Local regularity for texture segmentation: Combining wavelet leaders and proximal minimization
Nelly Pustelnik, Herwig Wendt, Patrice Abry

To cite this version:

HAL Id: hal-01148354
https://hal.archives-ouvertes.fr/hal-01148354
Submitted on 4 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: http://oatao.univ-toulouse.fr/
Eprints ID: 12607

To link to this article: DOI: ICASSP.2013.6638684
URL: http://dx.doi.org/10.1109/ICASSP.2013.6638684


Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr
LOCAL REGULARITY FOR TEXTURE SEGMENTATION: COMBINING WAVELET LEADERS AND PROXIMAL MINIMIZATION

Nelly Pustelnik (IEEE Member)\textsuperscript{1}, Herwig Wendt (IEEE Member)\textsuperscript{2}, Patrice Abry (IEEE Fellow)\textsuperscript{1}

\textsuperscript{1} Physics Dept. - ENSL, UMR CNRS 5672, F-69364 Lyon, France, firstname.lastname@ens-lyon.fr
\textsuperscript{2} IRIT - ENSEEIHT, UMR CNRS 5505, F-31062 Toulouse, France, herwig.wendt@irit.fr

ABSTRACT

Texture segmentation constitutes a classical yet crucial task in image processing. In many applications of very different natures (biomedical, geophysics,...), textures are naturally defined in terms of their local regularity fluctuations, which can be quantified as the variations of local Hölder exponents. Furthermore, such images are often naturally embedded in the class of piece-wise constant local regularity functions. The present contribution aims at proposing and assessing a segmentation procedure for this class of images. Its originality is twofold: First, local regularity is estimated using wavelet leaders, a novel multiresolution quantity recently introduced for multifractal analysis but barely used in local regularity measurement, comparisons against wavelet coefficient based estimation are conducted; Second, the challenging minimal partition problem underlying segmentation is convexified and conducted within a customized proximal framework. The estimation of the number of regions and their target regularity is obtained from a total-variation estimate that enables the actual use of proximal minimization for texture segmentation. Performance is assessed and illustrated on synthetic textures.

Index Terms— Texture segmentation, piece-wise constant, local regularity, wavelet leaders, proximal minimization.

1. MOTIVATION, RELATED WORKS, CONTRIBUTIONS

Texture segmentation from local regularity fluctuations. Texture characterization and segmentation constitutes a challenging task in image processing (see e.g., [1, 2, 3, 4, 5] and references therein). In a variety of applications of possibly very different natures, the relevant information characterizing textures, and thus conveying the information to be analyzed, consists of the fluctuations across space of their local regularity. This is notably the case for biomedical textures such as bone [6] or breast [7] (see also [8] for a review), for surface imagery [9], for satellite imagery [10], or for image registration [11, 12].

Often, local regularity is assessed in terms of the so-called Hölder exponent $h(x)$ [13]. In essence, $h(x)$ consists of the power law exponent quantifying the decrease of the powers of local fluctuations, measured via multiresolution quantities $T_X(a, x)$ (e.g., wavelet coefficients) around a given space position $x$ when the analysis scales $a$ goes towards fine scales, i.e., $E(T_X(a, x)^p) \approx C(a) a^{-\tau_X(x)}$ when $a \to 0$.

Related works: Hölder exponent estimation and segmentation. Local regularity based analysis of textures can naturally be organized into two sub-questions: i) Accurately estimating the Hölder exponents for each location; ii) Splitting the image into regions where estimates can be considered constant, so as to achieve texture segmentation. Though these issues have received significant efforts over the past two decades for 1D signals (cf. e.g., [14, 15]), much less contributions were dedicated to 2D fields or images (see a contrario [12] and references therein).

Estimation relies on two key choices, with crucial impact on performance: Multiresolution quantities $T_X(a, x)$ from which the $h(x)$ are measured; Range of analysis scales $a$ to be practically involved in the estimation. For multiresolution quantities, most works conducted so far relied on either increments, oscillations or wavelet coefficients. Bilinear time-scale representations were also used [15]. The choice of the range of scales is naturally framed in a classical bias-variance trade-off. While theoretically, estimation should be conducted in the limit of fine scales, practically a finite range of scales $a_L \leq a \leq a_U$ has to be used. Choosing a large range of scales yields smaller estimation variance at the price of increasing the bias because the $T_X(a, x)$ at large scales are less well localized (cf. e.g., [16, 15]).

