Rectangular Building 3D Reconstruction in Urban Zones

Thierry QUIGUER
ONERA-DES/STD, BP 72
92322 Chatillon CEDEX, FRANCE
quiguer@onera.fr

Commission III 1992-1996
Theory and Algorithms

KEY WORDS : Photogrammetry, Urban, Vision, Reconstruction, Algorithms, Edge, Stereoscopic, Three-dimensional

ABSTRACT :

Our paper concerns 3D reconstruction of buildings in urban and sub-urban zones by stereovision using vertical aerial images (resolution is in 40 cm range). Our images are well registrated, image lines correspond to epipolar lines. We limited our investigations to rectangular buildings because it is not an obvious problem. We propose a semi-automatic method in order to avoid major drawbacks of low-level processes. In effect, in low-level vision algorithms we need to introduce a priori knowledge (i.e. thresholds). So, in many cases we have to adapt thresholds to images. In order to overcome this particular unpleasant aspect, we focus our works on high-level process and we propose an original method to recognize building in an image. Our algorithm is semi-automatic because we select manually a corner then we apply our high-level algorithm. Results are very interesting because we obtain a good precision of detection and reconstruction. We compare our results with BDTOPO® (TOPOgraphic Data Base of French National Geographic Institute) which are truth data.

1. INTRODUCTION

Our paper concerns photogrammetry which consists in computing object dimensions by measures realized on perspective views of this object. We can find a large collection of papers concerning this domain, basic notions being available in the manual of photogrammetry [PhotoG 80]. Photogrammetry is a vast research domain so we deliberately restricted our investigation to rectangular building reconstruction which is not an obvious problem, see Figure 1 in order to illustrate this assertion. The size of this image is 2000 by 2000 pixels.

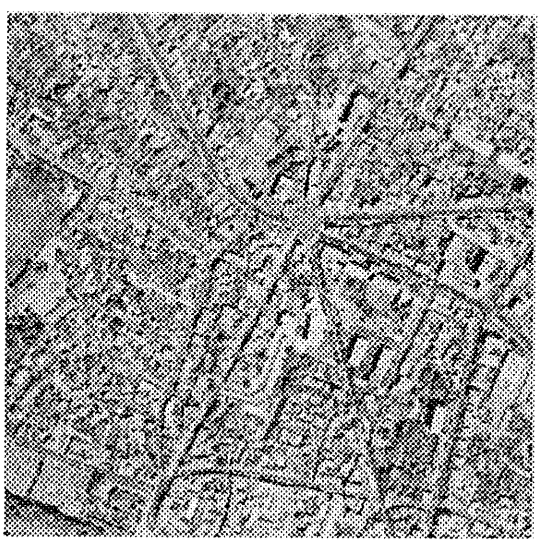


Figure1: Sub-Urban of Paris

Recent papers tackle this very difficult problem [Dang 94] [Dissart 95] [Gabet 94] [Huertas 88] [McKweon 93] [Maître 92] [Mohan 88] [Shufellt 93] and show that this problematics still stays a subject of interest in the international community. A common characteristic, about all these algorithms and about vision algorithms in general, can be pointed out: results of high-level process and consequently of complete process are dependant of low-level one. With this assertion two communities appear: those who neglect low-level process and consume time computational during high-level treatment, and those who try to have a perfect detection and consequently develop easy high-level technics. We think that an intermediate position will be better. Any detection process is perfect even if you provide several a priori knowledge. Thus we think it is important to overdetect primitives in image in order to provide all pertinent elements to the recognition level process. The job of high-level will be to separate good detections from false detections. We suppose low-level process provide weighted detection, weight qualifying quality of an element. This quality measure helps us during the high-level process.

Nevertheless, in order to be sure that all pertinent elements will be detected, we have to choose between several a priori knowledge and interactivity. We choose the second option because we hope to climb automation ladder (see figure 2) when detection problem will be resolved. In figure 2 we qualify our approach using classical critical systemic parameters used in literature. So, interactivity overcomes low-level problems and then we decorrelate some behavior parameters like *automation* and *complexity of a priori knowledge*.

