Statistical Polygonal Snakes for 3D building reconstruction using High Resolution SAR data

Vincent Le Moigne
GET - Télécom Paris
LTCI-UMR 5141
Paris, France
Email: lemoigne@tsi.enst.fr

Florence Tupin
GET - Télécom Paris
LTCI-UMR 5141
Paris, France
Email: Tupin@tsi.enst.fr

Jean-Marie Nicolas
GET - Télécom Paris
LTCI-UMR 5141
Paris, France
Email: nicolas@tsi.enst.fr

***Abstract*— In this paper, the polygonal active contour method (proposed in [1], [2]) is adapted to the case of Synthetic Aperture Radar (SAR) images of urban areas. The statistical criterion is modified to be able to deal with multiple images in order to improve the segmentation of buildings.**

A criterion is proposed and is then implemented and tested over real high resolution SAR images containing urban areas. A discussion about the benefits of this approach is done and further work about $3D$ statistical active contour is introduced.

I. Introduction

Interpretation or $3D$ reconstruction of high resolution SAR images remains difficult in urban areas due to many phenomena: speckle presence and geometric distortions (lay-overs, shadows). Thus, it is often necessary to be able to segment the SAR image to help the interpretation step. Classical segmentation techniques like [3] give satisfying results but they do not take into account object shape. In this paper, we are interested in semi-interactive approaches like snakes [4], where starting from an initialization, an energy is minimized to find the object of interest. Here we deal with a specific family of statistical polygonal snakes first introduced in [1], [2]. They are characterized by two properties: first they are polygonal approximations of the region shape (optimization is done on position of the nodes); secondly, the energy is based on the region (inside / outside) statistics (instead of gradient on the contour position) which is particularly adapted to noisy images where the edges are almost unnoticeable.

In the statistical active contour introduction [1], [2], the grey level fluctuations are described with statistical laws. The kind of distributions followed in the regions is supposed to be known and is supposed to be the same for all the regions. Only the law parameters permit to distinguish between the two areas. The aim of the statistical active contour is to maximize the Likelihood over regions and thus to minimize a criterion which is defined as the opposite of the log-likelihood.

In this paper we present a $3D$ building reconstruction chain, based on the image segmentation using statistical active contours. We use the intensity SAR-image segmented using a Gamma-criterion or the phase SAR-image using a Gaussian-criterion to estimate building shape.

The two ways can lead to good results but no one is always better than the other. Thus, we propose a criterion merging the information of the two images in order to improve robustness of the shape determination.

The article is organized as follows. The first part presents the framework and the energy used in different presented cases. The second part presents the results and a discussion is done in the conclusion, before introducing $3D$ statistical active contour as further work.

II. Framework and formulation

In this section we will concisely introduce our $3D$ building reconstruction chain before discussing about the different active contour energies we have tested.

A. 3D building reconstruction chain

The process of extraction and reconstruction of the buildings follows these steps:

- Shadow detection and creation of a shadow map using a Markovian process over the whole SAR amplitude image [3].
- Initialization of the contours using the shadow map. Each significant shadow or group of shadows corresponds to a contour. All shadows do not inevitably lead to the set up of a contour, the shadow must have a wide area and not be too close from an other big shadow .Contours must not cross each others.
- Creation of the small images where building is searched for. Only one building is considered per small image.
- Convergence of the contour using the statistical active contour [1], [2].
- Reduction of the number of contour nodes using an MDL criterion (minimum description length) [5].
- Very simple height estimation of the building using the interferometric phase image.
- Projection of the contour in a $3D$ local space for the automatic generation of a *VRML 2.0* model of the building.

On interferometric phase images, the shadows induced by buildings contain noise only. Thus, we have decided not to consider the pixels contained in these low coherence areas in the calculus of the criteria. The mask formed by the shadow map is also used to constrain the contour not to consider these areas.

B. Notations

The region of the background is denominated by B, the region of the object is O, the region of shadow is S. Let B_s be the background without shadow. (x, y) are the coordinates of the pixel and $s(x, y)$ its value, and $\mathcal{R} = \{O, B_s\}$. The number of pixels of a region r is denominated by N_r.

