PERFORMANCE ASSESSMENT OF SPATIO-TEMPORAL MAP-MRF CLOUD DETECTION FROM SEVIRI IMAGES

Paolo Addesso, Roberto Conte, Maurizio Longo, Rocco Restaino and Gemine Vivone

University of Salerno, D.I.E.I.I, Fisciano, Italy; {paddesso, rconte, longo, restaino, gvivone}@unisa.it

ABSTRACT

Cloud detection is an important preliminary step in most earth observation procedures, such as fire, sea and urban areas satellite monitoring. In this paper we elaborate on cloud detection from sequences of remotely sensed images with an aim to improve the accuracy of cloud masks by taking into account the temporal relationship between subsequent images. In particular, the estimated positions of cloud volumes can be used for constructing an additional temporal prior term within the Maximum A posteriori Probability - Markov Random Field (MAP-MRF) approach, that, classically, takes into account only the spatial correlation among neighboring pixels. In this perspective, FInite Set Statistics Theory (FISST) has been proven to be a powerful framework for multi-target tracking of clouds, whose number changes across images due to birth, death, merging and splitting phenomena. Its computational effort can be mitigated through the Probability Hypothesis Density (PHD) approximation, that permits viable implementations based on Sequential Monte Carlo (SMC) methods. The performance improvements achievable through this sequential Bayesian approach with respect to simpler algorithms, such as the region-matching one, are here assessed on simulated cloud masking superposed on real images acquired by the SEVIRI sensor.

INTRODUCTION

An accurate estimation of cloud dynamics by means of remotely sensed images is a very important source of information in sounding properties of the atmosphere (1). On the contrary, other applications can be heavily afflicted by the presence of clouds (2). Therefore an effective pre-processing of the remotely sensed scenes devoted to the *cloudy/non cloudy* pixel classification is of paramount importance, thus attracting a large number of scientific contributions (see (3) and the references therein).

The joint use of the Maximum A Posteriori Probability (MAP) framework, which minimizes the probability of pixel misclassification, and Markov Random Fields (MRF) (4), that take into account the spatial relationship among the neighboring pixel labels, has been proven effective for clouds detection (5). In addition, temporal correlation among successive images has been also considered (3). In (6) a *penalty term* properly accounts for previous acquisitions to improve the classification of the actual image. In this scenario, a crucial step is the propagation of the label information across the image sequence, that is performed by means of a Probability Hypothesis Density (PHD) filter. It relies upon the FInite Set Statistics Theory (FISST) and consists in considering only the first moment distribution (or intensity) of the posterior distribution (7). The effectiveness of this step depends on the particular motion model used to capture the targets (i.e. the cloud masses) dynamics, but also on several parameters that are strictly related to the particular implementation of the PHD filter, and that jointly affect the classification performances and the computational effort. Thus in this work we perform a simulation analysis in order to explore a wide range of possible algorithm settings. As a final step, we evaluate the classification algorithm on a real scenario by using a sequence of images acquired by the SEVIRI sensor.

The paper is organized as follows: in the next section we briefly introduce the tracking algorithm based on the PHD filter, and its approximation via Sequential Monte Carlo (SMC) implementation (8). Moreover we detail the proposed method for embedding the temporal information inside the MAP classification framework, with particular focus on the MRF-based techniques. Then we

present the performance evaluation, carried out both on simulated data and real images. Finally, in the last Section we report final considerations and suggest further lines of research arising from this study.

PROPOSED METHOD

The method is based on assigning an a priori probability that a given pixel be cloudy, derived from tracking cloud masses across the previous images. Practical implementation of this scheme rely on algorithms able to follow a variable number of targets, as the number of clouds may change due to events as births, deaths, splits and fusions. Such so called Multi Target Tracking (MTT) algorithms involve several peculiar problems, related to the statistical models of said events and to the association of observations to targets. Several approaches are present in the literature, as surveyed, for example, in (9).

MultiTarget Tracking based on Random Finite Sets

A successful framework for MTT problems consists in modeling the target state at instant *k*

$$X^{(k)} = \left\{ \mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, ..., \mathbf{x}_{M(k)}^{(k)} \right\}$$

as a Random Finite Set (RFS) that associates a set of variable finite cardinality to the elements of a probability space. In tracking applications the RFS is constituted by the chosen features (as, for example, position and velocity) of the targets present in the scene. In particular, for dealing with the non-point nature of clouds, we represent them as circumscribing rectangles (commonly named as Bounding Boxes – BB in the image and video processing literature), thus synthesizing the peculiar characteristics through the position and velocity of the center and the two length of sides. Accordingly, in our two-dimensional formalization of the problem, each vector of the RFS contains the six components

$$\mathbf{x}_j^{(k)} = [\, c_{j,1}^{(k)}, \dot{c}_{j,1}^{(k)}, c_{j,2}^{(k)}, \dot{c}_{j,2}^{(k)}, w_j^{(k)}, h_j^{(k)} \,].$$

At instant k the elements of $X^{(k)}$ are the union of the survived $T^{(k)}\left(X^{(k-1)}\right)$, targets the spawned targets $S^{(k)}\left(X^{(k-1)}\right)$ and the newly generated targets $\Gamma^{(k)}$.

