

# COMBINING SAR AND OPTICAL FEATURES IN A SVM CLASSIFIER FOR MAN-MADE STRUCTURES DETECTION.

*Gabrielle Lehureau<sup>1</sup>, Marine Campedel<sup>1</sup>, Florence Tupin<sup>1</sup>, Celine Tison<sup>2</sup>, Guillaume Oller<sup>3</sup>.*

<sup>1</sup>Institut TELECOM, TELECOM ParisTech, CNRS LTCI, 46 rue Barrault, 75013 Paris, France.

<sup>2</sup>CNES, 18 avenue Edouard Belin, 31401 Toulouse, France.

<sup>3</sup>Magellium, 24 rue Hermes, 31521 Ramonville Saint Agne, France.

## ABSTRACT

The increasing quality of satellite images has generated interests in extracting man-made structures in urban areas, such as buildings and roads. A classification adapted to urban areas can help to identify these structures. In this paper, SAR information are used to improve land-cover classification. We proposed a classification process using both radar and optical data, a segmentation and a classification with Support Vector Machines (SVM).

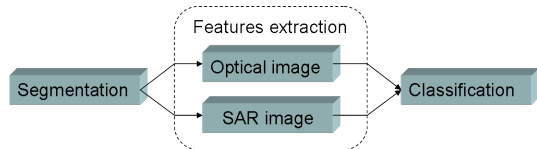
**Index Terms**— Classification, High-Resolution imagery, SAR, Optical imagery, urban areas, SVM.

## 1. INTRODUCTION

Among the Earth Observation sensors, optical and radar ones deliver images with different kind of information. Therefore, it is interesting to fuse such data. For this purpose, a registration process is required. But because of the particular geometric distortions in radar images, this step may be difficult. Indeed, depending on the height, some objects can be superposed in the SAR image with other. In order to improve the registration and help to interpret the data, we want to introduce a supervised classification of the optical data. It can help to identify objects above-ground and so, to calculate the real transformation between both images.

In high or very high resolution images, spectral characteristics of urban areas such as parking, grass field or building roofs, are so similar that they are often not sufficient to separate the different classes. For now, classification accuracy of panchromatic metric resolution data is not fully satisfactory. Spectral characteristics, geometrical information and multi spectral bands are usually used in classification. It may be useful to introduce additional information into the classifier. Some methods introduced multiband images or digital elevation model[1] [2]. In this paper we study the interest of SAR data combined with optical images.

We propose an object-based method (Figure 1): first a segmentation of the optical image of urban areas is performed; second, textural and geometrical characteristics are extracted from each segment; then, statistical information of



**Fig. 1. Classification process.**

SAR images are incorporated to the classifier; finally a supervised classification is performed. The model is evaluated using Support Vector Machines in a cross-validation process.

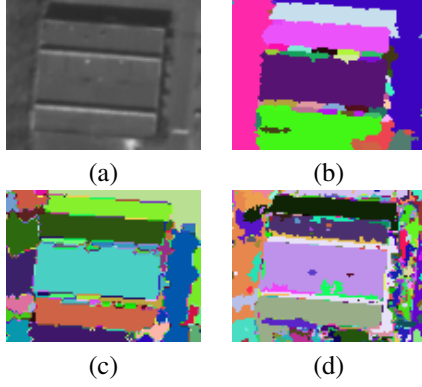
## 2. CLASSIFICATION PROCESS

### 2.1. Registration

The first step to classify the data is to register them. There is different way for automatic SAR/optical registration. Some use features [3] or pixel intensity [4], others use sensors parameters. In pre-processing we used the method proposed in [5], which does not require any sensor parameters. First a rigid deformation is computed based on Fourier-Mellin invariant. Second, the registration is refined using a polynomial transformation. Some particular points are extracted in optical image and their corresponding are determined in the SAR by mutual information. Finally we obtain a registration with an accuracy around ten pixels.

### 2.2. Segmentation

The proposed framework relies on a preliminary segmentation of the data to extract shape characteristics of the segments and textural features at object scale. Due to the specific radiometry of SAR imagery, the segmentation is performed on the optical image. Different segmentation methods were tested : Watershed [6], Guigues segmentation [7] and the segmentation tool presented in [8] by Luo, which is based on Fast Level Set Transform [9] for the decomposition and then regularized considering the noise present in optical images (Figure 2). These segmentations are then superposed to the SAR data, as if both images are well registered.



