

TEMPORAL REPETITION DETECTION FOR GROUND VISIBILITY ASSESSMENT

R. Grompone von Gioi^{1,*}, C. Hessel¹, T. Dagobert¹, J. M. Morel¹, C. de Franchis^{1,2}

¹ Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, 94235, Cachan, France

² Kayrros SAS

KEY WORDS: ground visibility detection, cloud detection, satellite time series, a contrario framework

ABSTRACT:

Assessing ground visibility is a crucial step in automatic satellite image analysis. Nevertheless, several recent Earth observation satellite constellations lack specially designed spectral bands and use a frame camera, precluding spectrum-based and parallax-based cloud detection methods. An alternative approach is to detect the parts of each image where the ground is visible. This can be done by comparing locally pairs of registered images in a temporal series: matching regions are necessarily cloud free. Indeed, the ground has persistent patterns that can be observed repetitively in the time series while the appearance of clouds changes at each date. To detect reliably the “visible” ground, we propose here an *a contrario* local image matching method coupled with an efficient greedy algorithm.

1. INTRODUCTION

Assessing ground visibility is an important step in optical satellite images analysis. Indeed, the presence of clouds and haze concealing the surface of the Earth often causes detection errors in automatic image analysis. This task is usually addressed as a cloud detection problem, where the image pixels are classified into classes such as ground, clouds, cirrus, snow, haze, etc. (Chandran, Jojy, 2015, Mahajan, Fataniya, 2019).

Satellite cloud detection often exploits spectral bands specifically designed for cloud detection (Irish, 2000, Zhang et al., 2001, Irish et al., 2006, Scaramuzza et al., 2012, Chandran, Jojy, 2015, Taravat et al., 2015, Hollstein et al., 2016). Alternatively, the inter-channel delay in push-broom satellites allows cloud detection by parallax analysis of the color bands (Shin, Pollard, 1996, Panem et al., 2005, Manizade et al., 2006, Wu et al., 2016, Sinclair et al., 2017, Frantz et al., 2018). The past decade has seen the launch of constellations with numerous satellites, considerably reducing the revisit time. Very often the revisit is repeatedly made from the same attitude. This opens the way to reliable change detection and therefore to automatic Earth monitoring, provided that atmospheric changes have been discarded (Baetens et al., 2019). Unfortunately, for reasons of size or cost, recent Earth observation satellite constellations like PlanetScope lack the specially designed spectral bands and use a frame camera, precluding the classic cloud detection approaches. Fortunately the short revisit time of these satellites can, to some extent, compensate the lack of physical or geometric cues about cloud cover.

Instead of detecting clouds, an alternative approach is to detect ground visibility, that is, the parts of each image where the ground is visible (Dagobert et al., 2019). This can be done by comparing the corresponding parts of a time series and selecting matching regions. The rationale is that the ground has persistent patterns that are observed repetitively in the time series, while clouds are changing phenomena appearing differently each time. This approach is based on the following three hypotheses:

1. the images in the time series are well registered;
2. the region of interest changes slowly relative to the time series;
3. each part of the region of interest is visible at least twice in the time series.

Under these hypotheses, any pattern observed more than once at the same position must correspond to a visible region of the ground. Of course, the method cannot be applied when these conditions are not satisfied. For example, the sea is continuously changing and does not satisfy the second hypothesis. It will therefore always be marked as not visible and should be detected separately, for example by a NDWI index.

The proposed method works as follows. Given a set of N registered images, the ground visibility masks for the N images are all initialized as *not visible*; then the $\frac{N(N-1)}{2}$ pairs of images are compared and when a local match is found, the corresponding parts are marked as *visible* in both ground visibility masks.

There is a large literature on how to perform local image comparison (Lowe, 2004, Ke, Sukthankar, 2004, Bay et al., 2006, Arandjelovic, Zisserman, 2012). For our problem, three properties are desirable: robustness to contrast changes (as satellite image dynamics may change), a good control of false detections, and a quick process. The last two requirements are imposed by the considerable size of data. There are many fast local image comparison methods, and most achieve contrast invariance by relying on the image gradient orientation, like the celebrated SIFT algorithm (Lowe, 2004, Rey Otero, Delbracio, 2014). An effective ground visibility detection algorithm based on SIFT image descriptors was presented in (Dagobert et al., 2019).

Concerning the control of false detections, the *a contrario* statistical approach (Desolneux et al., 2000, Desolneux et al., 2008) was introduced to provide a well founded mechanism for setting detection thresholds while producing few false detections. This framework has been successfully applied to local image comparison (Cao et al., 2008, Rabin et al., 2009, Pătrăucean et al., 2013, Grompone von Gioi, Pătrăucean, 2015,

* Corresponding author (grompone@ens-paris-saclay.fr)

