

A Region-Based Matching Approach for 3D-Roof Reconstruction from HR Satellite Stereo Pairs

Nesrine Chehata^{1,2} Franck Jung¹ Marc-Pierrot Deseilligny¹
Georges Stamon²

¹ Institut Gographique National. 2/4 avenue Pasteur. 94165 Saint-Mand Cdex
E-mail:nesrine.chehata@ign.fr, franck.jung@ign.fr, marc.pierrot-deseilligny@ign.fr

² Universit Paris 5 - CRIP5/SIP. 45 rue des saints pres - 75006 - Paris
E-mail:Stamon@math-info.univ-paris5.fr

Abstract. This study is a part of a global project on urban scenes interpretation using high resolution satellite images. Actually, the research is focused on buildings and roads are used to delineate zones of interest. This study details automatic reconstruction of 3D facets using a region matching approach based on hierarchical segmentation of images. The novelty consists in dedicating algorithms to satellital context to make them more robust to noise and to deal with low stereopair Base to Height ratio. First of all, a hierarchical segmentation process is explained, then matching regions constraints are detailed. Afterwards, optimal cuts, which corresponds to a set of regions that are likely to represent building rooftops, are processed in both hierarchies. Cuts are processed given matching scores and using regions planarity constraint. In the third part, global matching of both cuts is processed in order to obtain final matchings which will allow 3D scene reconstruction. Eventually, some results are presented.

Keywords high resolution satellite images, automation, stereoscopic pair, hierarchical segmentation, region matching, automatic 3D facets reconstruction, planar regions.

Context

These last years, the launch of high resolution commercial satellites [Ikonos, Eros, QuickBird] provides higher scale satellite images whose resolution ranges in [60cm-1m]. Consequently, 3D building reconstruction from satellite images is receiving more and more attention from the scientific community. In our case, algorithms are planned to be applied on the forthcoming satellite data of PLEIADE HR which will succeed SPOT5. The input data consists of a panchromatic stereo pair of satellite images, at a high resolution of 50-60 cm and a low Base to Height ratio B/H [0.05 .. 0.2]. In this context, the use of different primitives (segments, corners, planar facets) is a key step to generate building hypothesis. Some authors [9] have shown the importance of planar primitives to describe urban scenes. So, this paper will focus on 3D-facets reconstruction.

1 Related work

Extracting planar facet from stereo pairs has been treated by different authors. Many approaches are based on region segmentation [10, 7]. Methods are based on three steps; region segmentation, matching and reconstruction. Segmentation can be processed either on DEM (Digital Elevation Model)[9] or on images. In [4], 3D-Hough transformation is used to generate planes hypothesis. In [2], half planes attached to 3D-segments are extracted. This approach is not adapted to satellital context since 3D-segments suffer from poor altimetric accuracy which is spread at every step of the process. Algorithms have to be adapted to high resolution satellital context with low B/H stereo pairs.

2 Methodology

For satellital context, we chose a region based approach. Moreover, the low B/H of the stereopair ensures few geometric distortions between images and also few hidden parts which will make matching process easier. In addition, segmentation is computed on stereo images. In fact, for DEMs, obtained by correlation techniques, [1] on built-up areas, a compromise has to be found between geometric accuracy and robustness. In our case, DEM suffers from poor accuracy so it is not suitable for segmentation purposes, however a tolerance band around DEM can be used to validate matching and reconstruction. Finally, since there are several meaningful scales to describe an image(pixel, roof, building, town), we use a multi-scale segmentation to partition image into meaningful 2D-regions. It also allows region matching across levels. Figure 1 illustrates the global scheme. It is based on three steps; image segmentation, region matching and finally 3D-facet reconstruction

