IMAGE PROCESSING FOR HIGH RESOLUTION SATELLITE AND AERIAL DATA

Josiane Zerubia

Ariana (CNRS/INRIA/UNSA joint research group)
INRIA, BP 93, 2004 route des Lucioles
06902 Sophia Antipolis Cedex, France
Tel: +33 4 92 38 78 65 Fax: +33 4 92 38 76 43
Email: Josiane.Zerubia@sophia.inria.fr
Web site: http://www.inria.fr/ariana/

ABSTRACT

The last few years have seen the launch of several high-resolution satellites, and many more are planned in the near future. In parallel, new high-resolution CCD cameras for use in aerial photography have been developed. In this paper, we focus on this new optical high-resolution data. We present the challenges arising from the new technology in the introduction, and then concentrate on three applications to SPOT5 and Pleiades HR simulations and aerial images: deconvolution, urban area extraction, and road network detection. In conclusion, we comment on future prospects.

1. INTRODUCTION

High-resolution data is becoming increasingly available in remote sensing applications. In this paper, we will concentrate on civilian optical sensors on board satellites or airborne vehicles.

For any application, the important specifications are the type of sensor (panchromatic or multispectral), its resolution, quantization and swath width, and the type of the available stereovision system if any. Table 1 lists this information (obtained from the web sites of the manufacturers or from specialized workshops) for a few high-resolution satellites. Why are these specifications so useful and what are the new challenges associated with them? First, the spectrum that characterizes the sensor is very important for classification purposes: working on multispectral data in the visible, near infrared or medium infrared domain could enable the detection of different types of vegetation for instance. One image-processing problem related to the spectrum is the fusion of the information provided by the different bands. The second specification is the resolution of the sensor: increasing resolution provides new opportunities for object detection and finer classification, which are useful in urban

planning for example. Several image-processing problems are directly connected to HR data: the detection of oriented textures while preserving homogeneous regions, for example, is a real challenge. Another image-processing problem is created by very HR data, namely taking into account not only the texture but also the geometry of each detected object. A third important specification is the swath width of the sensor (related to the size of the region represented in a single image): usually, higher resolutions imply smaller swath widths. But this will no longer be true, for instance, following the introduction of the Pleiades HR satellite in the near future. In that case, an image-processing problem is created by the sheer size of the data, in particular how to store the data efficiently and retrieve relevant information. The image-processing task is thus embedded in a larger problem known as information mining. A fourth useful specification is the quantization (i.e. the number of bits used to code the image). More and more HR data is coded using between 10 and 12 bits rather than 8. A direct consequence is that images possess a larger number of gray levels, which enables better image processing (better analysis of shadows, for instance) but poses a problem with respect to data storage capacity and data compression. The last important specification is the stereovision capacity, whether performed in track or cross track. One image-processing task related to stereo is 3D reconstruction. The aim is to recover a high-resolution Digital Elevation Model (DEM), which is necessary for mobile telecommunication simulation for instance.

Hereafter, due to space limitations, we restrict ourselves to three image-processing tasks, namely image deconvolution, urban area extraction and road network detection. For each task, we have tested the proposed method on SPOT5 and Pleiades HR simulations (real SPOT5 data will be shown during the conference if available) and/or on aerial images. A detailed description of SPOT5 satellite characteristics, with many examples of

Satellite	Country	Launch Date	Specifications				
			Sensors	Resolution	Quantization	Swath Width	Stereo
Ikonos	U.S.A.	24/9/99	PAN	1m	11 bit	11km	In & cross track ±45°
			MS (4 bands)	4m	11 bit	11km	
Eros A	Israel	5/12/00	PAN	1.8m	11 bit	12.5km	In track ±45°
Quickbird	U.S.A.	18/10/01	PAN	0.61m	11 bit	16.5km	In track ±38° Cross track ±30°
			MS (4 bands)	2.44m	11 bit	16.5km	
Spot 5	France, Belgium, Sweden	3/5/02	PAN	2.5m – 5m	11 bit	60km	In track ±20° Cross track ±27°
			MS (4 bands)	10m	11 bit	60km	
Orbview 3	U.S.A.	Planned 2002	PAN	1m	11 bit	8km	In & cross track ±45°
			MS (4 bands)	4m	11 bit	8km	
Cartosat (IRS-P5)	India	Planned 2003/4	PAN	2.5m	10 bit	30km	Yes +26° -5°
			PAN (bis)	2.5m	10 bit	27km	
Eros B	Israel	Planned 2003	PAN	0.85m	8 bit	16km	In track ±45°
Pleiades HR	France, Italy	Planned 2006	PAN	0.7 – 1m	10 – 12 bit	20 – 40km	In & cross track ±30°
			MS (4 bands)	2.8 – 4m	10 – 12 bit	20 – 40km	

Table 1

possible applications, can be found in [1], while [33] contains an interesting description of the CCD cameras developed by IGN.

