ISPRS Journal of Photogrammetry and Remote Sensing 64 (2009) 351–366

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Use of Markov Random Fields for automatic cloud/shadow detection on high resolution optical images

Sylvie Le Hégarat-Mascle [a,*], Cyrille André [b]

[a] IEF/Université Paris-Sud, 91405 Orsay Cedex, France
[b] CETP/IPSL, 10-12 av. de l'Europe, 78140 Vélizy, France

ARTICLE INFO

Article history:
Received 20 August 2007
Received in revised form
6 December 2008
Accepted 19 December 2008
Available online 13 March 2009

Keywords:
Markov Random Fields
Object detection
Cloud
Shadow

ABSTRACT

In this study, we propose an automatic detection algorithm for cloud/shadow on remote sensing optical images. It is based on physical properties of clouds and shadows, namely for a cloud and its associated shadow: both are connex objects of similar shape and area, and they are related by their relative locations. We show that these properties can be formalized using Markov Random Field (MRF) framework at two levels: one MRF over the pixel graph for connexity modelling, and one MRF over the graph of objects (clouds and shadows) for their relationship modelling. Then, we show that, practically, having performed an image pre-processing step (channel inter-calibration) specific to cloud detection, the local optimization of the proposed MRF models leads to a rather simple image processing algorithm involving only six parameters. Using a 39 image database, performance is shown and discussed, in particular in comparison with the Marked Point Process approach.

1. Introduction

High resolution optical remote sensing images (such as SPOT/HRVIR) are often affected by cloud presence. For surface studies such as vegetation monitoring, change detection or land cover/land use analysis, these clouds appear either as some noise or some erroneous measurements that conceal or distort the information corresponding to the surface. Moreover, even if they represent only a small percentage of the scene surface, this proportion may be not negligible with regard to the rate of the studied phenomena, such as the land cover change. Hence, even if it is generally not possible to retrieve the missing data, it is important to identify the clouds and their shadows in order to not consider their signals on the studied area.

A large number of cloud detection methods have already been proposed. However, these methods are generally dedicated to data with spatial resolution of about one kilometer square, such as the NOAA/AVHRR images. Indeed, these approaches are based on the high temporal (Cihlar and Howarth, 1994) and spectral (Rossow and Garder, 1993; Chen et al., 2002) resolutions of that kind of data. Dealing with higher spatial resolution, the fourth component of the 'Tasseled Cap' transform (Kauth and Thomas, 1976) was found to be a good indicator of the presence of mist or clouds (Richter, 1996). However, this orthogonal transformation of the spectral bands is not optimized for cloud detection, since it was developed to distinguish the radiometric contribution of vegetation from those of bare soil. Then, recent work (Zhang et al., 2002) proposed an extension to derive a mist indicator and perform pixel radiometric correction. Concerning the shadow detection, different approaches have also been proposed. Some are based on the projection of the cloud shapes on the surface knowing the sun direction and the cloud altitude (Simpson and Stitt, 1998). Other approaches exploit the matched filter concept (Richter and Muller, 2005). The matched filter is then evaluated using the spectral band covariance matrix. Once more when the aim is the correction of the radiometric signal, previous methods are exploited in collaboration with a radiative transfer model. Finally, Ho and Zhenlei (1996) proposed comparing clouds and shadows to perform mutual validation of their detections.

Our work exploits several of the previously presented ideas. They were adapted and combined in order to derive a cloud/shadow detection method which is both robust and automatic. Relatively to classical threshold techniques, the proposed method exploits three main features of clouds (and shadows):

- P1: Clouds and shadows are connex objects;
- P2: Knowing the geometry of the acquisition and the sun location, the image location of the shadow of a cloud is known, but for one parameter (the cloud altitude);
- P3: Each cloud and its associated shadow have the same shape and area (but for the deformations due to relief).

* Corresponding author. Tel.: +33 1 69 15 40 36; fax: +33 1 69 15 40 00.
E-mail address: sylvie.le-hegarat@ief.u-psud.fr (S. Le Hégarat-Mascle).

0924-2716/$ – see front matter © 2009 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.isprsjprs.2008.12.007

In this study we show that these properties (P1, P2, and P3) can be formalized using Markov Random Field (MRF) framework at two levels: one MRF over the pixel graph for P1 modelling, and one MRF over the graph of objects (clouds and shadows) for P2 and P3. In the case of the object MRF, the graph nodes represent the objects, whereas the graph edges model the interactions between objects. Hence, this approach assumes the a priori knowledge of the number of objects and their interactions. Practically, this implies an initial over-detection of clouds and shadows that may have an impact on the final result. Therefore we have compared our approach to a model in which the number of objects and relationships between objects may be a priori unknown, namely the Marked Point Processes (MPP). However, this last approach is also much heavier in terms of optimization process and computation time. Besides, since it uses a global optimization, it demands the precise setting of the model parameters, whereas the local optimization of our first model (MRF) allows avoiding such a fastidious process.

The rest of the article is organized as follows. Section 2 presents the considered SPOT/HRVIR image features and database acquired in the AMMA (African Monsoon Multidisciplinary Analysis) program framework. Section 3 presents the model based on MRF (whereas MPP main concepts and model are presented in the Appendices). Section 4 describes the implementation of our model, in particular specifying the used observation fields (derived from multispectral satellite measurements), the assumed interactions (derived from acquisition and scene geometries) and the algorithm. Section 5 shows the obtained results, and Section 6 gathers our conclusions.

2. Study context and database AMMA/SPOT

2.1. AMMA database

The important inter-annual variability of the monsoon in West Africa is a phenomenon – with sometimes dramatic consequences – known and observed for several decades. However, this variability still raises a large number of questions both about the involved physical processes and about their social and economic consequences. The research program AMMA (African Monsoon Multidisciplinary Analysis) has hence two aims (Redelsperger et al., 2006). On the one hand, it tries to improve the comprehension of the monsoon in West Africa and its impact on the biosphere both at global and at local scales. On the other hand, it looks for the relationships between the climatic variability and the problems of health, water resources and food safety.

AMMA includes four interacting scales of observations. The larger one, the global scale, deals with the interactions between monsoon phenomenon and the remainder of the Earth. The monsoon process scale is the regional scale. The meso-scale deals with the interactions between atmosphere and watershed hydrology. At this scale, three study sites have been selected within West Africa, namely the Ouémé watershed in Benin, the Gourma one in Mali, and the Hapex square degree in Nigeria. Finally, over each of these sites, some 'super-sites' have been selected for studies and measurements at local scale studying the impacts of the climate on agriculture and antropic activities and the associated retroactions. At this scale, the characterization and the monitoring of surface state from remote sensing data require high spatial resolution sensors (pixel size of about few tens of meters). This requirement is fulfilled by SPOT4/HRVIR images. The whole database includes 39 SPOT/HRVIR scenes, whose acquisition dates and places are given in Table 1.

Finally, for algorithm performance evaluation, some cloud masks have been obtained by photointerpretation. In the lack of more objective data, these masks have been used as 'ground truth' to estimate the performance of the proposed automatic image processing method.

2.2. SPOT/HRVIR image features

The SPOT satellites, dedicated to the observation of the emerged surfaces, have quasi-polar, circular and heliosynchronous orbits at 832 km altitude. The fourth one (SPOT4) has two identical optical sensors HRVIR (High Resolution Visible and InfraRed on board), having a 60 km swath and acquiring panchromatic or multispectral images. In this last acquisition mode, measurements are performed in four bands (corresponding to wavelengths from green, to mid-infrared through red and near-infrared) with pixel size equal to 20×20 m^2.

As for any visible/infrared optical sensor, the acquired images may be affected by the presence of clouds. We propose a generic automatic cloud/shadow detection method. The proposed one is generic and would be applicable to any image provided that the basic assumptions (P1, P2, and P3) are valid. First, note that the separation of clouds (respectively shadows, respectively mist) from the remainder of the image is a non-trivial problem in most application cases. In particular, it is generally not possible to determine a decision threshold by simple analysis of the image histogram. Fig. 1 points out this difficulty. It compares, for each of the four spectral bands, the histograms of the pixels of soil and those of clouds in the case of a typical scene (extracted from the image acquired at 04/29/2006 over Benin). Whatever the spectral band, there is an important overlapping between soil and cloud histograms, inducing important rates of false alarms or misdetection regardless of the chosen threshold value. The only case where a threshold approach would be efficient is when the images only present completely opaque clouds. Besides, note that, even in this case, the threshold value can generally not be obtained from image histogram analysis because the cloud pixels are minor relative to the soil pixels.

3. Markov Random Field modelling

In this section, we recall the main concepts of the Markov Random Fields (MRF). We also introduce the notations used in the following sections. In the following section the general model is presented whereas its practical implementation described in Section 4.

MRFs are widely used in image processing (Abend et al., 1965; Geman and Geman, 1984) providing a solution to the causality problem. Nowadays, research is still active in this domain. Some work has focused on the development of efficient optimization techniques, in particular using graph cut methods (Boykov et al., 2001; Boykov and Kolmogorov, 2004; Kolmogorov and Zabih, 2004). Some other approaches aim at considering models more and more flexible, in particular relaxing the stationarity assumption of the image (Benboudjema and Pieczynski, 2005; Pieczynski and Benboudjema, 2006; Le Hégarat-Mascle et al., 2007), or models at higher level, in particular using MRF to model the relationships between image objects (Descombes, 2004).

3.1. Markov fields over graphs

Let G be a non-oriented graph. We assume that the edges of G define a neighbourhood system where V_s is the neighbourhood associated to node s, i.e. a set of nodes all having an edge toward s and checking $t \in V_s \Leftrightarrow s \in V_t$. The cliques c associated to V_s are defined as the union of some subsets of V_s and s such that a clique is a complete subgraph of G (constituted by 2×2 nodes considered to be neighbours; for example, Fig. 2 shows the possible cliques containing a pixel s, in 8-connexity over the pixel graph). Then, C denotes the set of cliques c over G.

Table 1
Dates and geographic coordinates of the 39 SPOT/HRV images of the AMMA database (over Nigeria the region of interest is covered by two images at each date, except from 08/28/2005).

Site	Dates (month/day/year)	Geographic coverage	
		Longitude	Latitude
Benin	04/21/2005, 10/30/2005, 01/10/2006, 03/03/2006, 04/08/2006, 04/14/2006, 04/29/2006, 05/25/2006	1°16E– 2°06′E	9°26′N– 10°04′N
Mali	06/05/2005, 07/06/2005, 07/22/2005, 08/06/2005, 08/22/2005, 10/08/2005, 10/23/2005, 01/04/2006, 04/13/2006, 06/04/2006	1°59′W– 1°11′W	15°03′N– 15°39′N
Nigeria	06/06/2005, 06/27/2005, 07/13/2005, 08/03/2005	2°7′E– 3°06′E	13°11′N14°20′N
	08/28/2005	2°7′E– 3°06′E	13°11′N13°50′N
	09/12/2005, 09/28/2005, 10/14/2005, 10/25/2005, 05/05/2006, 05/10/2006	2°7′E– 3°06′E	13°11′N14°20′N

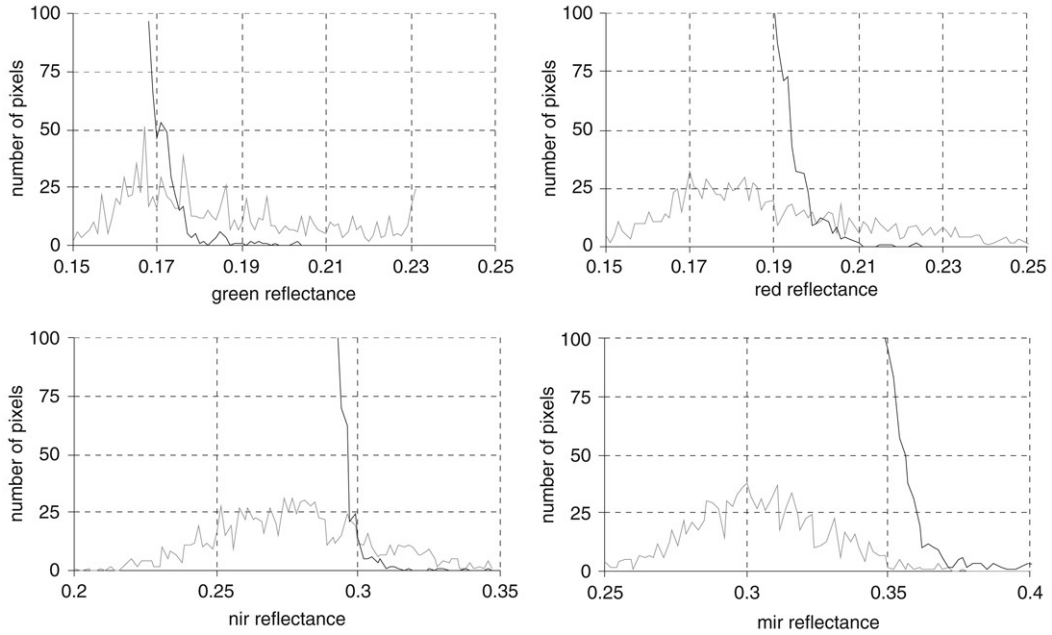

Fig. 1. Comparison of the histograms of the cloud pixels (in grey) and the soil pixels (in black) in the case of the four SPOT4/HRVIR bands. Concerning the soil histograms, only a subpart has been represented (for low reflectance values, the number of pixels takes very high values making the curve out of the figure).

