

# Hybrid Digital Elevation Model production guided by 3D-primitives: A global optimization algorithm using graph cuts

Nesrine Chehata

[chehata@math-info.univ-paris5.fr](mailto:chehata@math-info.univ-paris5.fr)

[nesrine.chehata@ign.fr](mailto:nesrine.chehata@ign.fr)

**Abstract :** This research work is a part of a global project related to urban scene modeling from high resolution satellite images with focus on building production. This work has been carried out as a collaboration among the IGN<sup>1</sup>, the CNES<sup>2</sup> and the University of Paris 5.

The input data consist of a panchromatic stereo pair of satellite images, with a submetric resolution of 45-70 cm and a low Base to Height ratio B/H [0.05 - 0.2]. Data are provided by CNES<sup>1</sup>.

Since a detailed extraction and description of building rooves is complex in a satellital context, we propose to describe the scene by means of a 3D-surface which provides either raster or vector information using different level description. From extracted primitives and raw correlation information, an hybrid 3D surface will be obtained by an energy minimization process via graph cuts.

The main contribution of our approach is the use of 3D-primitives such as 3D-segments and 3D-facets as well as the introduction of information from external database to guide the optimization process. The obtained product is an hybrid DEM (Digital Elevation Model) which provides the highest level of reliable primitives for each scene region.

**Keywords:** Image Processing, HR satellite images, stereo, 3D Building surface reconstruction, 3D-primitives, global optimization, graph cuts, hybrid DEM.

## 1. Introduction

This paper is focused onto 3D-building reconstruction from a high resolution satellital stereopair. For this research project, we assume that both the geometry and the camera calibration are already determined. The correspondance between epipolar lines in the two images is supposed to be known. Since a detailed extraction and a description of building rooves are complex in satellital context, we propose to describe the scene by means of a 3D surface using different description levels (3D-points, 3D-segments, 3D-facets).

## 2. Previous work

Stereo algorithms that produce dense disparity map can be interpreted as global optimization problem which minimizes an energy function that is a combination of a "data term" and a "smoothness" term. In this paper, we will focus on algorithms based on graph cuts.

Over the last few years, several algorithms based on graph cuts were presented [4,2] to minimize energy functions. However their energy is not discontinuity preserving. Boykov et al. [1] have presented approximate algorithm with a guarantee on bounds for discontinuity preserving energy functions. This has been generalized by Kolmogorov et al. [3] to enforce the uniqueness.

However, all these algorithms propose optimization on pixellar information. In this paper, the novelty consists in using 3D primitives as constraints to preserve discontinuities and ensure 3D-surface smoothness.

## 3. Global optimization

In this paper, we propose to use the maximum flow algorithm proposed by Roy et al. [4] since it allows the use of 3D-primitives in a simple way. In our previous work [5], a correlation volume was used to qualify 3D-facets. It will be used to construct our 3D graph.

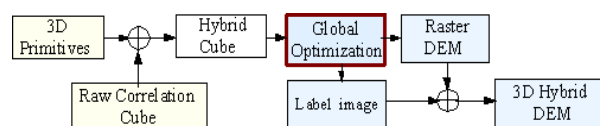


Figure 1: Global optimization pattern

A discretization step will project 3D-primitives into the cube. Since then, 3D-graph is embedded to the hybrid cube, which provides a disparity map of the scene, by means of minimal cut. DEM is obtained in raster format and the generated image of labels contains the primitives wherein the cut went through. A raster DEM combined to a label image

<sup>1</sup> IGN: French National Geographical Institute

<sup>2</sup> CNES : The French National Space Center

allow us to get a 3D representation of an hybrid DEM with different description levels.

### 3.1. 3D-graph construction

The 3D-graph is constructed using the precalculated correlation volume. Graph nodes correspond to cube voxels. In the object space, each voxel  $V$  is represented by 3D-coordinates  $(x,y,z)$  and its correlation score  $Corr(V)$ . Adding two terminals  $p$  and  $t$ , leads to a 3D pondered graph  $G = \langle v, \mathcal{E} \rangle$ ,  $v$  is the set

of nodes,  $\mathcal{E}$  the set of non-oriented edges  $v^2$ , with two terminal nodes  $s$  et  $t$ , the source and the sink.

The source  $s$  is connected to minimal altitude voxels and  $t$  to maximal altitude voxels as shown on figure 2-a.