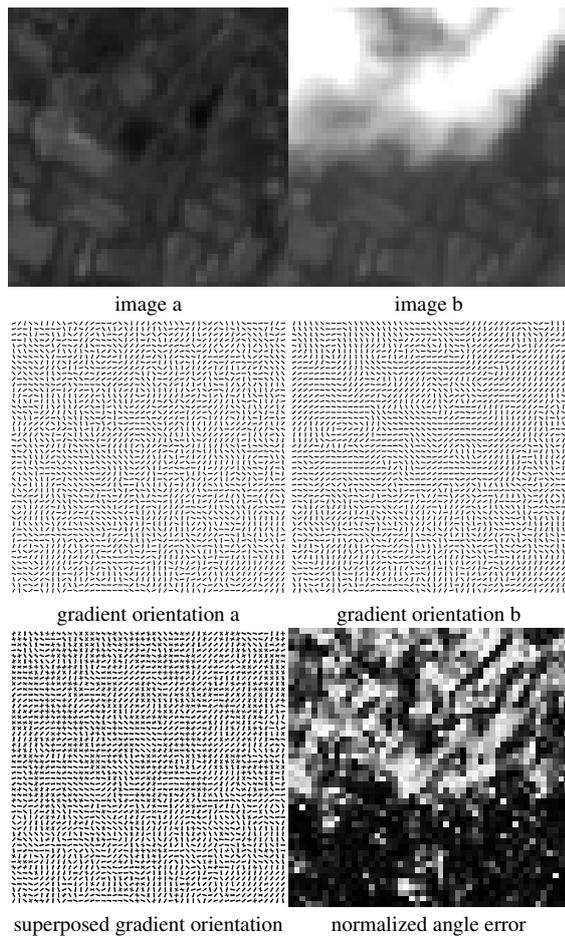


Figure 1. First row: Two images, a and b, to be compared (from Sentinel-2 constellation). Second row: The image gradient orientation field of both images. Third row left: the image gradient orientation fields of both images superposed; one can see how these orientations roughly agree in the visible zone while mainly differing where a cloud is present. Third row right: normalized gradient orientation error between the two images; black corresponds to zero error (same gradient orientation in both images) while white corresponds to the worst error (the gradient orientations have opposite direction).

Rodriguez, Grompone von Gioi, 2018), again based on the image gradient orientation for contrast change robustness. These methods work by comparing the gradient orientation in corresponding image patches, see Figure 1. While effective, there are two limitations when applied to our current problem. First, the patch shape and size need to be specified, and the optimal selection may be content dependent. Second, an exhaustive local patch comparison may be time-consuming.

Here we propose an *a contrario* local image matching method that can be applied to connected regions of arbitrary shapes. Second, we propose a greedy algorithm building quickly candidate matching regions. This procedure, coupled with a final morphological area filter, results in an effective ground visibility detection algorithm for satellite time series.

This paper is organized as follows. Section 2 introduces briefly the *a contrario* approach which is then applied in Section 3 for local image matching. Then Section 4 details the ground visibility detection algorithm. The method is illustrated with some experiments in Section 5. Section 6 present the conclusions.

2. THE *A CONTRARIO* APPROACH

The *a contrario* theory (Desolneux et al., 2000, Desolneux et al., 2008) is a statistical framework used to set detection thresholds automatically in order to control the number of false detections. It is based on the non-accidentalness principle (Witkin, Tenenbaum, 1983, Lowe, 1985) which informally states that an observed structure is meaningful only when the relation between its parts is too regular to be the result of an accidental arrangements of independent parts. In the words of D. Lowe, “we need to determine the probability that each relation in the image could have arisen by accident, $P(a)$. Naturally, the smaller that this value is, the more likely the relation is to have a causal interpretation” (Lowe, 1985, p. 39).

A stochastic background model \mathcal{H}_0 needs to be defined, where the structure of interest is not present and can only arise as an accidental arrangement. For example, many image matching methods (including the one proposed here) are based on the orientation of the image gradient (Lowe, 2004, Cao et al., 2008, Rabin et al., 2009, Pătrăucean et al., 2013, Grompone von Gioi, Pătrăucean, 2015, Rodriguez, Grompone von Gioi, 2018). Some geometrical feature detection methods such as line segments or elliptical arcs (Grompone von Gioi, 2014, Pătrăucean et al., 2017) are also based on the image gradient. In such cases, the background model \mathcal{H}_0 assumes that the gradient orientations at each pixel are independent random variables, uniformly distributed in $[-\pi, \pi)$. Under this *a contrario* model, a region of the image where the gradient orientation follows a regular structure would be a rare accident and is detected as such.

We also need to define a family of events of interest T . For feature detection the family of events is the set of all the geometrical events considered, i.e., all the line segments, elliptical arcs, etc., considered in the image domain. Then, we need to assess the accidentalness of a candidate feature. For example, if a line segment is present in an image, the gradient orientation at the corresponding position would be orthogonal to the line segment. Then, given a candidate line segment, one measures how well the image gradient corresponds to the candidate event, and evaluates the probability of observing such a good agreement by chance in the *a contrario* model. A rough agreement could arise just by chance and thus does not correspond to an interesting event; conversely, a very good agreement would be rare and suggest the presence of a structure instead of just a lucky accident. In other words, when this probability is small enough, there exists evidence to reject the null hypothesis and declare the event meaningful. However, one needs to consider that multiple candidates are tested. If 100 tests were performed, for example, it would not be surprising to observe among them one event that appears with probability 0.01 under random conditions. The number of tests N_T needs to be included as a correction term, as it is done in the statistical multiple hypothesis testing framework (Gordon et al., 2007).