Let X be a random field defined over G, and x a realisation of X. x_s denotes the value of x at the node s of G. x_s is a realisation of the random variable X_s. Then, by definition, X is a Markov field if and only if

$$P\left(X_s = x_s | x_t, \forall t \in G - \{s\}\right) = P\left(X_s = x_s | x_t, \forall t \in V_s\right). \quad (1)$$

If in addition there is no configuration of X with null probability: $\forall x, P(X = x) > 0$, then, according to the Hammersley–Clifford theorem (Besag, 1974) that states the equivalence between a Markov field and a Gibbs field, X is a Gibbs field having potential functions associated to the cliques c of C whose probability is

$$P(X = x) = \frac{1}{Z}.\exp\left(-U(x)\right),$$
$$\text{where } U(x) = \sum_{c \in C} U_c\left(x_s, s \in c\right), \quad (2)$$

and Z is a normalization constant equal to $Z = \sum_x \exp\left(-U(x)\right)$, called the 'partition function'.

Let us now assume that X is hidden and that, for the estimation of x we have at our disposal the realisation y of the random field of the observations Y, also defined over G, and related to X through the conditional probabilities $P(X = x|Y = y)$ and $P(Y = y|X = x)$. The Maximum A Posteriori (MAP) criterion corresponds to the estimation of x that maximizes $P(X = x|Y = y)$: $\tilde{x}_{MAP} = \arg\max_x P(X = x|Y = y)$, or, using the Bayes theorem:

$$\tilde{x}_{MAP} = \arg\max_x P(Y = y|X = x).P(X = x). \quad (3)$$

Finally, assuming first that no configuration of Y has a null probability: $\forall y, P(Y = y) > 0$, and second that the random variables Y_s associated to Y at the G nodes s are independent conditionally to X, then $P(Y = y|X = x) = \exp\left\{-\sum_{s \in G} - \ln\left(P\left(Y_s = y_s|X_s = x_s\right)\right)\right\}$, and the maximization of the a posteriori probability is equivalent to the minimization of the energy:

$$\tilde{x}_{MAP} = \arg\min_x \left[\sum_{s \in G} -\ln P\left(y_s|x_s\right) + \sum_{c \in C} U_c\left(x_t, t \in c\right)\right]. \quad (4)$$

In the following we focus on the minimization of the energy function, and we denote $U_s^0\left(y_s|x_s\right) = -\ln P\left(y_s|x_s\right)$ the 'data attachment' energy.

The success of MRF models is also due to the convenience of the minimization of the global (over the whole graph G) energy, that can be obtained by successively considering the G nodes separately processing each node s: (i) testing a new value x_s, (ii) computing the new s local energy: $U_s\left(y_s|x_s, x_t, t \in V_s\right) = U_s^0\left(y_s|x_s\right) + \sum_{c:s \in c} U_c\left(x_s, x_t, t \in c, t \neq s\right)$, and (iii) deciding in favour of x_s or not depending on the difference between the new value of $U_s\left(y_s|x_s, x_t, t \in V_s\right)$ and its former value. In the case of a local minimization, the tested value x_s is accepted only if it induces a decrease of the local energy. Conversely, the global minimization techniques, such as those based on simulated annealing (Kirkpatrick et al., 1983; Geman and Geman, 1984), allow some energy increases in order to overcome the local minima and reach the global minimum. In both cases, algorithms

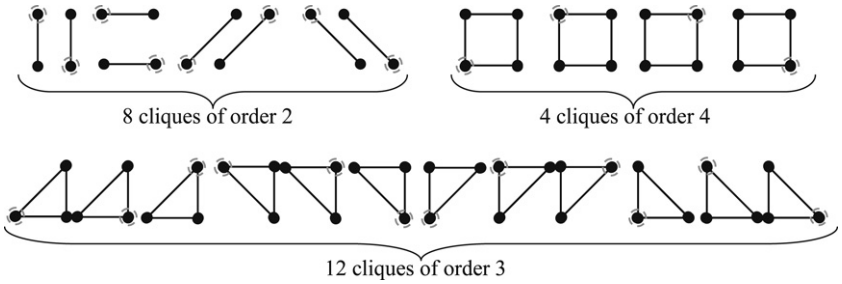

Fig. 2. Cliques over the pixel graph in 8-connexity that contain a pixel s (surrounded pixel).

are necessarily iterative. More recently, non-iterative techniques have been developed that propose efficient solutions to some energy minimization problems encountered in image processing (Greig et al., 1989). In particular, Kolmogorov and Zabih (2004) characterize the energy function that can be minimized by graph cut techniques.

We now describe the two particular cases of G used in the proposed automatic detection algorithm for cloud and shadow over the SPOT/HRVIR images, namely G is either the pixel graph, or the graph of the objects 'cloud', 'shadow' and 'mist'.

3.2. Pixel graph case

Let Ω be the image space, $|\Omega|$ is the cardinality of Ω, i.e. the number of image pixels. X is a random field defined over Ω ($G = \Omega$). x is the label or class image, and y is the data image. x_s denotes the class of a pixel s, i.e. of an element of Ω, and y_s its radiometric value (possibly a combination of different spectral bands, as will be detailed in Section 4.1). X_s is a random variable taking values in the class set, namely the class 'cloud or mist' denoted $\lambda_{c\cup m}$ and its complement denoted $\lambda_{\overline{c\cup m}}$ in the case of cloud detection, and the class 'shadow' denoted λ_{sh} and its complementary denoted $\lambda_{\overline{sh}}$ in the case of shadow detection.

Here, we aim at modelling P1, i.e. a kind of class regularity. Then, we propose the following model for the detection of cloud and mist:

$$\forall s \in \Omega, \quad U_s^0\left(y_s | x_s = \lambda_{c\cup m}\right) = \begin{cases} 0 & \text{if } y_s \geq t_h^c, \\ \beta_1 & \text{if } y_s \in [t_l^c, t_h^c), \\ \beta_2 & \text{if } y_s < t_l^c, \end{cases} \tag{5}$$

$$\forall s \in \Omega, \quad U_s^0\left(y_s | x_s = \lambda_{\overline{c\cup m}}\right) = \begin{cases} 0 & \text{if } y_s < t_l^c, \\ \beta_3 & \text{if } y_s \in [t_l^c, t_h^c), \\ \beta_4 & \text{if } y_s \geq t_h^c, \end{cases} \tag{6}$$

$$\forall s \in \Omega, \quad U_{c=(s,t)}\left(x_s, x_t\right) = \gamma \times \left[1 - \delta\left(x_s, \lambda_{c\cup m}\right) \times \delta\left(\lambda_{c\cup m}, x_t\right)\right], \tag{7}$$

where β_i, $i \in \{1, 2, 3, 4\}$, γ, t_l^c and t_h^c are parameters such that $\forall i \in \{1, 2, 3, 4\}\ \beta_i > 0$, $0 < \beta_1 - \beta_3 < \gamma < \beta_2/n_{connex}$, and $t_l^c \leq t_h^c - t_l^c$ and t_h^c correspond to thresholds high and low, respectively. Here, we only consider the cliques of order 2, n_{connex} is the neighbourhood connexity: 4 or 8. $\delta(i, j)$ equals 1 if $i = j$, and 0 otherwise. According to (7), the energy of a clique c is not null (equal to γ) except if its two pixels are labelled $\lambda_{c\cup m}$.

The data attachment energy piecewise constant functions (5) and (6) are rather basic. They were chosen in order to establish an equivalence, in terms of label decision, with the approaches based on either hysteresis thresholding or region growing. The label decision equivalence between our MRF model at pixel level and hysteresis thresholding is shown by Table 2. Indeed, for a given pixel s, Table 2 gives the energy values associated to each of the hypotheses $\lambda_{c\cup m}$ and $\lambda_{\overline{c\cup m}}$ depending on the value y_s relatively to the threshold parameters t_l^c and t_h^c and on the presence of a pixel labelled $\lambda_{c\cup m}$ in the s neighbourhood V_s. Grey cells point out the decisions in the case of the inequalities $\forall i \in \{1, 2, 3, 4\}\ \beta_i > 0$, $0 < \beta_1 - \beta_3 < \gamma < \beta_2/n_{connex}$. We see that they are identical to the decisions that would be given by a thresholding method using hysteresis between thresholds t_l^c and t_h^c. Thus, provided that the inequalities $\forall i \in \{1, 2, 3, 4\}\ \beta_i > 0$, $0 < \beta_1 - \beta_3 < \gamma < \beta_2/n_{connex}$ are verified, the local minimization of the global energy corresponds, in terms of label decision, to a threshold with hysteresis. Hence it is not necessary (assuming local minimization is sufficient) to carefully set the values of parameters β_i, $i \in \{1, 2, 3, 4\}$, and γ.

The case of the shadow detection is processed similarly with three main differences: the considered classes, namely 'shadow' λ_{sh} and its complementary $\lambda_{\overline{sh}}$, the sense of inequality, and the consideration of the spatial non-stationarity through a subset ω of Ω ($\omega \subseteq \Omega$):

$$\forall s \in \Omega, \quad U_s^0\left(y_s | x_s = \lambda_{\overline{sh}}\right) = \begin{cases} 0 & \text{if } y_s \geq t_h^{sh}, \\ \beta_3' & \text{if } \left(y_s \in [t_l^{sh}, t_h^{sh}) \text{ and } s \in \omega\right) \\ & \text{or } \left(y_s < t_h^{sh} \text{ and } s \notin \omega\right) \\ \beta_4' & \text{if } y_s < t_l^{sh} \text{ and } s \in \omega, \end{cases} \tag{8}$$

$$\forall s \in \Omega, \quad U_s^0\left(y_s | x_s = \lambda_{sh}\right) = \begin{cases} 0 & \text{if } y_s < t_l^{sh} \text{ and } s \in \omega, \\ \beta_1' & \text{if } \left(y_s \in [t_l^{sh}, t_h^{sh}) \text{ and } s \in \omega\right) \\ & \text{or } \left(y_s < t_h^{sh} \text{ and } s \notin \omega\right) \\ \beta_2' & \text{if } y_s \geq t_h^{sh}, \end{cases} \tag{9}$$

$$\forall s \in \Omega, \quad U_{c=(s,t)}\left(x_s, x_t\right) = \gamma' \times \left[1 - \delta\left(x_s, \lambda_{sh}\right) \times \delta\left(\lambda_{sh}, x_t\right)\right], \tag{10}$$

where the β_i', $i \in \{1, 2, 3, 4\}$ and γ' parameters check an analogue condition as the $\beta_{i,i\in\{1,2,3,4\}}$ and γ parameter one. Then, in the shadow case, the local minimisation of the energy is equivalent in terms of label decision to a region growing (rather than a hysteresis threshold) using as germs the pixels included in ω and having a y_s value lower than t_l^{sh} and as merging condition the inferiority to a 'high' threshold, t_h^{sh}. Note that here t_l^{sh} and t_h^{sh} are specific to the region (cf. Section 4.3). Indeed, using a table analogous to Table 2 but replacing the conditions on the y_s value relative to t_l^c and t_h^c in Table 2 by conditions on s belonging to a germ (belonging to ω and value inferior to t_l^{sh}) and value relative to t_h^{sh}, it can be shown that, provided that conditions on $\beta_{i,i\in\{1,2,3,4\}}'$ and γ' parameters are verified, a decision in favour of λ_{sh} occurs either when ($y_s < t_l^{sh}$ and $s \in \omega$) (germ pixel) or when ($y_s < t_h^{sh}$ and $\exists t \in V_s | x_t = \lambda_{sh}$), and decision in favour of $\lambda_{\overline{sh}}$ occurs otherwise.

3.3. Object graph case

Here we assume that the previous processing steps have led to the detection on the one hand of the clouds and mist, and on the other hand of the shadows. G is the object graph such that its nodes

Table 2
Values of the local energy associated to the different cases depending on data and neighbourhood n-connexity; the grey cells point out the minimal value for each case (column), the decisions in favour of $\lambda_{c\cup m}$ correspond to a grey case on the 3rd line, and those in favour of $\lambda_{\overline{c\cup m}}$ to a grey case on the 4th line.

y_s	$<t_l^c$		$\in [t_l^c, t_h^c[$		$\geq t_h^c$	
$\exists t \in V_s / x_t = \lambda_{c\cup m}$	No	Yes	No	Yes	No	Yes
$U_s(y_s \mid x_s = \lambda_{c\cup m}, x_t, t \in V_s)$	$\beta_2 + n.\gamma$	$\in [\beta_2, \beta_2 + (n-1).\gamma]$	$\beta_1 + n.\gamma$	$\in [\beta_1, \beta_1 + (n-1).\gamma]$	$+n.\gamma$	$\in [0, (n-1).\gamma]$
$U_s(y_s \mid x_s = \lambda_{\overline{c\cup m}}, x_t, t \in V_s)$	$+n.\gamma$	$+n\gamma$	$\beta_3 + n.\gamma$	$\beta_3 + n.\gamma$	$\beta_4 + n.\gamma$	$\beta_4 + n.\gamma$

are the different connex objects (derived from pixel level previous detections). y_s being the initial label of the object s, only two values are possible for y_s: 'cloud or mist', $\lambda_{c\cup m}$, or 'shadow', λ_{sh}. x_s being the searched label, four possible x_s values are considered: 'cloud', $\tilde{\lambda}_c$, 'mist', $\tilde{\lambda}_m$, 'shadow', $\tilde{\lambda}_{sh}$, or 'other', $\tilde{\lambda}_{\overline{c\cup m\cup sh}}$ (that corresponds to invalid initial detection).

