New Developments and Challenges in Remote Sensing, Z. Bochenek (ed.)


# Satellite SAR and urban remote sensing: status and perspectives

F. Dell′Acqua, P. Gamba & G. Trianni
*Remote Sensing Group, University of Pavia, Pavia, Italy*

Keywords: urban remote sensing, Synthetic Aperture Radar (SAR), Markov Random Field (MRF), texture

ABSTRACT: This paper aims at providing a few ideas about urban remote sensing and satellite SAR data. Instead of looking for future systems and providing hints about high resolution SAR, this paper shows that even with current satellite SAR data it is possible to obtain interesting results that stretch their usefulness beyond usual applications. In doing so, this paper briefly reviews the status of research in SAR urban remote sensing and highlights some of the perspectives still open for research.

## 1 INTRODUCTION

Urban environment is by far the most complex one that may appear in remotely sensed images, and its analysis requires to extract a wealth of information from the sensed data. The sensor that conveys the greatest amount of information about structural as well as dielectric properties of the urban materials is the synthetic aperture radar (SAR), whose ground resolution for spaceborne sensors is currently of a few meters. Additional advantages of satellite SAR remote sensing include all-weather capabilities, repeatability of the acquisition and, more recently, also polarimetric capabilities.

There is clearly a deep interest in the analysis of radar images over urban areas, because these seem to be the best candidate for structural characterization of the cities and towns, at the current available resolution. An interesting feature of the forthcoming satellite SAR instruments, however, will be the significantly improved spatial resolution, down to tens of centimeters. Low earth orbit (LEO) constellations of satellites like Cosmo/Skymed [1] or single satellites like TerraSAR-X [2] are expected to start delivering high-resolution SAR data at competitive prices; such data is naturally appealing where the scope of the analysis is the complex urban environment. In these times, the research on radar sensing over urban areas is undergoing a transition from a research aiming at exploiting the current resolution SAR data at its best towards a preparatory effort in sight of the new perspectives and troubles that higher-resolution data will open; only to mention some, single structural elements will become visible, while on the other hand traditional speckle noise models will not apply anymore due to a limited amount of scatterers in a single resolution cell, which implies a dramatically narrower statistical base for the statistical models [3].

The current situation allows the analysis of an urban environment from satellite data only at a rather coarse level. Indeed, studies have been made on the extraction of city limits using texture information [4] and different land use classes by means of a more adaptive

approach [5]. Moreover, until the most recent satellites (ENVISAT and RADARSAT-1) there was no flexibility on the radar viewing angle, so only a few analyses considering the variability of this parameter and its influence on urban area classification results were carried out [6, 7]. More extensive research was dedicated to data fusion between SAR and optical data for better classification [8] and also for feature enhancement [9].

In the following, we will try and sketch an outline of the current situation and perspectives projected from the current, dynamic scenario. SAR urban change detection for disaster monitoring [10] and SAR informal settlement mapping for water management in rural areas of developing countries [11] are two among the most stressed future applications of this line of research.

## 2 SAR DATA MAPPING IN URBAN AREAS

The current status of the use of SAR data for feature extraction, mapping and classification in urban areas is well described by the papers presented in the very recent EUSAR06 symposium, in two special sessions [12]. Most of the stress is on high resolution SAR data, possibly in conjunction with optical data, and the use of interferometric information for three-dimensional analysis of the urban environment.

Unfortunately, very fine spatial resolution is not within the reach of current satellite SAR sensors. Therefore, another line of research very active today is the use of available satellite SAR data for mapping the structure of urban areas. The huge amount of past SAR imagery is far from being completely analyzed, and it carries a wealth of information about the evolution of urban areas, still to be fully exploited. Moreover, enlarging the research view from urban areas -strictly speaking- to ‘‘human settlements’’ opens a whole new panorama of applications and possibilities. Usual counterexamples about these applications are related to the coarse spatial resolution of past SAR data and the unstructured character of most human settlements. Both problems deserve attention, but recent approaches have shown that, while they cannot be dismissed as ‘‘minor drawbacks’’, they can be reduced by means of more precise spatial pattern analysis of SAR data [13] and data fusion. In particular, it is worth stressing that in fusing SAR and optical data, for instance, very interesting results highlighting the above mentioned spatial patterns may be achieved [9], with a substantial increase in the ease of virtual interpretation and accuracy in land cover mapping.

