A ROBUST ALGORITHM FOR ESTIMATING DIGITAL TERRAIN MODELS FROM DIGITAL SURFACE MODELS IN DENSE URBAN AREAS

Nicolas champion Didier Boldo

MATIS Laboratory, Institut Géographique National
2, Avenue Pasteur. 94165 SAINT-MANDE Cedex - FRANCE
Firstname.Lastname@ign.fr

Commission III/2

KEY WORDS: DSM, DTM, Elastic Grid, Outliers, Robust Statistics

ABSTRACT:

This paper describes an algorithm in order to derive DTMs (Digital Terrain Models) from correlation DSMs (Digital Surface Models) and above-ground (buildings and vegetation) masks in dense urban areas. Among all the methods found in literature, the Elastic Grid method shows a good capability to reconstruct the topographic surface. This method consists in interpolating height values under above-ground masks by minimizing an energy. Nevertheless, this method is ill-adapted to outliers in input data (above-ground points out of above-ground masks). The main contribution of our study is the use of a method based on robust statistics in order to reject outliers from calculation so that the final DTM fits the "true" topographic surface for the best. For that purpose, the initial Elastic Grid has been noticeably changed. The results of the new method for 2 test sites with a pixel ground size of 20 cm (the first one is relatively flat and the second one is hilly) show the quality of the final DTM and the robustness of our method. Tests have been carried out with lower resolution DSMs and without any mask and show the feasability of extending the method to a more general context.

1 INTRODUCTION

1.1 Background

In the past few years, DTMs (Digital Terrain Models) have increasingly been used as an important tool for engineering works or environmental applications (water overflowing control for example).
In urban areas, especially in a change detection process, a DTM can be very useful. As a matter of fact, using only the radiometric and texture information from RGB images or orthophotos are generally not sufficient to perform a good detection of buildings. The buildings height, calculated by making the difference between a DSM and the corresponding DTM, is often necessary. A lot of techniques exist to calculate DSMs (lidar scanning, stereo-matching algorithms) but few techniques are available to calculate a reliable DTM. It is sometimes possible to use a reference DTM (generally built manually or semi-manually) but, especially when working on high resolution data, such a reference is often not as accurate as the corresponding DSM: that leads to classical detection problems, typically a high underdetection rate ("False Negative" rate) and a high overdetection rate ("False Positive" rate). In order to make the underdetection rate tend towards 0 and to have the overdetection rate as small as possible, a good DTM i.e a good approximation of the topographic surface is necessary. In this paper, a method for deriving a reliable DTM from a DSM, a building mask (derived from a database) and a vegetation mask is presented and evaluated.
In a DSM generation context, 2 families of techniques can be distinguished: lidar scanning and stereo-matching techniques. Lidar scanning methods have an undeniable advantage in rural areas, where they generally provide both DTMs and DSMs. In dense urban areas, the DTM can not be so easily obtained. Image-based DSM generation has then some advantages over lidar techniques. On the one hand, as images are most of time necessary for photogrammetric projects, generating a DSM with stereo-matching techniques does not implie additionnal costs. On the other hand, images provide a higher degree of internal geometric quality. The main challenge when deriving a DTM from a stereo-matching DSM is to filter and to discard outliers (blunders), i.e points that have too high an elevation compared with their surroundings (See Subsection 2.4 for a list of several kinds of outliers that can be found in a DSM). Almost all the methods found in literature try to deal with this problem, as shown in the following subsection.

