NEW HIERARCHICAL JOINT CLASSIFICATION METHOD FOR SAR-OPTICAL MULTIRESOLUTION REMOTE SENSING DATA

Ihsen HEDHLI[a,b], Gabriele MOSER[b], Sebastiano. B. SERPICO[b] and Josiane ZERUBIA[a]

[a] INRIA, AYIN team, France, firstname.lastname@inria.fr
[b] DITEN, Department of Electrical, Electronic, and Telecommunications Engineering and Naval Architecture – University of Genoa, Italy, firstname.lastname@unige.it

ABSTRACT

In this paper, we develop a novel classification approach for multiresolution, multisensor (optical and synthetic aperture radar), and/or multiband images. Accurate and time-efficient classification methods are particularly important tools to support rapid and reliable assessment of the ground changes. Given the huge amount and variety of data available currently from last-generation satellite missions, the main difficulty is to develop a classifier that can take benefit of multiband, multiresolution, and multisensor input imagery. The proposed method addresses the problem of multisensor fusion of SAR with optical data for classification purposes, and allows input data collected at multiple resolutions and additional multiscale features derived through wavelets to be fused.

Index Terms— Multisensor, multiresolution remote sensing images, supervised classification, hierarchical Markov random fields.

1. INTRODUCTION

Nowadays, a wide variety of remote sensing images is available. Therefore, it becomes more and more important to be able to analyze compound data sets consisting of different types of images acquired by different sensors, as they allow a spatially distributed and temporally repetitive view of the monitored area at the desired spatial scales. In particular, the opportunity of joint availability of synthetic aperture radar (SAR) and optical images offers high resolution (HR), all-weather, day/night, short revisit time data, polarimetric, and, multifrequency acquisition capabilities. Similarly, the strong differences in terms of wavelength range (microwave vs. visible and near infrared), sensitivity to cloud cover and sun illumination (strong for optical imagery vs. almost negligible for SAR), and noise-like properties (speckle in SAR vs. generally low noise variance in current HR optical sensors) make the joint use of HR optical and SAR imagery especially interesting for many applications to environmental monitoring and risk management. Within this framework, there is a definite need for classification methods that automatically correlate different sets of images taken on the same area from different sensors and at different resolutions. Two common ways are usually accepted to deal with multisensor data: (i) the combination of classifiers and (ii) the classification of preliminary fused images either explicitly or implicitly. Option (i) leads to an improved global classification by employing basic rules such as majority voting [1], or more specific methods, such as those relying on neural networks [2] or the Dempster-Shafer theory [1]. The second strategy (ii) banks on the direct fusion of the input images. In this case, the classification is obtained by applying, to the fused images, some widely used classification algorithms such as, for instance, methods based on Markov random fields (MRFs) [3, 4]. Due to their generally non-causal nature, these models lead to iterative inference algorithms that are computationally demanding (e.g., optimization via simulated annealing) [5]. One way to circumvent this problem is to resort to a Markov model on a quadtree [6] where in-scale causality permits non-iterative inference with acceptable computational time, and offers the well-known advantages of standard hierarchical techniques that improve robustness and ability to deal with multiresolution data. Under this contextual assumption, multisensor fusion may be based on explicit statistical modeling by finding a joint probability distribution given the class-conditional marginal probability density functions (PDFs) related to each sensor. The joint statistics can be designed by resorting to meta-Gaussian distributions as shown in [7], by using more sophisticated multivariate statistics such as multivariate copulas [8], or by applying non-parametric density estimators [20]. However, employing heterogeneous data (SAR-optical in our case) makes the task of finding an appropriate multivariate statistical model complex, time demanding, and possibly prone to overfitting.

In the proposed method, the computation of joint statistics is avoided, and a novel approach, based on multiple quadtrees in cascade, to multisensor and multiresolution fusion is described. For each sensor, the input image is associated with a separate quadtree structure on the basis of its resolutions. The proposed approach formalizes a super-

The authors would like to thank the Italian Space Agency (ASI) for providing the COSMO-SkyMed image and GeoEye Inc. and Google crisis response for providing the GeoEye image. The authors would also like to thank the Italian Gov and INRIA for the partial financial support of the PhD of Ihsen Hedhli.

vised Bayesian classifier within this multiple quadtree topology that combines a class-conditional statistical model for pixel-wise information and a hierarchical MRF for multisensor and multiresolution contextual information.

