A HYBRID METHOD FOR DERIVING DTMS FROM URBAN DEMS

C. Baillard

SIRADEL, 3 allée Adolphe Bobière CS 24343, 35043 Rennes, France -
cbaillard@siradel.com

Commission III, WG III/2

KEY WORDS: DEM/DTM, urban, aboveground, correlation, segmentation, classification, triangulation, Random Markov Field

ABSTRACT:

This study focuses on the automatic extraction of DTMs from urban DEMs produced by image correlation. Many methods have been proposed in the literature yet processing correlation DEMs in any kind of areas remains a challenge. This paper presents a hybrid approach that combines complementary aspects of both TIN-based and segmentation-based techniques. Unlike previous work involving two complementary modules, the two techniques closely interact during the process. The DEM is first segmented and classified into ground and aboveground regions using contextual information. A DTM is then derived from the ground regions using a TIN-based technique. The classification and the DTM estimation are finally iteratively performed until stability. The hybrid approach for DTM extraction has been tested over several representative data sets and compared with the TIN-based and region-based approaches applied independently. It clearly shows that coupling complementary approaches improves the quality of the resulting DTM, as well in dense urban areas as in rural or hilly areas.

1. INTRODUCTION

The production of Digital Elevation Models in urban areas has undergone major advances thanks to the evolution of sensors and processing algorithms. Urban DEMs can be acquired using two main techniques: LIDAR scanning techniques, providing highly accurate DEMs, and image matching techniques, based on the correlation of at least two aerial or satellite images, providing DEMs at a competitive cost. DEMs can rarely be used as such but is a good basis for further analysis. A preliminary step to building reconstruction or change detection is often the derivation of a Digital Terrain Model, which only describes the ground surface.

This study focuses on the automatic extraction of a DTM from a high-resolution DEM produced by image correlation in urban or rural areas. No external data is used. The method must be easy to tune, cost-effective, and able to cope with noise and occlusion areas. The resulting DTM should be accurate and smooth whilst preserving breaklines.

Section 2 first presents the current state-of-the art for deriving a DTM from a correlation or LIDAR DEM. Section 3 then focuses on a hybrid method based on two complementary techniques. Finally, qualitative results obtained on several data sets are presented and analysed in section 4.

2. RELATED WORK

All filtering algorithms are based on assumptions about input elevation data and ground characteristics. Some methods make use of external data, like building or vegetation masks. Most recent filters have been designed for LIDAR data: a thorough study of performance has been carried out in (Sithole, G. and Vosselman, G., 2004). We can distinguish several categories of methods that can be applied to LIDAR or correlation DEMs.

The first category of methods is based on local operators. This is the case of morphological filters or slope-based operators (Weidner, U. and Förstner, W., 1995; Eckstein, W. and Munkelt, O., 1995; Schiewe, J., 2000; Roggero, M., 2001; Sithole, G.,

2001; Zhang *et. al.,* 2003). These techniques often encounter problems when an aboveground object is disconnected from the ground (like a building surrounded by others), and they do not always preserve breaklines.

A second type of techniques relies on the optimisation of a piecewise continuous surface representing the ground. The DTM is initialised using all the input points then refined within an iterative process. A popular method is the hierarchical robust interpolation of the bare surface using linear prediction (Pfeifer *et. al.,* 2001). The ground surface can also be estimated using the principle of active shape models (Elmqvist, M., 2002) or by means of a global parametric model (Belli, T. *et. al.,* 2001). In (Champion, N. and Boldo, D., 2006), the elastic grid technique is coupled with M-estimators to reduce effects of outliers. In general, the quality of the resulting DTM depends on the initialisation and on the various parameters characterizing the surface. The limits of these techniques can be the computing time, the number of parameters, and often the difficulty to handle terrain discontinuities and sharp ridges.

The TIN-based approaches can be seen as a particular case of the surface-based approaches. The ground surface is modelled as a TIN (Triangular Irregular Network) and iteratively densified. Seed points chosen from the input DEM initialise the TIN, then consistent points are iteratively added to the current TIN model (Axelsson, P., 2000; Sohn, G. and Dowman, I., 2002). These methods provide accurate DTM preserving breaklines, however they are very sensitive to errors present in the initial TIN. In addition, they do not simultaneously cope with extended aboveground objects and steep hills.

