BUILDING FOOTPRINT DATABASE IMPROVEMENT FOR 3D RECONSTRUCTION: A DIRECTION AWARE SPLIT AND MERGE APPROACH

Bruno Vallet and Marc Pierrot-Deseilligny and Didier Boldo

IGN - Laboratoire MATIS 2/4 avenue Pasteur - 94165 Saint-Mand Cedex, France
bruno.vallet@ign.fr - http://recherche.ign.fr/labos/matis

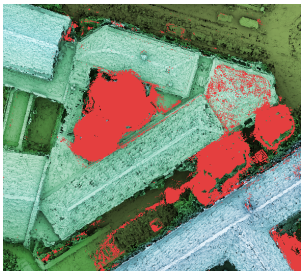
Commission III/3

KEY WORDS: Photogrammetry, 3D reconstruction, building footprint, split and merge, segmentation

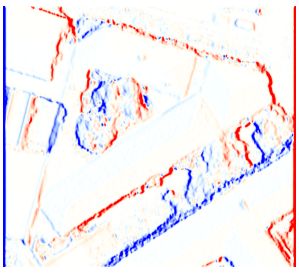
ABSTRACT:

In the context of 3D reconstruction of wide urban areas, the use of building footprints has shown to be of great help to achieve both robustness and precision. These footprints however often present inconsistencies with the data (more than one building in the footprint, inner courts, superstructures...) This paper presents a fast and efficient algorithm to enhance the building footprint database in order to make subsequent 3D reconstructions easier, more accurate and more robust. It is based on a segmentation energy that is minimized by a split and merge approach. The algorithm is demonstrated on a wide urban area of one square kilometer.

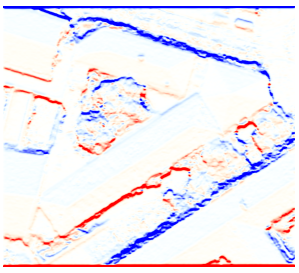
(a) Orthophotography and footprint



(b) Shaded DEM and vegetation mask



(c) Horizontal gradient



(d) Vertical gradient

Figure 1: Input to our algorithm

1 INTRODUCTION

The production of 3D models of urban areas has received a lot of attention from the scientific community in the last decade because of the broad range of its applications and the increase in both quality and quantity of data. In this setup, it becomes more and more crucial to design flexible tools to help human operators achieving efficient and accurate reconstruction of wide urban areas.

1.1 Problem statement

The problem of urban reconstruction consists in finding a 3D model (in general a polygonal surface) that is as coherent as possible with the input data. In our case where the footprints of the buildings are given, we can use the efficient and robust approach proposed in (Durupt and Taillandier, 2006). However, this approach relies heavily on the quality of the building footprint database, and might fail if the building to be reconstructed contains altimetric discontinuities that are not present in its footprint. This often happens in practice, and especially when:

- Two (or more) adjacent buildings with different roof heights share the same footprint.
- The real footprint of a building is only a portion of the footprint in the database (gardens, inner courts,...)
- The building has some superstructures which sizes and heights are not negligeable with respect to the expected precision of the reconstruction. This problem becomes increasingly difficult as reconstructions gain in precision, and has already been tackled in the context of photogrammetry (Bredif et al., 2007) (Dornaika and Bredif, 2008).

More difficult cases are often a combination of the three cited above, and require a manual intervention to enable a further reconstruction. In general, this intervention consists in subdividing the footprint by cutting through all (or most of) the altimetric discontinuities. In a production framework, where large areas need to be extensively reconstructed, it appears that this building footprint database enhancement step is one of the most time consuming. Hence, the problem that we tackle in this paper is that of automatizing this enhancement as a required preprocessing step to 3D reconstruction. More precisely, our problem is to segment a polygonal footprint into a set of non-overlapping polygonal sub-footprints that cover it entirely, such that the interface between the sub-footprints corresponds to altimetric discontinuities. This is a problem of segmentation of vector data (building footprints database) guided by raster data (photos, DEM,...)

