong and Shan, 2013; Sun and Salvaggio, 2013). Generally, the buildings and its elements possess unique geometric characteristics. Hence, the 3D geometric features are being used as the fundamental information in building detection and categorisation of its sub-elements (Rottensteiner et al., 2014; Xiong et al., 2013). 3D point clouds are well suited to infer the geometric characteristics of the objects. Particularly, the multi-view airborne oblique images are a suitable source to generate 3D point cloud for building analysis as they can provide information of both the roofs and facades of the building (Liu and Guo, 2014). Unmanned Aerial Vehicles (UAVs) are attractive platforms which can capture the images with suitable characteristics such as multi-view, high overlap and very high resolution to generate very dense 3D point cloud in minimal time and cost (Colomina and Molina, 2014).

Generally, the building detection process from 3D point clouds has been carried out through identifying planar segments as most elements of general buildings are planar surfaces (Dorninger and Pfeifer, 2008). Planar segments with its geometric features could help to detect and delineate buildings in the scene. However, an accurate segmentation of individual elements of the building is not always feasible, especially with the geometric features from image-based 3D point cloud. This is

due to various reasons such as 1) presence of low-textured planar surfaces might lead to sparse 3D point cloud generation with significant gaps. In such case, a single planar segment might get fragmented into multiple small segments, or even partly missed, leading to an inaccurate segmentation; 2) Outliers or random errors which are inherent in image based 3D point clouds, especially when the image block configuration is not optimal might also leads to artefacts or inaccurate segmentation of building elements (Rupnik et al., 2014); 3) Regions affected by poor visibility (e.g., only visible in single images due to occlusions) will have no 3D points and those areas cannot be segmented; 4) 3D points belonging to non-planar objects will not be segmented by plane-based methods and it is difficult to recognize the complex objects even using other methods such as model-driven approach from sparse and erroneous 3D point clouds (Xiong et al., 2014); 5) Objects that share the same plane geometry, e.g., windows in the roof and façade plane, might not get segmented as individual entity, hence leading to under segmentation.

The segmentation based on radiometric features alone might delineate the building regions that possess similar spectral or textural characteristics. However, elements of different category with similar spectral characteristics cannot be differentiated, e.g., roof and façade of the building with same surface characteristics and colour might be segmented as a single element. Also the segments found based on spectral features cannot be categorised into roofs, facades, etc., without inferring its geometric characteristics. Hence, it is obvious that both geometric and spectral features are important for an accurate segmentation and recognition of distinct elements of the building.

---

\* Corresponding author.

Many studies used radiometric features such as colour along with geometric features and shape descriptors for recognition of objects in 3D point clouds through segmentation (Aijazi et al., 2013; Strom et al., 2010). However, the image-based 3D point cloud might be erroneous and incomplete with missing 3D points for some regions. Hence, performing the segmentation in image space by utilizing the geometric information from 3D point cloud could be an alternative strategy.

Previously many studies have been reported for image segmentation by using the combination of 2D radiometric and 3D geometric features e.g., segmentation of depth images (RGB-D) (Mirante et al., 2011; Yin and Kong, 2013). The surface normal, gradient of depth, and residuals of plane fitting are the widely used geometric features in depth image segmentation (Enjarini and Graser, 2012; Hulik et al., 2012). Spectral and spatial features such as colour, texture, edges and shape are widely used image-based features for segmentation (Tian, 2013). Among them texture features from GLCM are often reported as key features to infer the radiometric characteristics of the surface (Rampun et al., 2013). Numerous segmentation approaches are used in practice such as region-based approach (e.g., region growing, split and merge), clustering-based approach (e.g., k-means, mean shift), and graph-based approach (e.g., graph-cut) (Boykov and Funka-Lea, 2006; Narkhede, 2013). However, the choice of segmentation approach depends on the application and kind of features available for segmentation. Region-based approaches are often preferred for segmentation based on multiple image and 3D-features as it implicitly utilizes the spatial connectivity as a constraint (unlike clustering methods). In contrast to graph-based approaches region growing is computationally cheap and multiple features can be combined straightforward.