The FISST framework permits to introduce a further degree of randomness also in the set of observations. In fact the latter is not restricted to contain just a single observation for each target, but, rather, it can include missed and false measurements. This naturally occurs in the real practice where the available data are unavoidably corrupted by clutter.

In principle, the Maximum A posteriori Probability estimation of the state trajectory can be recursively implemented (7); however this procedure is quite impractical and more viable alternatives are in order. One method consists in the Bayesian sequential estimation of the Probability Hypothesis Density (PHD) of $X^{(k)}$, that accounts for the intensity of the Random Finite Set (7). It has the property that, when integrated over a region R, yields the expected number of targets present in R, and thus can be exploited to extract the needed positions of clouds. In particular, a robust implementation of the Bayesian estimator for the PHD involves a Monte Carlo approximation (8), based on particle filters (PF). In this case the PHD $D_{k|k}(\mathbf{x})$ of the state posterior distribution at time *k* is represented by a set of support points (or *particles*) $\left\{\mathbf{x}_i^{(k)}\right\}_{i=1,...,L_k}$ and respective weights $\left\{\omega_i^{(k)}\right\}_{i=1,...,L_k}$, namely it can be written as

$$D_{k|k}(\mathbf{x}) \approx \sum_{i=1}^{L_k} \omega_i^{(k)} \delta(\mathbf{x} - \mathbf{x}_i^{(k)})$$

The set of L_k particles at time *k* is composed by the L_{k-1} particles present at time *k*-1, propagated according to the target dynamical models, and J_k particles representing the newborn

targets. The weights $\{\omega_i^{(k)}\}_{i=1,\ldots,L_k}$ are obtained by evaluating the corresponding likelihoods, given the current observations (see (6,9) for details).

Pixel classification

The separation of pixels belonging to cloud volumes from the clear sky is a binary hypothesis testing problem. Its solution, under the minimum error probability criterion, encompasses the comparison of the posterior density of the two hypotheses, given the available data. The classification of a single pixel should not be performed independently from the others, being the images strongly correlated. Accordingly, the MAP estimator maximizes the *joint* a posteriori probability of the label set $\mathbf{l} = \{l_i\}_{i\in 1,\ldots,I}$, where I is the number of pixels, and thus is the solution of

$$\hat{\mathbf{l}}_{MAP} = \arg\max_{\mathbf{l}} p(\mathbf{l}\,|\,\mathbf{d}) = \arg\max_{\mathbf{l}} p(\mathbf{d}\,|\,\mathbf{l})\,p(\mathbf{l}),$$

in which $p(\mathbf{d}\,|\,\mathbf{l})$ is the *likelihood* of the image $\mathbf{d} = \{d\}_{i=1,\ldots,I}$ and $p(\mathbf{l})$ the *a priori probability* of the labels. The latter incorporates the available information regarding the label distribution within the illuminated scene and is commonly used to model the spatial dependencies. A successful statistical model is based on Markov Random Fields (MRF) that corresponds to a Gibbs density for the labels joint pdf (4)

$$p(\mathbf{l}) \propto \exp\left(-\frac{U(\mathbf{l})}{T}\right).$$

$U(\mathbf{l})$ is the *energy function* and includes interactions among groups of pixels, or *cliques*. The widely employed Ising statistical model encompasses only the cliques composed by two adjacent pixels, and thus the energy includes second order terms of the kind

$$V_2(i,j) = \beta_s \delta(l_i, l_j),$$

in which β_s is the (spatial) interaction coefficient, $\delta(x,y)$ is the Kronecker delta function and the indexes i and j range on the whole image and within a neighborhood N_i of pixel I, respectively. Common choices for N_i is the set of adjacent pixels in the up, down, left and right directions (4-neighborhood) or the set of all surrounding pixels (8-neighborhood).