1.2 Related Works

Several methods to derive a DTM from a DSM have been considered.
The first method for estimating DTMs is based on morphological operators. A description can be found in (Weidner, 1996). This method is not robust when DSMs contain outliers. To solve this problem, (Eckstein and Munkelt, 1995) introduces the "Dual Rank Filter". Unfortunately, the structuring element remains difficult to define without any a priori knowledge about the study area (urban / industrial ...). Moreover, the method can fail because of big aggregations of vegetation or big buildings (typically a cathedral) in city centres. Eventually, such a tool generally relocates ridges and thalwegs.
An other strategy consists in using parametric methods in order to reconstruct the topographic surface. The final DTM is supposed to belong to a family of parameterized surfaces and these parameters have to be derived from observations . Unfortunately, as shown in (Jordan and Cord, 2004), all the kinds of surface can not be reconstructed and the reconstruction is all the more difficult and inaccurate as the study area is big.
A large set of methods based on triangulation have been found in literature (See (Baillard, 2003) for an example). The main challenge here is to choose ground points and then to triangulate them in order to interpolate height in the whole scene. As the final surface depends on this choice, finding good criteria to select true ground points (and not outliers!) is determinant. An other weak point is that the final surface is not regular (i.e not differentiable) and so not "natural".
A good method that gives regular surfaces is the Elastic Grid. Former works when producing the French Elevation Database have shown its capability to represent the topographic surface naturally and correctly (Masson d'Autume, 1978).

1.3 Presentation

The Elastic Grid method has always been used in order to derive a DTM from a set of extracted points (for example, contour lines). There are no ouliers in input data in that case. Tests have shown the limits of the algorithm in presence of outliers: when applied too roughly, the algorithm creates artificial blobs (See figure 3 in Section 4). The main goal of this study is also to show the feasability of adapting the method to such a context.
In Section 2, input data are first described. In Section 3, our method is detailed. In Section 4, the results of our method are presented and qualitative and quantitative results are given. Eventually, forthcoming research axes are given in concluding remarks

2 INPUT DATA

The algorithm presented in Section 3 uses a DSM, a building mask and a vegetation mask to estimate the final DTM.

2.1 DSM

In our study, 2 stereo-matching algorithms are used to compute the initial DSM. The first one is described in (Baillard and Dissart, 2000) and is based on cost minimization along epipolar lines. This cost takes discontinuities in heights and radiometric similarities into account. The second one is based on a multi-resolution implementation of Cox and Roy optimal flow image matching algorithm. More details are given in (Pierrot-Deseilligny and Paparoditis, 2006). DSMs have a resolution of 20 cm.

2.2 Buildings Masks

The buildings mask is directly derived from a cadastral database. This database is a vector database where buildings ground footprints are represented in 2D. As it is produced manually, it has a good precision but contains some discrepancies (for example demolished buildings), as shown in Subsection 2.4.

2.3 Vegetation Masks

The vegetation mask is produced by applying a threshold on NDVI images (Normalized Difference Vegetation Index). This index is high for vegetation due to the fact most of the visible light is absorbed and nearly all infrared light is reflected.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

This index is computed on orthophotos so that vegetation masks can be easily superimposed on buildings masks and DSMs. RGB and IR images used for orthophotos are calibrated but, to avoid problems linked to the "hot-spot" phenomenon, source images are corrected with an algorithm that performs a radiometric equalization. The model used for this correction is a parameterized and semi-empirical BRDF model. More details can be found in (Paparoditis et al., 2006).

2.4 Comments

Three types of outliers can be distinguished in input data. Firstly, some buildings are not represented in masks as the database is not necessarily up-to-date (Case 1 in Figure 1). Secondly, some above-ground points are out of them: cars, street furniture, newspapers kiosks...(Case 2 in Figure 1). Eventually, a lot of outliers are located at the edges of buildings (Case 3 in Figure 1). That comes from the fact that buildings footprints are given by the outlines of the walls in the cadastral database and that the walls limits do not necessarily fit the roof limits given in a DSM.

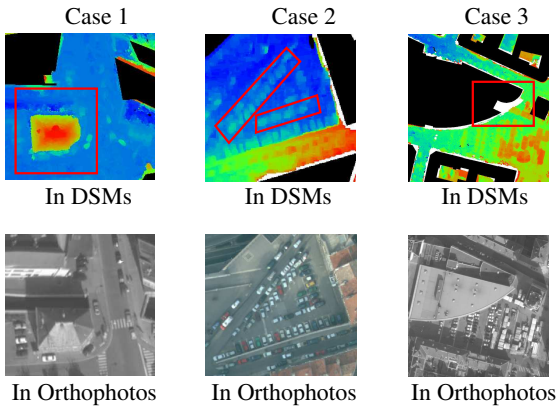

Figure 1: Problems in buildings masks. Outliers (highlighted in red boxes) in above-ground masks sumperimposed on DSMs (upper images) and corresponding orthophotos (bottom images)

3 METHOD

In this section, a short mathematical description of the Elastic Grid method is given. This method is based on a functional (that contains a regularization term and a data term) to minimize. Firstly, the importance of the norm ρ to use in the data term is shown. Secondly, the 3 parameters to be tuned (a tuning constant c intrinsic to the norm ρ, the standard deviation σ and the smoothing coefficient λ) are introduced and justified. Eventually, the general strategy used for the process is detailed.