Another aspect concerns the question whether a more data- or a more model-driven approach should be pursued. The key question is to which extent assumptions about the object structure and properties can be made. While model-driven methods help to mitigate the effect of insufficient observations by applying strong assumptions (knowledge) about the object, they might generalize quite strongly. If such knowledge is not available, a data-driven method should be used, being aware of the fact that uncertainties and errors in the observed information might lead to wrong results.

The objective of this paper is to develop a methodology to identify the distinct segments of buildings by 1) detecting the buildings from the 3D point cloud from the UAV-based oblique images and 2) performing a sub-segmentation within the building area in image space using both the spectral and corresponding geometric features from 3D point cloud. For both steps we aim to use as less assumptions (model knowledge) as possible, hence we are pursuing a strong data driven approach. The motivation for this is that one main application of our method is building damage assessment and for this task only vague assumptions should be made to avoid any kind of misinterpretation.

It is also important to note that so far we do not exploit multi-image observations for the segmentation except for the 3D point cloud information which is derived from image matching. Here again the damage mapping context justifies this decision: in many circumstances some parts of a damaged building are only well visible in single images. In this case still we want to be able to derive segmentation information.

## 2. METHODS

The methodology for image segmentation includes two processes, 1) building delineation from a 3D point cloud to define the region of interest for performing image segmentation and 2) image segmentation using the spectral information from the image and 3D geometric features from the 3D point cloud.

### 2.1 Building delineation from 3D point cloud

The building delineation is carried out by finding the connected 3D planar roof segments from the 3D point cloud. A straightforward, quite simplistic approach is used, which, however, turned out to be quite successful, see result section. We only briefly describe this method here, because actually it is just a pre-processing step which allows restricting the processing area for the main step – the segmentation.

- The 3D points are segmented into disjoint planar segments using the plane-based segmentation method as described in Vosselman (2012).
- The 3D points are classified into terrain and off-terrain points using the method proposed by Axelsson (2000) which is implemented as part of the software *lastools* (<http://lastools.org>). The height normalized 3D points are computed by differencing the height of each off-terrain 3D points to its closest terrain 3D point.
- The planar segments that are above certain height ( $T_H$ ) and have surface normal z-component ( $n_z$ ) greater than threshold ( $T_z$ ) are classified as roof segments.
- A connected component analysis is used to identify the spatially connected roof segments of a single building.
- A convex hull is used to define a 2D boundary of the connected roof segments that gives an approximate 2D boundary of the building.
- All 3D points that lie within the defined boundary are registered as the 3D points of the building.

### 2.2 Segmentation

The image segmentation process is carried out based on feature similarity between the spatially connected pixels.

It is a scenario where the 3D planar segments which are derived for detecting buildings from the 3D point cloud are available in addition to the image for segmentation. In this study, an image segmentation algorithm based on region growing concept is developed by utilizing both image spectral and 3D geometric features from the planar segments for finding the distinct segments in the building.

The success of region growing based image segmentation highly depends on three key elements,

- Selection of seed points:** The mid points of 3D planar segments (which are already identified as segments in 3D space) are taken as the seed points for region growing in image space. Here, we assume that at least

a small region of all elements of the building will have 3D points.

**b) Features used for pixel similarity analysis:**

- **Spectral features:** In this study colour features and gray level co-occurrence matrix (GLCM) based texture features are considered to measure the pixel similarity for region growing. A small experiment is conducted to identify the radiometric features that show maximum variation between the pixels belonging to surfaces with different radiometric characteristics. The identified feature is then used in the region growing process.
- **Geometric features:** The 3D points are projected onto the image and the geometric properties such as normal vector and XYZ coordinate of each projected 3D points are assigned to the corresponding image pixel.

**c) Criteria for region growing:** Each image pixel will have a feature vector that represents the spectral characteristics of the pixel and may have geometric features in addition.

Three criteria are used for region growing:

1. The distance between the feature vector of a new pixel to the mean feature vector of the region being grown (Spectral distance)  $< T_{SD}$ .
2. The dot product of normal vector of a new pixel and the plane-normal of the region being grown (Normal difference)  $< T_{angle}$ .
3. The distance between the 3D point corresponding to the new pixel to the plane of the region being grown (point to plane distance)  $< T_{distance}$ .