C. Criterion for interferometric phase image

We have chosen a Gaussian model to represent regions in the phase image. We can directly use the Gaussian-criterion proposed in [2] :

$$\mathcal{C}_{gau} = \sum_{r\in\mathcal{R}} \left[\frac{N_r}{2}. \left(log(\tilde{\theta_r}) - log\left(2\sqrt{2\pi}\right) + 1 \right) \right]$$

$$with: \ \tilde{\theta_r} = \frac{1}{N_r} \sum_{(x,y)\,\in\, r} s(x,y)^2 - \left(\frac{1}{N_r} \sum_{(x,y)\,\in\, r} s(x,y) \right)^2$$

To improve the segmentation on the phase image we have re-sampled to 256 gray levels the original image. Then we apply a threshold to delete the 1% brighter pixels corresponding to noise only pixels with high values. The 1% threshold has been manually defined on the interferometric phase image of the whole scene.

D. Criterion for intensity image

Under the Goodman speckle hypothesis, the noise of the intensity images can be represented by a Gamma law. Then we can also directly use the Gamma-criterion proposed in [2] assuming that the number of looks L equals 1:

$$\mathcal{C}_{gam} = \sum_{r\in\mathcal{R}} \left[N_r. \left(log(\tilde{\theta_r}) + 1 \right) \right]$$

$$with: \ \tilde{\theta_r} = \frac{1}{N_r} \sum_{(x,y)\,\in\, R_r} s(x,y)$$

Extension using Fisher criterion has been investigated in [6] but the computational burden is not worth compared to Gamma Hypothesis.

E. Criterion for both images

We calculate a new criterion using the two images (with the previous model, Gamma-noised for intensity and Gaussian-noised for phase). Assuming that pixels of the intensity image (s_{int}) and pixels of the phase image (s_φ) are not correlated, while calculating the opposite of the profile log-likelihood over the two regions, on both images, we find that the new criterion can be simply written as follows:

$$\begin{aligned} \mathcal{C} \ = \ & \sum_{r\in\mathcal{R}} \frac{N_r}{2}. \left(log(\tilde{\theta}_{r,s_\varphi}) - log\left(2\sqrt{2\pi}\right) + 1 \right) \\ & + N_r. \left(log(\tilde{\theta}_{r,s_{int}}) + 1 \right) \end{aligned}$$

$$with: \ \tilde{\theta}_{r,s_\varphi} = \frac{1}{N_r} \sum_{(x,y)\,\in\, r} s_\varphi(x,y)^2 - \left(\frac{1}{N_r} \sum_{(x,y)\,\in\, r} s_\varphi(x,y) \right)^2$$

$$and: \ \tilde{\theta}_{r,s_{int}} = \frac{1}{N_r} \sum_{(x,y)\,\in\, R_r} s_{int}(x,y)$$

In fact this is exactly the same as simply making the contour to evolve using a criterion corresponding to the addition of the criterion taken over both images.

F. Working with small images

It is possible to segment the whole scene in one pass only but this leads to a very big calculation time. Moreover the density of probability of the pixels of the image corresponding to the ground areas are variable over the whole image. Thus we decided to segment small images instead, with the strong assumption that each small image contains one building only.

III. Results

Below are presented the results of the $3D$ building reconstruction chain introduced in the framework section. We show the segmentation results obtained with the three methods previously introduced. Then we propose the $3D$ reconstruction of the building using result from the last segmentation method.

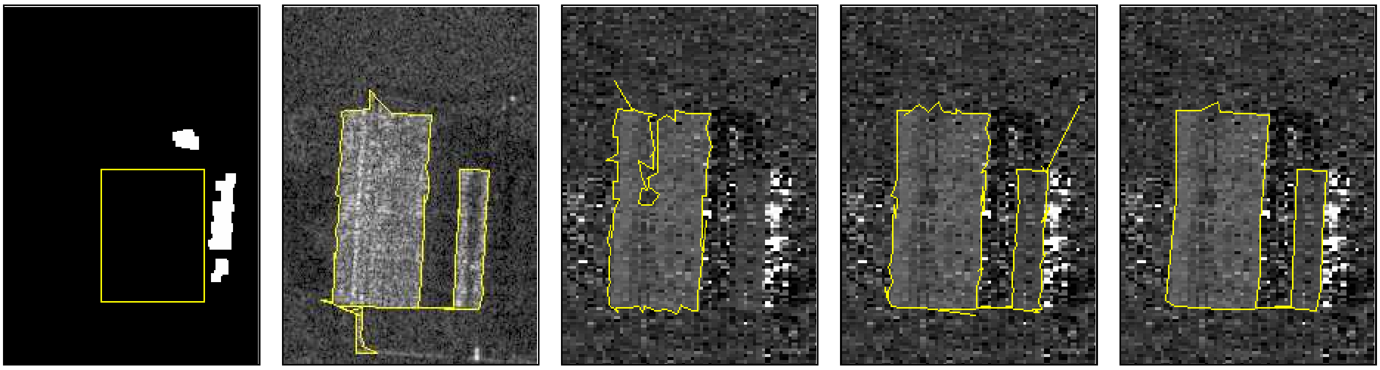
The output of the shadow map process and the contour creation steps over the whole scene are not presented, we will show instead, the extracted regions of maps of shadows and contours which are related to the building extracted images.

