

Unsupervised Segmentation of low Clouds From Infrared METEOSAT Images Based on a Contextual Spatio–Temporal Labeling Approach

Christophe Papin, Patrick Bouthemey, and Guy Rochard

Abstract—The early and accurate segmentation of low clouds during the night-time is an important task for nowcasting. It requires that observations can be acquired at a sufficient time rate as provided by the geostationary METEOSAT satellite over Europe. However, the information supplied by the single infrared METEOSAT channel available by night is not sufficient to discriminate between low clouds and ground during night from a single image. To tackle this issue, we consider several sources of information extracted from an infrared image sequence. Indeed, we exploit both relevant local motion-based measurements, intensity images and thermal parameters estimated over blocks, along with local contextual information. A statistical contextual labeling process in two classes, involving “low clouds” and “clear sky,” is performed on the warmer pixels. It is formulated within a Bayesian estimation framework associated with Markov random field (MRF) models. This comes to minimize a global energy function comprising three terms: two data-driven terms (thermal and motion-based ones) and a regularization term expressing *a priori* knowledge on the label field (expected spatial contextual properties). We propose a progressive minimization procedure of this energy function starting from initial reliably labeled pixels and involving only local computation. Thermal parameters associated to each class are estimated according to an unsupervised learning scheme enabling the handling of spatiotemporal nonstationarities. Our method produces segmentation maps displaying temporal coherency along the image sequence. Experimental results on representative meteorological situations are reported and favorably compared with NOAA/AVHRR cloud classifications which serve as reference results. They demonstrate the accuracy and efficiency of the proposed approach.

Index Terms—Contextual labeling, low clouds detection, Markov random field (MRF) models, METEOSAT satellite images, spatio-temporal image segmentation, thermal parameters estimation.

I. INTRODUCTION

LOW CLOUDS, such as stratus and fog, often cause serious concerns for air and road traffic safety. The earliest possible detection of the occurrence of low clouds during night-time is required to enable nowcasters to deliver reliable weather forecasts early in the morning. During the day, low clouds can be easily detected using both visible and infrared satellite images,

unless higher clouds hide them. At night, the visible channel is of course not available, while water vapor channel is inappropriate to analyze low layers of the atmosphere. If we consider polar satellite NOAA/AVHRR images, such a detection remains possible by exploiting the different emissivity ϵ of opaque water clouds (as fog) in the $3.7 \mu\text{m}$ channel ($\epsilon \approx 1$) and in the $11 \mu\text{m}$ channel ($\epsilon \approx 0.8\text{--}0.9$) (usually denoted by channel number resp. three and five) [8]. Contrary to the earth surface, the captured brightness temperature of a cloud structure is lower in channel three than in channel five, and using the difference between these two temperature fields allows one to straightforwardly solve this issue. However, due to the polar orbit of this satellite, this can only be performed at distant time points. Short-time forecasting requires a higher frequency of observations, which is provided by the geostationary METEOSAT satellite covering Europe. However, low clouds are often quite difficult to delineate in an image derived from the single thermal infrared METEOSAT channel available at night. Indeed, there is usually a poor thermal contrast between the top of such clouds and the underlying earth surface.

Numerous cloud classification methods relying on multi-spectral and textural information have already been developed [12], [22]. The use of statistical texture information along with spectral measurements increases the dimension of the feature space but can yield pertinent supplementary information. On the other hand, integrating temporal information in a cloud detection scheme as presented in [3] has been seldom investigated.

The most common approach is to classify each pixel on an individual basis. Two kinds of methods have been explored. Supervised cloud classification exploits a set of samples considered to be characteristic of the various classes by experts. A training stage is carried out beforehand, providing a partitioned feature space. Such techniques can be parametric or not. The parametric approach often involves the maximum likelihood estimation (MLE) of normal distribution parameters [7], [9]. Among the nonparametric methods, some local approaches rely on successive threshold tests applied to different combination of channels [5], [26], whereas others exploit multi-dimensional thresholding to partition the feature space [15], [18] or make use of the nearest-neighbor criterion [2], [17]. Another popular cloud classification scheme is the use of multi-layer perceptron neural networks [23]. They have led to accurate cloud classifications of high resolution images NOAA/AVHRR, with a computation time compatible with an operational usage (once the training stage is carried out).

Manuscript received October 23, 2000; revised May 14, 2001. This work was supported in part by Météo-France.

C. Papin and P. Bouthemey are with Irisa/Inria Rennes, France (e-mail: bouthemey@irisa.fr).

G. Rochard is with the Centre de Météorologie Spatiale, Météo-France, Lannion.

Publisher Item Identifier S 0196-2892(02)01421-3.

The second class of approaches are unsupervised methods. A statistical partition method called dynamic clustering was applied to the METEOSAT channels to provide a cloud classification into a given number of classes [6].

However, these pixelwise methods are prone to produce noisy results without postprocessing. Moreover, such local techniques appear not really appropriate to handle the issue of image segmentation into regions. Contextual labeling approaches [10], [16] have proven more efficient in that sense. In [16], two distinct decision rules embedded in a contextual classification scheme applied to NOAA/AVHRR images have been proposed. The dependence between neighboring pixels is taken into account, which consequently reduces noise effect while preserving small regions. In [10], the authors developed a hybrid model exploiting information both of structural and contextual nature. Spatial contextual information is formalized through a Markovian texture model. A structural model describes the image patterns from a frequency representation of the image.

To deal with the detection of low clouds by night from the METEOSAT infrared channel, we consider several kinds of inputs: motion-based measurements, information on the thermal structure of the scene, and contextual information. We propose an original cloud classification scheme into two classes corresponding to “low clouds” and “clear sky.” This process is performed only over areas of interest supplied by a preliminary stage. This classification issue is formulated as a statistical contextual labeling problem and is embedded in a Bayesian estimation framework associated with Markov random field (MRF) models. This labeling problem comes to minimize a global energy function involving several terms. The first two are data-driven terms (thermal and motion-based), the third one is a regularization term expressing *a priori* knowledge on the label field (expected contextual properties). We have designed an efficient minimization procedure of this energy function starting from initial reliably labeled pixels. The thermal parameters associated to each class are estimated in an unsupervised way and spatiotemporal nonstationarities are correctly handled.

The remainder of the paper is organized as follows. In Section II, we present the motivations of this study and we outline the main principles of our approach. Section III introduces the considered image features (or observations), and is particularly devoted to the analysis of the local motion-based measurements. In Section IV, we specify each component of the proposed model, we address the estimation of the thermal parameters, and we describe the minimization procedure applied to the designed energy function. In Section V, experimental results on a set of representative meteorological situations are reported, discussed, and compared with classification maps issued from NOAA/AVHRR images. Section VI contains concluding remarks.

II. PROBLEM STATEMENT

The only METEOSAT infrared channel usable for detecting low clouds at night is the [10.5–12.5 μm] thermal one. In that

case, multispectral analysis is not possible. Detecting stratus and fog by night from a single infrared image is generally quite difficult. However, the animation of several successive images allows one to observe visible displacements usually in the vicinity of the borders of such clouds. Therefore, to distinguish low clouds from earth surface, we aim at combining temporal information, issued from relevant motion-based measurements, with intensity and contextual information.

Our goal is not to produce a complete classification map involving all the different kinds of clouds. Many classification schemes applied to METEOSAT images have already been developed and run in an operational mode. Nevertheless, if most of the cloud types are satisfactorily identified, low clouds, and especially the stratiform low clouds (SLC) at night, are quite difficult to extract from the surrounding earth surface due to their frequent similarity in temperature.

We restrict ourselves to segment low level clouds from ground. Moreover, we will not try to distinguish between a radiation fog and stratus. We start from cloud classification maps provided by CMS (Centre de Météorologie Spatiale of Météo-France) [2], and we apply our two-class segmentation method only within areas classified as low clouds or clear sky, the other classes being assumed well-labeled. Let us note that in a preliminary version of this work, we introduced an additional class called “other clouds” as described in [25], and subsequently, the three-class segmentation process was performed over the entire image. It turned out that considering only two-classes led to similar results, while simplifying the modeling stage, suppressing the introduction of supplementary parameters and saving CPU time.

