

Merging of SAR and optical features for 3D reconstruction in a radargrammetric framework

Florence Tupin

Dépt. TSI, Ecole Nationale Supérieure des télécommunications
46 rue Barrault, F-75636 PARIS Cedex 13

Email: Florence.Tupin@enst.fr

Abstract—The aim of this paper is to propose a framework for the use of both SAR and optical data in a 3D reconstruction process. The SAR data provide height information either by interferometric or radargrammetric process and the optical data provides building shapes. The method is based on a Markov random field defined on a Region Adjacency Graph. The regions are obtained using a segmentation of the optical image. The graph is then fed by the height information (either interferometric or radargrammetric) computed with the SAR data. The Markovian regularization takes height discontinuities into account thanks to an implicit edge process.

I. INTRODUCTION

SAR images are nowadays widely used for Digital Terrain Model (DTM) production[1]. The interferometric potential of the Synthetic Aperture Radar (SAR) data permits to obtain accurate DEM and the recent SRTM mission with an almost full coverage of the earth is a new proof of this capability. Nevertheless, SAR data are also able to produce height information using the classical stereo-vision principle [2]. Both techniques are powerful in the case of satellite images (ERS, RadarSat, EnviSat,...). With the increase of the resolution of aerial campaigns and for the future satellite sensors like CosmoSkyMed, people are now interested in the Digital Elevation Model (DEM) production, specially in urban or semi-urban areas. But for both techniques in these areas, the object (building) definition remains rather difficult due to the speckle noise and to the particular backscattering phenomena. Therefore, the DEM quality remains often insufficient [3] [4] [5] and for instance multiple views must be acquired to improve the result [6].

In this paper, we are interested in the introduction of an external knowledge to help the 3D reconstruction process. A methodology to introduce an optical data in the DEM construction is proposed. The optical image (which can be of lower resolution than the SAR data) provides information on the building shapes which is hard to retrieve directly from the SAR data. A Markov random field defined on a Region Adjacency graph (RAG) seems to provide a well adapted framework. For each region, height hypotheses are given by the SAR data (either interferometric or radargrammetric). These hypotheses are then regularized using contextual knowledge on the RAG and taking into account height discontinuities, thanks to an implicit edge process. The paper is organized as follow: section 2 defines the RAG; the generation of height hypotheses

in the case of radargrammetric data is described in section 3; section 4 presents the DEM reconstruction framework and the Markovian energy. The used data set and results are given in section 5.

II. RAG DEFINITION

SAR images are very difficult to interpret in urban areas. This is due to the combination of many phenomena. First, the distance sampling process induces many geometrical distortions (layovers and shadows, inversion of the apparition order of the points). Second, the well-known speckle phenomena spoils the image appearance. Third, since the SAR signal strongly relies on the geometrical properties of the objects, and due to multiple bounce scatterings, the SAR image is characterized by very bright features on a darker background. An example is shown in figure 1.



Fig. 1. Example of a building in the SAR (slant range) and optical image.

Due to the difficulty of SAR image interpretation in urban areas, the object shape information is extracted on the optical image. Any kind of segmentation providing an over segmentation could be used (the better the segmentation, the better the

reconstruction process). Here, a split and merge segmentation algorithm is used, with parameters tuned to obtain an over-segmented image [7] (figure 2).

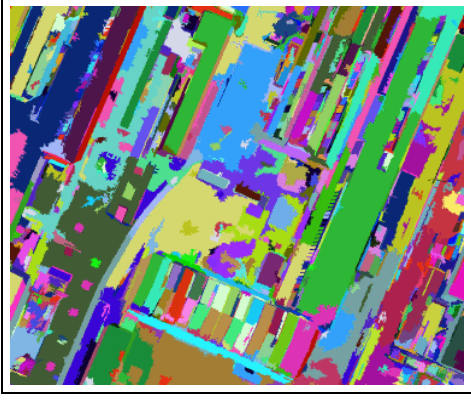


Fig. 2. Example of the segmented image by Suk algorithm

Starting from this set of regions, a graph is defined. Each region corresponds to a node of the graph and the relationship between two regions is given by their adjacency, defining a set E of edges. Let us denote by S the set of regions. The graph G is then $G = (S, E)$. For each region $s \in S$, \mathcal{R}_s^{opt} is the part of the optical image corresponding to s .

III. HEIGHT EXTRACTION

This information is given by the SAR data. The proposed framework is very general and could be applied both for interferometric or radargrammetric applications. Let us denote by E_{SAR} the set of features and their associated height extracted on the SAR image. This set of features can be pixels of high coherence in interferometry context or some extracted features in radargrammetric context. Indeed, classical approaches like correlation based approaches give very sparse and noisy results applied on SAR data [8]. Besides, the highly correlated values usually correspond to very bright features related to ground / wall corners, balconies, chimneys, etc. Therefore, instead of computing a dense disparity map with many mismatched pixels, a figural approach has been developed. It is based on a detection step of the bright punctual and linear features, followed by a matching step.

