SAR/multispectral image fusion for the detection of environmental hazards with a GIS

Angela Errico[a] and Cesario Vincenzo Angelino[a] and Luca Cicala[a] and Dominik Patryk Podobinski[a] and Giuseppe Persechino[a] and Claudia Ferrara[b] and Massimiliano Lega[b] and Andrea Vallario[c] and Claudio Parente[c] and Giuseppe Masi[d,e] and Raffaele Gaetano[d,e] and Giuseppe Scarpa[d,e] and Donato Amitrano[d] and Giuseppe Ruello[d] and Luisa Verdoliva[d,e] and Giovanni Poggi[d,e]

[a]CIRA, the Italian Aerospace Centre, Capua, Italy;
[b]Department of Engineering, University Parthenope, Naples, Italy;
[c]DiST, University Parthenope, Naples, Italy;
[d]DIETI, University Federico II, Naples, Italy;
[e]CNIT, National Inter-University Consortium for Telecommunications, Italy

ABSTRACT

In this paper we propose a GIS-based methodology, using optical and SAR remote sensing data, together with more conventional sources, for the detection of small cattle breeding areas, potentially responsible of hazardous littering. This specific environmental problem is very relevant for the Caserta area, in southern Italy, where many small buffalo breeding farms exist which are not even known to the productive activity register, and are not easily monitored and surveyed. Experiments on a test area, with available specific ground truth, prove that the proposed systems is characterized by very large detection probability and negligible false alarm rate.

Keywords: Remote sensing, GIS, multispectral image, synthetic aperture radar, image analysis.

1. INTRODUCTION

The use of multi-sensors and multi-temporal data has emerged as a very successful approach for many remote-sensing problems.[1,2] However, the integration of such data with other complex data, that are difficult to represent verbally or sometimes even visually, plays a key role in the analysis of complex phenomena. For these purposes, the use of Geographic Information Systems (GIS) is steadily increasing. In these systems, geographic data describing features on the earth's surface are managed, displayed, manipulated, and analyzed. Moreover, the use of GIS together with digital remote sensing makes it possible to rapidly collect and analyze spatial data, yielding a powerful set of tools for the analysis of the environment.

Some related works can be found in the literature. A GIS-assisted system for hazardous waste site monitoring based on the integration of multispectral and lidar data with numerous types of thematic information has been proposed.[3] Other papers[4,5] demonstrate the potential of multi-temporal Landsat images for landfill site monitoring. Moreover, the use of hyperspectral images to characterize vegetation at hazardous waste sites, with different analysis methods (vegetation indices, red-edge positioning, and machine learning) has been proposed.[6] In parallel with optical remote sensing, research on the use of radar remote sensing data for environmental crimes has also been going on, with several papers[7–9] focusing on the use of synthetic aperture radar (SAR) interferometry for the detection and monitoring of landfills. In all these papers, interesting cues for future developments are proposed, concerning in particular the integration of various types of information. However, this integration is still limited as, for example, optical and SAR data are never exploited jointly in the context of a GIS-supported system.

In this paper we propose a GIS-based methodology, using optical and SAR remote sensing data, together with more conventional sources, for the detection of small cattle breeding areas, potentially responsible of hazardous littering. This specific environmental problem is very relevant for the Caserta area, in southern Italy, where many small buffalo breeding farms exist which are not even known to the productive activity register, and are

not easily monitored and surveyed. Indeed, the use of multi-sensors[10] and multi-temporal[11] data has emerged as a very successful approach for many remote-sensing problems. Starting from the analysis of a small number of companies already surveyed or known a priori, we extract a general description of some typical features of such facilities, and of their signatures in remote sensed imagery, both optical and SAR. This information is then used in a GIS-based processing workflow to detect new facilities of the same type unknown to the official census. Experiments on a test area, with available specific ground truth, prove that the proposed system is characterized by very large detection probability and negligible false alarm rate. This paper is organized as follows: Section 2 describes the proposed workflow. In Section 3 experimental results are presented, and conclusions are drawn in Section 4.

2. FUSION WORKFLOW

The Buffalo breeding facilities (BBFs) have some features that allow developing an efficient workflow that proved quite effective in the experiments. BBFs are mostly characterized by sheds and by the fenced uncovered spaces nearby used for breeding and buffaloes and for accumulating animal waste. Features like sheds are distinctly visible both in optical and SAR images, even if the responses in both kinds of images are not at all specific. Indeed they can be confused with asphalted roads, which represents another highly reflective cover in optical images, or even worst in case of generic buildings that are a problem for both sources. Additionally sheds can be also far from the fenced spaces where buffaloes live, on the other hand, the spectral signature of the manure is highly specific, easily discriminated from bare soil in the near infrared (NIR) band, and stable to changes in solar illumination.

