CHARACTERISTIC SCALE IN SATELLITE IMAGES

Bin Luo [1], *Jean-François Aujol* [2], *Yann Gousseau* [1], *Saïd Ladjal* [1], *Henri Maître* [1]

[1] GET/Télécom Paris, CNRS UMR 5141, France
CNES-DLR-ENST Competence Center
[2] CMLA, ENS-Cachan, CNRS UMR 8536, France

ABSTRACT

We study the problem of finding the characteristic scale of a given satellite image. We want to define this feature so that it does not depend on the spatial resolution of the image. Our approach is based on the use of a linear scale space and the total variation. The critical scale is defined as the one at which the normalized total variation is maximum.

1. INTRODUCTION

In recent years, huge databases of remote sensing images have been created (enlarged by gigabytes every day). One property of these databases is that, most of the time, images come from different satellites and therefore have different (usually known) resolutions. It is therefore interesting to use features which do not depend on the resolution to index such images.

In this paper, given an image, we aim at finding a characteristic scale which does not depend on the spatial resolution of the image. This work is inspired by the approach of Lindeberg [1], who has suggested the use of a normalized derivative operator in the linear scale space to get local scale information. An alternative approach also using the linear scale space is presented in [2]. In the literature, there are many works proposing scale invariant features [1, 3, 4]. But as far as we know, they do not take into account the acquisition process. Contrarily to such approaches, we explicitly consider the convolution and the sampling involved in resolution changes. As we will see in this paper, this turns out to be essential with satellite images. To deal with the convolution step, our method is based on the use of a linear scale-space [5]. Moreover, since we seek a characteristic scale related to the geometry of the image, we use the total variation [6].

The plan of the paper is the following. We give a first definition of the characteristic scale in Section 2. We adapt it in Section 3 by taking into account the acquisition process. We present some experiments with synthetic images in Section 4 and then with data provided by the CNES in Section 5. We conclude in Section 6.

2. BASIC TOOLS AND SCALE DEFINITION

In this section, we recall the two main mathematical tools to be used in this work. Let us denote by $f : \mathbb{R}^2 \mapsto \mathbb{R}$ the original image. The linear scale space [5] is defined as $f_t(x, y) = k_t * f$, where :

$$k_t(x, y) = \frac{1}{2\pi t^2} \exp\left(-\frac{x^2 + y^2}{2t^2}\right) \qquad (1)$$

The basic idea of the proposed approach is to quantify the evolution of geometric structures of the image in the linear scale space. Therefore, we consider the total variation of an image, defined (when the image is regular enough) as $TV(f) = \int |\nabla f|$. This semi-norm is related to the geometry of the image through the coarea formula (see [6]), which implies for instance that for a binary image, $TV(f)$ is equal to the perimeter of the objects in the image.

We then define the characteristic scale as the maximizer of the normalised total variation $NTV(f; t)$ of the multi-scale space representation of f :

$$NTV(f; t) = t\, TV\,(k_t * f) = t \int |\nabla k_t * f|. \qquad (2)$$

That is, the characteristic scale is defined as

$$t_{\max} = \operatorname{argmax}_{\mathbb{R}_+^*} NTV(f; t). \qquad (3)$$

Observe the normalization factor t in Equation (2), which is needed to have a maximum of $NTV(f; t)$ as a function of t. This is in fact equivalent to the normalization used by Lindeberg [1] with differential operators. Observe also that if f is zoomed by a factor λ then t_{max} is multiplied by λ. However this property is not sufficient for our purpose (because of blurring and sampling). The next section shows how to adapt this normalization when dealing with resolution changes and discrete images.

3. RESOLUTION INVARIANCE

In order to ensure that the computed characteristic scale does not depend upon the resolution, we first have to model

the acquisition process. Recall that f is a continuous function corresponding to a given scene. We assume that the acquisition system performs a convolution followed by a sampling. We approximate the convolution kernel by a Gaussian kernel (k_σ). The sampling rate (which is also the resolution of the discrete image) is $r = \alpha\sigma$. The α parameter is a characteristic of the acquisition process (the larger α, the more aliased the image). In the numerical experiments presented in the paper, we use $\alpha = 1$. The following formula summarizes the acquisition process :

$$g_0^r = \Psi_r(k_\sigma * f) \qquad (4)$$

where g_0^r is the sampled version of f at resolution r (Ψ being the sampling operator).

