MULTI-RIDGE DETECTION AND TIME-FREQUENCY RECONSTRUCTION

René A. Carmona [*] Wen L. Hwang [*][†] Bruno Torrésani

Department of Mathematics
University of California at Irvine
Irvine, CA 92717, USA

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

June 21, 1995

Abstract

The ridges of the wavelet transform, the Gabor transform or any time-frequency representation of a signal contain crucial information on the characteristics of the signal. Indeed they mark the regions of the time-frequency plane where the signal concentrates most of its energy. We introduce a new algorithm to detect and identify these ridges. The procedure is based on an original penalization of the transitions of the random walk in a bounded domain of the plane. We show that this detection algorithm is especially useful for noisy signals with multi-ridge transforms. It is a common practice among practitioners to reconstruct a signal from the skeleton of a transform of the signal (i.e. the restriction of the transform to the ridges). After reviewing several known procedures we introduce a new reconstruction algorithm and we illustrate its usefulness on speech signals.

[*]Partially supported by ONR N00014-91-1010

[†]Supported by NSF IBN 9405146

1 Introduction and Notations

A wide class of signals may be conveniently described in terms of time-dependent amplitude and frequency, or sums of such amplitude and frequency modulated components. The main problem is often that of the numerical evaluation of such time-dependent quantities. Time-frequency representations are in general well adapted to this situation. However, although the evaluation of local amplitude and frequency is easy in the nois free case with only one component, the situation gets much more complex in the presence of noise and/or of several components. For example, bilinear representations such as the affine Wigner distributions can yield extremely precise results in the one-component case, but may completely fail in the multicomponent situation because of the presence of interference terms.

We proposed in [4] and [5] a new constrained optimization approach for processing one-component signals even in very noisy situations. We describe here another approach, capable of handling multiple component signals. This new approach is mainly based on two dinstinct steps: time-frequency representation of the signal, and generation of an associated random walk on the time-frequency plane. In the first step, the energy of the signal concentrates around "ridges", i.e. curves in the time-frequency plane. In the second step, the random walk is constructed in such a way that the random walkers (hereafter named "crazy climbers") are attracted by such one-dimensional structures. A third "resynthesis step" may be added if one wishes to reconstruct the "part" of the signal that led to the ridge.

Our discussion will be most of the time restricted to the case of the Gabor transform. The reason is our desire to consider applications to speech signals for which the Gabor transform is better suited to. It is important to notice that since our detection algorithm is only with post-processing of a time-frequency transforms, it can be used with other time-frequency energetic representations (such as the wavelet transform considered in [4]).

We close this introduction with a short summary of the contents of the paper. Our detection algorithm is presented in a general setting in Section 2, and the particular case of time-frequency transforms is studied in section 3. Examples of noisy signals are used to show the efficiency of the algorithm. Section 4 gives the details of a reconstruction procedure of the original signal from the estimates of the ridges of the Gabor transform of the observed signal. The principle of this reconstruction is not new. It was implemented for the wavelet transform in the companion paper [4].

The novelty of the presents result is twofold. We give a new ridge detection algorithm which can efficiently detect multiple ridges and we reconstruct a signal from the knowledge of the skeleton of its Gabor transform on arbitrary points of its ridges. The benefits of these two novelties are illustrated in the last section in which we apply our algorithms to the case of speech signals.

2 Drawing the Ridges of a Surface

2.1 Generalities

The goal of this section is to set up an abstract formalism for the numerical detection of ridges associated with surfaces or functions of two variables. Following [11], ridges may be defined as curves on a surface $z = f(x, y)$. They are completely characterized by their projection on the (x, y) plane. We also call these curves ridges by a convenient abuse of language. We shall be interested in special kinds of ridges, which we describe below.

We start with a subset D in the upper (time-frequency) half plane. D will be bounded in the applications but we can think of D as the whole upper half-plane for the purpose of the present discussion. We shall use the notation (b, ω) for the points of the domain D. The variable b is very common in works using the continuous wavelet transform while the use of the variable ω is standard in works involving Fourier transform and/or the Gabor transform.

We consider a nonnegative function $M(b, \omega)$ defined on a subset D of the upper (b, ω) half-plane. We define the ridge set R as the set of local maxima in ω of the functions $\omega \hookrightarrow M(b, \omega)$ when the "time variable" b is held fixed. We assume that the surface $M(b, \omega)$ is smooth enough so that the ridge set is the finite union of the graphs of smooth functions slowly varying on their respective domains. In other words we assume that:

$$R = \cup_{\ell=1}^{L} R_\ell \tag{1}$$

where each R_ℓ is the graph of a smooth function:

$$[b_{\ell,min}, b_{\ell,max}] \ni b \hookrightarrow \omega_\ell(b)$$

defined on a (possibly strict) subset of the domain of the variable b. In the practical applications which we have in mind, the ridge functions $\omega_\ell(b)$ are slowly varying. But notice that we do not make any assumption on the lengths of the individual ridges R_ℓ or even the fact that they could cross.

2.2 The Crazy Climbers Detection Algorithm

The main idea of the "crazy climbers" algorithm is as follows. A certain number of particles (the climbers) are initially randomly seeded on the domain D at step 0. Then each climber starts a random walk on D influenced (in a way similar to the simulated annealing algorithm) by the local values of the $M(b,\omega)$ function. In a nutshell, the algorithm combines simulated annealing in the ω direction and symmetric random walk in the b direction. The climbers are then encouraged to "climb on the hills" to reach the ridges.

In order to give the details of the algorithm we consider a discretized (and finite) version of the setup and we use the terminology of the Gabor transform to illustrate the meaning of the discretization.

We assume that the time interval over which the signal is analyzed is discretized into a finite set $\{b_0, b_1, \cdots, b_{B-1}\}$ with B elements. We also assume that the values of the frequency variable ω are discretized into a finite set $\{\omega_0, \omega_1, \cdots, \omega_{K-1}\}$. Setting $M(j,k) = |G_f(b_j, \omega_k)|$ we reduce the analysis of the modulus of the Gabor transform to the analysis of a finite $B \times K$ matrix with nonnegative entries.

2.2.1 Crazy climbers

At time $t = 0$ we initialize the positions $X_\alpha(0)$ of N climbers on the grid $\Gamma = \{0, \cdots, B-1\} \times \{0, \cdots, K-1\}$. The climbers are labelled by the parameter $\alpha = 1, \cdots, N$. The initial positions are chosen independently of each other, uniformly over the grid Γ. The climbers evolve independently of each other according to the same distribution. This distribution can be characterized in the following way. If a climber is at the point (j,k) at time t, i.e. if $X_\alpha(t) = (j,k)$, then its position at time $t+1$, say $X_\alpha(t+1) = (j', k')$ is determined according to the following law: $j' = j-1$ with probability $1/2$ and $j' = j+1$ with probability $1/2$. We do not discuss the particular cases $j = 0$ and $j = B-1$ involving boundary conditions not to confuse the issue. Then when the climber has decided to move to the left (when $j' = j-1$) or to the right (when $j' = j+1$) in the horizontal direction, the possible vertical move is considered. As for the horizontal component, the climber tries to move up (i.e. $k' = k+1$) or down (i.e. $k' = k-1$) with equal probabilities. Again we ignore the boundary conditions for the sake of simplicity. Unlike in the case of the horizontal direction, the move does not always take place. The transition from (j', k) to (j', k') takes place if the value of the function increases, i.e. if its so-called Delta

$$\Delta M = M(j', k') - M(j', k)$$

is nonnegative. On the other hand, the move does not necessarily take place if the function decreases, i.e. if $\Delta M < 0$. Indeed, in this case the transition is made i.e. $X_\alpha(t+1) = (j', k')$ with probability $\exp[\Delta M/T(t)]$ and the climber does not move vertically, i.e. $X_\alpha(t+1) = (j', k)$ with probability $1 - \exp[\Delta M/T(t)]$.

At each time t we consider two occupation measures. The first one is defined by:

$$\mu_t^{(0)} = \frac{1}{N}\sum_{\alpha=1}^{N} \delta_{X_\alpha(t)}.$$

It is obtained by putting a mass $1/N$ at the location of each of the climbers on the grid. In other words, $\mu_t^{(0)}(A)$ is the proportion of climbers in A at time t. The second one is the "weighted" occupation measure μ_t defined by:

$$\mu_t = \sum_{\alpha=1}^{N} M(X_\alpha(t))\delta_{X_\alpha(t)}$$

and obtained by putting a mass equal to the value of the function M at the location of the climber.

