A MULTIRESOLUTION AND OPTIMIZATION-BASED IMAGE MATCHING APPROACH: AN APPLICATION TO SURFACE RECONSTRUCTION FROM SPOT5-HRS STEREO IMAGERY

M. Pierrot-Deseilligny, N. Paparoditis

MATIS laboratory, Institut Géographique National
2, avenue Pasteur 94165 Saint-Mandé Cedex - Firstname.lastname@ign.fr

Commission I, WG I/5 & WG I/6

KEY WORDS: High resolution mapping, satellite imagery, SPOT5, stereo, automation, image matching, optimization, surface reconstruction, digital surface models

ABSTRACT:

This paper addresses the multi-resolution implementation of a Cox&Roy optimal flow image matching algorithm. This minimization algorithm aims at solving the surface reconstruction problem formalized as a minimization of an energy. This multi-resolution approach is necessary for achieving reasonable processing times on extended areas and improving robustness by restraining matching ambiguities. Some very good results are shown on a 10m ground sample distance (GSD) SPOT5-HRS stereopair. Intermediate results at different steps of the reconstruction process show all the Level Of Details provided by the different pyramid scale levels. The tuning of the regularization parameter is also discussed for different landscapes.

1. INTRODUCTION

Digital Surface Models are a key product for many applications ranging from orthoimage rectification to the generation of 3D city models but also are crucial for applications such as change detection for emergency mapping. Latest earthquakes have shown again the importance of obtaining this surface information on extended areas in the shortest time as possible in conditions in which airborne platforms are not adapted. The 120 km swath and the high revisit capacity and the stereoscopic imaging capabilities along the track (reduces diachronic artifacts that usually alter the matching quality) thus makes SPOT5-HRS a very good tool for these emergency missions.

Many robust image matching and surface reconstruction techniques have been developed in the last years particularly adapted for digital aerial cameras in the case of multi-view imagery [Paparoditis&al 2001]. Unfortunately, in these emergency cases, most of the time only two images are available. Thus the necessary robustness has to be achieved through dense optimization–based matching techniques.

2. SURFACE RECONSTRUCTION AS AN ENERGY MINIMIZATION PROBLEM

Regularization-based algorithms

In the set of all image matching-based surface reconstruction techniques, we will focus on those who can be formalized under the form of the minimization of an energy function. The methodological advantages of these approaches is to explicitly separate the what (the energy function) from the how (the minimization algorithm). These approaches define a global function on the image field which can be written as:

$$E(Z) = \sum A(x, y, Z(x,y)) + \alpha * F(\vec{G}(Z))$$

where

- Z is the unknown altitude function we are looking for;
- $A(x, y, Z(x,y))$ is the data attachment term measuring the image consistency/similarity at the image projections of point (x, y, Z) ; for example

$$A(x, y, Z(x,y)) = 1 - Corr(x, y, Z(x,y))$$

where $Corr(x, y, Z(x,y))$ is the normalized cross-correlation score ;

- $F(\vec{G}(Z))$ is a positive function which depends on the variations of Z, it is the regularization term which expresses the *a priori* knowledge of the surface regularity;
- α is a parameter weighting the relative importance of data attachment and regularization.

In this approach, we can distinguish to subclasses. On the one hand, the differential approaches which make the assumption that the relief and $F(\vec{G}(Z))$ are differentiable

(e.g. a quadratic function) and suppose the existence of an initial solution close enough to the final solution. On the other hand, the combinatorial approaches, that take for example for $F(\vec{G}(Z))$ a term such as $F(\vec{G}(Z)) = |Z(x+1,y)-Z(x,y)| + |Z(x,y+1)-Z(x,y)|$, that we will note $\Xi(Z)$; these approaches are looking for a global minimum with graph theory algorithms, they are implemented by sampling the possible values of Z (voxels in object space) ;

We will focus on the combinatorial approaches because, on the one hand we do not always have these close enough initial solutions and on the other hand, we can not assume that the relief is differentiable in urban areas. We will note E_α the energy function defined by:

$$E_\alpha(Z) = \sum 1 - Corr(x,y,Z) + \alpha * \Xi(Z)$$

Optimal flow algorithms

Up to recently, no algorithm allowing the calculation of a global optimum of the function E_α had been proposed ; in general a sub-optimal solution was calculated by analyzing the image line by line. Indeed looking for the optimum within a line (or a column) can be seen as a classical dynamic programming problem (e.g. [Baillard 97]); the drawback of these methods is to introduce a very strong dissymmetry in the processing of image lines and columns which has as a consequence the introduction of important artifacts in the result.

