Extrapolation of wavelet features for the indexing of satellite images with different resolutions

Bin Luo [1], Jean-François Aujol [2], Yann Gousseau [1], Saïd Ladjal [1]
[1] GET/Télécom Paris, CNRS UMR 5141, France
CNES-DLR-ENST Competence Center
Email : {bin-luo, gousseau, said.ladjal}@enst.fr
[2] CMLA, ENS-Cachan, CNRS UMR 8536, France
Email : aujol@cmla.ens-cachan.fr

***Abstract*— In this paper, we propose a new scheme to extrapolate wavelet features with respect to the resolution. By explicitly taking into account the acquisition process of satellite images, we compute how wavelet features behave when the resolution changes. This approach is validated by classifying satellite images with different resolutions.**

I. Introduction

Institutions such as the CNES (the French spatial agency), have recently expressed the need to develop automatic indexing schemes to deal with huge databases of satellite images. One particularity of these databases, compared to e.g. natural images databases, is that most of the time images have been acquired by different satellites and therefore have different and usually known resolutions. To index such images, one is therefore naturally led to consider resolution invariant features or to develop schemes to compare features at different resolutions. Although many scale invariant features have been proposed in the literature, see e.g. [1], [2], [3], resolution invariant features have hardly been studied. Indeed, this last problem is more involved since it necessitates to take the image acquisition process into account. In [4], this process is modeled as a convolution followed by a sampling and its effect on the computation of a characteristic scale is studied. In this paper, we make use of the same model and propose a method to relate wavelet features obtained at different resolutions.

Many features have been proposed to index satellite images. In this work, only mono-spectral images are considered and therefore texture features are chosen to classify them. Wavelet features are chosen since they have been proved suitable for texture indexation or classification [5], [6], [7].

The plan of this paper is the following. The specific features extracted from wavelet decompositions are presented in Section III. In Section IV, a method is given to extrapolate these features from a given resolution to another one.We validate these results with some numerical experiments in Section V. We then conclude in Section VI.

II. Model of the acquisition process

We assume that the scene under study is represented by a continuous function f, and that the digital image f_r at resolution r is obtained by convolution and sampling. Moreover, the convolution kernel is always assumed to be Gaussian, with a standard deviation proportional to the resolution. This can conveniently be modeled as

$$f_r = \Pi_r . f * k_{rp}, \tag{1}$$

where

$$k_\sigma(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{2}$$

Π_r is the Dirac comb on $r\mathbb{Z}^2$, that is,

$$\Pi_r = \sum_{i,j\in\mathbb{Z}} \delta_{(ir,jr)},$$

and the parameter p is a characteristic of the acquisition process (the smaller p, the more aliased is the image).

III. Wavelet features for texture indexation

Based on numerical experiments, Mallat [8] proposed to model the empirical distributions h of wavelet coefficients of natural textured images by Generalized Gaussian Distributions (*GGD*) :

$$h(u) = K e^{-(|u|/\alpha)^\beta}. \tag{3}$$

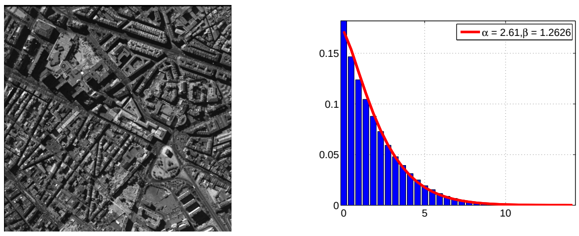
Parameter β is usually called a *shape* parameter, since it modifies the slope of the distribution, and α is a *scale* parameter, directly related to the variance of the distribution. It is shown in [5], [7] that the parameters α and β of GGD can be used as efficient features for texture classification. It is possible to compute these parameters from the estimation of the first and second order moments of $|u|$ [8] : we denote them respectively by $m_1 = \int |u| h(u) du$ and $m_2 = \int u^2 h(u) du$.

In this paper, for simplicity, we address the problem of relating features m_1 and m_2 to resolution changes. Since parameters α and β may be computed only using m_1 and m_2, extrapolating these features with respect to the resolution is straightforward. This can be useful when using the Kulback-Leibler distance in a classification task, see [7].

We denote by $\Theta_{r,t} = \{m_1(r,t), m_2(r,t)\}$ the wavelet features at scale t extracted from f_r.