A. Feature detection

CFAR detectors adapted to linear and punctual features are used. They are based on the ratio of empirical means computed on adapted windows. The line detector has been introduced in [9] and the target detector in [10]. For both detectors the detection thresholds have been set empirically due to the difficulty of statistical modeling of SAR data in urban areas (although Fisher distributions could have been used to do the study [11]). The set S_l^A is the set of linear features extracted from the SAR image A, S_p^A the set of punctual features extracted in image A, and S_l^B and S_p^B the equivalent sets for image B.

B. Feature matching

Starting from the binary images, the matching step is done by computing the cross-correlation criteria between S_l^A and S_l^B and S_p^A and S_p^B . Thanks to the epipolar geometry the search area can be restricted in the azimuth direction. For a given association, the associated disparity can then be converted into an height using the sensor parameters.

As can be seen in figures 1 and 3, most of the bright linear features correspond to the ground / wall corner reflectors. Therefore they are lying on the ground and provide information about the ground height. To prevent wrong matching between features on the ground and features on the building roof, a preliminary step is done which aims at detecting (and then suppressing) all the ground corner reflectors. The association between S_l^A and S_l^B giving the best correlation coefficient is selected and the corresponding height is computed. Then the most predominant height h_g is deduced. As said before it has a high probability to correspond to an height close to the ground height (and it was verified in our experiments).

Then a second matching step is performed but constrained by h_g . Matching positions around h_g (a variation of 2 meters is allowed) and with a correlation coefficient higher than a fixed threshold are selected. The subset of linear matched primitives is denoted by S_{lg}^A (for images A) and S_{lg}^B for image B. An example is shown figure 3.

The underlying hypothesis is that the ground is rather flat. Then the matched primitives are suppressed from both sets S_l^A and S_l^B . The punctual features belonging to these lines are also suppressed (they should not have been detected but in practice it can be the case). Let us denote by S_{le}^A and S_{pe}^A the remaining linear (which should correspond to elevated structures) and punctual features (and equivalently S_{le}^B and S_{pe}^B for image B).

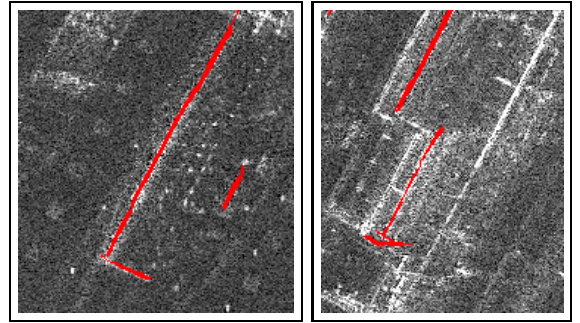


Fig. 3. Examples of features of S_{lg}^A (corresponding to building footprints). The remaining bright lines are elevated structures in S_{le}^A .

For each primitive a new matching is computed using the set of features of S_{le}^A and S_{le}^B (resp. S_{pe}^A and S_{pe}^B), and the best one is selected. A threshold on the correlation value is also applied to avoid wrong associations. Let us denote by \hat{E}_{SAR} the set of primitives (either for image A or B) p for which an association has been found, and by y_p the associated height.

Then the set E_s is defined for each optical region s by:

$$E_s = \{p \in \hat{E}_{SAR} / I^{opt}(p, y_p) \in \mathcal{R}_s^{opt}\}$$

with $I^{opt}(p, y_p)$ the image of the SAR primitive projected in the optical image using the height y_p . The projection is done using the equation system described for instance in [12]. Therefore E_s is the set of SAR primitives associated to each optical region. The observation associated to s is $y_s = \{y_p, p \in E_s\}$.

IV. DEM RECONSTRUCTION

The problem is modeled as the recovery of an height field H defined on the graph G , given a realization y of the observation field $Y = (Y_s)_{s \in E}$.

