AUTOMATIC UNSUPERVISED CLASSIFICATION OF SNOW-COVERED AREAS BY DECISION-TREE CLASSIFICATION AND MINIMUM-ERROR THRESHOLDING

Giorgia Macchiavello[1,2], *Gabriele Moser*[1], *Giorgio Boni*[2], *and Sebastiano B. Serpico*[1]

[1] University of Genoa, Dept. of Biophysical and Electronic Eng. (DIBE), via Opera Pia 11a, I-16145 Genoa (Italy), e-mail: sebastiano.serpico@unige.it

[2] CIMA Foundation, Savona University Campus, via A. Magliotto 2, I-17100 Savona (Italy)

1. EXTENDED ABSTRACT

In mountain areas such as Alpine and Apennine regions snow cover plays an important role in the hydrological modeling of extreme events such as floods and droughts and in the production of hydroelectric power. Information on snow cover are fundamental for monitoring and forecasting the available water during the snowmelt seasons. Indeed, an accurate snow-cover detection is fundamental to perform correct parameterization, calibration, and validation of snow models [1]. In this context, the availability of remotely sensed images represents a powerful tool for mapping and monitoring the extension and typology of snow cover over the mountainous territory. This potential is further enforced by the fact that *in-situ* measurements of snow cover and snow depth may often be difficult; in particular, they are usually impractical on large-scale areas. This points out the need for accurate and automatic remote-sensing image analysis techniques that allow snow-cover classes to be discriminated among one another and from other land-covers. From a methodological viewpoint, this goal could be achieved by resorting to supervised classification, but this approach would require training data to be available for the considered classes [2]. Due to the aforementioned intrinsic difficulty in collecting *in-situ* information, such a requirement may be critical, thus possibly strongly limiting the practical applicability of remote-sensing image classification in snow-cover mapping operational chains. This claims for the development of accurate and automatic unsupervised classification techniques, involving no training data in the generation of snow-cover maps.

In the present paper a novel method is proposed to automatically map the snow cover from optical multispectral images. The key-idea of the proposed technique is to integrate a “decision tree classifier” (DTC) approach and a Bayesian unsupervised thresholding algorithm, aiming at a complete automation of the image-classification process. Given a classification problem involving several classes, the DTC approach decomposes the problem in a suitable tree-structured collection of binary classification subproblems, for which simple (e.g., threshold-based) decision rules can be defined [3, 4]. DTCs are known to represent a powerful and flexible family of classification methods, allowing to incorporate both training data and prior knowledge, when available, in the definition of the tree structure. For instance, when working in a supervised framework, an optimal tree may be automatically constructed according to the distribution of the training data of the considered classes [3]. Alternatively, prior (either physical or semiheuristic) knowledge about the spectral responses of the channels of the adopted sensor may be exploited to define a case-specific tree especially tailored to the discrimination of given classes [4]. Here, the latter strategy is used by adopting the tree introduced in [4], which discriminates several snow-covered and non-snow-covered classes (i.e., “full snow,” “forest-snow,” “broken snow,” “snow-free,” “clouds”), by decomposing the related multiclass problem into a set of binary thresholding subproblems involving the multispectral channels and the corresponding normalized difference snow index (NDSI) [4].

The effectiveness of this DTC technique to tackle the snow-cover classification problem has already been verified [4]. However, a drawback of the method is the need to choose suitable values for the threshold parameters involved in the binary subproblems. Due to the lack of training samples, these thresholds should typically be set by the user by empirical “trial-and-error” procedures. However, this approach prevents the method from being automatic and may also be time-consuming and human-error prone.

In order to address this drawback, a novel automatic technique is proposed that combines the aforementioned DTC with the Kittler and Illingworth’s minimum-error thresholding algorithm (K&I for short) [5]. K&I is an unsupervised Bayesian thresholding technique that analyzes the histogram of a given (scalar) image in order to compute an optimal threshold value to be applied for binary classification. The method is based on a Gaussian model for the statistics of the pixel intensities

conditioned to each of the two classes to be discriminated and defines a cost function according to the Bayesian rule for the minimum probability of classification error [5, 6]. A functional representing the average cost function is derived and an optimal threshold is obtained by minimizing this functional [5]. The method has been employed in remote-sensing in the context of change-detection on multispectral images [7], proving a higher effectiveness as compared to alternate unsupervised thresholding methods based on discriminant analysis or fuzzy theory [8]. Here, the proposed method automatically optimizes the threshold value to be used in each binary subproblem by applying K&I to the histogram of the scalar image to be thresholded, thus fully automating the DTC technique for snow-cover mapping.

The proposed technique has been tested on an extensive one-year MODIS data set including 1-2 images per day, acquired over the Italian Alpine region. The experimental results are validated both qualitatively (by a visual photo-interpretation of the resulting classification maps) and quantitatively (by comparing with nivometric ground data). The experimental validation suggests the method to accurately discriminate snow-covered areas without requiring training samples for the considered classes. A visual comparison also suggests the proposed technique to provide more accurate results, at least in the considered geographical area, as compared with the standard MODIS global snow-cover product. A further evaluation of the results has also been carried out by integrating the snow-cover maps into snowmelt models and by deriving the snow contribution in the runoff process [1]. We remark that the proposed method is fully automatic and exhibits very short computation times. Indeed, thanks to the histogram-based formulation of K&I, the execution time of the automatic threshold-optimization process is also nearly independent of the image size.

11. REFERENCES

[1] T. A. Fontaine, T. S. Cruickshank, J. G. Arnold, R. H. Hotchkiss, "Development of a snowfall-snowmelt routine for mountainous terrain for the soil water assessment tool (SWAT)," J. Hydrol., 262, 209-223, 2002.

[2] D. Landgrebe, "Signal theory methods in multispectral remote sensing", John Wiley and Sons, 2003.

[3] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, "Classification and regression trees," Wadsworth International Group, 1984.

[4] P. Malcher, D. Floricioiu, and H. Rott, "Snow mapping in alpine areas using medium resolution spectrometric sensors", Proc. of IGARSS-2003, Toulouse (France), 21-25 July 2003, vol. 4, pp. 2835-2837.

[5] J. Kittler, J. Illingworth, "Minimum error thresholding", Pattern Recognition, vol. 19, no. 1, pp. 41-47, 1986.

[6] R. O. Duda, P. E. Hart, D. G. Stork, Pattern classification, Wiley, 2001.

[7] F. Melgani, G. Moser, S. B. Serpico, "Unsupervised change detection methods for remote sensing images", Optical Engineering, vol. 41, no. 12, pp. 3288-3297, 2002.

[8] Z. Chi, H. Yan, T. Pham, "Fuzzy algorithms: with applications to image processing and pattern recognition," World Scientific, 1996.