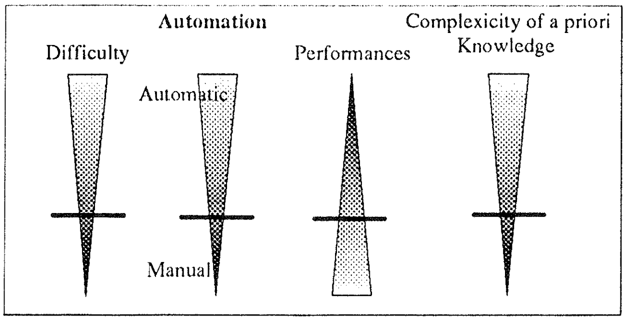


Figure 2 : Systemic parameters

Which form of interactivity will be introduced? In the form of an human operator who choices one corner of a building. By this way we hope we could give behavior independence between *automation* and *performances* and thus increase them.

We develop in details our approach in following parts of this paper. The next one concerns specificities of our approach. Then we explain our detection recognition reconstruction methodology. We will present our results along methodology explanation in order to be very clear. Then we will conclude and develop some perspectives.

2. OUR APPROACH

Today, most of problems encountered in recognition processes derive from poor performances of detection methods. Our approach consists in helping low-level process in order to control and understand high-level process and to find which modifications we will have to do in detection methods to increase performances. As we explain in introduction, we decided to tackle this particular problem by using interactivity. Our hope is that this interactivity could be exchange by an effective detection process. In conclusion we will present some perspective ideas in this sense. Immediate interest of interactivity is to reduce combinational by designing an object of interest. Our process follows four steps :

1- we select manually a building corner, we call it seed point in the following,

2- we detect two first sides of building passing by the seed point,

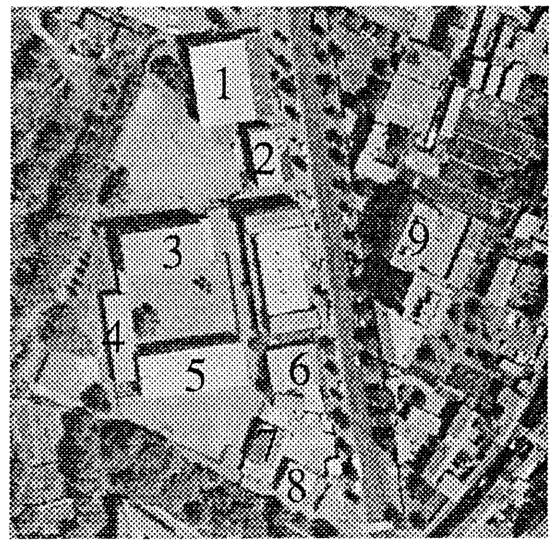
3- we detect the best parallelogram taking account a cost function using the two sides already detected,

4- then we search for homologous parallelogram in homologue image.

We present successively all details of these four steps in the next part.

3 - DETECTION AND RECOGNITION PROCESS

At first, we present area of interest (see figure 3) which we used in order to present our methodology. This image contains nine buildings which are marked and numbered from 1 to 9. We will use these numbers in the following in order to compare results.



(© French National Geographic Institute)

Figure 3: *Area of Interest*

3.1 Manual Seed Point Selection

Seed point (i.e. building corner) is chosen inside a zoomed area of image centered around the object of interest in order to localize precisely one of four building corners (see figure 4).