We finally consider the corresponding "integrated occupation measures", defined by ergodic averages as follows:

$$\mu_I^0 = \frac{1}{T}\sum_{t=1}^{T} \mu_t^0 \qquad (2)$$

and

$$\mu_I = \frac{1}{T}\sum_{t=1}^{T} \mu_t \qquad (3)$$

The occupation measure μ_I^0 is only given here for the sake of completeness. Indeed its main shortcoming is the fact that it assigns nonzero mass to regions without ridges if the lengths of the ridges are smaller than the length of the window. This is due to the very nature of the unrestricted horizontal motions of the climbers. Because the modulus of the *denoised* versions of the functions M which we use in the applications are essentially zero away from the ridges, the occupation mesure μ_I gives much better results when it comes to detecting ridges.

Further Remark:

Notice that the climbers evolve independently of each other without interaction. Moreover, their time evolution has the same statistical distribution. This means that the computer code generating the motions of the climbers is the same for all the climbers. This is an indication that the algorithm can naturally be parallelized on a SIMDIM machine (such as for example the massively parallel computers MASPAR I & II). We shall not report of such an implementations here.

2.2.2 A simple example

To illustrate the ridge detection algorithm, we present in Figure 1 the simple example of a sine wave multiplied by a Gaussian envelope (top left). The modulus of its Gabor transform with a Gaussian window is displayed at the top right of the figure, and the two integrated occupation measures ar displayed in the bottom of the figure. We clearly see on the figure the different meanings of the two measures. The weighting of the measure plays the role of a thresholding for the ridge detection.

2.2.3 Chaining

The output of the algorithm described above is a measure on the domain D. We identify it with its density which is a function on D. The next step of the algorithm is to identify the various ridges R_ℓ. This is done via a chaining procedure which replaces the occupation density by one-dimensional curves. This procedure is based on the following two steps:

1. Thresholding of the density function given as the output of the crazy climbers algorithm.

2. "Propagation" in the b direction: given a point (b,ω) belonging to a given ridge R_ℓ, look for the "best neighbor" among $(b+\epsilon_b,\omega)$ and $(b+\epsilon_b,\omega\pm\epsilon_\omega)$ (here ϵ_b and ϵ_ω are parameters fixed in advance); then iterate the process until only values below the threshold can be reached.

The result is a series of ridges, which are graphs of curves $\omega=\omega_\ell(b),\ b=b_0^\ell\cdots b_{k_\ell}^\ell$.

We shall not go into further details concerning the chaining algorithm here.

3 Ridges of Time-Frequency Representations

The goal of this section is to implement the detection algorithm introduced above in the special cases of the modulus of the continuous wavelet and Gabor transforms. The generalization to other time-frequency representations is straightforward.

We first describe the notations and the conventions we use in the sequel. We work with the complex Hilbert space $L^2(\mathbb{R})$ of square-integrable functions. Our convention for the Fourier transform is:

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \tag{4}$$

and consequently the Plancherel formula reads $||\hat{f}||^2 = 2\pi ||f||^2$.

3.1 The Case of the Continuous Wavelet Transform

We first consider the case of the continuous (complex-valued) wavelet transform. We work with a fixed function $\psi \in L^1(\mathbb{R})$ such that:

$$0 < c_\psi = \int_0^\infty |\hat{\psi}(\xi)|^2 \frac{d\xi}{\xi} < \infty, \tag{5}$$

i.e. fulfilling the wavelet admissibility condition. The wavelet transform of a signal $f(x) \in L^2(\mathbb{R})$:

$$T_f(b,a) = \frac{1}{a} \int_{-\infty}^{\infty} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx. \tag{6}$$

We use complex-valued wavelets providing a description of the complex Hardy space $H^2(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}),\ \hat{f}(\xi) = 0\ \forall \xi \leq 0 \right\}$. We use the notation:

$$\psi_{(b,a)}(x) = \frac{1}{a} \psi\left(\frac{x-b}{a}\right) \tag{7}$$

for the wavelet with scale a and location b. With such notations, it directly follows from Taylor's formula that the wavelet transform of signals of the form:

$$f(x) = \sum_{k=1}^{N} A_k(x) \cos(\phi_k(x)) \tag{8}$$

can be written in the form:

$$T_f(b,a) = \frac{1}{2} \sum_{k=1}^{N} A_k(x) e^{i\phi_k(b)} \overline{\hat{\psi}(a\phi_k'(b))} + r(b,a) \tag{9}$$

provided the amplitudes $A_k(x)$ and the phases $\Phi_k(x)$ are differentiable and twice differentiable respectively. Here

$$r(b,a) \sim O(|A_k'|, |\phi_k''|).$$

As a consequence, assuming that the wavelet $\psi(x)$ is localized near a certain value $\xi = \omega_0$ in the frequency domain, the wavelet transform modulus $M(b,a) = |T_f(b,a)|$ is localized near the N curves with equations $a = \varphi_k(b) = \omega_0/\phi_k'(b)$. These curves of the time-frequency plane are called the ridges of the transform.

We introduced in [4] a couple of algorithms capable of detecting ridges in the modulus of a wavelet transform. We illustrated the desirable features and the shortcomings of these algorithms in [4] and [5]. We illustrate in figures 2 the use of the crazy climber algorithm for the detection of ridges in the modulus of a wavelet transform. These examples are intended to show that this detection algorithm can overcome the limitations of the algorithms introduced in [4].

3.2 The Case of the Gabor Transform

We next consider the case of the Gabor transform. Although Gabor's original representation was discrete, we nevertheless call the continuous version which we use a Gabor transform. So the Gabor transform of a signal $f(x)$ of finite energy is defined as:

$$G_f(b,\omega) = \int_{-\infty}^{\infty} f(x) g(x-b) e^{-i\omega(x-b)}\, dx \ , \tag{10}$$

where $g(x)$ is a window function with a good time-frequency localization. We use the Gaussian kernel:

$$g(x) = g_s(x) = \frac{1}{s\sqrt{2\pi}} e^{-x^2/2s^2}, \tag{11}$$

where s is a scale parameter but other choices such as the Hamming windows which are very popular in speech processing would be as convenient. We shall use the notation:

$$g_{(b,\omega)}(x) = g(x-b) e^{i\omega(x-b)} \tag{12}$$

for the time-frequency atoms used in the definition of the Gabor transform. By the same argument as before, the continuous Gabor transform of signals of the type (8) may be written in the form:

$$G_f(b,\omega) = \frac{1}{2} \sum_{k=1}^{N} A_k(x) e^{i\phi_k(b)} \overline{\hat{g}(\phi_k'(b) - \omega)} + r(b,\omega). \tag{13}$$

Again the remainder term $r(b,\omega)$ depends upon the derivatives of the amplitudes and the local frequencies. Assuming for simplicity (this is the case for the Gaussian windows as well as for the Hamming windows) that the Fourier transform of the window has fast decay away from the origin of frequencies, we end up again with a Gabor transform modulus $M(b,\omega) = |G_f(b,\omega)|$ exhibiting a certain number of ridges.

Characterization of the signal's instantaneous frequency by the Gabor transform can be achieved by extracting numerically the ridges as the set of local maxima of the modulus of the Gabor transform. More precisely, for each fixed value of the time variable b we consider the local maxima of the function $\omega \hookrightarrow |G_f(b,\omega)|$ and we assume that this set can be represented as in (1).

3.3 Numerical Results for the Two Transforms

The crazy climber algorithm was tested on several signals containing multiple ridges. We restrict the present discussion to two examples, one treated with the wavelet transform and the other with the Gabor transform.

The first signal (displayed at the top of figure 2) is the sum of a (real) sonar signal emitted by a bat (this particular signal, displayed at the very top of Figure 2, together with noisy versions, was intensively studied in [4]) and a "linear chirp", i.e. a function of the form $A(x)\cos(\phi(x))$ with $A(x)$ a Gaussian function and $\phi(x)$ a quadratic phase. Figure 2 shows the modulus of the wavelet transform of the signal, together with the three different ridges found by the crazy climber method: the main ridge of the bat signal, the first harmonic component, and the chirp signal. The horizontal axis is the time axis, and the vertical axis corresponds to the logarithm of the scale. The modulus is represented with gray levels proportional to the wavelet transform modulus. The different ridges are displayed with different gray levels.

The second signal is a speech signal, namely 250 ms of the word /one/ (sampling frequency: 8kHz). The signal is displayed at the very top of Figure 6, and the modulus of its Gabor transform is the third item of the figure (here we used a Gaussian window of size approximately equal to 16ms (following the lines of [13]), and we computed the Gabor transform over the range 0Hz–4000Hz (with 100 different values for the frequency)). The horizontal axis is the time axis, and the vertical axis is the frequency axis (the conventions for the modulus and ridge displays are the sames as before). The crazy climbers algorithm (run with 500 climbers, and 10000 time steps each) found 18 different ridges, which are displayed at the bottom of Figure 6.

We shall come back to these examples when discussing the reconstruction from ridges.

