Semisupervised Image Classification With Laplacian Support Vector Machines

Luis Gómez-Chova, Gustavo Camps-Valls, *Senior Member, IEEE,*
Jordi Muñoz-Marí, and Javier Calpe, *Member, IEEE*

Abstract—This letter presents a semisupervised method based on kernel machines and graph theory for remote sensing image classification. The support vector machine (SVM) is regularized with the unnormalized graph Laplacian, thus leading to the Laplacian SVM (LapSVM). The method is tested in the challenging problems of urban monitoring and cloud screening, in which an adequate exploitation of the wealth of unlabeled samples is critical. Results obtained using different sensors, and with low number of training samples, demonstrate the potential of the proposed LapSVM for remote sensing image classification.

Index Terms—Kernel methods, manifold learning, regularization, semisupervised learning (SSL), support vector machines (SVMs).

I. INTRODUCTION

IN REMOTE sensing image classification, we are usually given a reduced set of labeled samples to develop the classifier. Supervised classifiers such as support vector machines (SVMs) [1], [2] excel in using the labeled information, being (regularized) maximum margin classifiers also equipped with an appropriate loss function [3], [4]. Nevertheless, these methods need to be reformulated to exploit the information contained in the wealth of unlabeled samples, which is known as semisupervised classification. In semisupervised learning (SSL), the algorithm is provided with some available supervised information in addition to the unlabeled data.

The SSL framework is very active and has recently attracted a considerable amount of research [5], [6]. Essentially, two different classes of SSL algorithms are encountered in the literature. First, generative models involve in estimating the conditional distribution by modeling the class-conditional distributions explicitly, such as expectation-maximization algorithms with finite-mixture models [7], [8]. Second, discriminative models, in contrast to generative models, estimate the conditional distribution directly, and one does not have to specify the class-conditional distributions explicitly. The following two subgroups of SSL algorithms can be distinguished within these models: 1) low-density separation algorithms maximize the margin for labeled and unlabeled samples simultaneously, such as transductive SVM (TSVM) [9], [10], and 2) graph-based methods, in which each sample spreads its label information to its neighbors until a stable state is achieved on the whole data set [11], [12].

In the last years, TSVM and graph-based methods have captured great attention. However, some specific problems are identified in both of them. In particular, TSVM is sensitive to local minima and requires convergence heuristics by using an (unknown) number of unlabeled samples. Graph-based methods are computationally demanding and, generally, do not yield a final decision function but only prediction labels.

In this letter, we present a recently introduced semisupervised framework that incorporates labeled and unlabeled data in any general-purpose learner [13]. We focus on a semisupervised extension of the SVM, which introduces an additional regularization term on the geometry of both labeled and unlabeled samples by using the graph Laplacian [11], thus leading to the so-called Laplacian SVM (LapSVM) [13]. This methodology follows a noniterative optimization procedure, in contrast to most transductive learning methods, and provides a closed-form classification function for testing on novel samples not used in the training phase (out-of-sample predictions), in contrast to graph-based approaches. In addition, hard-margin SVM, directed-graph methods, label propagation methods, and spectral clustering solutions [3], [5] are obtained for particular free parameters of the LapSVM.

The performance of the LapSVM is illustrated in two challenging problems: the urban classification problem using multispectral [Landsat Thematic Mapper (TM)] and radar (ERS2 SAR) data [14] and the cloud-screening problem using the Medium Resolution Imaging Spectrometer (MERIS) instrument onboard the ESA Environmental Satellite [15]. On the one hand, the classification of heterogeneous urban areas is an important yet complex problem, particularly when different sensor sources are used, as they induce a highly variable input feature space. On the other hand, the amount of images acquired over the globe every day by Earth Observation satellites makes inevitable the presence of clouds. However, very few labeled cloud pixels are typically available, and cloud features change to a great extent depending on the cloud type, thickness, transparency, height, or background. In addition, cloud screening must be carried out before atmospheric correction, being the input data affected by the atmospheric conditions.

