MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

15—6

Extraction of 3D primitives from stereopairs of satellite images for automatic reconstruction of buildings

Nesrine Chehata*† Marc Pierrot-Desseilligny* Franck Jung* Georges Stamon†

Abstract

This study is a preliminary operation of a global project which goal is to interpret urban scenes from satellite images in order to recognize and classify roads and buildings. Actually, the research is focused on buildings. We plan to apply these algorithms on the forthcoming satellite data of PLEIADE HR which will follow SPOT5. The input data consists of a panchromatic stereopair of satellite images, at a high resolution of 50-60 cm and a low Base to Height ratio B/H= 0.1,0.2.

In this paper, we describe three algorithms involved in a global strategy for extracting reliable 3D primitives which is a key step and an important preliminary operation to make the generation of building hypothesis easier. All algorithms use photometric constraints at every step to take advantage from low B/H and handle both views in a symmetric way. However, each algorithm can be used independently providing the specific primitives as input data.

Key words: HR satellite images, stereo, 3D building reconstruction, 3D-primitives, planar patches, roofs delimitation

1 Introduction

The aim of the project is to interpret urban scenes from satellite images in order to recognize and classify roads and buildings and extract 3D geographical objects. There is not such a system yet which enables to solve this problem automatically. At the moment, we assume that we have a reliable road network to divide the scene into houses blocks and we will focus only on buildings. In this study, the goal is to extract reliable 3D-primitives which are a key step in the automatic building reconstruction.The global scheme is illustrated Figure 1.

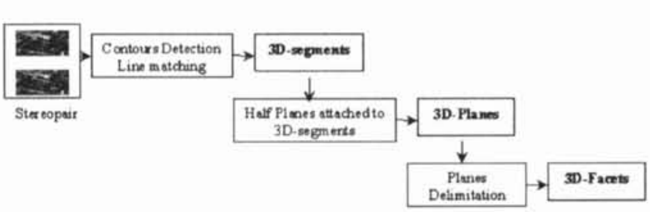

Figure 1: Global Scheme of 3D-primitives extraction

*Address: IGN/MATIS: Institut Gographique National, 2-4 Avenue pasteur-94165 Saint Mandé France. E-mail: name.surname@ign.fr

†Address: SIP/CRIP5 ; University of Paris 5. E-mail: Stamon@math-info.univ-paris5.fr

First of all, 3D-segments are extracted then a research is launched for half planes attached to the extracted segments. Only a single 3D-line with textured neighbourhood is required for a plane hypothesis. Finally, a new algorithm which delimits automatically roof edges is implemented. It will facilitate plane fusion and reduce the combinatory.

2 3D-segment extraction

The first algorithm consists in extracting 3D-segments.The 2D lines are obtained by applying the Canny edge detector [Deriche89](with subpixel accuracy)and a straight line estimation by orthogonal regression. Individual segment matching and reconstruction will be detailed. In order to reduce the combinatorial complexity(cf.Figure 2), besides the epipolar constraint, which is applied to segment extremities, we use a Digital Elevation Model DEM (obtained by a correlation process) [Baillard97] to reduce the search space provided by the minimum and maximum height available for each pixel x_i which corresponds to reciprocally to x_{imin} and x_{imax} in the second image.

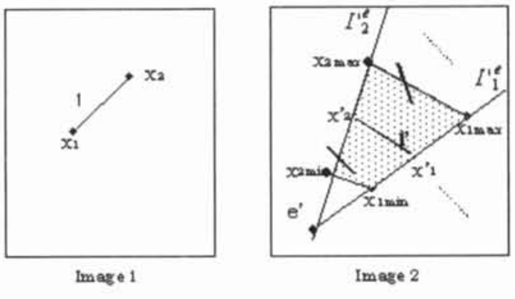

Figure 2: Epipolar and DEM Constraints

The geometric constraint is reduced to an overlap constraint. For each pair of lines which satisfies the geometric constraint above mentionned,we use the lines photometric neighbourhood for disambiguation. The basic idea is to consider each segment as a list of points. The epipolar geometry provides a point to point correspondence on matched lines. The similarity of neighbourhoods is then assessed by cross-correlation at the corresponding points. The correlation scores are reliable thanks to a low B/H. The photometric matching score is computed as the average of the individual correlation scores for the pixels of the line. Then the algorithm computes the final score taking into account geometric and photometric matching scores weighted reciprocally by 0.4 and 0.6 .A consistent match set is therefore determined by "a winner takes all " scheme.

