HAL
open science

Characterization of signals by the ridges of their wavelet transforms

René Carmona, Wen Liang Hwang, Bruno Torresani

To cite this version:

René Carmona, Wen Liang Hwang, Bruno Torresani. Characterization of signals by the ridges of their wavelet transforms. IEEE Transactions on Signal Processing, 1997, 10.1109/78.640725 . hal-01223134

HAL Id: hal-01223134

https://hal.science/hal-01223134v1

Submitted on 2 Nov 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.

CHARACTERIZATION OF SIGNALS BY THE RIDGES OF THEIR WAVELET TRANSFORMS

René A. Carmona *
Statistics & Operations Research Program
Princeton University
Princeton, N.J. 08544

Wen L. Hwang*†
Department of Mathematics
University of California at Irvine
Irvine, CA 92717, USA

Bruno Torrésani
CPT, CNRS-Luminy, Case 907
13288 Marseille Cedex 09
FRANCE

October 9, 1996

Abstract

We present a couple of new algorithmic procedures for the detection of ridges in the modulus of the (continuous) wavelet transform of one-dimensional signals. These detection procedures are shown to be robust to additive white noise. We also derive and test a new reconstruction procedure. The latter uses only information from the restriction of the wavelet transform to a sample of points from the ridge. This provides with a very efficient way to code the information contained in the signal.

*Partially supported by ONR N00014-91-1010

†Supported by NSF IBN 9405146

1 Introduction

The characterization and the separation of amplitude and frequency modulated signals is a classical problem of signal analysis and signal processing. Applications can be found in many situations, such as for instance radar/sonar detection and speech processing [10]. Many methods have been proposed in the past few years to analyze the time-frequency localization of signals. The most noticeable are the family of bilinear representations such as the Wigner representation and its generalizations (see [1, 7] for a review) and the linear representations such as the wavelet and Gabor transforms.

In 1990 by the Marseille group proposed a new algorithm (see [6] for a survey) based on the study of the phase of the wavelet (or Gabor) transform. The present work is an attempt to extend the latter to noisy situations. The main thrust of this correspondence is to use the localization properties of the modulus of the transform (which is generally more robust than the phase, eventhough the latter provides more precise estimates [6]). In the case of frequency-modulated signals, the wavelet transform is "concentrated" in the neighborhood of curves (the ridges of the transform). We develop a scheme in which these curves are searched as such, in a (high dimensional) space of ridges, via a stochastic relaxation procedure. This alternate characterization of the ridges is better suited to the needs of noisy signal analyses. We also propose a stable method for signal reconstruction from the numerically computed ridges. This method is also based on an L^2-minimization procedure.

For the sake of simplicity, our discussion is restricted to the case of the wavelet transform. But since our algorithms deal only with post-processing of time-frequency transforms, they can be extended to any time-frequency energetic representations. The case of the Gabor transform will be considered in the companion paper [4] where still another stochastic search algorithm, adapted to different situations, will be introduced.

We close this introduction with a short summary of the contents of the paper. Section 2 below is devoted to the statement of the problem and the definition of the ridges, Section 3 presents the main features of the variational problems which we propose and solve to estimate the ridges. We also give a Bayesian interpretation of this approach and we describe how one can modify the penalization functional in order to accommodate the presence of an additive noise in the signal. This section ends with a discussion of an example of a bat sound signal which we embed in noise. Section 4 is devoted to a quick account of the reconstruction problem. Namely, given a set of points in the time-scale domain, find the signal most likely to have a ridge going through these points. As before we outline the mathematical derivations and we illustrate the efficiency of the method on a numerical example.