The image pixels that do not have 3D features will be considered for region growing based on first criteria alone.

The global definition of spectral distance threshold  $T_{SD}$  is not appropriate for segmenting elements of the building corresponding to varying surface characteristics. For example, the pixels corresponding to a rough surface show high spectral variation between them, hence a high  $T_{SD}$  is required to avoid over-segmentation whereas a low  $T_{SD}$  is suitable for smooth surfaces to avoid under-segmentation. Therefore, instead of a global threshold, all seed points are assigned with an adaptive local threshold for region growing. The local threshold for each seed point is computed as the maximum spectral difference between the pixels corresponding to the 3D points that lie within a certain distance from the seed point in the 3D planar segment. Always, the region growing process is initiated by choosing the seed point corresponding to lowest local threshold value in the lists, in order to segment the smoother regions first to avoid under-segmentation.

**Procedure for segmentation of individual elements in the building:**

**a) Data preparation for image segmentation:**

1. Individual buildings in the scene are delineated from the 3D point cloud using the procedure described earlier.

2. Select one delineated building and an appropriate image where the building is visible for segmentation. We are not posing any requirements for image selection, since this decision should be made by the actual application, e.g. the image where a certain damage region is best visible.
3. The 3D points of the planar-segments of the delineated building which are visible in the selected image (camera view) are found using the hidden point removal (HPR) operator e.g., Katz et al. (2007) as described in Gerke and Xiao (2014). The visible points are then projected over the image.
4. The image pixels that correspond to the projected 3D points are assigned with their plane-normal vector and XYZ value.
5. A majority filter is used to assign the 3D features for pixels that do not have corresponding 3D points from their adjacent pixels that have 3D points.
6. The boundary of the building in image is defined by constructing a convexhull for the projected 3D points which forms the region of interest (ROI) for segmentation.
7. The spectral feature such as colour and texture are derived for each pixel.
8. The midpoints of all 3D planar segments are considered as the seed points and each seed point is assigned with four parameters: a) normal vector of the plane, b) distance of the plane to the origin and c) local spectral distance threshold ( $T_{SD}$ ), and d) feature vector of the seed point as mean spectral feature vector.

**b) Image segmentation:**

1. The seed points are sorted by local spectral distance threshold.
2. Remove the topmost seed point (i.e. the one with lowest  $T_{SD}$ ) in the list and initiate region growing using this seed point.
3. Consider the un-segmented neighbouring pixels to the pixels in the region as new pixels for growing.
4. Grow the region by adding the new pixels to the region if they satisfy the growing criteria (refer to (c) under section 2.2) and they lie within the ROI.
5. Update the mean spectral feature vector of the region based on the newly added pixels.
6. Continue steps 3 to 5 until no new pixel is added to the region.
7. Compute the boundary of the new region using a boundary tracing algorithm.
8. Find the seed points that lie within the boundary of the obtained region and remove them from the list.
9. Continue steps 2 to 8 until the seed point list becomes empty.
10. Find the boundary of the regions with significant size (number of pixels) that remain un-segmented.
11. Consider the midpoint of the un-segmented regions as seeds for region growing and perform steps 2 to 9.

The overall workflow is depicted in Figure 1.