Another category of methods relies on the segmentation and the classification of the input DEM. Segmentation can be performed using a growing and merging algorithm (Baillard, C. and Maître, H., 1999), or profile analysis (Sithole, G. and Vosselman, G., 2005). A label is given to each segmented region using geometric or contextual criteria. These methods provide a classification of the scene but the DTM still needs to be estimated.

Some examples of combinations coupling different techniques can be found in the literature. Typically a classification of the

scene is first performed, then the ground surface is estimated from the ground points only. In (Abo Akel, N., *et. al.*, 2004), the road regions are first detected then used to initialize a DTM before a robust interpolation based on polynomial polygons. In (Van de Woestyne I. *et. al.*, 2004), the scene is segmented using 3D spheres then a parametric surface model is applied. In (Bretar, F. *et. al.,* 2005), the points are first classified using a multiple pass classification, then the ground surface is estimated using a deformable model where "ground" laser points are considered as local attractors.

3. PRESENTATION OF THE HYBRID METHOD

Most previous studies focus on LIDAR DEMs. Correlation DEMs have different characteristics: height discontinuities are less accurately located and data are generally noisier, possibly with some occlusion areas that do not contain any 3D information. The extraction method must be robust enough to deal with noise and outliers.

A hybrid approach has been designed that combines complementary aspects of both TIN-based and segmentation-based techniques. It stems from the work described in (Baillard, C. and Maître, H., 1999), which was extended to cope with steep slopes, breaklines and complex shapes. Unlike previous work involving two complementary types of techniques, the TIN-based and segmentation-based techniques closely interact during the process.

Section 3.1 presents the selected TIN-based method for estimating the DTM. Section 3.2 presents the method for classifying the scene according to the estimated DTM and contextual information using a Markov random field model. Section 3.3 shows how the two techniques interact.

3.1 DTM estimation with a TIN-based approach

The approach is an extension of (Axelsson, P., 2000), which is implemented in the commercial software TerraScan. A sparse TIN is created from seed points and iteratively densified. At every iteration, points can be added to the TIN if they are below data derived thresholds. These thresholds are distances to TIN facets and angles to the facet nodes. The original algorithm proceeds as follows:
- Estimation of initial thresholds using all the data;
- Selection of seed points within a user-defined grid;
- Iterative densification of the TIN :
    o Updating thresholds using the current TIN
    o Adding to the TIN the points that meet constraints on distance and angles.

The selection of seed points is critical, as they are never questioned afterwards. Wrong seed points can be selected in case of low outliers in the input DEM, or in the presence of large building blocks or wooden areas. Besides, breaklines are preserved only if the final TIN density is close to the original density of points, which is computationally expensive.

The original algorithm has therefore been extended. Seed points are selected from a grid as the 95% lowest point of each grid cell, which makes initialisation more robust to low outliers. Selected seed points are used to create a TIN. The slopes of the TIN edges are computed. Each seed point is associated to the median slope value determined from all the TIN edges connected to it. If the slope value is too big (local maximum) or too small (local minimum) then the seed point is rejected. In order to reduce computational times, the final TIN density can be reduced to 1 to 5m according to the application. A new stage has therefore been added at the end of the original process: A point can be added to the TIN if its height is located between the lowest vertex and the highest vertex of the triangle that contains it. This post-processing helps in preserving breaklines and reduces effects of low outliers in the initial TIN.

The adapted Axelsson filter offers very interesting results in terms of computation time and breaklines, although particular attention must still be paid to the initialisation of the TIN.