A. Segmentation

We present here on figures fig.1 and fig.2 the segmentation results of two buildings in a SAR image using the three previously presented ways with active contours and the images we have. Both amplitude and interferometric phase image have been modified to be easier to observe.

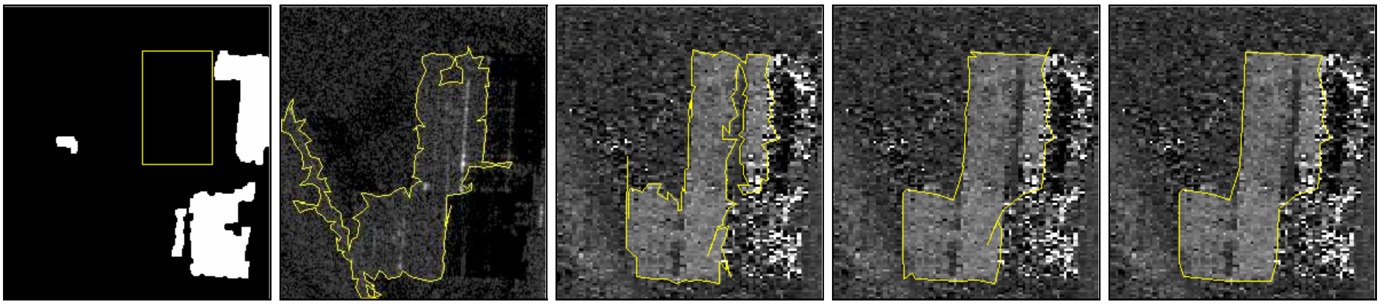
Figures have been organized from top-left to bottom-right as follows:

- **(a)** The initialization is shown on the image of the shadow map cropped to the small building image size.
- **(b)** The result of the segmentation with intensity image only.
- **(c)** The result of the segmentation with phase image only.
- **(d)** The result of the segmentation with the two images are superimposed on the interferometric image.
- **(e)** The number of nodes of the last result is then reduced with a criterion based on the MDL over interferometric image.



(a) (b) (c) (d) (e)

Fig. 1. Segmentation results of building n°1



(a) (b) (c) (d) (e)

Fig. 2. Segmentation results of building n°2

B. 3D models

After having segmented the building, an estimation of the mean height is done at this point on the phase image in order to roughly estimate the height of the building.

Because corner reflectors have the same height as the ground, they add a bias in the height estimation. Therefore the pixels of the phase image corresponding to the corner reflector are not taken into account to estimate the mean height. These pixels are obtained using a bright line detector (cf. [7]) on the amplitude image.

We also estimate the height of the ground which is subtracted to the building's one. Thus we have the relative to ground height of the building.

A *VRML 2.0* model of the building is then automatically generated after the building coordinates have been projected in the $3D$ local space with the help of the radar flight parameters.