III. DEFINITION OF OBSERVATIONS AND LABELS

We consider as observations (or feature vectors) image intensities $I(p)$ (numerical counts) and motion-based measurements $\omega(p)$, where p designates a pixel. We can easily deduce the corresponding temperatures from intensity values $I(p)$ by using the calibration tables provided by EUMETSAT.

The motion-related feature $\omega(p)$ is derived from the optic flow constraint equation (OFCE) [14]

$$\mathbf{v}(\mathbf{p}) \cdot \nabla I(p) + I_t(p) = 0 \quad (1)$$

where \mathbf{v} denotes the two-dimensional (2-D) image flow vector, ∇I the spatial intensity gradient, and I_t the temporal intensity derivative. This equation relates the spatio-temporal intensity derivatives to the apparent motion in the image by assuming the constancy of image brightness along 2-D motion trajectories. We can infer from equation (1) the expression of the projection ω_n of the 2-D motion vector on the direction of the spatial intensity gradient as follows:

$$\omega_n(p) = -\frac{I_t(p)}{\|\nabla I(p)\|} \quad (2)$$

ω_n , also called normal component of the flow vector or normal flow, is the only locally computable motion component without additional constraints.

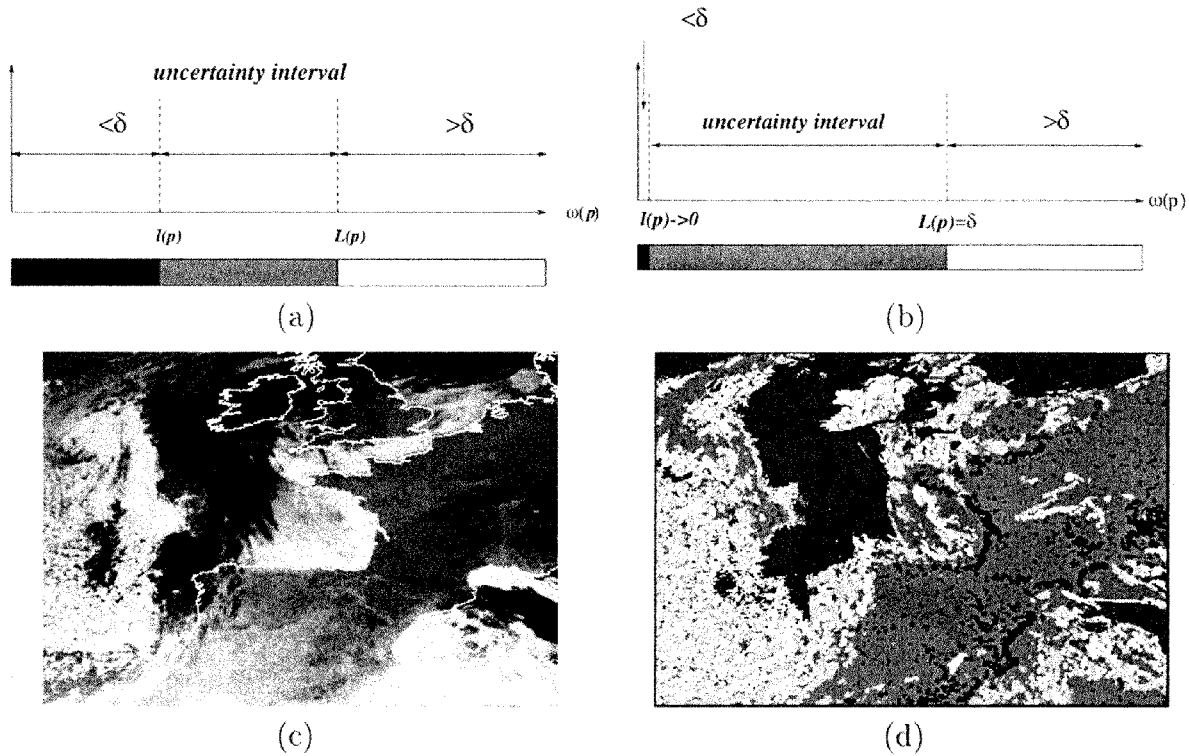


Fig. 1. Evolution of the uncertainty interval $[l(p), L(p)]$ according to the spatiotemporal context. (a) Case of an isotropic distribution of the spatial intensity gradient: the gap between the two bounds is the lowest and most of the motion-based measurements can be reliably interpreted. (b) Case of uniform areas: the gap between bounds is the largest and in particular small values of $\omega(p)$ cannot be unambiguously interpreted since $l(p)$ tends to zero. (c) Infrared Meteosat image. (d) Computed confidence map for $\omega(p)$ values: the map is displayed with the grey level code depicted in (a) or (b). Dark grey areas are not processed.

The motion-based feature $\omega(p)$ we introduce is defined as a local weighted average of $\omega_n(p)$. It is given by [21]

$$\omega(p) = \frac{\sum_{r \in \mathcal{W}_p} (\|\nabla \mathbf{I}(r)\| \times |I_t(r)|)}{\max \left(NG_m^2, \sum_{r \in \mathcal{W}_p} \|\nabla \mathbf{I}(r)\|^2 \right)} \quad (3)$$

where \mathcal{W}_p is a window centered on p containing N pixels (typically, a 3×3 or 5×5 local window) and G_m is a predetermined constant which can be assimilated as noise level.

The interest of this motion-based feature $\omega(p)$ is that it is still locally computable as $\omega_n(p)$, but it is far less affected by image noise due to the averaging process, it can be evaluated at any pixel p and above all, it is possible to assess the actual information conveyed by $\omega(p)$ according to the local distribution of the spatial intensity gradient as explained in Section III-A. These local measurements are nevertheless not valid in case of large displacements or within occlusion areas. We have then to take into account these two restrictions, as addressed in Sections III-B and III-C.

A. Evaluation of the Motion-Based Measurements

SLC are characterized by a rather smooth and featureless appearance. Therefore, it is important to determine the actual information conveyed by the motion-based measurement $\omega(p)$ at each pixel according to the spatial intensity gradient distribution

in its vicinity. This can also be viewed as a reliability or confidence assessment. As demonstrated in [21], lower and upper bounds on the observation $\omega(p)$ can be determined for a given minimal displacement magnitude δ to be detected.

A value of $\omega(p)$ greater than the upper bound $L(p)$ indicates that the underlying displacement magnitude is at least equal to δ . In contrast, a value of $\omega(p)$ smaller than the lower bound $l(p)$ reveals a displacement magnitude lower than δ . The two bounds $l(p)$ and $L(p)$ depend on the spatial intensity gradient distribution in the vicinity of p and on the considered displacement magnitude δ as illustrated in Fig. 1. On the other hand, nothing can be straightforwardly stated if $\omega(p)$ lies within $[l(p), L(p)]$, and then, contextual information must be exploited.

B. Handling of Large Displacements

It is reasonable to state that low clouds usually undergo small motions, and that pixels belonging to low clouds generally verify the OFCE equation.

Nevertheless, if the OFCE is locally not valid due to large displacements, a multiresolution approach can be adopted, and a coarse-to-fine scheme from the lowest to the original image resolution is usually employed to compute motion fields [19]. In our case, we only need to select the proper image scale at each pixel p to correctly compute the observation $\omega(p)$.