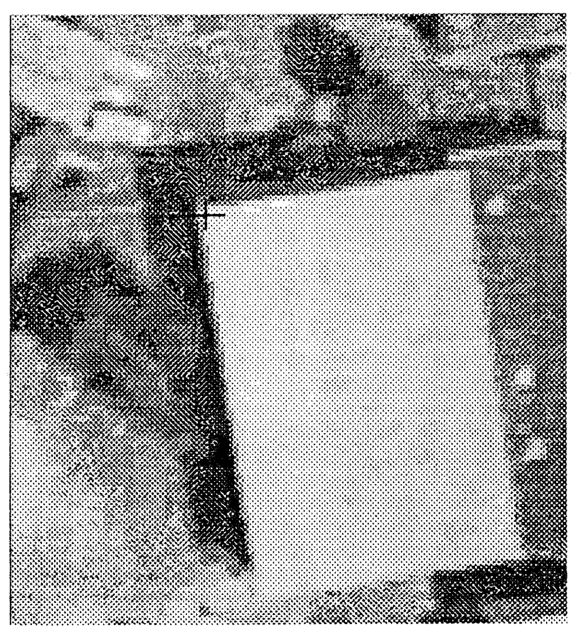


Figure 4: *Zoom of Building n°1*

Due to manual process selection, during the step 2 we will authorize to relax its position into a window of 5 pixels by 5 pixels in order to compensate bad manual selection.

3.2 - Two First Sides Detection

As soon as a seed point is chosen we proceed to monocular detection of building sides. For each seed potential position we apply a line detection process based on a criteria of radiometry discontinuity (i.e. gradient) and sign continuity of this discontinuity.

Selection criteria of the first side is based on biggest gradient along a line and on sign continuity of this gradient along the same line. Thus, for each line D_k (its equation being $Y=A_kX+B_k$) passing by the seed point we compute a cost function G_{D_k} which we try to maximize. This cost function takes the form of :

$$G_{D_k} = \sum_{i=0}^{i=n} grad[X_i][A_kX_i + B_k].S(i) \qquad (1)$$

$$\begin{cases} S(i) = 1 \\ if\ sign(grad[X_0][A_kX_0+B_k]) = sign(grad[X_i][A_kX_i+B_k]) \\ ifnot\ S(i) = 0 \end{cases}$$

Index i limits computation inside area of interest. S(i) express sign continuity along D_k line. When we have extracted the first side, it is very easy to find the second one because it is perpendicular to the first one. We used the same function cost to detect perpendicular side.

We used two types of gradient in order to maximize our function cost, the classical and the declivity ones. Results show that the second one provides best localization of the two sides detected. In effect, some detected sides are not lines with real building sides (see figure 5) when we used classical gradient, so we will keep declivity gradient (see figure 6) in the following (for more details about declivity operator see [Quiguer 91]).

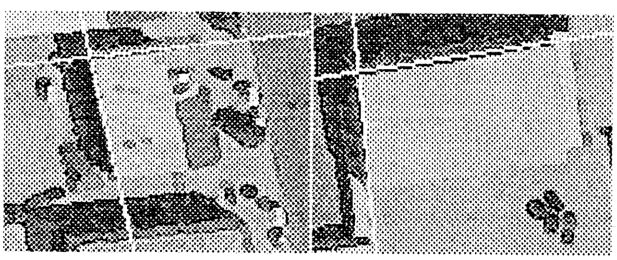


Figure 5: Use of Classical Gradient on Bdg n°2 & n°3

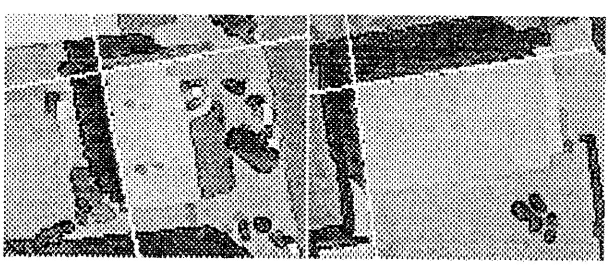


Figure 6: Use of Declivity Gradient on Same Bdg

3.3 - Parallelogram Closing

We used criterion of parallelism in order to close parallelogram which constitutes a building. So, we apply a set of parallel lines of the two first sides we have already detected. We realize closing by minimizing a cost function F_r. This function integrates homogeneity and discontinuity notions. Homogeneity appears inside roof of buildings and discontinuity on their sides. Homogeneity expresses likeness between grey levels inside building along two parallel sides (see figure 7). It has to be low ; it is computed by a difference of two means.