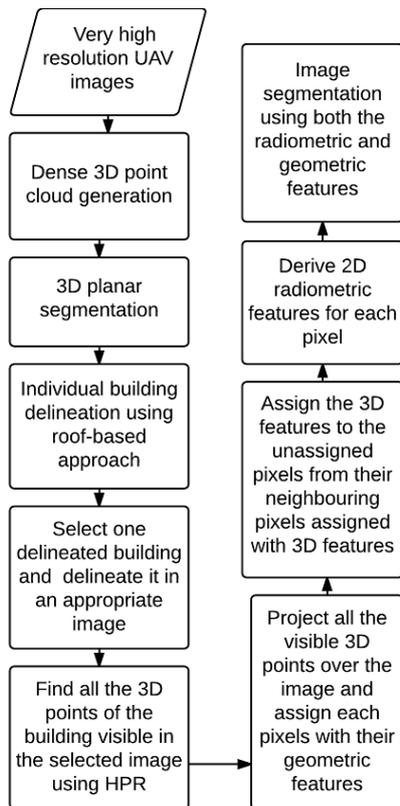


Figure 1. Overall workflow

### 3. EXPERIMENTAL RESULTS

The proposed methodology was tested on two data sets captured by a UAV platform. One important aspect of this kind of image analysis task is the question on how far thresholds and parameters are transferrable. Therefore – besides the standard evaluation of the method – this issue is checked further. It is done by fixing threshold values for the first data set and using the same values for the second.

#### 3.1 Data set 1 and results

The UAV images captured over a small region around the Church of Saint Paul in Mirabello, after the earthquake in 2012, were considered. The images were captured by a VTOL (vertical take-off and landing) UAV from various heights, positions and views (oblique and nadir). The average GSD of the captured images is around 1 cm. A dense 3D point cloud of the scene was generated from 152 images with an average point density of 650 points per  $m^2$  by automatic orientation of the images, followed by dense matching using the software pix4Dmapper (<http://pix4d.com>). The selected region contained six buildings. Among them the larger one comprised of various complex sub-components was considered for testing the developed segmentation method. The selected building consists of different segments such as roofs composed of planar faces with different orientations and different radiometric characteristics, façades painted with different colour, windows in the façade, non-planar objects on the roof, balconies, etc.

#### 3.1.1 Building delineation in 3D point cloud and in image of data set-1:

The 3D point cloud was segmented into disjoint planar segments. The thresholds  $T_H = 3$  meters and  $T_Z = 0.6$  were used to filter out the roof segments through the procedure described in section 2.1. All six buildings in the 3D point cloud of the scene were detected and delineated with close approximate to the actual boundary of the building. The major objective of this research is to segment the building into its various sub-components in image space, once it has been delineated in the 3D point cloud. Hence, detailed information about the conducted experiments, results and analysis related to building delineations from the 3D point cloud is not in the focus of this paper. An example for building delineation from the 3D point cloud and the delineation of the same in the image is shown in Figure 2 and Figure 3. The planar segments that are obtained from the 3D point cloud and lie within the boundary of the delineated building were projected onto the image. Their geometric features were assigned corresponding image pixels. Figure 5 shows that segments are not accurately delineated from 3D planar segments. Also many portions of the building do not have projected 3D points, particularly the façade regions contain sparse 3D points hence these portions have radiometric features alone for segmentation.

#### 3.1.2 Radiometric features and various threshold values used in segmentation:

The colour features such as red, green, blue, hue, and saturation, and GLCM texture features such as mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation were considered. The potential of each feature in separating the pixels of different elements of the building was analysed. Five small image regions corresponding to various elements of the building with different radiometric characteristics were considered as shown in Figure 3. The above mentioned radiometric features were derived for each region. A silhouette value ((Wang et al., 2009) which gives the measure of how well each pixel in one cluster matches with the pixels in the other clusters was used to identify the features that show maximum variation (high silhouette value) between the pixels corresponding to different clusters. The GLCM features showed a higher silhouette value than the colour features (c.f. Figure 4). Particularly, the contrast and homogeneity features produced higher silhouette values than when used independently than used in combination with other GLCM features. Therefore, contrast and homogeneity of GLCM features were used as the radiometric features for image segmentation. The adaptive local spectral threshold method (c.f. section 2.2) provided better results than a global threshold. However, in few regions, an over-segmentation was observed which was then resolved by adding a constant to the local threshold value. As we have the radiometric features as additional constraint for segmentation, the geometric constraints were relaxed by setting higher threshold values for  $T_{\text{angle}}$  (0.9) and  $T_{\text{distance}}$  (0.75 m) to achieve better results even with erroneous 3D point measurements.