3.1 Theoretical Aspects

The Elastic Grid method estimates the reconstructed topographic surface by fitting an elastic surface to a finite sample of observation points (i.e points considered as ground points in the DSM i.e points out of above-ground masks). Mathematically, this is equivalent to the minimization of this functionnal:

$$E(z) = K(z) + \lambda G(z, \sigma) \tag{2}$$

- the Regularization Term $K(z)$ corresponds to the discrete approximation of the second derivative of the surface to reconstruct. This term minimizes the mean quadratic curvature (i.e height variations) of the final DTM.

$$K(z) = \sum_{l=1}^{M}\sum_{c=1}^{N}\left(\frac{\partial^2 z_{c,l}}{\partial c^2}\right)^2 + \sum_{l=1}^{M}\sum_{c=1}^{N}\left(\frac{\partial^2 z_{c,l}}{\partial l^2}\right)^2 \tag{3}$$

$$= \sum_{l=1}^{M}\sum_{c=2}^{N-1}\left(z_{c-1,l} - 2z_{c,l} + z_{c+1,l}\right)^2 + \sum_{l=2}^{M-1}\sum_{c=1}^{N}\left(z_{c,l-1} - 2z_{c,l} + z_{c,l+1}\right)^2 \tag{4}$$

- the Data Term $G(z, \sigma)$ corresponds to the distance between the model to estimate and observations.

$$G(z, \sigma) = \sum_{l=1}^{M}\sum_{c=1}^{N}\rho\left(\frac{z_{c,l} - obs_{c,l}}{\sigma}\right) \tag{5}$$

where $z_{c,l}$ is the value of the estimated model at the (c,l) pixel, $obs_{c,l}$ the corresponding observation value and ρ, the norm used in order to calculate the distance.

- the factor λ is used in order to balance both terms. The higher λ is, the better the model fits observations. The smaller λ is, the smoother the model is.

3.2 Description of the grid parameters

The initial Elastic Grid method uses the Least-Squares method to minimize the difference between the model to estimate and observations. Therefore, a classical euclidean norm is introduced in the data term.

$$\sum_{l=1}^{M}\sum_{c=1}^{N}\left(z_{c,l} - obs_{c,l}\right)^2$$

Such an approach is not robust to outliers in input data and can become very unstable. The method used in our work in order to reject outliers from calculation is derived from the M-estimator technique. This technique reduces the effect of outliers by replacing the sum of squared differences (residuals) by a certain function ρ that is symmetric, positive-definite, with a minimum at zero and less increasing than square.

$$\sum_{l=1}^{M}\sum_{c=1}^{N}\rho(z_{c,l} - obs_{c,l})$$

Name of the Tested Norm	Tested Norm $\rho(x)$
L1L2	$2 \times (\sqrt{1 + \frac{x^2}{2}} - 1)$
Cauchy	$\frac{c^2}{2} \times \log\left(1 + \left(\frac{x}{c}\right)^2\right)$
Geman-McLure	$\frac{\frac{x^2}{2}}{1+x^2}$
Huber $\begin{cases} if\|x\| < c \\ if\|x\| \geq c \end{cases}$	$\frac{x^2}{2}$ $c \times (\|x\| - \frac{c}{2})$
Tukey $\begin{cases} if\|x\| < c \\ if\|x\| \geq c \end{cases}$	$\frac{c^2}{6} \times (1 - (1 - (\frac{x}{c})^2)^3)$ $\frac{c^2}{6}$