As usual in Markovian approaches, supposing that the solution H is Markovian, it can be shown that the MAP solution minimizes an energy which combines both a regularization term and a data fidelity term:

$$U(h|y) = U_{like}(y|h) + U_{cont}(h)$$

A. Likelihood term

This term is a data fidelity term and is built using the observations given by the SAR data. With an independence assumption, we can write $U_{like}(y|h) = \sum_s U_s(y_s|h_s)$. Many SAR primitives can be projected to the same optical region s . But among them some heights can be wrong due to false associations. Having no prior on the conditional probability of the SAR observations for a given height, but to take outliers into account, a truncated quadratic expression has been chosen:

$$U_s(y_s|h_s) = \sum_{p \in E_s} \min((h_s - y_p)^2, c)$$

B. Regularization term

This term implies that two adjacent regions have a higher probability of having the same height than two different heights, except for strong discontinuities inside the buildings. To take into account this case, once again a truncated quadratic function is used. Besides, an “edge” attribute is introduced to weight the potential between two regions. Indeed, in case of a high radiometric contrast between two regions of the image, they are likely to belong to two different areas (ground and building, roof and facade, etc...). Therefore, the regularization between the two regions should be suppressed. This is done using a coefficient γ_{st} between two regions s and t which is inversely proportional to the gradient (for a high gradient $\gamma_{st} \rightarrow 0$ and for a low gradient, $\gamma_{st} \rightarrow 1$). Eventually, the contextual energy is the following:

$$U_{cont}(h) = \sum_{(s,t)} \min[\gamma_{st}(h_s - h_t)^2, k]$$

C. Optimization

An Iterated Conditional Mode (ICM) algorithm is used [13]. Although only a local minimum of the energy is reached, it gives satisfying results with a good initialization and converges very fast.

The initialization is done by taking the minimum of the likelihood term for each region. It corresponds to the most representative height of the SAR primitives associated to an optical region.

V. RESULTS

The data set at our disposal is constituted by a radargrammetric couple of SAR images and an optical image with resolutions lower than 1 meter. The test area is an industrial area of Dunkerque, France.

The DEM analysis is done using a “ground truth” computed by classical stereovision with two optical images.

Figure 4 shows the optical image and the reconstruction result. The following remarks can be made :

- Some of the buildings are not reconstructed due to the lack of 3D information in the y_s data. This drawback could be avoided in interferometry where more information (although possibly noisy) is available. Indeed, some roofs have a very homogeneous response (radiometrically close to the one of the ground) in the case of smooth materials compared to the wavelength λ . In this situation, no height information can be extracted.
- This is also true for some parts of the buildings which are shadowed in the SAR images (for instance for the very big building on the left of figure 4). Due to radiometric contrasts and the edge process, the result is not regularized.
- Most of the building heights of figure 4 are in accordance with the ground truth. Some problems occur when the extraction of radargrammetric information fails (specially for small in size and height buildings).

Let us remark here that the use of the optical image is limited to the shape building extraction. In this case, an almost vertical view is well adapted since the building roof are well detected. We could also imagine to use the optical and SAR images to extract some 3D information (like in [14]).

VI. CONCLUSION

A very flexible framework to merge optical and SAR features has been proposed. It is based on a region adjacency graph built on the over-segmented optical image. The use of a Markovian field is powerful since a lots of prior knowledge can be introduced about region relationships. Two kinds of discontinuities have been taken into account: radiometric ones through a graph edge attribute (equivalent to a deterministic edge process) and height discontinuities inside radiometrically homogeneous regions through an implicit edge process modeled by the truncated quadratic function.

The use of interferometric and optical data in this framework, and tests on other areas, are both subjects of further work. Besides, other informations like lay-overs and shadows could be used to improve the reconstruction step [15]. It could be easily done by adding energetic terms in the global energy. In the same way, 3D informations coming from SAR and optical structures matching could also be introduced [14].

ACKNOWLEDGMENT

Thanks to ONERA for providing the SAR images and to Thales for the epipolar geometry re-sampling of the SAR images.