We denote by $\tilde{k}_t$ the discrete version of the gaussian kernel with standard deviation t (t being given in pixels). We therefore have $\tilde{k}_t \approx k_{rt}$ (up to some normalization constant which we can drop). Let us define the discrete scale space as :

$$g_t^r = \tilde{k}_t * g_0^r = \tilde{k}_t * (\Psi_r(k_\sigma * f)) \approx \Psi_r(k_{rt} * (k_\sigma * f)) \quad (5)$$

The last approximation means that (at least for non-aliased images such as $k_\sigma * f$) inverting convolution and subsampling is possible. In addition we can assume (for well-sampled images) that the total variation of the continuous and discrete versions are the same up to a normalization due to the zooming of factor r (this will be confirmed by the numerical experiments in the following sections). This leads to :

$$TV(g_t^r) \approx \frac{1}{r}TV(k_{rt}*k_\sigma*f) = \frac{1}{r}TV\left(k_{\sqrt{r^2t^2+\sigma^2}}*f\right) \quad (6)$$

Now we seek a normalization of the discrete total variation in order to easily relate it to the continuous normalized total variation NTV (defined in Equation (2)). Let us define :

$$G^r(t) = h(t)\,TV(g_t^r) \qquad (7)$$

The question is now how to choose the factor $h(t)$.

$$
\begin{aligned}
G^r(t) &\approx \frac{1}{r}h(t)\,TV\left(k_{\sqrt{r^2t^2+\sigma^2}}*f\right) \\
&= \frac{1}{r}\frac{h(t)}{\sqrt{r^2t^2+\sigma^2}}NTV\left(f;\sqrt{r^2t^2+\sigma^2}\right)
\end{aligned}
$$

Hence :

$$G^r(t) \approx \frac{1}{r^2}\frac{h(t)}{\sqrt{t^2+\frac{1}{\alpha^2}}}NTV\left(f;\sqrt{r^2t^2+\sigma^2}\right) \quad (8)$$

If we choose : $h(t) = \sqrt{t^2+\frac{1}{\alpha^2}}$
then we have :

$$G^r(t) \approx \frac{1}{r^2}NTV\left(f;\sqrt{r^2t^2+\sigma^2}\right) \qquad (9)$$

Recall that $t_{\max} = \mathrm{argmax}_{\mathbb{R}_+^*}(NTV(f;t))$. We denote by $t_{\max_r}$ the maximizer of G^r. We then have :

$$t_{\max} = \sqrt{r^2t_{\max_r}^2 + \sigma^2} \qquad (10)$$

For a discrete image of resolution r we thus compute $t_{\max_r}$ and derive the value of $t_{\max}$ using Equation (10). Notice that it is impossible to find a characteristic scale $t_{\max}$ smaller than σ (which is comparable to r). More generally, when the resolution of the image is larger than the actual characteristic scale $t_{\max}$ the computation becomes unreliable. Experiments show that $t_{\max}$ is retrievable as long as $r < t_{\max}$.

From now on, the values of NTV will be deduced from Equation (9).

Remark about the normalization : In view of equation (2), the intuitive normalization would be :

$$A^r(t) = tTV(g_t^r) \qquad (11)$$

Nevertheless, if we set $t_{\max_r} = \mathrm{argmax}_{\mathbb{R}_+^*}(A^r(t))$ and $t_{\max} = r \times t_{\max_r}$, then we can check numerically that $t_{\max}$ depends much more on the resolution than with the definition of equation (10) (see Figure 6 of Section 5).

Notice also that when $t >> 1$, then the definitions from Equations (7) and (11) are equivalent. Our normalization choice is important when r approaches $t_{\max}$.