Our aim here is to model P2 and P3, i.e. the relationship between the classes 'cloud' and 'shadow'. The neighbourhoods are defined in the direction given by the position of the shadow relatively to the projection of the cloud over the surface (cf. Section 4.2). The clique order is equal to 2 (one cloud and its shadow), but the cardinality of the neighbourhood is not set a priori (we only assume a maximal value $|V_s|_{\max}$ to calibrate the other model parameters).

The following model was then proposed:

$$U_s^0\left(y_s | x_s = \tilde{\lambda}_c\right) = \begin{cases} \delta & \text{if } y_s = \lambda_{c\cup m} \text{ and } \Pi \text{ true} \\ +A & \text{otherwise} \end{cases}, \tag{11}$$

$$U_s^0\left(y_s | x_s = \tilde{\lambda}_{sh}\right) = \begin{cases} \delta & \text{if } y_s = \lambda_{sh} \\ +A & \text{otherwise} \end{cases}, \tag{12}$$

$$U_s^0\left(y_s | x_s = \tilde{\lambda}_m\right) = \begin{cases} 0 & \text{if } y_s = \lambda_{c\cup m} \text{ and } \Pi \text{ false} \\ +A & \text{otherwise} \end{cases}, \tag{13}$$

$$U_s^0\left(y_s | x_s = \tilde{\lambda}_{\overline{c\cup m\cup sh}}\right) = \begin{cases} \alpha & \text{if } (y_s = \lambda_{c\cup m} \text{ and } \Pi \text{ true}) \text{ or } (y_s = \lambda_{sh}) \\ +A & \text{if } y_s = \lambda_{c\cup m} \text{ and } \Pi \text{false} \end{cases}, \tag{14}$$

$$U_{c=(s,t)}(x_s, x_t) = \begin{cases} 0 & \text{if } \left(\left(x_s = \tilde{\lambda}_c \text{ and } x_t = \tilde{\lambda}_{sh}\right)\right. \\ & \quad \left. \text{or } \left(x_s = \tilde{\lambda}_{sh} \text{ and } x_t = \tilde{\lambda}_c\right)\right) \text{ and } H \text{ false} \\ -\delta & \text{otherwise} \end{cases}, \tag{15}$$

where α, δ are parameters such that $0 < \alpha < \delta$, A is a constant taking a very large value ($|V_s|_{\max}.\delta < A < +\infty$, A is not infinite to prevent null probability for any configuration), Π is a property allowing the distinction between clouds (Π true) and mist (Π false) and H is a predicate corresponding to the criterion of overlapping of the translated objects. Π and H will be detailed in Section 4.4.1.

Table 3 allows checking that the local minimization of the global energy over G corresponds to the mutual validation of the detections of respectively the clouds and the shadows. Assuming the previous paragraphs ordering relationships between parameters, the grey cells point out the decisions. We see that they are identical to the decisions that would be given by a direct cross-validation between detected clouds and associated shadows. Thus, provided that the inequalities $0 < \alpha < \delta$ and $|V_s|_{\max}.\delta < A$ are verified, the local minimization of the global energy corresponds, in terms of label decision, to a 'mutual validation' of clouds and shadows, and once more as far as local optimization is concerned, the precise tuning of model parameters is not necessary. Finally, note that, according to (15) cloud and shadow are valid by default and they may be invalidated only if H predicate can be estimated, i.e. in particular if a cloud and a shadow have been associated. In other words, H is true by default if it cannot be estimated. It allows dealing with the cases where a cloud masking a shadow prevents the mutual validation and the cases where cloud or shadow is outside the image.

4. Implementation and algorithms

The implementation of previous models requires specifying:

- The choice of the data image: for cloud detection, having inter-calibrated the Green and MIR bands, we use a combination of these bands; for shadow detection, only the MIR band is used;
- The estimation of the direction of the shadow location relatively to the image projected cloud location;
- The setting of the model parameters.

In this section, we first detail these three points, and then we present the global algorithm.

4.1. Observation data

In the proposed model, the observations are taken into account through the 'data attachment' energy term (5)–(6) (8)–(9). First note that there is no reason a priori that the more pertinent observation (i.e. the one inducing the smallest false alarm and non-detection rates) be the same for cloud and for shadow detections. Indeed, the used observation data are different and will induce different performance. However, the key point is, the less the individual performance of each detection considered separately, the more the absence of redundancy between errors. In particular, as far as the individual (cloud or shadow) detections are concerned, the false alarms, provided that they are complementary, are much preferable to the non-detections. Indeed, through the cloud/shadow interaction model on the object graph, the false alarms will be strongly reduced.

For shadow detection, we use the MIR band (Richter and Muller, 2005). Indeed, even if some signal measurements may be confused with the signal measurement over dark soils, we assume these confusions will decrease either considering the limitation to the subset ω of Ω in the pixel graph, or thanks to the mutual validation cloud/shadow in the object graph. In the following we present the processing of the data that will be used for cloud detection.

4.1.1. Cloud case: Approach

Generally, clouds are distinguished from the surface thanks to their reflectance in one or several spectral bands. However, such an approach presents several drawbacks. First, the clouds may not be the sole targets that present high reflectance values in the visible and near-infrared inducing a real difficulty to set a detection threshold. Second, some clouds may be not completely opaque presenting then reflectance values that depend on the soil features below them. These two difficulties are visible in Fig. 3 that shows, in the plane (MIR, Green) the reflectance values of some pixels selected in a supervised way. Similar results have been obtained considering different images.

Considering only one band, we note that the spreading of the measurements of both soil and cloud pixels makes the threshold determination difficult. However, a linear separation of the two

Table 3
Values of the local energy associated to the different cases depending on the data and the neighbourhood; the grey cells point out the minimal value for each case (column), the decision in favour of $\tilde{\lambda}_c$, $\tilde{\lambda}_{sh}$, $\tilde{\lambda}_m$, or $\tilde{\lambda}_{\overline{c\cup m\cup sh}}$ correspond to a grey field on the 3rd, 4th, 5th, or 6th line, respectively.

y_s	$\lambda_{c\cup m}$ and Π true		λ_{sh}		$\lambda_{c\cup m}$ and Π false	
H true?	No	Yes	No	Yes	No	Yes
$U_s(y_s \mid x_s = \tilde{\lambda}_c)$	$+\delta$	$\leq \delta - \delta = 0$	$+A$	$\geq A - \lvert V_s\rvert_{\max}.\delta > 0$	$+A$	$\geq A - \lvert V_s\rvert_{\max}.\delta > 0$
$U_s(y_s \mid x_s = \tilde{\lambda}_{sh})$	$+A$	$\geq A - \lvert V_s\rvert_{\max}.\delta > 0$	$+\delta$	$\leq \delta - \delta = 0$	$+A$	$\geq A - \lvert V_s\rvert_{\max}.\delta > 0$
$U_s(y_s \mid x_s = \tilde{\lambda}_m)$	$+A$	$+A$	$+A$	$+A$	0	0
$U_s(y_s \mid x_s = \tilde{\lambda}_{\overline{c\cup m\cup sh}})$	$+\alpha$	$+\alpha$	$+\alpha$	$+\alpha$	$+A$	$+A$

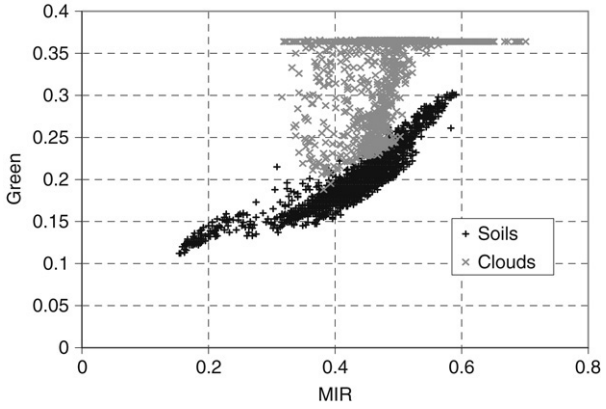

Fig. 3. Reflectance values of pixels respectively corresponding to clouds and soils, in the (Green–MIR) plane.

kinds of signals seems possible. Let us then consider the image y_{diff} such that in each pixel s,

$$y_{diff,s} = y_{G,s} - a \cdot y_{MIR,s} - b, \tag{16}$$

where a and b are the parameters of a line modelling the correlation between MIR and Green reflectance values over the soil, $y_{G,s}$ is the Green s pixel reflectance, and $y_{MIR,s}$ the MIR s pixel reflectance. A similar approach was proposed in Zhang et al. (2002). It allows taking into account the soil reflectance that affects the cloud reflectance in the case of semi-transparent clouds or clouds that do not saturate the signal.

4.1.2. Cloud case: Inter-calibration of the Green–MIR bands

The derivation of the image difference y_{diff} requires the estimation of the line parameters (a, b). We propose an automatic estimation of these latter. Now, the difficulty is due to the presence of some clouds.

In the (MIR, Green) plane, each pixel is characterized by the reflectance values $(y_{MIR,s}, y_{G,s})$. The linear regression of a point cluster in this plane according to a quadratic error minimization criterion takes into account the whole set of points and is then very sensitive to the presence of 'aberrant' points such as the cloud pixels (or the shadow ones) for the estimation of the soil line. The first way to remove automatically these 'aberrant' points would consist in considering, in the (MIR, Green) plane, the bi-dimensional (2D) histogram of the image (that is assumed representing the joint probability) and approximating the relationship between Green and MIR reflectances by the function f giving for each MIR value, $y_{MIR,.}$, the modal value of the 1D histogram of the pixels $(y_{MIR,.}, y_{G.})$:

$$\tilde{y}_{G,.} = f\left(y_{MIR,.}\right) = \arg\max_{y_{G,.}} \left\{P\left(y_{MIR,.}, y_{G,.}\right)\right\}, \tag{17}$$

where $P(t, u)$ is the joint probability of the random variables T and U at $T = t$ and $U = u$, and $y_{G,.}$ and $y_{MIR,.}$ are the Green and MIR reflectance values; since these values are pixel values (not an image), but this pixel may be any pixel, the s reference in the subscript 's' in (16) for example was replaced by a dot '.'.

However, nothing secures the constraint that the obtained function f be a straight line even in first approximation. Moreover, for the 'high' MIR values, even the modal value of the conditional histogram may be distorted by the presence of clouds whose proportion for these 'high' reflectance values may be important. This clearly appears in Fig. 4 that represents, in the case of the shown image, the soil and cloud pixel values and the inter-calibration relationship $f(y_{MIR,.})$ estimation using (17).

Then, in order to improve the estimation robustness, we constrain f to be linear at least in first approximation and we propose then the following 'cost' function that is to maximize:

$$C_N = \max_{(\tilde{y}_{G,.}(i))_{i\in[1,N]}:\forall i\in\{1,\ldots,N-1\},\left|(\tilde{y}_{G,.}(i))-(\tilde{y}_{G,.}(i+1))\right|\leq K} \times \left(\sum_{i=1}^{N} P\left(y_{MIR,.}(i), \tilde{y}_{G,.}(i)\right)\right), \tag{18}$$

where N is the number of classes of the 1D histogram associated to the random variable $Y_{MIR,.}$, $y_{MIR,.}(i)$ is the MIR reflectance of the class i, $\tilde{y}_{G,.} = f\left(y_{MIR,.}\right)$ the i class MIR reflectance given by the function f, $P(t, u)$ the joint probability of t and u, that are random variable realisations, K is a parameter modelling the 'rigidity' constraint set on f (absolute value of the maximal local slope). The C_N maximization consists in the search for the most probable way Γ through the (MIR, Green) plane. Such a way can be evaluated in two iterations using the Viterbi algorithm (Viterbi, 1967) that is also known as the dynamic programming principle. In our case, the partial costs are derived by recurrence as follows:

$$\begin{cases} \forall j \in [1, M], \quad C(1, j) = P\left(y_{MIR,.}(1), y_{G,.}(j)\right), \\ \forall i \in [1, N-1], \quad C(i+1, j) = P\left(y_{MIR,.}(i+1), y_{G,.}(j)\right) \\ \quad + \max_{k\in[-K,K]} (C(i, j+k)), \end{cases} \tag{19}$$

where M is the number of classes of the 1D histogram associated to the random variable $Y_{G,.}$, $y_{G,.}(i)$ the Green reflectance of the class i, and K is the maximal jump accepted (in number of Green histogram classes) between two successive points on Γ (i.e. two successive MIR histogram classes). Once the partial cost is derived, the Γ way is obtained by a backward recurrence:

$$\begin{cases} \Gamma(N) = \max_{j\in[1,M]} (C(N, j)), \\ \forall i \in [1, N-1], \quad \Gamma(i) = \max_{k\in[-K,K]} (C(i-k, \Gamma(i+1))). \end{cases} \tag{20}$$

The constraint set by the maximal jump amplitude K prevents Γ from drifting toward the high Green values even for high MIR values for which clouds are most probable. Finally, Γ is approximated by its linear regression, assumed equal to the researched (MIR, Green) soil line.

The improvement of the f estimation in terms of regularity is obvious from Fig. 4.