In 1999, the "founder" paper of Roy&Cox [Roy&Cox 98] showed that one could construct a graph such as:

- the nodes are the possible values of Z;
- the edges are the pairs of neighboring Z (in the plane or in altimetry);
- the planimetric edges are assigned the regularization cost and the altimetric edges are assigned the data attachment cost;

One can than demonstrate that the surfaces are the set of graph cuts between the set of nodes of maximal Z and nodes of minimal Z; moreover the weight of a cut associated to a surface is exactly equal to E_α. Finding the surface minimizing E_α can be seen as finding a minimal cut in a graph. Thus the problem can be solved in polynomial time with classical minimal cut and maximal flow graph theory algorithms.

3.MULTI-RESOLUTION IMPLEMENTATION

Even though polynomial, the major drawback of the Roy&Cox is its relative slowness and its memory greediness. In practice, it is very difficult to use it such as for massive image matching on airborne or spaceborne imagery. It would take several months to complete the matching of a complete HRS stereopair (2 images of 700 Mo) with a Pentium 4 PC.

We have thus implemented a multi-resolution variant of the Roy&Cox algorithm. We construct a pyramid of N images of resolution $2^K_{K \in [1;N]}$. At the coarsest resolution, we look for an optimal surface in the object/terrain space "matching cube" (composed of a similarity score [Paparoditis&al 00] for each voxel (X,Y,Z)) and in the complete Z research interval. In practice on a HRS stereopair, we choose N=16 or 32, leading to images of few Mo and disparity intervals of few dozens of pixels. Then at the current step, of resolution 2^K, we use the previous step, of resolution 2^{K+1}, as a predictor enabling a combinatory reduction. More precisely, the multi-resolution prediction is operated in the following way. Let $Z_{K+1}(x,y)$ be the surface at step $K+1$ calculated with a sampling distance $\partial Z2^{K+1}$. To calculate a surface at step K with a sampling distance $\partial Z2^K$, we first define an estimation $Z_K^0(x,y) = 2 * Z_{K+1}(\frac{x}{2}, \frac{y}{2})$. Let Δ_{xy} be the planar dilatation and Δ_z the altimetry dilatation; We note $\circ \lambda$ the morphological dilatation by a square of size λ ($\bullet \lambda$ for erosion). We define $Z_K^+(x,y) = Z_K^0(x,y) \circ \Delta_{xy} + \Delta_Z$ and $Z_K^-(x,y) = Z_K^0(x,y) \bullet \Delta_{xy} - \Delta_Z$. We then limit possible values of Z to the ones bounded by the upper surface Z_K^+ and the lower surface Z_K^-.
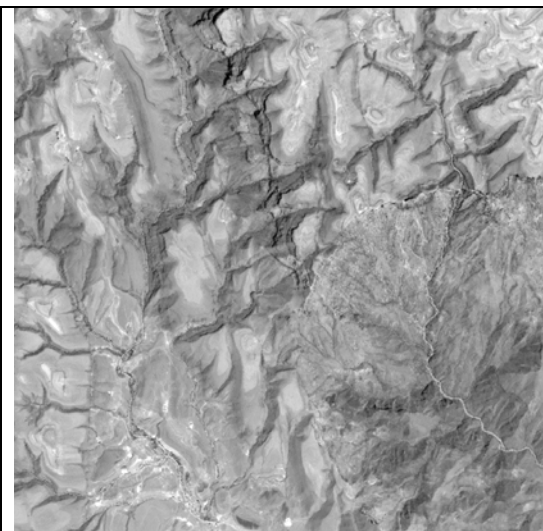
4. RESULTS