3.2 Scene classification with a region-based approach

The method was originally proposed in (Baillard, C. and Maître, H., 1999). The input DEM is segmented into ground and aboveground regions using a Markov random field model. It relies on the following definitions: an aboveground region is a homogeneous part of the scene higher than the surrounding ground, from a critical value δ_0. More precisely, the original technique can be described as follows:
- Segmentation of the input DEM into homogeneous regions using a growing and merging algorithm; creation of the corresponding adjacency graph, called "3D graph";
- Initialisation: each node s (region) is associated to the label "ground";
- Iterative classification of the node labels as "ground" or "aboveground" (see Figure 1):
    o Estimation of a DTM from the "ground" labelled nodes, using basic sampling within a 25m grid;
    o New estimation of the node labels using a Markov random field model

The process stops when the classification is stable.
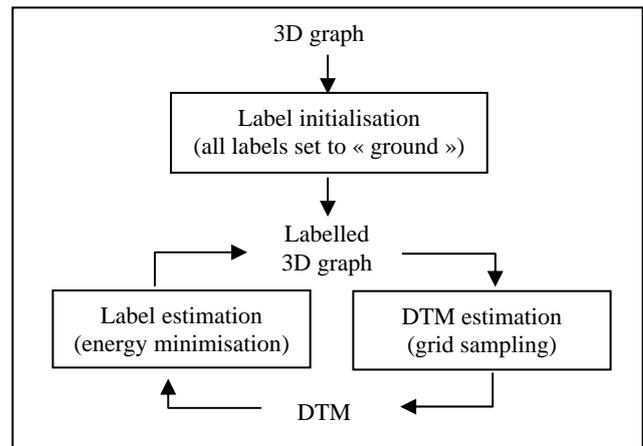
Figure 1: Iterative process for the region-based approach

The estimation of the labels is formulated as the minimization of a potential function $U_\mathcal{G}$ given by:

$$U_\mathcal{G} = \alpha \cdot U_d + \beta \cdot U_c \qquad (1)$$

The potential U_d is the data related potential linking elevation and label at each node. It is dependent on the current DTM.

The potential U_c is the contextual potential depending on the difference in height between each pair of neighbouring nodes.

Weighting parameters α and β are defined such as $\alpha + \beta = 1$. The potential functions are made of arcs of a Gaussian and need only one user-defined parameter δ_0, which is the maximal difference in height between two neighbouring "ground" nodes. The potential values computed at a given site are comprised between -1 and $+1$. A value equal to -1 indicates a possible configuration, whilst positive values correspond to unlikely configurations.

More precisely, the potentials are defined as follows:

- $$U_d = \frac{1}{N} \sum_s V_d(s) \qquad (2)$$

where $V_d(s)$ is the data related potential function at node s and N is the number of nodes.

Figure 2 shows the appearance of Vd(s) for a "ground" node and an "aboveground" node, as a function of the relative height h(s) = z(s) – zdtm(s).
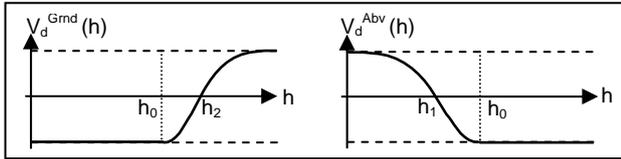
Figure 2: Data related potential functions

The functions are parameterised by data-derived parameters $h_0(s)$, $h_1(s)$ and $h_2(s)$ defined by:

- o $h_0 = zdtm_{Avg}(s) + \delta_0$
- o $h_1 = MIN(zdtm_{Min}(s) + \delta_0, zdtm_{Avg}(s))$
- o $h_2 = MAX(zdtm_{Max}(s) + \delta_0, zdtm_{Avg}(s) + 2\,\delta_0)$

where $zdtm_{Avg}(s)$, $zdtm_{Min}(s)$ and $zdtm_{Max}(s)$ are respectively the average, the minimum and the maximum elevations of the DTM over the node s. The elevation h_0 is the critical value above which $V_d(s)$ is favourable to aboveground label.