## 3 URBAN AREA MAPPING BY SAR IN CHALLENGING ENVIRONMENTS

As outlined in the previous section, formal and informal settlement mapping using SAR data will stretch current methodologies for urban mapping, developed for industrialized countries and mostly European and U.S. areas, and an extensive validation is definitely required. As a first example, we offer in this section the evaluation of current state-of-the-art techniques for SAR mapping in two very different areas, namely Central Asia and Eastern Africa. A brief recall of the methodology is first offered, then the achieved results will be used for comments and for drawing perspectives in the conclusion section.

### 3.1 *Fixed and variable width textures*

When dealing with urban areas in different parts of the world, one point really interesting is that they have a very different structure, and this means, from a remote sensing point of view, that there is a strong need for a scale-dependent approach. Formal and informal settlement have different spatial scales, and the ability to somehow "guess" the best scale for mapping a given area is therefore one of the most challenging and interesting research themes in, generally speaking, urban remote sensing. However, SAR images are more related to geometric properties of the scene than any other sensor, and therefore, they carry more information about this than any other source of information by satellites.

Technically speaking, many methods have been proposed in technical literature to provide a measure for the spatial relationships among neighboring pixels, but the most widely successful is the co-occurrence texture analysis [14]. Co-occurrence texture measures are computed starting from the co-occurrence matrix. There are a few parameters driving this process, namely the modulus of the distance between the two positions we are jointly considering, its direction, and the width of the window used for computation. The distance provides an important way to discriminate among textures using the element spacing. Direction is especially important when anisotropy in the texture is present [15]. Finally, the window width is usually neglected.

Remote sensing images, however, reveal usually only very compact patterns. Inside an urban area, for instance, usual distances are around a few meters so that basic texture elements are in adjacent pixels (or even in the same one) for most satellite images. Moreover, although many environments have preferred directions, such anisotropy is often immaterial for very fine textures and/or small-scale texture patches. Sometimes no clear texture segment can be seen, and patterns are continuously changing. So, the co-occurrence window width remains the only really important parameter. The window width, indeed, defines the area around a pixel where we assume that texture patterns are statistically stable. In turn, this number needs to be bound to the mean physical dimension of the textured areas we are looking for. This explains why this parameter has been found as the most important one in urban remote sensing [16], and the only parameter that depends on the spatial scale of the image.

A methodology for textures and multiple scales has been proposed in the same paper and was labeled *multi-scale texture fusion*. The approach is based on a supervised neural classification, fed by a feature extraction step. This step exploits the same training set of the classifier, and is based on the computation of a discrimination index, the Histogram Distance Index (HDI) [17]. The whole set of textures computed from the co-occurrence matrix for different values of the parameters is considered, and the subset highest in HDI ranking is used as input to a Fuzzy ARTMAP multi-band classifier [18]. Adaptive Resonance Theory (ART) networks, basically introduced for solving pattern recognition problems, have indeed shown to be very efficient in multi-band remote sensing data analysis (see also [19]). This is particularly true when we deal with bands whose statistical properties are very different, as is the case with texture measures. ART networks store in their memories information about the training samples, and compare test patterns with these memories. Any match assigns the pattern to an output category, e.g. a land use class. We proved that there is a strong correlation among the HDI rank order and the progression of the overall accuracy values after classification. In particular,

we observed that this step can provide the minimum number of textures measures **and** scales that boost to its best the classification performance.