2 Ridges

The first goal of this section is to set up an abstract formalism for the mathematical definition of the ridges of functions of two variables. Then, we propose two Monte Carlo algorithms to detect and identify these ridges. Let $\psi \in L^1(\mathbb{R})$ be such that $0 < c_\psi = \int_0^\infty |\hat{\psi}(\xi)|^2 \, d\xi/\xi < \infty$, *i.e.* fulfills the *wavelet admissibility condition.* The corresponding wavelet transform of $f(x)$ is given by:

$$T_f(b,a) = \langle f, \psi_{(b,a)} \rangle = \frac{1}{a} \int_{-\infty}^{\infty} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx = e^{-\varphi} \int_{-\infty}^{\infty} f(x) \overline{\psi\left(e^{-\varphi}(x-b)\right)} dx \ , \tag{1}$$

where we have introduced the auxiliary variable $\varphi = \log(a)$. We are mostly interested in frequency modulated signals which can be written as the sum of finitely many components of the form:

$$f(x) = A(x) \cos \phi(x) \ , \tag{2}$$

but for the purpose of the present correspondence we shall restrict ourselves to monocomponent signals. See [4] for a detailed analysis of the multicomponent case. It is convenient to use the so-called "progressive wavelets", *i.e.* with vanishing negative frequencies. If $\psi(x)$ is such a wavelet, then the wavelet coefficients of $f(x)$ are given by: $T_f(b,a) = \langle f, \psi_{(b,a)} \rangle = \frac{1}{2} \langle Z_f, \psi_{(b,a)} \rangle$ where $Z_f(x)$ is the "analytic signal" of $f(x)$ given by: $Z_f(x) = \frac{1}{\pi} \, \mathrm{P.} \int f(x+y) \frac{dy}{y}$, where P denotes principal value integral. It is well-known [1] that a signal of the form (2) with $A(x)$ and $\phi'(x)$ slowly varying gives $Z_f(x) \approx A(x) \exp\{i\phi(x)\}$. If we assume that the Fourier transform $\hat{\psi}(\xi)$ is peaked near a particular value $\xi = \omega_0$ of the frequency, like for instance the Morlet wavelet given in the Fourier domain by $\hat{\psi}(\xi) = \exp\{-(\xi - 2\pi)^2/2\}$, it follows from standard arguments [6] that the wavelet transform may be approximated as:

$$T_f(b,a) \approx \frac{1}{2} A(b) \exp\{i\phi(b)\} \overline{\hat{\psi}(a\phi'(b))} + O(|A'|/|A|, |\phi\phi''|/|\phi'|^2) \ . \tag{3}$$

From the localization properties of the wavelet in the Fourier domain one can see that the modulus $|T_f|$ of the wavelet transform is essentially maximum in the neighborhood of a curve $a = a_r(b) = \exp \varphi(b)$, the *ridge of the wavelet transform*, related to the instantaneous frequency of the signal by $a_r(b) = e^{\varphi_r(b)} = \omega_0/\phi'(b)$. In [6], the phase coherence of the wavelet transform was used to get a numerical estimate of the ridge. Since the phase can be somewhat difficult to control in noisy situations, we shall mainly focus here on the localization of the maxima of the modulus of the wavelet transform.

3 Ridge Detection: Variational Approaches

The purpose of this section is to give two examples of ridge detection algorithms both derived from variational problems. In both cases the ridge is searched in a high-dimensional space of curves and the

ridge estimate appears as the graph of the argument of the minimization of a suitable penalty function. Unlike the methods in [6], the penalty function is mainly on the square modulus of the wavelet transform. The loss of accuracy is weak, since the signals for which the methods are designed are supposed to have slowly varying frequencies. In the first case, the ridge is the graph of a function $b \to \varphi(b)$ while it is the graph of a parametric curve in the second case. The results of this section can be used beyond the single component case (the wavelet transform has a single ridge) when the ridges can be separated by a preprocessing localization procedure and then analyzed separately.

3.1 A Direct Search Algorithm: We first assume that the ridge of the wavelet transform of the signal f can be parametrized by a function $b \hookrightarrow \varphi(b)$ defined for all the values of b. For the sake of the present discussion we denote by Φ the space of all the twice differentiable functions with square integrable derivatives. We then define the penalty function F_f on the set Φ of ridge candidates φ by:

$$F_f(\varphi) = -\int |T_f(b, e^{\varphi(b)})|^2 db + \int \left[\lambda \varphi'(b)^2 + \mu \varphi''(b)^2\right] db \tag{4}$$

Such a penalty function clearly implements the two following features: the smoothness of the ridge and the localization in the time-frequency plane (for $\lambda = \mu = 0$, minimizing $F_f(\varphi)$ is equivalent to searching maxima of $|T_f|^2$ in the a direction). Our estimate of the unknown ridge of the wavelet transform of the signal f will be the function $\varphi(b)$ which minimizes $F_f(\varphi)$. The Euler equation associated with this minimization problem can easily be obtained, and once discretized into a finite difference equations, solved numerically. However, such an approach is efficient only for weak noise. The presence of a strong noise component implies the existence of many local extrema in which the algorithm may get trapped. We need a procedure which can jump over the local extrema to reach the global one(s). A natural candidate for this is the simulated annealing algorithm [9].