These problems constitute clear examples of classification in complex manifolds.[1] Learning in these conditions is more challenging when dealing with ill-posed problems, i.e., working with less number of labeled samples than the input dimension of the pixels. The use of LapSVM is motivated in this scenario,

[1]A manifold is a topological space that is locally Euclidean but in which the global structure may be more complicated.

since it permits one to use the labeled samples and efficiently exploiting the information contained in the high number of available unlabeled pixels to characterize the marginal distribution of data. This is particularly interesting in practical remote sensing applications, since obtaining additional unlabeled samples is cheap and easy, while labeling is expensive, difficult, or even impossible. In addition, the LapSVM constitutes a general framework for SSL, in which supervised SVM and other methods can be regarded just as particular cases. This letter presents, for the first time, an exhaustive analysis of the LapSVM in remote sensing problems.

The rest of this letter is outlined as follows. Section II reviews the framework of SSL, paying attention to the minimization functional and the need for different types of regularization. It also presents the formulation of the LapSVM. Section III shows the experimental results. Finally, Section IV concludes and outlines further work.

II. SSL FRAMEWORK

Regularization is necessary to produce smooth decision functions and, thus, to avoid overfitting to the training data. Since the work of Tikhonov [16], many regularized algorithms have been proposed to control the capacity of the classifier [1], [17]. The regularization framework has been recently extended to the use of unlabeled samples [13] as follows.

Notationally, we are given a set of l labeled samples $\{x_i, y_i\}_{i=1}^{l}$ and a set of u unlabeled samples $\{x_i\}_{i=l+1}^{l+u}$, where $x_i \in \mathbb{R}^N$ and $y_i \in \{-1, +1\}$. Let us now assume a general-purpose decision function f. The regularized functional to be minimized is defined as

$$\mathcal{L} = \frac{1}{l} \sum_{i=1}^{l} V(x_i, y_i, f) + \gamma_L \|f\|_{\mathcal{H}}^2 + \gamma_M \|f\|_{\mathcal{M}}^2 \quad (1)$$

where V represents a generic cost function of the committed errors on the labeled samples, γ_L controls the complexity of f in the associated Hilbert space $\mathcal{H}$, and γ_M controls its complexity in the intrinsic geometry of the marginal data distribution. For example, if the probability distribution is supported on a low-dimensional manifold, $\|f\|_{\mathcal{M}}^2$ penalizes f along that manifold $\mathcal{M}$. Note that this functional constitutes a general regularization framework that takes into account all the available knowledge.

The previous SSL framework allows us to develop many different algorithms just by playing around with the loss function V and the regularizers $\|f\|^2$. In this letter, we focus on the LapSVM formulation, which basically uses an SVM as the learner core and the graph Laplacian for manifold regularization. In the following, we review all the ingredients of the formulation.

LapSVM uses the same hinge-loss function as the SVM

$$V(x_i, y_i, f) = \max\{0, 1 - y_i f(x_i)\} \quad (2)$$

where f represents the decision function implemented by the selected classifier and predicted labels are $y_* = \text{sgn}(f(x_*))$.

We use as the decision function $f(x_*) = \langle w, \phi(x_*) \rangle + b$, where $\phi(\cdot)$ is a nonlinear mapping to a higher (possibly infinite) dimensional Hilbert space $\mathcal{H}$, and w and b define a linear discriminant function in that space. By means of the Representer Theorem [1], weights w can be expressed in the dual problem as the expansion over labeled and unlabeled samples $w = \sum_{i=1}^{l+u} \alpha_i \phi(x_i) = \Phi\alpha$, where $\Phi = [\phi(x_1), \ldots, \phi(x_{l+u})]^T$ and $\alpha = [\alpha_1, \ldots, \alpha_{l+u}]$. Then, the decision function is given by

$$f(x_*) = \sum_{i=1}^{l+u} \alpha_i K(x_i, x_*) + b \quad (3)$$

and K is the kernel matrix formed by kernel functions $K(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$. The key point here is that, without considering the mapping ϕ explicitly, a nonlinear classifier can be constructed by selecting the proper kernel. In addition, the regularization term can be fully expressed in terms of the kernel matrix and the expansion coefficients

$$\|f\|_{\mathcal{H}}^2 = \|w\|^2 = (\Phi\alpha)^T(\Phi\alpha) = \alpha^T K \alpha. \quad (4)$$