Table 1: Robust norms tested in our study

All the norms tested in our study are listed in Table 1. In most norms, there is a tuning constant c. It is all the more important as it determines points whose influence will be reduced in the process. In (Zhang, 1997), the author considers that noise follows a gaussian law $\mathcal{N}(0,\, 1)$ and gives, for each norm, the value for the tuning constant c in order to reach the 95 percent asymptotic efficiency on the standard normal distribution. (Zhang, 1997) shows that $c = 4.6851$ for the Tukey's norm for instance.
In our work, the difference $z_{c,l} - obs_{c,l}$ is assumed to follow a gaussian law but is not standardized $(z_{c,l} - obs_{c,l}) \sim \mathcal{N}(0,\, \sigma)$, what prevents us to apply the previously mentioned values directly. Therefore, a standard deviation σ must be calculated. It is calculated with the classical estimator, in a clean and horizontal area (typically a square without any tree, car...) so that it is not biased because of the presence of outliers.

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(r_i - \overline{r}_i)^2} \qquad (6)$$

where n is the number of pixels in clean areas, $r_i = z_i - obs_i$ is the difference between the estimated model and correponding observations and $\overline{r}_i = \frac{1}{n} \times \sum_{i=1} r_i$ is the mean value of differences in clean areas.

3.3 General Strategy

As can be seen in Figure 2, several steps are necessary in order to compute the final DTM. The process is divided into 4 steps: Initialization, Paving, Elastic Grid and Mosaicking.

As the process to minimize $E(z)$ is iterative, a good way to decrease the number of iterations is to calculate an initial solution. This approximate solution is given by a method based on a dual rank filter. This tool has some imperfections mentioned in Subsection 1.2 but is fast and easy to implement. Moreover, as the convergence is all the more long as the study area is big (the ratio "calculation time" / "study area size" is not linear), a paving strategy (with a 1000×1000 tile) has been set up. A mosaicking process is consequently necessary.

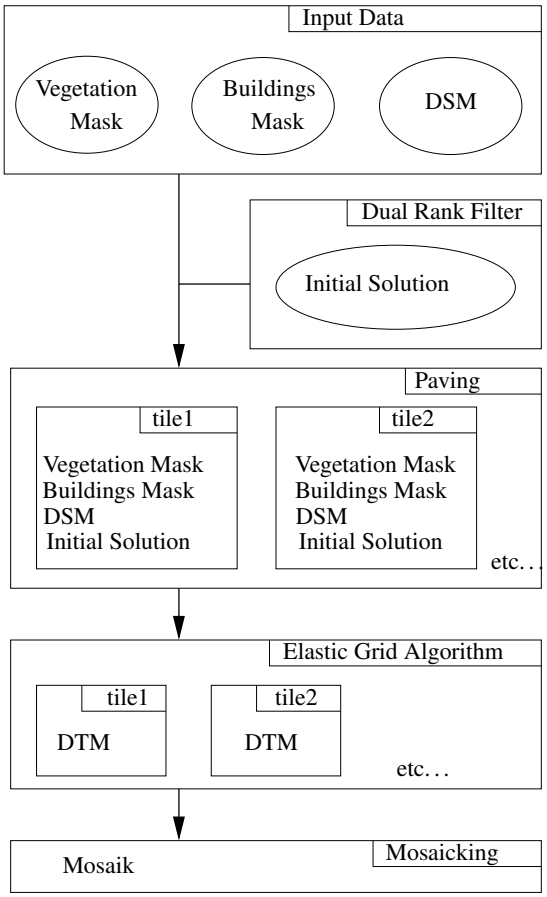

Figure 2: General Strategy

4 RESULTS AND DISCUSSION

4.1 Test areas and data

2 test sites are presented:

- Amiens City Centre, France
  - Pixel Ground Size = 20cm
  - Area = 800m $\times$ 800m $\simeq 0.64km^2$
  - Terrain Type: relatively flat
  - Land Cover Type: dense urban area
  - Matching Algorithm: (Baillard and Dissart, 2000)
- Marseille City Centre, France
  - Pixel Ground Size = 20cm
  - Area = 950m $\times$ 950m $\simeq 0.90km^2$
  - Terrain Type: hilly

– Land Cover Type: dense urban area

– Matching Algorithm: (Pierrot-Deseilligny and Paparoditis, 2006)