After such considerations we decided to use as main source of information the 4-band multispectral GeoEye-1 optical images, focusing on the manure signature. In this process the main purpose is the classification of the images, which are available at different dates and are processed separately to create maps of candidate BBFs. The first task carried out on the multiresolution images is pansharpening, which provides a datacube with high spatial resolution of the panchromatic image and full spectral features. Then image segmentation is used in order to improve performance and reduce complexity. Given the need to extract region contours as accurately as possible for subsequent vectorization, we resort to edge-oriented segmentation techniques, in particular we use the Canny edge detector[12] which provides a very good performances. But in the presence of such complex images as those provided by high-resolution remote sensing, even the best edge detector provides edge maps that are largely inconsistent, with real edges that go undetected because of insufficient contrast, and false edges that are generated due to noise and imperfections. So to obtain a close contours a reasonable choice is to use watershed transform,[13] which was not enough because the application of watershed to real-world remote sensing images provides an exceedingly large number of regions, many of which are due to minor imperfections of the edge map, or just to the discrete geometry of the images and should be obviously merged together. That is why we decided to use a more sophisticated segmentation algorithm called Edge Mark and Fill (EMF).[14,15] EMF carries out a marker-controlled watershed segmentation, markers that are regions superimposed to the original image that force all pixels covered by a given mark to belong to the same segment.

The classification task on the other hand has the purpose to classify each segment of the area of interest as either manure or not manure based on the spectral response vectors of the component pixels. The manure class is well represented by a single-mode probability density function (pdf), in particular a multivariate Gaussian with mean μ_c and covariance matrix Σ_c. Moreover, all other classes that are easily recognized, that is green vegetation, dry vegetation, gray soil or stone, brown soil, asphalt, water, and red roofs are also well described as the manure class, but they are not in our interest in this research so they are merged all together as not manure class. In principle, given the spectral vector $X(s) = x$ associated with pixel s, the label or class $\hat{c}(s) \in \mathcal{C} = 1, 2, ..., C$ is chosen according to the MAP (Maximum A-posteriory Probability) rule, that is

$$\hat{c}(s) = \arg \max_{c \in \mathcal{C}} \Pr(c|x) = \arg \max_{c \in \mathcal{C}} f(x|c)\Pr(c)/f(x) \quad (1)$$

where $f(x)$ and $f(x|c)$ are the unconditional and conditional pdfs of the data. However, since we assume the various classes to be equally likely a priori, lacking any information about them, the MAP rule becomes a

Maximum Likelihood rule which, in the Gaussian hypothesis, reduces to

$$\hat{c}(s) = \arg \min_{c \in \mathcal{C}} \left[\ln |\Sigma_c| + (x - \mu_c)^T \Sigma_c^{-1} (x - \mu_c) \right] \tag{2}$$

Finally, to reduce the influence of noise, the decision is made on regions rather than pixels. For each homogeneous region singled out by segmentation, the average spectral signature is computed, and this is used in the classification step. Thanks also to this choice, the classification appears to be quite reliable, despite the simple multivariate gaussian model adopted.

Since we are dealing with images coming from multiple sources, the coregistration process is necessary to avoid inaccuracies that can hardly affect the quality of the final image analysis product.[16] This task is necessary to guarantee a precise spatial correspondence among physical areas and objects in various images. To achieve a good quality of coregistration, the rectification was applied using Rational Function Model (RFM),[17] which requires a DTM (Digital Terrain Model) of the whole area and at least 39 Ground Control Points (GCPs) of which besides the image coordinates, also the 3D (altimetric and planimetric) position in a geodetic-cartographic reference system must be known. A 5m $\times$ 5m DTM of the region of interest was built by means of linear interpolation on vector maps in scale 1:5000, and over 100 GCPs were considered for RFM application. The quality of coregistration was tested by considering the positional error, that is, the difference between the exact and estimated coordinates, computed both on the GCPs and on a disjoint set of Check Points (CPs). The standard deviation of the error turned out to be always below 1m for GCPs, and slightly more than 1m for CPs.