Following the *a contrario* methodology (Desolneux et al., 2000, Desolneux et al., 2008), we define the *Number of False Alarms* (NFA) of an event e observed up to an error $k(e)$ as:

$$\text{NFA}(e) = N_T \cdot \mathbb{P} \left[K_{\mathcal{H}_0}(e) \leq k(e) \right], \quad (1)$$

where the right hand term is the probability of obtaining in the background model \mathcal{H}_0 an error $K_{\mathcal{H}_0}(e)$ smaller or equal to the observed one $k(e)$. The smaller the NFA, the more unlikely the event e is to be observed by chance in the background

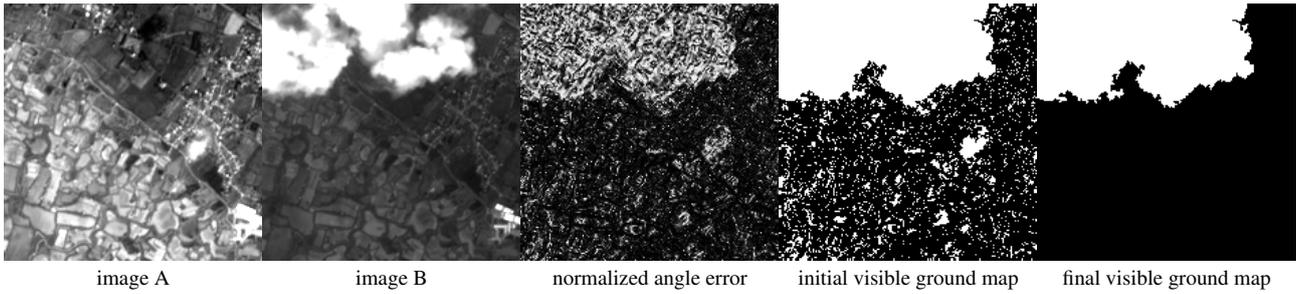


Figure 2. Two images to be compared and the intermediate and final result of the image matching method. In the normalized gradient orientation error map, black corresponds to zero error (same gradient orientation in both images) while white corresponds to the worst error (the gradient orientations are opposite direction). In the two ground visibility maps, black pixels correspond to **visible** ground while white corresponds to **not visible** ground. In the initial ground visibility mask one can see some “ground holes”, that is small white spots surrounded by black pixels. Those ground holes are filled for the final ground visibility mask. In this simple example with only two images, a common mask can be produced, and the zones that do not match are declared as not visible. However, when more images are included in the series, zones that are not at first detected as visible in the first image might match with other images in the series and therefore be declared as visible. This shows the relevance of the third hypothesis of the method: each zone should be visible at least twice in the time series.

model \mathcal{H}_0 ; thus, the more meaningful. The *a contrario* approach prescribes accepting as valid detections the candidates with $NFA < \varepsilon$ for a predefined value ε . It can be shown (Desolneux et al., 2000, Desolneux et al., 2008) that under \mathcal{H}_0 , the expected number of tests with $NFA < \varepsilon$ is bounded by ε . As a result, ε corresponds to the mean number of false detections under \mathcal{H}_0 . In many practical applications, including the present one, the value $\varepsilon = 1$ is adopted. Indeed, it allows for less than one false detection on an image or on a set of images, which is usually quite tolerable.

3. A CONTRARIO LOCAL IMAGE MATCHING

Given two images u and v of the same size $X \times Y$, and a set of pixels R , we would like to know whether both images are similar in the region R . To this aim, we will use the distance

$$d_{u,v}(R) = \sum_{i \in R} \frac{|\text{Angle}(\nabla u(i), \nabla v(i))|}{\pi}, \quad (2)$$

namely the sum of all normalized gradient angle errors in R (Pătrăucean et al., 2013, Grompone von Gioi, Pătrăucean, 2015). This distance can take values between zero and $|R|$ (the size of the region). Zero corresponds to a perfect match, while $|R|$ corresponds to the case where the gradient orientation in the two images are opposite at every pixel of R . Using only the image gradient orientation renders this distance invariant to illumination changes.

For a given region R , we need to decide whether the distance $d_{u,v}(R)$ is small enough, indicating whether the corresponding parts of the images are similar or not. We propose to use for this an *a contrario* formulation. As explained before, a natural background model \mathcal{H}_0 is that the gradient orientations at each pixel are independent random variables, uniformly distributed in $[-\pi, \pi)$. (This will happen, for example, if one of the images contains a cloud covering the region.) Following the *a contrario* framework, we will define the NFA associated to a candidate region match as:

$$NFA(u, v, R) = N_T \cdot \mathbb{P}[D_{\mathcal{H}_0}(R) \leq d_{u,v}(R)], \quad (3)$$

where $D_{\mathcal{H}_0}(R)$ is a random variable corresponding to the dis-

tance $d_{U,V}(R)$ for random images U and V whose gradient orientation follow \mathcal{H}_0 . But under \mathcal{H}_0 the gradient orientations are uniformly distributed in all directions, which implies that the normalized angle error at each pixel are independent random variables following a uniform distribution in $[0, 1]$. As a result, $D_{\mathcal{H}_0}(R)$ corresponds to the sum of $|R|$ independent and uniformly distributed random variables taking values in $[0, 1]$. Thus, $D_{\mathcal{H}_0}(R)$ follows the Irwin-Hall distribution (Johnson et al., 1995) and for a given d , with $0 \leq d \leq |R|$, we obtain:

$$\mathbb{P}[D_{\mathcal{H}_0}(R) \leq d] = \frac{1}{|R|!} \sum_{i=0}^{\lfloor d \rfloor} (-1)^i \binom{|R|}{i} (d-i)^{|R|}, \quad (4)$$

where $\lfloor d \rfloor$ is the integer part of d and $\binom{n}{i}$ is the binomial coefficient.