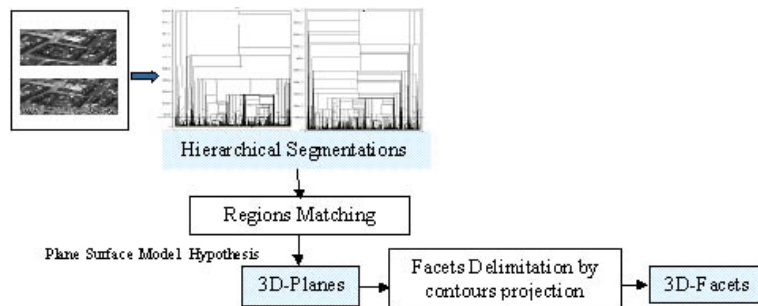


Fig. 1. Global Scheme of 3D facets reconstruction

In the following, we will recall the global scheme of hierarchical segmentation which is directly derived from work related in [3]. In section 4, region matching and different constraints will be detailed. Section 5 shows how to obtain optimal

segmentations by hierarchy cuts using matching scores. Eventually, global hierarchy matching and 3D facet reconstruction are detailed. Results are presented on simulated images obtained from aerial ones and adapted to satellital context.

3 Multi-scale Segmentation and Hierarchy

Our approach is based on hierarchical segmentation which is directly derived from [3]. Nested Segmentations $\Omega_0, \Omega_1, \dots, \Omega_n$ are produced proceeding bottom-up (fine to coarse) from an initial level by merging regions satisfying geometric and radiometric constraints. Base regions are obtained by watershed process.

Merge conditions To partition image into planar regions which correspond to building rooftops, we use a classical formulation of segmentation where two energies are confronted: geometric and radiometric ones. The objective consists in finding the partition which minimizes the global energy for a given λ .

$$E_\lambda(P) = \sum_{R \in P} E_\lambda(R) = \sum_{R \in P} \lambda E_{\text{Geo}}(R) + E_{\text{Rad}}(R) \quad (1)$$

λ is a regularization term : for a high λ , geometric energy will be predominant leading to a simple model (few regions with simple forms), in the contrary, with a weak λ , radiometric energy is more influential which leads to small homogenous regions satisfying radiometric model. In practice, geometric model has a linear cost proportional to image gradient. For radiometric energy, a gaussian MDL coding is used [5].

Practically, image hierarchy is processed at different scales by increasing λ and merging adjacent regions whose fusion decreases the global energy. Thus, each node of the hierarchy is indexed by its apparition scale and corresponds to a region defined by a set of connected base nodes.

4 Region based stereo matching

The matching process associates regions in the first hierarchy with regions in the second one which are likely to correspond to the same physical object. Matching may occur across levels of segmentation. The algorithm begins by matching independently each region of both hierarchies. Then, using matching scores and planarity constraint, cuts are processed in hierarchies in order to obtain planar regions. Eventually, both cuts will be matched globally, ensuring one-to-one matching that provides the best matching cost. In the following, Ω_l (resp Ω_r) represent the set of regions in left (resp right) image. The homologous regions set of a region R will be denoted $H(R)$.

4.1 Single region matching

To begin, each region of one hierarchy is matched independently with the second hierarchy. This process is computed on both hierarchies in a symmetric way. For the matching process, several constraints are used to reduce the number of homologous regions; size constraint, epipolar and altimetric constraints, similarity measure, and overlap constraint.

Size constraint This constraint is used at every step of the algorithm to avoid very small regions or big ones. Minimum and maximum size thresholds T_{min}^s, T_{max}^s are chosen depending on desired interpretation level (building, roof, chimney).

Epipolar constraint Epipolar constraint is used to ensure that homologous regions of a region R^l have their centres of gravity within an epipolar rectangle. Epipolar band is obtained by a fixed tolerance to R^l centre of gravity's corresponding epipolar line and the maximum and minimum average height of R^l (cf. Fig 2(a)). Heights are provided by a DEM [1] to reduce the search space. For a region R , the homologous regions set under epipolar constraint is denoted $HEpi(R)$.