- $$U_c = \frac{1}{M} \sum_{(s,s')} w(s,s') V_c(s,s') \qquad (3)$$

where $V_c(s,s')$ is the contextual potential function at pair (s,s'), $\omega(s,s')$ is a weighting function, and M a normalization constant defined by:

$$M = \sum_{(s,s')} w(s,s') \qquad (4)$$

Figure 3 shows the appearance of Vc(s,s') for the four possible configurations of labels, as a function of the difference in height $\delta = z(s) - z(s')$.
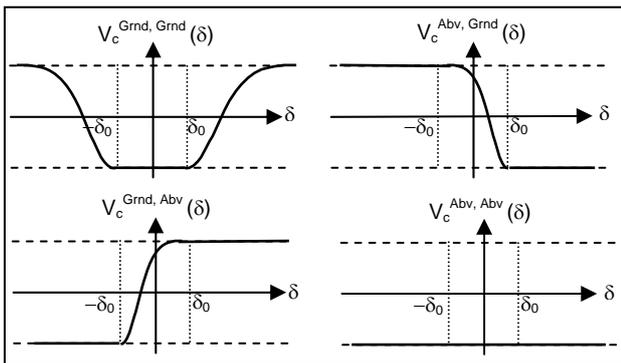
Figure 3: Contextual potential functions

The parameter δ_0 is the critical difference in height mentioned earlier. The standard deviations of the Gaussians are derived from the data.

The original method was modified in order to better deal with steep slopes or disconnected regions. The standard deviations were made dependent on the local slope of the input DEM. If θ is the local slope of the DEM computed between $[0;\pi/2]$, then the standard deviations are defined by:

$$\sigma_{Grnd,Grnd} = \frac{\delta_0}{\sqrt{8\ln 2}} \frac{\pi/2}{\theta} \qquad (5)$$

$$\sigma_{Abv,Grnd} = \sigma_{Grnd,Abv} = \frac{\sigma_{Grnd,Grnd}}{2} \qquad (6)$$

The standard deviations are inversely proportional to the slope.

Importantly, at the particular value $\theta = \frac{\pi}{4}$ we have:

$$V_c^{Grnd,Grnd}(2\delta_0) = V_c^{Abv,Grnd}(\frac{\delta_0}{2}) = V_c^{Grnd,Abv}(\frac{-\delta_0}{2}) = 0 \qquad (7)$$

Additionally, the weighting function $\omega(s,s')$ is equal to the length of the borderline between two neighbouring regions to limit the effect of small outliers.

These new potentials have a better behaviour in case of slope or extended aboveground areas. However low aboveground objects are not always detected, and the resulting DTM is irregular because no consistency is taken into account between the sampled points (basic sampling within a 25m grid).

3.3 Global optimisation with a hybrid approach

In order to cope with the limitations of both methods, a hybrid combination has been designed. The scene is first segmented and classified with the region-based approach using contextual information only (no DTM estimation is involved). Then the classification and the TIN-based DTM estimation are iteratively performed until stability (see Figure 4). The two techniques closely interact during the process: the region classification is used to filter non-ground points before the DTM estimation; the DTM estimation is used to compute the data related potential related to the graph.

Unlike the work previously described in (Baillard, C. and Maître, H., 1999), the DTM is finely estimated at each iteration of the process resulting in a accurate and consistent description of the ground elevation and the scene content. The first estimation stage based on contextual energy makes the TIN densification less dependent on initialisation because it is always based on a realistic classification.
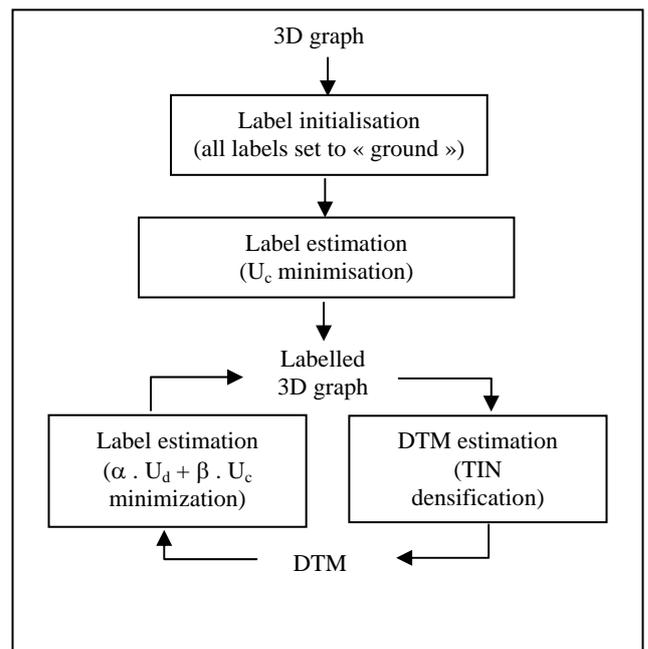
Figure 4: Iterative process for the hybrid approach

4. EXPERIMENTATION AND DISCUSSION

The hybrid approach for DTM extraction has been tested over several representative data sets and compared with the TIN-based and region-based approaches applied independently.