This paper is organized as follows: In Section 2, we focus on the proposed hierarchical model. Section 3 explains the methodological choices that come with the modeling of the class-conditional statistics. Results of the use of this new multisensor hierarchical model are presented in Section 4. Finally, we conclude and present some future work in Section 5.

2. HIERARCHICAL MULTISENSOR MODEL

The objective of this study is to develop a novel method for SAR-optical and multiresolution classification based on a hierarchical Markovian model. This hierarchical structure should be highly parallel in order to handle the heterogeneity of data acquired by different sensors with different physical properties and should provide a topology that simplifies the interactions between different images in the compound data set. In this context, the pyramid structure [9] is a type of signal representation in which images are organized according to their resolutions. Among others, quadtrees [6] have been proposed as attractive candidates for modeling the scale-to-scale interactions in the aforementioned pyramid structure. The choice of a quadtree allows taking benefit from its good analytical properties [10] (e.g., causality) and to apply non-iterative classification algorithms such as the marginal posterior mode (MPM) [6]. The aim is to maximize recursively the posterior marginal at each site s, which associates the most probable class label x_s given the entire input multisource information y :

$$\hat{x}_s = arg\ \max_{x_s} p(x_s \mid y) \tag{1}$$

An especially novel element of the proposed approach is the use of multiple quadtrees in cascade (see Fig. 1), each associated with the set of images given by one specific sensor. This approach aims at exploiting the multiscale information that is typically associated with either SAR or optical HR imagery.

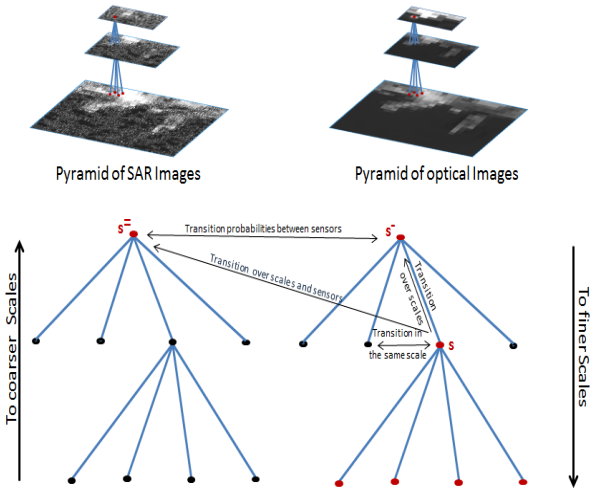

Fig. 1. Multisensor hierarchical structure

It is worth noting that the proposed hierarchical structure implies a constraint among the resolutions of the images at the various levels of the pyramid. In fact, the choice of quadtree topology imposes a factor of 2 between the spatial resolutions of the images in two successive levels. This implies that missing levels can occur in the pyramid structure if the input images do not fill all levels of the quadtree by themselves. In the proposed method, these "empty" levels are filled in through a wavelet decomposition of the images associated with the finer levels [11]. Quadtree structure allows, in a very natural way, the use of an explicit statistical modeling through a hierarchical Markov random field formulation using a series of random fields at varying scales, resolutions, and sensors, on the basis of the transitions defined on the quadtrees as shown in Fig. 1.