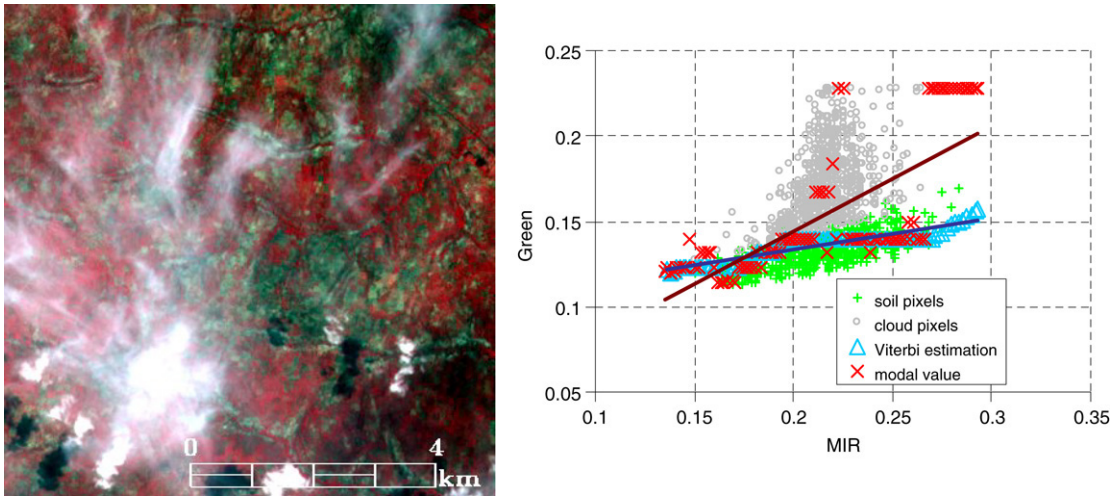

Fig. 4. Comparison for the left image (acquired on 04/29/2006 over Benin) of the results of the Green–MIR band inter-calibration using modal value estimation and the Viterbi estimation (linear approximations are plotted with the corresponding dark colour) respectively.

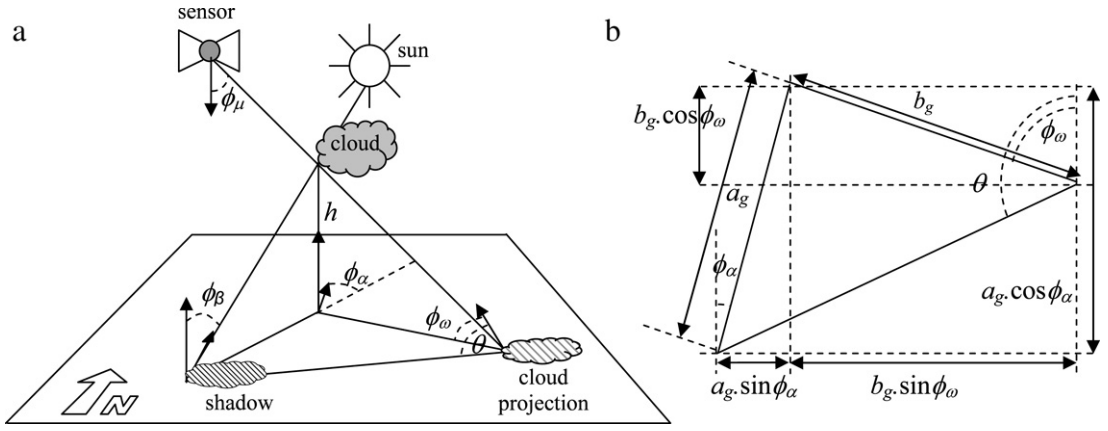

Fig. 5. Sun/Cloud/Shadow geometry, actual and apparent on the remote sensing image: (a) 3D representation, (b) 2D soil projection (with $a_g = h.\tan\phi_\beta$, $b_g = h.\tan\phi_\mu$).

4.2. Estimation of the location of the shadow relative to the cloud

The shadow of each cloud is located in the opposite direction relative to the sun apparent azimuth θ. As illustrated in Fig. 5, this direction depends at once on the azimuth ϕ_α, the sun incidence angle ϕ_β, the looking angle ϕ_μ and the orbit tilt $\pi/2 - \phi_\omega$.

According to Fig. 5 geometry:

$$\tan\theta = -\frac{\sin\varphi_\alpha \tan\varphi_\beta + \sin\varphi_\omega \tan\varphi_\mu}{\cos\varphi_\alpha \tan\varphi_\beta - \cos\varphi_\omega \tan\varphi_\mu}. \quad (21)$$

Note that it is possible at this stage to detail the shadow location only in terms of direction, since the cloud altitude is unknown.

4.3. Model parameter setting

Let us first recall that there is no need to set precisely the parameters $\beta_{i\in\{1,2,3,4\}}$ and γ of Eqs. (5)–(7) and similar parameters of Eqs. (8)–(10) provided that the conditions $\forall i \in \{1,2,3,4\}$ $\beta_i > 0, 0 < \beta_1 - \beta_3 < \gamma < \beta_2/n_{connex}$ (and similar conditions for β_i' and γ' parameters) are met – which will be done in the following – and as far as local minimization is concerned. Indeed, the local minimization of the global energy corresponds to a threshold with hysteresis which only depends on the values of the threshold parameters. Note that the global optimum is a priori dependent on Markovian model parameters (β_i, β_i', $i \in \{1,2,3,4\}$, γ, γ') and one should first set these parameters (whatever the way the global optimization is then performed). From our point of view and for this study, the advantage of avoiding the step of estimation of Markovian model parameters prevails upon the drawback of getting a local optimum instead of a global one.

Concerning the 'threshold' parameters involved in the data attachment terms, we distinguish the cloud and the shadow cases. In the cloud detection case, these parameters, denoted t_h^c and t_l^c, are stationary over the image and directly estimated from the image histogram analysis. In the shadow detection case, these parameters are local to the estimated 'shadow' object.

More precisely, t_h^c and t_l^c are estimated from the y_{diff} pixel distribution. Indeed, we observe that the cloud shadows do not really affect the y_{diff} distribution, and in the absence of clouds the distribution is symmetric. The thresholds are hence parameterized using the histogram's left part (values lower than the median value) assumed to be 'non-contaminated' by some possible clouds. Let z_p be the p-percentile (i.e.$P(y_{diff} \le z_p) = p/100$). t_p is the z_p symmetric relative to the distribution mean (that is very close to 0). Then, t_h^c and t_l^c are such that: $t_h^c = c_h \cdot t_p$, and $t_l^c = c_l \cdot t_p$, where c_h and c_l are two empirical coefficients such that $c_l \le 1 \le c_h$. In our application, $p = 0.1$, $c_h = 1.25$, $c_l = 0.95$. Fig. 6 illustrates the location of z_p, t_p, t_h^c and t_l^c in the histogram.

In the case of the shadows, the parameters to define are ω, t_h^{sh} and t_l^{sh}. These parameters are specific to each cloud shadow. The model (8)–(10) will then be applied as many times as there are cloud detections. The introduction of ω is motivated both by the fact that the shadow detection applied with stationary parameters

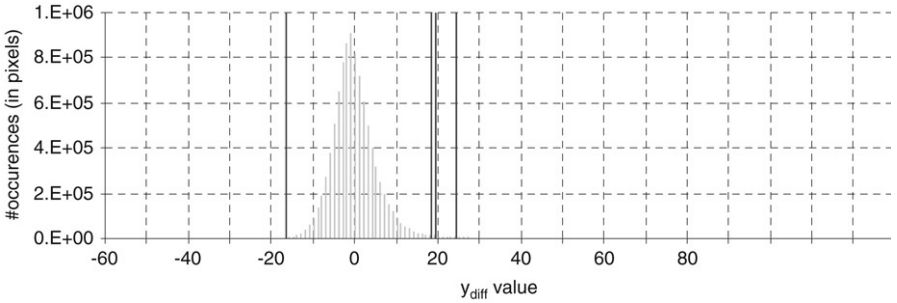

Fig. 6. Histogram of the image acquired on 04/29/2006 over Benin and values of $z_p = -17$, $t_p = 19$, $t_h^c = 24$ and $t_l^c = 18$.

to the whole image induces important false alarm rates, and by the fact that the model at object level (using the object graph) will invalidate any shadow not possibly associated to a cloud, i.e. having not a cloud in the direction θ given by (21) (cf. Section 4.2). Hence, practically and in order to reduce both false alarms and computation time, the shadow research areas are restricted to the areas consistent with the location of the detected clouds (cloud detection is performed first). The ω estimation is based on the a priori similarity (except for the cloud/shadow overlaps) between shadow and cloud shapes. Note that this assumption is true only in first approximation (because of the soil relief), and the 'exact' shape of the shadow will be detected from the data observations. For each cloud, the corresponding shadow is then researched by translating a binary mask B_C of the cloud shape in the direction θ. Practically, for each geometric translation step i in θ direction, $B_C(\theta, i)$ being the translated mask, and $\bar{y}_{B_C(\theta,i)}$ the mean y pixel reflectance over $B_C(\theta, i)$, the shadow location is given by the minimal argument of the reflectances $\bar{y}_{B_C(\theta,i)}$: $\tilde{i} = \arg\min_{i \in I} \bar{y}_{B_C(\theta,i)}$, and $\omega = B_C(\theta, \tilde{i})$. Note that when the intersection of the cloud and $B_C(\theta, i)$ is not empty, the pixels belonging to this intersection are not considered for the estimation of $\bar{y}_{B_C(\theta,i)}$. Finally, let us specify that the i search domain I excludes the geometric translations inducing a new selection of a shadow already associated to a cloud (so that two clouds aligned in θ direction get a priori two different shadows), and that the presence of another cloud in I or an overlapping of I with the image outside may invalidate the shadow search (making the H predicate valid, by default). The order of the estimation of the ω areas associated to the different clouds has thus an effect on the result: in an ad hoc way, we process the clouds by decreasing the size order. Besides $\omega = B_C(\theta, \tilde{i})$, the previously estimated shadow location provides approximate statistical parameters of the shadow, namely a minimal radiometric value, $y_{\min}(\omega)$, and a standard deviation, $\sigma(\omega)$ - values that depend on the soil features. The thresholds are then given by: $t_l^{sh} = y_{\min}(\omega)$, and $t_h^{sh} = y_{\min}(\omega) + c_{sh}\sigma(\omega)$, where c_{sh} is a real coefficient empirically set to 1 in our case.

In summary, concerning shadow detection, first the location of the shadow is estimated assuming its shape (equal to those of the cloud) and knowing the direction of the cloud-shadow axis. This gives the first rough estimation of the shadow called ω. Secondly, shadow shape is more accurately estimated using region growing process with germ given by the pixel(s) reaching the minimal value over ω and merging criterion involving a threshold t_h^{sh} also derived from ω.

4.4. Algorithm

4.4.1. Definitions of property Π and criterion H

Here we focus on the Markovian model at object level. We have to specify firstly the property Π that allows distinguishing mist from clouds, and secondly the predicate H corresponding to the overlapping criterion for the shadow and translated clouds.

The mist being a priori more 'transparent' than clouds, for an object labelled 'cloud or mist' $R_{c\cup m}$, Π is defined based on the proportion of pixels taking values between the two thresholds t_h^c and t_l^c: Π is true if $\left|\left\{s \in R_{c\cup m} : y_{diff,s} \in [t_l^c, t_h^c)\right\}\right| / |\{s \in R_{c\cup m} : y_{diff,s} \geq t_h^c\}| \geq t_b$, and is false otherwise. In our case, $t_b = 10^3$.

H is estimated by comparison of the shadows detected (as explained in Section 4.2) to the masks ω of the translated clouds. Let us consider a clique composed of an object 'shadow', denoted R_{sh}, and an object 'cloud', denoted R_c. $\omega(R_c)$ is the region corresponding to the $\tilde{i}$ geometric translation of R_c (in θ direction), but the pixels belonging to the cloud (in the same way as for ω estimation, when a cloud overlaps its shadow, the pixels belonging to the cloud are not considered). n_{sh}, n_c, and $n_{c\cap sh}$ are the numbers of pixels belonging respectively only to R_{sh}, only to $\omega(R_c)$, and both to $\omega(R_c)$ and R_{sh}: $n_{sh} = |\{t \in \Omega : t \in R_{sh}, t \notin \omega(R_c)\}|$, $n_c = |\{t \in \Omega : t \notin R_{sh}, t \in \omega(R_c)\}|$, $n_{c\cap sh} = |\{t \in \Omega : t \in R_{sh}, t \in \omega(R_c)\}|$. Then H is true if the proportions $\frac{n_{sh}}{n_{c\cap sh}}$ and $\frac{n_c}{n_{c\cap sh}}$, that are the parameters used for the mutual validation of clouds and shadows, are lower than the threshold t_v. Fig. 7 gives examples of false alarm detection or validation depending on H.

In this study, the threshold t_v was set equal to 0.75. Such a value may appear high, however, in the case of small clouds the surface relief may introduce important deformations in the shadow shape relatively to the cloud one, and the false alarm cases are mainly due to very 'bright' fields or buildings and the erroneously associated shadows are very large inducing ratios $n_{sh}/n_{c\cap sh} \gg 1$.

4.4.2. Global algorithm

Having defined the image model allowing the representation of properties P1, P2 and P3 in a Markovian framework, we showed that in the simple case of the proposed potential functions, the energy minimization (that corresponds to the maximization of the a posteriori probability) is obtained by hysteresis thresholding, region growing, and mutual validation of clouds and shadows as far as the local minimization is concerned. Besides, we recall that the shadow detection uses the cloud detection results (for ω estimation) so that the algorithm shall consider these two steps sequentially. Roughly, the proposed algorithm has three main steps respectively corresponding to cloud and mist detection, shadow detection, and cloud and shadow mutual validation. A priori parameters (p, c_h, c_l, c_{sh}, t_b, t_v) are set to the default values (0.1, 1.25, 0.95, 1, 10^3, 0.75) in our case.

The global algorithm synopsis is shown in Fig. 8.