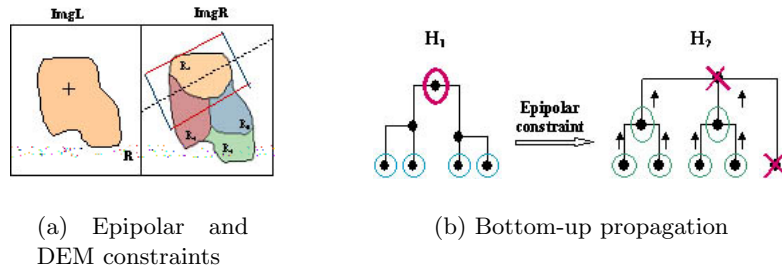


Fig. 2. Epipolar constraints

The novelty is the bottom-up propagation of constraint (cf. Fig 2(b)). First, it is processed on elementary nodes, which are marked then constraint is spread to higher levels. "Mother" regions are marked as homologous regions if all their "children" regions are marked. This method makes the constraint robust to noise, over-segmentation and ensures not missing regions.

Similarity constraint At this stage, each region R has a homologous regions set $HEpi(R)$. To reduce the combinatory, we compute a measure of similarity [8]. The homologous regions set under similarity constraint will be denoted $HSim(R)$. The similarity measure is defined as:

$$Sim(R_l, R_r) = \sum_{p=1}^q w_p s_p(R_l, R_r) \quad (2)$$

s_p is the similarity score for each parameter or region attribute and it's weighted by w_p . In practice, $w_p = 1$.

$$s_p(R_l, R_r) = 1 - \frac{\min(A_p(R_l), A_p(R_r))}{\max(A_p(R_l), A_p(R_r))} \quad (3)$$

A_p is some attribute of the region. In our case, three attributes are used : surface, bounding box, spatial moment.

Overlap constraint For a couple of matched regions, the corresponding 3D-plane $\Pi_{(R_l, R_r)}$ is estimated. Reconstruction will be detailed in section 5. The homologous regions set under overlap constraint will be denoted $HLap(R)$.

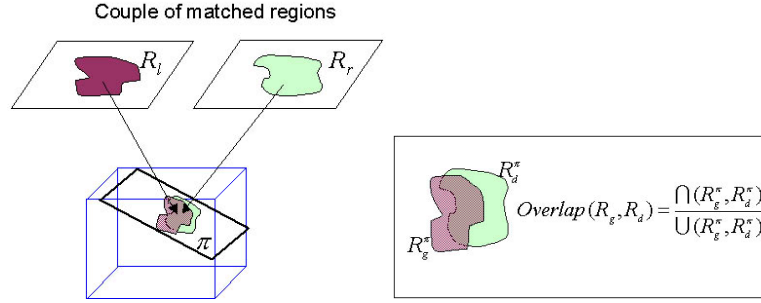


Fig. 3. Overlap constraint

Both regions are projected onto plane(cf.Fig 3). The overlap score is the ratio of rectified regions intersection by their union.

Matching relation Matching relation \mathcal{H} is defined as follows:

$$\forall R_l \in \Omega_l, \forall R_r \in \Omega_r \quad R_l \mathcal{H} R_r \iff \begin{cases} 1. T_s^{min} \leq R_l \leq T_s^{max} \text{ et } T_s^{min} \leq R_r \leq T_s^{max} \\ 2. R_r \in HEpi(R_l) \\ 3. R_r \in HSim(R_l) \iff Sim(R_l, R_r) \geq T_{sim} \\ 4. R_r \in HLap(R_l) \iff Overlap(R_l, R_r) \geq T_r \end{cases} \quad (4)$$

Hence, homologous regions have to verify size, epipolar, similarity and overlap constraints.

$T_{sim} = 0.3$ is chosen as a large threshold as it is a critical parameter, hard to evaluate since it depends on segmentation. $T_r = 60\%$ is easier to interpret and do not depend on segmentations quality.

5 3D-plane reconstruction from a couple of matched regions

3D-plane reconstruction is a key stage in our approach. It is necessary to process overlap constraint, to qualify the matching score (cf.section 6.2) and to reconstruct the 3D scene. Reconstruction can be geometric or photometric, using radiometric attributes [10]. In our case, plane reconstruction is based on region contours matching, which depends of segmentation quality and thus is more discriminative and adapted to qualify matching scores.