2. EXAMPLE 1: IMAGE DECONVOLUTION

The deconvolution of blurred and noisy satellite images is an ill-posed problem. Direct inversion leads to unacceptable noise amplification. Usually, the problem is either regularized during the inversion process, or the noise is filtered after deconvolution and decomposition in the wavelet domain. In the case of high-resolution data, the problem is worse, because we wish to preserve sharp edges and have good quality reconstructions of both homogeneous regions and areas of oriented texture.

Many methods have been proposed for regularizing the deconvolution problem by introducing a priori constraints on the solution [6,7,9,17]. However, most of them do not preserve textures. To achieve a better deconvolution, [12,13,25,26,30,31] have proposed variants on the second type of solution (i.e. denoising after deconvolution without regularization).

We have proposed a method based on Complex Wavelet Packets (CWP) [23,24], which possess better translation invariance than standard wavelets, and have more directional sensitivity. The coefficients of the CWP transform are thresholded fully automatically within a Bayesian framework. The results obtained on both SPOT5 and Pleiades HR simulations, as well as on real aerial images exhibit correctly restored textures and a high Signal to Noise Ratio (SNR) in homogeneous areas. Compared to concurrent algorithms [22], the proposed method is faster, rotationally invariant, and better restores textures by taking into account their orientation. The deconvolved images can be used as they are, or can provide a starting point for an adaptive regularization method in order to obtain sharper edges [22,23]. An example is given below: Fig. 1 represents the original image (SPOT 5 simulation), Fig. 2 corresponds to the blurred and noisy data, and Fig. 3 shows the result obtained by the proposed method.

3. EXAMPLE 2: URBAN AREA EXTRACTION

Various applications require the extraction and the analysis of urban areas from HR satellite data. Such results are needed for land-use classification in order to draw thematic maps, as well as for the study of urban pressure on forests and agriculture. Urban area detection is also used to follow urban expansion in developing countries. Military intelligence also requires this type of information. There are many more examples.

Fig. 1

Fig. 2

Fig. 3

The methods used to discriminate urban areas are often based on texture analysis. In [19], Gouinaud provides a comparison of many different approaches. Hereafter, we will mention just a few methods. In [20], Haralick derives features (e.g. energy, entropy, and correlation) from Grey Level Co-occurrence Matrices (GLCM), which are constructed from the image by estimating the pairwise statistics of pixel intensity. He applies this

method to the segmentation of LANDSAT images. However, in [3], Baraldi and Parmiggiani underline the inadequacy of these features for the analysis of urban scenes in SPOT images. In [8], Conners, Trivedi and Harlow define new features derived from GLCM but better adapted to urban scene segmentation. Oriented filter banks such as Gabor filters [15] have also been used for this purpose. In [21], Houzelle and Giraudon use radar images to select patches from which to estimate the texture parameters of urban areas in SPOT images. This method supposes that pairs of images (optical and radar) are available for the same urban area. In [10], Descombes and Prêteux use a texture parameter called temperature to extract urban masks. This method gives good results on SPOT3 images but not on SPOT5 simulations. In particular, the geometric shape of the detected urban areas is not correct due to oriented textures such as orchards and vineyards. In [36], Winter, Maître, Cambou and Legrand develop a multi-scale method based on the observation that an optimal detection scale exists for each object. This method uses entropic measures defined on wavelet pyramids. The results on SPOT5 simulations do not localize the boundary of the urban area accurately.

The method we have proposed in [27,28] is based on eight different 1D Markov Chain models as shown in Fig. 4. Instead of considering a classical isotropic neighborhood as usual for a Markov random field (with 4 or 8 neighbors), we just define in the direction d the