REFERENCES

- [1] T. Toutin and L. Gray, "State of the art of elevation extraction from satellite SAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 55, pp. 13–33, 2000.
- [2] H. K. Ramapriyan, J. P. Strong, Y. Hung, and C. Murray, "Automated matching of pairs of SIR-B images for elevation mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24, no. 4, pp. 462–472, 1986.
- [3] R. Bolter and F. Leberl, "Detection and reconstruction of human scale features from high resolution interferometric SAR data," *IGARSS'00*, 2000.
- [4] P. Gamba, B. Houshmand, and M. Saccani, "Detection and extraction of buildings from interferometric SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 1, pp. 611–618, 2000.
- [5] E. Simonetto, H. Oriot, and R. Garelo, "Radargrammetric processing for 3-D building extraction from high-resolution airborne SAR data," *IGARSS'03, Toulouse (France)*, pp. 780–782, 2003.
- [6] U. Soergel, U. Thoennessen, and U. Stilla, "Iterative building reconstruction from multi-aspect InSAR data," *Workshop '3D reconstruction from airborne laserscanner and InSAR data*, 2003.
- [7] M. Suk and S. Chung, "A new segmentation technique based on partition mode test," *Pattern Recognition*, vol. 16, no. 5, pp. 469–480, 1993.
- [8] F. Tupin, "Matching criteria for radargrammetry," in *IGARSS 2002*, vol. 5, Toronto, Canada, mar 2002, pp. 2608–2610.
- [9] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: application to road network extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 434–453, Mar. 1998.
- [10] A. Lopes, R. Touzi, and E. Nezry, "Adaptive speckle filters and scene heterogeneity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 6, pp. 992–1000, 1990.
- [11] C. Tison, J. Nicolas, and F. Tupin, "Accuracy of Fisher distributions and log-moment estimation to describe amplitude distributions of high resolution SAR images over urban areas," vol. 3, jul 2003, pp. 1999–2001.
- [12] F. Tupin and M. Roux, "Detection of building outlines based on the fusion of SAR and optical features," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, pp. 71–82, jun 2003.
- [13] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.
- [14] F. Tupin and M. Roux, "3D information extraction by structural matching of SAR and optical features," *ISPRS'2004, Istanbul, Turkey*, 2004.
- [15] F. Tupin, "Extraction of 3D information using overlay detection on SAR images," in *URBAN 2003, IEEE/ISPRS Joint Workshop*, Berlin, Germany, mar 2003.

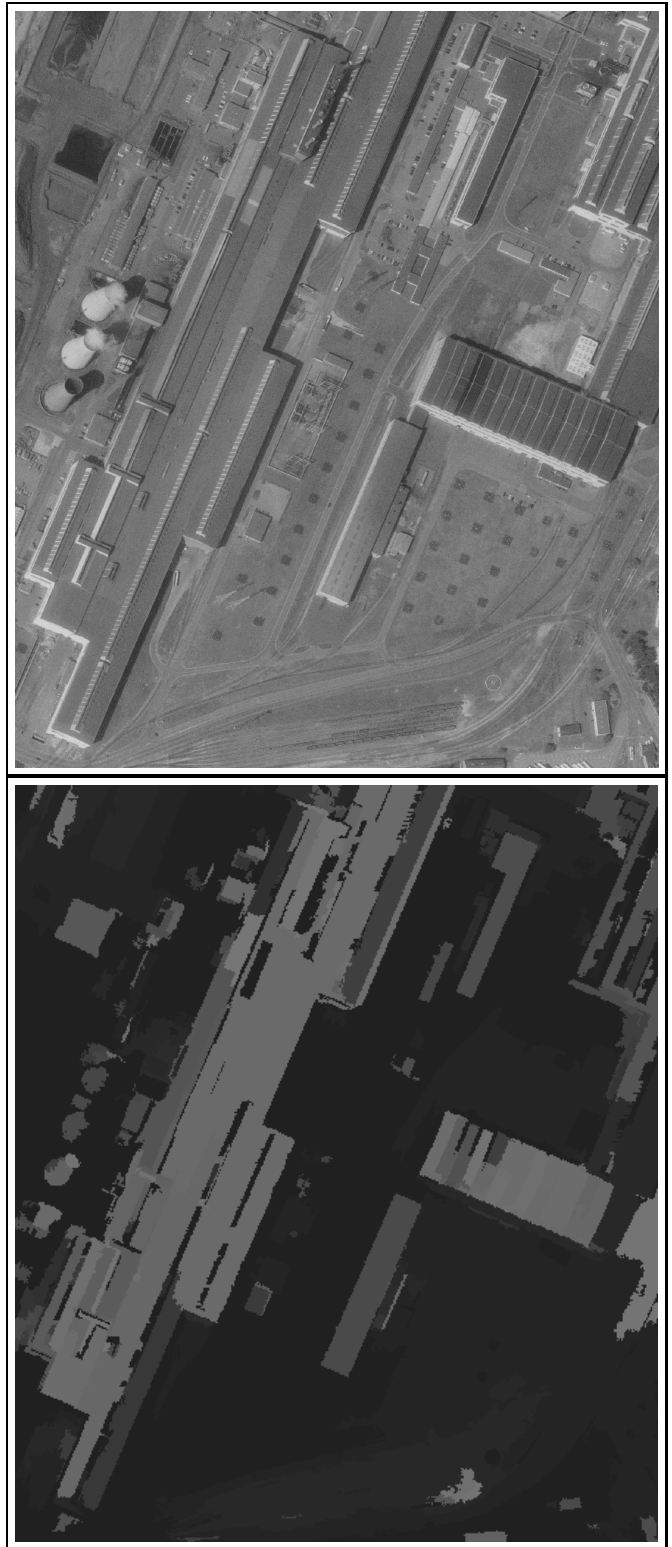


Fig. 4. Optical image of the test area and the corresponding reconstructed DEM.