1.2 Available data

The data available in our study mainly consisted of:

- A set of 10 centimeter resolution aerial images with a high recovery ratio around 60% (intraband + interband) in order to ensure that each ground point is seen in at least 4 images, covering an area of one square kilometer. The images are in RGBI (the infrared channel is used to obtain the vegetation mask).
- A vectorized cadastral map giving building footprints for the same area. It consists in a set of polygonal footprints given by their ordered list of points in ground coordinates (Figure 1(a), green).

From this initial data, existing algorithms can be run to extract:

- A Digital Elevation Model (DEM) over the whole area (Figure 1(b)). It was obtained by dense correlation following (Roy and Cox, 1998) and the implementation described in (Pierrot-Deseilligny and Paparoditis, 2006).
- The gradient of the DEM (Figures 1(c) and 1(d)) computed using a standard Canny-Deriche filter (Deriche, 1987).
- An orthophotography of the area (Figure 1(a)).
- A vegetation mask (Figure 1(b), red) obtained by the method exposed in (Iovan et al., 2007).

The initial data and extracted data form the *input* to our algorithm.

1.3 Previous works

The idea of using a 2D building footprint to enhance 3D building reconstruction first appeared in (Pasko and Gruber, 1996), and was developed in (Roux and Maitre, 1997), (Brenner, 2000) and (Jibrini et al., 2000). This idea is also at the core of the reconstruction method (Durupt and Taillandier, 2006) for which we designed our building footprint enhancement algorithm, and to the more general framework (Taillandier, 2005) from which it derives. In the context of laser data, it is also central to the works of Vosselman *et. al.* (Vosselman and Dijkman, 2001) (Vosselman and Suveg, 2001) (Suveg and Vosselman, 2001).

To the best of our knowledge, segmentation of building footprints has never been decoupled from the reconstruction itself as done in this paper, but used to find directly planar regions.

1.4 Proposed approach

In this paper we call $\mathcal{P}$ the polygonal footprint to segment, $\mathcal{P}_i$ the polygonal sub-footprint resulting from the segmentation and $I_i^j = \mathcal{P}_i \cap \mathcal{P}_j$ the *interface* between two sub-footprints (it is an edge or set of edges in some cases). The result of our algorithm is a *segmentation* of $\mathcal{P}$ that is given indifferently by the set of sub-footprints $\mathcal{P}_i$ or by the *interface* $\mathcal{I} = \cup_{i<j} I_i^j$ between the $\mathcal{P}_i$ (it is a set of edges).

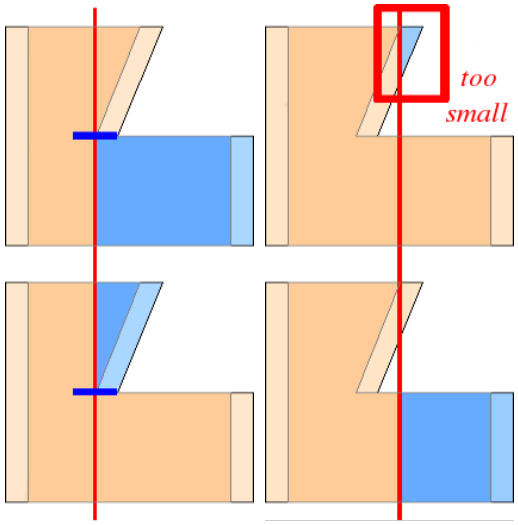
The approach that we propose consists in defining an energy that is negative (resp. positive) on edges that are likely (resp. unlikely) to be altimetric discontinuities, and to find the segmentation that minimizes the sum of this energy over the edges of $\mathcal{I}$. We start by choosing a gradient threshold T_∇ such that we consider that a point where the gradient value is above (resp. below) T_∇ is likely (resp. unlikely) to be on an altimetric discontinuity. The energy on an edge e can then be defined as:

$$E(e) = \int_{P \in e} T_\nabla - |\nabla z(P).\overrightarrow{n}(e)| dP \qquad (1)$$

where z is the height at point P given by the DEM and $\overrightarrow{n}(e)$ is a unit vector normal to e. As required, $E(e)$ is negative when the mean absolute gradient across e is greater than T_∇.