neighborhood of pixel s to be the set of the two nearest pixels to s in that direction. Then, the conditional probability for the model defined in the direction d depends on the mean m^d_s of the two neighbors of s in direction d. It can be shown that the key texture features are the local conditional variances σ_d^2, which can be estimated easily for each of the eight chain-based models. As the lattice presented in Fig. 4 is anisotropic, a normalization of the estimated parameters is necessary in order to correct the bias introduce by the anisotropy. A renormalization group approach [18] is used (renormalization by decimation). Then we classify the eight renormalized local conditional variances in increasing order and keep only the mean of the two central values. It can be shown [28] that this parameter characterizes urban areas while excluding water, forest, and fields (in particular orchards and vineyards), as well as green-houses for instance [27]. Next, a classification is obtained using a fuzzy C-means algorithm operating on the original objective function plus an entropy term. Finally, we do some post-processing: first, the result is regularized using a Potts model, then false alarms are removed using the minimum of the local conditional variance. Many tests have been conducted on SPOT5 simulations as well as on aerial images, and are described in [27]. It can be shown that the proposed technique gives better results than previously existing methods (in particular, comparisons have been made with Gabor filters and co-occurrence matrices [28]). An example is shown in Fig. 5 (SPOT 5 simulation). This method works well for a resolution between 20m [SPOT3] and 2.5m [SPOT5 Supermode]. Below 2m of resolution, the results are no longer that satisfactory due to the importance of building shape. A more appropriate method could be developed based on marked point processes for instance [11].
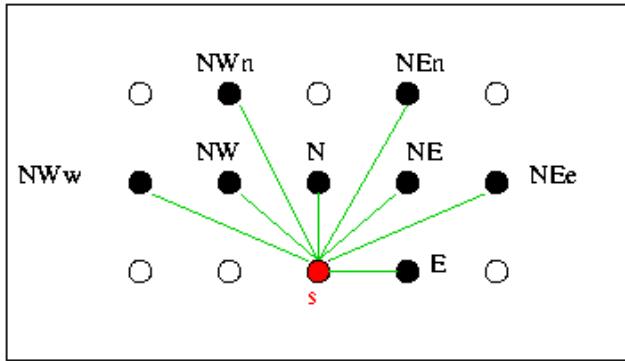

Fig. 4
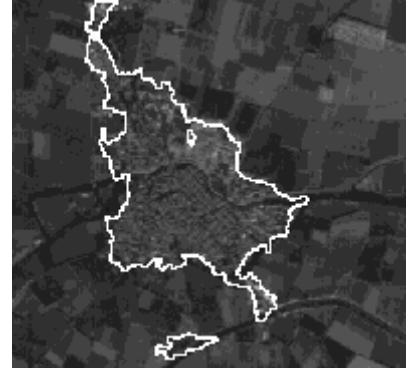
Fig. 5

4. EXAMPLE 3: ROAD NETWORK DETECTION

In this section, we focus on the automatic extraction of linear features such as road networks from HR satellite or aerial images. This is an open problem of great interest in image processing [5,14,16,29], but all these methods require starting points (usually chosen by hand).

Furthermore, most methods are pixel-oriented. Pixel-oriented approaches are quite sensitive to noise, local minima, and false alarms generated by similarity to other objects in the image. F. Tupin *et al*. in [35] adopt an object-based approach to the solution of this problem, using a random field on a graph of segments, and show the necessity for a global model of these features. The difficulty lies in the definition and extraction of the graph, which should contain the road network.

A new direction in image processing that overcomes the drawbacks of these pixel-oriented techniques is to analyze the objects and their interactions in the image directly, using stochastic geometry [2,4] and point processes.

In mathematical terms, practical and flexible line segment processes are required. Unfortunately, until recently, the only model that had been studied was the Poisson line segment process [32]. This model does not take into account the dependencies that exist in road networks for example.

To extract road networks using an object-based approach [34], we make the hypothesis that a road network is formed by several connected segments. In fact the segments represent ribbons, because roads have a strictly positive width. The segments have a center, an orientation, a width and a length (all these quantities are random variables). We make the hypothesis that a line network is the realization of a marked point process. The probabilistic model of this point process has two components. The first component is the interaction model,

which describes the interactions between segments: connection, attraction, repulsion, orientation, and the dimension of the line network. The second component is the data model, which describes the location of the different segments in the image. Thus, the problem of road network detection in an image can be stated as a marked point process [11]. We define the image of a segment by the geometric figure given by the label of the segment (i.e. the length, width and orientation) and its position (i.e. the center).

Using the marked point process framework, segments can appear or disappear during the optimization, and their locations can vary. We do not need to initialize a segment graph.