In order not to be restricted to dyadic resolution changes, continuous wavelet transform ([9]) is used instead of the more classical discrete wavelet transform. Moreover, we consider mother wavelets obtained as derivatives of a Gaussian kernel in horizontal, vertical and diagonal directions. This is motivated by the simplified model for resolution changes presented in the previous section, as will be shown by the computations of Section IV-A.

Figure III shows a histogram of absolute values of wavelet coefficients, illustrating the soundness of the use of GGDs to model such distributions.



(a) Marseille (b) Histogram

Fig. 1. (a) Image of Marseille at resolution $0.707m$ (©CNES); (b) Histogram (blue bars) of (a) at scale 5 (horizontal) and the approximation by GGD (red curve).

IV. Extrapolation of wavelet features through resolutions

A. Resolution Invariance

The discrete version of the Gaussian kernel with standard deviation t (t being given in pixels) is denoted by $\tilde{k}_t$ We therefore have $\tilde{k}_t \approx k_{rt}$. Let us define the discrete wavelet coefficient as (recall that the wavelets we use are derivative of the Gaussian kernel) :

$$w_{q,r,t} = \Delta_q \tilde{k}_t * f_r = \tilde{k}_t * \Delta_q f_r \tag{4}$$

where q is 0 or 1 and Δ_q stands for the difference between adjacent pixels in the horizontal ($q = 0$) or vertical ($q = 1$) direction. Next, we assume that the inversion between convolution and sub-sampling is licit for non-aliased images such as $k_\sigma * f$. The validity of this assumption on real images has been checked in [10]. In addition it is assumed (again, for well-sampled images) that the derivative of the continuous and discrete versions are the same up to a normalization due to the zooming of factor r. The validity of this assumption will be confirmed by the numerical experiments of the following sections. Writing $\partial_1 = \partial_x$ and $\partial_2 = \partial_y$, this yields :

$$w_{q,r,t} \approx r k_{rt} * k_\sigma * \partial_q f = r k_{\sqrt{r^2t^2+\sigma^2}} * \partial_q f.$$

We therefore deduce that :

$$\frac{w_{q,r,t}}{r} \approx k_{r\sqrt{t^2+p^2}} * \partial_q f. \tag{5}$$

Assume now that we have two images f_{r_1} and f_{r_2} of the same scene at resolutions r_1 and r_2. From (5), we deduce that if we choose scales t_1 and t_2 such that :

$$r_1\sqrt{t_1^2+p^2} = r_2\sqrt{t_2^2+p^2} \tag{6}$$

then :

$$w_{q,r_1,t_1}/r_1 \approx w_{q,r_2,t_2}/r_2 \tag{7}$$

Furthermore, we also have that :

$$m_1(r_1,t_1)/r_1 \approx m_1(r_2,t_2)/r_2 \tag{8}$$
$$m_2(r_1,t_1)/r_1^2 \approx m_2(r_2,t_2)/r_2^2 \tag{9}$$

with

$$m_1(r,t) = \frac{1}{|f_r|}\sum_q |w_{q,r,t}|,$$

where $|f_r|$ is the size of the discrete image f_r, and

$$m_2(r,t) = \frac{1}{|f_r|}\sum_q |w_{q,r,t}|^2.$$

Remark : A naive assumption could be drawn that for the same scene f, if we keep

$$r \times t = C \tag{10}$$

where C is a constant, the parameter set will also be constant (after the correct normalization). However, this assumption is not sufficient (especially on remote-sensing images) because it considers the resolution change simply by a zooming, which is not consistent with the acquisition process modeled in Section II.

B. Extrapolation of wavelet features

The aim of this paper is to propose a way to extrapolate wavelet features (i.e. the first and second order moments m_1 and m_2) from a resolution r_1 to a different resolution r_2. From equations (5)–(9), we deduce the following scheme : assume that we have f_{r_1} the image at resolution r_1 of a given scene, and that we want to predict its features at resolution r_2.

– Compute the wavelet coefficients for f_{r_1} at scales t_i, $i = 1, 2, 3, ..., N$;
– Estimate the parameters Θ_{r_1,t_i} from the wavelet coefficients at scales t_i for resolution r_1 ;
– For resolution r_2, compute the scales t'_i corresponding to t_i with the help of the function (see Equation (6))

$$t'_i = \sqrt{\frac{r_1^2}{r_2^2}(t_i^2+p^2) - p^2} \tag{11}$$

– Define $\tilde{\Theta}_{r_2,t'_i} = \Theta_{r_1,t_i}$ at scales t'_i.