4.1 Input Data

Four data sets are presented in this paper. The DEMs of Borup, Marseille1 and Marseille2 are shown in Figure 5a, 6a and 7a. They were computed by image correlation with software developed at Siradel and presented in (Baillard, C., 2003). The data was derived from stereo images with 60% overlap and includes noise and occlusion areas. The DEM over Lyngby is shown in Figure 8a. It was provided via EuroSDR (http://buildingsdetection.free.fr/) by the Danish National Survey and Cadastre (KMS) and was computed using LIDAR data. Table 1 summarizes the characteristics of the input DEMs.

Name	DEM Res (m)	Area size (m²)	ΔZ	Type
Borup	1	1000x1000	35	Peri-urban
Marseille1	0.37	464x655	55	Urban
Marseille2	0.37	591x653	35	Dense urban
Lyngby	1	1000x1000	45	Urban

Table 1. Characteristics of the input DEMs.

All the DEMs were processed with the same set of parameters. The final sampling resolution of the DTM is 5m for the three methods.

4.2 Qualitative results

Figure 5 and 6 show the results over Borup and Marseille1. The DTM was computed with different methods: TIN densification (Figures 5b and 6b), region classification and basic sampling (Figures 5c and 6c), and hybrid method (Figures 5d and 6d). It is clear that the TIN-based method tends to erode the top of hills (see Marseille1) and does not cope with extended aboveground regions (see forest of Borup). The region-based approach has a better behaviour with these objects however fails at detecting low above-ground objects (see south part of Borup). The hybrid approach takes advantage from both methods. Low aboveground objects as well as extended ones are correctly detected, and the relief is preserved even in case of strong slope. Even when classification fails, the robust TIN densification allows producing a correct DTM.

More examples can be found in Figure 7 and 8 on the areas of Marseille2 and Lyngby. On Marseille2 raised car parks have been correctly rendered. On Lyngby most ground features are correctly represented, and the breaklines are preserved. Only the raised road has been partially removed from the DTM. This shows a limit of the method that cannot distinguish between a raised road and a low and large building despite the global optimisation of the potentials. Some errors can also occur at low building or vegetation areas disconnected from the ground, which can be included in the final DTM.

The average computation time was 30s with the TIN-based method, 1mn30s for the region-based method, and 2mn30s for the hybrid method. The hybrid method is obviously more expensive than the other ones because classification and TIN estimation are iteratively performed, but the computation time is still acceptable for operational DTM production.

5. CONCLUSION

In this paper, we have presented a hybrid method for deriving DTMs from urban DEMs. TIN-based DTM estimation and region-based classification closely interact during the process. Improvement has been qualitatively demonstrated over several data sets. The method needs no tuning and is able to cope with noise and occlusion areas. Resulting DTMs are accurate and preserving breaklines. The method can be applied to correlation DEMs or LIDAR DEMS. The good results show the feasibility of such a method for the operational production of DTM over various kinds of areas: dense urban cities, suburban areas with forests, hilly areas, etc. A quantitative analysis must be performed to complete the study.

Residual errors often come from low buildings disconnected from the ground, which should be corrected by detecting isolated maxima in the final TIN. A pre-processing step detecting road regions as proposed in (Abo Akel, N., *et. al.*, 2004) could solve the delicate problem of raised roads and bridges. The main drawback of the method stems from the TIN representation itself, which is not smooth enough between breaklines. Further studies will focus on the final estimation of the DTM using a surface-based technique.

REFERENCE

Abo Akel N., Zilbersteil, O., Doytsher, Y., 2004. A robust method used with orthogonal polynomials and road network for automatic terrain surface extraction from lidar data in urban areas. In *IAPRS Vol. XXXIII, Part. B3*, XIX ISPRS Congress, Amsterdam, the Netherlands.