To this end, we have developed a local multi-resolution analysis involving a log-likelihood ratio test to verify the validity of the OFCE. This test was introduced in [13]. We briefly outline it

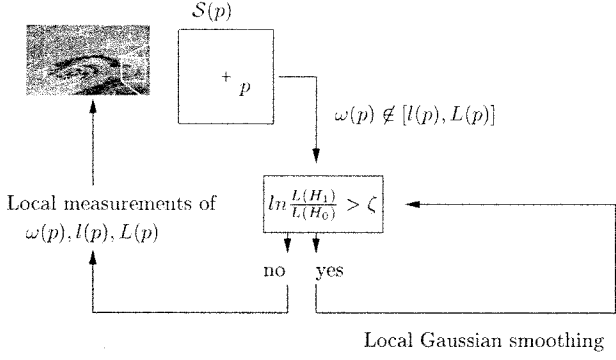


Fig. 2. Overview of the scheme associating a local multiresolution analysis with a log-likelihood ratio test to verify the validity of the OFCE in order to compute the motion-based observation $\omega(p)$ at the appropriate scale (Gaussian filtering with increasing variance).

hereafter. A first-order model of the intensity function I is considered, respectively, at time t and $t + 1$ in the vicinity of pixel $p = (x, y)$

$$\begin{cases} I(x + \delta x, y + \delta y, t) \\ = I(x, y, t) + a_t \delta x + b_t \delta y + n_1 \\ I(x + \delta x, y + \delta y, t + 1) \\ = I(x, y, t + 1) + a_{t+1} \delta x + b_{t+1} \delta y + n_2 \end{cases} \quad (4)$$

where n_1 and n_2 represent two independent Gaussian noises of zero mean and same variance. The point is to check if the first order parameters in (4) are the same at time instants t and $t + 1$, which will assess the validity of the OFCE at pixel p . Two competitive hypotheses H_0 and H_1 are defined [13]

$$\begin{cases} H_0: \{a_t = a_{t+1} \text{ and } b_t = b_{t+1}\} \text{ in } S_p, \text{ OFCE valid} \\ H_1: \{a_t \neq a_{t+1} \text{ or } b_t \neq b_{t+1}\} \text{ in } S_p, \text{ OFCE not valid} \end{cases} \quad (5)$$

where S_p designates the estimation support centered on p (typically, a 3×3 or 5×5 window). A log-likelihood ratio test is then designed to decide which hypothesis must be selected.

The motion-based measurements $\omega(p)$ computed at the original image resolution verifying $\omega(p) \notin [l(p), L(p)]$ will be considered as exploitable. The procedure described below and illustrated on Fig. 2 is only applied to pixels p such that $\omega(p) \notin [l(p), L(p)]$ at the original image resolution. We start from the original image resolution. If $\ln(L(H_1)/L(H_0)) > \zeta$, where $L(H_i)$ denotes the likelihood function associated to hypothesis H_i , $i = 0, 1$, and ζ is a predetermined threshold, the OFCE is not valid. Then we proceed to the next lower image resolution by locally smoothing the intensity function using a Gaussian filter of variance σ , and the test is evaluated again. The process is iterated until the test is not satisfied. When the image resolution level verifying the OFCE is reached, we calculate $\omega(p)$ along with the interpretability bounds $l(p)$ and $L(p)$.

C. Treatment of Occlusion Areas

Occlusion areas correspond to situation where, from one time instant to the next one at a given image location, a ground part previously hidden by clouds is uncovered, or conversely, a previously visible ground part is then covered by clouds. In these configurations, the OFCE is obviously not valid and consequently, it is necessary to identify such occlusion areas.

This problem is of particular importance since most of the reliable motion-based measurements verifying $\omega(p) \geq L(p)$ are located in the vicinity of the SLC boundaries. On the maritime zones, we exploit strong *a priori* knowledge stating that the temperature of the cloud top is always greater or equal to the surrounding sea temperature [3]. The sign of the temporal derivative $I_t(p)$ of the intensity at pixel p is then sufficient to determine the nature of an occlusion situation (occluding or occluded situation) and to decide if the cloud is present at time t or time $t + 1$. In addition, it is possible to determine, as explained in Section IV-A, which of a low cloud or ground is the warmer entity in terrestrial zones.

More precisely, the presence and type of an occlusion area at time t can be determined by an elementary test relying on the sign of the temporal intensity difference $I_t(p) = I(p, t) - I(p, t - 1)$ and applied to all pixels p for which the OFCE is not valid

$$\begin{cases} \text{if sign } [I_t(p)] = \text{sign } (\mu_{cs} - \mu_{lc}), \\ \text{then } p \text{ can be assumed to belong to a } lc \text{ region} \\ \text{if sign } [I_t(p)] \neq \text{sign } (\mu_{cs} - \mu_{lc}), \\ \text{then } p \text{ can be assumed to belong to a } cs \text{ region} \end{cases} \quad (6)$$

where μ_{cs} and μ_{lc} respectively correspond to locally estimated representative temperatures of respectively classes “*cs*” and “*lc*” (as further explained in Section IV-A), where “*cs*” stands for clear sky and “*lc*” for low cloud.

IV. STATISTICAL MODELING

As previously stated, we have formulated the detection of low clouds as a contextual labeling issue. More formally, we have adopted a statistical approach relying on Markov Random Field (MRF) models and the Bayesian maximum *a posteriori* (MAP) criterion. Here, the observations o comprise the intensity values $I(p)$ and the motion-based measurements $\omega(p)$. The label field to be determined $e = \{e(p)\}$ involves two classes “*lc*” (low clouds) and “*cs*” (clear sky). According to Bayes rule and due to the equivalence between MRF and Gibbs distributions [11], maximizing the posterior conditional probability $p(e|o)$ leads to the minimization of a global energy function $U(o, e)$ with respect to the unknown variables e . We have defined an energy function consisting of two data-driven terms, respectively related to temperature observations $I(p)$ and motion-based measurements $\omega(p)$, and a regularization term involving *a priori* information in terms of contextual information. This energy function is given by

$$\begin{aligned} U(o, e) &= \underbrace{U_1(I, e) \text{ (temperature term)} + U_2(\omega, e) \text{ (motion term)}}_{\text{data-driven terms}} \\ &= \underbrace{\sum_{p \in \mathcal{S}} V_1(I(p), e(p)) + \sum_{p \in \mathcal{S}} V_2(\omega(p), e(p))}_{\text{data-driven terms}} \\ &+ \underbrace{\alpha \sum_{\langle p, r \rangle \in \mathcal{C}} V_3(e(p), e(r))}_{\text{regularization term}} \end{aligned} \quad (7)$$

where S denotes the set of sites p (or pixels), and \mathcal{C} represents all the binary cliques associated with the considered second-order neighborhood system on the set of sites. Let us recall that a binary clique is formed by two pixels which are mutual neighbors. V_i , $i = 1, 2, 3$ stand for so-called potential functions, which are locally defined (on sites or cliques). Parameter α weights the respective influence of the data-driven terms and the regularization term. Before describing the different components of this model, let us first present the unsupervised learning procedure to estimate the parameters involved in the temperature data-driven term $U_1(I, e)$. They will be called “thermal” parameters in the subsequent sections.

A. Unsupervised Estimation of Thermal Parameters

We must tackle three main difficulties. First, spatial ambiguities may occur on the thermal information. Indeed, two distant pixels belonging to two different classes may have similar temperature. On the other hand, a given class may present different thermal emissivity in two different regions. Secondly, we have to cope with temporal temperature variations stemming from the cooling phenomena during night. This can be handled by performing the parameter learning step for each image of the sequence provided the parameter estimation procedure is not too time consuming. Thirdly, since the learning step will be solved in a statistical way, the amount of usable observations must be significant.

To account for the spatial temperature variability, the image is divided into a regular grid of blocks, and the parameter estimation is performed in each block. A tradeoff must be found between a block size large enough to provide accurate estimates, and small enough to deal with spatial nonstationarities. Numerical experiments were conducted in [24] to study the influence of the choice of the partition scale. In the experiments reported later, block size varies from 35×35 to 90×90 pixels.

1) *Computation of Relevant Histograms:* We have to compute relevant histograms for each class (i.e., “ lc ” class and “ cs ” class) in each block to properly estimate the thermal parameters corresponding to each class (if present in the considered block). Motion information can be fruitfully exploited to this end. Indeed, a pixel undergoing motion is necessarily a cloud pixel, and a clear sky pixel is supposed to be static. This can be accounted for by taking into account only temperature values at pixels verifying $\omega(p) \notin [l(p), L(p)]$. Thermal parameters to be estimated are the mean μ_k and variance σ_k^2 in each block \mathcal{B}_k for each class. Moreover, the occurrence of several modes (in practice two modes) is also considered for each class in every block. The “clear sky” class also comprises land and sea areas which have different physical properties. Thus, we separately process these two types of areas by exploiting a binary ground mask \mathcal{G} segmenting land and sea.

Sea Areas: Due to the low spatial resolution of the METEOSAT infrared images, no apparent motion can be perceptible on the sea surface. Consequently, sea areas can be distinguished from cloud areas according to the computed motion measurements. However, the poor thermal contrast over sea regions prevents us from getting unambiguously interpretable motion measurements as previously explained [i.e., lower bound $l(p) \rightarrow 0$].