Thanks to the previous process, it is now possible to compare, on similar bases, images taken at different resolutions and, for instance, to train classification methods on a set of images at only one resolution and to apply the recognition criteria to images at different resolutions.

V. Experiments and results

A. Validity of the extrapolation of features

In this section, we validate the proposed extrapolation scheme with some numerical experiments. The CNES (the French spatial agency) has provided us with images of several scenes (such as fields, forests and cities, see Figure V-A(a)-(c)) at various resolutions. It is important to note that convolution kernels used by the CNES are far from being Gaussian.

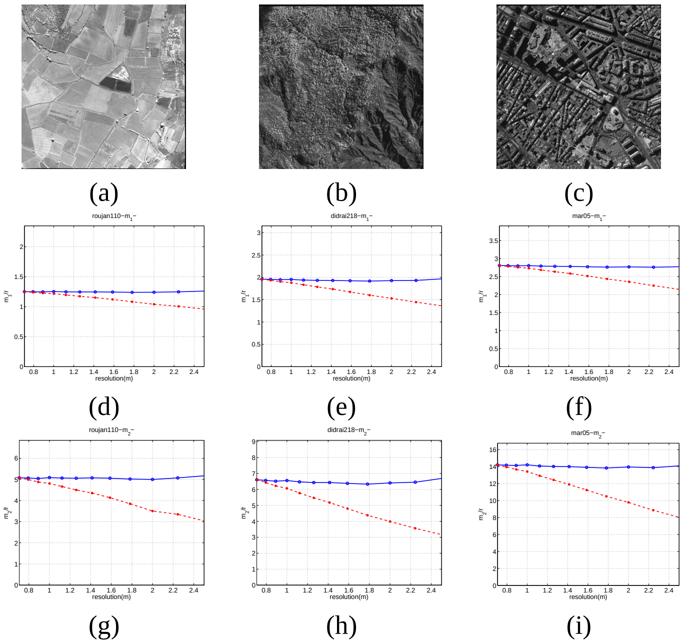


Fig. 2. **(a)-(c)** Three images (©CNES); **(d)-(f)** graph of $m_1(r,t)/r$ as function of r, and **(g)-(i)** graph of $m_2(r,t)/r^2$ as function of r. We display the case when $r^2\sqrt{t^2+p^2}$ is kept constant (with $p = 1.3$) with solid lines, and the case when rt is constant with dash lines

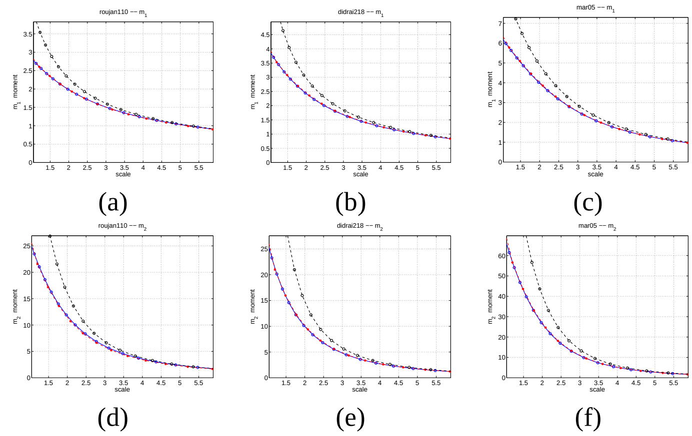


Fig. 3. Extrapolation of m_1 (**(a)-(c)**) and extrapolation of m_2 (**(d)-(f)**), starting from resolution $1m$ up to $3.175m$ for the images shown in Figure V-A(a)-(c). Dash lines show the extrapolation results obtained by using the scheme proposed in Section IV-B but replacing Equation (11) by $t_i' = r_1 t_i / r_2$ (the naive normalization). Solid lines : in each figure, there are in fact two lines which perfectly coincide with each other. They represent respectively the moments estimated directly on the images at resolution of $3.175m$ (the ground truth) and the extrapolation results obtained by using the scheme proposed in Section IV-B.

However, we will see that the approximate acquisition model of Section II yields good numerical results.