The geometry of the data is modeled with a graph in which nodes represent both labeled and unlabeled samples connected by weights W_{ij} [5]. Regularizing the graph follows from the smoothness (or manifold) assumption and, intuitively, is equivalent to penalize the "rapid changes" of the classification function evaluated between close samples in the graph

$$\|f\|_{\mathcal{M}}^2 = \frac{1}{(l+u)^2} \sum_{i,j=1}^{l+u} W_{ij}(f(x_i) - f(x_j))^2 = \frac{f^T L f}{(l+u)^2} \quad (5)$$

where $L = D - W$ is the graph Laplacian; D is the diagonal degree matrix of W given by $D_{ii} = \sum_{j=1}^{l+u} W_{ij}$, and $D_{ij} = 0$ for $i \neq j$; the normalizing coefficient $1/(l+u)^2$ is the natural scale factor for the empirical estimate of the Laplace operator [13]; and $f = [f(x_1), \ldots, f(x_{l+u})]^T = K\alpha$, where we have deliberately dropped the bias term b.

By plugging (2), (4), and (5) into (1), we obtain the regularized function to be minimized

$$\min_{\substack{\xi \in \mathbb{R}^l \\ \alpha \in \mathbb{R}^{l+u}}} \left\{ \frac{1}{l} \sum_{i=1}^{l} \xi_i + \gamma_L \alpha^T K \alpha + \frac{\gamma_M}{(l+u)^2} \alpha^T K^T L K \alpha \right\} \quad (6)$$

subject to $y_i(\sum_{j=1}^{l+u} \alpha_j K(x_i, x_j) + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$, $i = 1, \ldots, l$, where ξ_i are slack variables to deal with committed errors in the labeled samples. Introducing restrictions into the primal functional (6) through Lagrange multipliers, β_i and η_i, and taking derivatives with respect to b and ξ_i, we obtain

$$\min_{\alpha, \beta} \left\{ \frac{1}{2} \alpha^T \left(2\gamma_L K + \frac{2\gamma_M}{(l+u)^2} K^T L K \right) \alpha \right.$$
$$\left. - \alpha^T K J^T Y \beta + \sum_{i=1}^{l} \beta_i \right\} \quad (7)$$

where $J = [I \ 0]$ is an $l \times (l+u)$ matrix with I as the $l \times l$ identity matrix (the first l points are labeled) and $Y = \text{diag}(y_1, \ldots, y_l)$. Taking derivatives again with respect to α, we obtain the solution [13]

$$\alpha = \left(2\gamma_L I + 2\frac{\gamma_M}{(l+u)^2} L K \right)^{-1} J^T Y \beta^*. \quad (8)$$

Now, substituting again (8) into the dual functional (7), we obtain the following quadratic-programming problem [13]:

$$\beta^* = \max_{\beta \in \mathbb{R}^l} \left\{ \sum_{i=1}^{l} \beta_i - \frac{1}{2}\beta^T \mathbf{Q}\beta \right\} \qquad (9)$$

subject to $\sum_{i=1}^{l} \beta_i y_i = 0$ and $0 \le \beta_i \le (1/l)$, $i = 1, \ldots, l$, where

$$\mathbf{Q} = \mathbf{YJK} \left(2\gamma_L \mathbf{I} + 2\frac{\gamma_M}{(l+u)^2}\mathbf{LK} \right)^{-1} \mathbf{J}^T \mathbf{Y}. \qquad (10)$$

Therefore, the basic steps for obtaining the weights α_i for the solution in (3) are as follows: 1) build the weight matrix $\mathbf{W}$ and compute the graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$; 2) compute the kernel matrix $\mathbf{K}$; 3) fix regularization parameters γ_L and γ_M; and 4) compute α using (8) after solving (9).