### 3.2 *Classification via Markov Random Fields*

Alternatively to the Fuzzy ARTMAP approach, a Markov Random Field (MRF) approach is also used in this work. MRF and fuzzy ARTMAP approaches share the same advantages, i.e. they can handle multiple images, even with very different statistics, and may be easily adapted to different inputs. This is the reason why MRF's are widely used in data fusion methodologies, where more data sets, coming for instance from radar and optical sensors, are used to classify the same scene [20]. The second fact in common between MRF and the above mentioned neuro-fuzzy classifier is their ability to consider spatial relationships among neighboring pixels in the classification framework. This is implicit in the MRF formulation, as we will see in a moment, but may be also introduced in the neuro-fuzzy classification chain through a refinement of the spectral classification based on a spatial re-analysis scheme, as already discussed in [18].

To briefly summarize the MRF framework, let us consider a set of features or images coming from $n$ sensors; then, let us consider the $M \times N$ pixel image acquired by sensor $r$ as made up of $M \cdot N$ pixels or feature vectors $X_r(1,1), \ldots, X_r(M,N), r = 1, 2, \ldots n$, where $X_r(i,j) = (x_r(i,j,1), \ldots x_r(i,j,B_r))$, and $B_r$ is the number of spectral bands or features for sensor $r$. We assume we know that $K$ classes $c_1, c_2, .., c_K$ are present in the images, and that they have prior probabilities $P(c_1), P(c_2), \ldots, P(c_K)$. Let us denote with *C(i, j)* the class for pixel *(i, j)*; we indicate with $X_r$ the set of the pixels of the whole image $X_r = \{X_r(i,j); 1 \leq i \leq M, 1 \leq j \leq N\}$ and with $C = \{C(i,j), 1 \leq i \leq M, 1 \leq \mathrm{j} \leq \mathrm{N}\}$ the set of labels for the same scene; in practice for a given pixel $(i,j), C(i,j) \in \{c_1, c_2, \ldots, c_K\}$.

If we indicate with $P(X_1, \ldots, X_n|C)$ the conditional probability density of feature vectors $X_1, X_2, \ldots, X_n$ given the scene labels set $C$, and with $P(C|X_1, \ldots, X_n)$ the posterior probabilities, the classification task consists of assigning each pixel to that class that maximizes the posterior probabilities. There is naturally a relation between the data (measurements or features) and the prior information, which can be represented in a Bayesian formulation as:

$$P(\mathrm{C}|X_1, \ldots, X_n) = \frac{P(X_1, \ldots, X_n|C)P(\mathrm{C})}{P(X_1, \ldots .X_n)} \tag{2}$$

where *P(C)* represents the prior model for the class labels.

Thus, we want to maximize the likelihood function $L(X_1, \ldots, X_n|C) = P(X_1|C)^{\alpha l} \ldots P(X_n|C)^{\alpha n} P(C)$ where $\alpha_r, 0 \leq \alpha_r \leq 1$ is the reliability factor for sensor $r$.
Furthermore, denoting $G_{ij}$ the local neighborhood of pixel *(i, j)* we can write:

$$P(C(i,j)|C(k,l); \{k,l\} \neq \{i,j\} = P(C(i,j)|C(k,l); \{k,l\} \in G_{ij}) = \frac{1}{Z} e^{-U(C)/T} \tag{3}$$

where $U$ is the so called energy function, Z is a normalizing constant factor and T is a temperature term often used in statistical physics. In our model we always considered a second order neighborhood, that is, the eight pixels closer to each single pixel of the image. If we want to maximize $P(C(i,j)|C(k,l); \{k,l\} \in G_{ij})$, we find that we need to minimize

$U(C)$, where $U(C(i,j)) = \sum_{\{k,l\} \in G_{ij}} \beta I(C(i,j), C(k,l))$ and $I(C(i,j),\ C(k,l)) = -1$ if $C(i,j) = C(k,l)$, 0 if $C(i,j) \neq C(k,l)$.