Several norms have been tested in our work. For each norm, the tuning constant c is firstly found in literature. Secondly, as our data are not standardized, a standard deviation must be calculated in a clean area to standardize them and to be able to apply the value of c found in literature. Once these 2 factors fixed, a sensitivity study is carried out in a small area (typically, a 300 $\times$ 300 area) to determine the best value to give to the smoothing coefficient λ. Results are assessed by visual inspection (difference between the DSM and the DTM) and by editing profiles along lines in the DSM and corresponding lines in the DTM. Experiments have shown the terrain is best reconstructed with a Tukey's norm. This norm has also been used with the 3 factors (c, σ and λ) previously determined to process the whole area. The corresponding results are given in the next subsection.

4.2 Results

As the process to minimize $E(z)$ is long, an optimized numerical library (GNU Scientific Library) is used. The calculation time with a 1.8 GHz PC is about 30 hours in Amiens (size of the whole scene: 4000 $\times$ 4000 / 36 tiles) and 40 hours in Marseille (size of the whole scene: 4600 $\times$ 4600, 49 tiles). The qualitative and quantitative results of our algorithm are now given.

4.2.1 Qualitative Results The benefit of introducing a robust norm (instead of the classical euclidean norm) is clearly shown in Figure 3 where results in a small test area are presented. The initial DSM is displayed on the upper left image. Above-ground masks are sumperimposed on the initial DSM and are displayed on the upper right image. Final DTMs are displayed on bottom images (on the left, the one processed with the classical Elastic Grid, on the right, with our algorithm). All the figures are displayed by using the same scale in height as the bottom right DTM. In this way, readers can have a first visual idea of the quality of DTMs. In the bottom left DTM, artificial blobs are created. That comes from outliers present in input data: as they are considered by the classical Elastic Grid algorithm as true ground points, they have the same influence in the process and deviate the reconstructed topographic surface upwards. As the new algorithm introduces a robust norm in order to reject outliers, such a bad effect does not occur.

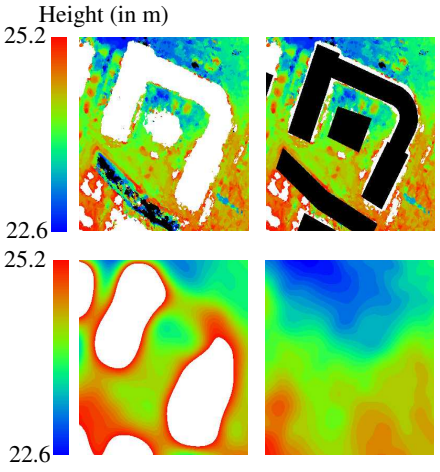

Figure 3: Comparison between the classical Elastic Grid algorithm and the new algorithm. In white, points higher than 25.2m

The qualitative results when applying our algorithm are now given in Figures 5 - 8 (Amiens) and Figures 9 - 12 (Marseille). As detailed in the introduction, a good detection of above-ground points implies a good reconstruction of the topographic surface. Therefore, the difference $DSM - DTM$ is a good indicator for assessing the quality of final products and is given in Figure 8 (Amiens) and Figure 12 (Marseille).

4.2.2 Quantitative Results In order to assess results quantitatively, a statistical analysis is firstly performed by manually extracting ground points from initial DSMs and by comparing them with corresponding points extracted from DTMs. A bias, a standard deviation and a RMS are then calculated and are shown in Table 2. Secondly, profiles along arrows in DSMs and corresponding arrows in DTMs are edited (See Figures 4 - 13 - 14).

Area	Bias	σ	RMS	Nb Pts
Amiens City Centre	0.424	0.447	0.616	1104
Marseille City Centre	0.307	2.606	2.604	886

Table 2: Stastistical Analysis

The topographic surface is well reconstructed in Amiens. The bias is slightly positive. That means the reconstructed surface is slightly disturbed by the presence of outliers. Nevertheless, as shown in the profile (Figure 4) along the green arrow (Figures 5 and 7), the final DTM perfectly clings to points in streets and courtyards and reconstructs all the small undulations of the terrain.