The major drawback of this workflow is the presence of a lot of false alarms caused by shadows projected by buildings over bare soil that are plentiful in urban area because the high density of buildings. To solve this problem we will use the multi-temporal SAR stack, which is processed jointly, to detect and create a map of man-made areas. The simplest way to separate man-made areas from natural ones based on SAR images is to exploit the interferometric coherence between successive acquisition because of their different scattering formation physical principles.[18] SAR images must be coregistered with one-another in order to estimate the average interferometric coherence and, subsequently, the urban/rural mask. The coregistration comprises three steps[19] during which the alignment is progressively refined until a precision in the order of fractions of a pixel is reached. The map is obtained by firstly co-registering the SAR images with one-another, and then computing the stack of interferometric phases and the corresponding coherence map, which is eventually thresholded to provide the desired urban mask.

The last step of the workflow is to process in a GIS environment both the cadastral map, including prior information on the location of facilities officially registered and all these outputs that firstly have to be converted from raster to a vectorial format. In the data fusion block, the classification maps corresponding to the various dates (two in our case) are combined through a logical AND, discarding in advance, however, regions too small and isolated. Detections occurring in urban areas, singled out thanks to the SAR-domain processing, are removed as well. Eventually a map of possible BBFs unknown to the official registry will be the result of the entire process.

3. EXPERIMENTS

We will consider in the experiments the measures, used in two-class hypotheses test, called precision P, recall R, and the synthetic F1 measure F. Indeed, we are interested in detecting the presence of a given target class, manure in classification, and buffalo breeding facility in detection. These measures are defined, w.r.t. a generic class T, as

$$P = \Pr(c = T | \hat{c} = T) \tag{3}$$

$$R = \Pr(c = T | \hat{c} = T) \tag{4}$$

and

$$F = \frac{2PR}{P + R} \tag{5}$$

where c and $\hat{c}$ indicate the true and selected class/hypothesis.

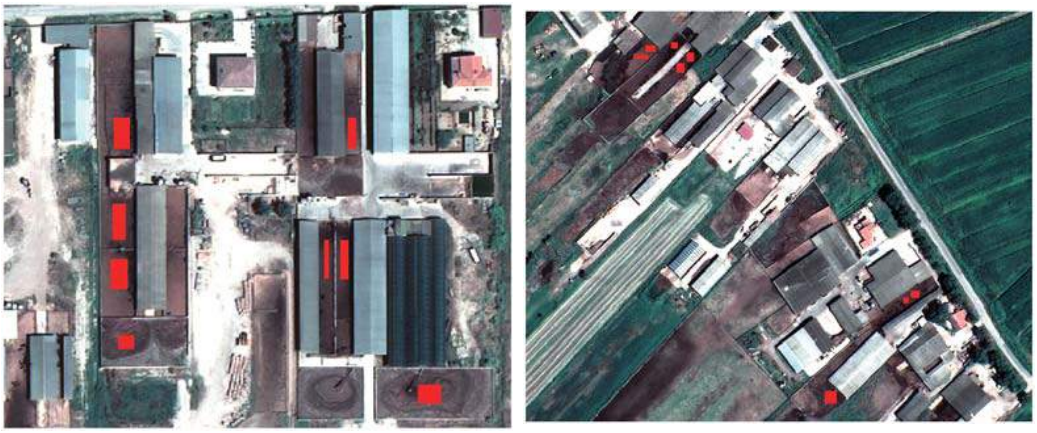

Figure 1. Training (left) and test (right) sets for classification. Red boxes correspond to manure areas.

training	**classification**	**precision**	**recall**	**F-measure**
pixel	pixel	0.862	0.975	0.915
pixel	segment	0.902	0.995	0.946
segment	segment	0.854	0.994	0.919

Table 1. Comparison of training/classification combinations.

For the classification task, we resort to supervised classification, given the wealth of information available. A relatively small area of the image, shown in Fig. 1a, is selected as training set, taking care to include all the features that will be found in the test set where the analysis will be eventually carried out. Classification performance is then evaluated on a different area of the image, the test set, shown in Fig. 1b. Our classifier is trained on the pixels drawn from the training set, while the decision is made on segments, namely on the average spectral response computed over all pixels belonging to a segment. This mixed solution was chosen after comparing performance with the other meaningful alternatives, where training and classification are performed both on pixels or both on segments. Results are reported in Table 1, w.r.t. the target class manure, and are computed pixel-wise irrespective of how the decision is made. Although the performance is definitely good in all cases, the selected mixed solution guarantees an appreciable gain in precision, and therefore in the F-measure.