However, the a priori probability $p(\mathbf{l})$ is strongly influenced also from the labels distribution at previous instants. We propose to take into account this correlation by adding a term derived from the tracking step of past images. This *penalty function* $v = \{v_i\}_{i=1,\ldots,I}$ is obtained from the bounding boxes positions estimated at the current time. The most straightforward method (see (10) for more sophisticated options) consists in associating, to each bounding box, a rectangle with the same center and with sizes augmented by a factor $A > 2$. A triangular function that decreases from at the center of the rectangle to 0 at the perimeter quantifies the required penalty term that is thus guaranteed to assume value greater than 0.5 inside the bounding box and smaller than 0.5 outside. Possible ambiguities in the definition of the penalty term are resolved by taking the maximum of the potential values. When including this term into the a priori probability a desirable opportunity is offered from preserving the Gibbs distribution form that allows for efficient implementations of the search (4). Thus we choose to insert the penalty function as a further term into the a priori energy

$$V_1(f_i) = \beta_t[1 - v_i]\delta(l_i,0) + \beta_t v_i \delta(l_i,1)$$

where β_t is the (temporal) interaction coefficient, that allows to balance the temporal and the spatial interaction contributions. The final expression for the a priori energy is thus the following

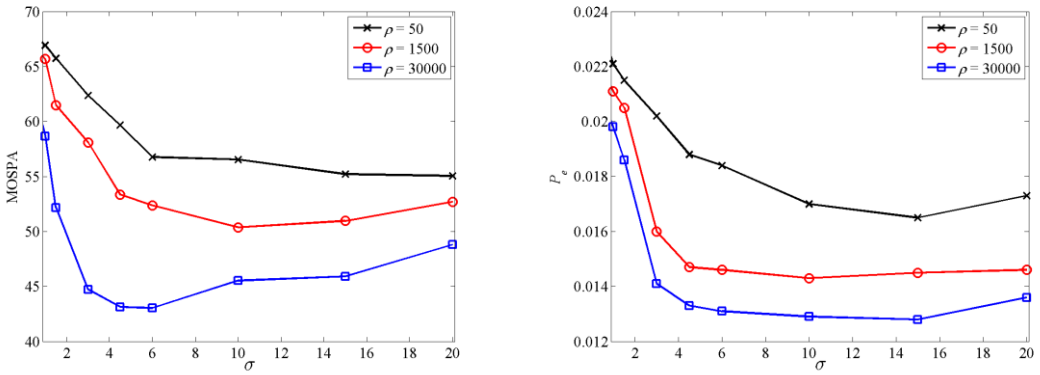
$$U(l) = \sum_{i=1}^{I} V_1(f_i) + \sum_{i=1}^{I} \sum_{N_i} V_2(l_i, l_j) = \sum_{i=1}^{I} \beta_t [1 - v_i] \delta(l_i, 0) + \beta_t v_i \delta(l_i, 1) + \sum_{i=1}^{I} \sum_{N_i} \beta_s \delta(l_i, l_j).$$

EXPERIMENTAL RESULTS

We present the results achieved on synthetic images generated by a properly developed cloud simulation software. Clouds shapes and radiometry are simulated by means of a two dimensional fractional Brownian motion (fBm), as fractals are well suited to mimic natural surfaces. In particular we set the Hurst coefficient H to 0.65 as it is typical for clouds (11). The cloud objects, that are identified as the regions of the fBm that exceed a suitable threshold, are superposed on a background SEVIRI image. Then they evolve according to the well-known *constant velocity model* (9), that is characterized by an acceleration Gaussian noise with standard deviation σ_S, and by a sampling time T.

Our purpose is to investigate the tracking algorithm effectiveness. So we analyse the consequences of a mismatch between the cloud motion model used in the simulator (i.e. the state of the nature) and the one used in the algorithm (that is characterized by a standard deviation σ). In particular, in the generation of synthetic images we employ a linear state equation with a white Gaussian acceleration noise with standard deviation $\sigma_S = 3$ pixels/(sample interval)^2 (in the following we will use the abbreviation p/s$_i^2$). We test the robustness of the cloud tracking algorithms with respect of variation of σ, that ranges from 0.5 to 20 p/s$_i^2$. The pixel values are modeled as a product of independent Laplace densities.

In order to assess the tracking performances, we use the Optimal SubPattern Assignment (OSPA) metric, introduced in (12), that is able to quantify estimation errors for MTT algorithms. In Fig. 1(a) we show the Mean OSPA (MOSPA) of order $p = 2$ and cut-off parameter $c = 100$ computed for three different numbers of particles per target ρ (i.e. 50 , 1500, 30000) on 50 iterations, each



(a) *(b)*

Figure 1: MOSPA (computed with $p = 2$ and $c = 100$) (left plot) and Error rate P_e (right plot) Vs. acceleration noise standard deviation σ measured in p/s$_i^2$. For each plot the three curves are computed for a different number of particles ρ for each target and for a simulated dynamic with $\sigma_S = 3$ p/s$_i^2$.

composed by 20 frames. It is evident that the use of a value of σ greater than σ_S improves the performances mainly for small values of ρ while, when the latter parameter increases, values closer to σ_S also become influential. To tie the performances of the tracking phase and the accuracy of the overall cloud classification procedure, in Fig. 1(b) we present the misclassification rate P_e that is, in fact, shown to be characterized by the same behaviour.