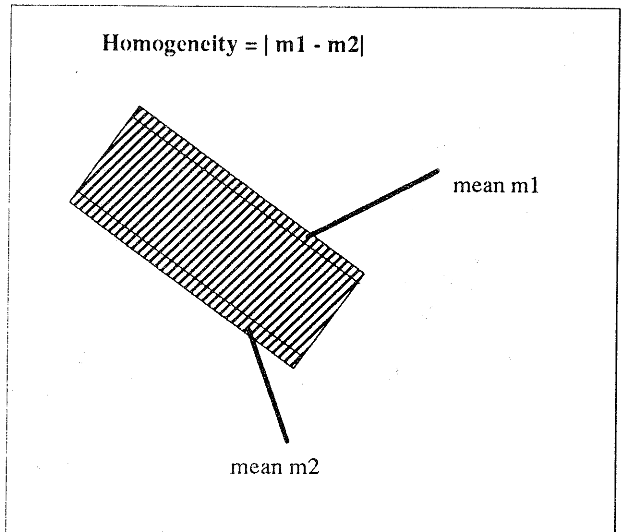


Figure 7: Computation of Homogeneity

Fr takes the form of :

$$F_r = Homogeneity\ /\ Gradient \qquad (2)$$

In effect, using only gradient is insufficient because urban zones are complex scenes and include several parallel sides belonging to different buildings. So, we can separate buildings using luminance criterion.

Nevertheless, using only this criterion is insufficient too because it can't exist local minimum of function F_r inside building. So we compute F_r with ranked gradient into a decreasing order. Optimum corresponds to the first local minimum (see figure 8).