3.4 Case of Noisy Signals

In many applications the observed signal $f(x)$ appears as the sum of a pure component $f_0(x)$ and a noise component $n(x)$. In some situations, "a-priori" knowledge of the noise is available (for instance the cases where the power spectrum of the noise is known, or the cases where a piece of the signal is known to contain only noise,which gives us the chance to learn about the statistics of this noise). Then the detection algorithm may be improved, by "renormalizing" the time-frequency representation,i.e. subtracting what is supposed to be the "typical" contribution of the noise. This contribution could be chosen to be the expectation $\mathbb{E}[M_n(b,\omega)]$ over all the possible realizations of the additive noise n, where $M_n(b,\omega)$ is the considered time-frequency representation of $n(x)$. If an a-priori model for the noise is available, such a quantity may be estimated by Monte-Carlo simulations, or sometimes by a direct computation.

Example: Assume that $n(x)$ is a second order stationary noise with power spectrum of the form $p(\xi) \sim \sigma^2\xi^\alpha$. Then it is easy to derive that if one considers its wavelet transform $T_n(b,a)$, $\mathbb{E}[|T_n(b,a)|^2] \sim K_\alpha\sigma^2 a^{-\alpha-1}$ provided the wavelet $\psi(x)$ is such that $K_\alpha = \int u^\alpha|\hat{\psi}(u)|^2du < \infty$.

In most practical applications, we only have one realization of the noise component and it is impossible to compute directly this expectation. A simple ergodic argument justifies the use of the estimate:

$$V(a) = \frac{1}{B}\int_0^B |T_n(b,a)|^2 db. \tag{14}$$

Then in the penalty term used to define the ω motion of the climbers, the squared modulus of the time-frequency transform may be replaced with:

$$\tilde{T}_f[b,\omega] = |T_f(b,\omega)|^2 - V(\omega) \tag{15}$$

The effect of such a modification is to avoid "trapping" the ridge in regions dominated by the noise.

As an illustration, we display in figure 6 the ridges of the same signal as before, embedded into a Gaussian white noise, with signal to noise ratio equl to 0dB. We can see that the main ridge is quite well reconstructed, and that the ridge of the chirp is also recovered, although only a part of it has been detected, namely the most energetic part. The first harmonic component of the bat signal has not been detected (the corresponding wavelet transform modulus was too low compared to the typical size of the noise).

4 Reconstructions from the Ridges

We now address the problem of the reconstruction of a signal from its ridge(s). Though the ridge detection part of the algorithm was independent of the chosen time-frequency representation, this is not the case for the reconstruction any more. In particular, the approach developed in [4] is not adapted to bilinear time-frequency representations such as those given by the Wigner-Ville transform. We shall restrict ourselves to the cases of the wavelet and the Gabor transforms. Nevertheless, our approach extends to linear representations such as those obtained from matching pursuit (see [17]) or those developed in [21]). Since [4] contains already the major elements of the reconstruction procedure of a signal from the ridge skeleton of its wavelet transform, we shall concentrate in this section on the specificities of the difficulties due to the presence of multiple ridges and of the sampling/compression issues which were not addressed in [4].

4.1 General Discussion

In order to put our reconstruction algorithm in perspective we first review the procedures currently used. The methods described in subsections 4.1.2 and 4.1.3 will be developed with more details in section 4.2.

4.1.1 The Transform Skeleton

The first reconstruction is the simplest one (once the ridges have been estimated). It consists in restricting the transform (whether we are working with the wavelet transform of the Gabor transform) to the ridges. It is motivated by the approximate formulas (9) and (13). More precisely, using the notation used throughout the paper, this reconstruction is given in the Gabor case by:

$$\hat{f}(x) = \sum_{\ell=1}^{L} G_f(x, \omega_\ell(x)) \tag{16}$$

where the summation in the right hand side is restricted to the ℓ's for which $\omega_\ell(x)$ makes sense (in particular, $\hat{f}(x) = 0$ when there is no ridge at "time" x), and by a similar formula in the wavelet case. The restriction of a transform to a ridge is sometimes called a "skeleton" of the transform [7].

As Figure 3 shows, the results of this naive reconstruction can be very good. Its main shortcoming is that it requires the knowledge of the transform at ALL the points of the ridges . This limitation makes it impossible to subsample the ridge (for compression purposes for example).

4.1.2 Ridge Penalization

The second approach was introduced an used in [4] in the case of the wavelet transform with a single ridge. We use freely the notations of this paper. The rationale of the method was based on the desire to mimic the values of the transform at the points of a sample from the ridge while trying to make sure that the modulus of the transform was maximum on the estimated ridge and that this modulus was smooth and slowly varying along the ridge. The way we made sure that these two requirements were satisfied was to introduce an L^2-penalty for the transform and a smoothness penalty for the modulus of the transform along the ridge. The results of this method are illustrated in Figure 6 in the case of the Gabor transform and Figure 3 in the wavelet case. This method allows for a sampling of the ridge, but the nature of the L^2 penalty forces the modulus of the tranform to be small in between two successive sample ridge points, especially when the distance separating these points is large enough for the smoothness penalty not to overwhelm the overall costs. This may create undesirable oscillations in the modulus of the transform of the reconstructed signal. The latter cannot be ruled out because the second component of the cost introduced in our variational formulation is actually penalizing the smoothness more than the oscillations.

4.1.3 Sampling Penalization

In order to alleviate the problems noticed above in the limitations of the subsampling of the ridge (and consequently the compression rate) we can replace the L^2-norm component of the cost by a penalty involving only the L^2-norms of the functions $\omega \hookrightarrow M(b_j, \omega)$ for the values $b = b_j$ of the first components of the sample points of the ridges. The effect of this change is to replace the identity matrix (as the first component of the matrix Q) by a diagonal matrix $\tilde{I}$ given by formula:

$$\tilde{I}(x, y) = \frac{1}{||g||^2}\,\delta(x-y)\sum_{j=1}^{J}|g(x-b_j)|^2 \tag{17}$$

4.1.4 Non-linear Reconstruction

The main drawback of the above mentioned reconstruction schemes is that they are fundamentally linear, and as such, constrained by the theory of frames of wavelets and/or Gabor functions (see [6] for a review). This puts severe limitations on possible subsampling.

For the sake of completeness we quote a non-linear reconstruction scheme that has been successfully used in speech processing in the framework of the so-called sinusoidal model. See [13] for a review. The main observation is that although the wavelet and the Gabor transforms restricted to ridges are often very oscillatory (and then uneasy to compress), the corresponding amplitudes $A(b)$ and frequencies $\nu(b)$ are sometimes slowly varying (this is the case for the Gabor transform of speech signals when the window is broad band). The rationale is then to try to represent an amplitude and frequency modulated signal by coding $A(b)$ and $\nu(b)$ with a few values. The reconstructed signal is then obtained by interpolating the sampled amplitude and frequency, integrating the frequency to recover the phase

$$\phi(b) = \phi(b_0) + \frac{1}{2\pi}\int_{b_0}^{b} \nu(x)dx$$

and then computing $A(b)\cos(\phi(b))$.

4.2 The Penalization Approach

We now focus on the penalization approach, and on the Gabor case (since the wavelet case has been described in [4], we shall only mention it briefly). Our purpose is to present a reconstruction algorithm which produces, from the mere knowledge of the Gabor transform at sample points of the ridges, a very good approximation of the original signal. This reconstruction procedure was implemented and tested in [4] and [5] in the case of the wavelet transform of a signal when the latter had only one ridge. More than presenting the details of the computations in the case of the Gabor transform, the purpose of the present discussion is to illustrate the new difficulties occurring because of the presence of multiple ridges and the singularities and/or instabilities due to the possible confluences of the various ridge pieces.

We assume that the ridges can be parametrized by continuous functions $[b_{\ell,min}, b_{\ell,max}] \ni b \hookrightarrow \omega_\ell(b) \in (0,\infty)$ where $\ell \in \{1,\cdots,L\}$ is the ridge label. These ridges are usually constructed as smooth functions resulting from fitting procedures (spline smoothing is an example we are using in practical applications) from the sample points obtained from ridge estimation algorithms such as the hive algorithm presented in this paper or the snake annealing described in [4].

4.2.1 Statement of the Problem

We assume that the values of the Gabor transform of an unknown signal of finite energy $f_0(x)$ are known at sample points $(b_{\ell,j}, \omega_{\ell,j})$ which are regarded as representative of the ridge of the modulus of the Gabor transform of the (unknown) signal $f_0(x)$. We use the notation $g_{\ell,j}$ for the value of the Gabor transform of f_0 at the point $(b_{\ell,j}, \omega_{\ell,j})$. The set of sample points $(b_{\ell,j}, \omega_{\ell,j})$ together with the values $g_{\ell,j}$ constitute what we call the skeleton of the Gabor transform of the signal to be reconstructed.

As we already explained, we use smooth functions $b \hookrightarrow \omega_\ell(b)$ which fit the sample points and we look at the graphs of these functions as our best guesses for the ridges of the modulus of the Gabor transform of f_0.