Then, we only build in each block included in those regions the temperature histogram defined by $\mathcal{H}_k^l = \{I(p), p \in \mathcal{B}_k / \omega(p) \geq L(p)\}$. By assuming that low clouds over the sea are colder than the sea surface, we can readily determine uncovered background from moving pixels with the following decision test

$$\begin{cases} \text{if OFCE not valid, } \omega(p) \geq L(p), \text{ and } \text{sign}(I_t(p)) > 0, \\ \quad \text{then } p \text{ belongs to a } lc \text{ region} \\ \text{if OFCE not valid, } \omega(p) \geq L(p), \text{ and } \text{sign}(I_t(p)) < 0, \\ \quad \text{then } p \text{ belongs to a } cs \text{ region.} \end{cases} \quad (8)$$

Hence, we can finally build two histograms $\mathcal{H}_k^l(lc)$ and $\mathcal{H}_k^l(cs)$.

Ground Areas: Continental zones are usually more contrasted than sea areas. It becomes possible to take into account temperature of pixels whose $\omega(p)$ values are lower than $l(p)$ since in that case $l(p) > 0$. We construct the temperature histogram $\mathcal{H}_k^l(cs) = \{I(p), p \in \mathcal{B}_k / \omega(p) < l(p)\}$ corresponding to points belonging to ground areas with a very small value of δ (typically, $\delta = 0.5$) accounting for noise. For these static pixels, we can use the temperature values at p at both time instants $t-1$ and t . We also need to determine the occlusion areas. To this end, we detect occurrences of temperature inversion over ground. Histograms $\mathcal{H}_k^l(cs)$ are compared to histograms \mathcal{H}_k^l of the intensity values in the same block \mathcal{B}_k , $\mathcal{H}_k^l = \{I(p), p \in \mathcal{B}_k\}$. A temperature inversion can be inferred in \mathcal{B}_k if the average temperature deduced from $\mathcal{H}_k^l(cs)$ is lower than the one determined from \mathcal{H}_k^l . Such an identification process allows us to build histograms $\mathcal{H}_k^l(lc) = \{I(p), p \in \mathcal{B}_k / \omega(p) > L(p)\}$. Only pixels verifying the OFCE or satisfying the appropriate constraint on the sign of the temporal intensity derivative (i.e., $\text{sign}[I_t(p)] > 0$ in case of thermal inversion or $\text{sign}[I_t(p)] < 0$ otherwise) are exploited to construct histograms $\mathcal{H}_k^l(lc)$.

2) *Estimation of the Thermal Parameters:* From the computed histograms defined earlier, we evaluate the thermal parameters attached to each class in every block \mathcal{B}_k . More precisely, we estimate parameters of a mixture of distributions assumed to be Gaussian. Let us denote $\theta_i^j(k) = [\mu_i^j(k), (\sigma_i^j)^2(k)]$ in a given block \mathcal{B}_k for class i and thermal modes j , where $\mu_i^j(k)$ and $(\sigma_i^j)^2(k)$ denote the mean and variance, respectively, of the corresponding Gaussian distribution.

If the empirical variance of the intensity distribution represented by \mathcal{H}_k is lower than a given threshold ξ , we consider that the intensity distribution is not complex enough to introduce more than one Gaussian law, and then $\mu_i^j(k)$ and $(\sigma_i^j)^2(k)$ are respectively given by the empirical mean and variance. Otherwise, we introduce a mixture of two Gaussian laws. Knowing the number of Gaussian components, the parameter estimation is achieved using the “Expectation-Maximization” (EM) algorithm [4]. A test of bi-modality [20] is then performed. If satisfied, the estimated parameters $\theta_i^j(k)$, $j = 1, 2$ will be used in the temperature data-driven term.

It can sometimes occur that the histogram of temperatures related to a given class i is not significant enough to provide pertinent estimates of the parameters in some blocks. In that case, we search for the nearest block where the corresponding parameters of class i have been reliably estimated, and we assign the same parameter values to the considered block.

B. Definition of the Potential Functions

We will now describe the expressions of the potential functions introduced in the global energy function defined in relation (7).

Potential V_1 expresses the adequacy between the label field e and the intensity image (or equivalently temperature map) I . Potential V_1 is given by

$$V_1(I(p), e(p)|\mathcal{G}(p)) = \min_{[j=1, N_i]} \left[1 - \exp \left(- \frac{(I(p) - \mu_i^j)^2}{(\sigma_i^j)^2} \right) \right] \quad (9)$$

where i stands for one of the two considered classes “ lc ” and “ cs ” and j for the intensity mode within class i . We have dropped the index k attached to block \mathcal{B}_k in variables N_i^j , σ_i^j and μ_i^j for notation convenience. Pixel p determines the appropriate block \mathcal{B}_k to be considered. We handle in this potential function the possible presence of several characteristic temperatures related to a given class within the same area. Such multimodal situations can be caused for instance by local transparency phenomena or be due to particular conditions such as cold snowy mountains within warmer uncovered areas. The introduction of the robust Leclerc estimator $\rho(u) = 1 - \exp(-u^2/\zeta)$ in the potential definition allows us to cope with large deviations (outliers) from the assumed thermal model. We have plotted an example of such a potential function in Fig. 3. The warmest mode stands for sea surface temperature whereas low clouds are characterized by two modes in that example.

Potential V_2 accounts for the intrinsic mobility and the smooth intensity appearance of SLC, while taking into account interpretability information on motion-based measurements $\omega(p)$. It is defined by

$$V_2(\omega(p), e(p)) = \begin{cases} \gamma_{l(p)} \frac{(a\omega(p) - b)^2}{1 + (a\omega(p) - b)^2} & \text{if } e(p) = lc \text{ and } \omega(p) < L(p) \\ 0 & \text{if } e(p) = lc \text{ and } \omega(p) \geq L(p) \\ 1 - \exp\left(-\frac{\omega(p)^2}{\zeta_{cs}}\right) & \text{if } e(p) = cs \end{cases} \quad (10)$$

with

$$\gamma_{l(p)} = \frac{2l(p)}{\delta}, \quad a = \frac{b}{L(p)} \quad \text{and} \quad b = \frac{1}{3} \frac{\sqrt{3}L(p)}{l(p) - L(p)}.$$

If $\omega(p)$ lies within $[l(p), L(p)]$, i.e., no direct reliable interpretation of $\omega(p)$ is available, it is preferable not to favor any of the two labels. On the other hand, if $\omega(p) \geq L(p)$, assignment of label “ lc ” must induce no cost. Continuity constraint on V_1 leads to set $a = (b/L(p))$. In the expression of $V_2(\omega(p), lc)$, the robust McClure M-estimator [1] is introduced to cope with large deviations from the defined motion-based model. It is weighted by factor $\gamma_{l(p)}$. In case of an isotropic distribution of gradient

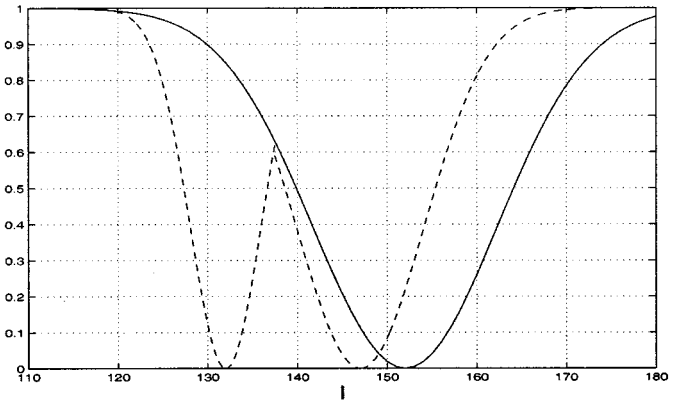


Fig. 3. Plot of an example of potential function V_1 attached to a block located in a sea area. Two modes have been introduced (dashed line) for the low cloud class and one mode (solid line) for the clear sky class.

intensity in the vicinity of pixel p , which is a favorable configuration, $l(p)$ is equal to $\delta/2$, and then $\gamma_{l(p)} = 1$, while degenerated intensity gradient distribution will make $\gamma_{l(p)}$ tends to zero. According to the spatial intensity gradient context, low values of $\omega(p)$ will indeed imply more or less pronounced penalties when assigning label “ lc .” $V_2(\omega(p), cs)$ is modeled as a Leclerc estimator which is more restrictive than McClure function. Satisfying the constraint $V_2(l(p) + (l(p) + L(p))/2, cs) = V_2(l(p) + (l(p) + L(p))/2, lc)$ leads to determine the expression of scale factor ζ_{cs} . Expression of parameter b is obtained by locating the point of inflection of $V_2(\omega(p), lc)$ in the middle of interval $[l(p), L(p)]$.