For segmentation, i.e., to detect changes amongst Hölder exponents measured at different locations, this trade-off also turns crucial. This is why both estimation and change detection have mostly been conducted in model dependent frameworks. Fractional Brownian motion (fBm) [17], the paradigm model for scale invariance, implies a constant and unique Hölder exponent across sample fields. Multifractional Brownian motion extends fBm by allowing piece-wise smooth variations of $h(x)$ across space (while preserving joint Gaussianity) [18] and constitutes essentially the only model with space-dependent $h$ that has been studied and used. The problem has mostly been addressed for 1D signals and barely for images (see a contrario [12, 7] which aimed at identifying particular points in the image defined by specific values of $h(x)$). Furthermore, to our knowledge, little work has been published aiming at detecting changes in $h(x)$ in a context where an exact model is not a priori assumed (see a contrario [15] for 1D signals).

With respect to the generic issue of image segmentation, efficient variational methods were proposed [19, 20, 21, 22, 23, 24], however none of them guarantees convergence towards a global minimum, and thus, obtained solutions strongly depend on initialization. Instead, a convex relaxation of the minimal partition techniques were proposed in [25, 26]. In [26], the algorithmic solution is based on a Arrow-Hurwicz type primal-dual algorithm but requires inner iterations and upper boundedness of the primal energy in order to improve convergence speed. Also, a proximal solution in a context of disparity estimation was proposed in [27]. Alternatively, recent techniques are framed into variational principles, making use of comparisons between pairs of neighboring patches [28, 3]. Notably, in [28], a convex relaxation of the original problem was envisaged so as to ensure to find the global minimizer, while in [3] a more general but non-convex framework was proposed, which yields good results at the price of a well-chosen initialization. However, the computational cost is a limitation of these methods.
Goals and contributions. In this context, the first originality of the present contribution consists in performing region-segmentation of images, that excludes the recourse to a fully parametric model for the data, but instead only assumes that images belong to the class of piecewise constant Hölder exponent functions. Its second originality lies in a twofold departure from earlier works: For multisresolution quantities, wavelet leaders are chosen and shown to yield improved statistical performance for Hölder exponent estimation, compared to the classical wavelet coefficients (cf. Section 2); For segmentation, using estimated Hölder exponents as inputs, a proximal algorithm is devised. It essentially relies on the customization of the formulation proposed in [27], which expands on that proposed in [26] to efficiently handle the resulting large size problem. Further, we propose an original method to a priori estimate the number of regions as well as their mean regularity, an initialization step that turns compulsory and crucial for the actual use of the proximal approach in that context. These are detailed in Section 3.

This combination (use of wavelet leaders and of proximal algorithm with original initialization) enables us to propose an efficient data-model free piecewise constant Hölder exponent minimal partition image segmentation procedure, that furthermore depends on very few parameters that need to be tuned manually.

2. WAVELET COEFFICIENTS AND LEADERS FOR HÖLDER EXPONENTIAL ESTIMATION

Local regularity and Hölder exponent. Let $X$ denote the bounded 2D function (image) to be analyzed. The local regularity around position $x_0 \in \mathbb{R}^2$ is measured by the so-called Hölder exponent $h(x_0)$, defined as the largest $\alpha > 0$, such that there exists a constant $C > 0$ and a polynomial $P_c$ of degree less than $\alpha$, such that $|X(x) - P_c(x)| \leq C|x - x_0|^\alpha$ in a neighborhood $\delta(x_0)$. When $h(x_0)$ is close to 0, the image is locally irregular and close to discontinuous. Conversely, a large $h(x_0)$ corresponds to a locally smooth field. For example, when $h(x_0)$ increases towards 2 then the fields becomes smoother and smoother up to being differentiable. Fig. 1-b below displays a Gaussian texture, with piecewise constant Hölder exponents, the region with the lowest $h$ (outer ring) is more irregular then the one with the largest $h$ (inner ring).

Wavelet coefficients. Let $\phi(x)$ and $\psi(x)$ denote the scaling function and mother wavelet defining a 1D multiresolution analysis. Let the 2D wavelets be defined as: $\psi^{(0)}(x) = \phi(x_1)\phi(x_2)$, $\psi^{(1)}(x) = \psi(x_1)\phi(x_2)$, $\psi^{(2)}(x) = \phi(x_1)\psi(x_2)$, $\psi^{(3)}(x) = \psi(x_1)\psi(x_2)$ with $x = (x_1, x_2)$. The collections $\psi^{(j)}(x_0)$ is $2^{-j}\psi^{(j)}(2^{-j}x - k)$ of dilated (to scales $2^a$) and translated (to space positions $k = 2^b$) templates of $\psi_0$ form a basis of $L^2(\mathbb{R}^2)$ for well chosen functions $\psi$. Let $d\phi^{(j)}(x) = \langle X, \psi^{(j)}(x) \rangle$ denote the (L1-normalized) coefficients of $X$ of the so-called discrete wavelet transform (DWT) of $X$. Readers may refer to [29] for further details.