To perform the classification, we need to minimize $U(X_1, \ldots, X_n, C) = \alpha_s U_{spect}(X_s) + U_{sp}(C)$, where $U_{sp}$ is given by previous equation and $U_{spectr}$ is defined as below.

$$U_{spectr}(X_s(i,j), C(i,j)) = \frac{B_s}{2} \ln |2\pi \sum\nolimits_k| + \frac{1}{2}(X_s(i,j) - \mu_k)^T \sum\nolimits_k^{-1} (X_s(i,j) - \mu_k) \tag{4}$$

$\sum_k$ and $\mu_k$ are, respectively, the class-conditional covariance matrix and mean vector for class k and $B_{\mathrm{r}}$ is the number of spectral bands or features for source $r$.

Many algorithms are available for MRF implementation, but they are often demanding. The simplest one, but still effective, is ICM, Iterated Conditional Mode, which rapidly allows to reach a local minimum of the energy function.

## 4 SOME RESULTS

The first test site for the proposed procedure is the town of Bam, Iran, affected by a disastrous earthquake in December 2003. To analyze the area, a set of two acquisitions by the ASAR sensor on board of the ENVISAT-1 satellite is considered. In particular, the first image was recorded on December 3rd, 2003, and the second on February 11th, 2004. The goal for mapping this urban area was to provide a detailed land use map, similar to those available on line and provided, because of the emergency situation, by means of a visual interpretation of optical high-resolution images of the area. In particular, five classes were considered, i. e. mixed urban/vegetation, vegetation, urban residential, dense urban and desert. The colors used in the map (see the ground truth in Figure. 1) are light green, dark green, yellow, red and white, respectively.

The maps shown in Figure.2, together with the overall accuracy values with respect to a digitized version of the above mentioned manual map, gives us a feeling of the current achievable results and the current drawbacks of the technique. More precisely, this Figure allows to compare the analysis based solely on MRF or neuro-fuzzy classification of the two images with the map obtained considering a set of texture features found to be suitable for a better classification. While the overall accuracy does

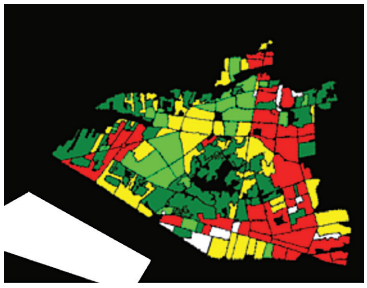


Figure 1. Ground truth for Bam land use mapping.