Accordingly, a novel formulation of MPM is proposed through multiple quadtrees in cascade. Specifically, the posterior marginal $p(x_s \mid y)$ of the label of each site in the quadtree related to the optical sensor is expressed as a function not only of the posterior marginal of the parent node $p(x_{s^-} \mid y)$ in the same quadtree but also of the posterior marginal $p(x_{s^=} \mid y)$ of the parent node in the quadtree associated with SAR images as shown in (2), with the aim to characterize the SAR-optical correlations associated, at different scales, with distinct images in the input multisource data i.e. :

$$\boldsymbol{p}(\boldsymbol{x_s} \mid \boldsymbol{y}) = \sum_{x_{s^-},\, x_{s^=}} \frac{p(x_s,\, x_{s^-},\, x_{s^=} \mid y_{d(s)})}{\sum_{x_s} p(x_s,\, x_{s^-},\, x_{s^=} \mid y_{d(s)})} \cdot \boldsymbol{p}(\boldsymbol{x_{s^-}} \mid \boldsymbol{y})\, \boldsymbol{p}(\boldsymbol{x_{s^=}} \mid \boldsymbol{y}), \tag{2}$$

where bold denotes the marginal posteriors of interest to MPM. (2) involves two conditional independence assumptions: (i) the label x_s depends only on the data of the site and their descendants; and (ii) the label at parent s^- is independent of the label at parent $s^=$ at the previous date, when conditioned to the data y (details can be found in [13]).

This formulation allows calculating recursively the posterior marginal $p(x_s \mid y)$ at each site s while the probabilities $p(x_s,\, x_{s^-},\, x_{s^=} \mid y_{d(s)})$ are made available. Thus, the computation of these joint probabilities boils down to the determination of the other probabilities involved in (3):

$$p(x_s,\, x_{s^-},\, x_{s^=} \mid y_{d(s)}) = p(x_s \mid x_{s^-},\, x_{s^=}) \cdot \frac{p(x_{s^-} \mid x_{s^=}) \cdot p(x_{s^=})}{p(x_s)} \cdot p(x_s \mid y_{d(s)}), \tag{3}$$

where the factor $p(x_s \mid x_{s^-},\, x_{s^=})$ corresponds to the child-parent transition probability; $p(x_s)$ is the prior probability; $p(x_{s^-} \mid x_{s^=})$ is the transition probability between sensors at the same scale and $p(x_s \mid y_{d(s)})$ is the partial posterior marginal probability.

To compute these probabilities, we take benefit from the hierarchical structure defined above and we use two recursive passes on the quadtree, referred to as "bottom-up" and "top-down" passes (see Fig. 2.).

2.1. Top-down pass: Prior probability estimation

A preliminary classification is done at the first quadtree composed by SAR images using the classical MPM using a single quadtree, where the segmentation is obtained recursively over scales (details can be found in [6]). Then we use the resulting classification map to estimate the prior in the root R of the second quadtree composed by optical images (blue arrow labeled with the number 1 in Fig. 2.). To estimate the prior in the root given the resulting classification map, we use a spatial Markovian model which takes into account the contextual information given by SAR images, and therefore, leads to better prior estimation for the optical images by adding spatial information in the classification process. Thus, employing the Hammersley-Clifford theorem [3], we can define a local prior for each site s as a Potts model as shown in (4).

$$p\left(x_s \,|x_{s'}, s' \sim s\right) \propto exp\left(-\beta \sum_{s \sim s'} \delta\left(x_s, x_{s'}\right)\right), \qquad (4)$$

where $\delta(.)$ is the Kronecker symbol, $s \sim s'$ denotes that s and s' are neighbors with respect to a given neighborhood system, and β is a smoothness parameter [3].

Then, a top-down pass (blue arrow labeled with the number 2 in Fig. 2.) is performed for each finer level, and the prior-probability distribution is derived as a function of the prior-probability distribution calculated at the parent level and of the transition probabilities from the parent to the current level, i.e.:

$$p\left(x_s\right) = \sum_{x_{s^-}} p\left(x_s \mid x_{s^-}\right). p\left(x_{s^-}\right), \qquad (5)$$

to favor an identical parent-child labeling and to model the statistical interactions between consecutive levels of the quadtree. We model the transition probability in the form introduced by Bouman et al. [12], i.e. $(n = 0, 1 \cdots R-1)$

$$p(x_s \mid x_{s^-}) = \begin{cases} \theta & x_s = x_{s^-} \\ \dfrac{1-\theta}{M-1} & x_s \neq x_{s^-} \end{cases}, \qquad (6)$$