To each segment are associated two attractive areas around the extremities and a repulsive area around the center of the segment. If a segment intersects the attractive area of another segment, we have an attractive interaction (favored configuration) with an increasing intensity as the orientation difference between the two segments decreases. If a segment intersects the repulsive area of another segment, we have a repulsive interaction (penalized configuration) with a decreasing intensity as the orientation difference between the two segments tends to $\pi/2$. Some penalizing constants are added for segments having one or two unconnected extremities. These three terms allow us to control the local curvature of the network, the number and shape of the crossings, and the connectivity of the network respectively. To compute the data term, we define the silhouette of a segment as its projection onto the image lattice. We then define a mask consisting of three parts: the segment silhouette, pixels to the left of the segment and pixels to the right of the segment. The likelihood, under a Gaussian assumption, of the three hypotheses: "the mask corresponds to one, two or three different areas" is used to define the external field. To optimize the model, we consider a reversible jump Markov Chain Monte Carlo (MCMC) algorithm embedded into a simulated annealing scheme. The definition of the transition kernel is crucial for obtaining reasonable convergence rates. We have to define transitions that are consistent with the constraints of the model. For the road extraction model, we first have a birth and death kernel. These transitions correspond to adding or removing a segment with uniform probability. A birth or a death is also allowed within the neighborhoods of the segments in the current configuration. This transition allows the prolongation of the network. We also have some local perturbations of one segment (translation, rotation). This model has been tested on aerial images and on SPOT5 simulations. A result (Fig. 7) is shown for an aerial image (Fig. 6).

Fig. 6
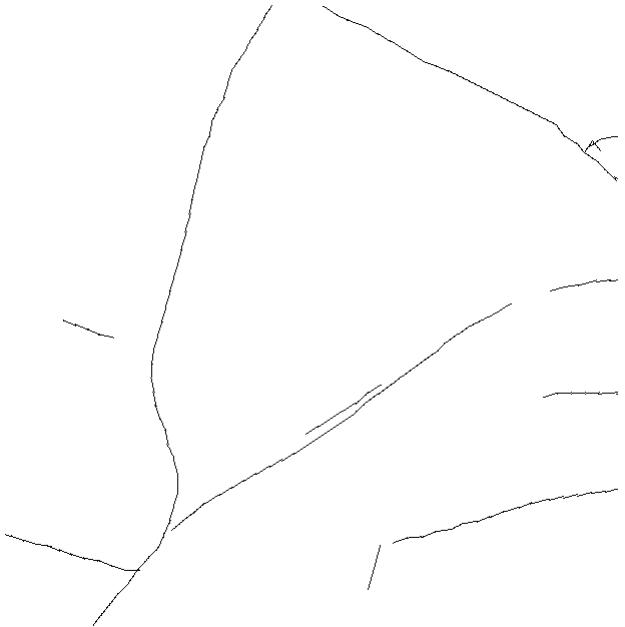
Fig. 7

5. CONCLUSION

In this paper, we have presented some challenges arising in the processing of high-resolution satellite and aerial images. We have briefly described three methods based on various mathematical models (complex wavelet packets, Markov random fields and marked point processes) in order to achieve, respectively, better image

deconvolution, more precise urban area extraction, and automatic road network detection.

Due to the increase in the resolution and swath width of the new sensors, as well as the number of images available, one of the main problems in the near future will be fast and efficient information retrieval.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] "SPOT5: Towards new applications," special issue (in French)-*bulletin SFPT* N°164-165, 2001.

[2] A.J. Baddeley and M.N.M. van Lieshout, "Stochastic geometry models in high-level vision", In *Statistics and Images, Vol. 1, K.V. Mardia and G.K. Kanji (Eds.) Advances in Applied Statistics, a supplement to Journal of Applied Statistics*, Vol. 20, pp. 231-256, Abingdon, Carfax, 1993.

[3] A. Baraldi and F. Parmiggiani, "Urban area classification by multispectral spot images. *IEEE Trans. on GRS*, 28(74):674-680, 1990.

[4] O.E. Barndoff-Nielsen, W.S. Kendall and M.S.M. van Lieshout, editors, "Stochastic geometry, likelihood and computation", *Chapmann and Hall*, London, 1999.

[5] M. Barzohar and D.B. Cooper, "Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation", *IEEE Trans. on PAMI*, Vol. 18, pp. 707-721, 1996.

[6] B. Chalmond, « Image restoration using an estimated Markov model *», Signal Processing*, 15/115-129, 1988.

[7] P. Charbonnier, L. Blanc-Feraud, G. Aubert and M. Barlaud, « Deterministic Edge-Preserving Regularization in Computed Imaging *», IEEE Trans. on IP*, 6(2):298-311, 1997.

[8] R.W. Conners, M.M. Trivedi and C.A. Harlow, "Segmentation of high-resolution urban scene using texture operators", *CVGIP*, 25:273-310, 1984.

[9] G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems", *IEEE Trans. on ASSP*, 37(12)/2024-2036, 1989.