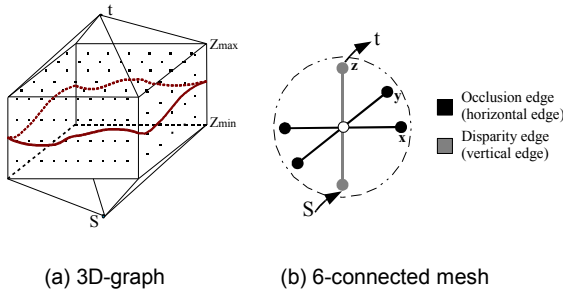


Figure 2 : 3D-graph construction pattern

The mesh is 6-connected (cf. figure 2-b). A cut  $C_G \in \mathcal{E}$  is a set of edges such that the terminals are separated into two vertices  $v^s, v^t$ . The cost of  $C_G$ , denoted  $|C_G|$ , is equal to the sum of its edge weights. The minimum cut graph problem is to find the cut with the minimal cost.

$G$  is a non-oriented graph and all edges capacities are positive.

In order to control the smoothness of the recovered surface, two types of edges are distinguished. Vertical edges along  $z$  axis, which determine disparity value, have  $C_{disp}$  capacity. All the other edges have occlusion capacity,  $C_{occ}$  which controls the surface smoothness .

$$C(u, v) = \begin{cases} 0 & \text{si } (u, v) \notin G \\ \infty & \text{si } u = s \text{ ou } v = t \\ C_{disp} & \text{si } (u - v) = (0, 0, \Delta z) \\ C_{occ} & \text{si } (u - v) = (\Delta x, \Delta y, 0) \end{cases} \quad (1)$$

Edge capacities are related to correlation scores. A higher correlation score increases voxel reliability and minimizes the corresponding edge capacity. Let  $u$  be a node in  $G$ , the node cost is expressed as following:

$$Cost(u) = 100 * (1 - Corr(u)) \quad (2)$$

$Corr(u)$  is a centred normalized correlation coefficient of  $u$ .

$$\begin{cases} C_{disp}(u, v) = \frac{Cost(u) + Cost(v)}{2} \\ C_{occ}(u, v) = K C_{disp}(u, v) + Cste \end{cases} \quad (3)$$

- $u$  and  $v$  two adjacent nodes in  $G$ .
- $C_{disp}(u, v)$  is vertical edge capacity (disparity capacity)
- $C_{occ}(u, v)$  is horizontal edge capacity (occlusion capacity) .
- $K$  is a smoothness parameter.

A higher  $K$  increases the smoothness of recovered surfaces while inversely, a lower  $K$  facilitate depth discontinuities. The constant  $Cste$  penalizes all discontinuities in the same way, independently from correspondent edge capacities. This parameter is useful for taking into account 3D-segment discontinuities as described in section 3.2.2.

### 3.2. 3D-primitives constraints

The novelty of the proposed approach is the use of 3D extracted primitives from images to constraint the global optimization.

3D-facets are reliable primitives since their size is larger than correlation windows so resulting correlation scores should be more reliable. On the other hand, with a low stereopair Base to Height ratio  $B/H$ , 3D-planes are more reliable than segments since robust estimation is performed over more 3D-points. Moreover, 3D-facets correspond to building rooves and so provide two important informations: roof smoothness and delimitation.

#### 3.2.1. 3D-facets

Injecting 3D-facet into the correlation volume is shown in figure 3 by means of a 2D view graph.

In this paper, we assume that for any planimetric region, there is only one corresponding 3D-facet. Facet overlapping will not be treated at this stage. The graph is initialized by a raw correlation score volume. In order to keep a simple graph structure, which allows the use of any type of external data, 3D-primitives will be discretized and projected into the 3D volume.

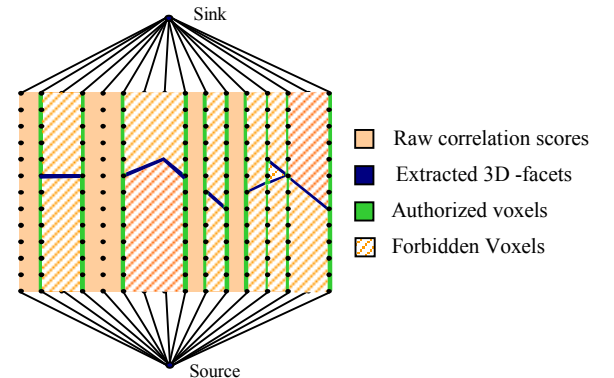


Figure 3 : 3D-facet constraints

3D-facet is discretized on a 6-connected neighborhood. Corresponding voxel cost is the average correlation score of the facet, obtained by the precalculated correlation volume.

The objective is to force the maximum flow to go through 3D-facets. Voxels that are located above and under these facets are forbidden. The cost of forbidden edges (vertical and horizontal) is maximal  $C_{Vf}=C_{Hf}=100$ . Paths are created on facet borders (Eq 4) in order to allow switches between facets.

$$C_F = \bigcup_i^n \bigcup_{k=z_{\min}^i}^{z_{\max}^i} \{(x_i, y_i, k) / (x_i, y_i) \in \oplus Cont(F)\} \quad (4)$$

where  $n$  is pixels number on dilated  $F$  contour and  $z_{\min}^i, z_{\max}^i$  are the minimum and maximum altitudes of pixel  $(x_i, y_i)$ .