2.2. Bottom-up pass: joint probabilities estimation

A bottom-up pass recursion is performed to estimate the joint probabilities $p\left(x_s, x_{s^-}, x_{s^=} \mid y_{d(s)}\right)$ starting from the leaves of the quadtree corresponding to the optical sensor and proceeding until the root is reached based on the factorization in (3). In addition to the prior that is calculated in the previous top-down pass, there are two kinds of transition probabilities that are needed to compute the factorization in (3). The first one is the transition probability between sensors at the same scale $p\left(x_{s^-} \mid x_{s^=}\right)$, and the second is the child-parent transition probability $p\left(x_s \mid x_{s^-}, x_{s^=}\right)$. Details on the calculations of these probabilities are shown in [13].

The bottom-up pass also involves the estimation of the partial posterior marginals $p\left(x_s \mid y_{d(s)}\right)$ where $y_{d(s)}$ denotes the collection of the observations including s and all its descendants. Concerning the former probabilities, Laferté et al. [6] proved that:

$$p\left(x_s \,|y_{d(s)}\right) \propto p(y_s|\, x_s).\, p\left(x_s\right).\prod_{\tilde{s}\in s^+}\sum_{x_{\tilde{s}}}\left[\frac{p\left(x_{\tilde{s}}\,|y_{d(\tilde{s})}\right)}{p(x_{\tilde{s}})}.\, p\left(x_{\tilde{s}}\,|x_s\right)\right] \qquad (6)$$

$$p\left(x_s \,|y_s\right) \propto p(y_s|\, x_s).\, p\left(x_s\right), \qquad (7)$$

Thus, the bottom-up pass is a recursion that estimates $p\left(x_s \mid y_{d(s)}\right)$ starting from the leaves of the quadtree where the partial posterior marginals are computed via (7) (green arrow labeled with the number 1 in Fig. 2) and then, proceeding until the root which is reached using (6) (green arrow labeled with the number 2 in Fig. 2). (6) involves the pixelwise class-conditional PDFs $p(y_s|\, x_s)$ of the image data at each node of each quadtree (see next section).

2.3. Top-down pass: posterior probabilities and optimization

The algorithms described in Sections 2.1 and 2.2 provide all needed probabilities to compute $p\left(x_s, x_{s^-}, x_{s^=} \mid y_{d(s)}\right)$ at each level of the quadtree (green arrow labeled with the number 2 in Fig. 2). After initializing the posterior at the root (red arrow labeled with the number 1 in Fig. 2) where $p\left(x_r \mid y_{d(r)}\right) = p\left(x_r \mid y\right)$, one could easily compute, recursively in a top-down pass (red arrow labeled with the number 2 in Fig. 2), the posterior $p\left(x_s \mid y\right)$ at each pixel s for all tree levels using (2). Then, the maximization of $p\left(x_s \mid y\right)$ is done by employing a modified Metropolis dynamics (MMD) algorithm, which exhibits good properties with respect to both computational cost (which is relatively low) and accuracy of the results [15].

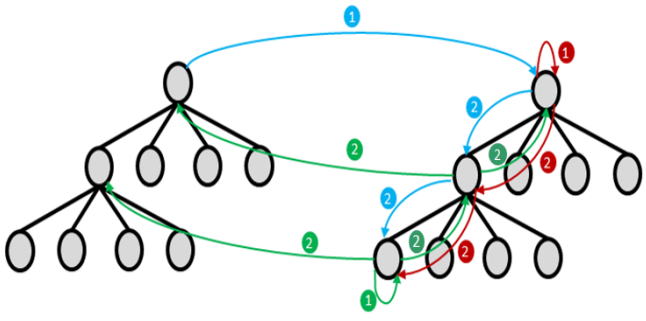

Fig. 2. Recursive passes on the Quadtrees

3. PIXELWISE CLASS-CONDITIONAL PDFS

Given a training set for each input data, for each class m, scale n and sensor, we model the corresponding class-conditional marginal PDF $p(y_s|x_s = m)$ using finite mixtures of independent grey level distributions:

$$p(y_s|x_s = m) = \sum_{i=1}^{K^n} \pi_i^n \, F_i^n(y_s|\theta_i^n), \qquad (8)$$

where π_i^n are the mixing proportions, θ_i^n is the set of the parameters of the i$^{\text{th}}$ PDF mixture component of the m^{th} class at the n^{th} scale level F_i^n.