The reconstruction problem is to find a signal $f(x)$ of finite energy whose Gabor transform $G_f(b,\omega)$ satisfies:

$$G_f(b_{\ell,j}, \omega_{\ell,j}) = g_{\ell,j}, \qquad \ell = 1, \cdots L, \;\; j = 1, \cdots, n_\ell \tag{18}$$

and has the union R of the graphs of the functions $\omega_\ell(b)$ as set of ridges. Recall that this last statement means that for each b, the points $(b, \omega_\ell(b))$ of the time-frequency plane are the local maxima of the function $\omega \hookrightarrow |G_f(b,\omega)|$. We are about to show how to construct such a signal. We will also show that it is a very good approximation of the original signal $f_0(x)$.

4.2.2 The Penalty function for the wavelet case

For completeness, we reproduce here the description of the reconstruction procedure in the case of wavelet transform. More details can be found in [4]. As explained in [4], the signal reconstructed from a ridge is obtained as the function $\tilde{f}(x)$ minimizing a quadratic functional $F(f) = \langle f, Q \cdot f\rangle$ with the constraints $T_{\tilde{f}}(b_j, a_j) = z_j$. Here Q is an integral operator, defined by the kernel

$$Q(x,y) = \delta(x-y) + \epsilon Q_2(x,y)\ .$$

where $Q_2(x,y)$ is a term that intends to enforce smoothness of $|T_{\tilde{f}}|$ on the ridge.

$$\begin{aligned} Q_2(x,y) \;\; = \;\; & \int \frac{db}{\varphi(b)^4} \Big(\psi(\frac{x-b}{\varphi(b)})^* \psi(\frac{y-b}{\varphi(b)})[\varphi'(b)^2 - \omega_0^2] \\ & + \psi'(\frac{x-b}{\varphi(b)})\psi'(\frac{y-b}{\varphi(b)})[\frac{(x-b)(y-b)}{\varphi(b)^2} + 1 + \frac{x-2b+y}{\varphi(b)}] \\ & + \psi(\frac{x-b}{\varphi(b)})\psi'(\frac{y-b}{\varphi(b)})\varphi'(b)[1 + \frac{y-b}{\varphi(b)}] + \psi'(\frac{x-b}{\varphi(b)})\psi(\frac{y-b}{\varphi(b)})\varphi'(b)[1 + \frac{x-b}{\varphi(b)}] \Big)\ . \end{aligned}$$

4.2.3 The Penalty function for the Gabor case

The argument given for the (continuous) wavelet transform of frequency modulated signals can be reproduced here with only minor changes. This leads to the minimization of the cost function:

$$\mathcal{G}(f) = \mathcal{G}_1(f) + \epsilon \sum_{\ell=1}^{L} \int \left[\left| \frac{d}{db} G_f(b, \omega_\ell(b)) \right|^2 - \omega_\ell(b) |G_f(b, \omega_\ell(b))|^2 \right] db. \tag{19}$$

where, owing to the above discussion

$$\begin{cases} \mathcal{G}_1(f) &= \int \left(\int |G_f(b,\omega)|^2 d\omega \right) db \qquad \text{or} \\ \mathcal{G}_1(f) &= \sum_j \int |G_f(b_j,\omega)|^2 d\omega \end{cases} \tag{20}$$

The first term, together with the constraints (18), forces the localization of the energy distribution of the Gabor transform on the ridges as given by the graphs of the functions $b \hookrightarrow \omega_\ell(b)$. The second term is a quadratic form which provides a good approximation of

$$\epsilon \sum_{\ell=1}^{L} \int \left| \frac{d}{db} |G_f(b, \omega_\ell(b))| \right|^2 db$$

As explained in the companion paper [4], this part of the cost function is designed to insure that the tops of the ridges remain smooth and slowly varying. The free parameter $\epsilon > 0$ can be chosen to control the relative importance of the two contributions to the penalty. Our reconstruction procedure is given by the solution of the minimization of $F(f)$ subject to the linear constraints (18) which we rewrite in the form $L_j(f) = g_j$ where the linear functionals L_j are defined by:

$$L_j(f) = \int f(x) g\left(x - b_j\right) e^{-i\omega_j(x-b_j)} dx. \tag{21}$$

4.3 Solution of the Optimization Problem

A simple computation shows that the second term in the right hand side of equation (19) reads:

$$\mathcal{G}_2(f) = \int \int \mathcal{G}_2(x,y) f(x) f(y)\, dx dy$$

where the kernel $\mathcal{G}_2(x,y)$ is defined by the formula:

$$\begin{aligned} \mathcal{G}_2(x,y) \;=\; \sum_\ell \Bigg(& \int \Big(g'(x-b)g'(y-b) + g(x-b)g(y-b)[(x-b)(y-b)\omega_\ell'(b)^2 \\ - \; & (x+y-2b)\omega_\ell(b)\omega_\ell''(b)] \Big) \cos\left(\omega_\ell(b)(x-y)\right)\, db \\ + \; & \int \Big(g'(x-b)g(y-b)[\omega_\ell(b) - (y-b)\omega_\ell'(b)] \\ - \; & g(x-b)g'(y-b)[\omega_\ell(b) - (x-b)\omega_\ell'(b)] \Big) \sin\left(\omega_\ell(b)(x-y)\right) \Bigg)\, db \end{aligned} \tag{22}$$

The total cost function $\mathcal{G}(f)$ is then expressed as the sum of two terms

$$\mathcal{G}(f) = \mathcal{G}_1(f) + \mathcal{G}_2(f) = \int\int \mathcal{G}(x,y) f(x) f(y)\, dx dy$$

and is a quadratic functional given by the kernel:

$$\mathcal{G}(x,y) = \mathcal{G}_1(x-y) + \epsilon\mathcal{G}_2(x,y). \tag{23}$$

where we have set

$$\left\{\begin{array}{lcl} \mathcal{G}_1(x,y) & = & \delta(x-y) \qquad\qquad\qquad\qquad \text{or} \\ \mathcal{G}_1(x,y) & = & \dfrac{1}{||g||^2}\,\delta(x-y)\displaystyle\sum_{j=1}^{J}|g(x-b_j)|^2 \end{array}\right. \tag{24}$$

Notice that the kernel $\mathcal{G}(x,y)$ becomes a finite matrix for the purpose of practical applications. Notice further that formulas (23), (24) and (22) give a practical way to compute the entries of the matrix $\mathcal{G}$.

The reconstruction may be conveniently reformulated as a minimization problem in the real domain rather than the complex domain by noticing that since we stick to the case of real-valued signals, the kernel $\mathcal{G}(x,y)$ is real valued, and by replacing the n complex constraints (18) by the $2n$ real constraints

$$R_j(f) = r_j, \qquad\qquad j = 1,\cdots,2n, \tag{25}$$

where:

$$\left\{\begin{array}{lcl} R_j(f) & = & \displaystyle\int f(x) g(x-b_j)\cos(\omega_j(x-b_j))\, dx \ \text{ and} \\ R_{n+j}(f) & = & \displaystyle\int f(x) g(x-b_j)\sin(\omega_j(x-b_j))\, dx \end{array}\right. \tag{26}$$

for $j = 1,\cdots,n$ and $r_j = \Re z_j$ and $r_{n+j} = \Im z_j$.

Consequently, there exist real numbers $\lambda_1,\cdots,\lambda_n$, $\lambda_{n+1},\cdots,\lambda_{2n}$ (the Lagrange multipliers of the problem) such that the solution $\tilde{f}(x)$ of the optimization problem is given by:

$$\tilde{f}(x) = \sum_{j=1}^{2n}\lambda_j\tilde{g}_j(x) \tag{27}$$

where the functions $\tilde{g}_j$ are defined by:

$$\left\{\begin{array}{lcll} \tilde{g}_j(x) & = & \Re Q^{-1} g_{(b_j,\omega_j)}, & j = 1,\cdots,n, \\ \tilde{g}_j(x) & = & \Im Q^{-1} g_{(b_j,\omega_j)}, & j = n+1,\cdots,2n, \end{array}\right. \tag{28}$$

The Lagrange multipliers are determined by requiring that the constraints (25) be satisfied. In other word, by demanding that the wavelet transform of the function $\tilde{f}$ given in (27) be equal to the z_j's at the sample points (b_j, ω_j) of the time-scale plane. This gives a system of $(2n) \times (2n)$ linear equations from which the Lagrange multipliers λ_j's can be computed.

Remark: An alternative to the constraints (26) consists in imposing a constraint only on the real part of the time-frequency transform on the ridge, on twice as more ridge samples.

4.4 The Reconstruction Algorithm

The results of the discussion of this section may be summarized in an algorithmic walk through our solution to the reconstruction problem.