An example of potential V_2 is plotted in Fig. 4. Case (1) corresponds to an ambiguous situation (e.g., the core of a stratiform cloud). Interval $[l(p), L(p)]$ here exhibits its largest width $[l(p) \rightarrow 0, L(p) \rightarrow \delta]$. For a low $\omega(p)$ value, we coherently get a rather small difference between potential values corresponding to label “ lc ” and to label “ cs .” In case (2), the local intensity contrast is supposed to yield closer bounds ($[l(p) \rightarrow \delta/2, L(p) \rightarrow \delta/\sqrt{2}]$), and label “ lc ” is then more penalized for a motion-based observation $\omega(p)$ smaller than $(l(p) + L(p))/2$. Whatever the local intensity contrast, potential V_2 attached to label “ lc ” takes lower values for $\omega(p) \geq L(p)$. Indeed, we can assume that $\omega(p) \geq L(p)$ indicates a displacement of magnitude at least equal to δ , which is usually perceptible in the vicinity of low clouds boundaries or inside textured clouds. Let us note that the only parameter to set is the minimal displacement magnitude δ to be considered for the class of low clouds, which can be easily fixed by the user.

Regularization potential V_3 is defined on two-site cliques $\langle p, r \rangle$. It models the expected spatial properties of the label field. It is given by

$$V_3(e(p), e(r)|\nu_T(p)) = \begin{cases} +\beta, & \text{if } e(p) \neq e(r) \\ \beta \operatorname{atan}(\nu_T(p) - \tau), & \text{if } e(p) = e(r) = lc \\ -\beta, & \text{if } e(p) = e(r) = cs. \end{cases} \quad (11)$$

The goal is to favor by a negative potential value the extraction of homogeneous regions with rather smooth boundaries and to eliminate spurious isolated labels. This is of key importance

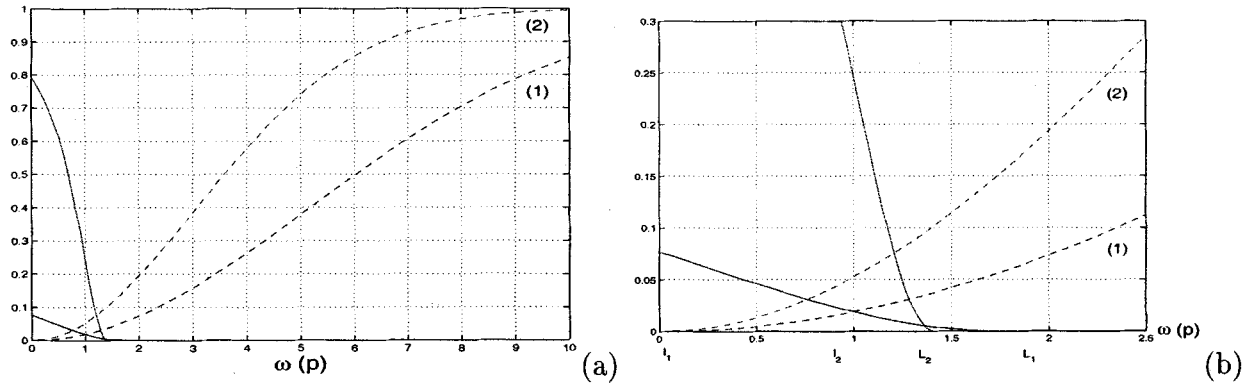


Fig. 4. Plot of the motion-related data-driven potential function V_2 for two different configurations with $\delta = 2$. 1) Homogeneous area ($[l(p) = 0.15, L(p) = \delta]$). 2) Isotropic distribution of spatial intensity gradients ($[l(p), L(p)] = [\delta/2, \delta/\sqrt{2}]$). Dashed line corresponds to the case $e(p) = cs$, solid line to $e(p) = lc$. (b) Enlargement of plot (a) close to the origin.

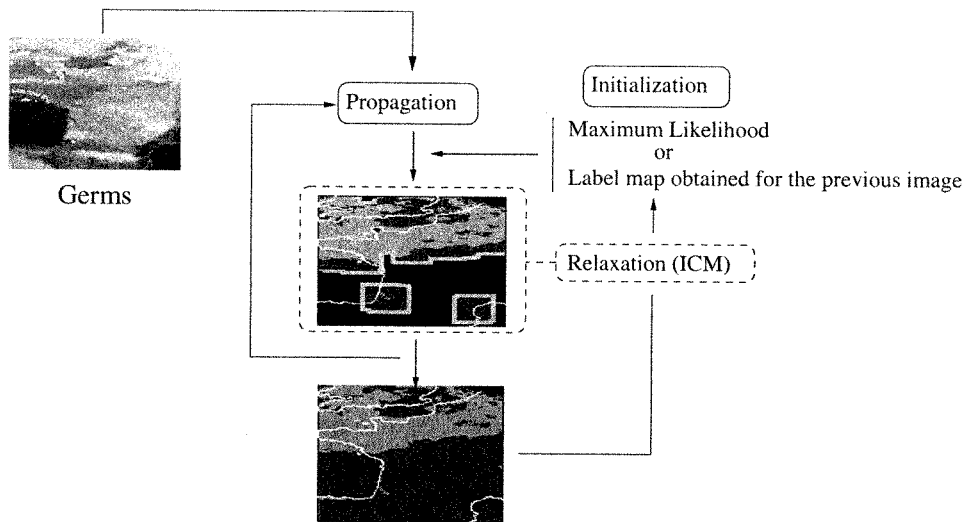


Fig. 5. Overview of the minimization scheme with spatial propagation. Initialization of newly introduced areas in the propagation stage, is performed by using an ML criterion in the first image of the sequence, and then by taking labels obtained in the previous image. Classification in two classes also exploits *a priori* contextual information.

inside areas supplying unreliable motion-based measurements. Potential expression corresponding to “ lc ” label involves the parameter $\nu_T(p)$, which denotes the local temperature variance at pixel p . The smaller the local infrared variance ν_T , the more favored configuration $e(p) = e(r) = lc$ ($V_3 < 0$). Here, we account for the usual intensity uniformity of low clouds. τ determines the value of ν_T , for which potential value becomes positive.

C. Minimization Scheme

Our formulation of the cloud classification problem leads to the minimization of the global energy function $U(o, e)$. To save computation time, we solve this issue in a deterministic way, which involves the risk to be stopped by local minima. A proper handling of the initialization step is then particularly important. Thus, we have designed an original minimization scheme based on the iterated conditional mode (ICM) deterministic iterative relaxation algorithm embedded in a spatially “progressive” scheme (as illustrated in Fig. 5). The principle is to propagate information from reliable points toward areas presenting uncertain motion-based measurements. The minimization process

first considers the neighborhoods of pixels which can be locally reliably labeled at this initial level, the latter are called seeds. The set \mathcal{R} of such pixels in a block \mathcal{B}_k is defined by

$$p \in \mathcal{R} \text{ if } \{p/\omega(p) \geq L(p) \text{ and } |I(p) - \mu_{lc}(k)| \leq \sigma_{lc}(k)\}. \quad (12)$$

Initial label assignment for seeds and neighboring sites is performed according to an ML criterion. The minimization of the global energy function is then performed in the neighborhood of these initially labeled sites, and is progressively extended over the image. This minimization scheme ends when the entire image has been processed.

V. RESULTS

We only report here the most representative examples of meteorological situations we have treated. More results can be found in [24]. Let us mention that all the obtained segmentation results have been assessed by forecasters from Meteo-France.