The LapSVM is intimately related to other unsupervised and semisupervised classifiers. This is because the method incorporates both the concepts of kernels and graphs in the same classifier, thus having connections with transduction, clustering, graph-based, and label propagation methods. The minimizing functional used in the standard TSVM considers a different regularization parameter for labeled and unlabeled samples, which is the case in the proposed framework [cf. (1)]. In addition, LapSVM is directly connected with the soft-margin SVM ($\gamma_M = 0$), the hard-margin SVM ($\gamma_L \to 0$, $\gamma_M = 0$), the graph-based regularization method ($\gamma_L \to 0$, $\gamma_M > 0$), the label propagation regularization method ($\gamma_L \to 0$, $\gamma_M \to 0$, $\gamma_M \gg \gamma_L$), and spectral clustering ($\gamma_M = 1$). In conclusion, by optimizing parameters γ_L and γ_M over a wide enough range, the LapSVM theoretically outperforms the aforementioned classifiers [13].

LapSVM is formulated for binary classification problems. Extensions to multiclass problems can be devised by using classical one-against-one or one-against-all strategies. However, the following considerations have to be taken into account. On the one hand, when using one-against-one strategy, one analyzes each pair of classes separately by removing the other classes from both the labeled set and the unlabeled set. However, this approach is not realistic for real-world multiclass SSL, since the labels of unlabeled data are (obviously) unknown. On the other hand, the one-against-all strategy can be easily implemented for multiclass problems, as in [13].

III. EXPERIMENTAL RESULTS

This section presents the experimental results of the proposed method in two challenging scenarios: urban monitoring and cloud screening. These two examples are well suited, because efficient exploitation of unlabeled samples becomes strictly necessary to attain satisfactory results.

A. Model Development and Experimental Setup

We used both the linear kernel, $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i, \mathbf{x}_j \rangle$, and the radial basis function (RBF) kernel, $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/2\sigma^2)$, where $\sigma \in \mathbb{R}^+$ is the kernel width for the SVM, LapSVM, and TSVM, which is used in order to compare results provided by a related semisupervised SVM. The graph Laplacian $\mathbf{L}$ consisted of $l + u$ nodes connected
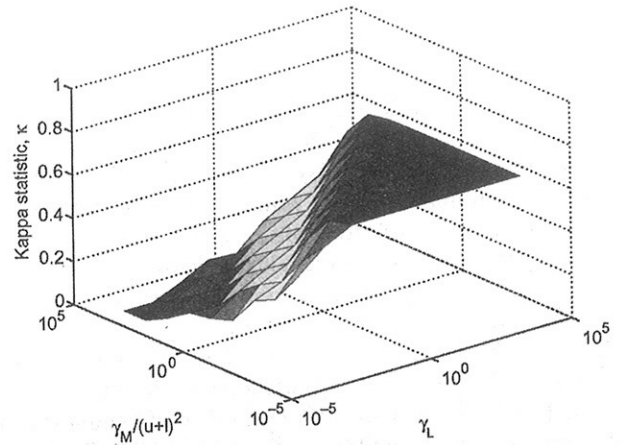

Fig. 1.   Illustrative example of a kappa statistic surface over the validation set for the LapSVM as a function of regularization parameters γ_L and γ_M.

using k nearest neighbors and compute the edge weights W_{ij} using the Euclidean distance among samples.

For the experiments, we generated training and validation sets consisting of $l = 400$ labeled samples (200 samples per class) and added $u = 400$ unlabeled (randomly selected) samples from the analyzed images to the training set for the LapSVM and TSVM. We focus on the ill-posed scenario and vary the rate of both labeled and unlabeled samples independently, i.e., 2%, 5%, 10%, 20%, 50%, and 100% of the labeled/unlabeled samples of the training set were used to train the models in each experiment. In order to avoid skewed conclusions, we run all experiments for a number of realizations, where the used training samples were randomly selected. All classifiers are compared using the overall accuracy (OA, in percent) and the estimated kappa statistic κ as a measure of robustness in the classification.