To complete the formulation we still need to specify the family of tests T and the corresponding number of tests N_T . Instead of using rectangular patches as in (Pătrăucean et al., 2013, Grompone von Gioi, Pătrăucean, 2015, Rodriguez, Grompone von Gioi, 2018), here we are interested in connected regions of any shape. We consider *4-connectivity*, in which a pixel (x, y) is connected to the pixels at coordinates $(x \pm 1, y)$ and $(x, y \pm 1)$. Groups of pixels connected under 4-connectivity correspond to the figures called *polyominoes* (Golomb, 1994, Gardner, 1960). The exact number b_n of different polyomino configurations of given size n is not known in general, but there are good approximations of this number (Jensen, Guttman, 2000). In our case, it is enough to have an estimation of the order of magnitude, so the approximate formula given in (Jensen, Guttman, 2000) is sufficient for our needs. It reads

$$b_n \approx B \frac{\tau^n}{n}, \quad (5)$$

where $B \approx 0.316915$ and $\tau \approx 4.062570$. We need to consider that each particular polyomino may be placed at any position in the image, thus we need to multiply b_n by XY to consider all possible placements (this estimation is not exact as it counts some polyominoes extending outside the image domain; nevertheless, again, it gives the right order of magnitude). Also, we consider connected regions of different sizes, from 1 to XY , the latter being the case where all the pixels are part of one polyomino. This calculation is not exact due to the restrictions

Algorithm 1: Ground visibility detection algorithm

input : N images u_1, \dots, u_N defined on $\Omega (X \times Y)$
parameter: region growing angular tolerance $\rho = \frac{1}{5}$
parameter: grain filter region size λ
output : N ground visibility masks V_1, \dots, V_N

- 1 $B \leftarrow 0.316915$ polyominoes count constants B and τ
- 2 $\tau \leftarrow 4.062570$
- 3 $V_k(\Omega) \leftarrow$ **not visible** initialize all masks as not visible
- 4 $g_k \leftarrow$ GradientAngle(u_k) compute gradient orientation
- 5 **for** $(a, b) \in$ Pairs(N) **do** loop on every pair of images
- 6 $\gamma \leftarrow$ NormalizedAngleError(g_a, g_b)
- 7 **for** $i \in \Omega$ **do** loop on seed pixels
- 8 $R \leftarrow$ region(i, γ, ρ) neighbors of i with $\gamma(j) < \rho$
- 9 $d \leftarrow \sum_{j \in R} \gamma(j)$ normalized angle error sum
- 10 $NFA \leftarrow \frac{N(N-1)}{2} X^2 Y^2 B \frac{\tau^{|R|}}{|R|} \frac{d^{|R|}}{|R|}$ NFA approx.
- 11 **if** $NFA < 1$ **then** meaningful match found
- 12 $V_a(R) \leftarrow$ **visible** mark visible region in a
- 13 $V_b(R) \leftarrow$ **visible** mark visible region in b
- 14 $\gamma(R) \leftarrow 1$ pixels in R will never be visited again
- 15 $V_k \leftarrow$ GrainFilter(V_k, λ) fill visibility holes of size $< \lambda$

imposed by the total size of the image; e.g., among the b_{XY} polyominoes of size XY , only one is actually possible, the one using all the pixels in the images; again, this rough estimation is useful for our case. The number of tolerated false detections ε is divided into XY categories, and to each one we allow to produce $\frac{\varepsilon}{XY}$ false detections. Thus, for a given region size, the test becomes

$$XY b_{|R|} \cdot \mathbb{P} \left[D_{\mathcal{H}_0}(R) \leq d \right] \leq \frac{\varepsilon}{XY}. \quad (6)$$

This is equivalent to setting the number of test to $X^2 Y^2 b_{|R|}$. Finally, in our setting, if there are N images in the time series, we will try to find matches between all the possible pairs. Thus, an extra factor $\frac{N(N-1)}{2}$ needs to be included. The final number of tests is then

$$N_T = \frac{N(N-1)}{2} X^2 Y^2 b_{|R|}. \quad (7)$$

All in all, we define the NFA of a candidate match by

$$NFA(u, v, R) = \frac{N(N-1)}{2} X^2 Y^2 b_{|R|} \mathbb{P} \left[D_{\mathcal{H}_0}(R) \leq d_{u,v}(R) \right], \quad (8)$$

where $d_{u,v}(R)$ is the distance defined in Equation 2. We set $\varepsilon = 1$ and all candidates with $NFA(u, v, R) < 1$ are considered detected local matches.

4. GROUND VISIBILITY DETECTION

The former section described the *a contrario* model to decide whether a couple of images are similar in a given region or not. But trying every possible connected region in the image domain is obviously impossible. Instead, we propose a greedy algorithm for performing the comparison of all pairs of images in a reduced time.¹ Algorithm 1 provides a full overview and the steps are discussed in what follows. Figure 2 illustrates the procedure on two images.