To validate the obtained results without available ground-truth data, we have compared them to cloud classifications, derived from NOAA/AVHRR images, computed using the pixel-

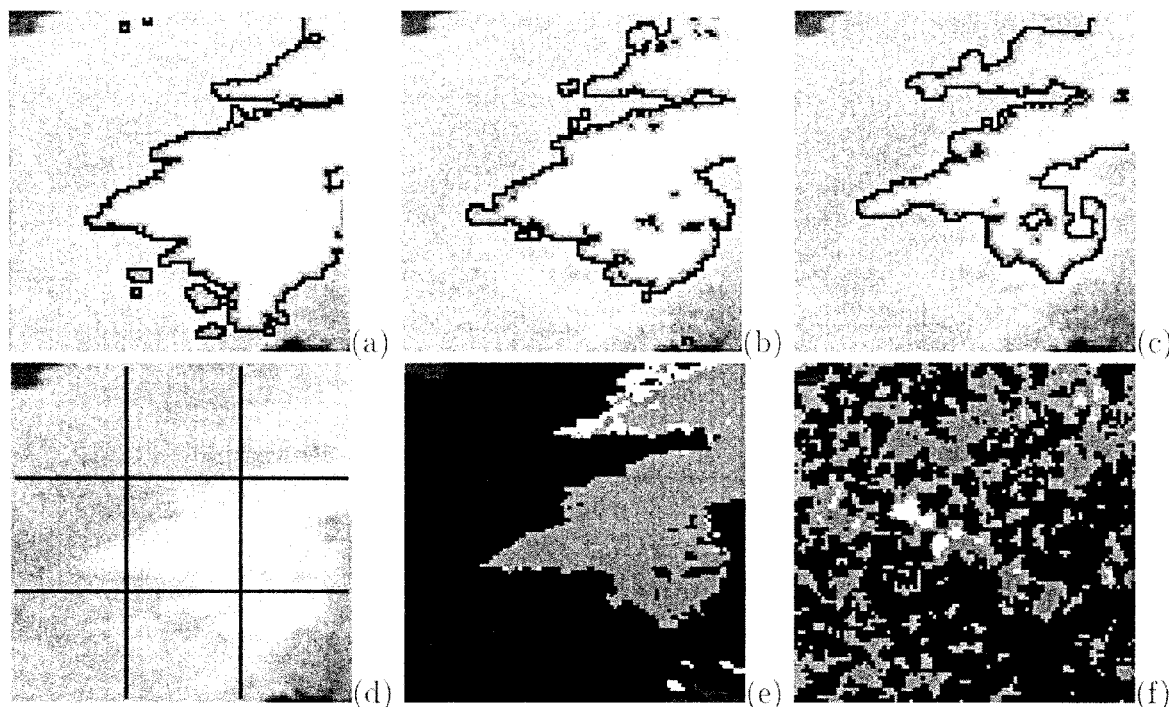


Fig. 6. Classification into low clouds and clear sky obtained for infrared METEOSAT images acquired on January 23, 1998 (a) at 1h30 UTC, (b) at 3h00 UTC, and (c) at 4h30 UTC. Clear sky areas are delineated in black and low clouds in grey. (d) Block partition used in the thermal parameter estimation stage. (e) Cloud classification (Brisson *et al.* 1997) issued from the processing of NOAA/AVHRR image acquired at 3h06 UTC. White and grey areas correspond, respectively, to “cloud boundaries” and sea areas. (f) Confidence map of motion-based measurements for the image acquired at 3h00 UTC. White areas designate pixels verifying $\omega(p) > L(p)$, black areas $\omega(p) < l(p)$, and grey areas $\omega(p) \in [l(p), L(p)]$.

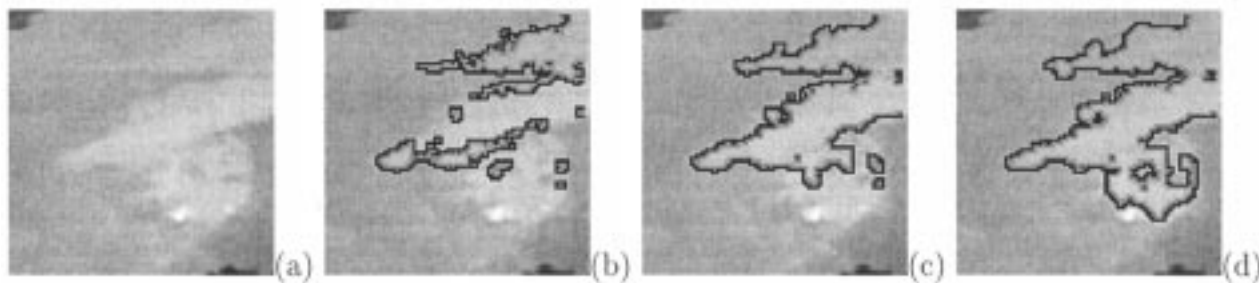


Fig. 7. Influence of the regularization term according to the value of the weighting factor α . (a) METEOSAT enhanced infrared image acquired on January 23, 1998, at 4h30 UTC. (b)–(d) Obtained classification, low clouds areas are delineated in grey and clear sky in black. (b) $\alpha = 0$, (c) $\alpha = 2$, and (d) $\alpha = 4$ with $\beta = 0.025$.

wise multispectral method described in [5]. Exploiting NOAA satellite images permits discriminating low clouds from ground in an easy way. Therefore, we consider such NOAA-based classification maps as “reference maps.” We mean that supplying a Meteosat-based segmentation map close to the NOAA-based classification can be assessed as a relevant result. The major benefit of our method is then to be able to provide accurate low cloud segmentations every 30 min (with Meteosat satellite images) instead of twice a night (with NOAA satellite images), which is of crucial interest for forecasters. Let us note that the latter are polarstereographic, while our results are displayed according to the satellite view point, and that the spatial resolution of NOAA/AVHRR images is about eight times higher.

The motion-related measurements involved in our segmentation method are computed between two successive Meteosat satellite images, i.e., they correspond to a temporal interval of 30 min.

The first example is displayed in Fig. 6(a)–(c). Results on three images acquired on January 23, 1998, are reported. The segmentation map of Fig. 6(c) can be compared to the NOAA classification shown in Fig. 6(e). These two maps are indeed quite similar. Let us stress that the low clouds located in the upper right corner of the image present a very low difference in temperature (between 1° and 2°) with the ground. Moreover, only few points of this area are initially reliably stated as moving points [Fig. 6(f), white points]. Nevertheless, the correctness of the final results to the corresponding reference map is quite satisfactory, and the obtained segmentation is accurate. We can also point out that small relevant regions are preserved and correctly labeled, which means that the regularization term is not too prominent. The parameter values were set as follows: $\delta = 2$, $G_m = 1$, $\alpha = 4$, $\beta = 0.025$ and $\tau = 2$.

In Fig. 7(b)–(d), the influence of the regularization energy term introducing local spatial contextual information is evalu-

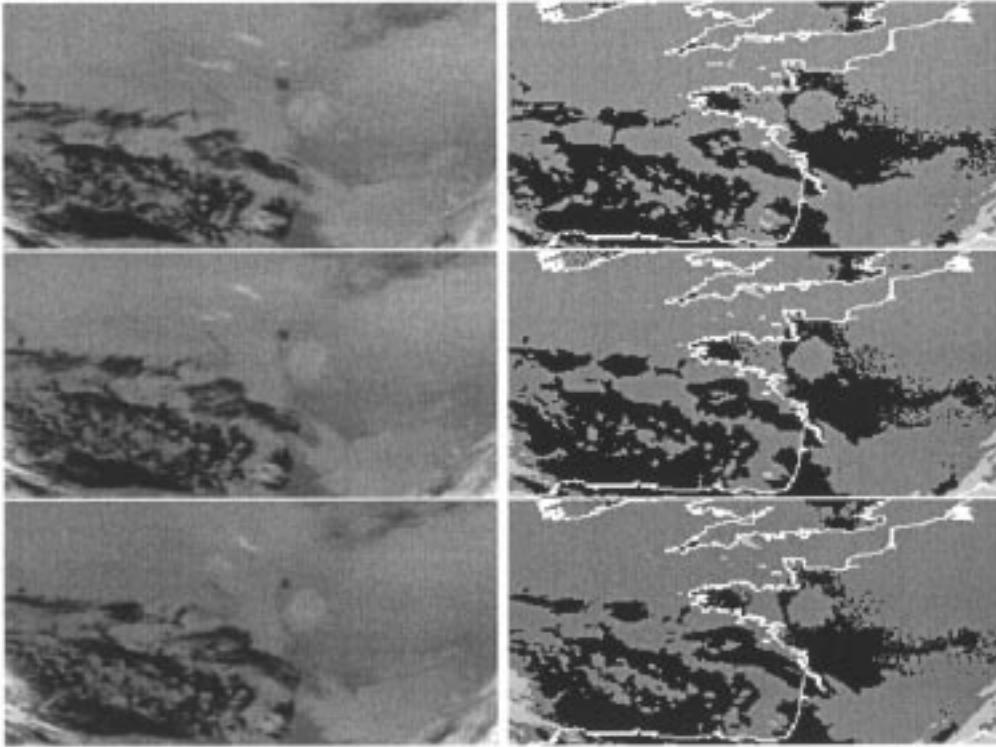


Fig. 8. Meteorological situation on January 7, 1997 from 0h30 UTC to 1h30 UTC from top to bottom. In the left column, METEOSAT enhanced infrared images; in the right column, obtained segmentation maps. Low clouds areas are in grey, clear sky in black, and other clouds in light grey.