[10] X. Descombes and F. Prêteux, "Topology and parameter estimation in MRF modeling", *In SPIE, Neural and Stochastic Methods in Image and Signal Processing II*, Vol. 2032, pp. 156-166, San Diego, USA , 1993.

[11] X. Descombes and J. Zerubia, "Marked point processes in image analysis*", special issue of IEEE SP Magazine*, to appear in 2002.

[12] D. Donoho, "Denoising by soft thresholding", *IEEE Trans. on IT*, 41:613-627, 1995.

[13] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage", *Biometrika*, 81:425-455, 1994.

[14] M.A. Fischler, J.M. Tenenbaum and H.C. Wolf, "Detection of roads and linear structures in low-resoltuion aerial imagery using a multisource knowledge integration technique", *Computer Graphics and Image Processing*, Vol. 15, pp. 201-223, 1981.

[15] I. Fogel and D. Sagi, "Gabor filters as texture discriminator", *Biological Cybernetics*, 61:103-113, 1989.

[16] D. Geman and B. Jedynak, "An active testing model for tracking roads in satellite images", *IEEE Trans. on PAMI*, Vol. 18, pp. 1-14, 1996.

[17] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuity", *IEEE Trans. on PAMI*, 14(3):367-383, 1992.

[18] B. Gidas, "A renormalization group approach to image processing problems", *IEEE Trans. on PAMI*, 11(2):164-180, 1989.

[19] C. Gouinaud, "Traitement d'images satellitaires pour la detection d'agglomérations", *PhD thesis, Ecole Nationale Supérieure des Télécommunications, Telecom Paris 96E035*, France, 1996.

[20] R.M. Haralick, « Statistical and structural approaches to texture *», Proc. IEEE*, 67(5):786-804, 1979.

[21] S. Houzelle and G. Giraudon, "Data fusion using SPOT and SAR images for bridge and urban area extraction", In *IGARSS*, Vol. 3, pp. 1455-1458, Helsinki, Finland, 1991.

[22] A. Jalobeanu, "Modèles, estimation bayésienne et algorithmes pour la déconvolution d'images satellitaires et aériennes", *PhD thesis, Université de Nice-Sophia Antipolis*, France, 2001.

[23] A. Jalobeanu, L. Blanc-Féraud and J. Zerubia, "Satellite Image Deconvolution Using Complex Wavelet Packets", *INRIA Research Report, no 3955*, 2000.

[24] A. Jalobeanu, L. Blanc-Féraud and J. Zerubia, " Satellite image deconvolution using complex wavelet packets", *Proc. ICIP*, Vancouver, Canada, 2000.

[25] J. Kalifa, "Restauration minimax et déconvolution dans une base d'ondelettes miroirs", *PhD thesis, Ecole Polytechnique*, France, 1999.

[26] J. Kalifa and S. Mallat, "Bayesian inference in wavelet based methods", chapter Minimax restoration and deconvolution, *Springer*, 1999.

[27] A. Lorette, "Analyse de texture par méthodes markoviennes et par morphologie mathématique: application à l'analyse des zones urbaines sur des images satellitaires", *PhD thesis, Université de Nice-Sophia Antipolis*, France, 1999.

[28] A. Lorette, X. Descombes and J. Zerubia, "Texture analysis through a Markovian modeling and fuzzy classification : application to urban area extraction from satellite images », *IJCV* 36(3), 221-236, 2000.

[29] N. Merlet and J. Zerubia, "New propects in line detection by dynamic programming", *IEEE Trans. on PAMI*, Vol. 18, pp. 426-431, 1996.

[30] B. Rougé, "Théorie de la chaîne image et restauration d'image optique à bruit final fixé », *Mémoire d'habilitation à diriger des recherches*, France, 1997.

[31] B. Rougé, "Fixed chosen noise restoration", *IEEE*, Philadelphia, USA, 1995.

[32] D. Stoyan, W.S. Kendall and J. Mecke, "Stochastic geometry and its applications", *John Willey and Sons*, 1987.

[33] C. Tom and J.P. Souchon, "Le point sur les cameras numériques de l'IGN", *bulletin SFPT* N° 149, pp.12-20, 1998.

[34] R. Stoica, X. Descombes and J. Zerubia, "Road extraction in remote sensed images using a stochastic geometry framework", Proc. MaxEnt, Gif sur Yvette, France, 2000.

[35] 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 Trans. on GRS*, Vol. 36, no. 2, 1998.

[36] A. Winter, H. Maître, N. Cambou and E. Legrand, « An original multi-sensor approach to scale-based image analysis for aerial and satellite images", In *ICIP*, Vol. 2, pp. 234-237, Santa Barbara, USA, 1997