The input is a time series composed of N registered images

¹ Code at <https://github.com/rafael-grompone-von-gioi/visibility>

u_1, \dots, u_N defined on the same domain Ω of the size $X \times Y$. The algorithm could be easily adapted to work on images defined on different domains, provided that the common parts are well-registered, but using a common domain makes the formulation much simpler.

The algorithm depends on two parameters, ρ and λ , to be detailed below. The output is a set of N ground visibility masks V_1, \dots, V_N , defined on the same domain Ω ; at a given pixel i , the mask value $V_k(i)$ can take two values, **visible** or **not visible**, indicating the ground visibility of pixel i of the image u_k .

The first two steps set the two constants B and τ described in Section 3 and used to count polyominoes. Then, step 3 initializes all the pixels of all the ground visibility masks to **not visible**. Step 4 computes the gradient orientation $g_k(i)$ at each pixel i of each image u_k . In our implementations the gradient is computed using central differences.

Now, all the $\frac{N(N-1)}{2}$ pairs of images u_a and u_b are evaluated (step 5). The normalized gradient angle error map γ between u_a and u_b is computed (step 6) and defined as

$$\gamma(i) = \frac{|\text{Angle}(\nabla u_a(i), \nabla u_b(i))|}{\pi} = \frac{|\text{Angle}(g_a(i), g_b(i))|}{\pi}. \quad (9)$$

A greedy algorithm is used to propose candidate regions R , which are then evaluated using the *a contrario* formulation described before. For this, every pixel i in the domain is a possible seed pixel for a new region (step 7). Starting from the pixels i , a region growing algorithm (step 8) iteratively adds neighbor pixels in which the normalized gradient angle error $\gamma(j)$ is smaller than the parameter ρ . A 4-connectivity is used to define neighbors. Thus, the result is a region R in which all the pixels are connected to each other by a 4-connection chain of pixels of R and in which all its pixels have $\gamma(j) < \rho$. The $NFA(u_a, u_b, R)$ is computed in steps 9 and 10. When its value is smaller than one, a meaningful match is found between images u_a and u_b at region R . Thus, both ground visibility masks V_a and V_b are updated to **visible** at the pixels of R (steps 11 to 13).

After the test and to accelerate the procedure, regardless of whether a match is validated or not, the normalized error map γ is set to one at all the pixels of the region R ; as a result, none of these pixels would satisfy the condition of being smaller than ρ and would not start nor be added to a new region in step 8. As a consequence, each pixel is part of no more than one region.

Some remarks are required. First, in step 9, the distance $d_{u_a, u_b}(R)$ is evaluated as defined in Eq. 2. Then, the NFA is computed according to Eq. 8 but using an approximation to the probability term $\mathbb{P} \left[D_{\mathcal{H}_0}(R) \leq d \right]$. It can be shown that the Irwin-Hall distribution in Eq. 4 can be upper-bounded by the first term of the sum,

$$\mathbb{P} \left[D_{\mathcal{H}_0}(R) \leq d \right] \leq \frac{d^{|R|}}{|R|!}. \quad (10)$$

The latter expression is much simpler to compute and due to the upper-bound, the condition $N_T \frac{d^{|R|}}{|R|!} < 1$ implies that the NFA is also smaller than one. Detected regions are guaranteed to be meaningful. However, some regions could be incorrectly rejected when its exact value is near one. Nevertheless, numerical experiments confirm that this effect is minor. Indeed, when