**Fig. 2. Example of building segmentation:** (a) Optical image, (b) Watershed, (c) Guigues segmentation and (d) Luo segmentation.

Most of the difficulty when using Watershed or Guigues segmentation is to determine the good parameters (the flood level or scale parameters) for the segmentation. The segmentation based on Fast Level Set Transform permits to overcome the problem, since it does not require any parameters. Figure 2 represents an exemple of segmentation obtained by each of the three methods. Influence of the segmentation is discussed in section 3.1.

### 2.3. Feature Extraction

Feature extraction can be seen as finding a set of vectors that represents a good observation of each target objects and helps to discriminate their classes. For this purpose, textural and geometrical characteristics are used. Gabor wavelets [10] are used to extract textural characteristics, with 4 orientations and 3 dimensions. For each object, variance and mean of the filtered image are computed; this will constitute the  $4 \times 3 \times 2 = 24$  textural features. Other characteristics are extracted, like mean and variance of the original image.

Textural characteristics are not sufficient to discriminate the different classes in high resolution panchromatic data. Thus, some geometrical features of each object have been computed according to their shape. For this purpose, the perimeter, surface and a term of scale (scale = surface/perimeter) of each objects have been computed.

In this paper we have investigated how the introduction of SAR data can improve the classification process. Some specific features like bright scatterers are visible on buildings, whereas roads appear dark; thus these images are particularly relevant in order to discriminate the classes of urban areas. In [11] it is shown that the 3 first log-cumulants are very useful to capture discriminative information. Log-cumulants on the SAR image over the segmented objects provide the last feature set. Moreover these parameters are really easy to compute: the log-cumulants  $k_n$  are function of the log-moments

$m_n$ , where  $m_n$  is defined as follow.

$$\begin{aligned} m_n &= \int_0^\infty (\log u)^n p_x(u) du \\ k_1 &= m_1 \\ k_2 &= m_2 - m_1^2 \\ k_3 &= m_3 - 3m_1m_2 + 2m_1^3 \end{aligned}$$

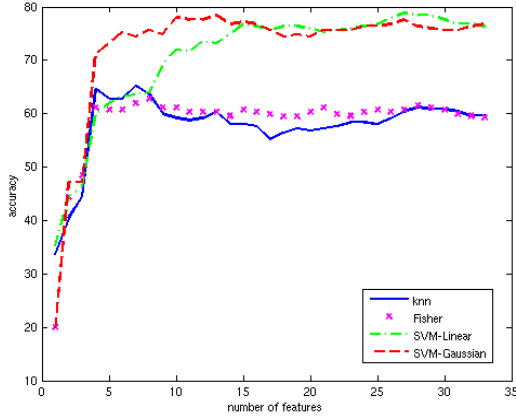
### 2.4. Choice of the classifier

The third step of the classification process is the choice of the classifier. Several urban classification methods have been proposed in the remote sensing literature. Classification algorithms such as maximum likelihood [12], neural network [1], fuzzy models [13] or Support Vector Machines (SVM) were frequently used for the classification. All these methods have their own characteristics and advantages. For the purpose of supervised classification of optical and SAR data, one of the main difficulties of statistical methods is the small ratio between the number of training samples and the number of features. The less training samples we need, the better it is for the user. But it is proved that a low number of training samples makes it impossible to obtain a good estimation of the probability density function [14]. In consequence, we choose to use a classifier based on “geometrical” criterion instead of “statistical” criterion in this process. Several methods were compared regarding their classification result and the number of features required.