4. TESTS ON SYNTHETIC IMAGES

We test the method on spatial periodic images with period D, using different patterns such as squares, disks, Gaussians ...For instance, Figure 1(a) is made of periodic squares with spatial period $D = 40$. On Figure 1(b), we display the graph of the discrete normalized total variation. By computing t_{max} on images with different patterns at different spatial periods, we experimentally find that : $t_{max} \sim 0.15D$, and that neither the shape nor the size of the pattern seem to influence much the value of t_{max}.
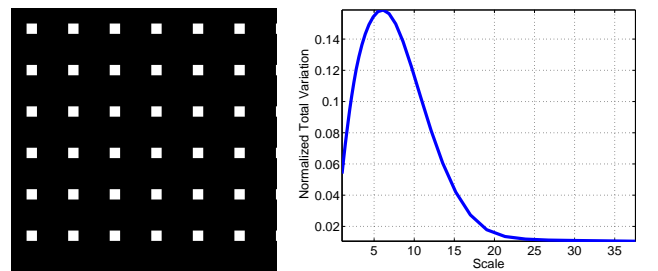
(a) Periodic function of squares   (b) Graph of NTV

Fig. 1. An image composed of squares with spatial period $D = 40$ (the length of each square being 10 pixels) and the graph of the normalized total variation. The maximum is reached for $t = 6.1$.

5. APPLICATION ON REAL IMAGES PROVIDED BY THE CNES

The CNES (Centre National d'Etudes Spatiales) has provided us with images of several scenes (such as cities, forests and fields, see Figure 2). For each scene, we have 33 images at different resolutions ranging from 25 cm to 10.08 m (with a geometrical progression). In particular, we thus have the resolution of the following satellites : Orfeo (70.7 cm), Spot 5 THR (252 cm), SPOT 5 (504 cm), SPOT 2 and 3 (1007.9 cm). These images have been obtained by numerical simulations performed by the CNES.
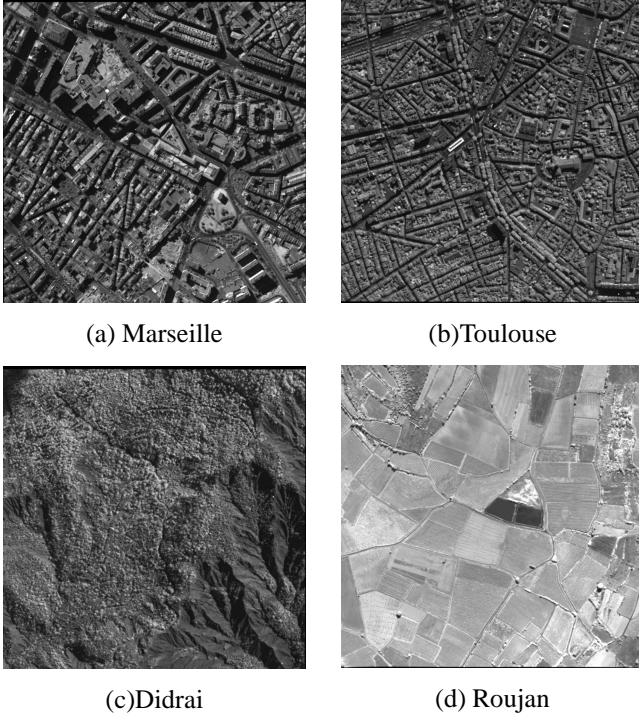

(a) Marseille   (b)Toulouse

(c)Didrai   (d) Roujan

Fig. 2. Aerial images provided by the CNES

On Figure 3, we display the graphs of the normalized total variation for the images shown in Figure 2 (at resolution $r = 25cm$). There is at least one local maximum in each case. In the case of cities (Marseille or Toulouse), the characteristic scale is related to the size of the buildings and of the streets. In the case of the Didrai image (forest), the scale is related to the vegetation. Notice that in the case of the Roujan image (field), there are two local maxima, the small one (which is zoomed in Figure 3(e)) being related to the furrows inside the fields, and the large one being related to the fields themselves.