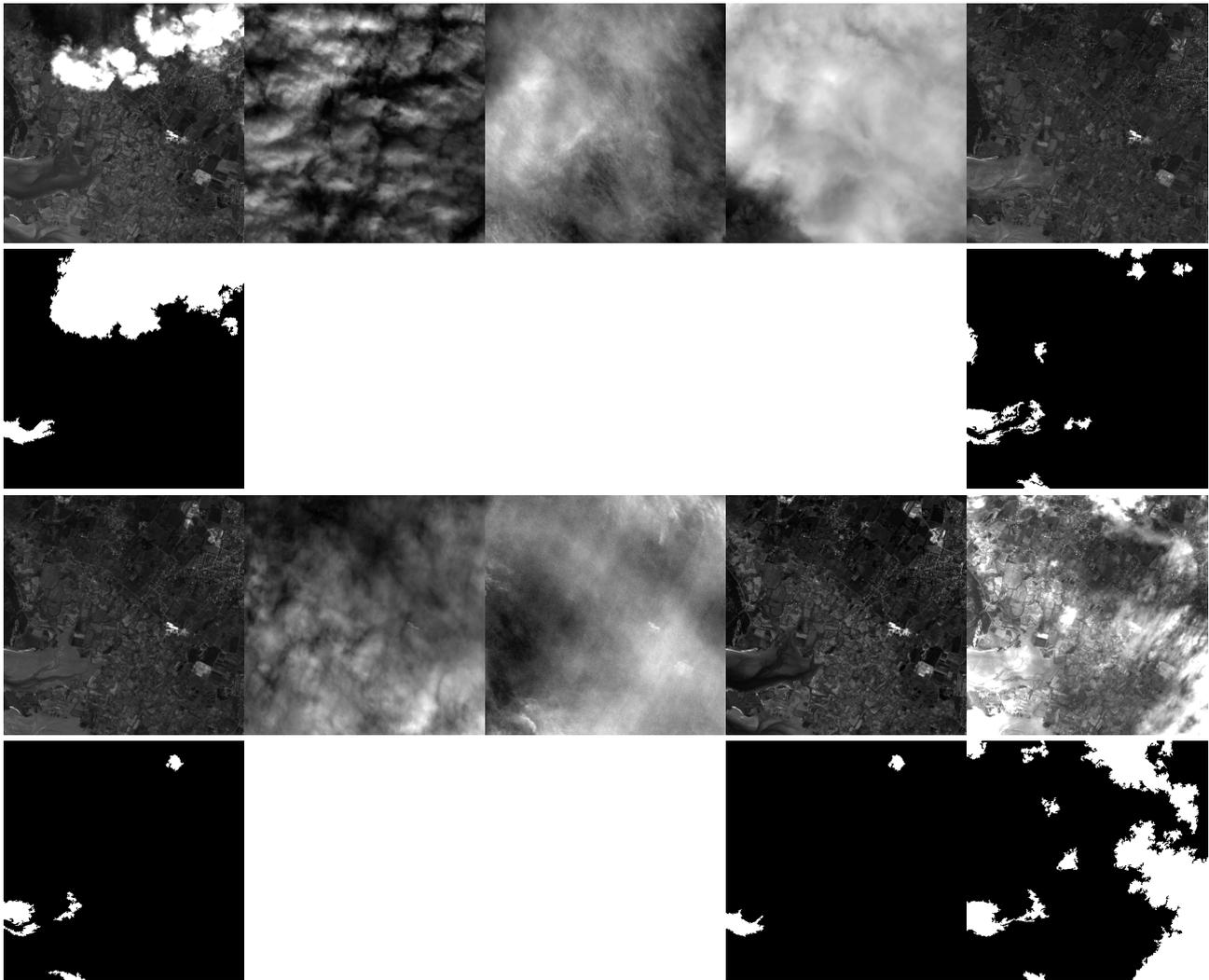


Figure 3. Ground visibility computed with the proposed algorithm in a time series of ten images from the Sentinel-2 constellation. First and third rows: input images. Second and fourth rows: resulting ground visibility maps, where black pixels correspond to **visible** while white corresponds to **not visible**. The images are 496×496 and the ten ground visibility masks were jointly computed in about one second on an 1.2 GHz Intel Core M processor.

similar structures are present in two images, the NFAs are usually very small, and even this upper-bound is also very small.

The second remark concerns the parameter ρ . This value has an impact on the greedy method used to propose the candidate regions R . However, it has no impact on the final validation step based on Eq. 8. The parameter ρ is fixed to a very relaxed value $1/5$, which implies that gradients angles with a difference of up to 36 degree are grouped together. If all the pixels in the region have an error of about 36 degree, this would rarely lead to a meaningful match; in good matches, however, many pixels have a much smaller errors, leading to meaningful validations.

It is important to notice that the partition of the domain Ω into the candidate regions does not depend on the particular order in which the region growing algorithm works; the result is fully determined by the angle error map γ , the use of 4-connectivity, and the tolerance parameter ρ .

As can be seen in Figure 2, the initial ground visibility masks have many “ground holes”, which are seen in the figure as isolated white spots. These holes are filled by a simple grain filter that removes connected **not visible** regions of less than λ pixels.

This parameter is adjusted according to the satellite resolution, the size of the clouds, and the size of the structures of interest.

5. EXPERIMENTS

Figure 3 shows the result of the ground visibility detection on a time series of ten Sentinel-2 images. Here the hole filling parameter λ was set to 500 pixels. The first and third rows show the ten input images and the second and fourth rows show the resulting ground visibility masks. Here white corresponds to **not visible** and black to **visible**. At first glance one can see the correct result for images 2, 3, 4, 7 and 8; the images are fully covered by clouds and so indicate the complete white masks. The other images need more careful observation. In most of the cases these masks are mainly black, correctly indicating the general ground visibility of those images. The white spots correspond mostly to the presence of clouds or to zones where the topography does not satisfy the slow change hypothesis, essentially zones covered by water. But there are also **not visible** parts which are due to the lack of repetitive structure; this is the case in some small fields where no relevant texture can be matched from one image to another. The algorithm is quite