In Figure V-A(d)-(f)(resp. (g)-(i)) graphs of $m_1(r,t)/r$ (resp. $m_2(r,t)/r^2$) as functions of r are presented when rt is kept constant (that is when using the naive normalization of Equation (10)) and when $r\sqrt{t^2+p^2}$ (here $p = 1.30$) is kept constant (see Equation (6)). The resolution r ranges from $0.707m$ to $2.5m$. For the image at resolution $0.707m$ (the highest available resolution), m_1 and m_2 are computed at scale 5 and in the horizontal direction. It may be seen that using Equation (10) (that is forgetting the convolution step in the model of resolution change) does not yield a constant parameter set, especially when the resolution change is large. In such cases, one must use Equation (6) to extrapolate features.

Next, Figure V-A(a)-(c)(resp. (d)-(f)) show the extrapolations of m_1(resp. m_2) from a resolution of $1m$ to a resolution of $3.175m$ according to the scheme presented in Section IV-B, as well as the results obtained by the same scheme except that Equation (11) is replaced by $t_i' = r_i t_i / r_2$. It can be seen that when taking into account the convolution step in the resolution change, the parameter sets extracted from two different resolutions are nearly identical, which enables the classification of images at different resolutions.

B. Classification

In this subsection, we carry out the supervised classification of satellite images. The features we use are the first and second order moments m_1 and m_2, and we use the extrapolation scheme proposed in Section IV-B to compare parameter set extracted at different resolutions.

We built a database composed of images at different resolutions (see Figure 4 and Table I). We intend to classify it into three classes : fields, forests and cities. The examples are extracted manually from 4 kinds of satellite images (respectively at 3 resolutions) : Quickbird Panchromatic ($0.61m$) [11], Quickbird Multi-spectral ($2.44m$) [11], SPOT 5 THR ($2.5m$) [12] and SPOT5 HMA ($5m$). First and second order moments of wavelet coefficients are used for characterizing these images. We recall that derivatives of Gaussian kernels are used as wavelets (in the horizontal, vertical, and two diagonal directions). Since rotation invariance is important (objects of the same type may have different orientations), the mean values in the four directions are taken as features.

First, the performance of classification is tested on the SPOT5 HMA images by cross validation, using wavelet coefficients at scales 1, 2 and 4. Then all the images of SPOT5 HMA are used for training the classifiers in order to classify the other images (i.e. the Quickbird Panchromatic images, the Quickbird Multi-spectral images and the SPOT 5 THR images). For this purpose, the first and second order moments m_1 and m_2 for these three kinds of images are computed at scales ranging from 1 to 64. They are then compared with the features extracted from SPOT5 HMA images for classification.

We compare the classification results obtained respectively when $p = 0.00$ (i.e. when using Equation (10)) and $p = 0.50$ (when using (6)). Notice that the value of p is smaller than in Section V-A, because the images for classifications are not obtained by the same captors as the images used in Section V-A. The images used here are more aliased (consequently less blurred) than those shown in Figure V-A.