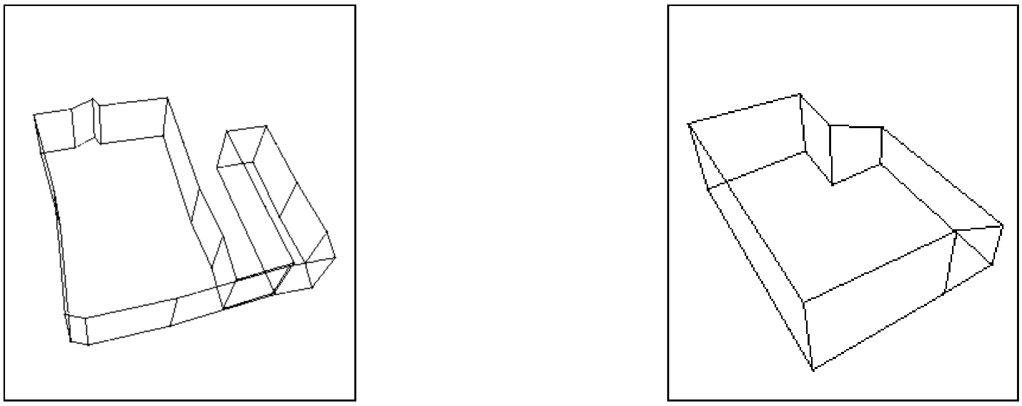
Final models of both buildings are presented bellow in figures fig.1 and fig.2. The figures shows on their left side *(a)* the $3D$ reconstructed model and on their right side *(b)* the $3D$ models calculated using IGN's[1] ground truth.

C. Interpretation

On the first building, the shape obtained using the interferometric image only (fig.1.c) is incomplete, the small part on the right is missing, although it is present on the shape obtained using the other image (fig.1.b). In addition, there is a hole in the detected roof on image fig.1.c due to the corner reflector which height is the height of the ground. The segmentation results using both image is greatly better than the results with interferometry only. This is comparable to results achieved using amplitude only, but the part of the shape due to the reflective corner in the amplitude image is reduced while using both images for segmentation.

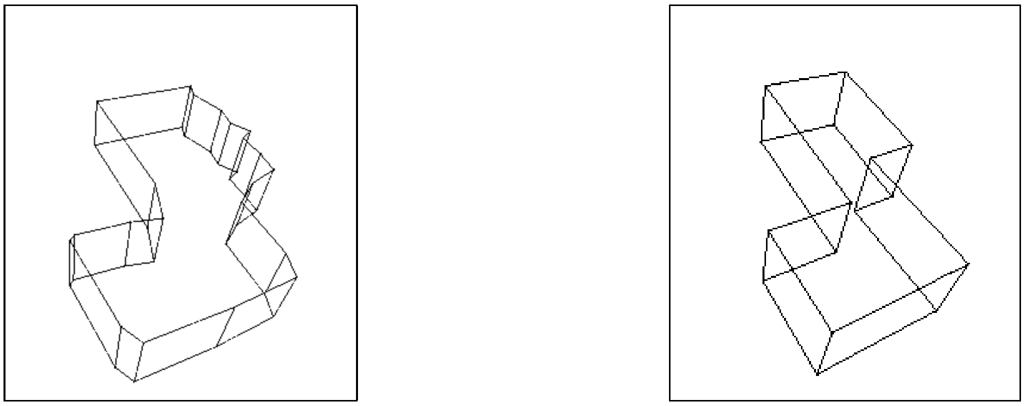
On the second building, the right part on the roof is better connected to the rest of the roof on image fig.2.d (both images used for segmentation) than on image fig.2.c (interferometric image only). Using both image for segmentation also permit to

[1]IGN: Institut Géographique National



(a) (b)

Fig. 3. $3D$ model of building n$^\circ 1$



(a) (b)

Fig. 4. $3D$ model of building n$^\circ 2$

avoid the fence integrated to the result shape while using amplitude image only (fig.2.c). Moreover the contour is smoother using both images than in the two other cases. But on the contrary, on the result presented on fig.1.d the reflective corner still have a contribution to the final shape (little L shaped excrescence on fig.2.d).

Using both image based criterion for the segmentation of the building leads to a smoother contour with a better shape. Complicated shape are better segmented.

We still have topology incoherence do to the nature of the sensor e.g on fig.1 the shadow induced by the left part of the roof on the second little part leads to a black area on amplitude image, and to a lower phase area on interferometric image.

Finally the $3D$ models (fig.3 and fig.4) reconstructed using the segmentation results are relevant: the shapes are correct except the gap between the two parts of the building n$^\circ 1$ (fig.3.a). The heights reconstructed are proportionately coherent but there a significant difference between estimated and real height: 12 meters for building n$^\circ 1$ for a real height of 14 meters and 14 meters for building n$^\circ 2$ for a real height of 18 meters. This might be partly due to the fact that the height is estimated on overlay and simple roof regions instead of estimating it using regions where the ground and the walls have no contribution.

D. Possible Improvements

There is a few improvements which could lead to a better buildings reconstruction which are presented in this paragraph.