In order to better explore the effect of particles, we compute the error rate P_e for a variable number ρ of particles per target (see Table 1), when the motion model is matched to the ground truth dynamic, i.e. when $\sigma = \sigma_S = 3$ p/s$_i^2$. As expected, the performances benefit from the increase of ρ but also exhibit a saturation behaviour: the improvement decreases with increasing ρ . On the other side one should account for the computational load of the tracking phase which grows linearly with ρ

Table 1: Error rate P_e vs number of particles ρ (the noise standard deviation σ and the simulated dynamic σ are equal to 3 p/s$_i^2$).

ρ	50	300	1500	7500	30000	150000
P_e	2.024 E-2	1.788 E-2	1.602 E-2	1.476 E-2	1.406 E-2	1.375 E-2

Finally, the algorithm is tested on one real SEVIRI time series. The dataset represents a sequence of Central Italy images acquired on 07/21/2008 from 12:15 am to 13:30 am. The tracking parameters for this real case are chosen as a reasonable compromise between classification performances and computational effort.

The detection phase is tuned by setting the temporal interaction coefficient β_t to 35, while the spatial interaction coefficient β_s is estimated using a Least-Squares supervised technique although others approaches, such as Pseudo-Likelihood estimators, can be used as well (4). The result of the proposed algorithms is shown in Fig. 2. In particular in Fig. 2(a) the original image (i.e. the last acquisition of the time sequence) to be classified is presented, while in Fig. 2(b) a “naked-eye” ground truth is shown. The classification is presented in Fig. 2(c): false alarms are shown as green pixels and missed detection are red ones. The overall error rate P_e is 0.1.

(a) *(b)* *(c)*

Figure 2: Classification using the Central Italy dataset: (a) the original image; (b) the corresponding ground truth; (c) the classification results.

CONCLUSIONS

In this paper we report an evaluation of a multitemporal MAP-MRF cloud detection algorithm that exploits the time correlation among subsequent acquisitions by means of a powerful tracking phase based on a PHD-SMC technique. Both tracking and classification performances have been evaluated for different settings of the process parameters. Then an example of classification performed on a time sequence of real SEVIRI images has been shown. Future developments of this technique will be focused mainly on the particle clustering phase that still remains an issue. Investigations on the use MTT algorithms for other applications, such as the oil slicks monitoring, are also planned.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. P.G. Marchetti and the Service Support and Ground Segment Technology Section at ESA-ESRIN, Frascati (Italy) for many helpful discussions and for providing the SEVIRI data.

REFERENCES

1 Gómez-Chova L, G Camps-Valls, J Calpe-Maravilla, L Guanter & J Moreno, 2007. Cloud-screening algorithm for ENVISAT/MERIS multispectral images. IEEE Trans. Geosci. Remote Sens., 45 (12): 4105–4118

2 Cadau E & G Laneve, 2008. Improved MSG-SEVIRI images cloud masking and evaluation of its impact on the fire detection methods. In 28th IEEE Geoscience and Remote Sensing Symposium (IGARSS, Boston), 2: 1056–1059.

3 Papin C, P Bouthemy, & G Rochard, 1993. Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatiotemporal labeling approach. IEEE Trans. Geosci. Remote Sens., 40 (1): 104–114

4 Li S, 2009. Markov Random Field Modeling in Image Analysis, 3rd Ed. (Springer), 323 pp.

5 Le Hégarat-Mascle S & C André, 2009. Use of Markov Random Fields for automatic cloud/shadow detection on high resolution optical images, ISPRS J Photogramm, 64(4): 351-366

6 Addesso P, R Conte, M Longo, R Restaino & G Vivone, 2012. MAP-MRF Cloud Detection Based on PHD Filtering. IEEE J. Sel. Topics Appl. Earth Observ., 5 (3): in press

7 Mahler R, 2007. Statistical Multisource-Multitarget Information Fusion, (Artech House), 888 pp.

8 Vo B-N, S Singh & A Doucet, 2005. Sequential Monte Carlo methods for multi-target filtering with random finite sets. IEEE Trans. Aerosp. Navig. Electron., 41 (4): 1224–1245

9 Bar-Shalom Y, X R Li & T Kirubarajan, 2001. Estimation with Applications to Tracking and Navigation. (Wiley-Interscience), 584 pp

10 Addesso P, R Conte, M Longo, R Restaino & G Vivone, 2011. Cloud Detection in SEVIRI Images by Non-Homogeneous Temporal Markov Random Field. In: 31th EARSeL Symposium (EARSeL, Prague), 255-265

11 Hentschel H G E & I Procaccia, 1984. Relative diffusion in turbulent media: The fractal dimension of clouds, Phys. Rev. A, 29 (3): 1461-1471

12 Schuhmacher D, B-T Vo & B-N Vo, 2008. A consistent metric for performance evaluation of multi-object filters. IEEE Trans.Signal Process., 56: 3447–3457