In Table II, the classification results using Knn (K nearest

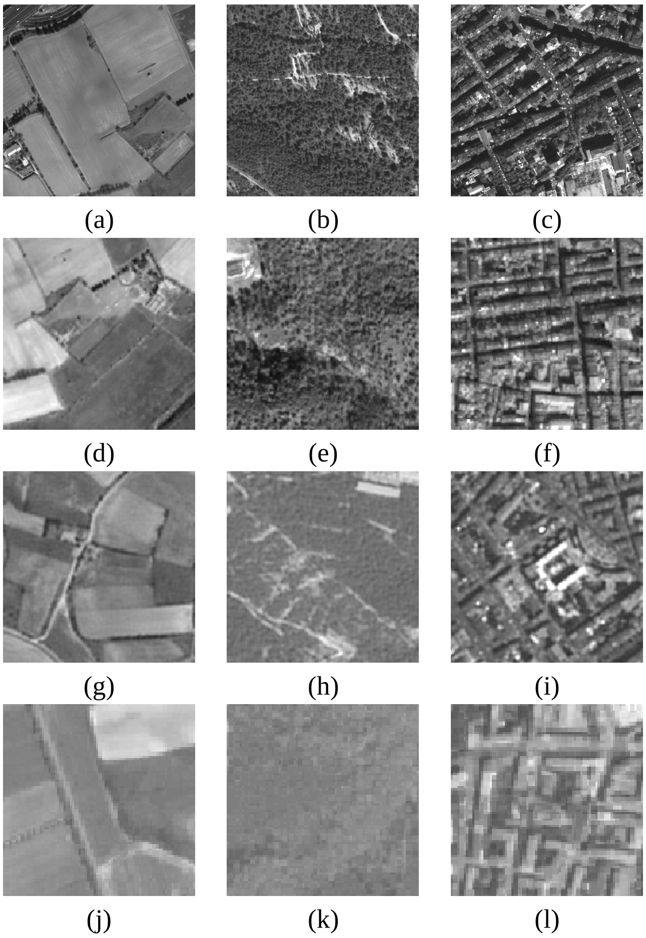


Fig. 4. Image samples : (a)-(c) Images of Quickbird Panchromatic (0.61m) ; (d)-(f) Images of Quickbird Multispectral (channel 1, 2.44m) ; (g)-(i) Images of SPOT5 THR (2.5m) ; (j)-(l) Images of SPOT5 HMA (5.0m).

TABLE I
NUMBER OF SAMPLES FOR EACH TYPE

image type	QB_PAN	QB_MUL	SPOT5_THR	SPOT5_HMA
field	7	8	38	100
forest	17	8	25	100
city	39	31	12	305
total	63	47	75	505

neighbor, that is, each sample is labeled according to the nearest training sample) are displayed. It can be seen that the use of Equation (6) always (though slightly) improve the classification results when compared with those obtained with the help of Equation (10). This improvement is less significant when resolution changes increase, because when $r_2 \ll r_1$, we have $t_i \gg p$ in Equation (11), yielding $t_i' \approx r_1 t_i / r_2$. This can be noticed in the first column of Table II.

VI. CONCLUSIONS

In this paper, we have proposed a scheme for extrapolating the wavelet features, which allows us to compare images taken at different resolutions. We first checked experimentally the validity of this scheme. This approach is then applied for the

TABLE II
CLASSIFICATION RESULTS (KNN, $k = 1$)

image type resolution	QB_PAN $0.61m$	QB_MUL $2.44m$	SPOT5_THR $2.5m$	SPOT5_HMA $5.0m$(c.v.)
wavelet ($p = 0.5$)	84.12%	82.97%	85.33%	96.63%
wavelet ($p = 0.0$)	84.12%	80.85%	82.67%	96.63%

classification of images at several resolutions. The classification performances are slightly improved by our scheme, compared to a naive approach where resolution change is simply modeled by a zoom. We believe that these improvements can be much more significant on larger databases and plan to carry out such experiments. We are also currently comparing the proposed approach with the use of different texture features (such as Haralicks features, [13]).

Acknowledgements : We thank Mihai Datcu, Alain Giros and Henri Maître for their advice and comments.

REFERENCES

[1] Fernand S. Cohen Zhigang Fan and Maqbool A. Patel, “Classification of rotated and scaled textured images using gaussian markov random field models,” *ieeepami*, vol. 13, no. 2, pp. 192–202, February 1991.

[2] Jianguo Zhang and Tieniu Tan, “Affine invariant texture signatures,” in *ICIP*, 2001.

[3] 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.

[4] B. Luo, J-F. Aujol, Y. Gousseau, S. Ladjal, and H. Maître, “Characteristic scale in satellite images,” in *ICASSP 2006*, 2006.

[5] J-F. Aujol, G. Aubert, and L. Blanc-Féraud, “Wavelet-based level set evolution for classification of textured images,” *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1634–1641, 2003.

[6] Gert Van de Wouwer, *Wavelets for Multiscale Texture Analysis*, Ph.D. thesis, Universiteit Antwerpen, 1998.

[7] Minh N. Do and Martin Vetterli, “Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance,” *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, February 2002.

[8] Stephane G. Mallat, “A theory for multiresolution signal decomposition : The wavelet representation,” *IEEE PAMI*, vol. 11, no. 7, pp. 674–693, July 1989.

[9] S.G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[10] B. Luo, J-F. Aujol, Y. Gousseau, S. Ladjal, and H. Maître, “Resolution independent characteristic scale in satellite images,” *submitted*, 2006.

[11] DigitalGlobe, *QuickBird Imagery Products - Product Guide*, DigitalGlobe, Feb. 2006.

[12] C. Latry and B. Rouge, “SPOT5 THR mode,” in *Proc. SPIE Vol. 3439, p. 480-491, Earth Observing Systems III, William L. Barnes ; Ed.*, W. L. Barnes, Ed., Oct. 1998, pp. 480–491.

[13] R.M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems Man and Cybernetics*, vol. SMC-3(6), 1973.