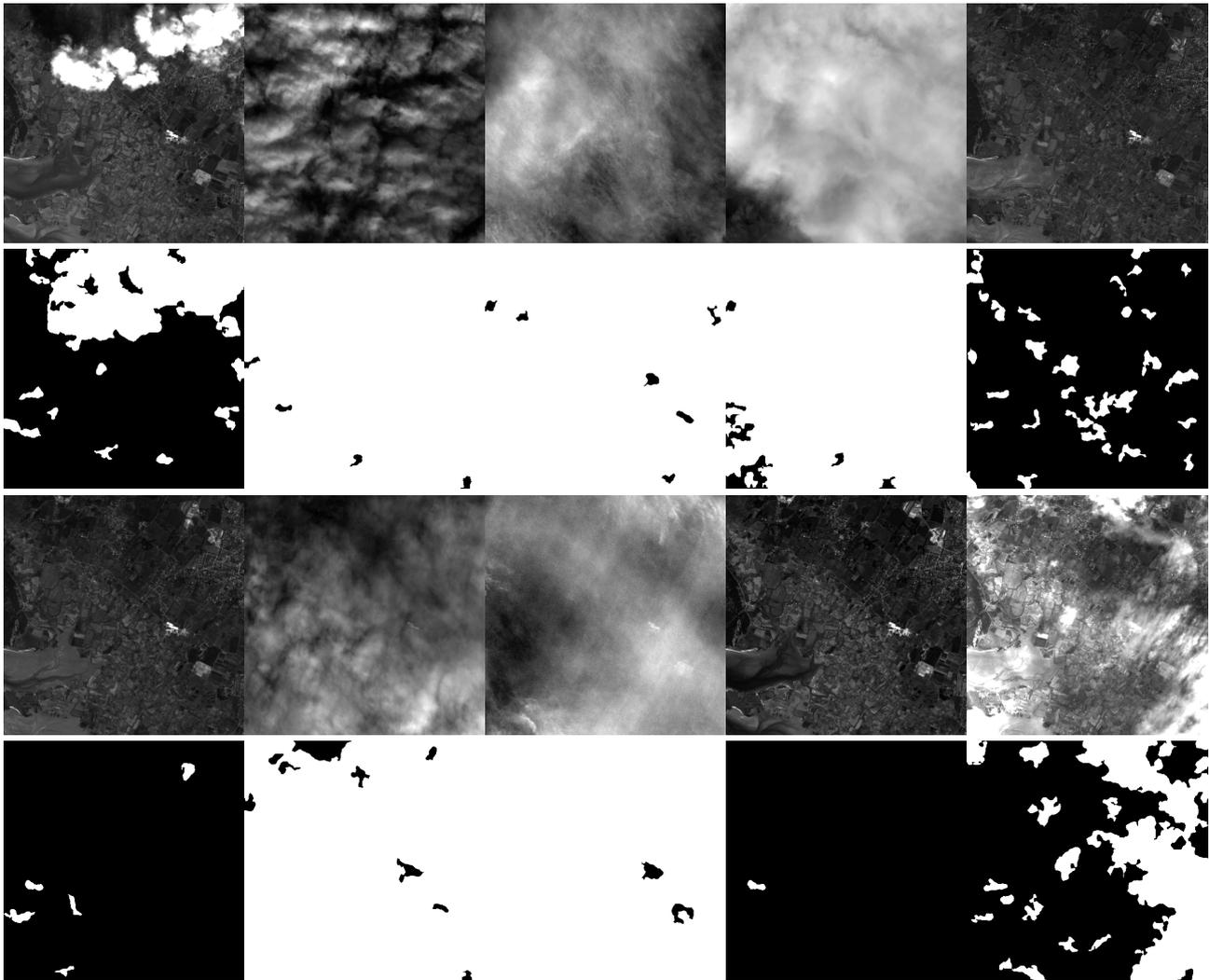


Figure 4. Ground visibility computed by the algorithm of (Dagobert et al., 2019) in a time series of ten images from the Sentinel-2 constellation. First and third rows: input images. Second and fourth rows: resulting ground visibility maps, where black pixels correspond to **visible** while white corresponds to **not visible**. The images are 496×496 and the ten ground visibility masks were jointly computed in about one minute and 20 seconds on an 1.2 GHz Intel Core M processor.

fast as the ten ground visibility masks were jointly computed in about one second on an 1.2 GHz Intel Core M processor.

In this case, the algorithm managed to provide a fairly good evaluation of the general ground visibility. As a rule, the larger the number of images in the time series, the better the result.

For comparison, Figure 4 shows the ground visibility computed by the algorithm proposed in (Dagobert et al., 2019), for the same time series as in the previous experiment. The result is similar but with some spurious ground and cloud zones. Notice also that the computation time was one minute and 20 seconds.

6. CONCLUSIONS

An algorithm for temporal repetition detection in time series was proposed based on the *a contrario* statistical framework and using a greedy procedure to accelerate the result. The method produces good results when its functioning hypotheses are satisfied. Among them, the more demanding one is that the target zone changes slowly relative to the sampling frequency, which is not valid, for example, over the sea. Future work will

concentrate on the hole filling step, aiming at reducing the demand for setting manually the only sensitive parameter of the method.

ACKNOWLEDGEMENTS

We thank Thibaud Ehret for valuable discussions and suggestions. Work partly financed by Office of Naval research grant N00014-17-1-2552, DGA Astrid project “Filmer la Terre” no. ANR-17-ASTR-0013-01 and Kayrros.

REFERENCES

- Arandjelovic, R., Zisserman, A., 2012. Three things everyone should know to improve object retrieval. *CVPR*, 2911–2918.
- Baetens, L., Desjardins, C., Hagolle, O., 2019. Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure. *Remote Sensing*, 11(4).