For the selected solution, we computed over a total of $N = 113367$ pixels, the complete 15-class confusion matrix A with entries a_{ij} counting the number of pixels of class j that have been classified as belonging to class i. Based on a confusion matrix, several global quality indicators are usually computed. The *overall accuracy* (OA), which represents the percentage of sample pixels that are correctly classified, is defined as

$$\tau = \sum_i a_{ii}/N. \tag{6}$$

The *kappa coefficient*, defined as

$$\kappa = \frac{N \sum_i a_{ii} - \sum_i a_{i+} a_{+i}}{N^2 - \sum_i a_{i+} a_{+i}} \tag{7}$$

with $a_{i+} = \sum_j a_{ij}$ and $a_{+i} = \sum_j a_{ji}$, discounts successes obtained by chance, and is therefore more conservative (it can be also negative). The average accuracy (AA), also frequently used, is defined as the mean of per-class producers accuracies a_{ii}/a_{+i}. Finally, the normalized accuracy norm is computed on a confusion matrix modified in order to give equal importance to all classes, irrespective of the number of samples in each one. These indexes are all very high for our classifier: $\tau = 78.19\%$, $\kappa = 76.05\%$, $AA = 80.98\%$, $\tau^{norm} = 86.92\%$ especially considering the large number of classes considered some of which pretty similar to one another.

Detection performance is assessed on the image shown in Fig. 2. The small region in the red box will be used for visual inspection. Our goal is *i*) to detect BBFs, when present, and *ii*) to avoid declaring their presence otherwise. The ground truth was provided by an expert photointerpreter who analyzed thoroughly the whole

Figure 2. The image used in the experiments. The small region in the red box is used for detailed visual inspection of results.

Figure 3. Segment level decision on the small area of the image at the two dates. Green: correct, red: false alarms.

urban mask	images	precision	recall	F-measure
NO	T_1	0.950	0.197	0.326
	T_2	0.975	0.234	0.378
	$T_1 + T_2$	0.925	0.872	0.897
YES	T_1	0.950	0.307	0.464
	T_2	0.975	0.514	0.673
	$T_1 + T_2$	0.9325	0.959	0.942

Table 2. Detection performance with different variants of the proposed procedure.

image and detected eventually 40 BBFs, drawing their approximate contours in GIS as regular polygons, shown in yellow (nine of them) in the example clip of Fig. 3.

In the same figure we also show the segments that have been classified as wet soil, in green (correct decision) when they are mostly inside a BBF area, in red (false alarm) when they are mostly outside all of them. However, we are interested in detecting facilities, not segments. Therefore we use these data to label the 40 BBFs as either detected, when comprising at least a green segment, or missed, when no green segment falls within its bounding polygon. In the example clip, all 9 BBFs are detected at both dates. With this information, we can compute a meaningful recall indicator. As for precision, no similar conversion seems possible, and we are forced to operate at segment level, computing precision as the ratio between the number of segments (green) correctly declared wet soil, presumably manure, and the number of all segments (green or red) declared wet soil, irrespective of their real class, thus including errors. Although working at segment-level, this latter indicator provides a good insight on the quality of the whole procedure. If precision is too low, in fact, the technique points out many more targets than actually present, becoming basically useless. To reduce false alarms we resort to a suitable map of man-made areas, obtained starting from the SAR coherence map. However, we further process this image, taking advantage on information on build- ing density, readily available in the GIS environment, after a suitable GCP-guided co-registration, to extract the map of dense man-made areas shown in Fig. 4, which allows us to separate urban from rural regions in the scene. In the last row of Tab. 3 we report the performance indicators obtained with the proposed procedure (last row) including the masking of dense urban regions and the combination of multitemporal data. In the same table we also show results obtained with some variants where some of the available pieces of information are neglected. When only one date is considered, either T1 or T2, we obtain a lot of false alarms, causing a very low recall and F-measure. This is already clear in Fig. 3, where many red segments appear. However, while regions in BBFs are persistent, because they are continuously covered by manure, external regions are only occasionally classified as such, maybe because periodically fertilized, and can be eliminated through a multitemporal analysis. By using a simple logical AND, we obtain the pretty good indicators in the third row. Fig. 5 shows the effects of the logical AND on our example clip. However, this is not always the case, reducing slightly precision to 0.925. In the last three rows of the table we report the same data when the mask for dense urban areas is used. Recall improves significantly because bare soil areas, when shadowed by buildings, generate a large number of false alarms. On the other hand, precision is obviously not affected by masking, because BBFs are always rather far from large urban centers. The full fledged technique guarantee eventually both high precision (37 facilities detected out of 40) and high recall (only 12 false alarm segments out of 306).