In this technique, both images are handled in a symmetric way, and the set of possible correspondences is pruned under reciprocity constraint that keeps only reliable matches. Considering a couple of homologous segments, we project each of 2D-extremities onto the corresponding 3D-line then, we keep the two farthest ones which leads to the final segment reconstruction. In order to validate these 3D-segments, a search volume is then derived from the DEM dilatation. We can thus force the reconstructed segments to have their extremities in the search volume. As far as the 3D-segment extraction is concerned, the improvement consists in the use of photometric constraint and the use of a DEM during matching and reconstruction steps. Extracted lines consist one of 3D-primitives used in building reconstruction and will also guide 3D-plane extraction.

3 3D-plane extraction

The second algorithm is a method for automatically computing piecewise planar reconstruction based on extracted 3D segments. This method, essentially derived from [Baillard99], is adapted to our context. The method consists in computing reliable half planes defined by only one 3D-line and similarity scores of its neighbourhood computed over two views. This method was chosen since it includes both geometric and photometric constraints. Our objective is to determine, for each 3D-line, within one-parameter families $\pi(\theta)$ and $\pi(\theta')$ whether there are attached half planes or not. If there are, we want to compute the best estimate of θ and θ'. The two families represent half planes on each side of the vertical reference attached to the line. First, as illustrated in Figure 3, it is necessary to determine a textured point set in order to produce a discriminative similarity function. Thus, correlation will be computed only in the neighbourhood of textured Points Of Interest(POI)[Harris88]. Given θ, the plane π_θ defines a point to point map between both images. If the plane is correct, intensities at corresponding pixels will be highly correlated. The similarity score function Sim(θ) has to be selective, and robust to irrelevant points. It is defined as:

$$Sim(\theta) = \sum_{i=1,2} \iint_{POI} Cor^2(x, x') \tag{1}$$

and ranges between (0,1). The correlation term $Cor^2(x, x')$ is the centred normalized cross-correlation between x in the first view and its corresponding point x' in the second one. Cross correlation is used to be selective on θ over textured regions. The correlation is squared to give more weigth to high scores and therefore to be more selective.

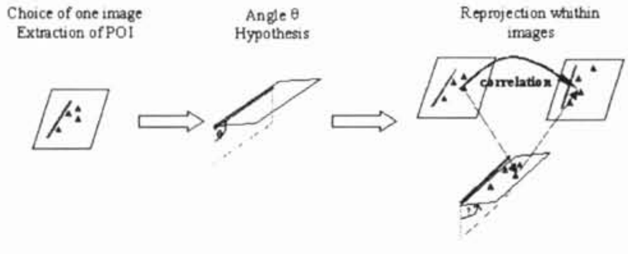

Figure 3: 3D-plane Extraction

We use adaptative windows that fits roof borders and extracted segments. In our case, POI are calculated on both images and are handled in a symmetric way so that the score correlation should be averaged over both views. This increases the reliability of the detected planes. We only keep scores higher than a fixed threshold ($S_{th} = 0.5$). Optimum angle θ is computed by searching the maximum of the similarity function Sim (θ) with θ included in [20°, 160°] with regards to a vertical reference. The process is iterated for θ' to extract the second half plane.

Afterwards, planar facets enable line grouping and the construction of parts of the wireframe (by plane intersections) which were missed due to the inevitable shortcomings of feature detection and matching.