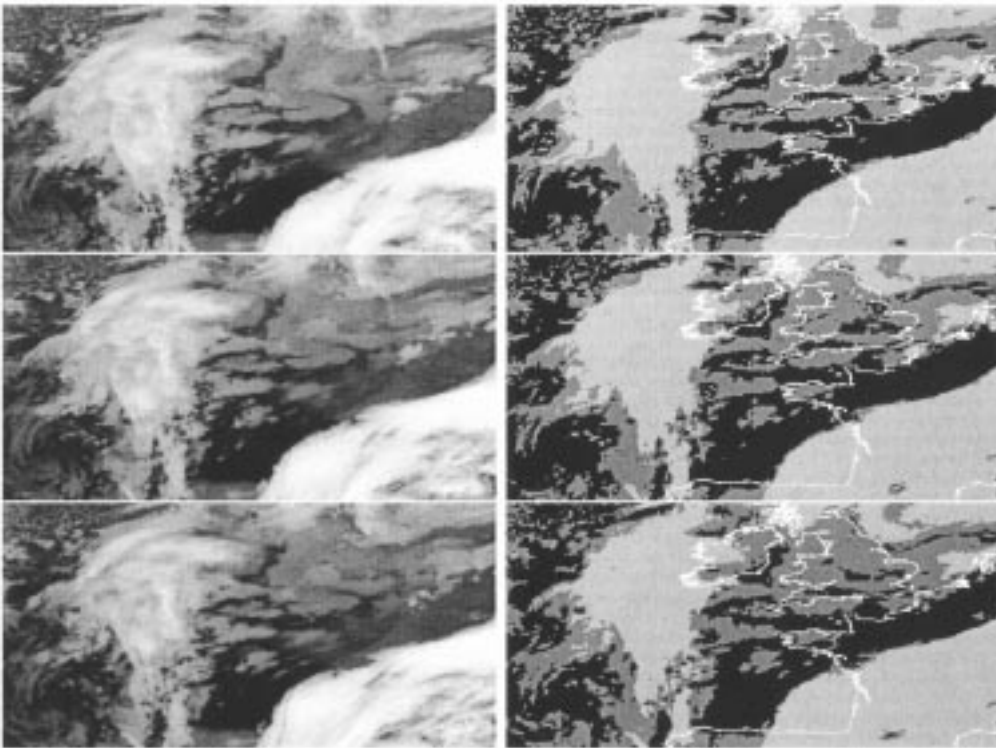


Fig. 9. Meteorological situation on January 23, 1997, from 2h00 UTC to 3h30 UTC from top to bottom. In the left column, METEOSAT enhanced infrared images; in the right column, obtained segmentation maps. Low clouds areas are in grey, clear sky in black, and other clouds in light grey.

ated by varying factor α . Three values are considered, $\alpha = 0$, 2, and 4. For $\alpha = 0$, i.e., without regularization, it is obvious that the data-driven term is not sufficient to correctly retrieve all the low clouds areas. It can only recover low clouds where the

motion-based measurements $\omega(p)$ are exploitable (center of the image) or in regions where temperature parameters have been well-identified in the learning stage (top of the image). The greater the value of α , the more contextual information is re-

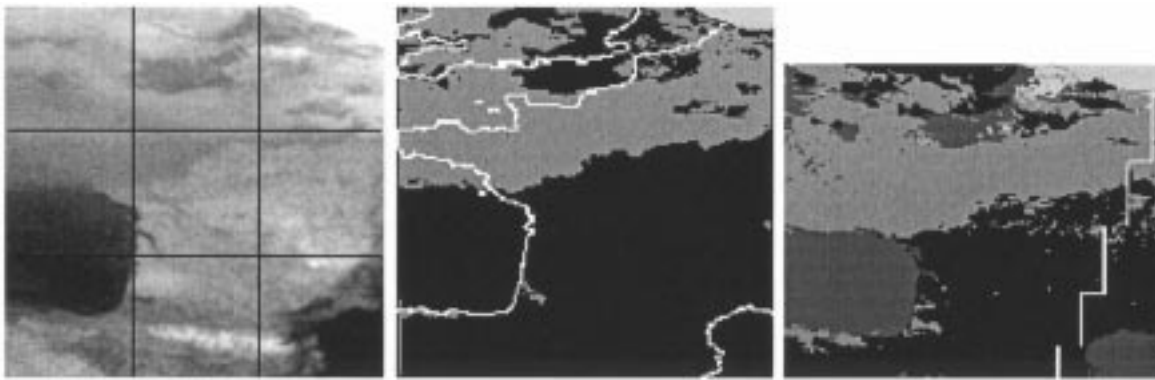


Fig. 10. Meteorological situation involving a thermal inversion over ground on October 2, 1997 at 3h30 UTC. In the left column, processed part of the METEOSAT enhanced infrared image with the block partition, considered to estimate thermal parameters, overprinted. In the center column, segmentation map obtained with our method. Low clouds areas are in grey, clear sky in black, and other clouds in light grey. In the right column, cloud classification [2] issued from the processing of NOAA/AVHRR image acquired at 3h26 UTC. Dark grey and light grey areas correspond respectively to sea areas and high clouds.

tained in the labeling process in areas of low temperature variance. In all cases, the spatial propagation effect diminishes as soon as ν_T becomes greater than τ (here, $\tau = 2$). In this example, the appropriate value of factor α is 4.

Let us point out that small areas of snowy Alps mountains not hidden by low clouds are well labeled as clear sky [see Fig. 7(d)]. They correspond to two spots of very high (white) intensity value near the bottom-right image corner in Fig. 7. Mountain areas are generally well-contrasted, and indeed provide a reliable support for the interpretation of motion-related measurements $\omega(p)$. Thus, such areas, often misclassified by classical methods, are correctly handled owing to our image segmentation method involving spatio-temporal measurements. This is again demonstrated in Fig. 10. Although the Pyrenees mountains in the South of France are snow-covered and then correspond to high intensity values, they are correctly labeled as clear sky.

If the contribution of the contextual information is decisive in most situations, its role should not be overestimated. It is in particular the case for the images shown in Fig. 8. This case includes a zone of fog in formation covering the Northeast part, and a group of stratus located over the South of France. These two cloud types are merged within the same class “low clouds.” If we are not able to distinguish them by means of temperature or texture information, they may nevertheless exhibit some distinct shape features: boundaries of stratus clouds are in general sharper than fog ones. If α is set to zero (Fig. 8), we can notice that contours of low cloud areas detected over the North part of France are then rather noisy, while those over the South part of France remains sharper. This could be exploited to further discriminate fog areas from stratus ones. However, it is reasonable to omit the regularization term, only if the thermal parameters can be quite correctly estimated in a sufficient number of blocks.

Another attractive property of our segmentation method is its intrinsic ability to supply coherent segmentation maps over time with the same parametrization, as shown in Figs. 8 and 9. This is straightforwardly achieved by initializing the label field to be determined in the current image with the final label field estimated in the previous image. Moreover, this saves computation time. In Fig. 9, the thermal contrast between low clouds and

ground is particularly low over Ireland and Great Britain. In that case, only a spatiotemporal image analysis allows us to deliver satisfactory results.