Region contours matching Contours matching is processed in epipolar geometry with subpixelar precision to ensure good 3D altimetric precision. The potential homologous point is re-sampled as many times as desired precision. Re-sampling is processed thanks to a bicubic interpolator. Thus, to one homolog pixel corresponds a correlation curve and its maximum returns a subpixelar position of the corresponding pixel.

Robust 3D-plane reconstruction Given a couple of matched regions, the 3D corresponding plane is defined by : $z = \Pi(x, y) = ax + by + d$. To deal with outliers, parameters a,b,d are estimated using a robust M-estimator L_p . In practice, $p=1.2$ which represents a good compromise between complexity and stability [11]. Although less robust than some estimators(RANSAC, Least Median of squares), these estimators behave well with good initialization. In our case, initialization depends on matched regions, so M-estimators behaviour will qualify the matching process.

6 Optimal segmentation in a hierarchy

Hierarchy cut is processed in a cooperative way with matching qualification. It is based on a top-down approach since matching is more reliable on larger regions. Each match is qualified by regions correlation score. A correlation volume in object space is used. The volume calculation and matching qualification are detailed below. Given matching qualification, we search for an optimal cut in the hierarchy which contains the maximum of planar regions which correspond to building rooftops.

6.1 Correlation volume

Correlation volume is processed in object space, using geographic coordinates [6]. Planimetric intervals correspond to visible scene from both stereopair points of view. Altitude interval corresponds to minimum and maximum elevation on DEM. A fine altimetric discretization ensures subpixelar precision. Each voxel of the volume and his horizontal neighbours (within a 5*5 window) are projected onto images (cf.Fig 4)and then correlated using central normalized correlation score.

6.2 Matching score

A couple of matched regions is qualified by the average of correlation score on the corresponding 3D facet. In fact, once the 3D-plane $\Pi(R_l, R_r)$ is reconstructed, both regions R_l and R_r are projected on the plane(cf.fig 4). Their intersection delimits the 3D-facet $\mathcal{F}_{(R_l, R_r)}$ denoted \mathcal{F} . The 3D-facet is valid if it belongs to a tolerance band derived from DEM. Each voxel of the facet has its correlation score precalculated in the volume. The average of 3D-facet voxels scores qualifies the matching of the corresponding couple of regions.

$$Q(R_l, R_r) = \frac{\sum_{\mathcal{V} \in \mathcal{F}} ScoreCorrel(\mathcal{V})}{Card(\mathcal{F})} \quad (5)$$

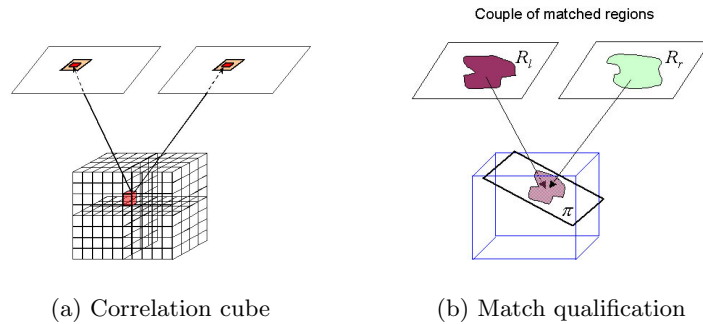


Fig. 4. Correlation cube

The use of correlation volume reduces computational time since correlation scores are processed once a time. This approach is adapted to a low B/H since there are few distortions between images, so horizontal vignettes are sufficient to qualify matches and there is no need to rectify regions and correlate them in object space.

6.3 Optimal cut process

The optimal cut process is detailed in the following section for one hierarchy. In practice, the process is iterated on both hierarchies in a symmetric way in order to obtain more robust matching. Each region R has a score $Q(R)$ to qualify all its matches. $Q(R)$ is defined as the maximum of correlation scores with homologous regions .

$$Q(R) = \text{Max}(Q(R, R'_i)) \quad R'_i \in H(R) ; i = 0, \dots, \text{Card}(H(R)) \quad (6)$$

Optimal cut is processed in a top-down way. It stops at a given level if following conditions are satisfied.

- *Condition 1.* The region matching score is higher than its children ones.

$$Q(R) \geq Q(R_c) \quad \forall R_c \in \text{children}(R) \quad (7)$$

- *Condition 2.* 3D-planes corresponding to children has to be quasi-coplanar.