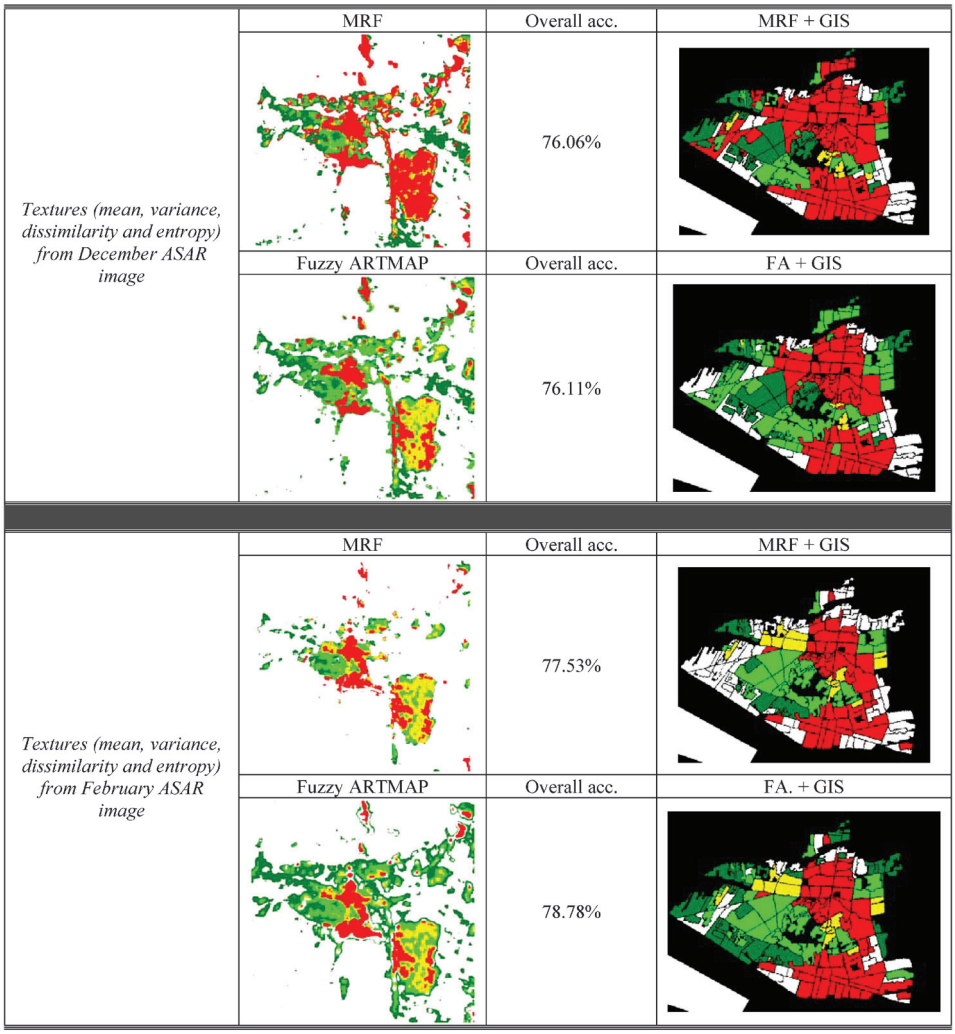
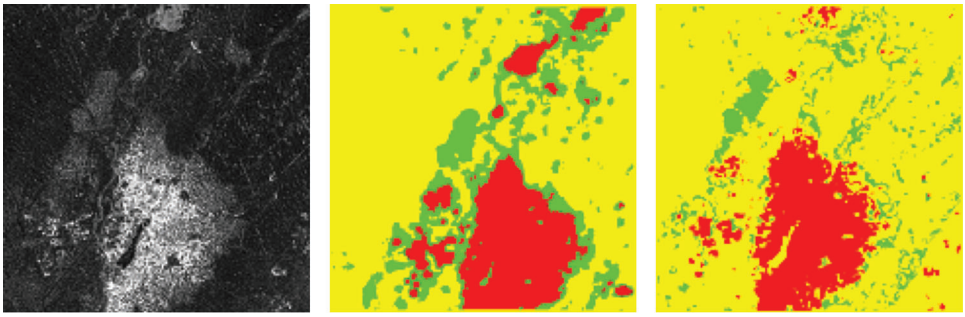



Figure 2. Classification maps and accuracy results for SAR data classification in the urban area of Bam (Iran). Five classes are used: mixed urban/vegetation (light green), vegetation (dark green), urban residential (yellow), dense urban (red) and desert (white).

not change very much from the first two rows to the following ones, Figure. 2 shows that texture features are more suitable to model the interior of the urban area and to discriminate among different land use classes. Moreover, the comparison between MRF and neuro-fuzzy results shows that MRF is more "conservative" and does not suffer too much from commission errors, while the fuzzy ARTMAP classifier is more prone to misclassification of pixels outside the settlement area, but it characterizes better the inner core of the town.

The second test area is the area around the town of Al Fashir, the largest human settlement in the North Darfur region in Sudan. The data on Al Fashir were recorded by



(a) (b) (c)

Figure 3. Classification maps and accuracy results for SAR data classification in the human settlement inside and around the town of Al Fashir (Sudan). Three classes are used: informal settlements (light green), formal settlements (red), rocks and bare soil (yellow): (a) original data; (b) map from classification of SAR textures; (c) map from joint classification of SAR and SPOT textures.