- determination of a finite set $\{R_\ell\}_{\ell=1,\cdots,L}$ of ridges and, on each of them, of a set of sample points $(b_1, \omega_{\ell,1}) = \omega_\ell(b_1), \cdots, (b_n, \omega_{\ell,n(\ell)} = \omega_\ell(b_n(\ell))$ on the ridge
- Construction of smooth estimates $b \hookrightarrow \omega_\ell(b)$ of the ridges from the sample points
- computation of the matrix $\mathcal{G}(x, y)$ of the smoothness penalty along the ridge estimate
- computation of the reconstruction time-frequency atoms $\tilde{g}_{\ell,j} = \mathcal{G}^{-1} g_{(b_{\ell,j}, \omega_{\ell,j})}$ localized (in the time-frequency plane) at the ridge sample points.
- Computation of the coefficients λ_j

The solution $\tilde{f}$ of the reconstruction problem is then given by formula (27).

4.5 The Case of Noisy Signals: Smoothing Spline Type Reconstruction

Section 4.2 addresses the problem of the reconstruction from observations of the wavelet or Gabor transform at a finite sample of points of the time/scal or time/frequency plane. This problem was considered in [4] when the values of the transform were assumed to be observed faithfully. In this paper we consider the possibility of an additive (possibly colored) noise in the observations of the input signal and the possibility of noise in the computation of the transform of the signal. As before our approach is motivated by the smoothing splines technique as presented in [22]. The generalization presented in this paper was alluded to as a possible extension to the reconstruction algorithm derived and used in [2] and [4]. The motivation of [2] was to simplify and shed light on the algorithm introduced by Mallat and Zhong to reconstruct a signal from the extrema of its dyadic wavelet transform. The motivation of [4] was to generalize this approach to the case of the continuous wavelet transform, the role of the extrema of the dyadic wavelet transform being played by the ridges of the continuous wavelet transform. Notice that the estimation of the ridge was taking into account the possible presence of noise while the reconstruction algorithm was assuming that the observations of the transform were correct. As in [4], the reconstruction which we present is based on a variational approach involving a penalty on the smoothness of the transform along the estimated ridges. But contrary to [4], the observations of the transform along the ridges are not brought into the problem in the form of "knowledge" - constraints. Instead they are used to define a second penality component. This form of the variational problem allows for a delicate balance between the fit of the transform of the solution to the observations and the smoothness of the modulus of the transform along the ridges. Moreover the generality of the present approach makes it easy to avoid penalizing a finite dimensional space of signals which one can choose apriori.

The purpose of this subsection is to derive the formulas for the reconstruction of the original signal from the observations of the values of the transform at sample points of time/scale or time/frequency plane. We use the notation $T_f(b,a)$ for the transform of the signal f. This notation stands for the wavelet transform $W_f(b,a)$ as well as for the Gabor transform $G_f(b,\omega)$.

We assume that we have observations $f(x)$ of an unknown signal $f_0(x)$ in the presence of an additive noise $\epsilon(x)$ with mean zero. In other words we work with the model:

$$f(x) = f_0(x) + \epsilon(x)$$

and we assume that the noise is given by a mean zero stationary process with covariance:

$$\mathbb{E}\{\epsilon(x)\epsilon(y)\} = \Gamma(x-y).$$

The case $\Gamma = I$ (i.e. $\Gamma(x-y) = \delta(x-y)$) corresponds to the case of an additive white noise. We now transform the observations, estimate the ridges of the transform and sample these estimates. We end up with a discrete set $\{(b_1,a_1),\cdots,(b_n,a_n)$ in the time/scale plane and observations $T_f(b_j,a_j)$ of the transform of the unknown signal at these points. We assume that the observations follow the usual linear model:

$$z_j = T_f(b_j,a_j) + \epsilon'_j$$

where the computational noise terms ϵ'_j are assumed to be identically distributed and uncorrelated between themselves and with the observation noise terms $\epsilon(x)$. Hence the final model is of the form:

$$z_j = L_j f_0 + \epsilon_j, \qquad j = 1, \cdots, n \tag{29}$$

where L_j is the linear form representing the value of the transform at the point (b_j, a_j) and where:

$$\epsilon_j = T_\epsilon(b_j, a_j) + \epsilon'_j.$$

The assumption that the two sources of noise are uncorrelated implies that the covariance matrix Σ of the ϵ_j is the sum of the covariance of the $T_\epsilon(b_j, a_j)$ and the covariance of the ϵ'_j. The latter being of the form $\sigma'^2 I$ we have:

$$\Sigma = \sigma'^2 I + \Sigma^{(1)}$$

where the entries of the matrix $\Sigma^{(1)}$ are given by the formula:

$$\Sigma^{(1)}_{j,k} = \frac{1}{a_j a_k} \int \psi\left(\frac{x - b_j}{a_j}\right) \Gamma(x - y) \psi\left(\frac{y - b_k}{a_k}\right) \, dx dy.$$

The reconstruction algorithm is formulated as the solution of the minimization problem:

$$\min_f \; \frac{1}{n} \|\Sigma^{-1/2}(Z - T_f(\cdot, \cdot)\|^2 + \lambda \langle Qf, f\rangle \tag{30}$$

where Z denotes the n-vector of observations z_j and $T_f(\cdot, \cdot)$ denotes the n-vector of values of the transform of the candidate function f at the points $(b - j, a_j)$ and where the constant $\lambda > 0$ is introduced to balance the effects of the two components of the penalty. Theorem 1.3.1 of [22] implies that the solution is given by the formula:

$$\hat{f}_\lambda(x) = \sum_{j=1}^{n} \lambda_j \tilde{\psi}_j(x) \tag{31}$$

where the dual wavelets/atoms are defined by:

$$\tilde{\psi}_j = Q^{-1} \psi_j$$

where we used the notation ψ_j for the wavelet/atom centered around the point (b_j, a_j) of the time/scale or time/frequency plane. The coefficients λ_j of the linear combination (31) are given by the formula:

$$\Lambda = A^{-1}\Sigma^{-1/2}Z \tag{32}$$

and the matrix A is defined as $A = n\lambda I + \tilde{\Sigma}$ where the matrix $\tilde{\Sigma}$ is defined by $\tilde{\Sigma}_{j,k} = \langle\tilde{\psi}_j, \psi_k\rangle$.

Remarks:

• Notice that we did not use the full generality of the smoothing spline problem as defined in [22]. Indeed, we could have chosen a quadratic penalty of the form $\|Q^{1/2}P_1 f\|^2$ where P_1 is the projection onto the orthogonal complement of a subspace of finite dimension. In this generality it is possible to avoid penalizing special subspaces of functions (for example, the space of polynomial functions of degree smaller than a fixed number, ...). Since the form of the solution is much more involved and since we did not find an application to this level of generality, we decided to use the smoothing spline approach in our simpler context. The reader interested in this specific feature of the smoothing splines technique can consult [22].

• The approach presented here was alluded to as a possible extension to the reconstruction algorithm derived and used in [2] and [4]. The latter corresponds to the case where the knowledge of the wavelet transform of the unknown signal is assumed to be perfect. In other words to the case where both Γ and σ'^2 are assumed to be zero. It is easy to see that, under these extra assumptions, the reconstruction procedure given by the above minimization problem reduces to the minimization of the quadratic form $\langle Qf, f\rangle$ under the constraints (18). This is the problem which was solved in [2] and [4]. It appears as a particular case of the more general procedure presented here. The advantages of the latter were explained in the introduction. We shall not reproduce this discussion here.

• Notice that the reconstructed signal appears as a linear function of the observations. Nevertheless, our whole analysis is nonlinear because of the ridge estimation and the sampling of the latter.

4.6 Supplementary Remarks Concerning the Numerical Implementation

• Because the computation time needed to solve a linear system grows like the cube of the number of equation, it is important to find ways to speed up the computations. In this respect, the following simple remark is very important. For the sake of the present discussion let us say that two elementary ridges R_ℓ and $R_{\ell'}$ are well separated if the supports of the elementary wavelets $\psi_{(b,a)}$ (resp. the Gabor atoms $g_{(b,\omega)}$) centered on the first ridge R_ℓ do not intersect the supports of the $\psi_{(b,a)}$ (resp. $g_{(b,\omega)}$) centered on the second ridge $R_{\ell'}$. Our interest in this notion is the following. When individual ridges are well separated, their contributions to the reconstructions are nonoverlapping and consequently, they can be separated and computed by solving systems of smaller orders. This simple remark can significantly reduce the computing time of the reconstructions.

• In practical applications the exact values the Gabor transform at the sample points of the ridges are not known exactly. The $g_{\ell,j}$'s are merely noisy perturbations of the true values. The reconstruction problem can be solved in a very similar way. See for example [2] and [22] for the modifications which are necessary in order to get the solution of this seemingly more general problem.

4.7 An Example

Let us return to the wavelet analysis of the bat signal with the additional chirp. We used $L = 3$ ridges, say R_1, R_2 and R_3, and we chose on each ridge estimate a number of samples proportional to the length of the ridge and inversely proportional to the corresponding scale, according to the sampling theory of wavelet transforms, see [6].