To simplify this problem, and gain in robustness and quality, we will restrict the directions of the interface edges to follow directions present in the original footprint, which is not a strong conditional assumption. This proved to be true on most examples that we have tested. In order to solve this problem, we propose a split and merge approach based on principal directions detected on the initial footprint $\mathcal{P}$:

1. Cluster the directions of the footprint's edges in a direction space taking their lengths into account.
2. Recursively split the footprint along lines of minimal energy.



(a) Two hypotheses for a snapped cut (b) Two hypotheses for a double cut

Figure 2: Cutting hypotheses. The eroded footprint is darkened.

3. Merge the resulting sub-footprints in order to minimize $E(\mathcal{I})$.

The first step is a simple clustering in the space of line angles (modulus π), and does not require special care. Simply notice that we should keep the number of direction clusters as small as possible, for instance by eliminating the clusters which edges' length sum is smaller than a given threshold, or a ratio of the "largest" cluster.

In our algorithm, we will often need to compute energies of the form given by (1) thus to access the gradient across edges that can only be in a limited number of directions. Thus for efficiency reasons, we will precompute the gradient for each direction on a grid aligned with the direction and with the same resolution than the DEM. These grids will serve a double purpose as they will also be used to discretize our cutting lines.

2 RECURSIVE SPLIT

2.1 Cutting hypotheses

For each direction, we will discretize the set of possible cut lines C_i as the lines passing through the (center of) rows of pixels in our grids for each direction. This way the integral of the gradient over an edge in this line's direction will simply be computed as a sum over pixels of the same row in the grid.

As our input footprint might not be convex, a cut might generate more than 2 sub-footprints. In this case, the same cut line C_i generates several cutting hypotheses, one for each edge of $P \cap C_i$ (see Figure 2(b)). Similarly, we snap our cuts by prolongating the initial footprint's edges, and generating a new cut hypothesis for each part of the cut (see Figure 2(a)). This way, each cutting hypothesis consists of the two footprints generated by the split, and their interface I which is a single edge.

This process however can introduce extremely poorly shaped footprints and small footprints that are not desired in the final solution. To prevent the occurrence of such bad geometries, we build an erosion $\mathcal{P}_e$ of the footprint $\mathcal{P}$ by a centered segment of length d orthogonal to the current direction (see Figure 3). This erosion is then used to discard the cutting hypotheses for which:

$$|I \cap \mathcal{P}_e| < |I|/2 \qquad (2)$$

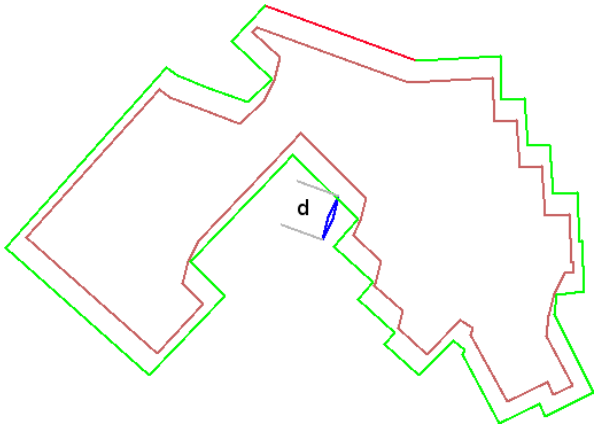


Figure 3: Erosion of the input footprint (green) by a flat rhombus (blue) of height d orthogonal to a main direction (red).

which means that the splitted footprints have a width of at least d on at least half of their length. Hence, the parameter d is used to indicate minimum expected size of a footprint. For instance the hypothesis in Figure 2(b) (top) is discarded because the top right triangle satisfies this criterion (it has $|I \cap \mathcal{P}_e| = 0$). This geometric criterion proved to be the most robust in our experiments, and it was implemented using the CGAL Minkowsky sums. Note that we replaced the segment by a flat rhombus to avoid degeneracies.