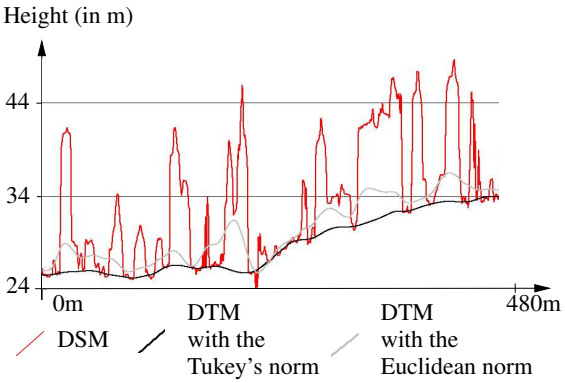

Figure 4: Profiles in Amiens along the green arrow in the DSM and DTM (See Figures 5 and 7). In red, the initial DSM. In black, the result with the new algorithm. In light grey, the result of the classical Elastic Grid. The blobs that correspond to buildings and vegetation in the DSM are filtered with the new method. Artificial blobs are created when using the euclidean norm.

Some problems occur in Marseille. The bias is small and proves that the computed DTM is a good approximation of the topographic surface. As shown in the profile (Figure 13) along the green arrow in the DSM and DTM (Figures 9 and 11), the final DTM generally clings to true ground points and filters blobs corresponding to buildings. In that case, results are similar to Amiens. Nevertheless, a high RMS outlines problems in specific areas, especially in breaklines areas. For example, some problems occur in the profile along the red arrow (Figure 14). The left breakline is well reconstructed, which proves the capability of our algorithm to reconstruct such terrain types. Nevertheless, the right breakline is completely eroded. This problem firstly comes from the terrain type (a 50m high cliff difficult to reconstruct), secondly from the combination of using a robust norm and a coarse initial solution: using a robust norm is an efficient means

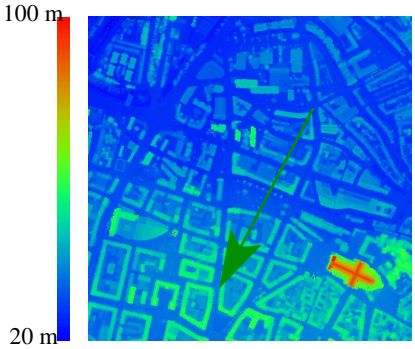

Figure 5: Amiens - Initial DSM (Top View)

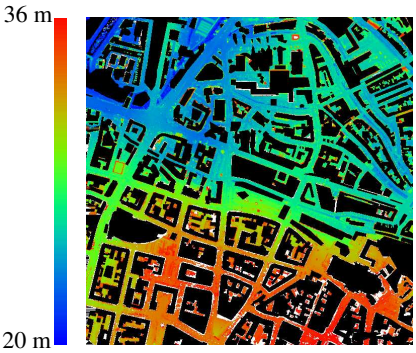

Figure 6: Amiens - Mask over DSM (Top View)

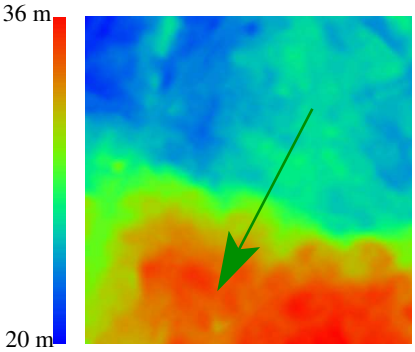

Figure 7: Results in Amiens - DTM (Top View)

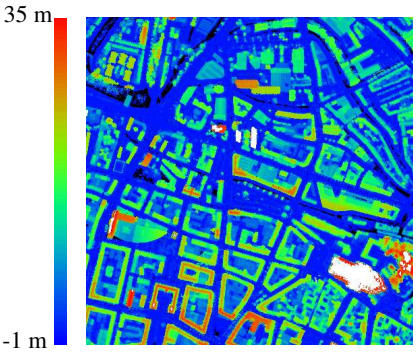

Figure 8: Results in Amiens - $DSM - DTM$ (Top View)