5. Application and discussion of results

5.1. Markovian model results

Fig. 9 shows qualitatively the cloud and shadow detection by the proposed algorithm. Fig. 9(a) is the original multispectral SPOT/HRVIR image, from which the images y_{diff} (Fig. 9(b)) and y_{MIR}

(a) $n_c/n_{c\cap sh}$ & $n_{sh}/n_{c\cap sh} > t_v \Rightarrow$ false alarm. (b) $n_c/n_{c\cap sh} > t_v \Rightarrow$ false alarm. (c) $n_c/n_{c\cap sh}$ & $n_{sh}/n_{c\cap sh} < t_v \Rightarrow$ validation.

Fig. 7. Mutual validation according to criterion H based on the overlap between 'translated' cloud and its assumed shadow.

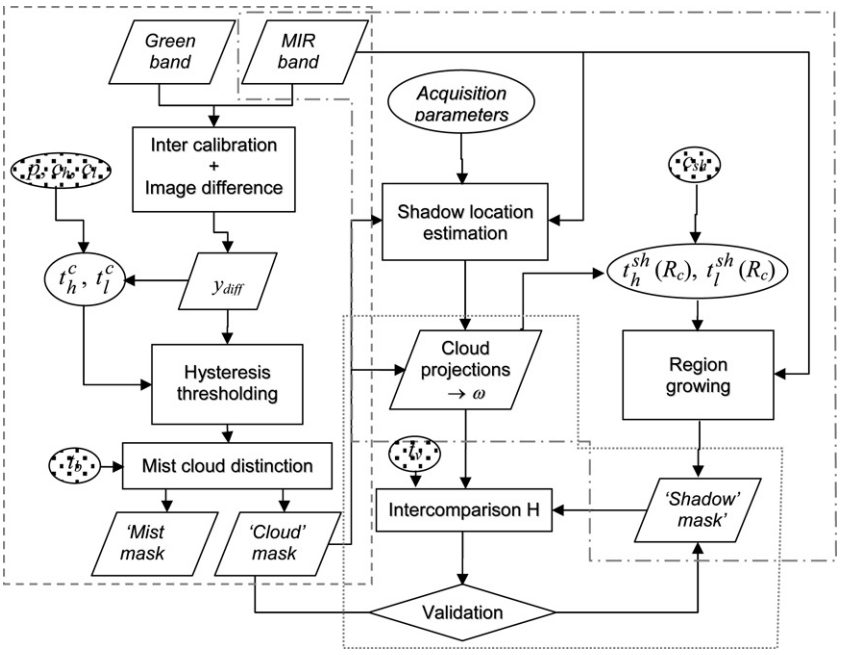

Fig. 8. Global algorithm synopsis. First Green and MIR bands are inter-calibrated (cf. Section 4.1.2) to get the image difference y_{diff}. Then, hysteresis thresholding of y_{diff} (that gives the local optimum of the pixel level MRF model for clouds) leads to the cloud mask (cf. Section 3.2). Then, for each detected cloud, the location of the associated shadow is determined knowing the acquisition parameters and the MIR image (cf. Section 4.2). The cloud is then projected to this location to get the germ and the parameters of the region growing process (that gives the local optimum of the pixel level MRF model for shadows) leading to the shadow mask (cf. Section 3.2). Finally, for each cloud, the intercomparison of its projection and its detected shadow allows validating it. If not validated, both cloud and associated shadow are removed from the cloud and the shadow masks, respectively.

Table 4
Possible decisions versus the truth: grey cells point out the erroneous decisions.

Decision	Truth	
	No cloud (A + C)	Cloud (B + D)
No cloud (A + B)	Right non-detection (A)	mis-detection (B)
Cloud (C + D)	False alarm (C)	Right detection (D)

(Fig. 9(c)) are derived. By processing the mist and cloud detection based on the pixel graph, the Fig. 9(d) image is obtained, and by processing the shadow detection, so is the Fig. 9(e) image. Finally, the cloud and shadow mutual validation processed using the object model provides the Fig. 9(f) image. The presented results are typical of those obtained with the different AMMA database images. Qualitatively, these results appear satisfying. Main remaining false alarms are due to bright areas for which the absence of shadow was justified by the location at the scene border or by the possible masking by another cloud. Non-detected clouds are small clouds over a bright soil.

Quantitatively, the results are evaluated in terms of 'right non-detections', 'false alarms' (there is no cloud in both cases that differ by the decision, either a non-detection, i.e. right decision, or a detection, i.e. erroneous decision), and in terms of 'mis-detections' and 'right detections' (there is a cloud in both cases that differ by the decision, either a non-detection, i.e. erroneous decision, or a detection, i.e. right decision). Table 4 summarizes these different cases.

Fig. 10(a)–(d) plot, in number of pixels, the 'right non-detections', the 'false alarms', the 'mis-detections' and the 'right detections'. Each point corresponds to a result from one of the 39 AMMA images. For a given Fig. 10(x) the total number of considered pixels (clouds for Fig. 10(a), (b), and clear pixels for Fig. 10(c), (d)) varies with the proportion of clouds within the images. We note the sum of the pixels corresponding to a right decision and those corresponding to an erroneous decision: graphically, for a given x-axis value, the sum of the y-axis value of the green cross and the y-axis value of the blue asterisk equals the y-axis value of the red cross. Results are hence good if the green crosses (right detections on Fig. 10(a), (b) or right non-detection on Fig. 10(c), (d)) are close to the red crosses, and the blue asterisks (erroneous decision) close to the x-axis. Then, we observe satisfyingly the previously described good results and that the errors are minor.

Now, most processed images include only a low percentage of clouds, that is represented in Fig. 10(a) by a point cluster close to the graph origin (most cloudy images correspond to about 3.5% and 1.5% of clouds on the image). Fig. 10(b) presents a zoom around the graph origin. The distribution of the 'mis-detection' rate ($B/(B + D) \times 100$) has median value equal to 8.33%, and 1st and

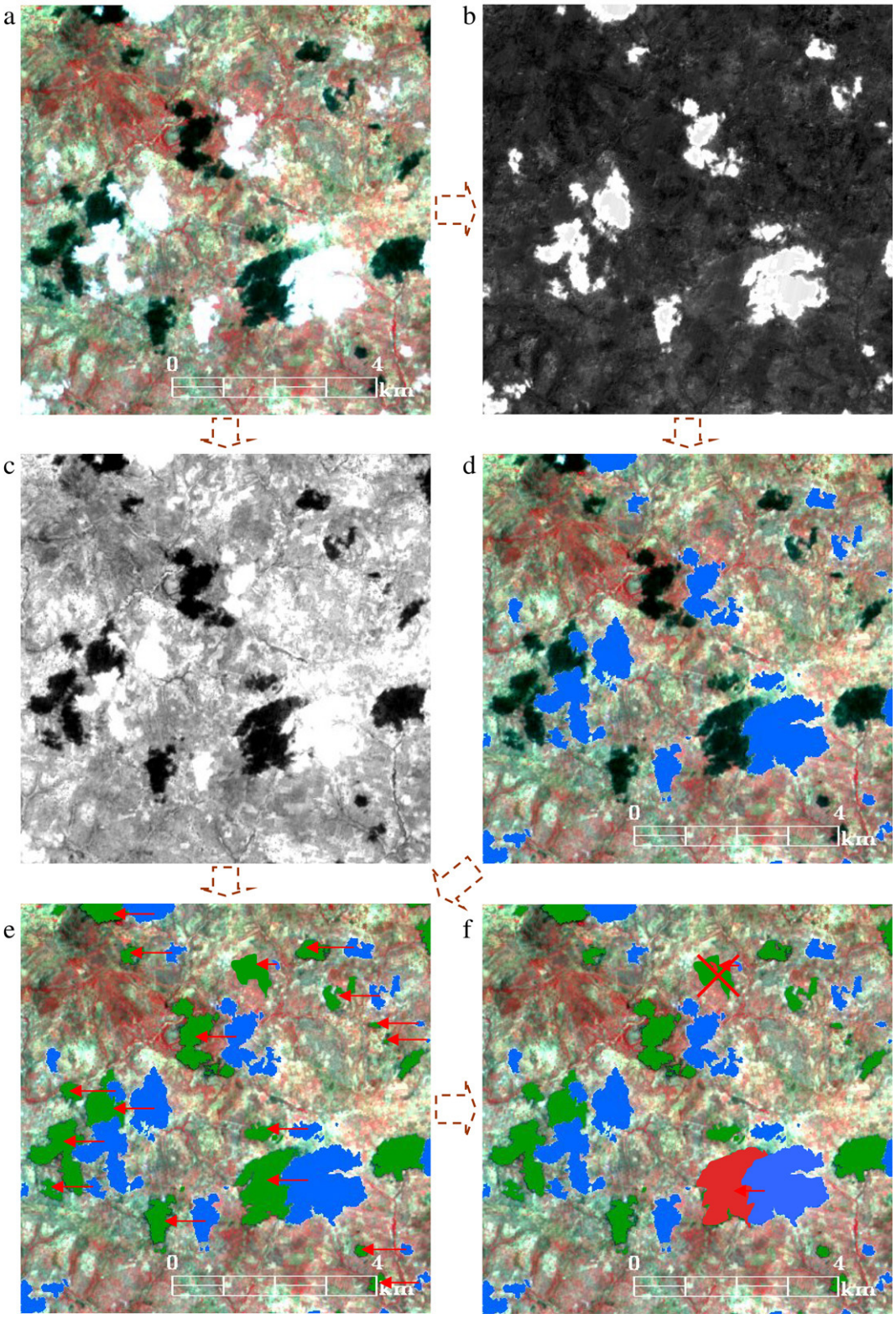

Fig. 9. Example of cloud and shadow detection: (a) HRV image, (b) y_{diff}, (c) y_{MIR}, (d) pixel level cloud detection (in blue), (e) pixel level shadow detection (in green) using the cloud projection (red arrows) in θ direction, (f) object level mutual validation.

3rd quartile values respectively equal to 2.33% and 12.23%. The proposed method mainly is efficiently reducing the false alarms; cf. Fig. 10(c), (d) (zoom). The distribution of the 'false alarm' rate relative to the non-cloud pixels ($C/(A + C) \times 100$) has median value equal to 0.0%, and 3rd quartile value equal to 0.0016% and the distribution of the 'false alarm' rate ($C/(C + D) \times 100$) has median value equal to 8.47%, and 1st and 3rd quartile values respectively equal to 0.16% and 100% (corresponding to the cases of clear sky images in which a 'small' cloud was detected, inducing 100% false alarms). Just for comparison of performance, we performed cloud

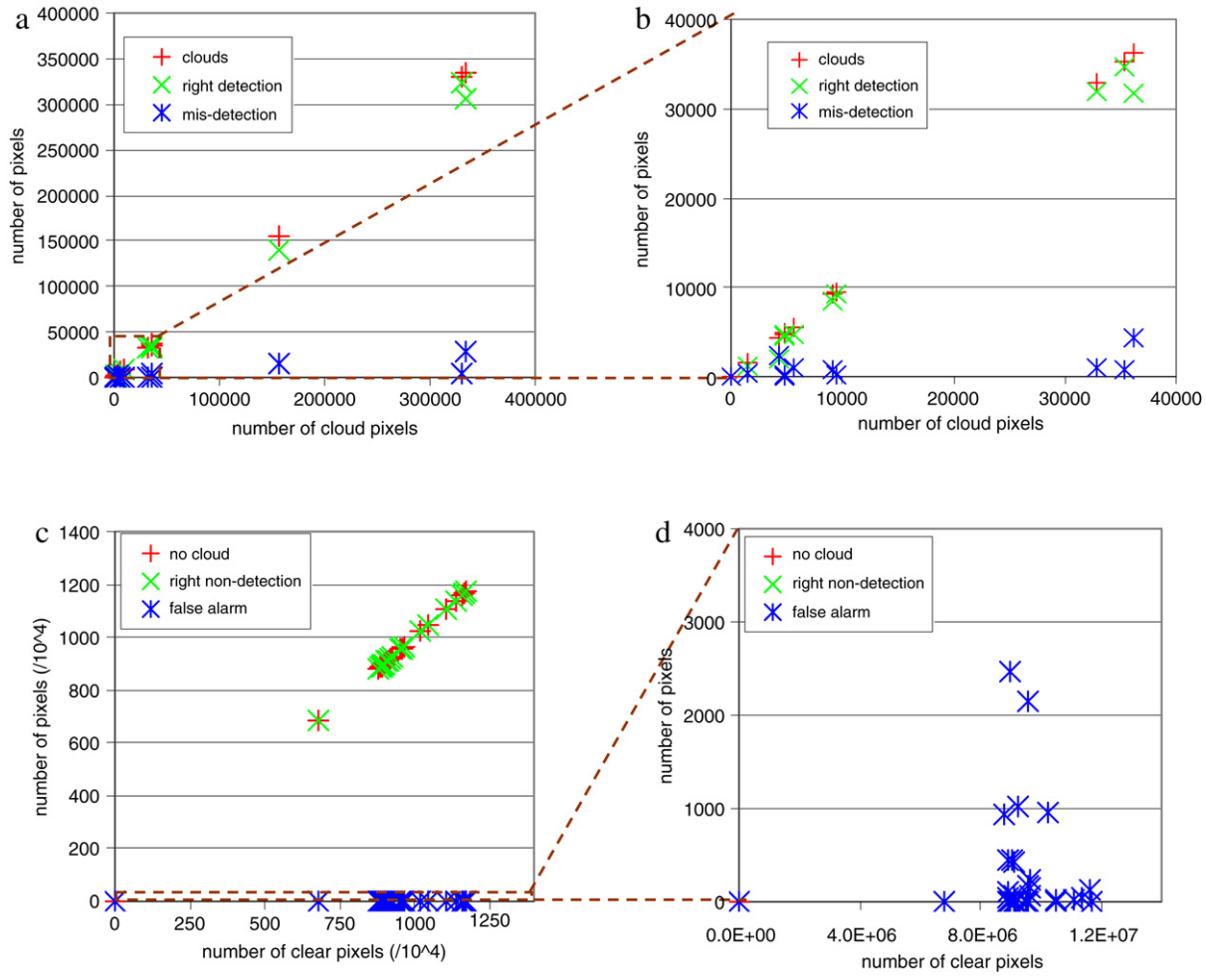

Fig. 10. MRF model results obtained on AMMA database: each point corresponds to an image, the 1st line presents the results (right detection and mis-detection) relative to the cloud pixels, the 2nd line presents the results (right non-detection and false alarm) relative to the clear pixels (no cloud); the 2nd column is a zoom over a subpart of the graph on the same line.