To check that $t_{\max}$ does not depend on the resolution, we have made the following experiments. For a given scene, we generate g^r, the image at resolution r, and we compute $t_{\max}$ (we take $\alpha = 1$). Figure 4 shows the graph of $t_{\max}$ as a function of r. We can observe that $t_{\max}$ is almost constant (as long as $r < t_{\max}$). Remark that in the case of Roujan, there are two

(a) Marseille   (b) Toulouse
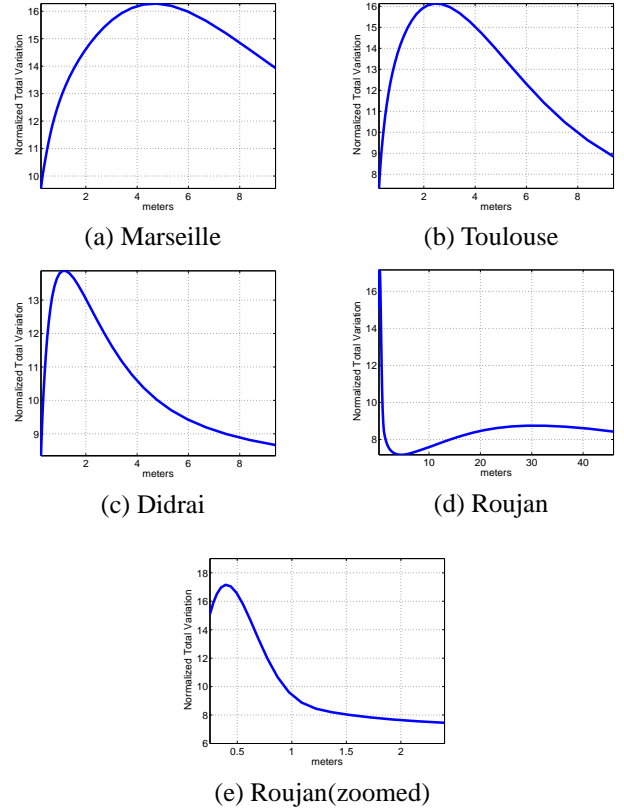
(c) Didrai   (d) Roujan

(e) Roujan(zoomed)

Fig. 3. (a) NTV computed on the image of Marseille ($r = 0.25m$); $t_{max} = 4.8m$; (b) NTV computed on the image of Toulouse ($r = 0.25m$); $t_{max} = 2.4m$; (c) NTV computed on the image of Didrai ($r = 0.25m$); $t_{max} = 1.2m$; (d) NTV computed on the image of Roujan ($r = 0.25m$); there are two local maxima. The first one appears at the resolution $0.4m$, the second one appears at the resolution $30m$; (e) Zoom around the first local maximum shown in (d).

different characteristic scales : this is coherent with the result shown in Figure 3(d). When the resolution r is fine enough, $t_{\max}$ is the characteristic scale corresponding to the furrows. But when r gets larger, then the furrows disappear (one no longer sees them in the images), and $t_{\max}$ is then related to the size of the fields.

We have also performed similar experiments on the images provided by the CNES. Figure 5 shows the graph of $t_{\max}$ as a function of the resolution. We get results that are very similar to the ones of Figure 4. We observe that $t_{\max}$ is almost constant (as long as the resolution $r < t_{\max}$). Notice that even though the kernel used by the CNES for its simulations is not Gaussian, the approximations made in Section 3 appear to be still valid.

On Figure 6, we show the results of the same experiment on the image of Toulouse with the intuitive normalization of Equation (11) which does not take into account convolution in the resolution change. As expected, $t_{\max}$ is very sensitive to the resolution.
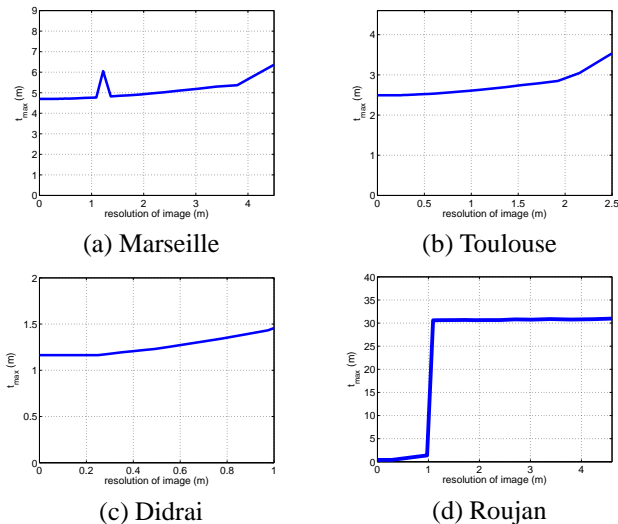
(a) Marseille (b) Toulouse

(c) Didrai (d) Roujan

Fig. 4. Characteristic scales t_{max} as a function of the resolution, on the four scenes shown in Figure 2. The images at different resolutions are obtained using the acquisition model presented in Section 3 (with $\alpha = 1$), that is with a Gaussian convolution kernel. Notice that the characteristic scale is almost independent from the resolution.