The obtained segmentation result for the above mentioned threshold values is shown in Figure 6. Based on visual analysis, it was inferred that the segmentation obtained in image space based on both radiometric and geometric features is more accurate than the segmentation in 3D object space without using radiometric features. The developed segmentation algorithm delineated all planar surfaces in the building with close approximate to their actual boundary. The non-planar objects

and regions that do not have 3D points were segmented using the radiometric features alone. However, in such cases, over- or under-segmentation was observed. For example, c.f. Figure 6, where the rooftop element and small portion of ground were segmented as single segment because of radiometric similarity and absence of 3D information. This clearly implies that both geometric and radiometric features are essential for accurate segmentation. The same building was segmented in another image with smaller scale and different orientation (Figure 7 a). The segmentation was largely similar (Figure 7 b). However, the segmentation in larger scale image is more accurate. This slight performance difference may be due to the variation in texture representation between different scales.

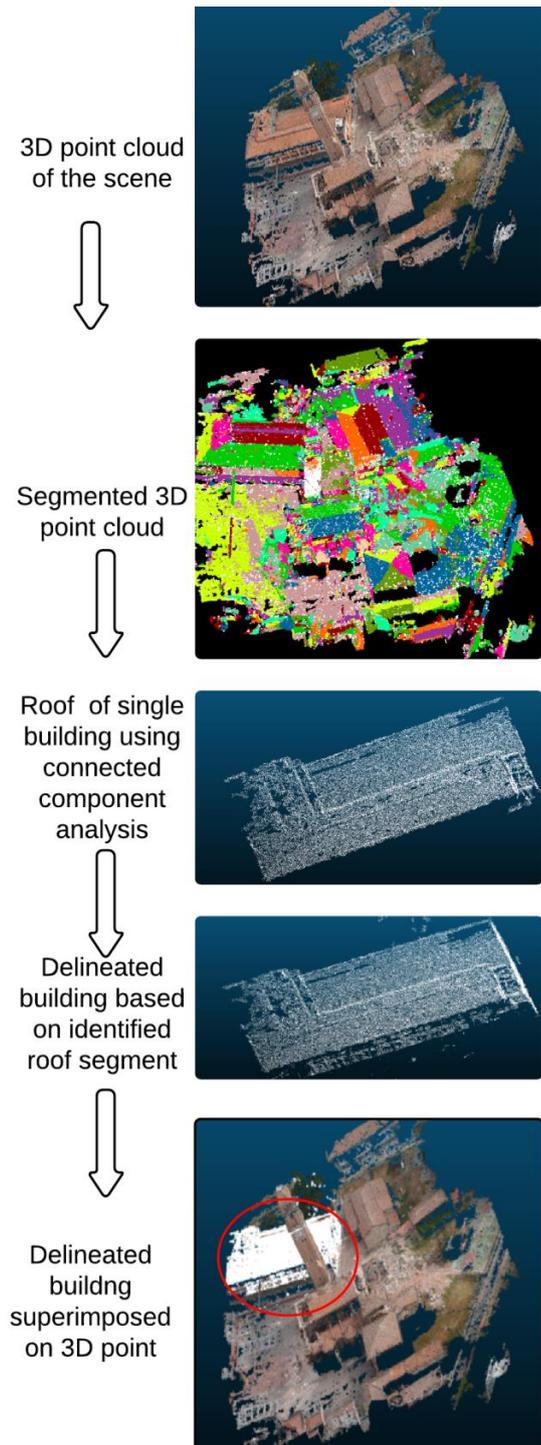


Figure 2. Building delineation from 3D point cloud



Figure 3. Building delineated in image and annotated regions (R1 –R5) are used for feature significance analysis as described in section 3.1.2

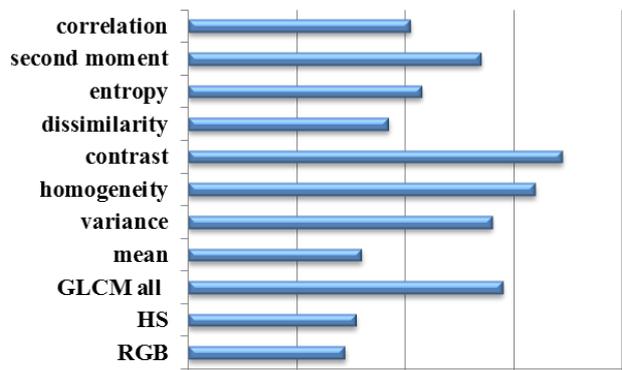


Figure 4. Silhouette value for analysing the feature significance in differentiating the image regions with different radiometric characteristics