Free parameters γ_L and γ_M were varied in steps of one decade in the range of $[10^{-4}, 10^4]$, the number of neighbors k used to compute the graph Laplacian was varied from three to nine, and the Gaussian width was tuned in the range $\sigma = \{10^{-3}, \ldots, 10\}$ for the RBF kernel. The selection of the best subset of free parameters was done by cross validation. Fig. 1 shows κ as a function of the regularization parameters γ_L and γ_M obtained by a LapSVM in the validation set for an illustrative example. This figure clearly shows that the best classification results are obtained with $\gamma_L > \gamma_M/(u+l)^2$, which suggests a preference for the regularization of the classifier (supervised information) than for the regularization of the geometry of the marginal data distribution (unsupervised information). Regarding the number of neighbors, the analysis of the results revealed a low sensitivity to k in most of the experiments, but for small training sets, lower k values were chosen.

B. Urban Monitoring

1) Data Description: The image used in this first experiment was collected in the Urban Expansion Monitoring ESA Project (for further details, visit: http://dup.esrin.esa.int/ionia/projects/summaryp30.asp). The considered test site was Naples, Italy, where images from ERS2 SAR and Landsat TM sensors were acquired in 1999.
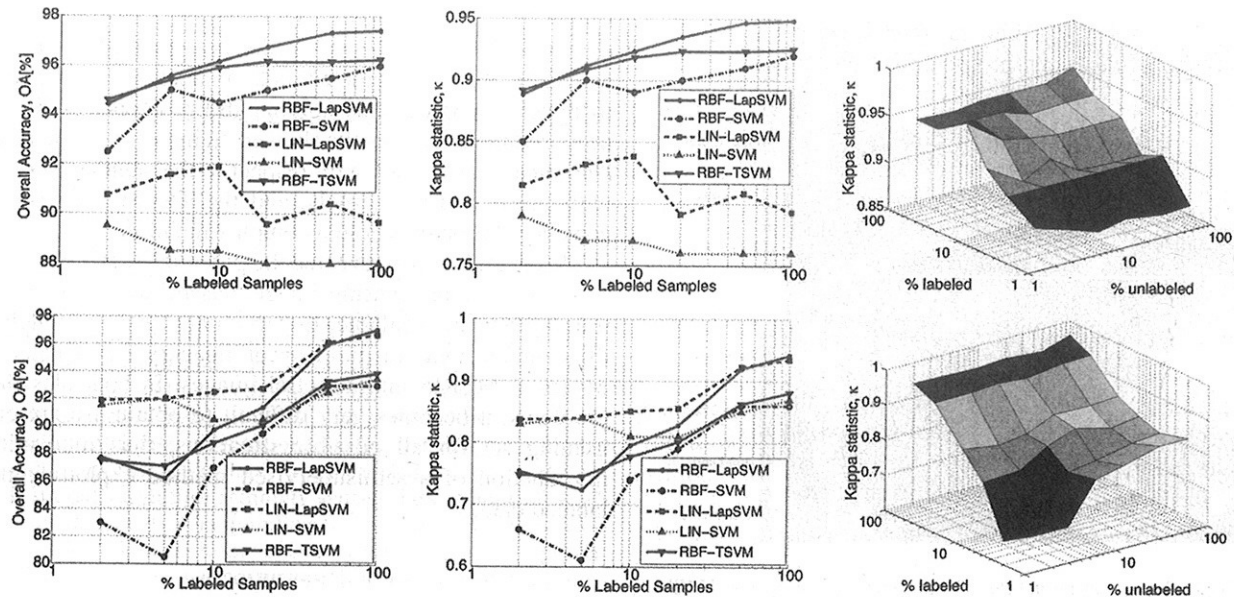
Fig. 2. Results for (top row) the urban classification and (bottom row) the cloud-screening problems. (Left) Overall Accuracy (OA, in percent) and (middle) Kappa statistic κ over the validation set as a function of the rate of labeled training samples used to build models. (Right) Kappa statistic surface over the validation set for the best RBF–LapSVM classifier as a function of the rate of both labeled and unlabeled training samples.

The available features were the seven Landsat bands, two SAR backscattering intensities (0–35 days), and the SAR interferometric coherence. Since these features come from different sensors, the first step was to perform a specific processing and conditioning of optical and SAR data and to coregister all images. After preprocessing, features were stacked at a pixel level (for full details, see [14]).