2.2 Cut score

For each cut hypostesis, we can compute a cutting score as the energy $E(\mathcal{I})$ restricted to the cut. To enhance this estimation we take into account the following facts:

- An existing edge corresponds to an altimetric discontinuity. Hence the gradient in its vicinity should not be taken into account for the score of a new cut. Thus the meaningful zone is defined by the erosion of the footprint by a centered segment. Ideally, the length of this segment should equal the size of the kernel used to compute the gradient. In practice, it should be even greater as the edges of the footprint are not exactly located on discontinuities. We chose the same length d as before, such that we only need to compute one erosion per footprint and per direction. We chose to compute the erosion with CGAL's exact arithmetics as we encountered failure cases using inexact computations. This is quite time consuming, such that the choice of taking the same parameter is really saving us time.

- Vegetation hides the geometry of the building so the DEM will be considered not pertinent within the vegetation mask.

- The DEM is more inaccurate in shadowed areas.

These three facts are integrated in the computation of $E(\mathcal{I})$ by weighting the gradients by a *confidence* term that is 0 outside the eroded footprint and in vegetation areas, and elsewhere proportional to luminosity.

2.3 Recursion

For the input footprint $\mathcal{P}$, we can build the cutting hypotheses (Section 2.1) and their scores (Section 2.2). We select the cutting hypothesis with the lowest score and apply it to the footprint $\mathcal{P}$, which splits it into two sub-footprints $\mathcal{P}_1$ and $\mathcal{P}_2$. We apply this process again to $\mathcal{P}_1$ and $\mathcal{P}_2$, and so on recursively.

To ensure that our cuts minimize E, we stop the recursion when the lowest score becomes positive. In that case the footprint is *final* and will not be splitted. Our shape criterion (2), guarantees that the width of the resulting sub-footprints is greater than d in each direction.

2.4 Results

As figure 4 shows, the segmentation resulting from the recursive split runs through most of the altimetric discontinuities. However, the segmentation presents many undesired cuts as our cuts are straight so they run through the whole footprint when they may correspond to much more local altimetric discontinuities. To achieve a better segmentation, and further minimize our energy, we need to remove these superfluous cuts by merging sub-footprints whenever this improves the energy $E(B)$.

3 MERGE

3.1 Geometric polygon merging

Merging the sub-footprints resulting from the splitting process can be tricky as numerical precision forces us to use thresholds to determine whether two edges from different polygons touch or not. To make the merge process independent from numerical precision and thresholds, we label all edges produced during the splitting process by (a pointer to) the cut line that produced it. This way, the merging algorithm is both robust and simple:

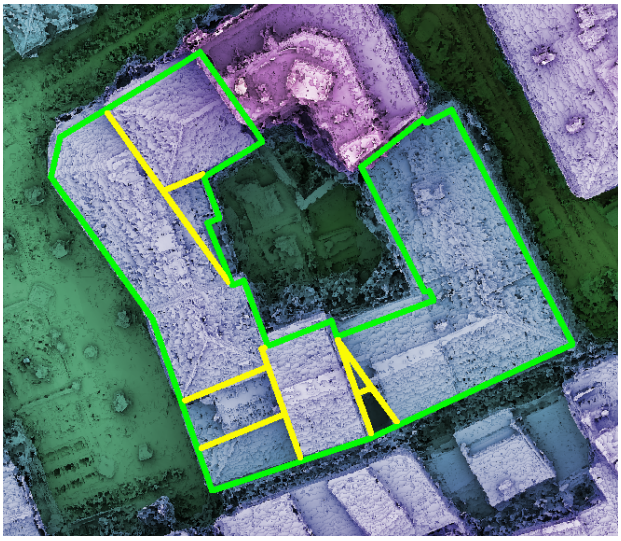
1. For each pair of edges $e_k^i \in P_i$ and $e_l^j \in P_j$ belonging to the same cut line:
   - Compute the intersection edge $e_{k,l} = e_k^i \cap e_l^j$
   - If $e_{k,l} \neq \emptyset$, add $e_{k,l}$ to $I_{i,j}$.
2. Build the connected components of $I_{i,j}$. If there are more than one, this means that the merged footprint has holes. We need to prevent these holes to appear as they are harder to handle in the reconstruction process. To do so, we keep only one connected component in $I_{i,j}$ (the longest or the one with lowest score).
3. Build the merged footprint $P_{i,j}$:
   - For each interface edge $e_{k,l} \in I_{i,j}$ tag e_k^i and e_l^j as interface edges.
   - Build the connected components C_i and C_j of edges of P_i and P_j not tagged as interface.
   - Connect the endpoints of C_i and C_j (this is unambiguous if P_i and P_j where properly oriented).