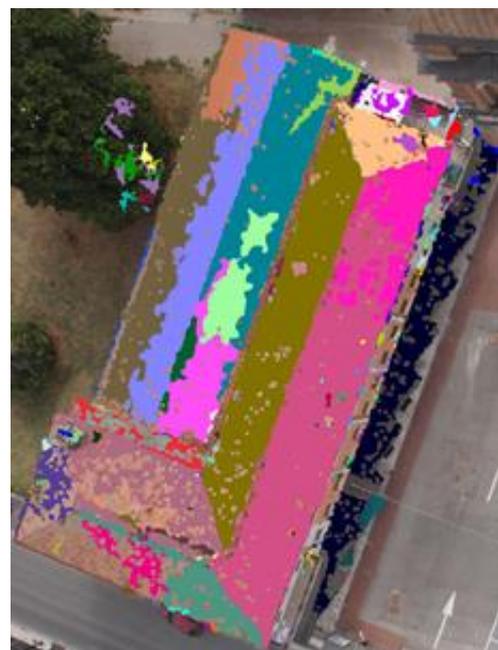
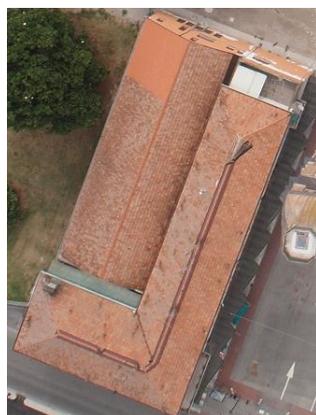


Figure 5. 3D planar segments of the delineated building from 3D point cloud are projected over the image

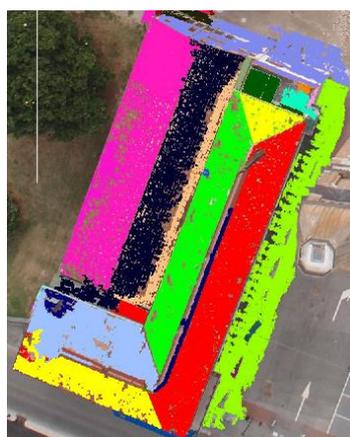


Ground and roof-top object segmented as single element due to radiometric similarity and no 3D information

Figure 6. Delineated building segmented using radiometric and 3D geometric features



(a)



(b)

Figure 7. (a) Building in image at smaller scale and different orientation compared to Figure 3, and (b) corresponding segmentation

The transferability of the thresholds to other datasets will be demonstrated with the following experiment.

### 3.2 Data set 2 and results

The developed segmentation algorithm was tested with the 3D point cloud generated from the UAV images of small urban area in the municipality of Nunspeet in The Netherlands (Hinsbergh et al., 2013). The images are captured in nadir view with an average GSD of 1.5 cm and the 3D point cloud was generated with an average point density of 250 points per  $m^2$ . The buildings in this region are less complex compared to the selected building from dataset 1. For example, the individual elements in the building are highly homogeneous and show high contrast with their neighbouring elements in terms of radiometric characteristics. Moreover, the buildings in the selected region are more identical and mainly made of planar surfaces. Among them two buildings that show different structure were considered for testing the segmentation algorithm (Figure 8a & 8d). The selected buildings have gabled roofs with different kinds of windows on them, such as flat windows that lie in the same roof plane and windows extruded above the roof. The façades are single planar surfaces with uniform colour and texture.

The 3D planar segments obtained from 3D point cloud were projected over the image. Many of the 3D segments were more accurately segmented when compared to the 3D segments obtained for the building in dataset 1 (Figure 8b & 8e). However, over-segmentation was observed in the façade and few places in the roof (Figure 8e). The flat windows over the roofs were not identified as separate segments in the 3D segmentation.