The mixture modeling is performed depending on the different types of remote sensing imagery used in this study. Indeed, when the input data at the n^{th} scale level is an optical image, class-conditional marginal *PDF* $p(y_s|x_s = m)$ related to each class m can be modeled by a Gaussian mixture model [16] with a set of parameters associated with the corresponding mean and variance. On the opposite, SAR acquisitions are known to be affected by speckle [17]. For this reason, we use appropriate SAR-specific models for such images, such as the generalized gamma distribution [18]. The parameters of the mixture model for both SAR and optical images are estimated through the stochastic expectation maximization (SEM) algorithm [19], which is an iterative stochastic parameter estimation algorithm developed for problems characterized by data incompleteness and approaching, under suitable assumptions, maximum likelihood estimates. For each scale, SEM is applied to the training samples of each class to estimate the related parameters.

4. RESULTS

Experiments are discussed regarding two datasets collected over Port-au-Prince (Haiti) including: (i) single-polarization single-look CSK (© ASI; Fig. 3. (a)) with HH polarization, stripmap acquisition modality, with 2.5-m pixel spacing; and (ii) a coregistered GeoEye RGB acquisition (© GeoEye; Fig. 3. (b)). Multiscale features are extracted through 2D discrete wavelet transform. Five land cover classes have been considered: urban (red), water (blue), vegetation (green), soil (yellow) and containers (purple).

We compare the proposed method with the classification results obtained using the following techniques:

1) Separate hierarchical classifications of the images provides by the two sensors (Fig. 3(d) for optical and Fig. 3(e) for SAR). In this case, the classification is obtained recursively over the scales where the use of quadtree structure in MPM schema yields "blockly" segmentations (details can be found in [6]).
2) The multisensor single-scale approach in [7] (see Fig. 3(g)), in which the likelihood term is constructed by merging generalized gamma (for SAR) and Gaussian (for optical) marginals into a meta-Gaussian distribution. The classification is obtained by the maximum likelihood rule.
3) The multi-sensor multiscale method proposed in [8], in which a model for the multivariate joint class-conditional statistics of the co-registered input images at each resolution is designed by resorting to multivariate copulas. The estimated joint probability density function is plugged into a hierarchical Markovian model based on a quadtree structure (see Fig. 3(f)).

A preliminary visual analysis of the resulting classification maps suggests that the proposed hierarchical method leads to accurate results, especially as compared to separate hierarchical classifications of the images provided by the two sensors (see Figs. 3(e) (d)). Indeed, experimental results obtained by using only SAR data accurately detect roads and containers, while the results generated by using only optical data better discriminate classes that are spatially homogenous. The proposed method effectively takes benefit from both SAR and optical imagery, and allows generating a classification result that visually well discriminates all classes in the considered HR data set. It is especially interesting to notice that the detection of the "containers" class, which is generally very overlapping with other urban or built-up classes in the feature space, improves significantly compared to the method in [8], in which the classification map is over smoothed (see highlighted region in fig. 3(f)). Compared to the previous multisensor classification method in [7] (see fig. 3 (g) and Tab. 1), which is based on transformations of the input features to a common jointly Gaussian domain, the proposed algorithm provides a spatially more regular classification result, thanks to contextual MRF modeling and wavelet feature extraction. There is an improvement using the proposed method compared to the hierarchical multisensor method in [8] (see fig. 3 (f) and Tab. 1) in terms of accuracy and computation time. For this method, the main source of misclassification is the container area, where asphalt is erroneously classified as vegetation. Moreover, with the proposed method, we avoid the computation of joint PDFs, while copulas are used in [7] and meta-Gaussian distributions are used in [8]. This choice results in reducing the computation time (see Tab.1).