Wavelet leaders. Wavelet leaders were recently introduced in the context of multifractal analysis [13, 30], their use in the context of local regularity measurement has however been barely considered. The wavelet leader $L_X(j, k)$, located around position $x_0 = 2^j k$ is defined as the local supremum of all wavelet coefficients taken within a spatial neighborhood across all finer scales $2^l \leq 2^j$:

$$L_X(j, k) = \sup_{m = 1, 2, 3 \ldots} |d\phi^{(2)}(j, k)|,$$

Local regularity estimation. It has long been known that, for large (but not all) classes of images [13], wavelet coefficients reproduce the Hölder exponent $h(x_0)$, in so far as, for all $x_0$ such that $2^j \lambda_x \approx x_0$, $\frac{1}{2} \sum_{m=1}^{2^j} |d\phi^{(j)}(x, \xi)| \approx C \delta_{2^j}^{(j)}(2^j x)$ when $2^j \to 0$. More recently, it has been shown [13, 30] that wavelet leaders systematically (that is: for all classes of images) reproduce the Hölder exponent $h(x_0)$: $L_X(j, k) \approx C \delta_X^{(j)}(2^j x)$ when $2^j \to 0$. This naturally leads to the estimation of $h(x_0)$ by means of linear regressions:

$$\hat{h}(x_0) = \sum_{j=1}^{2^j} \frac{w_j}{2} \sum_{m=1}^{2^j} |d\phi^{(j)}(x, \xi)|,$$

$$\hat{h}_L(x_0) = \sum_{j=1}^{2^j} \frac{w_j}{2} L_X(j, k),$$

where weights $w_j$ chosen to perform non weighted regression [15]. Section 4 compares estimates obtained from either wavelet coefficients or leaders and shows that the latter yield a substantial decrease in variance and are thus to be preferred.

3. MINIMAL PARTITION SEGMENTATION USING PROXIMAL TOOLS

Let $\Omega \subseteq \mathbb{R}^2$ denote a spatially continuous domain and $f \in L^2(\Omega)$ the image that we intend to label. Note that, in Section 4, $f$ will successively model the estimated Hölder exponents obtained from wavelet leaders or coefficients, i.e., $f = h_0$ or $f = h_L$. The Mumford-Shah approach in [20] consists in computing a piecewise constant approximation of data $f$. Assuming that $f$ can be approximated by $Q$ labels such that $(\Omega_1)_{1 \leq q \leq Q}$ models the $Q$ distinct regions, this problem can be written as

$$\min_{\Omega_1, \ldots, \Omega_Q} \sum_{q=1}^{Q} \left( f - v_q \right)^2 \text{d}x + \frac{1}{2} \sum_{q=1}^{Q} \text{Per}(\Omega_q),$$

s.t. $\bigcup_{q=1}^{Q} \Omega_q = \Omega$, $(\forall q \neq p), \Omega_q \cap \Omega_p = \emptyset$, (4)

where $v_q$ stands for the value of $f$ in region $\Omega_q$ (with, by convention, $v_q \leq v_{q+1}$), where the left-hand-side term consists of the sum of the variances of $f$ over all regions, and where Per$(\Omega_q)$ measures the perimeter of region $\Omega_q$, and where the constraints imposed on $\Omega_q$ ensure to obtain a non-overlapping partition of the image. The discrete analogue of Model (4) is the Potts model which is known to be NP-hard to solve. A solution to circumvent this difficulty consists in relaxing this non convex formulation into a convex approximation [26].