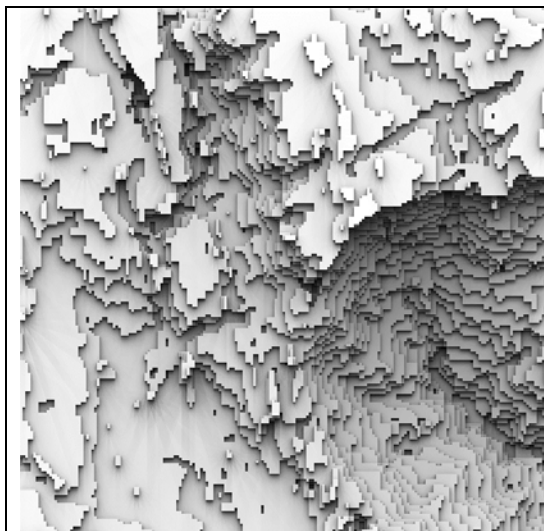**Results on SPOT5-HRS stereopair over a mountainous area**

The method described has been used to generate DSMs on 10 meter GSD SPOT-HRS stereopairs. The images were rectified in epipolar geometry. Fig.1. illustrates the process evolution on a mountainous and desert area. For this kind of landscape, the parameters used were:

- dilatation of 4 in altimetry and 8 in planimetry ;
- disparity sampling distance (in epipolar geometry) of 1 pixel ;
- template matching window sizes of 3x3 ;
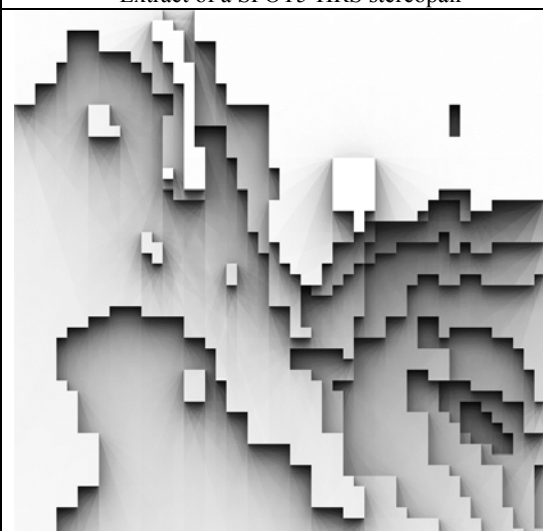- regularization parameter of 0.2 ;