First, in order to enhance the height, one could use a mean height calculated over roof area without taking into account the overlay areas, this way prevents to compute a height having a bias due to the ground and wall contributions. One can also use the shadows to estimate the height of the building using geometrical relation (*cf.* [8]).

Then, to add knowledge about the topology of the buildings (*eg.* building n$^\circ 1$), one can use an optical image of the same building, so we could be able to deal with buildings with multiple parts, able to know if roofs are adjacent or parts of different buildings. The building n$^\circ 1$ perfectly represents such a case; one building with two roofs with different heights which have been disconnected because of the shadow induced by the higher roof on the lower roof.

The actual roof model is very simple but sufficient to first approximate the building model. To enhance the roofs, one can consider an approach integrating a model of sloped roofs using a tangent plane detection over the phase image. (*cf.* [9]), but to do so one must have a very reliable interferometric phase.

IV. Conclusion

In this paper we have presented a criterion using both intensity and phase images at the same time, for the fast statistical active contour developed by Chesnaud et al.

The simple segmentation that could be achieved using only a Gamma-criterion on amplitude image, or a Gaussian-criterion on the interferometric phase image can be both improved by a joint optimization.

This new criterion using the two images improves and increases the robustness of the segmentation of the building in SAR images. The region found corresponds to the best compromise between intensity similarity over the regions and height similarity over the same regions.

As this active contours approach to the buildings reconstruction issue needs improvement concerning the height of the buildings estimation, roof topology problems and the roofs shapes, we have listed a few ideas that should lead to a better $3D$ reconstruction.

The aim of this study was to have a first reference result which can easily be achieved by using a basic active contour over interferometric phase image and amplitude image. The previously presented improvements will not be directly implemented but an overview of what we aim at is presented in the *Further work* section.

V. Further work

We would like to take into account the different regions of the building's signature in SAR images (Overlay, roof, shadow, reflective corner) over both interferometric and amplitude images. To do so, a $3D$ active contour will be set up. It will consist in a contour C set at an height H in the $3D$ local space. The $3D$ contour will evolve in this space and will be projected from the $3D$ local space into SAR image space using sensor parameters to compute the energy. The projection of the $3D$ contour in SAR image space will create several different contours corresponding to each significant part of the signature of the building (Overlay, roof, *etc.*).

The energy will have to gather all these significant regions of the SAR images and should improve the segmentation. We aim at optimizing both the height and the shape of the $3D$ contour at the same time, directly in the $3D$ space.

Furthermore we will be able to add an optical image, using the projection of the $3D$ contour into the optical image space. The final aim being the fusion of information existing in images acquired with both optical and SAR sources using the $3D$ active contour.

Acknowledgment

The authors are indebted to CNES[2] and ONERA[3] for providing the SAR and ground truth data.

References

[1] O. Germain and P.Réfrégier, "Optimal snake-based segmentation of a random luminance target on a spatially disjoint background," *Opt. Lett.*, vol. 21, no. 22, pp. 1845–1847, 1996.

[2] C. Chesnaud, P. Réfregier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 21, no. 11, pp. 1145–1157, 1999.

[3] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maître, "A new statistical model for markovian classification of urban areas in high-resolution SAR images," *IEEE transactions on geoscience and remote sensing*, vol. 42, no. 10, pp. 2046–2056, October 2004.

[4] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1987.

[5] F. Galland, N. Bertaux, and P. Réfrégier, "Minimum description length synthetic aperture radar image segmentation," *IEEE Transactions on Image Processing*, vol. 12, no. 49, pp. 995–1006, September 2003.

[6] V. Le Moigne, F. Galland, J.-M. Nicolas, and F. Tupin, "Statistical polygonal snakes with Fisher distribution," *EUSAR 2006, Dresden- Germany*, April 2006.

[7] F. Tupin, H. Maitre, 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, pp. 434 – 453, March 1998.

[8] A. Bennett and D. Blacknell, "Infrastructure analysis from high resolution SAR and InSAR imagery," *Remote Sensing and Data Fusion over Urban Areas, 2003. 2nd GRSS/ISPRS Joint Workshop*, pp. 230 – 235, May 2003.

[9] 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, pp. 611 – 617, January 2000.

[2]CNES: Centre National d'Etudes spatiales.

[3]ONERA: Office National d'Etudes et de Recherches Aérospatiales.