3.2 Snake Penalization We now consider a ridge as a parametrized curve $r : s \in [0,1] \to r(s) = [\rho_1(s), \rho_2(s)]$ in the time-scale plane. The ridge then takes the form of a "snake" (see [8] for a description of the method in an image processing context). We use a cost function which takes into account the modulus of the wavelet transform, as well as additional terms needed in order to ensure the smoothness of the ridge (both in the b and a directions). We set:

$$F_f(r) = -\int |T_f\left(\rho_1(s), \rho_2(s)\right)|^2 ds + \int \left[\lambda_a \rho_2'(s)^2 + \mu_a \rho_2''(s)^2 + \lambda_b \rho_1'(s)^2 + \mu_b \rho_1''(s)^2\right] ds \ , \tag{5}$$

where λ_a, λ_b, μ_a and μ_b are positive constants. In the "snake terminology" of [8] the second term is the "internal energy" of the snake. Its role is to control the smoothness and the rigidity of the snake. The first term is the "external energy" of the snake. It accounts for the interaction of the snake with the wavelet transform modulus. For the reasons mentioned in the previous section, we turn to

stochastic optimization techniques (see [8] for a direct solution of the corresponding Euler equations) for the numerical solution of such a minimization problem.

Remark: In many applications the signal $f(x)$ is the sum of a pure component $f_0(x)$ and a noise component $n(x)$. When some information on the noise is available, it may be included into the penalty function (see e.g [3, 4, 5] for more details on this point).

3.3 Bayesian Interpretation: Both ridge detection procedures have a Bayesian interpretation. Let us present it in the case of the direct search. Consider the prior probability measure defined formally by $\mu_{prior}(d\varphi) = Z_1^{-1} \ \exp\{-\int [\mu_a|\varphi'(b)|^2 + \lambda_a|\varphi''(b)|^2]\, db\}\ ''d\varphi''$ and the conditional probability $\mu_\varphi(df) = Z_2^{-1} \ e^{\int |T_f(b,\varphi(b))|^2}\ ''df''$. which gives the probalility, conditioned by φ, that the signal is in the infinitesimal $''df''$ in the space of finite energy signals. Then according to Bayes' rule, the conditional probability knowing the signal is given by $\mu_{posterior}(d\varphi|f) = \exp -F[\varphi]''d\varphi''/Z$ for some constant Z. Maximizing $\mu_{posterior}(d\varphi|f)$ is equivalent to minimizing (4).

3.4 Cost Minimization by Simulated Annealing: We included the ridge detection procedures described above in a package of S functions made available on the Internet [5]. The implementation was done by solving the variational problems by simulated annealing (see [9] for background on this combinatorial optimization technique). The details are spelled out in [3] and the book [5].

3.5 Examples: Numerical experiments have been made on various types of academic and real signals. We illustrate the method described above with a (real) sonar signal emitted by certain species of bats. The signal is frequency modulated, with approximately hyperbolic instantaneous frequency. The wavelet transform (w.r.t. Morlet's wavelet) was computed for frequencies ranging from $\nu_s/16$ to $\nu_s/2$, with ν_s the sampling frequency, in geometric progression (i.e. of the form $a = 2a_0^{n/20}, n = 0, \ldots 59$). Figure 1 shows the wavelet transform of the signal (left) and the wavelet transform of the same signal with additive Gaussian white noise, with input $SNR = -5dB$. Superimposed are the ridges estimated with the direct search procedure. Both transforms are coded with gray levels proportional to their modulus square. In Figure 2, we show the ridge estimated with the snake procedure (left, notice that the boundaries have been fairly well reproduced), and a comparison of ridge estimations in various situations (right).