The calculation time with a 3 GhZ PC, was 3 days for the complete stereopair (two images of 700 Mo of size 12000*60000). The quantitative evaluation is not yet finished but an experiment on a test site of "les dentelles de Montmirail » has shown an RMS of 3.5 meters. All Digital Surface Models shown have been omni-directionally shaded to increase the visibility of all morphological details.
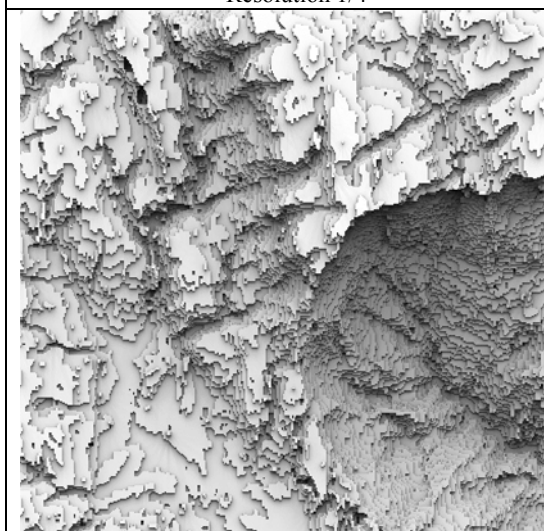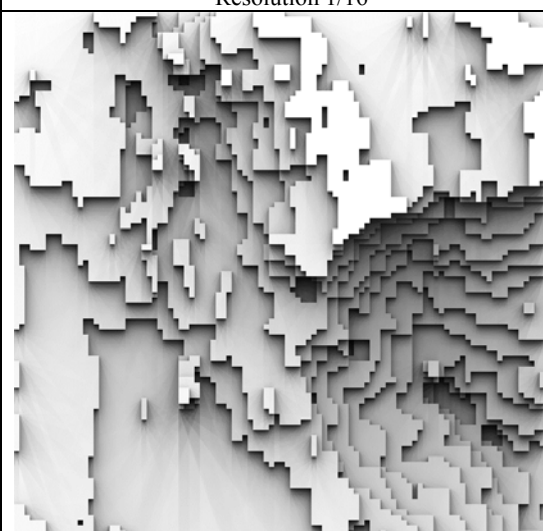
Extract of a SPOT5-HRS stereopair
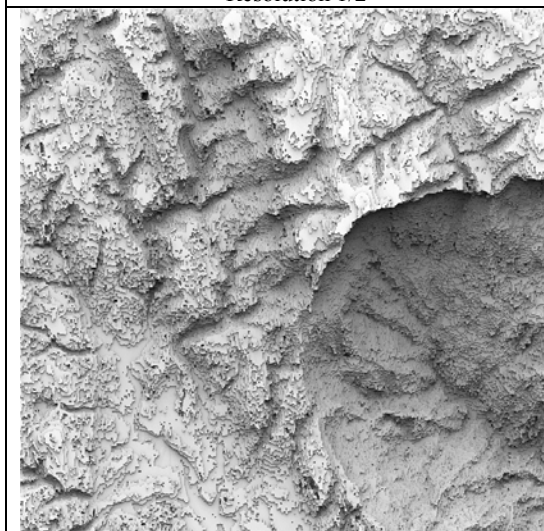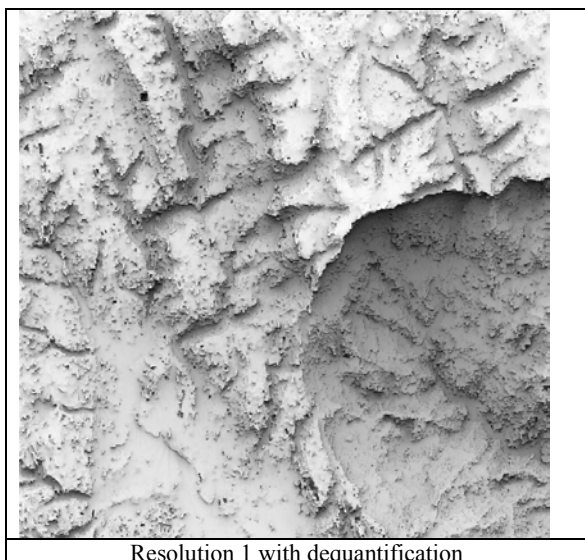

Resolution 1/4

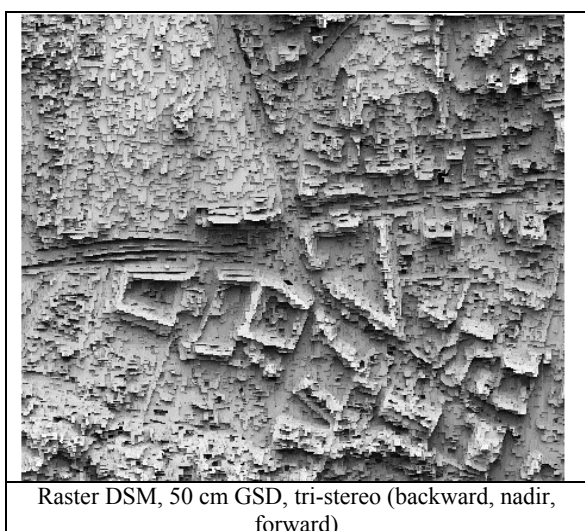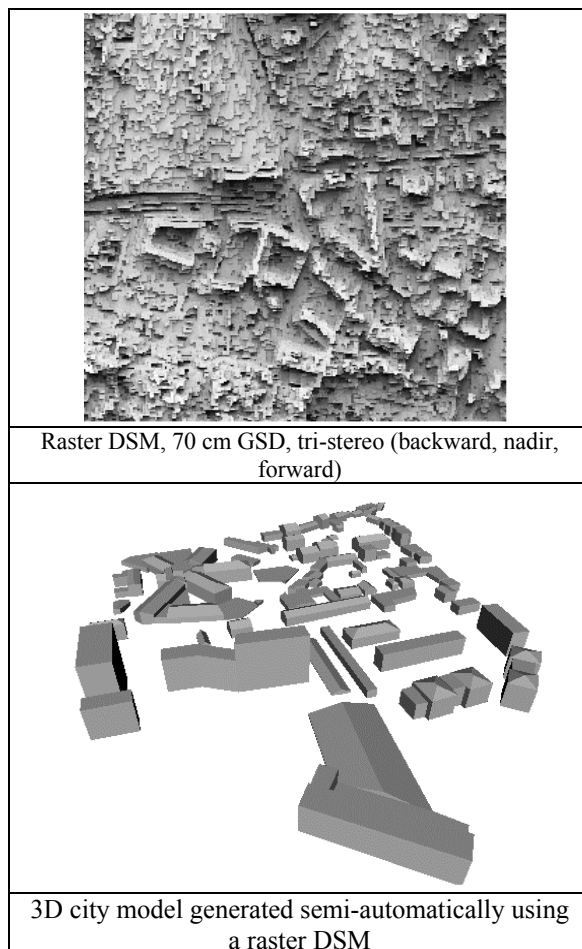Resolution 1/16