Table 5
Performance of a simple threshold approach for cloud detection: comparison with the proposed MRF based algorithm results (values between brackets).

Thresholding above $(t_h^c + t_l^c)/2$	'mis-detection' rate (B/(B + D) × 100)	'false alarm' rate (C/(A + C) × 100)	'false alarm' rate (C/(C + D) × 100)
1st quartile value	0.0% (<2.33%)	0.006% (>0.0%)	6.18% (>0.16%)
Median value	17.7% (>8.33%)	0.31% (>0.0%)	54.0% (>8.47%)
3rd quartile value	29.3% (>12.23%)	0.72% (>0.0016%)	100% (=100%)

Table 6
Sensitivity to parameters — cloud thresholds, shadow threshold, mist/cloud distinction parameter and cloud/shadow comparison one: L1 (L2) norm values of difference between obtained error rates with tested parameter values and error rates corresponding to parameter default value, varying the parameters within a range of ±10% around the proposed default values.

L1 (L2)	t_l^c	t_h^c	t_{sh}^c	t_b	t_v
Mis-detection rate (in %)	1.84 (3.32)	0.44 (0.94)	5.40 (19.9)	0.38 (1.44)	0.0097 (0.052)
False alarm rates (in %)	0.93 (2.74)	0.51 (1.79)	3.30 (17.7)	0.021 (0.065)	0.0007 (0.0032)

detection using thresholds equal to $(t_h^c + t_l^c)/2$. The 1st, 2nd and 3rd quartile values of 'mis-detection' rate relative to the cloud pixels, and 'false alarm' rate either relative to non-cloud pixels or relative to the cloud labelled pixels are given in Table 5.

Finally, Table 6 shows the sensitivity of our results to algorithm parameters. We note that it is rather robust since performance measured in terms of 'mis-detection' rate (B/(B + D) × 100) and 'false alarm' rate (C/(C + D) × 100) does not vary much (less than 1% in most cases, the shadow threshold being the most critical parameter).

5.2. Results using Marked Point Process

Marked Point Processes (MPP) can be viewed as a natural extension of the stochastic model, in particular Markov Random Field ones. Even if their use in image processing is relatively recent, they have already been applied to various problems (Descombes et al., 2004). They allow in particular making up for the lack of flexibility of the MRF over graphs as far as the number of considered nodes (i.e. objects) is concerned.

In this study, our aim is to use the MPP as an advanced model for comparison with the MRF model proposed. In this section only

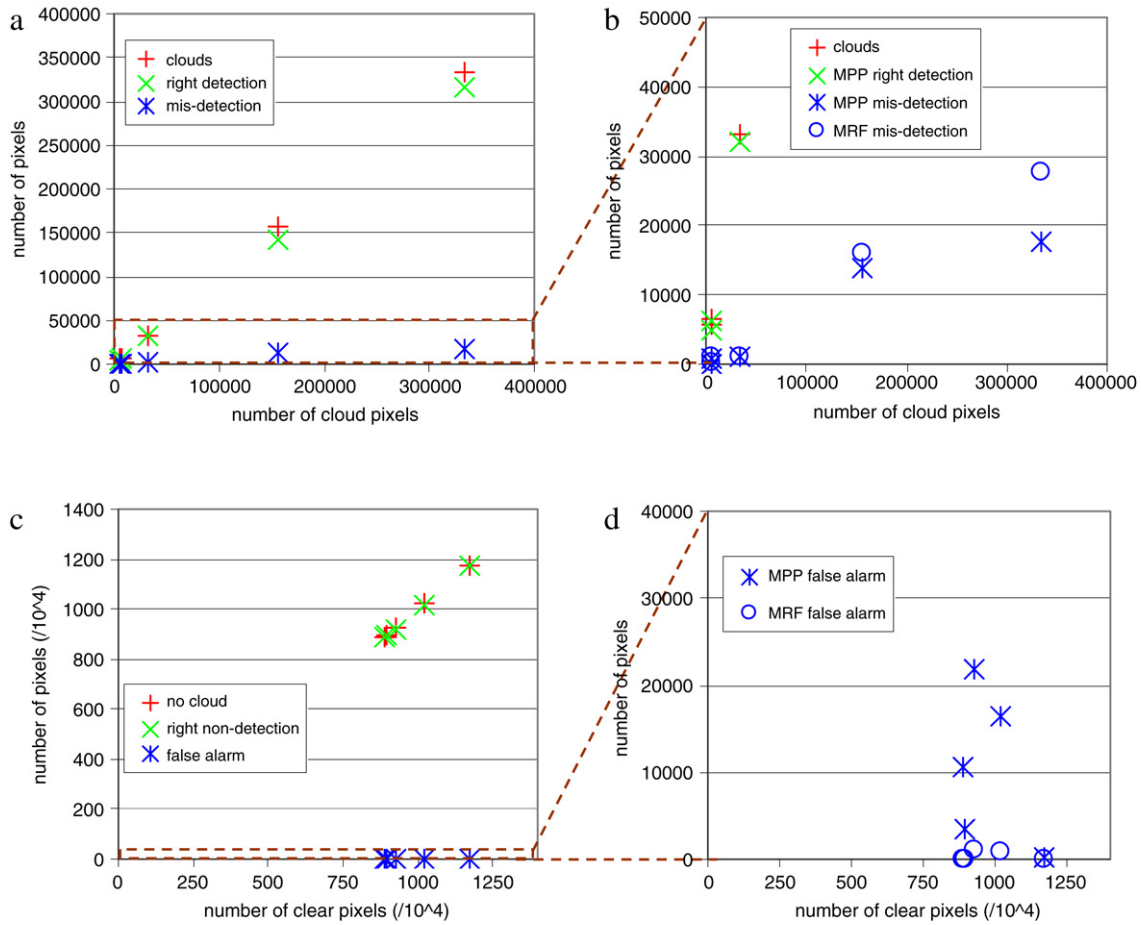

Fig. 11. MPP model results obtained on AMMA database: each point corresponds to an image, the 1st line presents the results (right detection and mis-detection) relative to the cloud pixels, the 2nd line presents the results (right non-detection and false alarm) relative to the clear pixels (no cloud); the 2nd column compares the results obtained using respectively the MRF proposed model and the MPP one — for figure legibility, the *y*-axis scale focuses on error (mis-detection or false alarm) value range.

the obtained results are presented and commented in comparison with the MRF ones, whereas the main features of the MPP are summarized in Appendices A and B presenting the way we apply them to our cloud/shadow detection problem.

Fig. 11 shows the results obtained using the MPP model and compares them to the results obtained using the MRF model over graphs (Section 5.1). Note that among the 39 data images, only 5 have been processed. These images are those where the MPP model performs better, in particular images without mist since the MPP developed model does not process the case of the mist. Indeed, as in the case of the MRF model, the mist shadow, difficult to detect, cannot be used for mutual validation with mist without taking the risk of highly increasing the mist mis-detection. Now, conversely to the MRF model, the MPP model does not distinguish the cloud and mist cases (the distinction done in the MRF model is at object level while the MPP model remains at pixel level) and therefore always looks for a shadow pixel associated to a 'cloud or mist' pixel. The second reason for processing only a few images is the algorithm running time that is much more important (multiplied by a factor about 1000) for the MPP model because of the use of a simulated annealing process for global optimization (while only local optimization is performed in the case of the MRF model). Note that we checked that local optimization of MPP model leads to very poor results.

Fig. 11(a), (c) are similar to Fig. 10(a), (c) except for the number of considered images (figure points). The MPP results obtained correspond to the same order of performance as MRF ones. To compare the two models in a more precise way, Fig. 11(b), (d) present both results (for each of the 5 considered images) focusing on the error value ranges (either mis-detection number or false alarm number). The main comments are the following.

From Fig. 11(b), we note that mis-detection cases are reduced using the MPP model instead of the MRF model. This is consistent with the fact that, using MRF over graphs, the graph node number cannot vary. Practically, in our application, if a cloud pixel is mis-detected during the first algorithm step (using the pixel graph), there is no way to recover it during the following steps: at object level, only previous step detections can be validated. Conversely, the MPP model is able to handle a graph with a variable number of nodes thanks to the birth and death transformations. Our results hence illustrate the advantage of such a more general model. However, we also note that the price for the mis-detection rate decrease is an increase of the false alarm rate, visible in Fig. 11(d). Now, if one recalls that, for applications such as quantitative estimation of land cover change, cloud/shadow false alarms are preferable to mis-detection, the advantage is for the MRF model approach (MPP will be more advantageous in the case of applications for which the respective costs of false alarm and mis-detection are almost equal).

Fig. 12 shows an example of results obtained in the case of the image acquired on 04/29/2006 on Benin: Fig. 12(a) is the data image, Fig. 12(b), (c) show, respectively, the MRF model result and the MPP one, and Fig. 12(d) is the image difference between MPP and MRF results. One can see a slight increase in the number of detected cloud pixels consistent with Fig. 11 previously commented on. However, the main point of Fig. 12 is the difference in shadow estimation. The shadows obtained from the MPP model are much closer in shape than those obtained using the MRF model.

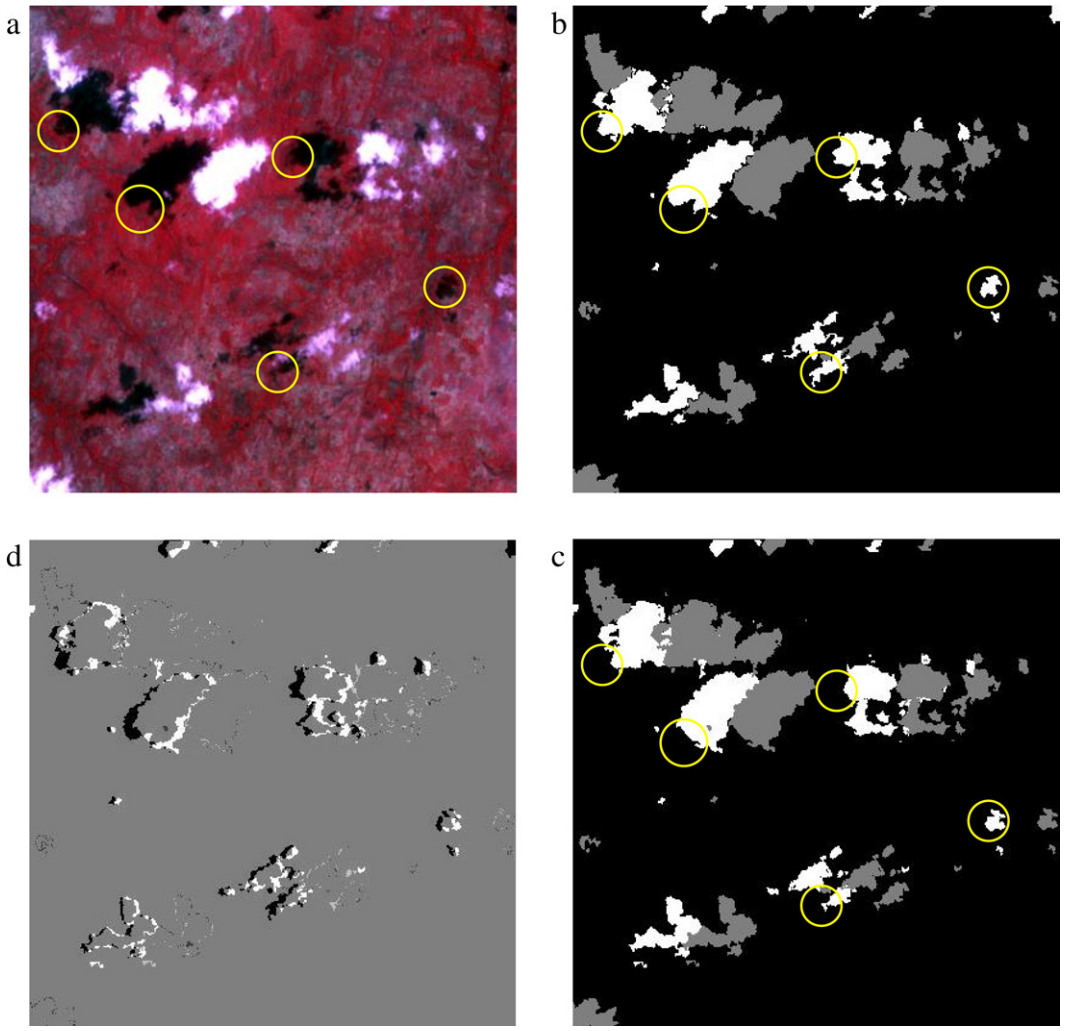

Fig. 12. Comparison of MPP and MRF models (subpart of the 04/29/2006 image acquired on Benin): (a) data image, (b) MRF model result, (c) MPP model result, (d) image difference between MPP and MRF results.