4 Reconstruction from the Skeleton on a Ridge

We present in this section a new algorithmic reconstruction of a signal from the knowledge of sample values of its wavelet transform on the ridges of its modulus. For the sake of simplicity we restrict ourselves to the case of a single ridge. See nevertheless [4] for the analysis (in the case of the Gabor transform) of the more general case of finitely many arbitrary ridges. Let us focus on ridges given in

the form $b \to \varphi(b)$. In practical applications one only knows sample points $(b_1, a_1), \ldots, (b_n, a_n)$ and the smooth function $b \hookrightarrow \varphi(b)$ which we use in lieu of the true (unknown) ridge function is merely a guess which one constructs from the sample points. We use a smoothing spline (but any other kind of nonlinear regression curve would do as well). From now on, $\varphi(b)$ is a smooth ridge function which is constructed from the n sample data points.

4.1 Statement of the Problem: We are concerned with the implementation of the folk belief that *a signal can be characterized by the values of the restriction of its wavelet transform to its ridges.* Illustrations can be found in [6], where it is shown that in the case of signals of the form (2) the restriction of the wavelet transform to its ridge of the wavelet transform behaves as $A(x)\exp[i\phi(x)]$ (see also [10] for similar remarks for the Gabor transform in the context of speech, yielding good quality reconstruction with high compression rate). Such an approach can be used in non-noisy situations, but it does fail in the presence of a significant noise component. We assume that the values of the wavelet transform, say z_j, are known at sample points (b_j, a_j). The set of sample points together with the values z_j constitutes the wavelet transform *skeleton* of the signal to be reconstructed. We look for a signal $f(x)$ of finite energy whose wavelet transform has the graph of the function $a_r(b)$ as ridge and satisfies:

$$T_f(b_j, a_j) = z_j, \qquad j = 1, \ldots, n \ . \tag{6}$$

4.2 The Penalization Approach: We use a strategy which was successfully used by Mallat and Zhong to reconstruct a signal from the extrema of its dyadic wavelet transform [11, 2]. In the present setting, we look for a signal $f(x)$ which satisfies the constraints (6) while the L^2-norm in the scale variable a of the modulus is kept to a minimum for each b. This may be achieved by minimizing

$$F_1(f) = \frac{1}{c_\psi} \int db \int \frac{da}{|a|} |T_f(b,a)|^2 \tag{7}$$

(note that when the integration is performed on the whole half-plane, $F_1(f) = ||f||^2$ by the energy conservation formula). Since the cost function $F_1(f)$ is a quadratic form in the unknown function f, the solution is easily computed by means of Lagrange multipliers. A solution can be constructed as a linear combination of the wavelets $\psi_{(b_j,a_j)}$ at the sample points of the ridge, the coefficients being given by the solution of a $n \times n$ linear system. This solution is not completely satisfactory, especially when the number of sample points is small. It ignores the empirical fact that (in most of the practical cases) the restriction of the modulus $|T_f(b,a)|$ to the ridge, i.e. the function $b \hookrightarrow |T_f(b, a_r(b))|$, is smooth and slowly varying. In order to force the solution of the constrained optimization problem to respect this requirement, we introduce the extra term $\tilde{F}_2(f) = \int_{b_{min}}^{b_{max}} \left| \frac{d}{db} |T_f(b, a_r(b))| \right|^2 db$, and consider the minimization of the cost function $\tilde{F}(f) = F_1(f) + \epsilon \tilde{F}_2(f)$ where the free parameter $\epsilon > 0$ is chosen to balance the two contributions to the penalty. Unfortunately, $\tilde{F}_2(f)$ is not quadratic in f. In order to remedy this problem we remark

that according to the analysis of [6], $\frac{d}{db}\Omega(b, a_r(b)) \approx F'(b) = \omega_0/a_r(b)$, where $\Omega_f(b,a) = \arg T_f(b,a)$. Then we replace $\tilde{F}_2(f)$ by a quadratic form which gives a good approximation of it, and $\tilde{F}$ with

$$F(f) = \frac{1}{c_\psi}\int db \int \frac{da}{|a|}|T_f(b,a)|^2 + \epsilon \int_{b_{min}}^{b_{max}} \left(|\frac{d}{db}T_f(b,\varphi(b))|^2 - \frac{\omega_0^2}{\varphi(b)^2}|T_f(b,\varphi(b))|^2 \right) db = \langle Qf, f\rangle \ . \quad (8)$$