3.2 Merging algorithm

The merging process goes as follows:

1. Compute all possible merges, their interfaces $I_{i,j}$ and scores $S_{i,j} = E(I_{i,j})$.
2. Build a priority queue of all merges, where the priority is the score $S_{i,j}$. Remember that a high score means it is likely that the interface is not an altimetric discontinuity so it should be removed from the final cut.
3. While the merge with highest priority is positive:
   - Apply the merge with highest priority $S_{i,j}$ between footprints P_i and P_j by replacing P_i and P_j by their union $P_{i,j} = P_i \cup P_j$.
   - Remove all merges involving P_i and P_j from the priority queue.
   - Compute all possible merges involving $P_{i,j}$, their interfaces, their scores, and add them in the priority queue.

(a)

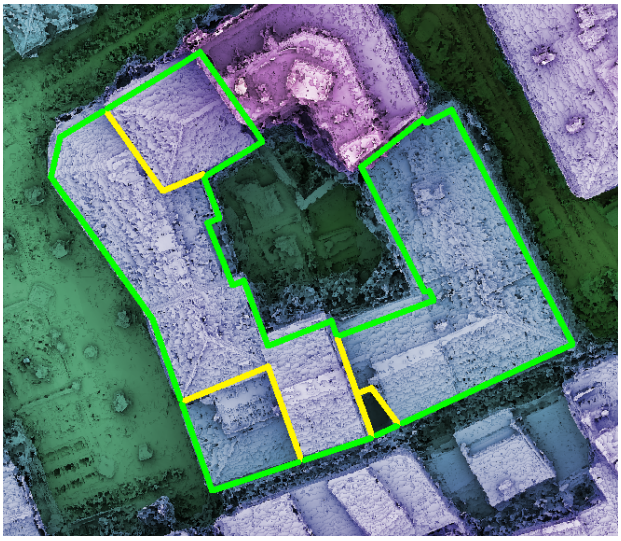


(b)

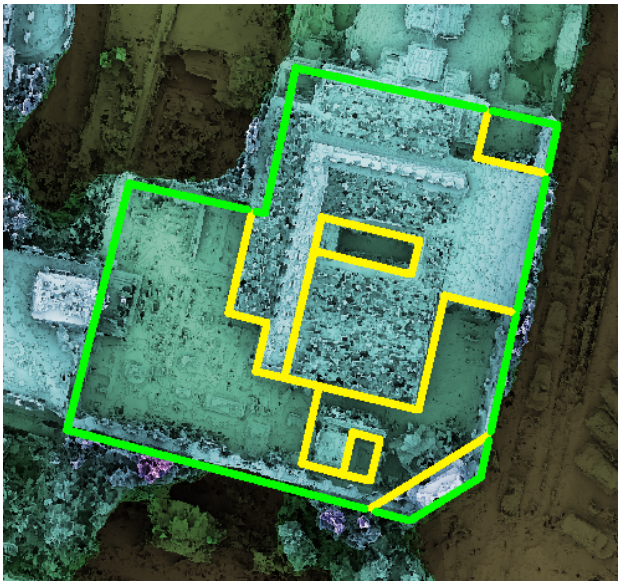
(c)

Figure 4: Results of the splitting process

(a)