2) Model Comparison: Fig. 2 (top) shows the validation results for the analyzed SVM-based classifiers. Several conclusions can be obtained from this figure. First, LapSVM classifiers produce better classification results than SVM in all cases (note that SVM is a particular case of the LapSVM for $\gamma_M = 0$) for both the linear and the RBF kernels. LapSVM also produces better classification results than TSVM when the number of labeled samples is increased. Differences among methods are numerically very similar when a low number of labeled samples is available. On the right plot, the κ surface for the LapSVM highlights the importance of the labeled information in this problem.

3) Visual Inspection: The best LapSVM classifier was used to classify the whole scene, which consists of a scene of 200×200 pixels with urban and nonurban labeled samples. The classification map is shown in Fig. 3. Excellent classification accuracy is obtained, and uniform classification covers can be observed, even with so few labeled training samples.

C. Cloud Screening

1) Data Description: Experiments were carried out using two MERIS Level 1b (L1b) images taken over Barrax, Spain, which are part of the data acquired in the framework of the SPECTRA Barrax Campaigns 2003 and 2004 (ESA–SPARC Project, Contract ESTEC-18307/04/NL/FF). These two images were acquired on July 14 of two consecutive years (2003 and 2004). For this letter, we used as input 13 spectral bands (MERIS bands 11 and 15 were removed, since they are af-

Fig. 3. (Top left) Landsat TM RGB color composite, (top right) SAR intensity map, (bottom left) true map, and (bottom right) classification map of the urban scene, indicating (gray) "urban" and (white) "nonurban" classes. (Black) Pixels masked out in classification maps correspond to pixels where the available ground truth does not provide the class label, and they were rejected in order to estimate the classification accuracy.

fected by atmospheric absorptions) and six physically inspired features extracted from MERIS bands in a previous work [18]: cloud brightness and whiteness in the visible and near-infrared spectral ranges, along with atmospheric oxygen and water vapor absorption features.

2) Model Comparison: Fig. 2 (bottom) shows the validation results for the considered methods. Again, LapSVM models produce better classification results than SVM in all cases. In fact, LapSVM performs particularly better than the standard SVM for the RBF kernel when a low number of labeled samples is available, and unlabeled samples help in estimating the

Fig. 4. RGB color composite and classification maps for the analyzed MERIS multispectral images using the best LapSVM classifiers.

geometry of the manifold. The linear kernel seems to be more appropriate for the presented cloud-screening application if the number of labeled samples is low, while the difference between the linear and the RBF kernels decreases as the number of labeled samples increases. It is well known that simple (linear) decision functions are more appropriate in extremely ill-posed situations. As in the urban classification problem, LapSVM produces better classification results than TSVM when a moderate number of labeled samples is used (both methods provide similar classification accuracies when a low number of labeled samples is available). Finally, the κ surface for the LapSVM in Fig. 2 (bottom right) confirms, in general terms, the importance of both labeled and unlabeled information in this problem.

3) Visual Inspection: We used the best LapSVM to classify the whole scenes, which consist of MERIS L1b images of 1153×1153 pixels (reduced to around $500\,000$ useful pixels after the projection in latitude/longitude coordinates). The classification map for the 2003 and 2004 images are shown in Fig. 4. In this figure, we use as ground truth a cloud classification made by an operator following the methodology described in [18]. Excellent classification accuracy of 95.51% and 96.48% are obtained for the 2003 and 2004 images, respectively. The value of κ is 0.78 for both images and reflects the misclassification of a significant number of false cloud pixels, while all cloud pixels are correctly classified, suggesting that LapSVM benefits from the inclusion of unlabeled samples and obtains a reliable estimation of the marginal cloud data distribution. The committed errors correspond to high-altitude locations and bright bare-soil covers which are not well represented in the small randomly selected training set.