4 3D-plane delimitation

Finally, a new method for delimiting 3D-planes is presented. It is based on relaxation method which is an iterative procedure for labelling objects based on their incomplete information. In this paper, we focus on labelling roof regions with a probabilistic model. Probabilistic relaxation was chosen as it ensures continuous and reliable surfaces based on 3D-planes [Alcharroun01]. This algorithm will facilitate plane fusion and reduce the combinatory. It can be used to delimit a priori the correlation surface when searching for half planes. Once the plane π_θ attached to a 3D-line computed, we will consider a rectangular surface whose height is the line length and a fixed width. The fixed width is over estimated. The initial surface might run over other roofs but the algorithm is robust in these cases. At the beginning, both rectangular surfaces in both views are projected onto the plane π_θ . As illustrated in Figure 4, the line is on the left side of the projected images. Then we calculate the sum of gradients in order to compute a watershed process and to generate a Region Adjacency Graph (RAG). Finally a probabilistic relaxation model is processed on the obtained RAG.

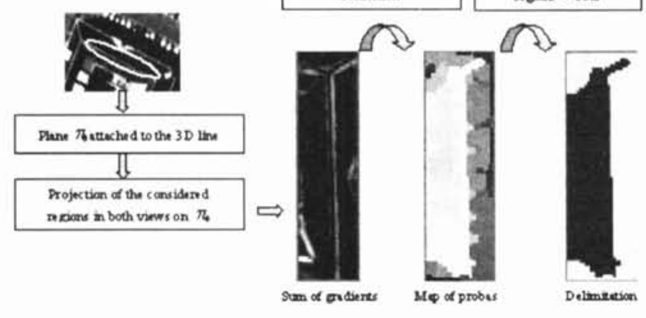

Figure 4: 3D-plane delimitation algorithm

The dynamics of the system depend on its governing equation. We use the Rosenfeld's updating rule [Fu&Yan97]. Let $p_i^{(k)}(\lambda), \lambda = 1, \ldots, m$, denote the probability of the i^{th} object having label λ at step k. The updating rule have this form:

$$p_i^{(k+1)}(\lambda) = p_i^{(k)}(\lambda).\frac{1 + s_i^{(k)}(\lambda)}{1 + s_i^{-k}} \tag{2}$$

where

$$s_i^{(k)} = \sum_{j=1}^{n} d_{ij} \sum_{\lambda'=1}^{m} c_{ij}(\lambda, \lambda') p_i^{(k)}(\lambda') \quad (3)$$

$$\overline{s}_i^{(k)} = \sum_{\lambda=1}^{m} p_i^{(k)}(\lambda) s_i^{(k)}(\lambda) \quad (4)$$

with

- A set of m labels: $\{\lambda_i, i = 1, \ldots, m\}$ for each object.
- $c_{ij}(\lambda, \lambda')$ represents the compatibility measure of object i with label λ when object j has label λ'. In our case, the compatibility measure is considered as a conditional probability which satisfies these conditions :

$$\forall\ \lambda, \lambda'\ 0 \leqslant c_{i,j}(\lambda, \lambda') \leqslant 1. \quad (5)$$

$$\sum_{\lambda} c_{i,j}(\lambda, \lambda') = 1 \quad (6)$$

- d_{ij} represents the influence coefficient of the i^{th} object from the j^{th} object which satisfies $\sum_j d_{ij} = 1$.

In practice, there are three parameters to define:

· The influence weigth d_{ij}

· The conditional probabilities c_{ij}

· The set of initial certainty measures denoted by $p_i^0(\lambda), \lambda = 1, \ldots m$.