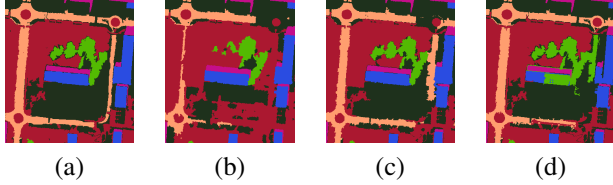
The first method is the k-nearest-neighbor [15] which is a reference in pattern recognition. It is based on counting the number of nearest neighbor in the training set. The second is the Fisher linear discriminant [16], the SVM was tested using a linear kernel (SVM-Linear) or a Gaussian kernel. The features extracted before are sorted according to their relevance. Feature selection and classification was computed with libsvm in the spider environment. Features were selected with a method based on Fisher discriminant. Figure 3 plots the accuracy of each classifier versus the number of features. It can be seen, that SVM gives better result than the other methods (76% instead of 59%). The ten first features are really significative and usefull for the classification. Amongst them, are the log-cumulants from the SAR image, that show the importance of both data. Moreover, when using a non-linear kernel, SVM gives better result. So for the classification process, SVM with Gaussian kernel were chosen.

## 3. EXPERIMENTAL RESULTS

The proposed classification process has been tested on a couple of Quickbird and Ramses images with size 2048x2048 and a resolution around 0.7m, the study area lies in Toulouse (France). The classifier returns five classes corresponding to urban areas: buildings, high vegetation, ground, roads and shadows.



**Fig. 3.** Classification result when using 1 to 33 features. In red: the SVM with Gaussian kernel, green: SVM-Linear, pink: Fisher and blue: 3-nearest-neighbor.



**Fig. 4.** Projection exemples of the ground truth depending on the segmentation: (a) original, (b) Watershed, (c) Guigues and (d) Luo segmentation.

In order to have a quantitative evaluation, ground truth was manually generated. Though, because of the object approach, it is necessary to project the ground truth on the segmentation. We assign the majority pixel class to each segment. Figure 4 shows the projected ground truth. 50 segments of each class were used as training set. The parameters of the SVM were evaluated through cross-validation on the training set with all the features proposed in 2.3.

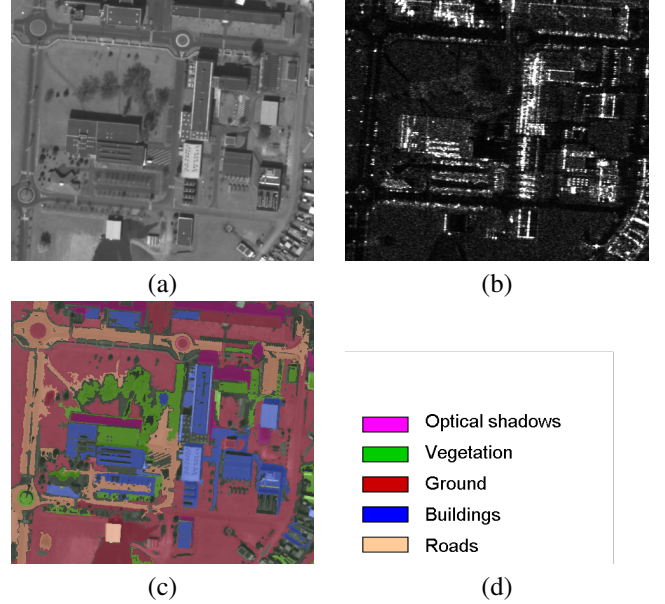
### 3.1. Robustness to the segmentation

In order to assess the robustness of the classifier to the segmentation, each of them have been tested. For each segmentation, the classification has been trained five times with different initialisation for the training. Table 1 shows the mean percentage of good classification considering the number of segments or the surface. Depending on the segmentation, a scene can be divided into a variable number of segments. Thus it is important to jointly consider the accuracy of the segments versus the surface of the scene.

As it can be seen, all segmentation algorithms have equivalent performances, but Luo's method is higher when considering the well classified surface. However when considering the segments, Luos method has less accuracy. Table 1 show

	Watershed	Guigues	Luo
segments	76.1%	79.23%	74.42%
surface	80.2%	81.44%	82.17%

**Table 1.** Percent of good classification with each segmentation method. The results are given considering the number of segments or the number of pixels well classified.



**Fig. 5.** Classification result on Toulouse (France): (a) Quickbird image, (b) Ramses image, (c) classified image superposed to the optical one.

that none of these segmentation methods is better than the other, they all seem to have the same accuracy. That proves a robustness of the classifier to the segmentation. Considering that, in the following we use Luo segmentation. The main benefit of this method is that there is no parameters to determine.