From now on, the sampled version of the image $f$ is considered and denoted as a vector of pixels $f = (f^{(n)})_{1 \leq n \leq N} \in \mathbb{R}^N$. The $Q$ regions $\Omega_q$ are labeled through an auxiliary variable $u = (u^{(n)})_{1 \leq n \leq N}$, such that $u^{(n)} = v_q$ if and only if the pixel $n$ belongs to the region $\Omega_q$. Further, one can introduce $Q$ binary functions $\theta = (\theta_1, \ldots, \theta_Q)$ such that, for every $q \in \{1, \ldots, Q\}$,

$$\theta_q^n = \begin{cases} 1 & \text{if } u^{(n)} \geq v_q, \\ 0 & \text{otherwise}. \end{cases}$$

(5)

It results that the labeling function $u$ can be recovered from $\theta = (\theta_1, \ldots, \theta_Q)$ through the relation

$$u^{(n)} = \sum_{q=1}^{Q} (v_q - v_{q-1}) \theta_q^n,$$

(6)
with \( v_k = 0 \). A bijection between \( u \) and \( \theta \) is guaranteed by constraining, for every \( n \in \{1, \ldots, N\} \):

\[ B_n = \{ \theta^{(n)} \in \{0, 1\} \times \ldots \times \{0, 1\}, 1 \geq \theta_1^{(n)} \geq \ldots \geq \theta_Q^{(n)} \geq 0 \}, \]

whose convex relaxation is

\[ B_n = \{ \theta^{(n)} \in \{0, 1\} \times \ldots \times \{0, 1\}, 1 \geq \theta_1^{(n)} \geq \ldots \geq \theta_Q^{(n)} \geq 0 \}. \]

Moreover, for every \( n \in \{1, \ldots, N\} \), \( \theta_1^{(n)} = 1 \).

**Convex criterion** The convexification of Model (4) leads to the following minimization problem involving the \( Q \) binary functions \( \theta = (\theta_1, \ldots, \theta_Q) [26] \):

\[
\begin{align*}
\min_{\theta_1, \ldots, \theta_Q} & \sum_{q=1}^{Q-1} \sum_{n=1}^{N} (\theta_1^{(n)} - \theta_{q+1}^{(n)}) (f^{(n)} - v_q)^2 \\
& + \sum_{n=1}^{N} \theta_Q^{(n)} (f^{(n)} - v_Q)^2 + \frac{1}{2} \sum_{q=1}^{Q} \rho_T (H\theta_q, V\theta_q) \\
\text{s.t.} & \begin{cases}
(\forall n \in \{1, \ldots, N\}), \theta_1^{(n)} = 1, \\
(\forall n \in \{1, \ldots, N\}), 1 \geq \theta_1^{(n)} \geq \ldots \geq \theta_Q^{(n)} \geq 0,
\end{cases}
\end{align*}
\]

where \( H \in \mathbb{R}^{N \times N} \) and \( V \in \mathbb{R}^{N \times N} \) are matrix representations of, respectively, the horizontal and vertical first-order discrete differences, and where the total variation reads, for every \( \eta = (\eta^{(n)})_{1 \leq n \leq N} \in \mathbb{R}^N \) and \( \zeta = (\zeta^{(n)})_{1 \leq n \leq N} \in \mathbb{R}^N \),

\[ \rho_{TV} : (\eta^{(n)})_{1 \leq n \leq N}, (\zeta^{(n)})_{1 \leq n \leq N} \mapsto \sum_{n=1}^{N} \sqrt{|\eta^{(n)}|^2 + |\zeta^{(n)}|^2}. \]

The functions involved in (7) are convex, lower semi-continuous, and proper. \( H \) and \( V \) are diagonalizable in the Fourier domain and the proximity operator of each function has a closed form. It results that the proximal algorithms PPXA+ [31] can be used to efficiently find the minimum of (7).

**Estimation of \((v_q)_{1 \leq q \leq Q}\)** The a priori choice of \((v_q)_{1 \leq q \leq Q}\) is likely to strongly impact the estimated \((\theta_q)_{1 \leq q \leq Q}\). Here, we propose to extract this value from a denoised estimator of \( f \) such that

\[ f^* = \arg \min_{f} \| f - g \|^2 + \lambda \rho_{TV} (Hg, Vg) \]

where \( \lambda > 0 \) denotes a regularization parameter, that impacts the quality of the denoised estimate. Section 4 further details the estimation of \((v_q)_{1 \leq q \leq Q}\) from \( f^* \).

4. RESULTS

**Numerical simulations.** The potential and performance of the combined estimation and segmentation procedures described in Sections 2 and 3 above are illustrated and assessed by application to independent realizations of synthetic images. They are produced numerically according to a 2D multifractional Brownian field model [14, 18], whose definition has been slightly modified here to ensure an homogeneous variance across the image:

\[ X(\xi) = C(\xi) \int_{\mathbb{R}^2} e^{i2\pi \xi \cdot \omega} \frac{dW(\omega)}{|\xi|^H + \sqrt{\frac{2}{\pi}}}. \]

Table 1. Estimation Performance. Coefficients vs leaders.