Region probabilities belonging to the roof are initialised to 99% for regions near the segment and to the similarity function value elsewhere. The similarity function is obtained by cross correlation between both reprojected images on the plane $\pi(\theta)$. The correlation score is weighted with $\frac{1}{d}$, where d is the distance of the region center to the considered line. In fact, regions near the segment are likely to belong to the roof so we choose a high probability in order to propagate this information to the neighbourhood. Moreover, the similarity function is reliable since regions belonging to the roof are well reprojected onto the 3D-plane so that the similarity score should be high. The choice of conditional probabilities is not critical. So we use

$$Labels:\ \lambda, \lambda' \in \{" \in roof", " \notin roof"\}$$
$$c_{ij}(\lambda, \lambda) = 0.6$$
$$c_{ij}(\lambda, \lambda') = 0.4$$

Probabilities are close to 0.5 to ensure a slow upgrade of the system. The influence weight between regions that belong to the roof is chosen following three criteria: increasing in accordance with the frontier length between regions, decreasing with the gradient module along the frontier. In fact if the gradient value between two regions is high, the regions are likely not to belong to the same roof. Finally the gradient condition has more weight than the first one. This leads to the following influence function:

$$d_{ij} = \frac{L_{ij}.e^{-\lambda <grad>_{ij}}}{\sum_{k=0}^{n} L_{ik}.e^{-\lambda <grad>_{ik}}} \quad (7)$$

where

- L_{ij} is the frontier length between two regions.
- $<grad>_{ij}$ is the gradient module along the common frontier between region i and j.
- In practice, $\lambda = -3.5$. It ensures no overdelimitation of roofs.

The upgrade of the system leads to a final map of probabilities. The highest values i.e the clearest regions in the figure 5. are likely to belong to the roof. Finally, by means of a cut in the RAG, regions whose probabilities are higher than a fixed threshold ($P_{th} = 0.75$), are assigned to belong to the roof. Then we select only the connected regions among roof regions. The greatest one is kept as a final delimitation of 3D-facet.

The delimitation algorithm of roofs will facilitate the plane fusion and reduce the combinatory. It can also be used to delimit a priori correlation surfaces while searching for half planes.

5 Results

Due to the lack of available data , first results are presented either on simulated data computed by the CNES/ QTIS [1] on Toulouse (France) at 60cm (Figure 5-6) or on aerial images with a low B/H =0.2 on Amiens (France) at 50cm (Figure 7).

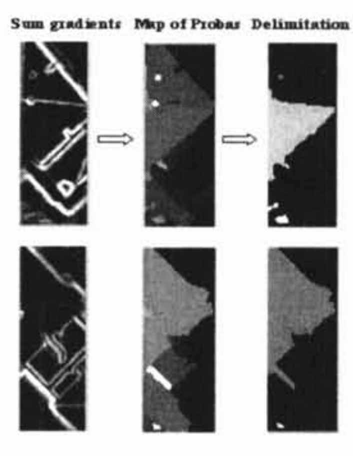

Figure 5: 3D-plane delimitation

Performance. In proposed approach, the quality of the reconstruction is governed by the completeness and the correctness of the input line set. Since a line is the only mechanism for instantiating 3D planes, if lines are missing, entire planes may be missed. Moreover, due to a low B/H of the stereopair , the lack of accuracy in the 3D reconstruction implies altimetric errors on extracted lines , which are spread at each step.

[1] Centre National d'Etudes Spatiales, France(The French National Space Center)
QTIS: Qualité et Traitement d'Images Spatiales