5. CONCLUSION

In the proposed method, multisensor and multiresolution fusion is based on explicit statistical modeling. It combines a marginal statistical model of the considered input optical and SAR images, through hierarchical Markov random field modeling based on quadtrees in cascade, leading to a statistical supervised classification approach. We have developed a novel multisource MPM-based hierarchical Markov random field model that takes into account both SAR and optical information and leads to improved robustness of the classifier. When applied to challenging high-resolution data sets associated with urban and semi-urban test sites, the proposed method gives high overall classification accuracy with a small computation time (a few minutes). A further advantage of the proposed classifier is that it can be generalized to the use of different satellites and/or acquisitions dates by extending the multiple quadtree structure suitably. This research work will be done in the near future.

REFERENCES

[1] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition,", *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 3, pp. 418–435, 1992.

[2] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural network approaches versus statistical methods in classifica-

tion of multisource remote sensing data," *IEEE Trans. Geosci. Remote Sens*., vol. 28, no. 4, pp. 540–552, 1990.
[3] S. Z. Li, *Markov random field modeling in image analysis 3rd edition*, Springer, 2009.
[4] G. Winkler, Image Analysis, *Random Fields and Markov chain Monte Carlo Methods 2nd edition,* Springer, 2003.
[5] Z. Kato and J. Zerubia. *Markov Random Fields in Image Segmentation*. Boston: NOW publishers, 2012.
[6] J.M. Laferté, P. Perez and F. Heitz. "Discrete Markov modeling and inference on the quad-tree." *IEEE Trans. Image Processing*, pp. 390–404, 2000.
[7] B. Storvik and G. Storvik and R. Fjortoft, "On the combination of multisensor data using meta-Gaussian distributions," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2372–2379, 2009.
[8] A.Voisin, V. Krylov, G. Moser, J. Zerubia and S. B. Serpico. "Supervised Classification of Multi-sensor and Multi-resolution Remote Sensing Images with a Hierarchical Copula-based Approach." *IEEE Trans. on Geoscience and Remote Sensing*, pp. 3346-3358, 2014.
[9] J. M. Jolion, A. Rosenfeld. *A Pyramid Framework for Early Vision: Multiresolutional Computer Vision*. Norwell, MA, USA: Kluwer Academic Publishers, 2004.
[10] M. Basseville, A. Benveniste and A. S. Willsky. "Multiscale autoregressive processes, Part I: Schur-Levinson parametrization." *IEEE Trans. on Signal Processing*, vol. 40. no. 8, pp. 1915-1934, 1992.
[11] M. Crouse, R. Nowak and R. Baraniuk. "Wavelet-based statistical signal processing using hidden Markov models." *IEEE Trans. Signal Processing*, vol. 46, pp. 886–902, 1998.
[12] C. Bouman, M. Shapiro, "A multiscale image model for Bayesian image segmentation." *IEEE Trans. on Image Processing*, vol. 3, pp 162-177, 1994.
[13] I. Hedhli, G. Moser, J. Zerubia, S. B. Serpico. "New cascade model for hierarchical joint classification of multitemporal, multiresolution and multisensor remote sensing data", *IEEE International Conference on Image Processing*, France, 2014.
[14] J. Besag, "Spatial interaction and the statistical analysis of lattice systems (with discussion),"*J. R. Statist. Soc. B*, vol. 36, pp. 192–326, 1974.
[15] Z. Kato, M. Berthod and J. Zerubia. "A hierarchical Markov random field model and multitemperature." *Graphical models and image processing*, pp. 18–37, 1996.
[16] J. A. Richards, *Remote Sensing Digital Analysis*, Springer, 2013
[17] H.-C. Li, W. Hong, Y.-R. Wu, and P.-Z. Fan, "On the empirical-statistical modeling of SAR images with generalized gamma distribution," *IEEE J. Sel. Top. Signal Process*, vol. 5, no. 3, pp. 386–397, 2011.
[18] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar images*. SciTech Publishing, 2004.
[19] G. Celeux, D. Chauveau and J. Diebolt. "Stochastic versions of the EM algorithm: an experimental study in the mixture case." *Journal of Statistical Computation and Simulation*, vol. 55, no. 4, 287–314, 1996.
[20] K. Fukunaga, *Introduction to statistical pattern recognition*, 2nd edition, Academic Press, New York, 1990.