(b)



(c)

Figure 5: Results of the merging process

3.3 Results

The merging process ensures that the result is a valid segmentation of the input footprint into a set of sub-footprints. As seen in Figure 5, the algorithm is general enough to allow for a broad range of possible sub-footprints, while being constrained enough (in particular by the allowed directions and minimum size d) to avoid overly complex shapes. The advantage is that such simple shapes are proper for reconstruction. The inconvenient is that if discontinuities do not follow the detected directions, they will not be detected and lead to inconsistencies. Finally, note that as we prevent holes from appearing, inner courts stay connected to the outer boundary (there are two examples of that behavior in Figure 5(c))

4 DISCUSSION

Our method allows for a much more accurate 3D reconstruction on footprints with inner altimetric discontinuities as shown in Figures 6 and 7. However, it sometimes misses some global cuts that are obvious to the eye but do not correspond to altimetric discontinuities. For instance Figure 6 show that a single (and small) handmade cut relying more on a global perception of the footprint shape than on an altimetric discontinuity allows for a great improvement of the result.

This method is proposed as a tool to support the reconstruction of wide urban areas. The splitting and merging results shown here are all obtained based on the same parameters. The tuning parameters are mainly the erosion width d that controls the minimum footprint size and gradient threshold T_∇ that serves to specify the limit between what is a discontinuity and what is not. They are intuitive and simple to tune. In practice, we used the same standard parameters ($d = 1.5m$, $T_\nabla = 3.5$) to process an entire 1km by 1km working area.

Step	(a)	(b)	(c)
Load inputs	0.27	0.4	0.27
Precompute	0.24	0.44	0.12
Erosions	0.2	0.19	0.36
Scores	0.15	0.2	0.1
Splits	0.23	0.17	0.15
Merge	0.01	0.01	0.04
Total	1.1	1.41	1.04

Table 1: Timings (in seconds on a 2.8GHz Pentium 4 processor) of the different steps of the algorithm. The three columns correspond to the examples shown on figures 4 and 5.

In terms of computation time, the algorithm is extremely fast (see table 1). This makes it possible to process very wide working zones rapidly, or to tune the parameters interactively.

The algorithm is heavily dependant on the quality of the input DEM, and only very weakly on the orthophotography and vegetation mask (the latter only serves when the footprint contains vegetation that has an important impact on the DEM, which is quite rare). The most important problems that we encountered are:

- The DEM has a poor quality on shadows as it requires a good contrast. As roughly half of the altimetric discontinuities generate a shadow at their bottom, half of the altimetric discontinuities are not accurately represented in the DEM. We simply added a confidence parameter to handle this issue, but we believe some more adequate solutions can be found.

- If the footprint contains an important altimetric discontinuity that is not aligned with one of the clustered direction, it will perturb the splitting as it will add an important factor to the energy of all cuts not exactly orthogonal to it. To limit this effect we penalized wrong gradient directions by weighting the gradient by a factor $\max\left(0, \cos(2(\vec{n}, \vec{\nabla z}))\right)$ that smoothly decreases from 1 (perfect direction) to 0 for angles greater that $\pi/4$.

- Superstructures cause altimetric discontinuities that are often close to or higher than discontinuities between different buildings. Thus they may generate cuts even with a fine tuning of T_∇. A possible remedy would be to implement a superstructure detection such as (Bredif et al., 2007) prior to cutting.

The energy that we use matches closely the Mumford and Shah segmentation formulation (Mumford and Shah, 1989) except that it has no data attachment term. This drawback is inherent to the problem that we pose, and its consequence will be that we lack of a global quality measure. This will sometimes lead to a lack of global coherence, such as missing a small cut that would enhance greatly the reconstruction (see Figure 6). A workaround would be to interact with the reconstruction method, and for instance only split footprints on which the reconstruction is bad (far from the DEM). As this estimation needs to be done many times, this would require the reconstruction to be very fast, which is not the case for the one that we were working with (at least for complex footprints).