### 3.2. Feature complementarity

Table 2 shows the mean classification recognition rates obtained with features extracted from each and both data. It can be seen on one hand that the optical image provides good information to identify the shadows, ground and high vegetation. But the textural and geometrical features are not sufficient to recognize roads. On the other hand, the SAR image helps at identifying buildings and roads, even if the images are not exactly registered. Using the information contained in both data improves the classification performance. Figure 5 show the registered Quickbird image, the Ramses image and the classification result. Because of the number of segments (around 200000), only the largest were classified.

	buildings	roads	high vegetation	ground	shadow
optical only (text. + geom.)	66%	0%	75%	76%	85%
radar only	76%	61%	64%	24%	0%
both features	90%	85%	94%	81%	90%

**Table 2.** Mean classification recognition rates obtained on the urban area of Toulouse, considering five classes and features extracted from optical image only, radar image only and from both data.

#### 4. CONCLUSION

Recent research has shown the interest of incorporating other information for classification of urban areas in panchromatic image. Usually, multi-spectral data and geometrical characteristics are used. In this paper, we proposed a supervised classification process using an optical based segmentation, SAR and optical information and non-linear SVM. The SAR information leads to significant improvement in classification when conjugated with geometrical and optical textural features even if the data are not exactly registered. For now, the accuracy of the classification is between 75 to 80%, but this result can be improved with a pixel-based registration. As future work, we propose to evaluate the class and the position of each segment in the same time through a classification process.

#### 5. ACKNOWLEDGMENTS

The authors are grateful to ONERA and DGA for providing the data.

#### 6. REFERENCES

- [1] J. A. Benediktsson, M. Pesaresi, and K. Arnason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, pp. 1940–1949, September 2003.
- [2] C. Baillard and H. Maitre, "3-D reconstruction of urban scenes from aerial stereo imagery: A focusing strategy," *Computer Vision and Image Understanding*, vol. 76, pp. 244–258, December 1999.
- [3] M.A. Ali and D.A. Clausi, "Automatic registration of SAR and visible band remote sensing images," 2002, vol. 3, pp. 1331–1333 vol.3.
- [4] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 10, pp. 2104–2120, Oct. 2004.
- [5] G. Lehureau, F. Tupin, C. Tison, G. Oller, and D. Petit, "Registration of metric resolution SAR and optical images in urban areas," in *Proc. 7th European Conference on Synthetic Aperture Radar (EUSAR)*, 2008.
- [6] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," in *International Workshop on image processing, real-time edge and motion detection/estimation*, 1979, pp. 17–21.
- [7] L. Guigues, *Modèles multi-echelles pour la segmentation d'images*, Ph.D. thesis, Ecole doctorale Sciences et Ingénierie de l'Université de Cergy-Pontoise, Cergy-Pontoise, 2004.
- [8] B. Luo, *Echelle et Resolution en Imagerie de Teledetection*, Ph.D. thesis, TELECOM ParisTech, 46 rue Barrault, 75013 Paris, France, December 2007.
- [9] P. Monasse, "Fast computation of a contrast invariant image representation," *IEEE Trans. on Image Proc.*, vol. 9, pp. 860–872, 1998.
- [10] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, pp. 837–842, 1996.
- [11] J.-M. Nicolas, C. Tison, and F. Tupin, "Comparison of beta1-beta2 diagram and logcumulant2-logcumulant3 diagram for SAR image texture characterisation," in *EUSAR 2004*, Ulm (Allemagne), June 2004.
- [12] M. Lienou, H. Maitre, and M. Datcu, "Is it possible to automatically produce a Corine Land Cover map from a single SPOT image?," in *Fourth conference on Image information mining, ESA-EUSC*, 2006.
- [13] J. Chanussot, J.A. Benediktsson, and M. Fauvel, "Classification of remote sensing images from urban areas using a fuzzy possibilistic model," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 1, pp. 40–44, 2006.
- [14] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, 1968.
- [15] E. Fix and J.L. Hodges Jr, "Discriminatory analysis, non-parametric classifications," *USAF Sch. Aviat. Med., Rep.*, vol. 4, 1951.
- [16] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning," 2001, Springer Series in Statistics.