Axelsson, P., 2000. DEM generation from laser scanner data using adaptive tin models. In *IAPRS Vol. XXXIII, Part. B4*, XIX ISPRS Congress, Amsterdam, the Netherlands.

Baillard, C. and Maître, H., 1999. 3-D Reconstruction of Urban Scenes from Aerial Stereo Imagery : A Focusing Strategy. *Computer Vision and Image Understanding, 76(3), pp. 244-258*, dec 1999.

Baillard, C., 2003. Production of urban DSMs combining 3D vector data and stereo aerial imagery. In *IAPRS Vol. XXXIV-3/W8*, Proc. of the ISPRS Workshop Photogrammetric Image Analysis, Munich.

Belli, T., Cord, M., Jordan, M., 2001. 3D data reconstruction and modeling for urban scene analysis, In *Automatic extraction of man-made objects from aerial and space images (III)*, Ascona, Switzerland.

Bretar, F., Roux, M., Pierrot-Deseilligny, M., 2004. Terrain modeling and Airborne laser data classification using multiple pass filtering. In *IAPRS Vol. XXXV, Part. B3*, XX ISPRS Congress, Istanbul, Turkey.

Champion, N. and Boldo, D., 2006. A Robust Algorithm for estimating Digital Terrain Models from Digital Surface Models in Dense Urban Areas. In *IAPRS Vol. XXXVI part 3*, Comm. III Symposium on Photogrammetric Computer Vision (PCV'06), Bonn, Germany.
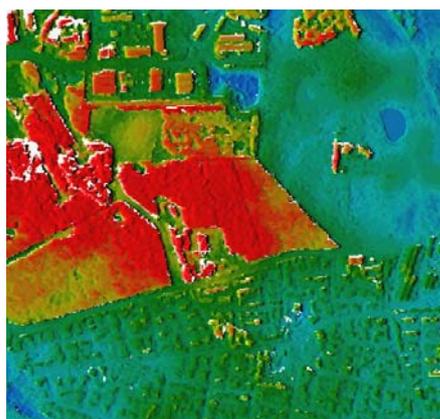
(a)
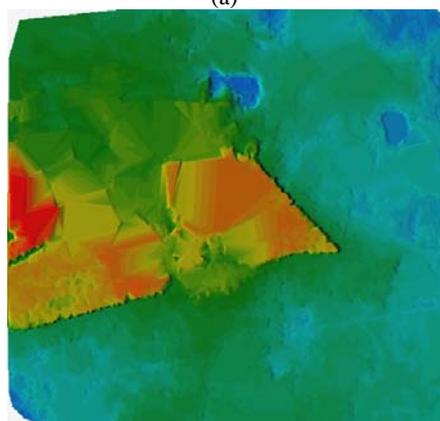
(b)
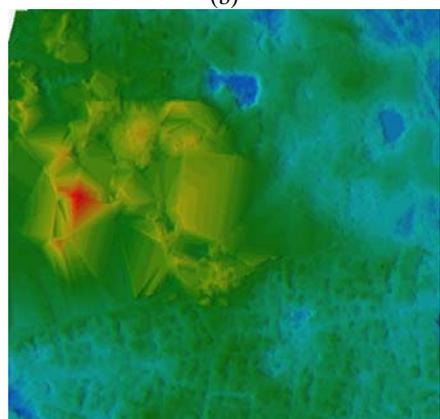
(c)
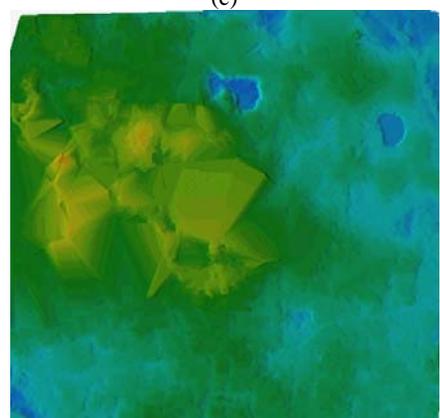
(d)