4. CONCLUSION

We proposed a methodology for detecting environmental hazards based on SAR and multispectral data and GIS-based processing. The proposed system shows quite satisfactory results, and it can be a valuable help to counter and control environmental crimes. Of course, the proposed methodology can be easily converted to tackle other environmental hazards by modifying some input data, and some processing tasks. A campaign of validation on a wider test area is under development.

REFERENCES

[1] Lega, M. and Napoli, R. M. A., "Aerial infrared thermography in the surface waters contamination monitoring," *Desalination and Water Treatment* **23**(1-3), 141–151 (2010).

Figure 4. Dense urban areas (in lilac) extracted by GIS-level refinement of the man-made map.

Figure 5. Segment level decisions based on multi temporal data. No false alarm occurs in the clip.

[2] Lega, M., Kosmatka, J., Ferrara, C., Russo, F., Napoli, R. M. A., and Persechino, G., "Using advanced aerial platforms and infrared thermography to track environmental contamination," *Environmental Forensics* **13**(4), 332–338 (2012).

[3] Jensen, J., Hodgson, M., Garcia-Quijano, M., Im, J., and Tullis, J., "A remote sensing and gis-assisted spatial decision support system for hazardous waste site monitoring," *Photogramm. Eng. Remote Sensing* **75**, 169 – 178 (2009).

[4] Faisal, K., Alahmad, M., and Shake, A., "Remote Sensing Techniques as a tool for Environmental Monitoring," in [*International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*], (2012).

[5] Lein, J., "Landfill monitoring using remote sensing: a case study of glina, romania," *Waste Management and Research* **31**, 1075 – 1080 (2013).

[6] Im, J., Jensen, J., Jensen, R., Gladden, J., Waugh, J., and Serrato, M., "Vegetation cover analysis of hazardous waste sites in utah and arizona using hyperspectral remote sensing," *Remote Sensing* **4**, 327 – 353 (2007).

[7] Ottavianelli, G., Hobbs, S., Smith, R., and Bruno, D., "Assessment of sar data and interferometric products for solid waste landfill management," in [*RSPSoc Annual Conference 2005 - Measuring, Mapping and Managing a Hazardous World*], (September 2005).

[8] Ottavianelli, G., Hobbs, S., Smith, R., Morrison, K., and Bruno, D., "Sar interferometric products and hyperspectral data for monitoring solid waste landfill operations," in [*Fourth ESA CHRIS Proba workshop*], (September 2006).

[9] Karathanassi, V., Choussiafis, C., and Grammatikou, Z., "Monitoring the Change in Volume of Waste in Landfill Using SAR Interferometry," in [*32nd EARSel Symposium and 36th General Assembly*], (2012).

[10] Farah, I., Boulila, W., Ettabaa, K., Solaiman, B., and Ahmed, M. B., "Interpretation of multisensor remote sensing images: multiapproach fusion of uncertain information," *IEEE Transactions on Geoscience and Remote Sensing* **46**(12), 4142 – 4152 (2008).

[11] Melgani, F., Serpico, S. B., and Vernazza, G., "Fusion of multitemporal contextual information by neural networks for multisensor image classification," in [*International Geoscience and Remote Sensing Symposium (IGARSS)*], **7**, 2952 – 2954 (July 2001).

[12] Canny, J., "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (June 1986).

[13] Beucher, S. and Lantuejoul, C., "Use of Watersheds in Contour Detection," in [*International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation, Rennes, France.*], (Sept. 1979).

[14] Gaetano, R., Masi, G., Scarpa, G., and Poggi, G., "A marker-controlled watershed segmentation: Edge, mark and fill," in [*Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*], 4315–4318 (2012).

[15] Masi, G., Scarpa, G., Gaetano, R., and Poggi, G., "A watershed-based segmentation technique for multiresolution data," in [*Image Analysis and Processing - ICIAP 2013 - 17th International Conference, Naples, Italy, September 9-13, 2013. Proceedings, PartI*], 241–250 (2013).

[16] Dai, X. and Siamak, K., "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Transactions on Geoscience and Remote Sensing* **36**(5), 1566 – 1577 (1998).

[17] Tao, C. V. and Hu, Y., "Image rectification using a generic sensor model - rational function model," in [*International Archives of Photogrammetry and Remote Sensing, Vol. XXXIII, Part B3*], 874–881 (2000).

[18] Ferretti, A., Prati, C., and Rocca, F., "Permanent scatterers in sar interferometry," *IEEE Transactions on Geoscience and Remote Sensing* **39**(1), 8–20 (2001).

[19] Li, Z. and Bethel, J., "Image coregistration in sar interferometry," in [*The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Services, XXXVII (B1)*], 433–438 (2008).