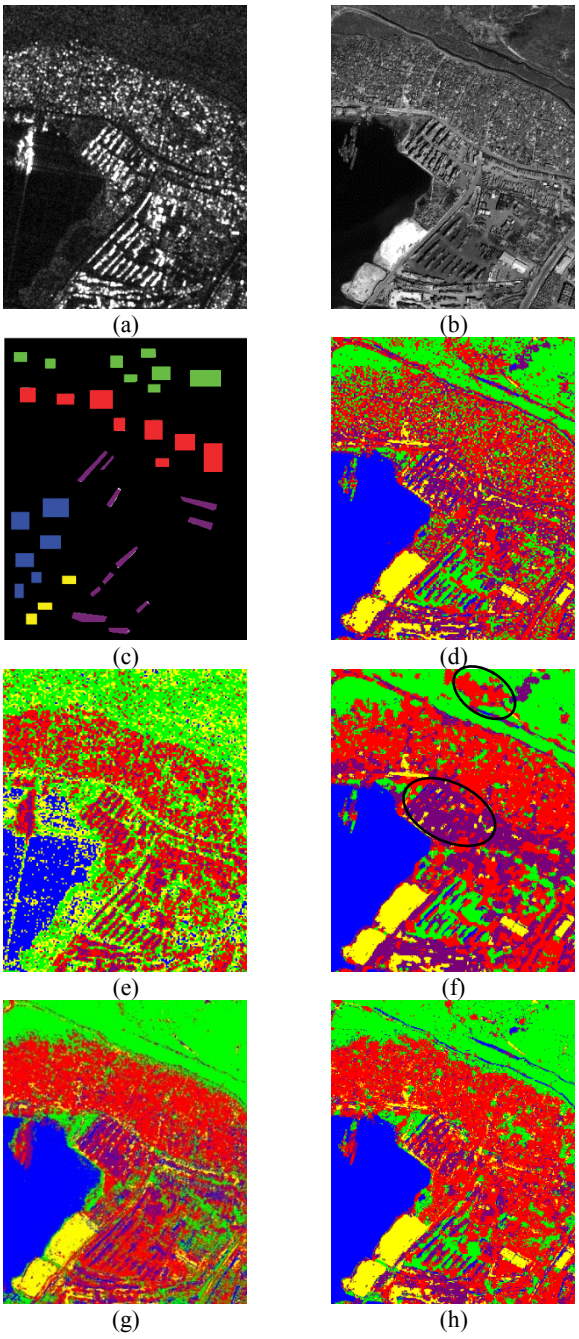

Fig. 3. (a) SAR image (© ASI), (b) one channel from the optical image (© GeoEye), (c) the available ground truth, (d) hierarchical MRF-based classification obtained from the optical image, using method in [6], (e) hierarchical MRF-based classification obtained for the SAR image, using method in [6], (f) result of the multisensor and multiresolution method in [8], (g) result of the multisensor method in [7], (h) hierarchical MRF-based classification obtained by the proposed cascade method.

	Water	urban	vegetation	soil	containers	overall	Computation time
The proposed method	100%	78,12%	89,46%	98,78%	47,12%	**82,69%**	**254s**
The method proposed in [7]	99,95%	97,32%	90,81%	96,22%	37,25%	79,44%	298s
The method proposed in [8]	100%	75.24%	87,16%	98,89%	49,31%	82,12%	668s

Table 1. Classification accuracies for the Port-au-Prince dataset
Experiments were conducted on an Intel i7 quad-core (2.40 GHz) 8-GB-RAM 64-bit Linux system.