IV. CONCLUSION

A semisupervised method has been presented for the classification of urban areas and clouds. This method brings together the ideas of spectral graph theory, manifold learning, and kernel-based algorithms in a coherent and natural way to incorporate geometric structure in a kernel-based regularization framework. The solution of the LapSVM constitutes a convex-optimization problem and results in a natural out-of-sample extension from the labeled and unlabeled training samples to novel examples, thus solving the problems of previously proposed methods. Results showed an increase on the LapSVM classification accuracy with respect to the standard SVM, both with linear or RBF kernels, and the TSVM, suggesting that considered problems hold a complex manifold.

This letter has also revealed the potential of this classification method in remote sensing image classification, when reduced training sets are available. In particular, it has accurately identified urban areas in multisensor imagery and located cloud covers in MERIS multispectral images. In both classification problems, it becomes very difficult to obtain a representative training set for all possible situations, thus motivating the introduction of a semisupervised method exploiting the unlabeled data.

REFERENCES

[1] B. Schülkopf and A. Smola, *Learning With Kernels—Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press, 2002.
[2] G. Camps-Valls, J. L. Rojo-Álvarez, and M. Martinez-Ramón, Eds., *Kernel Methods in Bioengineering, Signal, and Image Processing*, Hershey, PA: Idea Group Publishing, Jan. 2007.
[3] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[4] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
[5] O. Chapelle, B. Schülkopf, and A. Zien, *Semi-Supervised Learning*, 1st ed. Cambridge, MA: MIT Press, 2006.
[6] X. Zhu, "Semi-supervised learning literature survey," Comput. Sci., Univ. Wisconsin-Madison, Madison, WI, Tech. Rep. 1530, 2005. Last modified on June 24, 2007. [Online]. Available: http://www.cs.wisc.edu/~jerryzhu/pub/ssl_survey.pdf
[7] N. M. Dempster, A. P. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc., B*, vol. 39, no. 1, pp. 1–38, Jan. 1977.
[8] B. Shahshahani and D. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1087–1095, Sep. 1994.
[9] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
[10] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.
[11] F. Chung, *Spectral Graph Theory*. Providence, RI: Amer. Math. Soc., 1997.
[12] G. Camps-Valls, T. V. Bandos Marsheva, and D. Zhou, "Semi-supervised graph-based hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3044–3054, Oct. 2007.
[13] M. Belkin, P. Niyogi, and V. Sindhwani, "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *J. Mach. Learn. Res.*, vol. 7, pp. 2399–2434, Dec. 2006.
[14] L. Gomez-Chova, D. Fernández-Prieto, J. Calpe, E. Soria, J. Vila-Francés, and G. Camps-Valls, "Urban monitoring using multitemporal SAR and multispectral data," *Pattern Recognit. Lett.*, vol. 27, no. 4, pp. 234–243, Mar. 2006.
[15] M. Rast and J. Bezy, "The ESA Medium Resolution Imaging Spectrometer MERIS a review of the instrument and its mission," *Int. J. Remote Sens.*, vol. 20, no. 9, pp. 1681–1702, Jun. 1999.
[16] A. N. Tikhonov, "Regularization of incorrectly posed problems," *Sov. Math., Dokl.*, vol. 4, pp. 1624–1627, Jan. 1963.
[17] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Adv. Comput. Math.*, vol. 13, no. 1, pp. 1–50, Apr. 2000.
[18] L. Gómez-Chova, G. Camps-Valls, J. Calpe, L. Guanter, and J. Moreno, "Cloud-screening algorithm for ENVISAT/MERIS multispectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 4105–4118, Dec. 2007.

◆IEEE

(Contents Continued from Page 321)