The matching score threshold is a critical parameter to choose. For this reason, hierarchy cut uses relative criteria comparing region score to its children scores. If the "mother" region has a higher score than its children ones and if children corresponding planes are quasi-coplanar which means that children represent an over-segmentation of the region, cut has to be done at the "mother" level. 3D-Planes are compared using angular deviation between their normals.

The optimal cut quality depends on 3D-planes estimation. A bad estimation may lead to an "over-segmented" cut since planes angular deviation is high and cut goes on to lower levels.

7 Global hierarchy matching

Once cuts are obtained in both hierarchies, all cut regions have homologous regions so matches have to be compared to keep best ones. Thanks to a low B/H, occlusions and hidden parts are reduced and segmentations are almost similar so global matching behaves well. First of all, correspondences are merged between both cuts. A consistent match set is therefore determined by "a winner takes all" scheme which gives a decreasing importance respectively to score matching, overlap score and finally to similarity score.

8 Results

Due to the lack of available data, results are presented on simulated PLEIADE images or on aerial images satisfying satellital context. In order to focus only on buildings, a preliminary basic classification was processed to eliminate ground regions. Numerical results were calculated on 4 stereopairs on French cities with different B/H and resolutions.

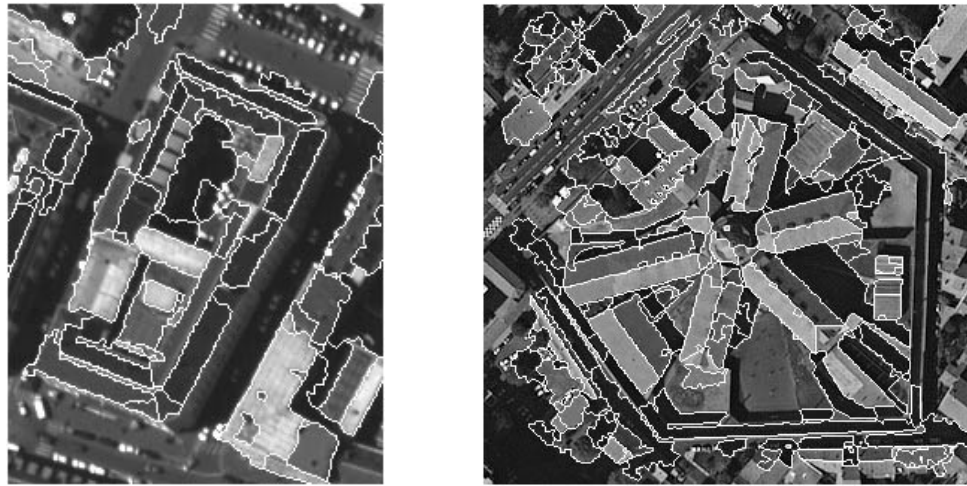
Matching evaluation Once all constraints are applied, only 23% of matched regions are kept. The mean number of homologous regions for kept regions, before cut process, is about 2.

Optimal cut evaluation Figure 5 illustrates optimal cuts on images with different resolutions and B/H. First results are satisfying, the scene is well segmented, major rooftops were found and under-segmented regions were avoided. Over-segmentations are due to initial differences between hierarchies. If matched regions are different, correlation score is low and hierarchy cuts goes on to lower levels. Over-segmented regions can be merged by post-treatments.

Optimal cut enables to extract planar regions in images which have the best matching score. Table 1.a shows that on regions belonging to optimal cut, 38% are reliable, i.e. have only one homologous region, 62% have many homologous regions. For the last case, in 97% of cases, homologous regions are nested in a descendancy relation. Table 1.b shows that 75% of extracted regions correspond to building rooftops. Missing rooftops are evaluated to 14%. To overcome this problem, matching can be spread using adjacency constraint. Missing rooftops can also be completed in the final step of 3D model reconstruction, using hypothesis like symmetric roofs for example.