Figure 5: Borut; (a) = Input DEM; (b) = DTM from TIN-based method; (c) = DTM from region-based method; (d) = DTM from hybrid method
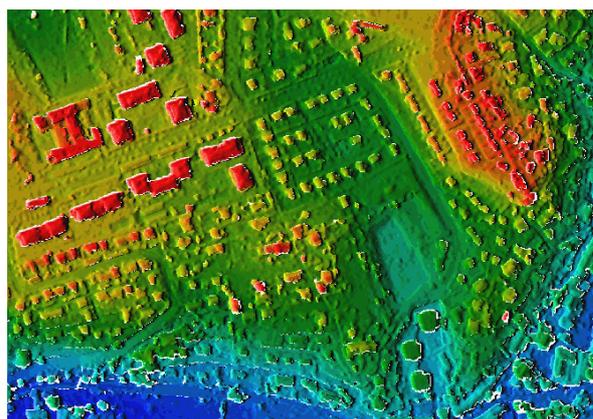
(a)
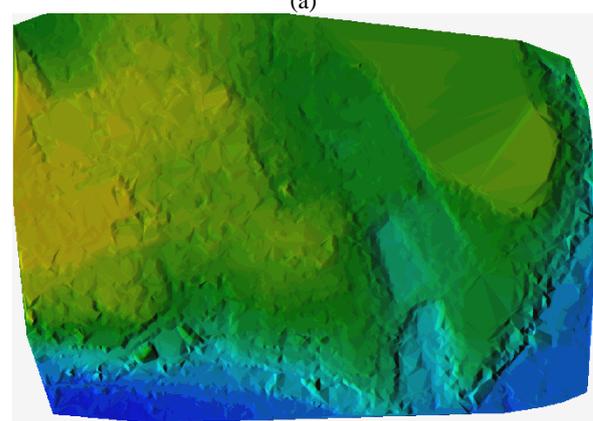
(b)
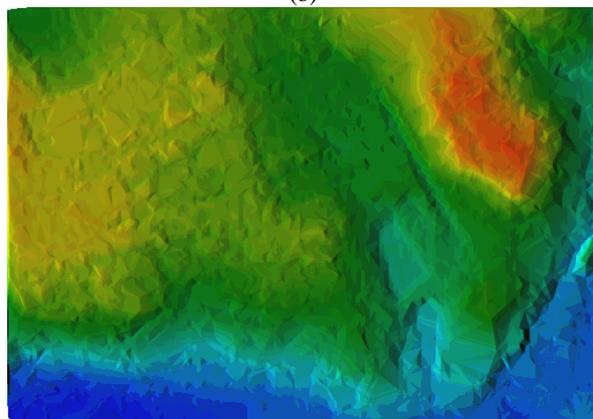
(c)
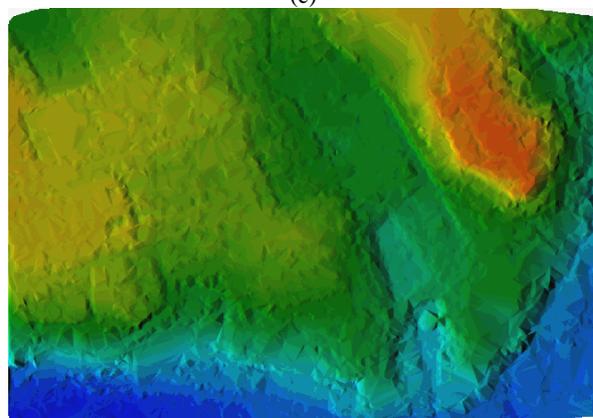
(d)