The image segmentation using the texture features along with the projected 3D features was carried out following the same procedure and thresholds used for the segmentation of building in dataset 1. The segmentation results are shown in Figure 8c & 8f.

The obtained results were better than the plane-based 3D point cloud segmentation where the over segmented regions in 3D space such as façades were well segmented in the image space (c.f. Figure 8e & 8f). Most of the windows and small non-planar above roof elements were also segmented as separate segments. However, in few places over-segmentation was observed due to the variation in radiometric characteristics within the same element. For example, the dirt in the corner of the segment resulted in over-segmentation even though they are geometrically recognized as single planar segment (c.f. annotated region in Figure 8-d and same region in Figure 8e & 8f). This is due to the weakness in the segmentation criteria where the geometric constraints are relaxed to a certain extent when the radiometric characteristics are similar but not the other way around. However, the radiometric constraint has to be relaxed when there is strong hold on geometric characteristics. For example in the above case, the segmentation based on geometric features results in uniform shape whereas the consideration of radiometric features results in a non-uniform shape. In such instance the radiometric constraint can be relaxed. This kind of analysis can be carried out even in post-processing.



(a)



Radiometric  
variation due  
to dirt

(d)

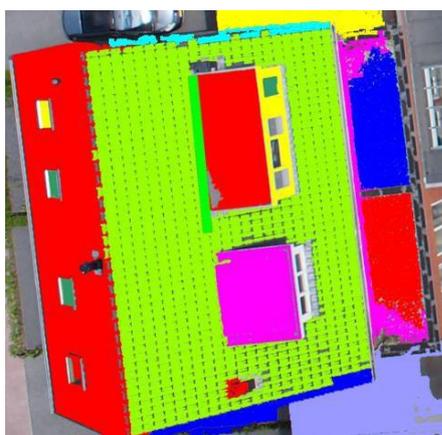


(b)

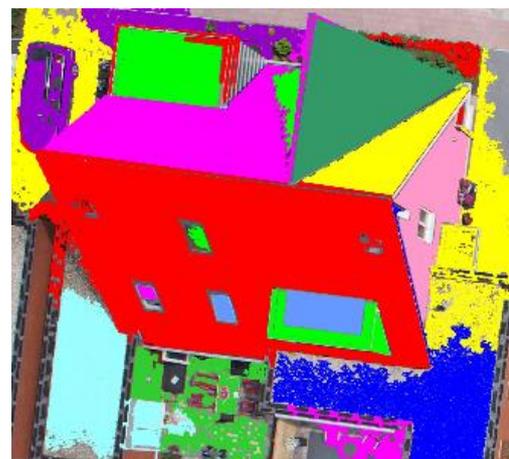


Over-segmented facade

(e)



(c)



(f)

Figure 8. (a) & (d): Buildings in UAV image for segmentation, (b) & (e): projected 3D segments over the images of (a) & (d) respectively, and (c) & (f): finally segmented images using the developed method

#### 4. DISCUSSION AND CONCLUSION

The overall results indicate that the radiometric features complement the 3D geometric features and a combination of the two produced significantly superior segmentation compared to the 3D geometric features based segmentation alone. The radiometric features seem to be advantageous in identification of single segments even though there is significant error in geometric measurements. The sub-segmentation of planar objects also might lead to over-segmentation, when the face contains shadows, dirt, etc., refer to Figure 8d, 8e & 8f. This is however, the correct behaviour since on purpose we chose this data driven approach. In the actual application – like damage mapping – those segments might give valuable information for the interpretation.

In this study, the 3D features such as normal orientation and planarity derived from plane-based segmentation in 3D space were used as geometric features in combination with radiometric features for segmentation in image space. The plane-based features alone are not sufficient in all cases. For example, plane-based features cannot accurately segment the curved surface which leads to over-segmentation. In such cases, other 3D features could be of help, such as the curvature feature which can be computed based on local neighbourhood 3D points. The inclusion of more 3D features such as curvature likely will improve the segmentation accuracy.