About the Cover: DInSAR mean deformation velocity map (in color) relevant to an area located in central Nevada, extending for about 600 × 100 km. The presented map, superimposed on an SRTM DEM (gray scale) of the zone, has been obtained by combining data relevant to six contiguous standard ERS SAR frames (track: 442, frames: 2781–2871) and by subsequently applying the SBAS-DInSAR algorithm to the overall data set composed by 264 frames (44 data sets for each frame). The investigated time period spans the 1992–2000 interval, and the spatial resolution of the DInSAR products is of about 200 × 200 m. The presented five plots show the deformation time series relevant to some of the main deformation patterns revealed by the SBAS-DInSAR analysis: (a) Lone Tree Gold Mine: subsidence due to water pumping in support of open-pit gold mining at the Lone Tree gold mine. (b) Crescent Valley: similar processes than for the Lone Tree Gold Mine occurring here at the Cortez Gold Mine. (c) Antelope Valley: subsidence due to water pumping in support of agricultural exploitation. (d) Eureka Valley Earthquake (M6.1 May 17, 1993). (e) Coso Geothermal Area: deformation associated with geothermal production and intense microseismicity. For more information, please see "SBAS-DInSAR Analysis of Very Extended Areas: First Results on a 60 000-km² Test Site," by Casu *et al.*, which begins on page 438.
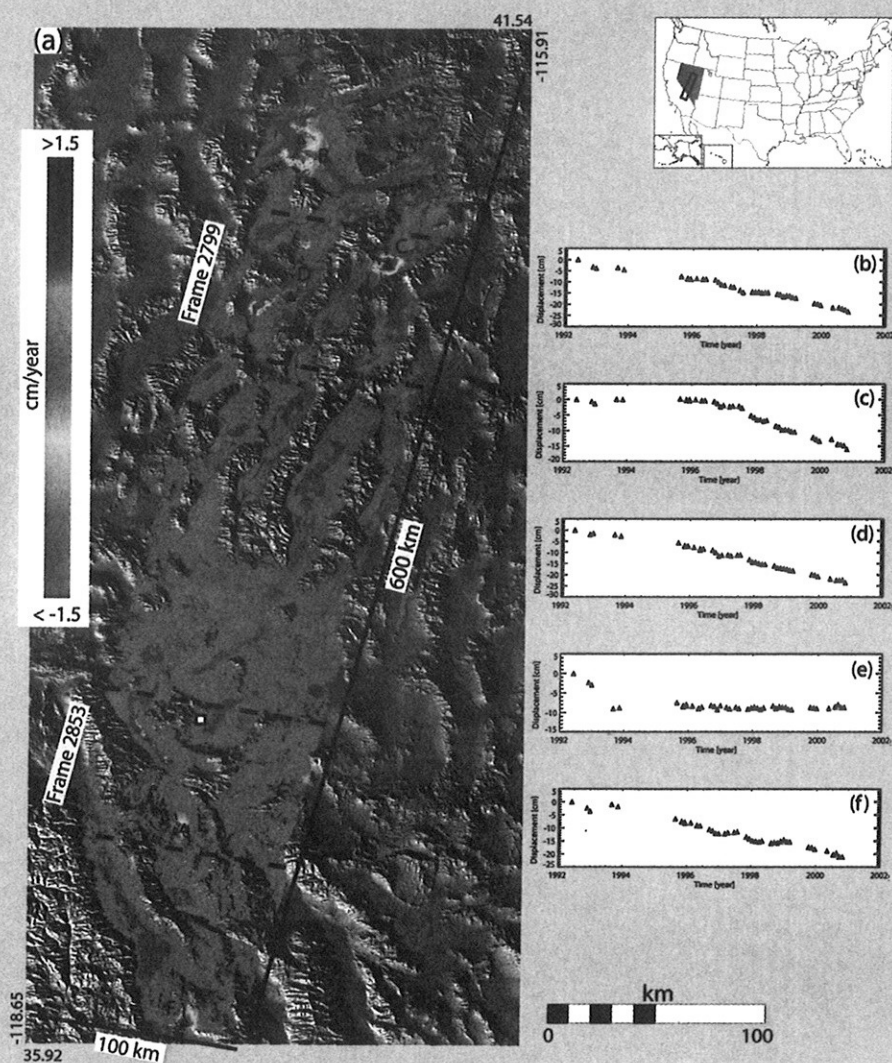
DInSAR results obtained by applying the Small BAseline Subset (SBAS) technique to a dataset of 264 ERS SAR frames relevant to an area of 600 × 100 km, located in central Nevada. (a) Retrieved mean deformation velocity map (in color) superimposed on an SRTM DEM (gray scale) of the zone. (b)–(f) DInSAR deformation time series for the pixels marked in (a) by triangles and labeled by the letters from B to F.

◆IEEE