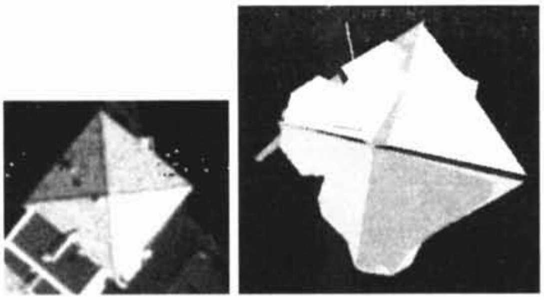

Figure 6: Toulouse (France) 60cm

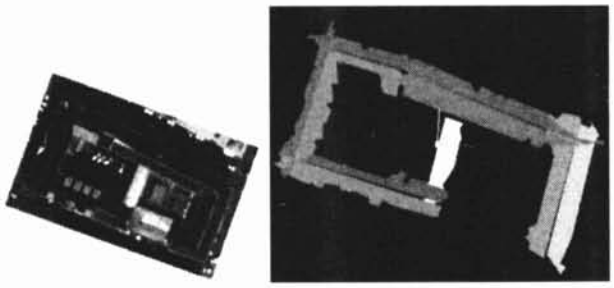

Figure 7: Amiens (France) 50cm

Numerical results on Amiens scene are presented : Segments are only evaluated with the angular deviation to the corresponding reference line. The segment length is not taken into account. A half plane Π_i is considered to be correct since it joins to his attached segment i another collinear segment. Finally delimitation is evaluated visually.

129 segments		
37 good (29%)		89 erroneous (71%)
13 segments 15 Π_i missing	24 segments 30 Π_i detected	7 erroneous planes

30 detected planes	
22 good (73%)	8 erroneous (27%)

22 good planes	
15 good delimit (68%)	7 erroneous delimit (32%)

Table 1: Results on Amiens Block

As shown in the Table 1, due to the lack of accuracy in the 3D reconstruction, only 29% of extracted segments are correctly positioned. Considering right segments, 65 % are found to be attached to a plane. 73% of detected planes are good, errors are due to facades. In fact, correlation scores are not reliable since planes are vertical. Finally good planes are well delimited in 68% of cases. Delimitation errors are essentially due to segments with an overestimated length where regions close to the segment extremities do not belong to the roof. Overdelimitation of roofs can be solved by 3D-facets intersections.

If we consider wrong segments, only 0.76% were attached to planes so the second algorithm is robust to errors.

As we can see, each algorithm gives very satisfying results since primitives are reliable.

6 Conclusion

In this paper, three algorithms used in the extraction of 3D-primitives are presented. Based on a low B/H and handling both views in a symmetric way, they all use photometric constraint at every step. The three detailed algorithms are part of a global strategy for extracting reliable 3D-primitives in order to generate 3D building model. Both last algorithms are based on primitives extracted in the previous step. However, each algorithm can be used independently providing specific primitives as input data. For instance the delimitation algorithm can be used either to delimit the surface correlation when searching for appropriate half planes or later for delimiting roof edges.

7 Future trends

The following step foresees fusion of planes which will validate extracted 3D-segments and find those missed during the first step.

Due to the low B/H ratio, half planes are often located with poor accuracy, so next goal is to use an algorithm based on region matching in order to reconstruct more reliable 3D facets and not only based on 3D-segments (for avoiding errors propagation). Both methods will be compared and might be used in a cooperative way.

Finally, studies will be carried out to evaluate the planimetric and height accuracy of 3D building extraction using the developed algorithms and thus specify the final 3D-object models that can be extracted automatically from such images.

Acknowledgments

This work has been carried out as a collaboration among the IGN[2] , the CNES and the University of Paris 5 (Research lab SIP/CRIP5).

References

[Alcharroun01] M,Alcharroun.*Positionnement de plans 3D par corrélation multi-vues*, stage IGN, 2001.

[Baillard97] C. Baillard. *Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains.* PhD thesis, ENST, 1997.

[Baillard99] C. Baillard, A. Zisserman. *Automatic reconstruction of piecewise planar models from multiple views.* In Proc, CVPR, 1999.

[Deriche89] R.Deriche. *Using Canny's criteria to derive a recursively implemented optimal edge detector.* International Journal of Computer Vision, 1(2) : 167-187, Mai 1987.

[Fu&Yan97] Fu.A,Yan.H. *A new probabilistic relaxation method based on probability space partition.* Pattern recognition, Vol.30, N11, Elsevier Science Ltd,1997.

[Harris88] C.Harris, M.Stephens. *A combined edge and corner detector.*In Proceedings 4^{th} Alvey Confer ence: 147-151, Manchester, August 1988.

[2]Institut Gographique National,France.(National Geographical Institute)