4.3 Solution of the Optimization Problem: The constrained minimization problem can be solved using Lagrange multipliers. The solution is given by

$$\hat{f}(x) = \textstyle\sum_{j=1}^{2n} \lambda_j Q^{-1}\psi_j(x) \quad (9)$$

where Q is the operator (matrix after discretization of the problem) defined in (8) and the functions ψ_j are defined by: $\psi_j(x) = a_j^{-1}\psi\left((x-b_j)/a_j\right)$, $j = 1, \ldots, n$, The Lagrange multipliers are determined by imposing the constraints (6). This gives a system of $(2n) \times (2n)$ linear real equations from which the Lagrange multipliers λ_j's can be computed.

4.4 Examples: To illustrate the reconstruction procedure, we selected a subset of $n = 500$ consecutive samples from the bat signal (Figure 3, left). We used 40 sample points on the estimate of the ridge and the value $\epsilon = .5$ to reconstruct the signal. The result of the reconstruction is given at the bottom of Figure 3. As may be seen, the reconstruction is of extremely good quality. An analysis (not presented here) of the modulus of the wavelet transform of the reconstructed signal shows that, because we chose a ridge estimate which ignored the existence of a secondary ridge, the latter is not present in the reconstruction. Further results (see [4, 5]) confirm the quality of the reconstruction method. This justifies a posteriori the approximation we made in the definition of the quadratic penalty function.

5 Conclusions

We presented a new approach to the problem of ridge detection in an energetic distribution of a signal. Our approach is based on the minimization of a penalty function on the set of all possible ridge candidates. The penalty function takes into account a-priori information on the signal (namely the time-frequency representation of the signal which is essentially localized around a curve, this curve is smooth,...) and possibly on the noise (through an a-priori noise model, or simulations). The minimization is achieved through Monte-Carlo type methods. We also proposed a new synthesis procedure which requires only a small number of values of the transform on the ridge and which is very robust to noise. We have focused here on the case where the time-frequency representation is given by the square modulus of the wavelet transform (the scale variable being interpreted as an inverse frequency variable). Any time-frequency representation can be used as well. The case of the the square modulus of the Gabor transform will be considered in a forthcoming publication. [4].

References

[1] B. Boashash (1992): Estimating and Interpreting the Instantaneous Frequency of a Signal, *Proc. IEEE* **80**; Part I, pp. 520–538. Part II: *Proc. IEEE* **80**, pp. 540–568.

[2] R. Carmona (1992): Spline Smoothing & Extrema Representation: Variations on a Reconstruction Algorithm of Mallat and Zhong. in *Wavelets and Statistics*, Lecture Notes in Statistics, A. Antoniadis & G. Oppenheim Eds.

[3] R. Carmona, W.L. Hwang and B. Torrésani (1995): technical report, available at `http://cptsx1.univ-mrs.fr/~ torresan/publi.html`.

[4] R. Carmona, W.L. Hwang and B. Torrésani (1995): Multiridge Detection and Time-Frequency Reconstruction (submitted for publication).

[5] R. Carmona, W.L. Hwang and B. Torrésani (1996): Practical Time/Frequency Analysis: Wavelet and Gabor Transforms with an Implementation in S. (book under review)
Swave: a Public Domain Package of S Functions for Time-Frequency Analysis.
Available at `http://soil.princeton.edu/ rcarmona/Publications`, or `anonymous ftp from chelsea.princeton.edu /pub/outgoing/Swave`

[6] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, B. Torrésani (1992): Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies. *IEEE Trans. Inf. Th.* **38**, pp. 644-664.

[7] P. Flandrin (1993): Temps-Fréquence. *Traité des Nouvelles Technologies, Hermès.*

[8] M. Kass, A. Witkin and D.Terzopoulos (1988): Snakes: Active Contour Models, Int. J. of Computer Vision, 321-331.