<table>
<thead>
<tr>
<th>( H = )</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
</tr>
</thead>
<tbody>
<tr>
<td>WavC</td>
<td>WavL</td>
<td>WavC</td>
<td>WavL</td>
</tr>
<tr>
<td>Bias</td>
<td>-0.36</td>
<td>-0.03</td>
<td>-0.19</td>
</tr>
<tr>
<td>Std</td>
<td>0.52</td>
<td>0.19</td>
<td>0.55</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.63</td>
<td>0.19</td>
<td>0.58</td>
</tr>
<tr>
<td>min</td>
<td>-0.14</td>
<td>0.27</td>
<td>0.27</td>
</tr>
<tr>
<td>max</td>
<td>-0.09</td>
<td>0.31</td>
<td>0.37</td>
</tr>
</tbody>
</table>

**Segmentation.** Synthetic data are produced according to the model in (9) with piecewise constant Hölder exponent function \( h(\xi) = H_{Q_{\xi}} \) for Region \( \xi \). Each region is thus regarded as a different texture and the goal is to achieve texture segmentation of the image.

As an instructive example, we use images consisting of \( Q = 3 \) different regions with piecewise constant Hölder exponents on concentric disks, as shown in Fig. 1-a. The three different textures can clearly not be identified from the histogram of the image pixel values shown in Fig. 1-c. In simulations, \( N = N_{Q_1} = 512 \times 512 \) and \( N_{Q_2} = 2 \) vanishing moments. Regularity is estimated using the scaling range \((j_1, j_2) = (1, 3)\).

**Estimation.** To compare estimation performance obtained with wavelet leaders against those produced with wavelet coefficients, the estimation procedure described in Section 2 is applied to 100 independent copies, of size \( N \times N \), with \( N = 1024 \) and with uniform Hölder exponent: \( h(\xi) \equiv H_{Q_{\xi}} \). Table 1 reports estimation performance for different values of \( H \) and clearly indicates that leader-based estimates of \( h \) systematically outperform wavelet coefficient based ones, with significant decrease in variances at the price of increased bias, overall yielding substantially decreased mean-square errors.

![Fig. 1. Synthetic data. (a) Piecewise constant Hölder exponent mask \((H_1, H_2, H_3) = (0.25, 0.50, 0.75)\), (b) Sample field, and (c) Corresponding histogram of the pixel values.](image-url)
consequence of the very large variance of the estimate $\hat{h}_d$. Despite
smaller variance, resulting in a non unimodal histogram (Fig. 3-b),
the leader-based estimates $\hat{h}_L$ (Fig. 3-a) do not clearly reveal the
three constant-$h$ regions either.

The denoising procedure described in Section 3 (cf. (8)), applied
to both $\hat{h}_d(x)$ and $\hat{h}_L(x)$, yields estimates $\hat{h}^+_d(x)$ and $\hat{h}^+_L(x)$ (resp.
Fig. 2-c and Fig. 3-c), whose histograms, shown in Fig. 2-d and
Fig. 3-d respectively, clearly display 3 modes. Segmentation based
on a direct thresholding of such histograms however yields poor re-
results as illustrated in Fig. 2-e and Fig. 3-e and Table 2. Instead, the
denoising procedure is used to initialize the proximal based seg-
mentation: The number of regions $Q$ is set to the number of local maxima
observed in the denoised estimates $\hat{h}^+_d(x)$ and $\hat{h}^+_L(x)$, here $Q = 3$;
The target mean intensity values $q_x, q = 1, 2, 3$ are set to values cor-
responding to the positions of the local maxima of the histograms of
estimates $\hat{h}^+_d$ and $\hat{h}^+_L$.

The proximal tool based segmentation procedure described in
Section 3 is now applied to both $\hat{h}_d(x)$ and $\hat{h}_L(x)$, initialized as de-
scribed above. Results are compared in Fig. 2-f and Fig. 3-f, (for one
single image) and quantified in Table 2 (averaged across 100 inde-
pendent copies). This shows that the number of mis-classified pixels
is systematically smaller when segmentation is based on $\hat{h}_L(x)$. Dif-
fences between wavelet coefficient and leader based performance
may appear small, this is however only a consequence of the cho-

...
6. REFERENCES