The fact that this energy is not necessarily positive makes it impossible to minimize with graph cuts based segmentation where the non-negativity of weights is a fundamental requirement (Kolmogorov and Zabih, 2004). However, this energy is very natural for segmenting with an unknown a priori number of regions, as minimizing this energy will naturally lead to an optimal number of region, without the need to specify a source/sink pair. For instance, not cutting is a solution like any other, and it has its own energy that can be optimal in the case that no segmentation is required (which is the case on many footprints that are adequate for reconstruction without enhancement). In contrast, graph cut energy is always lower for not cutting than for cutting, and the result is in fact the optimum over bipartition. The drawback is that we cannot use the very efficient graph cut algorithm and need a heuristic approach with no guarantee on optimality.

5 CONCLUSIONS AND FUTURE WORK

We have presented an algorithm to split cadastral maps into smaller regions proper for subsequent 3D reconstruction. The algorithm has only be tested for one reconstruction method but the authors believe it might be a useful preprocessing step to any 3D reconstruction method based on the cadastral map or any other vectorial footprint of the building to reconstruct. The algorithm is simple and fast, as it has been designed with the purpose of helping reconstruction of large urban areas.

In the future, we plan on running this algorithm in a production framework to have a better feedback on its large scale usability. We will also look into correcting the DEM in shadowed area, or maybe detection of altimetric discontinuities directly based on correlation in the aerial images. Finally, we will look into less heuristic means of minimizing our energy, especially in the merging phase.

ACKNOWLEDGEMENTS

The work reported in this paper has been performed as part of Cap Digital Business Cluster TerraNumerica project.

In addition, the authors wish to thank Grgoire Maillet for the important feedback on the usability of the algorithm for productive purposes, Mathieu Brédif for his sound scientific advice and Mélanie Durupt for her help with handling the data.

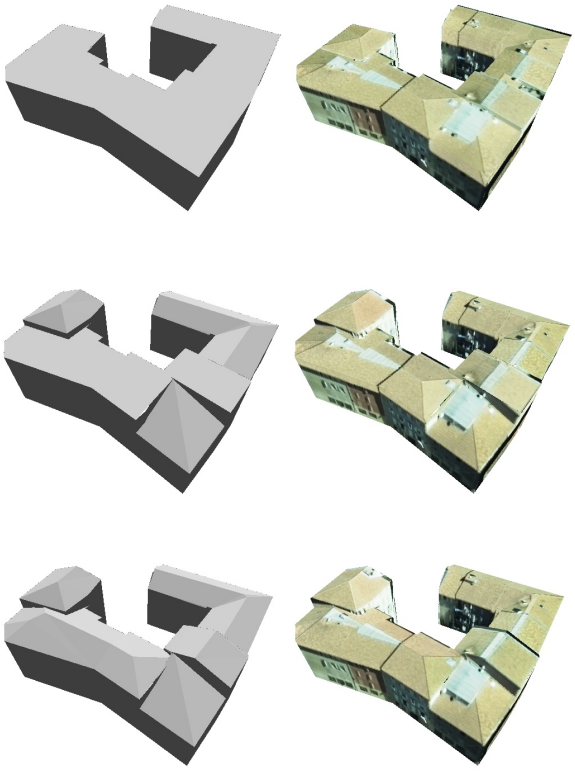


Figure 6: Reconstruction results on the example of Figures 4(b) and 5(b). From top to bottom: reconstruction without enhancement, with enhancement, with enhancement and a single manual cut. This manual cut improves greatly the result but cannot be detected based on our method as the altimetric discontinuity is too low.

REFERENCES

Bredif, M., Boldo, D., Pierrot-Deseilligny, M. and Maitre, H., 2007. 3d building reconstruction with parametric roof superstructures. In: Proc. of the IEEE International Conference on Image Processing.

Brenner, C., 2000. Towards fully automatic generation of city models. In: In: IAPRS, pp. 85–92.

Deriche, R., 1987. Using canny's criteria to derive a recursively implemented optimal edge detector. International Journal of Computer Vision 1(2), pp. 167–187.