ASAR on July 26th and August 13th, 2004. In that period, the war in progress in the area and the famine produced by the same events resulted in the sudden raising of a huge tent camp North West of the main town area. Therefore, the area was therefore the target of a wide effort of humanitarian aid, and a vast amount of remotely sensed data were recorded to help rescuers and NGO officers.

Figure.3 shows the mapping results obtained using both ASAR data sets and applying the above described texture analysis to them. Best performances are obtained using two sets of textures (mean, second moment and variance; or: mean, entropy, variance and dissimilarity), computed using a $21 \times 21$ window. The results of the second choice are shown in Figure. 3(b), next to one of the two original images. The tent camp is the low brightness area North-West of the main urban core, and it is actually split into two main tent blocks.

Figure. 3(b) shows that texture analysis may provide a first discrimination between formal and informal settlements, but that spatial analysis at a single frequency does not allow to discard uninteresting areas outside human settlement. There are indeed rock formations and bare soil areas with very similar radar response as some parts of these settlements. Therefore, a certain amount of misclassification is in order, testified by the red and green ‘‘blobs’’ all around the map, and especially in the top right area, where only rocky hills may be found.

### 4.1 *Integrating other sources of information*

As already noted, results in Figure. 2 are very interesting, and validate the approach we have been using for more structured urban areas in Europe. However, they can be made even more useful by integrating possibly available prior information about the urban area. Indeed, SAR images analyzed by means of texture classification do not provide a sharp definition of the area, because of the mix of their own coarse resolution and the effect of texture computation in boundary areas. This problem may be, however, very

simply solved if a Geographical Information System layer would be available. This is not always the case, especially for developing countries. If this information is not retrievable from a GIS, the boundaries may be obtained by printed maps, or by a first, possibly rough, interpretation of other sources of information. These maps (or GIS layers) are easy to fuse with pixel-based classification map. The most straightforward approach is to assign each sub-area individuated to the class to which the majority of mapped pixels belong. Results for Bam area are provided in the same Figure. 1 and show how effective this very simple processing step may be.

A similar approach may be provided for the Al Fashir area, but the high variability of the human settlement characteristics in the tent camp does not allow to get GIS layers accurate enough for the area. However, we may use acquisitions from the SPOT sensor for the same area in a data fusion procedure. To this aim, texture features from both SAR and SPOT data were considered and jointly classified by means of the neuro-fuzzy classifier. A comparison between Figure. 3(b) and (c) allows understanding the advantage of using both information sources. As a matter of fact, the misclassifications have been greatly reduced and better delineation of both the formal and the informal settlement areas has been obtained. Although some misclassification with rock soil still persists, this has been reduced, too.

## 5 CONCLUSIONS

Satellite SAR data, far from being outdated by VHR optical imagery, are expected to have a future in the urban remote sensing arena.This is not only a result of the increased ground spatial resolution of future systems, which is going to rival optical sensors. The relatively long time coverage of radar data all over the world and the increased ability to use even coarse data to discriminate among urban land use classes may be helpful for a number of different research lines.

In this work we have chosen to discuss the challenges coming from mapping formal as well as informal human settlements in different parts of the world. The two examples come from desert or semi-arid areas, and show the problems related with misclassifications with bare soil classes. It is expected, however, that similar problems arise in tropical areas, due to dense vegetation, and we hope that they may be solved with similar texture analysis approaches. This is a very open point and worth the efforts of future researches.

## ACKNOWLEDGEMENTS

The authors wish to thank B. Houshmand for providing the data over Bam, and the Joint Research Center in Ispra for sharing the ASAR and SPOT images of Al Fashir. Many thanks also to M. Stasolla for performing the classifications for this area.

## REFERENCES


Alparone, L., S., Baronti, A., Garzelli, F., Nencini, ''Landsat ETM + and SAR image fusion based on generalized intensity modulation'', *IEEE Trans. Geosci, Remote Sens*., vol. 42, n. 12, pp. 2832–2839, Dec. 2004.