We used the value $\epsilon = .5$ to reconstruct the signal. The result of the reconstruction is given in the second part of Figure 3. The last two plots of Figure 3 give the reconstructions of the two components: the bat signal, reconstructed from two ridges (to be compared with the top plot in Figure 2), and the chirp (the original chirp and the reconstructed one are displayed on the same plot, bottom of Figure 3). As we can see, the agreement is very good (except at the end of the chirp, where the ridge was a bit smaller than the true signal. In addition, we stress that the number of coefficients needed to characterize such a signal (i.e. twice the number of complex constraints) was approximatly one fifth of the number of samples. Although compression was not our goal, the method seems to have an interesting potential in such a direction.

5 Ridges and the Sinusoidal Model for Speech Signal

A popular representation of speech signals is to regard the signal as the output of a slowly time-varying filter excited by a glottal waveform. Here the filters model the resonant characteristics of the vocal tract. We won't go into details of speech modeling here (we refer to [13] and [15] for a detailed presentation), but we notice that the resulting model for speech signal is of the form given in equation (8). Hence it is natural to use a time-frequency representation in order to separate the components of the signal and express them separately. Since those components are close to harmonic components, the Gabor transform is better suited than the wavelet transform for the description of the speech signals which we want to consider. Indeed, since the wavelet processing may be viewed as a filter bank of constant *relative* frequency, it is not able to separate the high frequencies components (see nevertheless [15] for a method to separate the first low frequency components).

Hence, we use the continuous Gabor transform with a Gaussian window of length approximately equal to 160 ms. Our approach is, at least in spirit, similar to that of [13]. However there is a major difference: since the detection algorithm described in section 2 returns ridges, i.e. one-dimensional structures, the chaining method necessited by the Mc Aulay-Quatieri approach is not needed here.

We illustrate this discussion on the example of the /one/ signal displayed at the top of Figure 6. Our results were obtained using approximately 200 ridge samples, i.e. 400 real constraints, while the signal's length is 2048. As can be seen on the top two plots of the figure, the reconstructed signal is very close to the true one. Of course such a comparison is not significant from the speech processing point of view. However, we stress the fact that the main features of the signal are preserved (in particular the pitch).

A better comparison is obtained by listening the two sounds; they turn out to be almost undistinguishable.

6 Conclusions

We presented a new technique to detect the ridges in the Gabor transform of speech signals. This algorithm is based on the stochastic relaxation of a particle system of a new type. We showed that its realm of application is not limited to the examples of the paper. This detection technique performs extremely well, especially at very low SNR's. In particular, it can be used to detect ridges in all the energetic distribution representations of a signal. We also adapted the reconstruction procedure introduced earlier in the case of the wavelet transform to the case of the Gabor transform and we showed that it was performing very well on speech signals, even in the presence of significant noise disturbances.

The most important extension to the results presented in this paper would be a real time implementation. It is relatively easy to find approximations of the reconstruction procedure which would be amenable to on line implementations. It seems more difficult to modify the ridge detection algorithm to accommodate frequent updates. We are currently working on the design of such algorithms.

Acknowledgements: This work was done while the third named author (B.T.) was visiting the Department of Mathematics of the University of California at Irvine. Its warm hospitality should be acknowledged.

References

[1] M.-O. Berger (1993): Towards Dynamic Adaptation of Snake Contours. In *International Conference on Image Analysis and Processing, Como (Italy).* IAPR, September 1991.

[2] R. Carmona (1992): Spline Smoothing & Extrema Representation: Variations on a Reconstruction Algorithm of Mallat and Zhong. to appear in the proceedings of the conference *Wavelets and Statistics*, Villard de Lans, France (1994), A. Antoniadis & G. Oppenheim Eds, Lecture Notes in Statistics.

[3] R. Carmona (1993): Wavelet Identification of Transients in Noisy Signals. Proc. SPIE 15-16 June 1993, San Diego, California. Mathematical Imaging: Wavelet Applications in Signal and Image Processing. 392-400.

[4] R. Carmona, W.L. Hwang and B. Torrésani (1994): Characterization of Signals by the Ridges of their Wavelet Transforms. (preprint)

[5] R. Carmona, W.L. Hwang and B. Torrésani (1994): Identification of Chirps with Continuous Wavelet Transform. Preprint CPT.93/P.3163, to appear in the proceedings of the conference *Wavelets and Statistics*, Villard de Lans, France (1994), A. Antoniadis & G. Oppenheim Eds, Lecture Notes in Statistics.

[6] I. Daubechies (1992): Ten Lectures on Wavelets. *SIAM.*

[7] 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****, special issue on Wavelet and Multiresolution Analysis* 644-664.

[8] P. Flandrin (1993): Temps-Fréquence. *Traité des Nouvelles Technologies, série Traitement du Signal, Hermès.*

[9] S. Geman and D. Geman (1984): Stochastic Relaxation, Gibbs Distributions and Bayesian Restoration of Images. *IEEE Proc. Pattern Ana. Mach. Intell.* ***6***, 721-741.

[10] B. Gidas (1985): Nonstationary Markov Chains and Convergence of the Annealing Algorithm. *J. Statist. Phys.* ***39***, 73-131.

[11] P. Hall, W. Qian and D.M. Titterington (1992): Ridge Finding from Noisy Data. *J. Comput. and Graph. Statist.* ***1***, 197-211.

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

[13] 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.

[14] R.J. McAulay and T.F. Quatieri (1992): Low Rate Speech Coding Based on the Sinusoidal Model. In *Advances in Speech Signal Processing*, Edited by S. Furui and M. Mohan Sondui.

[15] S. Maes (1994): Dissertation, Rutgers University.

[16] S. Mallat and W.L. Hwang (1992): Singularities Detection and Processing with Wavelets. *IEEE Trans. Info. Theory* ***38#2***, 617–643.

[17] S. Mallat and Zhang (1993): Matching Pursuit with Time-Frequency Dictionaries, preprint.

[18] S. Mallat and S. Zhong (1992): Characterization of Signals from Multiscale Edges. *IEEE Trans. Pattern Anal. Machine Intel.* ***14***, 710-732.

[19] T.F. Quatieri and R.J. McAulay (1989): Phase coherence in speech reconstruction for enhancement and coding applications. *Proc. IEEE Int. Conf. Audio, Speech and Sig. Proc.*, Glasgow, 207-209.

[20] Ph. Tchamitchian and B. Torrésani (1992): Ridge and Skeleton Extraction from the Wavelet Transform. In *Wavelets and Applications, M.B. Ruskai & al Eds, Jones & Bartlett Pub. Comp. Boston,* 123–151.

[21] B. Torrésani (1992): Time-Frequency Distributions: Wavelet Packets and Optimal Decompositions. *Annales de l'Institut Henri Poincaré* ***56#2***, 215–234.

[22] G. Wahba (1988): Spline Models for Observational Data. CBMS-NSF Reg. Conf. Ser. in Applied Math. # 59. *SIAM.*