This can be explained by the fact that in the MPP model there is a term that compels the geometric translation (between cloud pixel and associated shadow pixel) of neighbouring pixels to be close (that is necessary to ensure the similarity of cloud and shadow global shapes). In the MRF model, having estimated approximate shadow locations, shadow shape is achieved independently of cloud (for comparison between cloud and shadow at object level). Now, even if we have not performed quantitative evaluation of the algorithm performance on shadow, areas circled in Fig. 12 point out areas where the shadows obtained in Fig. 12(b) seem qualitatively more consistent (relative to Fig. 12(c)) with those that would be obtained performing the manual photo interpretation of Fig. 12(a).

Finally, it should be recalled that the MPP approach requires the precise fitting of 12 parameters. On the one hand, even if numerous tests have been performed in order to determine (in a supervised way) the parameter values that lead to best performance, it is still possible that better results can be obtained using other parameters. On the other hand, the fact that the MRF model based algorithm requires only few parameters to fit (at least in the practical implementation that we propose) may be seen as a serious advantage in terms of robustness and user-friendliness of the algorithm.

6. Conclusions

In this paper, we present a new method to detect cloud and shadow on optical visible high resolution images. It is based on the use of the Markov Random Field formalism on graphs at two levels: the pixel level in which case the considered graph is the image lattice, and the object level in which case the considered graph nodes are the cloud/mist/shadow objects detected from the first processing steps at pixel level. The main advantages of the proposed algorithm are:

(i) good performance on the tested database (containing 39 SPOT/HRV images),
(ii) reduced false alarm rates (relative to classical pixel level approaches),
(iii) only six effective parameters (c_h, c_l, p, c_{sh}, t_b, t_v) that are rather robust and can be set at a default value (kept constant for the processing of the considered 39 images, that correspond to various acquisition, landscape and season conditions),
(iv) fast convergence (relatively to methods requiring global optimization).

The proposed model was compared to a more general and flexible model defined using the Marked Point Process framework, and the results of our simplest approach in terms of performance were shown. The development and the use of the MPP model pointed to two drawbacks: the absence of consideration of the object level (due to the difficulty to model simply the geometrical properties of clouds), and the fastidious feature of the task to set MPP model parameters. Then, future studies can try to overcome these drawbacks by constructing a more sophisticated model

including object features and exploiting unsupervised techniques (that automatically estimate the model parameters).

Acknowledgements

Based on a French initiative, AMMA was built by an international scientific group and is currently funded by a large number of agencies, especially from France, UK, US and Africa. It has been the beneficiary of a major financial contribution from the European Community's Sixth Framework Research Programme. Detailed information on scientific coordination and funding is available at the AMMA International web site http://www.amma-international.org.

Appendix A. Point processes — Notations, definitions and algorithms

In this section, the main definitions of Point Processes (PP) are given on the prospect of their application to the stochastic modelling of the objects (existence and interactions) present in the image (Van Lieshout, 2000; Descombes, 2004). The definitions are given assuming a space χ, that is generally a continuous bounded set included in R^2 representing the image, $L(\chi)$ the Lebesgue σ-algebra over χ, and having a metric, d, e.g. the Euclidian distance.

A.1. Marked Point Processes

A configuration $\mathbf{x}$ is a not-ordered set of χ points. It is locally finite if in any χ Borel bounded set, there is a finite number of points. A PP is an application X from a probability space (Ω, A, P) to the set N^{lf} of the locally finite configurations, such that for any Borel set $A \subseteq \chi$, the number of points in A, denoted $N(A)$, is a finite random variable (the PP realisations are random configurations of a finite number of points).

Let us now consider (χ, d) and (M, d'), two metric spaces complete and separable. A Marked Point Process (MPP) whose locations are in χ and the marks in M is a PP such that the process of the not-marked points is a well defined PP. Now, the marginalisation over χ of the joint random variable is a finite random variable in particular if M, finite, is discrete.

A.2. Density of a point process with respect to the Poisson process

The Poisson PP (most famous) corresponds to a uniform repartition of the points in χ. By definition, a Poisson PP of intensity $\nu(.)$ ($\nu(.)$ being a Borel measure over (χ, d) such that $\nu(\chi) > 0$ and $\nu(A) < +\infty$ for any bounded A) is such that: (P1′) for any Borel bounded $A \subseteq \chi$, $N(A)$ follows a Poisson law of mean $\nu(A)$: $P(N(A) = n) = \mathrm{e}^{-\nu(A)} \cdot \frac{(\nu(A))^n}{n!}$, and (P2′) for any k disjoint Borel sets $A_1, \ldots, A_k$, the random variables $N(A_i)$, $i = 1, \ldots, k$, are independent. (P2′) can be interpreted as a spatial non-correlation property.

The Poisson processes are used, in an analogue way to the Lebesgue measure over R^d, to define any PP by its density with respect to the distribution of the reference Poisson PP: Let $\pi(.)$ be the distribution of a finite intensity Poisson process over χ, and $h(.)$ a positive function from N^f, the set of finite configurations, to $[0, +\infty)$; if $Z(h, \pi) = \int_{N^f} h(\mathbf{x}) .\mathrm{d}\pi(\mathbf{x}) < +\infty$, then $Z^{-1}.h(\mathbf{x})$ may be interpreted as the density of a PP with respect to the reference Poisson process (Van Lieshout, 2000). In the following, PP are characterized by their density $h(.)$.

A.3. Markov and Gibbs point processes

Let $\sim$ be a symmetrical relationship over χ: Two points of χ, r and s are 'neighbours' if $r \sim s$. The neighbourhood of a bounded Borel set $A \subseteq \chi$ is the set of χ points having a neighbour in A: $\{x \in \chi | \exists a \in A : a \sim x\}$. A configuration $\mathbf{x}$ of N^f is a clique if any two $\mathbf{x}$ points are neighbours: $\forall (u, v) \in \mathbf{x} \times \mathbf{x}, u \sim v$. Then, Ripley and Kelly (1977) defined a Markov PP as follows. Let $\nu(.)$ be a Borel measure over (χ, d), $\pi_\nu(.)$ the Poisson PP law with intensity $\nu(.)$, and X a PP over χ defined by its density $h(.)$ with respect to $\pi_\nu(.)$. X is a Markov PP with respect to the relationship $\sim$ over χ if for any $\mathbf{x}$ of N^f such that $h(\mathbf{x}) > 0$, then (i) for any $\mathbf{y} \subseteq \mathbf{x}$, $h(\mathbf{y}) > 0$, and (ii) for any $u \in \chi$, $\lambda(\mathbf{x}, \{u\}) = h(\mathbf{x} \cup \{u\})/h(\mathbf{x})$ only depends on u and $\{x \in \chi | u \sim x\}$.

Using a theorem analogue to the Hammersley–Clifford theorem, the density h: $N^f \to [0, +\infty)$ of a Markov PP writes $h(\mathbf{x}) = \prod_{\text{cliques } \mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y})$, where $\phi(.)$ is a measurable function: $N^f \to [0, +\infty)$. More generally, $h(\mathbf{x}) = \prod_{\mathbf{y} \subseteq \mathbf{x}} \phi(\mathbf{y})$ where $\phi(.)$ is a function modelling the interactions that equals 1 except if its argument is a clique. Finally, by definition, a Gibbs PP has a density ($h(.)$ with respect to the reference PP) that writes in its most general form $\frac{1}{Z} \exp\{-U(\mathbf{x})\}$, where Z is a normalization constant and $U(\mathbf{x})$ the PP energy written as a sum of interaction potentials between couple pixels: $U(\mathbf{x}) = \sum_{i,j:} \sum_{1 \le i < j \le n} V(x_i, x_j)$. For example, when $\chi = R^2$ and $V(x_i x_j) = \begin{cases} -\ln \gamma & \text{if } d(x_i x_j) < r \\ 0 & \text{if } d(x_i x_j) \ge r \end{cases}$, where $d(.,.)$ is a spatial distance, $r > 0$, $\gamma > 0$, the PP is a Strauss PP: $\gamma = 1$ corresponds to the Poisson PP (i.e. without point interaction), $\gamma \in (0, 1)$ to a repulsion between neighbour points, and $\gamma > 1$ to an attraction between neighbour points.

A.4. Algorithm RJMCMC

Let us consider the space $S = \chi \times M$. Here, N^f is the set of finite configurations of marked points. $p(.)$ is the MPP law, $h(.)$ its density with respect to the reference Poisson PP $\pi(.)$ having intensity $\nu(.)$ over N^f: $\forall \mathbf{x} \in N^f, p(\mathrm{d}\mathbf{x}) = h(\mathbf{x}) .\pi(\mathrm{d}\mathbf{x})$

The RJMCMC algorithm aims at constructing a Markov chain $(X_n)_{n \ge 0}$ over N^f. The convergence toward $p(.)$ of a Markov chain having transition node $Q(\mathbf{x}, \mathrm{d}\mathbf{x})$ is ensured once Q is aperiodic, irreducible and satisfies $\int_A p(\mathrm{d}\mathbf{x}) . \int_B Q(\mathbf{x}, \mathrm{d}\mathbf{y}) = \int_B p(\mathrm{d}\mathbf{y}) . \int_A Q(\mathbf{y}, \mathrm{d}\mathbf{x})$ ('equilibrium' condition). Green (1995) proposes using a mixing proposition node $Q_m(\mathbf{x}, d\mathbf{y})$, where m represents a transformation (e.g. 'birth', i.e. addition of a point, 'death', i.e. suppression of a point, modification of the parameters of a point, etc.), such that: $Q(\mathbf{x}, \mathrm{d}\mathbf{y}) = \sum_m Q_m(\mathbf{x}, \mathrm{d}\mathbf{y})$ and $Q(\mathbf{x}, \chi) \le 1$ so that we stay in the same state with probability $\left[1 - \sum_m Q_m(\mathbf{x}, \chi)\right]$. $Q_m(\mathbf{x}, \chi)$, also denoted $p_m(\mathbf{x})$, is the probability to select transformation m being in state $\mathbf{x}$. $Q_m(\mathbf{x}, \chi) = 0$ for some m means that some transformations from a given state $\mathbf{x}$ are not possible. Let $\alpha_m(\mathbf{x}, \mathbf{y})$ be the probability to accept the new state $\mathbf{y}$ (using a transformation m). To estimate $\alpha_m(\mathbf{x}, \mathbf{y})$, we consider $f_m(.,.)$ the density of $h(.).Q_m(.,.)$ with respect to the measure ξ_m over $\chi \times \chi$: $\int_A \int_B p(\mathrm{d}\mathbf{x}) .Q_m(\mathbf{x}, \mathrm{d}\mathbf{y}) = \int_A \int_B f_m(\mathbf{x}, \mathbf{y}) .\xi_m(\mathrm{d}\mathbf{x}, \mathrm{d}\mathbf{y})$ ($f_m(.,.)$ is the Radon–Nikodym derivative). According to Green (1995), if ξ_m is symmetrical, the equilibrium condition is checked once $f_m(\mathbf{x}, \mathbf{y}) \cdot \alpha_m(\mathbf{x}, \mathbf{y}) = f_{m^{-1}}(\mathbf{y}, \mathbf{x}) \cdot \alpha_{m^{-1}}(\mathbf{y}, \mathbf{x})$ where the subscript m^{-1} denotes the m reverse transformation. Besides, to minimize the Markov chain autocorrelation $\alpha_m(\mathbf{x}, \mathbf{y})$ should be chosen as large as possible (Peskun, 1973). Green (1995) proposes:

$$\alpha_m(\mathbf{x}, \mathbf{y}) = \min\left\{1, \frac{f_{m^{-1}}(\mathbf{y}, \mathbf{x})}{f_m(\mathbf{x}, \mathbf{y})}\right\}. \tag{A.1}$$

The Reversible-Jump-Monte-Carlo-Markov-Chains (RJMCMC) algorithm is then as follows:

1. choose a transformation m with probability $p_m(\mathbf{x})$; either go to step 2, or stay in configuration $\mathbf{x}$ with probability $\left[1-\sum_m p_m(\mathbf{x})\right]$ and do not consider the following steps;
2. simulate $\mathbf{y}$ according to $Q_m(\mathbf{x}, \mathrm{d}\mathbf{y})$;
3. compute the Green ratio $R = f_{m^{-1}}(\mathbf{y}, \mathbf{x}) / f_m(\mathbf{x}, \mathbf{y})$;
4. accept the transformation $\mathbf{x} \to \mathbf{y}$ with probability $\alpha_m(\mathbf{x}, \mathbf{y}) = \min(1, R)$.

Appendix B. Marked Point Process: The proposed model

Generally used in the context of object detection in image processing, MPP allows for some stochastic modelling that takes into account the geometrical features of the objects as the interactions between objects. However such modelling requires that the objects be characterized by simple geometrical shapes: line segments for the roads, ellipses for the trees and rectangles or parallelepipeds for the buildings (Descombes, 2004). In our application case, unfortunately clouds do not have a sufficiently simple geometry to allow characterization by a small number of parameters or marks. Hence we apply the MPP not at the cloud object level, but at the cloud pixel level (or equivalently we consider only elementary clouds of size equal to 1 pixel). Then, each couple of a cloud pixel and a shadow pixel can be characterized by the location of the cloud pixel (at pixel level the space χ is the image space Ω) and the length l of the line segment between the two pixels (since the direction is determined by the sun and geometry acquisition conditions). The MPP is hence defined over $S = \Omega \times [l_{\min}, l_{\max}]$, where $l_{\min}$ and $l_{\max}$ are the minimal and maximal values of the cloud-shadow geometric translation norm. In the following, the length l is estimated with a given precision so that M is discrete.