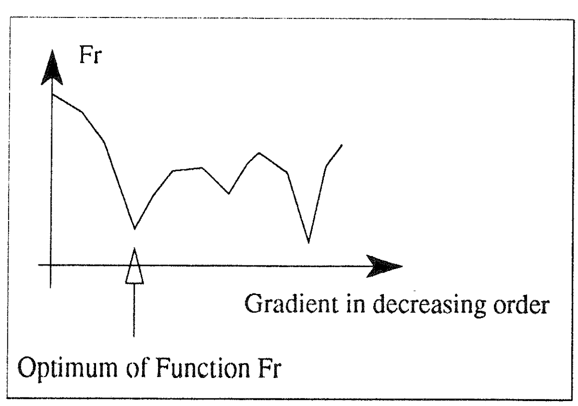


Figure 8: Optimum of Cost Function F_r

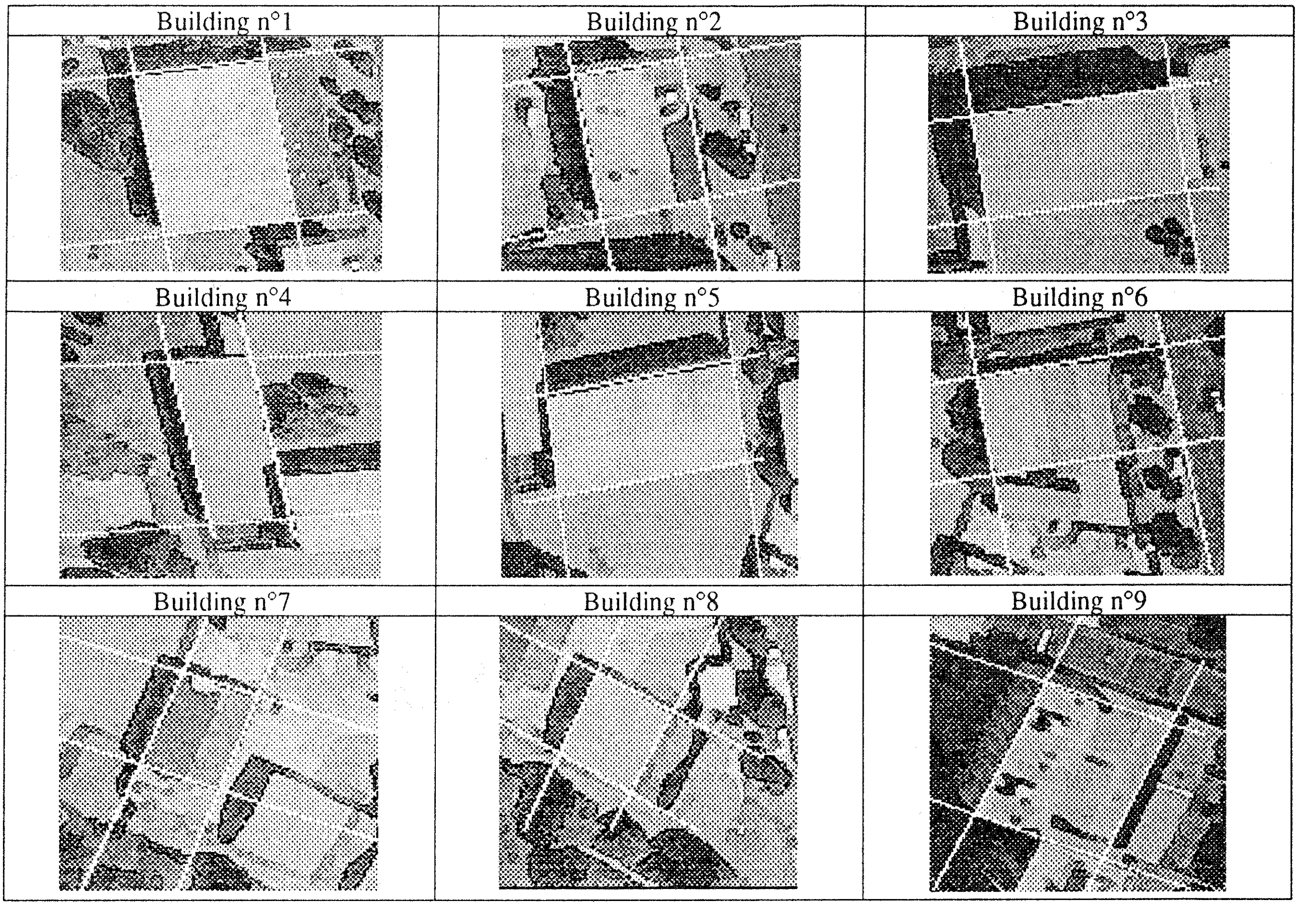


Figure 9: Closing of Rectangles

Results are very attractive (see figure 9). Nevertheless, some artifacts appear which can't be resolved either because our detection process is ineffective and we have to do several efforts to compensate errors, or because is due to our monocular approach and consequently this kind of error is redhibitory. In building n°2 we extract a false side due to bad continuity of gradient sign. In building n°9 due to luminances which are equal both on building roof and on ground we can't extract its side.

3.4 - Tridimensional Reconstruction

As soon as we recognize a building in an image by extracting its sides, we look for its homologous in the other view using normalized correlation. Maximum of correlation gives us disparity value of tested building and consequently its elevation. We compare our computed results with BDTopo® Data Base of French National Geographical Institute, and with manual measures of disparities (see table 1). We are under one pixel tolerance for major buildings.

Bdg	*n°3*	*n°4*	*n°5*	*n°6*	*n°7*	*n°8*	*n°9*
A	61	67	64	66	69	66	49
B	62	67	65	68	66	66	50
C	66,5	66,3	63,8	66,5	66,2	66,3	49,2

Table 1: Disparities Comparison

Line A of table 1 corresponds to our computed disparities, line B to manual disparities and C to BDTopo® ones. We do not provide disparities on buildings n°1 et n°2 because they don't exist in BDTopo®. Some examples are presented here after (see figure 10) and then two perspective views of our scene (see figure 11 and 12).

4 - Conclusion

Our approach presents some interesting aspects. First, it is possible to exchange quickly our interactive detection by an effective detection process. Nevertheless, this interactivity allowed us to realize a complete process without integration of low-level errors and consequently to better understand difficult points of low-level process. Second, the gradient we used (i.e. declivity one) allows better detection than classical one and consequently improves performances, we are under one pixel of error at the end of the process.

In perspective, we think that a binocular detection will better like [Lotti 94] thus it is possible to reconstruct buildings which have not horizontal roof.