Facet contours are dilated in order to create horizontal edges with minimal capacity which will make the flow go through these paths. Edge capacities in these paths are minimal but not null in order to take into account discontinuity magnitudes

If a 3D facet is under segmented or even missing, global optimization will be done over raw correlation cube voxels with respect to surface smoothness imposed by 3D facets. This ensures to obtain a final dense DEM.

### 3.2.2. 3D-segments

In a satellital context, due to low stereopair  $B/H$ , 3D extracted primitives have a good planimetric precision but suffer from poor altimetric reconstruction precision specially for linear primitives such as segments. As for 3D-segments, only planimetric coordinates will be used, by creating paths (Eq. 5) along these segments which makes the flow go through discontinuities. In fact, 3D-segments correspond either to facades or to roof intersections. Corresponding path is defined as following:

$$C_S = \bigcup_i^n \bigcup_{k=z_{\min}^i}^{z_{\max}^i} \{(x_i, y_i, k) / (x_i, y_i) \in \oplus P_S\} \quad (5)$$

Where  $P_S$  is a segment  $s$  set of pixels. Like 3D-facets, segments are dilated in order to create horizontal edges with minimal capacity to make the flow go through corresponding paths.

#### Segment edge capacity

In order to facilitate discontinuity depth along segments,  $C_{disp}=0$  on corresponding path edges.

To use 3D-segments, the parameter  $K$  has to be different from 0.

In fact, if  $K=0$ , all discontinuities are similarly penalized ( $C_{occ}(u,v)=Cste$ ). In this case, segments are useless.

If  $K \neq 0$  and  $C_{disp}=0$ ,  $C_{occ}$  of segment paths is equal to the constant  $Cste$  and is inferior to  $C_{occ}$  else where. The minimal cut is likely to pass through segment paths (Fig. 4.b).

However, if  $K \neq 0$  and  $C_{disp} \neq 0$ , horizontal edge capacity is not constant anylonger and disparity

uniqueness for one pixel is no more guaranteed. In this case, the estimated disparity will be the larger one (Fig. 4.c).

### 3.2.3. Use of Ground

To overcome 3D-facet lack of precision or errors on roof delimitation, we propose to modelize locally the ground from a DEM. A basic classification is processed after the use of segments, since they regularize building borders. The ground will be projected onto the correlation volume like a minimal 3D-facet which covers all the scene. The 3D-graph source will be linked to the ground. As a consequence, the ground will be considered in concurrence to any 3D-facet and delimitation errors will be corrected. Ground can also be obtained from external database like road database or cadastral maps.

## 4. Results

Results are presented on aerial images of the city of Amiens, France, with a 45 cm-resolution,  $B/H=0.2$ . Figure 4-a presents examples of extracted primitives [chehata02, chehata03], 3D-segments and 3D-facets. Facets are obtained from label image after the global optimization process (figure 1). It means that the minimal cut went through these facets.

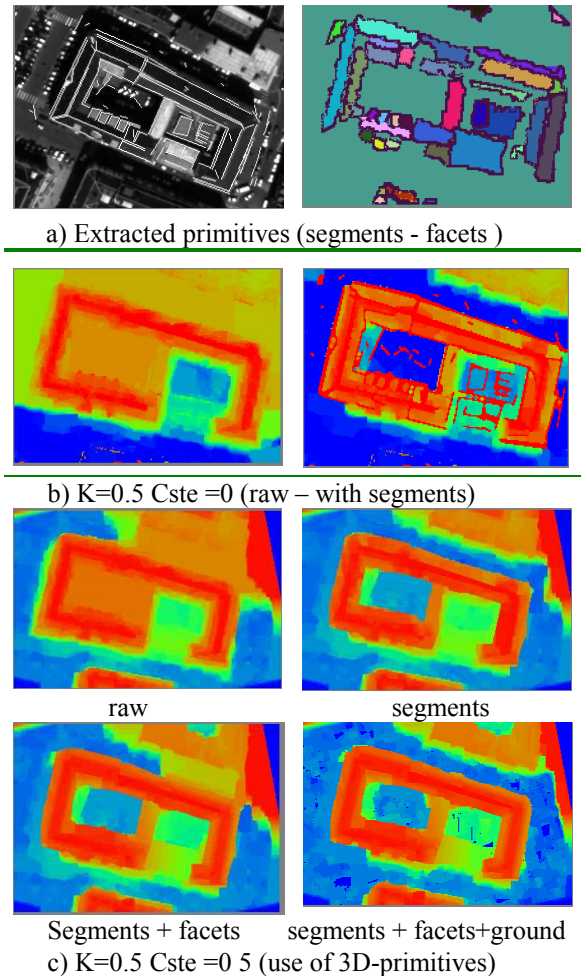


Figure 4: Global optimization results

Some results of global optimization are shown with different parameters:

- ( $k=0.5$ ,  $C_{ste}=0$ ), horizontal edge capacity for segments is null since  $C_{disp}=0$ . These capacities ensure the cut go through discontinuities. However, since the capacity along segment discontinuities is null, their magnitude is not taken into account and the estimated altitude is just the maximum altitude (figure 4-b).
- ( $k=0.5$ ,  $C_{ste}=0.5$ ), horizontal edge capacity for segments is not null, altitude along segments is better estimated. By using 3D-facets, the final DEM is smoother on roof surface and building borders are more regular depending on 3D-facet delimitation quality. However, some delimitation problems appear due to false 3D-facets. This problem can be overcome thanks to ground modeling.

## 5. Raster evaluations

Evaluations were processed on raster DEM comparing to a reference one. Statistics are focused on altitude differences between DEMs. In order to separate planimetric and altimetric errors, evaluations are processed by class. Three different classes were examined: "Building", "Building Border" and "Ground". Ten different parametrizations were used for global optimization. As shown on figure 5, average error on building altitude is about 20 cm. It is about 110 cm for building border. The last class is difficult to evaluate since it combines both planimetric and altimetric errors.

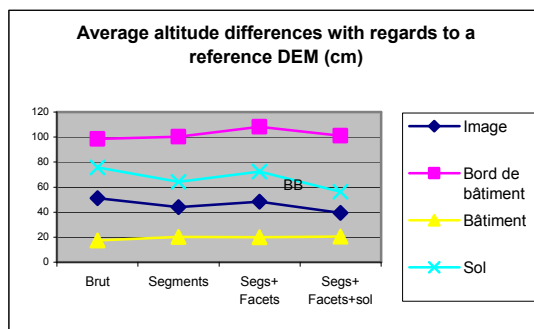


Figure 5: Average altitude differences with regard to a reference DEM

The use of 3D-segments improves results on "Building Border" and "Ground" classes since it facilitates depth discontinuities.

However, raster evaluations do not quantify neither 3D-surface regularity nor roof delimitations. The results on average altitude differences for buildings are not improved since the previous optimization without primitives is robust. To show the apport of 3D-facets, high-level evaluations need to be processed to quantify delimitation, surface regularity, facet slopes.

Finally, we can see that errors induced by 3D-facets are rectified by the modelization of ground.

## 6. Conclusion

In this paper, we presented an algorithm for 3D-surface modelling, formulated as an energy minimization problem constrained by 3D-primitives and solved via graph cuts. This method takes into account complementary primitives such as 3D-segments for discontinuities and 3D-facets for surface smoothness.

This method advantage enables the integration of any kind of primitives or external database. Moreover, this algorithm is independant from any primitive extraction processus and can be applied to any kind of aerial and HR satellital images. This algorithm can also be extended to a multiple view context since the optimization is processed into a correlation volume in the object space. However, it depends on 3D primitives quality and reliability. This point is not critical for segments since only planimetric coordinates are taken into account. It is more complex for facets since surface smoothness and roof delimitations depend on them.

## 7. Future work

This algorithm can also confront concurrent facets, obtained from different extraction process. Capacities of facet edges have to be changed to avoid too many switches between facets. High-level evaluation of facets is under progress.

## 8. Bibliography

- [1] Y.Boykov, O.Veksler and R.Zabih, "Fast Approximate Energy Minimization via Graph Cuts", ICCV, vol I, pp 377-384, 1999
- [2] H. Ishikawa and D. Geiger, "Occlusions, discontinuities, and epipolar lines in stereo", ECCV, pp 232-248, 1998.
- [3] V.Kolmogorov and R. Zabih, "Computing Visual Correspondence with Occlusions via Graph Cuts", ICCV, pp 508-515, 2001.
- [4] S. Roy and I. Cox, "A maximum-flow formulation of the n-camera stereo correspondance problem", ICCV, 1998.

## 9. Publications

### International conference

- [5] N.Chehata, M-P.Deseilligny, F.Jung, G.Stamon. "Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings", MVA2002, pp 636-639, Dec 2002, Japon.
- [6] N. Chehata, F. Jung, MP Deseilligny and G. Stamon. "A region-based matching approach for 3D-Roof reconstruction from high resolution satellite images", DICTA2003, Dec2003, Sydney, Australia.

### Journals

- [7] N. Chehata, M. Pierrot-Deseilligny and G. Stamon. "Génération de MNE hybrides (raster/ vecteur) : optimisation à base de flots de graphes contrainte par des primitive », revue de la SFPT, n°176(2004-4), pp 30-41, 2004.