Cut regions/matched regions 28%		Valid planar regions	
"Reliable" regions	Regions with many		wrong regions
1 homologous region	homologous regions	75%	25%
38%	62%		

Table 1. Optimal cut hierarchy evaluation



(a) Amiens - 45cm - B/H 0.2

(b) Toulouse - 60cm - B/H 0.08

Fig. 5. Optimal cut

Each region in the final optimal cut has only one homologous region. Afterwards, 3D-planes are processed from couples of regions. Eventually, 3D-Facets are obtained by projecting region contours on corresponding 3D-planes. An evaluation of 3D planes reconstruction depending on image resolution and B/H ratio is under process.

9 Conclusion

In this paper, a new approach for 3D-facet reconstruction in satellite context is presented. It is based on three stages which are adapted to a low B/H and handles both views in a symmetric way. First of all, a multi-scale hierarchy is used to segment images. Merging criteria are adapted to find planar regions and by means of cut in hierarchy, interpretation level of the scene can be chosen. Afterwards, matching region problem is treated. The novelty in our approach consists in constraint propagation bottom-up in hierarchy which makes algorithms more robust to noise and images over-segmentation, besides reducing computational time. Moreover matching process is adapted to satellital context by the use of an initial correlation volume with subpixelllar precision and subpixelllar contours matching for 3D-plane reconstruction. Finally, a cooperative way is used to process optimal segmentation using matching scores and planarity constraints.

10 Future Work

First results are satisfying. Improvements are under process. We aim to implement a robust contour matching of homologous regions using subpixel accuracy to overcome problems due to a low ratio B/H . In a second part, in order to make hierarchy cut more robust, we propose to validate the cut from fine to coarse. Criterias should be different from those used to process the cut.

Eventually, for the final 3D scene reconstruction, a refinement stage of 3D facets should be processed to overcome matching contours problems.

Acknowledgment

This work has been carried out as a collaboration among the IGN¹, the CNES² and the University of Paris 5 (Research lab SIP/CRIP5).

References

1. C. Baillard. *Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains*. PhD thesis, ENST, 1997.
2. C. Baillard and A. Zisserman. Automatic reconstruction of piecewise planar models from multiple views. *CVPR*, pages 559–565, June 1999.
3. L. Guigues, H. Le-Men, and J-P. Cocquerez. Scale-sets image analysis. *ICIP'03*, October 2003. Barcelone.
4. H. Jibrini, M-P. Deseilligny, N. Paparoditis, and H. Matre. Reconstruction de modles de btiments partir de donnes cadastrales vectorises et d'images ariennes. *RFIA2000*, III:299–306, 2000. Paris.
5. T. Kanungo, B. Dom, W. Niblack, and D. Steele J. Sheinvald. Mdl-based multi-band image segmentation using a fast region merging scheme. Research Report RJ 9960, IBM, May 1995.
6. N. Paparoditis, C. Thom, and H. Jibrini. Surface reconstruction of urban area from highly overlapping digital aerial images. *International Archives of Photogrammetry and Remote Sensing, XXXIII ISPRS*, Geo Information for all, July 2000. Amsterdam, Netherlands.
7. S. Randriamasy and A. Gagalowicz. Region-based top-down segmentation controlled by stereo matching. *IAPR:MVA1992*, 1992. Japan.
8. P. Sander, L. Vinet, L. Cohen, and A. Gagalowicz. Hierarchical region based stereo matching. In *Proceedings of the Sixth Scandivian Conference on Image Analysis*, pages 71–78, June 1989.
9. C. Vestri and F. Devernay. Using robust methods for automatic extraction of buildings. In *Proceedings of the International Conference on Computer Vision and Pattern Recongition, IEEE*, 2001. Hawaii.
10. J-M. Vzien and J-P. Tarel. A generic approach for planar patches stereo reconstruction. In *Proceedings of the Scandivian Conference on Image Analysis*, pages 1061–1070, 1995. Uppsala, Sweden.
11. Z.Zhang and O.Faugeras. *3D scene analysis :a stereo based approach*. Springer, 1992.

¹ Institut Gographique National,France.(National Geographical Institute)

² Centre National d'Etudes Spatiales,France.(The French National Space Center)