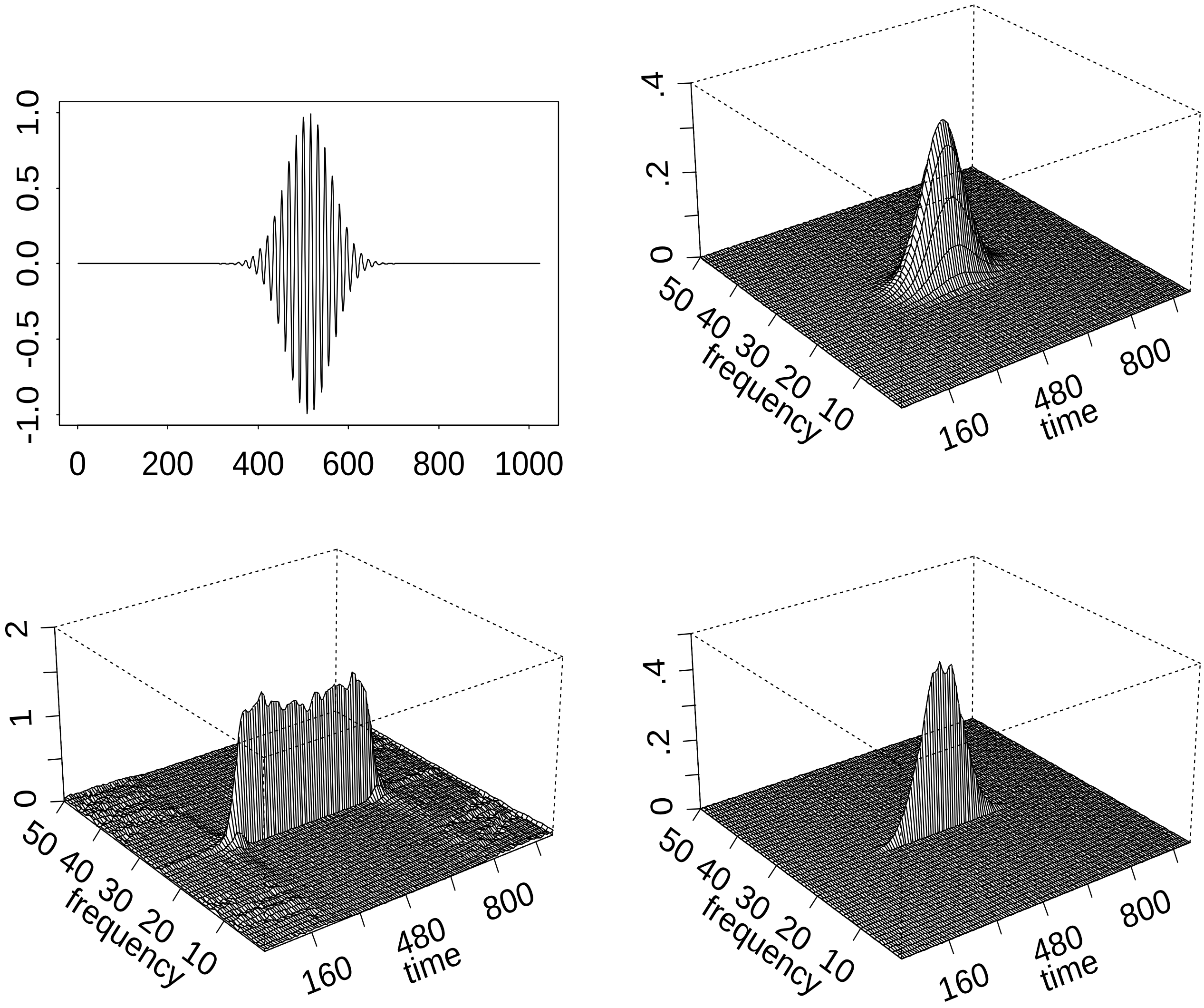

Figure 1: The occupation measures for a simple windowed sine wave: left top, the signal; right top: its Gabor transform modulus; left bottom: unweighted occupation measure; right bottom: weighted occupation measure.

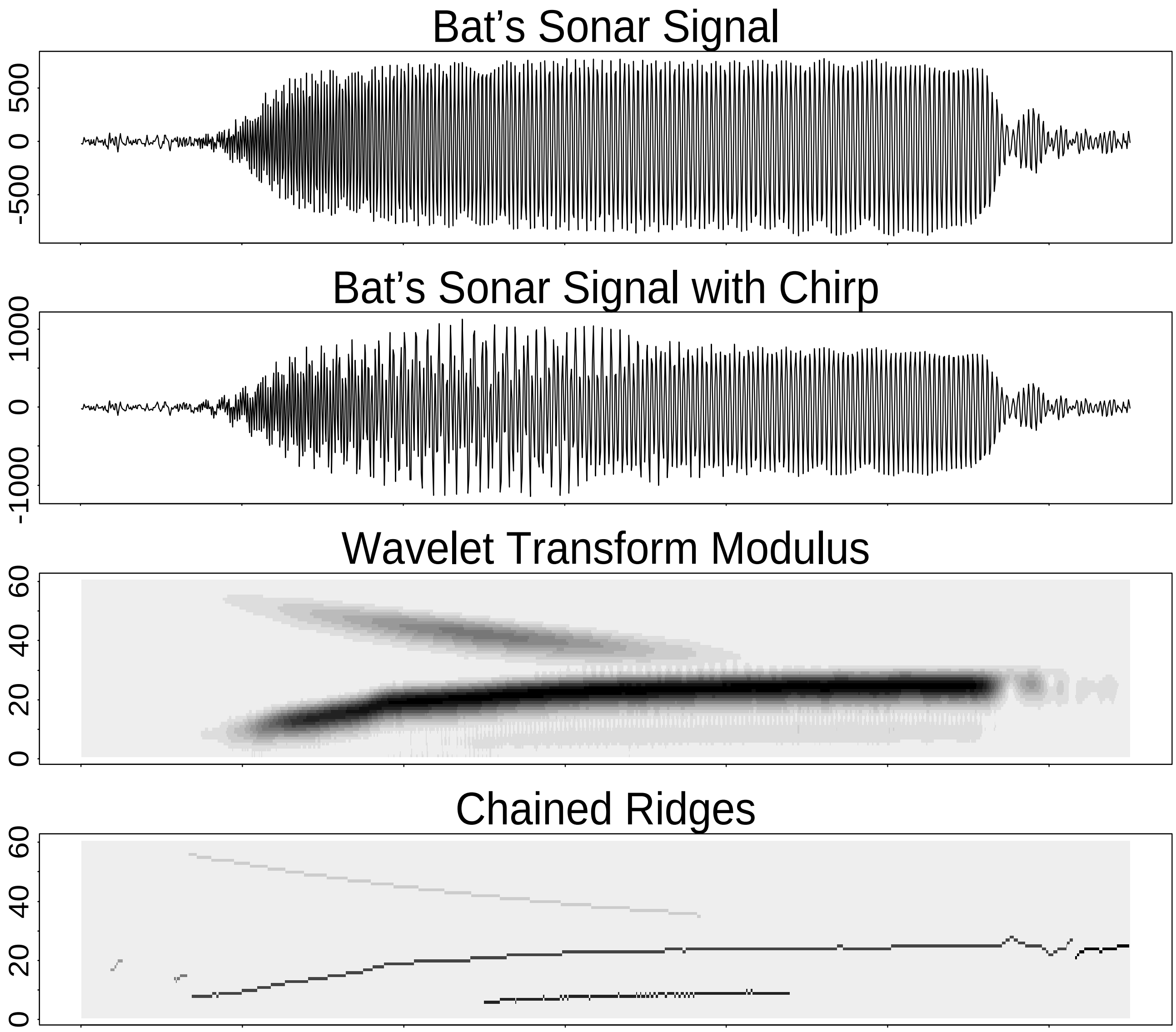

Figure 2: Bat sonar signal with an additional chirp.

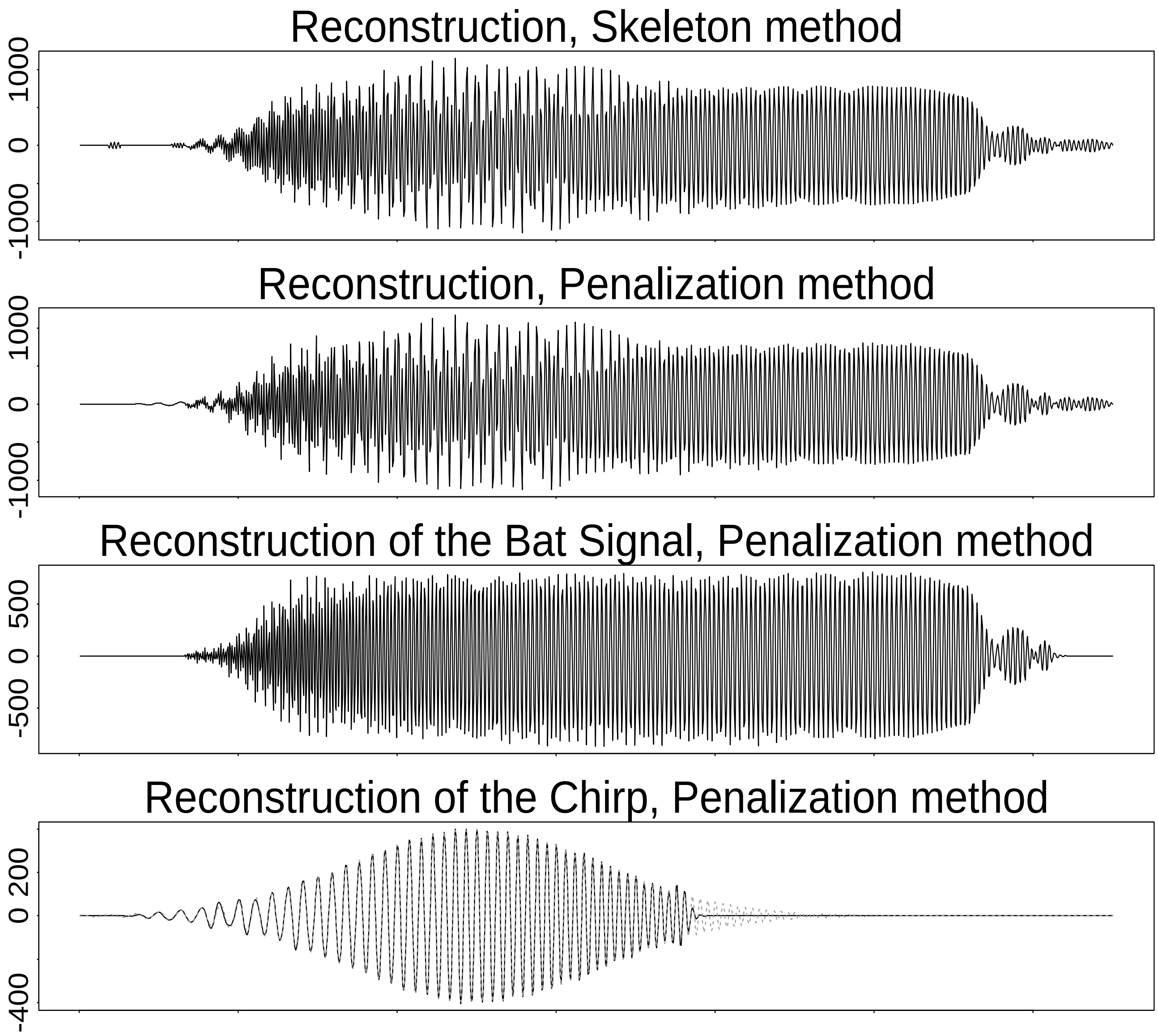

Figure 3: Reconstruction from the ridges: last plot: full curve: reconstructed chirp; dashed curve: the original chirp.