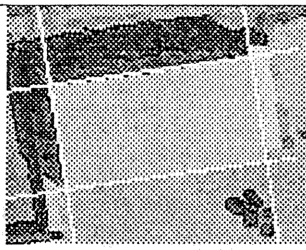
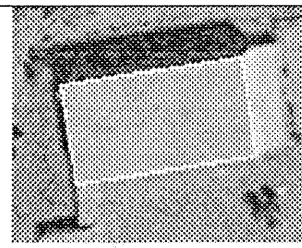
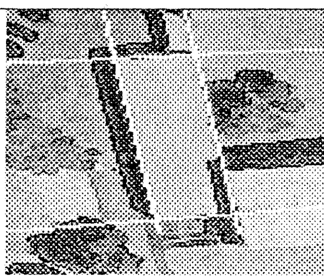
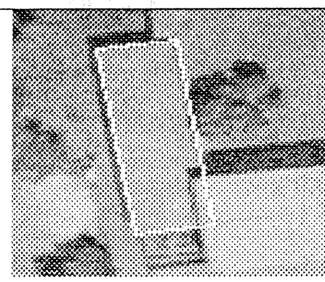
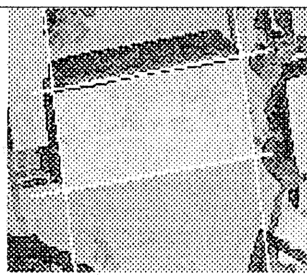
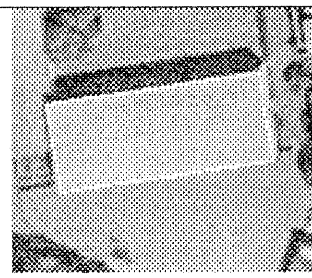
Left Image	Right Image
Building n°3	
	