Barber D.G. and E.F., LeDrew, ''SAR sea ice discrimination using texture statistics: a multivariate approach,'' *Photogrammetric Engineering and Remote Sensing*, Vol. 57, No. 4, pp. 385–395.

Caltagirone, F., G., Manoni, P., Spera, and R., Vigliotti, ''SkyMed/COSMO mission overview,'' in Proc. *IGARSS98*, 1998 pp. 610–621, June 1988.

Available online at: http://www.caf.dlr.de/tsx/start_en.htm, accessed on the 22th of May, 2006.

Carpenter, G.A., M.N., Gjaja, S., Gopal, and C.E., Woodcock, ''ART neural networks for remote sensing: vegetation classification from Landsat TM and terrain data'', *IEEE Trans. Geoscience and Remote Sensing*, Vol. 35, pp. 308–325, 1997.

Costamagna, E., P., Gamba, P., Lombardo, G., Chinino, ''Statistical analysis and neuro-fuzzy classification of polarimetric SAR images of urban areas'', Proc. of the *ERS/ENVISAT Symposium*, Gotheborg, Sweden, Oct. 2000.

Dekker, R.J. ''Texture analysis and classification of SAR images of urban areas'', Proc. of *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pp. 258–262, Berlin, May 2003.

Dell'Acqua, F., P., Gamba, ''A preliminary statistical analysis of textural features in RADARSAT-1 images of an urban area'', Proc. of *URSI Commission F Symposium*, Ispra, Apr. 2005.

Dell'Acqua, F., P., Gamba, A., Iodice, G., Lisini, D., Riccio, G., Ruello: ''Simulation and analysis of fine resolution SAR images in urban areas,'' Proc. of the 2nd *GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas* (URBAN2003), Berlin, Germany, 22–23 May 2003, pp. 132–135.

Dell'Acqua, F., P., Gamba, ''Discriminating urban environments using multi-scale texture and multiple SAR images'', Proc. of the *Workshop on on Advances in Techniques for Analysis of Remotely Sensed Data,* CD-ROM, Greenbelt (MD), Oct. 2003.

Dell'Acqua, F., P., Gamba, ''Remote sensing of human settlements for water management in Africa'', Department Of Electronics Internal Report, May 2006.

www.dlr.de/hr/EUSAR2006/Program_Session/, accessed as of May 22th, 2006.

Fatone, L., P., Maponi, F., Zirilli, ''Fusion of SAR/optical images to detect urban areas'', Proc. of *IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pp. 217–221, Rome, 2001.

Gamba, P., F., Dell'Acqua, ''Discriminating urban environments using multi-scale texture and multiple SAR images'', *Int. J. Remote Sensing,* in press.

Gamba, P., F., Dell'Acqua: ''Improved multiband urban classification using a neuro-fuzzy classifier,'' *Int. J. Remote Sensing*, Vol. 24, n. 4, pp. 827–834, 2003.

Gamba, P., F., Dell'Acqua, G., Trianni: ''Semi-automatic choice of scale-dependent features for satellite SAR image classification'', *Pattern Recognition Letters*, Vol. 27, n. 4, pp. 244–251, Mar. 2006.

Haralick, R.M., K., Shanmugam, and I., Dinstein, ''Texture features for image classification,'' *IEEE Trans. on Syst. Man Cybern.*, Vol. 3, pp. 610–621, Nov. 1973.

Matsuoka M. and F., Yamazaki, ''Use of interferometric satellite SAR for earthquake damage detection,'' in Proc. 6th. *Conf. Seismic Zonation*, 2000.

Pesaresi, M. ''The remotely sensed city,'' Final report of the post-doctoral fellowship at the JRC Ispra, Italy, March 2000.

Schistad Solberg, A.H., A.K., Jain, and T., Taxt, ''Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images'', *IEEE Trans. Geoscience and Remote Sensing*, Vol. 32, no. 4, pp. 768–778, July 1994.