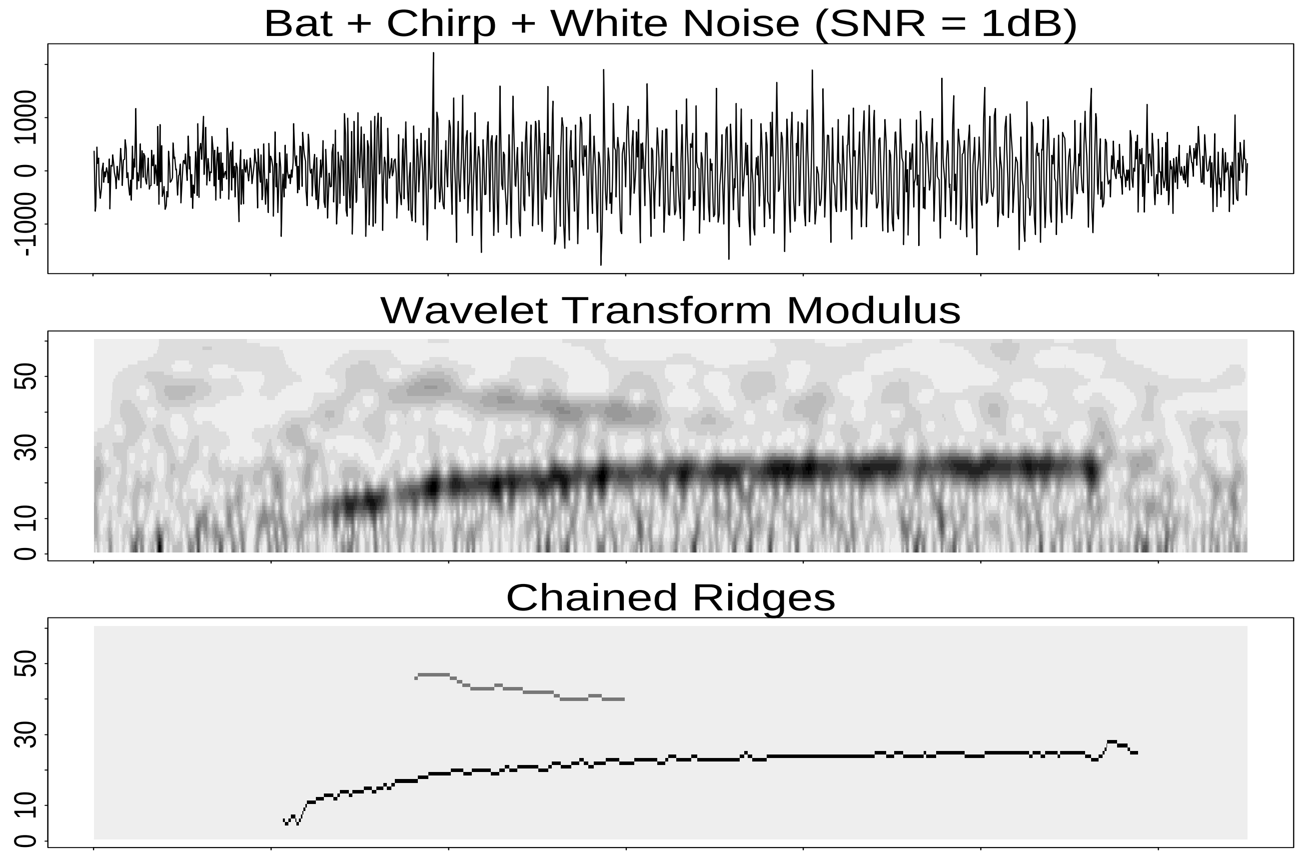

Figure 4: Bat's sonar signal and chirp embedded into a Gaussian white noise (Signal to Noise Ratio: 1dB)

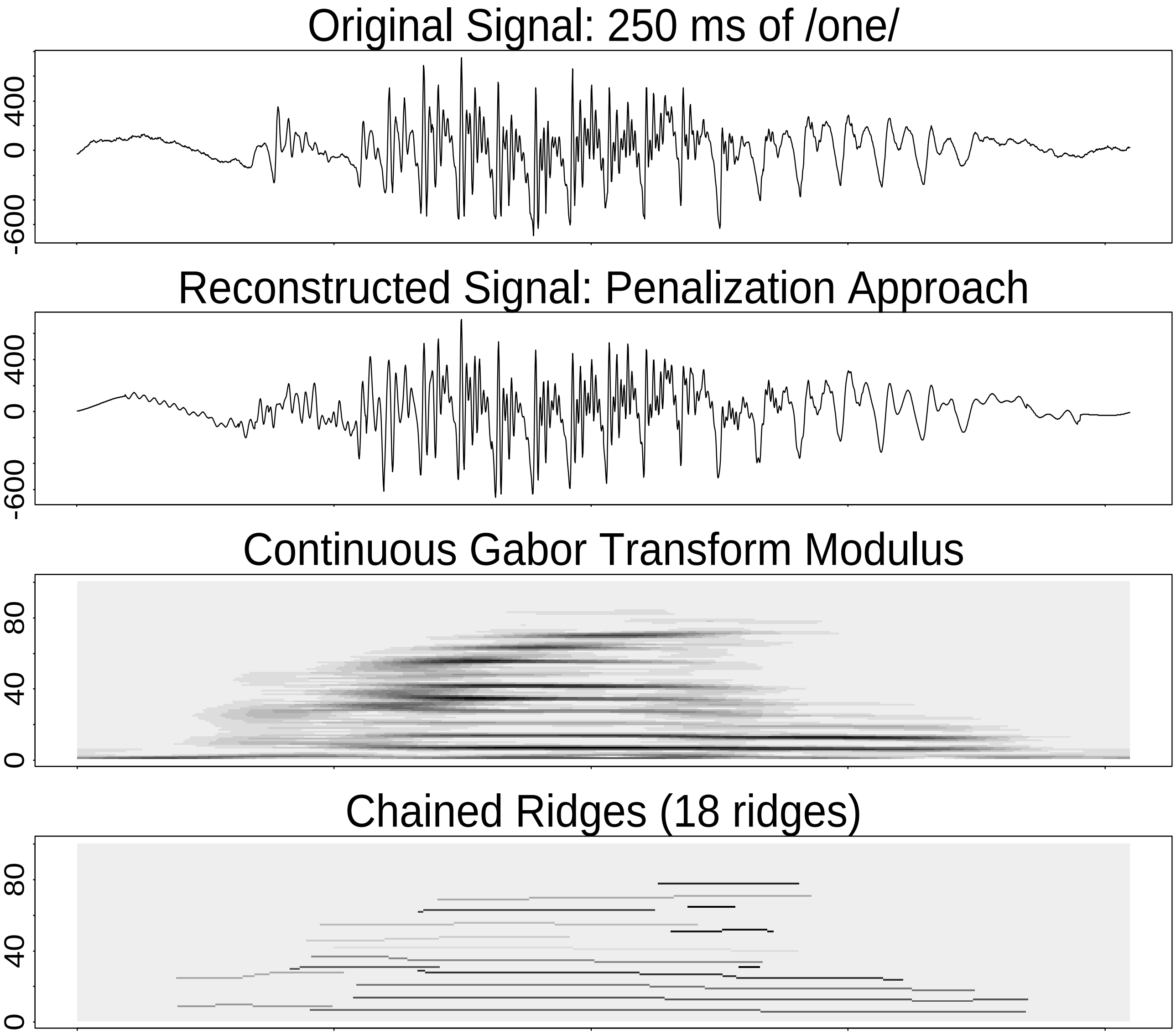

Figure 5: 250 ms of the speech signal /one/ (sampling frequency: 8 kHz)