[9] P.J.M. van Laarhoven and E.H.L. Aarts (1987): Simulated Annealing: Theory and Applications. Reidel Pub. Co.

[10] R.J. McAulay and T.F. Quatieri (1986): Speech Analysis/Synthesis Based on a Sinusoidal Representation. *IEEE Trans. on Audio, Speech and Sign. Proc.* **34 #4** 744–754.

[11] S. Mallat and S. Zhong (1992): Characterization of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. Machine Intel.* **14**, #7.

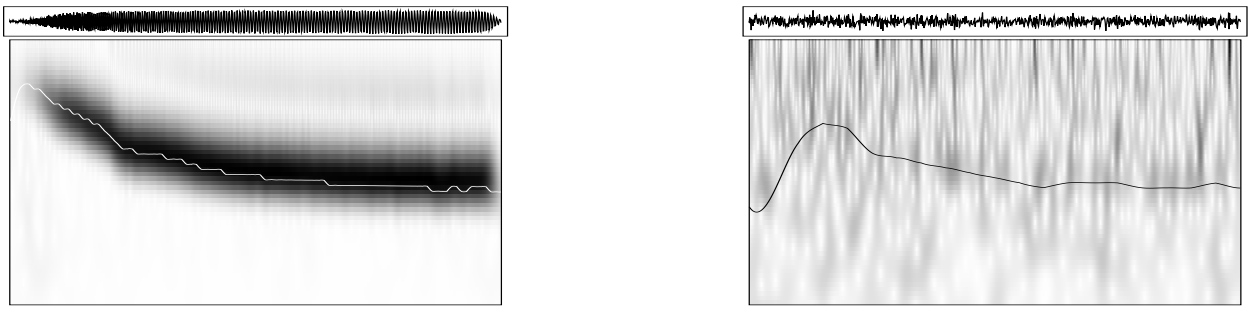

Figure 1: Left) Intensity plot of the modulus square of the wavelet transform of the bat signal. The ridge superimposed. Right) Ridge estimate, annealing method; $SNR = -5dB$.

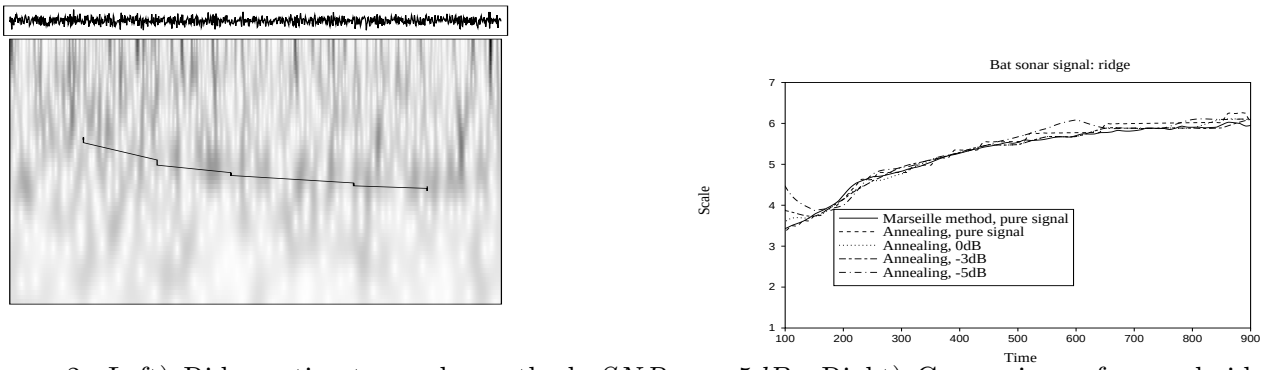

Figure 2: Left) Ridge estimate, snake method; $SNR = -5dB$. Right) Comparison of several ridge estimates for the bat sonar signal. Notice that the scale is now increasing upward as opposed to all the other plots where it is increasing downward.

Figure 3: Top) Bat signal used to illustrate the reconstruction procedure. Bottom) Result of the reconstruction from the values of the wavelet transform at the 40 points of the estimate of the ridge.