Pixel level allows getting a mark space of low dimension even if it induces some drawbacks. In particular, it will not be possible to introduce (at least in a simple way) some criteria about the respective shapes and sizes of a cloud object and its assumed shadow one, or some criteria about the object proportion checking a property (useful to distinguish the cloud from the mist). Thus, even if the energy functions for the proposed model have been chosen to make it as close as possible to the previous Markovian model (Section 3), both models will not be equivalent: The gain in flexibility of the modelling related to the possibility to get a graph of dimension variable (number of points of configuration $\mathbf{x}$) contrasts with the loss of the possibility to consider two modelling levels (pixel and object).

B.1. Energy term definition

In Appendix A, we saw that for a Gibbs PP, the density $h(.)$ is determined by the energy $U(\mathbf{x})$. Generally $U(\mathbf{x})$ writes as the sum of a data attachment energy $U_d(\mathbf{x})$ that involves each point of $\mathbf{x}$ separately and a prior energy $U_p(\mathbf{x})$ that represents the a priori information about point interactions and that involves some subsets points of $\mathbf{x}$ belonging to a same clique: $U(\mathbf{y}_c, \mathbf{y}_s|\mathbf{x}) = U_d(\mathbf{y}_c, \mathbf{y}_s|\mathbf{x}) + U_p(\mathbf{x})$.

Concerning the prior energy $U_p(.)$, it is defined in order to model some spatial regularity and consistence between cloud object and associated shadow object. More precisely, it is composed of two terms: an attraction term $U_a(\mathbf{x})$ favouring the gathering together of the MPP points (in agreement with P1 property according to which clouds are connex objects), and a likeness term $U_l(\mathbf{x})$ favouring close mark (cloud-shadow geometric translation norm) values for neighbour points (that are assumed belonging to a same cloud and therefore having consistent shadow location in agreement with properties P2 and P3). Then, the neighbourhood is chosen to coincide with the n_{connex}-connexity ($n_{connex} \in \{4, 8\}$) at least one of the two extremities of the cloud-shadow line segment: $u \sim v \Leftrightarrow (\|u - v\|_2 \le d_{\max}$ or $\left\|u + \vec{t}_u - v - \vec{t}_v\right\|_2 \le d_{\max})$, where $d_{\max} \in [1, \sqrt{2})$ if $n_{connex} = 4$, and $d_{\max} \in [\sqrt{2}, 2)$ if $n_{connex} = 8$. V_u is the set of points v of $\mathbf{x}$, configuration of Ω, such that $u \sim v$: $V_u = \{v \in \mathbf{x}, v \neq u : v \sim u\}$, having cardinal $|V_u|$. Finally, we propose $U_p(\mathbf{x}) = U_a(\mathbf{x}) + U_l(\mathbf{x})$ with:

$$U_a(x) = -\beta \cdot \sum_{u \in \mathbf{x}} |V_u|, \tag{B.1}$$

$$U_s(\mathbf{x}) = \delta \cdot \sum_u \left| \frac{l_u - \mu\left(l_{V_u}\right)}{\sigma_{\left(l_{V_u}\right)}} \right|, \tag{B.2}$$

where β is a parameter, $\mu\left(l_{V_u}\right)$ and $\sigma\left(l_{V_u}\right)$ are the mean and the standard deviation of l over V_u.

Concerning the data attachment energy $U_d(.)$, it is decomposed into two terms concerning respectively the cloud and the shadow. Let $y_{c,u}$ denote the data value at pixel u ($u \in \mathbf{x}$) for the cloud detection observation field and $y_{s,t}$ the data value at pixel t for the shadow detection observation field. The shadow pixel coordinate is determined from the coordinate of the cloud pixel u and the associated mark l_u that gives the cloud-shadow geometric translation vector $\vec{l}_u$: $t = u + \vec{l}_u$. Now, assuming independence of the observations conditionally to the label field, and mimicking the Markovian modelling over the Ω graph (pixel level), we propose:

$$\begin{aligned} U_d(\mathbf{y}_c, \mathbf{y}_s|\mathbf{x}) = & \sum_{u \in \mathbf{x}} [U_d^c\left(y_{c,u}|u \in \mathbf{x}\right) + U_d^s\left(y_{s,u+\vec{l}_u}|u \in \mathbf{x}\right)] \\ & + \sum_{u \notin \mathbf{x}} [U_d^c\left(y_{c,u}|u \notin \mathbf{x}\right) + U_d^s\left(y_{s,u+\vec{l}_u}|u \notin \mathbf{x}\right)], \end{aligned} \tag{B.3}$$

where $U_d^c(y_{c,u}|u \in \mathbf{x})$ is $U_s^0(y_s|x_s = \lambda_{c \cup m})$ of (5) for $s = u$, $U_d^c(y_{c,u}|u \notin \mathbf{x})$ is $U_s^0(y_s|x_s = \lambda_{\overline{c \cup m}})$ of (6) for $s = u$, $U_d^s(y_{s,v=u+\vec{l}_u}|u \in \mathbf{x})$ is $U_s^0(y_s|x_s = \lambda_{sh})$ of (9) for $s = v$ and $\omega = \Omega$, and $U_d^s(y_{s,v=u+\vec{l}_u}|u \notin \mathbf{x})$ is $U_s^0(y_s|x_s = \lambda_{\overline{sh}})$ of (8) for $s = v$ and $\omega = \Omega$.

B.2. Used transformations

We now aim at specifying, for each transformation m, $f_m(\mathbf{x}, \mathbf{y})$ versus $h(\mathbf{x})$ and $\mathbf{x}$.

Let us first consider the basic transformations that are the birth (subscript b) and the death (subscript d) of a point. According to (e.g. Descombes (2004), for a death, the point to remove is chosen among the points of the current configuration $\mathbf{x}$ following the uniform law $1/n(\mathbf{x})$, and, for a birth, the point to add is chosen following the $\nu(.)/\nu(\chi)$ law for its location and the uniform law $1/|M|$ for its marks. Then, for $\mathbf{y} = \mathbf{x} - \{u\}$ (u death): $p(\mathrm{d}\mathbf{x}) \cdot Q_d(\mathbf{x}, \mathrm{d}\mathbf{y}) = h(\mathbf{x}) \cdot \pi(\mathrm{d}\mathbf{x}) \cdot p_d(\mathbf{x}) \cdot 1/n(\mathbf{x}) \Rightarrow f_d(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}) \cdot p_d(\mathbf{x}) / n(\mathbf{x})$, and for $\mathbf{y} = \mathbf{x} \cup \{u\}$ (u birth): $p(\mathrm{d}\mathbf{x}) \cdot Q_b(\mathbf{x}, \mathrm{d}\mathbf{y}) = h(\mathbf{x}) \cdot \pi(\mathrm{d}\mathbf{x}) \cdot p_b(\mathbf{x}) \cdot \nu(\mathrm{d}u) / \nu(\chi) \Rightarrow f_b(\mathbf{x}, \mathbf{y}) = h(\mathbf{x}) \cdot p_b(\mathbf{x}) / [\nu(\chi) \cdot |M|]$. The Green ratios $R_m = f_{m^{-1}}(\mathbf{y}, \mathbf{x}) / f_m(\mathbf{x}, \mathbf{y})$ are thus:

$$\begin{aligned} R_d &= \frac{h(\mathbf{x} - \{u\}) \cdot p_b(\mathbf{x} - \{u\}) \cdot n(\mathbf{x})}{h(\mathbf{x}) \cdot p_d(\mathbf{x}) \cdot (l_{\max} - l_{\min}) / \delta l \cdot \nu(\Omega)} \\ &\tilde{\propto} \frac{h(\mathbf{x} - \{u\})}{h(\mathbf{x}) \cdot (l_{\max} - l_{\min})}, \end{aligned} \tag{B.4}$$

$$\begin{aligned} R_b &= \frac{h(\mathbf{x} \cup \{u\}) \cdot p_d(\mathbf{x} \cup \{u\}) \cdot (l_{\max} - l_{\min}) / \delta l \cdot \nu(\Omega)}{h(\mathbf{x}) \cdot p_b(\mathbf{x}) \cdot (n(\mathbf{x}) + 1)} \\ &\tilde{\propto} \frac{h(\mathbf{x} \cup \{u\}) \cdot (l_{\max} - l_{\min})}{h(\mathbf{x})}, \end{aligned} \tag{B.5}$$

where δl is the l mark resolution, and the approximations are obtained approximating the ratio $p_d(\mathbf{x}) / p_b(\mathbf{y})$ by $n(\mathbf{x}) / \nu(\Omega)$ ($\mathbf{x}$ and $\mathbf{y}$ are two configurations related by a death or a birth).

Finally, it is also possible to consider transformations that only change the parameters of a point (e.g. modification of the value of its mark) and leave unchanged the total number of points of the current configuration **x**. Such transformations that are said to be 'non-jumping', subscribed *nj*, are performed in order to speed up the convergence. They induce a classical Metropolis–Hasting scheme with a probability to accept the change $\mathbf{x} \rightarrow \mathbf{y}$:

$$\alpha_{nj}(\mathbf{x}, \mathbf{y}) = \min\left\{1, \frac{h(\mathbf{y})}{h(\mathbf{x})}\right\}. \tag{B.6}$$

References

Abend, K., Harley, T.J., Kanal, L.N., 1965. Classification of binary random pattern. IEEE Transactions on Information Theory 11 (4), 538–544.

Benboudjema, D., Pieczynski, W., 2005. Unsupervised image segmentation using triplet Markov fields. Computer Vision and Image Understanding 99 (3), 476–498.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society, Series B 36 (2), 192–236.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (11), 1222–1239.

Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (9), 1124–1137.

Chen, P.Y., Srinivasan, R., Fedosejevs, G., Narasimhan, B., 2002. An automated cloud detection method for daily NOAA-14 AVHRR data for Texas. USA. International Journal of Remote Sensing 23 (15), 2939–2950.

Cihlar, J., Howarth, J., 1994. Detection and removal of cloud contamination from AVHRR images. IEEE Transactions on Geoscience Remote Sensing 32 (3), 583–589.

Descombes, X., 2004. Méthodes stochastiques en analyse d'image : des champs de Markov aux processus ponctuels marqués, habilitation à diriger des recherches. Université de Nice-Sophia Antipolis.

Descombes, X., Kruggel, F., Lacoste, C., Ortner, M., Perrin, G., Zerubia, J., 2004. Marked point processes in image analysis: From context to geometry. In: Proc. of International Conference on Spatial Point Process Modelling and its Application, SPPA, Castellon, Spain, pp. 4–8. April 2004, 8 p. (on CDROM).

Geman, S., Geman, D., 1984. Stochastic relaxation Gibbs distribution and Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6), 721–741.

Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82 (4), 711–732.

Greig, D., Porteous, B., Seheult, A., 1989. Exact maximum a posteriori estimation for binary images. Journal of Royal Statistical Society Series B 51 (2), 271–279.

Ho, A.T.S., Zhenlei, C., 1996. Cloud detection in satellite images for tropical regions. Proceedings of the SPIE 2818, 167–171.

Kauth, R.J., Thomas, G.S., 1976. The tasselled cap – a graphic description of the spectral-temporal development of the agricultural crops as seen by Landsat. In: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana, USA, pp. 4B41–4B50.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220 (4598), 671–680.

Kolmogorov, V., Zabih, R., 2004. What energy functions can be minimized via graph cuts? IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (2), 147–159.

Le Hégarat-Mascle, S., Kallel, A., Descombes, X., 2007. Ant colony optimization for image regularization based on a non-stationary Markov modeling. IEEE Transactions on Image Processing 16 (3), 865–878.

Peskun, P.H., 1973. Optimum Monte Carlo sampling Markov chains. Biometrika 60 (3), 607–612.

Pieczynski, W., Benboudjema, D., 2006. Multisensor triplet Markov fields and theory of evidence. Image and Vision Computing 24 (1), 61–69.

Redelsperger, J.L., Thorncroft, C., Diedhiou, A., Lebel, T., Parker, D.J., Polcher, J., 2006. African monsoon multidisciplinary analysis (AMMA): An international research project and field campaign. Bulletin of American Meteorological Society 87 (12), 1739–1746.

Richter, R., 1996. Atmospheric correction of satellite data with haze removal including a haze/clear transition region. Computers and Geosciences 22 (6), 675–681.

Richter, R., Muller, A., 2005. De-shadowing of satellite/airborne imagery. International Journal of Remote Sensing 26 (15), 3137–3148.

Ripley, B.D., Kelly, F.P., 1977. Markov point processes. Journal of the London Mathematical Society 15 (1), 188–192.

Rossow, W.B., Garder, L.C., 1993. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. Journal of Climate 6 (12), 2341–2369.

Simpson, J.J., Stitt, J.R., 1998. A Procedure for the detection and removal of cloud shadow from AVHRR data over land. IEEE Transactions on Geoscience and Remote Sensing 36 (3), 880–897.

Van Lieshout, M.N.M., 2000. Markov Point Processes and their Applications. Imperial College Press, London.

Viterbi, A., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory 13 (2), 260–269.

Zhang, Y., Guindon, B., Cihlar, J., 2002. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. Remote Sensing of Environment 82 (2–3), 173–187.