Resolution 1/2

Resolution 1/8

Resolution 1

Resolution 1 with dequantification

Figure 1: DSM at different pyramid levels

Results on a simulated Pléaides data set over urban areas

The same method has been used to generate urban areas DSMs from simulated data sets of the CNES Pléiades satellite. The main difference with the rural area test is in the tuning of the regularisation parameter; indeed in an HR urban context, due to facade depth discontinuities, there is no need for injecting a strong *a priori* knowledge on the regularity of the relief. The parameter was wery weak and tuned to 0.04. Fig. 2. shows some results on along-track backward, nadir, and forward tri-stereo. We have not yet performed quantitative evaluations on this data set. Nevertheless, the results are sufficiently robust and accurate to be used as an input for automated 3D city models reconstruction algorithms. The right image of Fig. 3. shows an example of the 3D models generated with such a framework [Maillet&al 06] using the 70 cm GSD DMS. The accuracy of the 3D building models is a little less than one meter.

Raster DSM, 50 cm GSD, tri-stereo (backward, nadir, forward)

Raster DSM, 70 cm GSD, tri-stereo (backward, nadir, forward)

3D city model generated semi-automatically using a raster DSM

Figure 2 : Results on simulated images of the future Pléiades satellite

5. CONCLUSION

We have presented a multi-resolution optimization-based approach for image matching. It provides robust reconstructions in reasonable processing time and reconstructed surfaces are morphologically preserved due to the fact that the template windows are chosen small (3x3) thanks to the optimization process. This approach has also been applied very successfully to multi-view imagery such as encountered in very high resolution imagery provided by aerial digital frame cameras.

Aknowledgments

References

Baillard, C. , 1997. Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains. Détection et caractérisation du

sursol, thèse de doctorat, ENST, laboratoire MATIS, IGN-SR-97-005-C-THE-CB.

Maillet, G., Flamanc., D., Buissart, H., Cantou, J-P. « Preparing the use of Pléaides-Hr images for mapping purposes: preliminary assessments at IGN-France», in Proc. of the Commission I workshop "Topographic mapping from space", Ed. Ugur Murat Leloglu (CD-Rom), Ankara, Turkey, February, 2006

Paparoditis N., Thom C., Jibrini H., 2000. « Surface reconstruction in urban areas from multiple views of aerial digital frames », in proc. *IAPRS*, vol. XXIII, Amsterdam, 2000.

Paparoditis N., Maillet G., Taillandier F., Jibrini H., Jung F., Guigues L. and Boldo D., 2001, Multi-image 3D feature and DSM extraction for change detection and building reconstruction. In E.P. Baltsavias, A. Gruen and L. Van Gool (eds) *AutomaticExtraction of Man-Made Objects from Aerial and Space Images III* (Lisse, The Netherlands: Balkema), pp. 217-230.

Roy, S., Cox, I.J., 1998. « A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem», *Proc. IEEE Intern. Conference on Computer Vision*, pp. 492-499, Bombay.