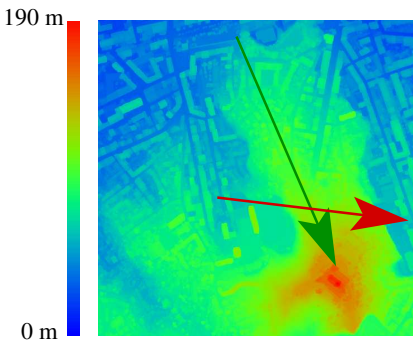

Figure 9: Marseille - Initial DSM (Top View)

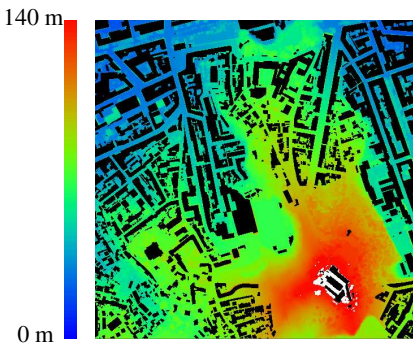

Figure 10: Marseille - Mask over DSM (Top View)

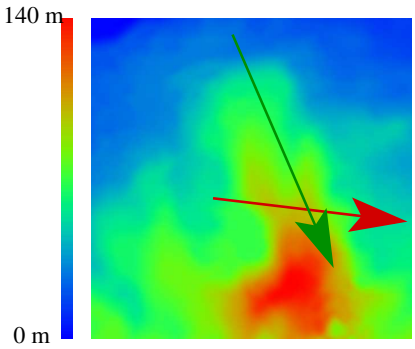

Figure 11: Results in Marseille - DTM (Top View)

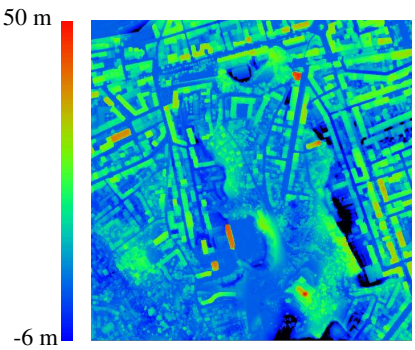

Figure 12: Results in Marseille - $DSM - DTM$ (Top View)

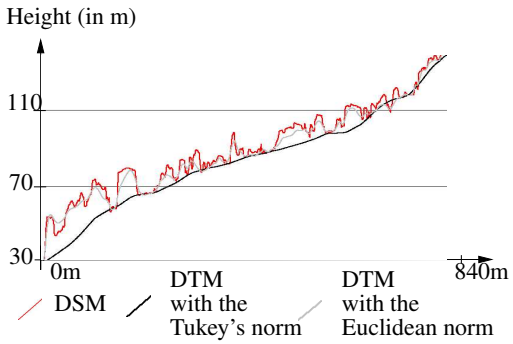

Figure 13: Profiles in Marseille along the green arrow. In red, the initial DSM. In black, the result of the new algorithm. In light grey, the result of the classical Elastic Grid. The blobs that correspond to buildings and vegetation in the DSM are filtered with the new method.

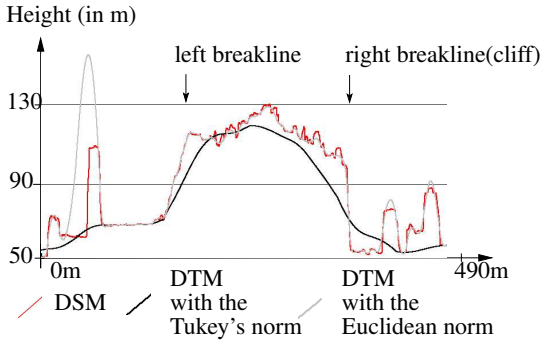

Figure 14: Profiles in Marseille along the red arrow. In red, the initial DSM. In black, the result of the new algorithm. In light grey, the result of the classical Elastic Grid. The left breakline is well reconstructed. The right breakline is too eroded (Problems in the initialization step).

to reject definitely outliers; problems occur when rejected points are inliers (true ground points). In Marseille, the right breakline is eroded in the initialization step (because of the use of a dual rank filter, see Subsection 1.2 for more explanations). As the difference $z_{c,l} - obs_{c,l}$ is then too big in that area, all the corresponding points (even inliers) are considered outliers: the process does not use these points to reconstruct the topographic surface and the breakline is not modelled very well in the end. A multiresolution coarse-to-fine approach is being considered to give our algorithm a more precise initial solution.