Figure 6: Marseille1; (a) = Input DEM; (b) = DTM from TIN-based method; (c) = DTM from region-based method; (d)=DTM from hybrid method
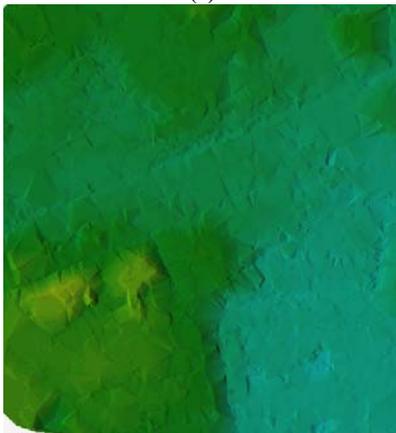
Figure 7: Marseille2
(a) = Input DEM; (b) = DTM from hybrid method
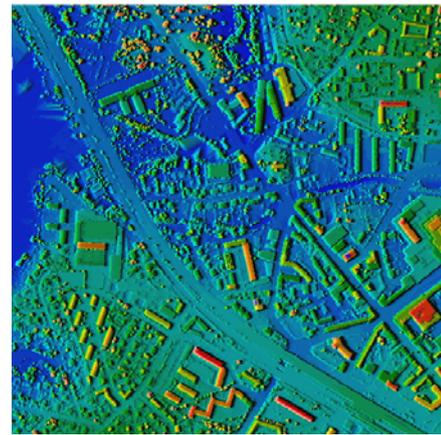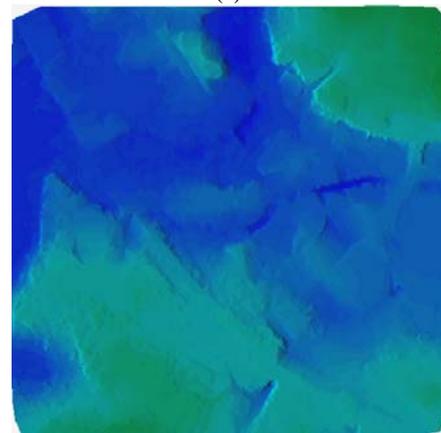
Figure 8: Lyngby
(a) = Input DEM; (d) = DTM from hybrid method

Eckstein, W. and Munkelt, O., 1995. Extracting objects from digital terrain models. In: T. Schenk (ed.), *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, SPIE.

Elmqvist, M., 2002. Ground surface estimation from airborne laser scanner data using active shape models. In *IAPRS Vol. XXXIV part 3A*, Comm. III Symposium on Photogrammetric Computer Vision (PCV'02), Graz, Austria.

Pfeifer, N., Stadler, P., Briese, C., 2001. Derivation of digital terrain models in the SCOPP+ environment. *Proc. of OEEPE workshop on Airborne Laserscanning and Interferometric SAR for Digital Elevation Models*, Stockholm, Sweden.

Roggero, M., 2001. Airborne laser scanning: clustering in raw data. In *IAPRS Vol XXXIV-3/W4*, WGIII/3-6 Workshop on Land surface mapping and characterization using laser altimetry, Annapolis, MD, USA.

Schiewe, J., 2000. Improving the integration of digital models. In *IAPRS Vol. XXXIII, Part. B3*, XIX ISPRS Congress, Amsterdam, the Netherlands.

Sithole, G., 2001. Filtering of laser altimetry data using a slope adaptive filter. In *IAPRS vol. XXXIV-3/W4*, WGIII/3-6 Workshop on Land surface mapping and characterization using laser altimetry, Annapolis, MD, USA.

Sithole, G. and Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59 (1-2), pp. 85–101.

Sithole, G. and Vosselman, G., 2005. Filtering of airborne laser scanner data based on segmented point clouds. In *IAPRS Vol. XXXVI-3/W19*, Proc. of the ISPRS workshop on Laser scanning 2005, Enschede, the Netherlands.

Sohn, G. and Dowman, I., 2002. Terrain surface reconstruction by the use of tetrahedron model with the MDL criterion. In *IAPRS Vol. XXXIV part 3A*, Comm. III Symposium on Photogrammetric Computer Vision (PCV'02), Graz, Austria.

Van de Woestyne I., Jordan, M., Moons, T., Cord, M., 2004. A software system for efficient DEM segmentation and DTM estimation in complex urban areas. In *IAPRS Vol. XXXV, Part. B3*, XX ISPRS Congress, Istanbul, Turkey.

Weidner, U. and Förstner, W., 1995. Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4):38-49.

Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J. and Zhang, C., 2003. A progressive morphological filter for removing nonground measurements from airborne lidar data. *IEEE Transactions on Geoscience and Remote Sensing* 41(4), pp. 872-882.