(a) Marseille (b) Toulouse

(c) Didrai (d) Roujan

Fig. 5. Characteristic scales t_{max} as a function of the resolution, on the four scenes shown in Figure 2. The images at different resolutions are obtained by the simulations of the CNES. The convolution kernel is no longer Gaussian. Notice that the characteristic scales are almost invariant when the resolutions change.

6. CONCLUSION AND FUTURE PROSPECTS

We have proposed a new method to compute a characteristic scale for a given image. Moreover, we have defined it so that it does not depend on the resolution (as long as the objects are bigger than one pixel).

In future works, we intend to use this scale as a feature for the problem of satellite image indexing. It is indeed a major asset that such a feature does not depend on the resolution. We are currently investigating the possibility to extend the approach to other characteristic scale definitions and scale invariant features. Moreover, we expect to find the texture/geometry behaviour of a scene [7], which indeed depends on the resolution and can be related to the characteristic scale. This could be useful for features selection. We also need to understand more deeply the effect of sampling on total variation, especially when the resolution is close to t_{max}.

Fig. 6. Characteristic scales $t_{\max}$ as a function of the resolution, on the scene of Toulouse. The scale $t_{\max}$ is computed with the naive normalization of Equation (11). Notice that it is far from being constant compared with Figure 5(b). In the range $[0.25, 2m]$, the variation of the value is 18% with our method and 40% with the naive normalization

7. REFERENCES

[1] T. Lindeberg, "Feature detection with automatic scale selection," *IJCV*, vol. 30, no. 2, August 1998.

[2] Martin Jägersand, "Saliency maps and attention selection in scale and spatial coordinates : An information theoretic approach," in *ICCV*, 1995, pp. 195–202.

[3] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *IJCV*, vol. 60, no. 1, pp. 63–86, 2004.
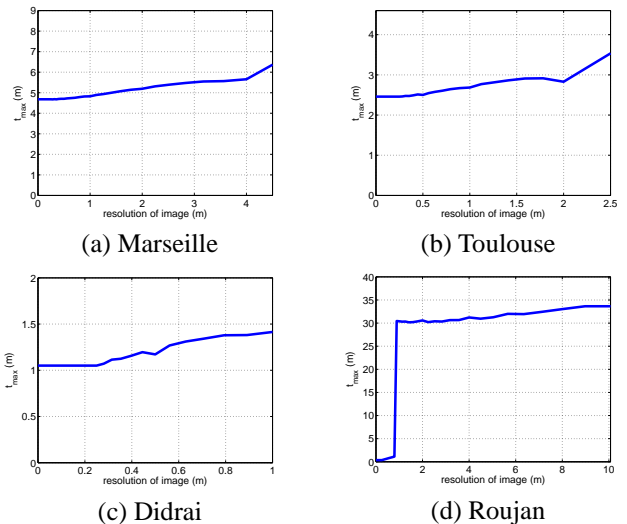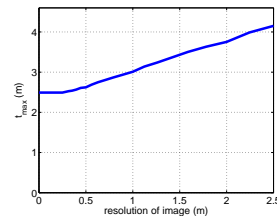
[4] C. M. Pun and M. C. Lee, "Log-polar wavelet energy signatures for rotation and scale invariant texture classification," *IEEE PAMI*, vol. 25, no. 5, pp. 590–603, 2003.

[5] J.J. Koenderink, "The structure of images," *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.

[6] L.C. Evans and R.F. Gariepy, *Measure Theory and Fine Properties of Functions*, vol. 19 of *Studies in Advanced Mathematics*, CRC Press, 1992.

[7] Y. Meyer, *Oscillating Patterns in Image Processing and in Some Nonlinear Evolution Equations*, AMS, March 2001, The Fifteenth Dean Jacquelines B. Lewis Memorial Lectures.