5 CONCLUSIONS AND FUTURE WORK

Our goal was to compute a DTM from a DSM and above-ground masks, in dense urban areas and in a difficult context (presence of outliers in input data). The initial Elastic Grid has been revised by introducing a robust norm (instead of the classical euclidean norm) and by setting the grid parameters (the tuning constant c, the standard deviation σ and the smoothing coefficient λ) suitably. The results presented in this paper and corresponding to different configurations (high resolution DSM, relatively flat / hilly city centres) show the robustness of our approach. In order to make our method as generic as possible, tests have been carried out with lower resolution DSMs (pixel ground size = 70cm and 5m). First results are promising and show the feasability of extending our method to such a resolution. An other research axis is to adapt our method so that the initial above-ground mask becomes optional. The challenge here is to reject above-ground points rapidly. On the one hand, the points where the difference between the model and observations is negative and that are also closer to the model to estimate (typically ground points) must have their influence in the calculation increased. On the other hand, points with a positive difference (typically above-ground points) must be rejected. In (Jordan et al., 2002), a dissymetric norm is used for that purpose: the euclidean norm is used wherever the difference $z_{c,l} - obs_{c,l}$ is negative and the Tukey's norm is used wherever it is positive. As the euclidean norm is more increasing than the Tukey's norm, lowest points are advantaged. The main idea is to introduce such a dissymetric norm (by replacing the non robust euclidean norm with the more robust L1L2 norm) in our grid data term.

6 ACKNOWLEDGEMENTS

The first author would like to thank his colleagues for useful and valuable comments, suggestions and discussions.

REFERENCES

Baillard, C., 2003. Production of urban DSMs combining 3D vector data and stereo aerial imagery. In: ISPRS Archives, Vol. XXXIV, Part3/W8.

Baillard, C. and Dissart, O., 2000. A stereo matching algorithm for urban Digital Elevation Models. ASPRS Photogrammetric Engineering and Remote Sensing 66(9), pp. 1119–1128.

Eckstein, W. and Munkelt, O., 1995. Extracting objects from digital terrain models. In: Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes, Vol. 2572, pp. 43–51.

Jordan, M. and Cord, M., 2004. Etude portant sur l'état de l'art, l'évaluation et la comparaison de méthodes de segmentation sol / sursol à partir de modèles numériques d'élévation. Technical report, ETIS, Equipe Image, ENSEA.

Jordan, M., Cord, M. and Belli, T., 2002. Building detection from high resolution digital elevation models in urban areas. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV-3B, Graz, Austria, pp. 96–99.

Masson d'Autume, G., 1978. Construction du modèle numérique d'une surface par approximations successives: application aux modèles numériques de terrain. Bulletin de la Société Française de Photogrammétrie et Télédétection 71, pp. 33–41.

McKeown, D., Bulwindle, T., Cochran, S., Harvey, W., McGlone, J. and Shufelt, J., 2000. Performance evaluation for automatic feature extraction. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII-B2, Amsterdam, The Netherlands, pp. 379–394.

Paparoditis, N., Souchon, J., Martinoty, G. and Pierrot-Desseiligny, M., 2006. High-end aerial digital cameras and their impact on the automation and quality of the production workflow. IJPRS. To Appear.

Pierrot-Deseilligny, M. and Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery. In: International Archives of Photogrammetry and Remote Sensing, Vol. XXXVI, Ankara, Turkey.

Weidner, U., 1996. An approach to building extraction from digital surface models. In: Proceedings of the 18th ISPRS Congress, Comm. III, WG 2, Vol. 43 - Building Detection from a Single Image, pp. 924–929.

Zhang, Z., 1997. Parameter estimation techniques: A tutorial with application to conic fitting. Image and Vision Computing Journal 15(1), pp. 59–76.