Building n°4	
	
Building n°5	
	

Figure 10: *Results of Matching Process*

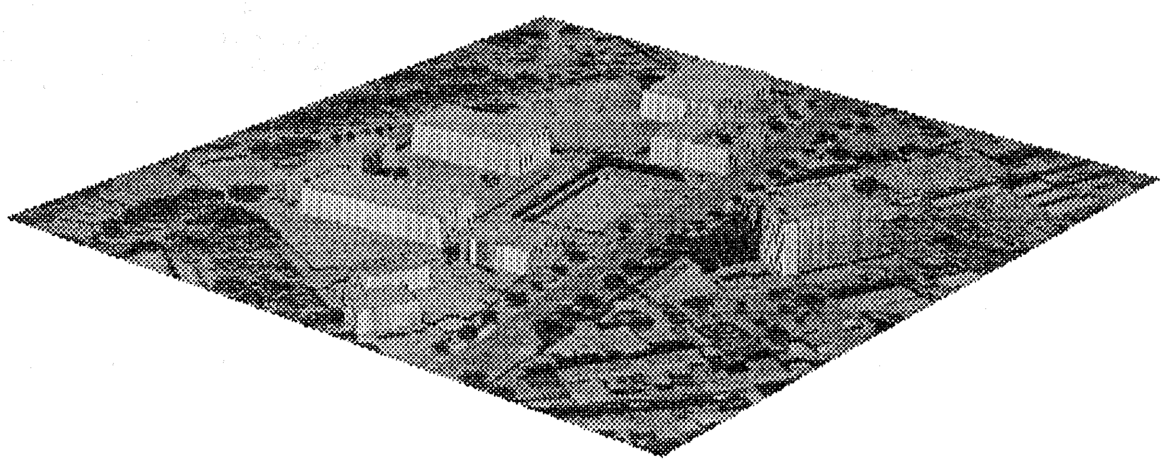


Figure 11: *First Perspective View*

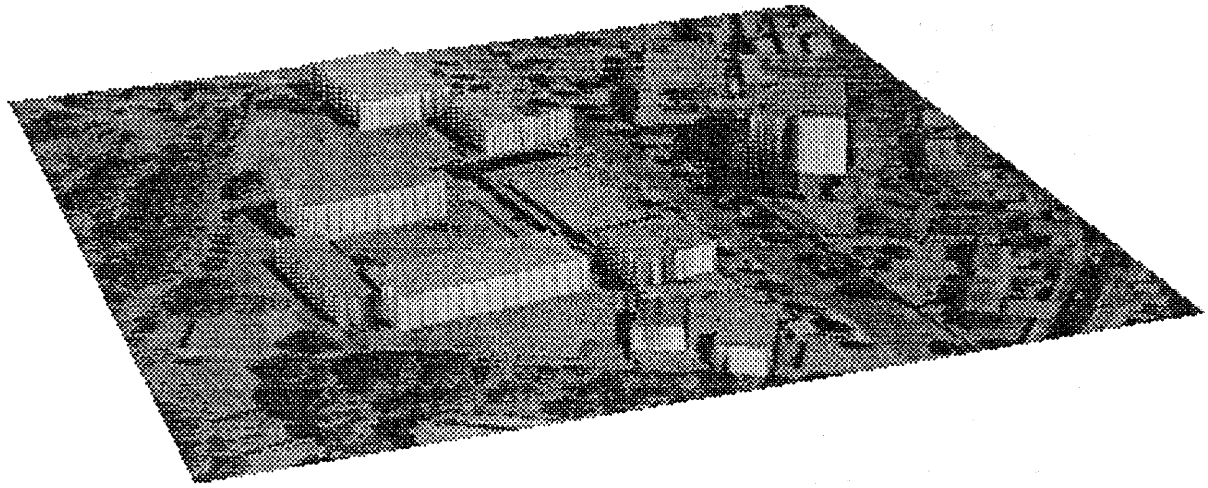


Figure 12: Second Perspective View

5- Acknowledgments

Thanks to MATIS Laboratory of French National Geographical Institute which provides us stereo images, particularly to O. DISSARD and O. JAMET.

6 - References

[Dang 94] T.Dang, H.Maître, O.Jamet, O.Dissard, Using Stereo Matching and Perceptual Grouping to Detect Buildings in Aerial Images, actes du congrès EUROPTO, 26-30 september 1994, ROME, pp 64-75

[Dissart 95] O.Dissart, O. Jamet, 3D reconstruction of buildings from stereo using both monocular analysis and stereo matching, SPIE Aerosense, Orlando, vol. 2486, 1995

[Gabet 94] L.Gabet, G.Giraudon, L.Renouard, Construction Automatique de Modèle Numérique de Terrain à Haute Résolution en Zone Urbaine, Bulletin SFPT n°135, pp 9-25, 1994

[Huertas 88] A.Huertas, R.Nevatia, Detecting Buidings in Aerial Images, CVGIP, Vol.41, n°2, pp 131-152, 1988

[Lotti 94] J.L. Lotti, G. Giraudon, Correlation Algorithm with adaptative window for aerial image in stereo vision, SPIE, vol. 2315, 1994

[McKeown 93] D.McKeown, J.McGlone, Integrating Photogrammetric Techniques with Scene Analysis aand Machine Vision, Proc. of SPIE, Vol.1944, 14-15 April 1993, Orlando, Florida, USA

[Maître 92] H.Maître, W.Luo, Using Models to Improve Stereo Reconstruction, IEEE trans. on PAMI, Vol.14, n°2, 1992, pp 269-277

[Mohan 88] R. Mohan, R. Nevatia, Perceptual grouping for the detection and description of structures in aerial images, Proc. of Image Understanding workshop, Ed. Morgan Kaufmann, San Mateo, CA, USA, 1988

[PhotoG 80] Manual of Photogrammetry, Fourth Edition, American Society of Photogrammetry, 1980

[Quiguer 91] T. Quiguer. P. Miché, R. Debrie, Segmentation by auto-adaptative thresholding, 6th International Conference on IAPR, COMO, ITALY, September, 1991

[Shufelt 93] J.Shufelt, D.Mckeown, Fusion of Monocular Cues to Detect Man-Made Structures in Aerial Imagery, CVGIP: Image Understanding, vol.57, n°3, pp 307-330, May 1993