#### ACKNOWLEDGMENTS

We thank Pix4D for providing us with the research license of pix4dmapper. All image orientation and dense matching tasks for this research were performed with pix4dmapper. We thank the Dutch Kadaster for providing us with the Nunspeet UAV dataset. The work was funded by the FP7 project RECONASS (Reconstruction and Recovery Planning: Rapid and Continuously Updated Construction Damage and Related Needs Assessment; grant no 312718).

#### REFERENCES

Aijazi, A.K., Checchin, P. and Trassoudaine, L., 2013. Segmentation based classification of 3D urban point clouds: A super-voxel based approach with evaluation. *Remote Sensing*, 5(4), pp. 1624-1650.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, 33(B4/1), pp. 111-118.

Boykov, Y. and Funka-Lea, G., 2006. Graph cuts and efficient ND image segmentation. *International Journal of Computer Vision*, 70(2), pp. 109-131.

Colomina, I. and Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, pp. 79-97.

Dong, L. and Shan, J., 2013. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 84, pp. 85-99.

Dorninger, P. and Pfeifer, N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors*, 8(11), pp. 7323-7343.

Enjarini, B. and Graser, A., 2012. Planar segmentation from depth images using gradient of depth feature, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4668-4674.

Gerke, M. and Xiao, J., 2014. Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87, pp. 78-92.

Hinsbergh, W.H.M.v., Rijdsdijk, M.J. and Witteveen, W., 2013. UAS for cadastral applications: testing suitability for boundary identification in urban areas. *GIM International*, 27, pp. 20-25.

Hulik, R., Beran, V., Spanel, M., Krsek, P. and Smrz, P., 2012. Fast and accurate plane segmentation in depth maps for indoor scenes, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1665-1670.

Katz, S., Tal, A. and Basri, R., 2007. Direct visibility of point sets. *ACM Trans. Graph.*, 26(3), pp. 24.

Liu, J. and Guo, B., 2014. Reconstruction and simplification of urban scene models based on oblique images. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1, pp. 197-204.

Mirante, E., Georgiev, M. and Gotchev, A., 2011. A fast image segmentation algorithm using color and depth map, *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1-4.

Narkhede, H., 2013. Review of image segmentation techniques. *International Journal of Science and Modern Engineering*, 1(8), pp. 54-61.

Rampun, A., Strange, H. and Zwiggelaar, R., 2013. Texture segmentation using different orientations of GLCM features, *Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*. ACM, pp. 17.

Rottensteiner, F., Sohn, G., Gerke, M., Wegner, J.D., Breitkopf, U. and Jung, J., 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, pp. 256-271.

Rupnik, E., Nex, F. and Remondino, F., 2014. Oblique multi-camera systems - orientation and dense matching issues. *ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-3/W1, pp. 107-114.

Strom, J., Richardson, A. and Olson, E., 2010. Graph-based segmentation for colored 3D laser point clouds, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2131-2136.

Sun, S. and Salvaggio, C., 2013. Aerial 3D building detection and modeling from airborne LiDAR point clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3), pp. 1440-1449.

Tian, D., 2013. A Review on Image Feature Extraction and Representation Techniques. *International Journal of Multimedia and Ubiquitous Engineering*, 8(4), pp. 385-396.

Vosselman, G., 2012. Automated planimetric quality control in high accuracy airborne laser scanning surveys. *ISPRS Journal of Photogrammetry and Remote Sensing*, 74, pp. 90-100.

Wang, K., Wang, B. and Peng, L., 2009. CVAP: Validation for cluster analyses. *Data Science Journal*, 8, pp. 88-93.

Xiong, B., Oude Elberink, S. and Vosselman, G., 2014. Building modeling from noisy photogrammetric point clouds. *ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3, pp. 197-204.

Xiong, X., Adan, A., Akinci, B. and Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction*, 31, pp. 325-337.

Yin, J. and Kong, S., 2013. Hierarchical image segmentation algorithm in depth image processing. *Journal of Multimedia*, 8(5), pp. 512-518.