Dornaika, F. and Bredif, M., 2008. An efficient approach to building superstructure reconstruction using digital elevation maps. In: IAPRS, Volume 37 (Part 3A).

Durupt, M. and Taillandier, F., 2006. Automatic building reconstruction from a digital elevation model and cadastral data: an operational approach. In: Proc. of the ISPRS Commission III Symposium on Photogrammetric and Computer Vision, ISPRS, Bonn, Germany.

Iovan, C., Boldo, D. and Cord, M., 2007. Automatic extraction of urban vegetation structures from high resolution imagery and digital elevation model. In: URBAN, GRSS/ISPRS Joint Workshop on Data Fusion and Remote Sensing over Urban Areas.

Jibrini, H., Pierrot-Deseilligny, M., Paparoditis, N. and Maitre, H., 2000. Automatic building reconstruction from very high resolution aerial stereopairs using cadastral ground plans. In: Proc. of the XIXth ISPRS Congress, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, ISPRS, Amsterdam, The Netherlands.

Kolmogorov, V. and Zabih, R., 2004. What energy functions can be minimized via graph cuts? IEEE Transactions on Pattern Analysis and Machine Intelligence.

Mumford, D. and Shah, J., 1989. Optimal approximations by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics 17(4), pp. 577–685.

Pasko, M. and Gruber, M., 1996. Fusion of 2d gis data and aerial images for 3d building reconstruction. In: Int. Archives of Photogrammetry and Remote Sensing, Vol. XXXI, Part B 3, Vienna.

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: Proc. of the ISPRS Conference Topographic Mapping From Space (With Special Emphasis on Small Satellites), ISPRS, Ankara, Turkey.

Roux, M. and Maitre, H., 1997. Three-dimensional description of dense urban areas using maps and aerial images. In: Extraction of Man-Made Objects from Aerial and Space Images, II, pp. 311–322.

Roy, S. and Cox, I., 1998. A maximum-flow formulation of the n-camera stereo correspondence problem. In: Proc. of the IEEE International Conference on Computer Vision, Bombay, India, pp. 492–499.

Suveg, I. and Vosselman, G., 2001. 3d building reconstruction by map based generation and evaluation of hypotheses. In: Proceedings of the British Machine Vision Conference, p. 643652.

Taillandier, F., 2005. Automatic building reconstruction from cadastral maps and aerial images. In: U. Stilla, F. Rottensteiner and S. Hinz (eds), Proc. of the ISPRS Workshop CMRT 2005: Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation, Vienna, Austria, pp. 105–110.

Vosselman, G. and Dijkman, S., 2001. 3d building reconstruction from point cloud and ground plans. In: Proc. of the ISPRS Workshop on land surface mapping and characterization using laser altimetry, International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV, Annapolis, U.S., pp. 37–43.

Vosselman, G. and Suveg, I., 2001. Map based building reconstruction from laser data and images. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), pp. 231–239.

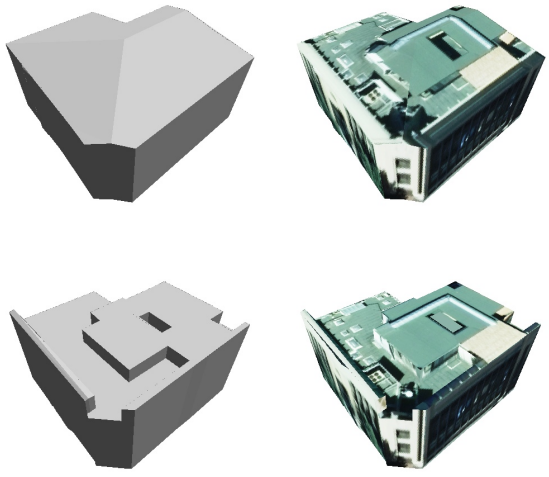


Figure 7: Untextured and textured 3D reconstruction results on the example of Figures 4(c) and 5(c). Top: reconstruction without enhancement, Bottom: with enhancement