Thermal inversion situations over ground can be handled by our segmentation method as illustrated also in Fig. 10. By “thermal inversion,” we refer to meteorological phenomena where low clouds are in fact warmer in temperature than the ground surface under the clouds for some types of ground surface. More details on that issue can be found in [24]. In the example of Fig. 10, such a situation occurs over Brittany where the temperature of low clouds is greater (low intensity levels) than the one of the ground areas underneath. The comparison with cloud classification maps provided from both Meteosat and NOAA imagery shows that the behavior of our segmentation method is adequate. The difference in temperature between these clouds and the ground is about 1.5 °C only (less than four in terms of image intensity levels). Let us add that there are initially few reliable seeds in that area, and that the progressive relaxation scheme, described in Section IV-C, is decisive here.

The computational time is in accordance with operational requirements knowing that Meteosat satellite images are acquired every 30 min. The whole analysis process takes about seven minutes for subimages of size 396×276 on a Sun Ultra 60 workstation (360 Mhz).

VI. CONCLUSION

We have proposed a novel contextual statistical labeling method to segment low clouds from the uncovered ground surface using infrared METEOSAT images during the nighttime. This is of particular importance for every day weather analysis and meteorological forecast in Western Europe. The use of local motion-based measurements allows us to enlarge the information supplied by the single infrared METEOSAT channel usable by night. An unsupervised learning stage has been developed to estimate the parameters of the thermal modes of the two considered classes (“low clouds” and “clear sky”). It can cope both with spatial nonstationarities and temporal evolution of these parameters. Moreover, local thermal inversions can be detected. Our formulation of this segmentation problem leads to the minimization of a global energy function comprising

two data-driven terms and a regularization term bringing useful contextual information. To achieve this minimization, we have designed an original and efficient progressive minimization scheme relying on the iterative deterministic relaxation algorithm ICM which involves only local computation. Temporal coherence of the obtained cloud classification maps along the image sequence is also straightforwardly provided by our method. Segmentation maps can thus be animated, and provide fruitful insights on the weather situation to be analyzed. We have demonstrated the accuracy of the proposed approach by reporting satisfactory results obtained on representative cases. Our results are favorably compared with cloud classification maps issued from NOAA/AVHRR images by the method reported in [5].

ACKNOWLEDGMENT

The authors would like to thank the forecasters of Météo-France, Rennes, France, for their useful advice and their evaluation of the experimental results, and the DEMOS Department, CMS-Lannion, France. The authors would also like to thank G. Robinson for his careful reading of the manuscript and the anonymous reviewers for their valuable comments.

REFERENCES

- [1] A. Blake, "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 2–12, Jan. 1989.
- [2] A. Brisson, M. Dioszeghy, P. LeBorgne, and A. Marsouin, "Classification nuageuse Meteosat: Expérience 1996," Tech. Rep., Centre Météorol. Spatiale Lannion, Météo, France, 1997.
- [3] J. F. Cayula and P. Cornillon, "Cloud detection from a sequence of SST images," *Remote Sens. Environ.*, vol. 55, pp. 80–88, 1996.
- [4] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc.*, vol. B39, pp. 1–38, 1977.
- [5] M. Derrien, L. Farki, B. Harang, H. Le Gleau, A. Noyalet, D. Pochic, and A. Sairouni, "Automatic cloud detection applied to NOAA-11/AVHRR imagery," *Remote Sens. Environ.*, vol. 46, no. 3, pp. 246–267, 1993.
- [6] M. Desbois, G. Sèze, and G. Szejwach, "Automatic classification of clouds on Meteosat imagery: Application to high level clouds," *J. Appl. Meteorol.*, vol. 21, pp. 401–412, 1982.
- [7] E. E. Ebert, "A pattern recognition technique for distinguishing surface and cloud types in the polar regions," *J. Clim. Appl. Meteorol.*, vol. 26, pp. 1412–1427, 1987.
- [8] J. R. Eyre, J. L. Brownscombe, and R. J. Allam, "Detection of fog at night using advanced very high resolution radiometer (AVHRR) imagery," *Meteorol. Mag.*, vol. 113, pp. 266–271, 1984.
- [9] L. Garand, "Automated recognition of oceanic cloud patterns—Part I: Methodology and application to cloud climatology," *J. Climate*, vol. 1, pp. 20–39, 1988.
- [10] L. Garand and J. A. Weinan, "A structural-stochastic model for the analysis and synthesis of cloud images," *J. Clim. Appl. Meteorol.*, vol. 25, pp. 1052–1068, 1986.
- [11] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, June 1984.
- [12] Z. Q. Gu, C. N. Duncan, P. M. Grant, C. F. N. Cowan, E. Renshaw, and M. A. Mugglestone, "Textural and spectral features as an aid to cloud classification," *Int. J. Remote Sensing*, vol. 12, pp. 953–968, 1991.
- [13] F. Heitz and P. Boutheymy, "Multimodal estimation of discontinuous optical flow using Markov random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1217–1232, Dec. 1993.
- [14] B. Horn and B. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.
- [15] T. Inoue, "A cloud type classification with NOAA-7 split window measurements," *J. Geophys. Res.*, vol. 92, pp. 3991–4000, 1987.
- [16] J. Kittler and D. Pairman, "Contextual pattern recognition applied to cloud detection and identification," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-23, pp. 855–863, 1985.
- [17] J. Lee, R. C. Weger, S. K. Sengupta, and R. M. Welch, "A neural network approach to cloud classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 846–855, Sept. 1990.
- [18] E. Liljas, "Automated techniques for the analysis of cloud imagery," in *Nowcasting*, K. Browning, Ed. Orlando, FL: Academic, 1983, pp. 167–176.
- [19] A. Mitiche and P. Boutheymy, "Computation and analysis of image motion: A synopsis of current problems and methods," *Int. J. Comput. Vis.*, vol. 19, no. 1, pp. 29–55, 1996.
- [20] Y. Nakagawa and A. Rosenfeld, "Some experiments on variable thresholding," *Pattern Recognit.*, vol. 11, pp. 191–204, 1979.
- [21] J.-M. Odobez and P. Boutheymy, "Separation of moving regions from background in an image sequence acquired with a mobile camera," in *Video Data Compression for Multimedia Computing*, H. H. Li, S. Sun, and H. Derin, Eds. New York: Kluwer, 1997.
- [22] G. S. Pankiewicz, "Pattern recognition techniques for the identification of cloud and cloud systems," *Meteorol. Applicat.*, vol. 2, pp. 257–271, 1995.
- [23] G. S. Pankiewicz, "Toward synoptic scale cloud classification: Combining spectral and textural features," in *Proc. Meteorological Satellite Data Users' Conf.* Vienna, Austria, Sept. 1996, pp. 111–118.
- [24] C. Papin, "Analyse spatio-temporelle d'images satellitaires météorologiques: Détection et suivi de structures nuageuses critiques," Ph.D. dissertation, Univ. Rennes, Rennes, France, Dec. 1999.
- [25] C. Papin, P. Boutheymy, and G. Rochard, "Detection of low clouds in Meteosat IR nighttime images based on a contextual spatio-temporal labeling approach," in *Proc. 9th AMS-EUMETSAT Conf. Satellite Meteorology and Oceanography*, vol. 2, Paris, France, May 1998, pp. 370–373.
- [26] R. W. Saunders and K. T. Kriebel, "An improved method for detecting clear sky and cloudy radiances from AVHRR data," *Int. J. Remote Sensing*, vol. 9, pp. 123–150, 1988.

Christophe Papin was born in France in 1972. He received the Ph.D. degree in signal processing and telecommunications from the University of Rennes, Rennes, France, in December 1999.

From November 1996 to April 2000, he was with INRIA at IRISA, Rennes. Since May 2000, he has been with the Electronic Imaging Research and Development Group, Kodak, Paris. His research interests include image sequence understanding, motion analysis, deformable object tracking, image segmentation, and motion picture restoration.

Patrick Boutheymy graduated from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1980, and received the Ph.D. degree in computer science from the University of Rennes, Rennes, France, in 1982.

From December 1982 until February 1984, he was with the Department of Visual Communications, INRS-Telecommunications, Montreal, PQ, Canada. Since April 1984, he has been with INRIA at IRISA, Rennes. He is currently "Directeur de Recherche" with INRIA, and Head of Vista Group. His major research interests are concerned with image sequence analysis: 2-D motion analysis and tracking (rigid motion and deformable motion), statistical models for image sequence processing (MRF models, robust estimation, Bayesian estimation), dynamic scene interpretation, and video indexing.

Dr. Boutheymy is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING.

Guy Rochard, photograph and biography not available at the time of publication.