- Bay, H., Tuytelaars, T., Van Gool, L., 2006. Surf: Speeded up robust features. *ECCV*, 1, 404–417.
- Cao, F., Lisani, J. L., Morel, J. M., Musé, P., Sur, F., 2008. *A Theory of Shape Identification*. Springer.
- Chandran, G., Jojy, C., 2015. A Survey of Cloud Detection Techniques For Satellite Images.
- Dagobert, T., Morel, J. M., de Franchis, C., Grompone von Gioi, R., 2019. Visibility detection in time series of planetscope images. *IEEE International Geoscience and Remote Sensing Symposium*, 1673–1676.
- Desolneux, A., Moisan, L., Morel, J.-M., 2000. Meaningful Alignments. *International Journal of Computer Vision*, 40(1), 7–23.
- Desolneux, A., Moisan, L., Morel, J.-M., 2008. *From gestalt theory to image analysis: a probabilistic approach*. 34, Springer.
- Frantz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 215, 471 - 481.
- Gardner, M., 1960. More about the shapes that can be made with complex dominoes (Mathematical Games). *Scientific American*, 203(5), 186–201.
- Golomb, S. W., 1994. *Polyominoes*. 2. ed. edn, Princeton University Press.
- Gordon, A., Glazko, G., Qiu, X., Yakovlev, A., 2007. Control of the mean number of false discoveries, Bonferroni and stability of multiple testing. *The Annals of Applied Statistics*, 1(1), 179–190.
- Grompone von Gioi, R., 2014. *A contrario line segment detection*. Springer.
- Grompone von Gioi, R., Pătrăucean, V., 2015. A contrario patch matching, with an application to keypoint matches validation. *ICIP*, 946–950.
- Hollstein, A., Segl, K., Guanter, L., Brell, M., Enesco, M., 2016. Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images. *Remote Sensing*, 8(8).
- Irish, R. R., 2000. Landsat 7 automatic cloud cover assessment.
- Irish, R. R., Barker, J. L., Goward, S. N., Arvidson, T., 2006. Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) Algorithm. *Photogrammetric Engineering & Remote Sensing*, 72(10), 1179–1188.
- Jensen, I., Guttmann, A. J., 2000. Statistics of lattice animals (polyominoes) and polygons. *Journal of Physics A: Mathematical and General*, 33(29), L257.
- Johnson, N. L., Kotz, S., Balakrishnan, N., 1995. *Continuous univariate distributions*. Distributions in statistics, 2. ed. edn, Wiley, New York, NY [u.a.].
- Ke, Y., Sukthankar, R., 2004. PCA-SIFT: A more distinctive representation for local image descriptors. *CVPR*, 2, 506–513.
- Lowe, D., 1985. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers.
- Lowe, D. G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2), 91–110.
- Mahajan, S., Fataniya, B., 2019. Cloud detection methodologies: variants and development—a review. *Complex & Intelligent Systems*. <https://doi.org/10.1007/s40747-019-00128-0>.
- Manizade, K. F., Spinhirne, J. D., Lancaster, R. S., 2006. Stereo Cloud Heights From Multispectral IR Imagery via Region-of-Interest Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9), 2481–2491.
- Panem, C., Baillarin, S., Latry, C., Vadon, H., Dejean, P., 2005. Automatic cloud detection on high resolution images. *Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.*, 1, 4 pp.–.
- Pătrăucean, V., Gurdjos, P., Grompone von Gioi, R., 2017. Joint A Contrario Ellipse and Line Detection. *IEEE transactions on pattern analysis and machine intelligence*, 39(4), 788–802.
- Pătrăucean, V., Grompone von Gioi, R., Ovsjanikov, M., 2013. Detection of mirror-symmetric image patches. *CVPRW*, 211–216.
- Rabin, J., Delon, J., Gousseau, Y., 2009. A Statistical Approach to the Matching of Local Features. *SIIMS*, 2(3), 931–958.
- Rey Otero, I., Delbracio, M., 2014. Anatomy of the SIFT Method. *Image Processing On Line*, 4, 370–396.
- Rodriguez, M., Grompone von Gioi, R., 2018. Affine invariant image comparison under repetitive structures. *ICIP*, 1203–1207.
- Scaramuzza, P. L., Bouchard, M. A., Dwyer, J. L., 2012. Development of the Landsat Data Continuity Mission Cloud-Cover Assessment Algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), 1140–1154.
- Shin, D., Pollard, J. K., 1996. Cloud height determination from satellite stereo images. *IEE Colloquium on Image Processing for Remote Sensing*, 4/1–4/7.
- Sinclair, K., van Dienenhoven, B., Cairns, B., Yorks, J., Wasilewski, A., McGill, M., 2017. Remote Sensing of Multiple Cloud Layer Heights using Multi-Angular Measurements. *Atmospheric Measurement Techniques Discussions*, 1–23.
- Taravat, A., Proud, S., Peronaci, S., Del Frate, F., Oppelt, N., 2015. Multilayer Perceptron Neural Networks Model for Meteosat Second Generation SEVIRI Daytime Cloud Masking. *Remote Sensing*, 7(2), 1529–1539.
- Witkin, A. P., Tenenbaum, J. M., 1983. On the role of structure in vision. J. Beck, B. Hope, A. Rosenfeld (eds), *Human and Machine Vision*, Academic Press, 481–543.
- Wu, T., Hu, X., Zhang, Y., Zhang, L., Tao, P., Lu, L., 2016. Automatic cloud detection for high resolution satellite stereo images and its application in terrain extraction. *{ISPRS} Journal of Photogrammetry and Remote Sensing*, 121, 143 - 156.
- Zhang, W. D., He, M. X., Mak, M. W., 2001. Cloud detection using probabilistic neural networks. *IGARSS 2001. Scanning the Present and Resolving the Future. Proceedings. IEEE 2001 International Geoscience and Remote Sensing Symposium (Cat. No.01CH37217)*, 5, 2373–2375 vol.5.