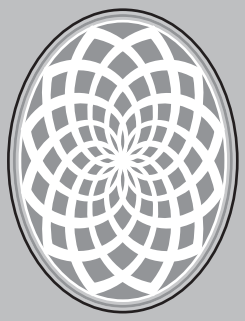

Industrial Mathematics Institute

2000:02

Fast ridge pursuit with a multiscale dictionary of Gaussian chirps

R. Gribonval

IMI

Preprint Series

Department of Mathematics
University of South Carolina

Fast Ridge Pursuit with a multiscale dictionary of Gaussian chirps

R. Gribonval*, Centre de Mathématiques Appliquées,
École Polytechnique, France

Abstract

We introduce a modified Matching Pursuit algorithm, called Fast Ridge Pursuit, to approximate N-dimensional signals with M Gaussian chirps at a computational cost $\mathcal{O}(MN)$ instead of the expected $\mathcal{O}(MN^2 \log N)$. At each iteration of the pursuit, the best Gabor atom is first selected, then its scale and chirp rate are locally optimized so as to get a good chirp atom. A ridge theorem of the Gaussian chirp dictionary is proved, from which an estimate of the locally optimal scale and chirp is built. The procedure is restricted to a sub-dictionary of local maxima of the Gaussian Gabor dictionary, so as to accelerate further the pursuit. The efficiency and speed of the method is demonstrated on a sound signal.
EDICS : 2-TIFR

1 Introduction

There has been a considerable interest last decade in developing analysis techniques to decompose non-stationary signals into elementary components, called *atoms*, that characterize their salient features. As many signals display both oscillatory phenomena, which time-frequency methods can extract, and transients or singularities, to which time-scale techniques [9, 19, 4] are better adapted [15, 21, 23], adaptive decompositions were developed, using *redundant* families of atoms that can characterize independently scale and frequency [7, 6, 31, 29].

Chirp atoms were introduced to deal with the non-stationary behavior of the instantaneous frequency of some signals [17]. Baraniuk and Jones [1] built orthonormal bases and frames of such chirp atoms, while Mann and Haykin [24] defined a "Chirplet Transform". Roughly speaking, this transform compares a signal $x(t)$ with each chirp atom

$$g_{(s,u,\xi,c)}(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\left(\xi(t-u)+\frac{c}{2}(t-u)^2\right)} \quad (1)$$

of a large family, the *chirp dictionary* $\mathcal{D}^+$, which is an extension of the Gabor multiscale time-frequency dictionary $\mathcal{D}$ [31] [29]. These atoms are characterized by their scale s, time u, frequency ξ and *chirp rate* c. Their instantaneous frequency $\omega(t) = \xi + c(t-u)$ varies linearly with time.

In an orthonormal basis of chirp atoms [1], a given signal can be efficiently decomposed into elementary chirps. However the elementary atoms are somehow too "rigid" for many applications, as their parameters s, $\xi \propto 1/s$ and $c \propto 1/s^2$ are not independent one from another. On the other hand, the chirplet transform is very redundant and does not have this intrinsic rigidity. It can thus provide with a large variety of viewpoints to look at the signal, in order to find meaningful structures in it. However its redundancy is also its weakness : the collection $\mathcal{D}^+$ is no longer a frame (hence there is no inverse chirplet transform to reconstruct the signal) and the computational complexity of the chirplet transform is very large.

Bultan [5] suggested to use the Matching Pursuit algorithm of Mallat and Zhang [22] to decompose a signal into elementary chirp atoms. He demonstrated the interest of this technique, but its practical use was limited by the large computational complexity $\mathcal{O}(MN^2 \log N)$ needed to get an M-term approximation of an N-sample signal. In order to limit the complexity, Bultan suggested to reduce the size of the dictionary by limiting the resolution of the chirp rate.

In this work we show that it is possible to get rid of such a limitation *and* get a low complexity $\mathcal{O}(MN)$ by modifying the underlying "Matching Pursuit" algorithm and using a *Gaussian* chirp dictionary. To get such a low complexity we introduce a (substantially) modified pursuit algorithm, by using some *ridge* techniques and the *local maxima* of the Gabor dictionary.

*currently in the Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA. e-mail : remi@math.sc.edu, phone : (803) 777-5299, fax : (803) 777-3783

The paper is organized as follows. In the next section, we review the definition of the multiscale time-frequency chirp dictionary $\mathcal{D}^+$ and show the numerical complexity implied by its very large size. In section 3 the definition and basic properties of the Matching Pursuit are recalled. Section 4 is devoted to the detailed study of the ridges of the Gaussian multiscale Gabor dictionary for analytic signals. We use those results to analyze the selection of the *locally optimal chirp atom*. In section 5 we summarize the Ridge Pursuit algorithm with the real-valued chirp dictionary, and show how it can be further accelerated with a sub-dictionary technique. Finally we analyze in section 6 the numerical results obtained with our new algorithm on an acoustic signal.

2 Multiscale Dictionary of Time-frequency chirp atoms

Every chirp atom (1) is obtained from an elementary window $g(t)$ by dilation, translation, frequency and chirp modulation. It can thus be described with its index (s, u, ξ, c). The window $g(t)$ is localized around 0 both in the time domain and the frequency domain. As a result $g_{(s,u,\xi,c)}$ is localized at time u with a temporal dispersion proportional to its scale s. The Wigner-Ville transform [12, 20] of a chirp atom defines a quadratic time-frequency energy distribution. It is localized around the line of instantaneous frequency $\omega = \xi + c(t-u)$. Its dispersion is proportional to $1/s$ in the ω direction. A Gaussian chirp atom is built from the unit Gaussian window $g(t) = \pi^{-1/4} e^{-t^2/2}$. Such an atom is displayed on figure 1 with its Wigner-Ville distribution.

Sampling the dictionary

The set $\left\{g_{(s,u,\xi,0)}, (s,u,\xi) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}\right\}$ of chirp atoms with chirp rate $c = 0$ is exactly the multiscale Gabor dictionary [31, 22, 29]. The discrete Gabor dictionary $\mathcal{D}$ is the collection of atoms $g_{(s,u,\xi,0)}$ (denoted for short by $g_{(s,u,\xi)}$) such that $(s, u, \xi) = (a^j, na^j\Delta u, ka^{-j}\Delta\xi), j, n, k \in \mathbb{Z}$, where Δu and $\Delta\xi$ are some constants. Watson and Gilholm [32] showed that this sampling of the scale, time and frequency parameters is uniform with respect to the natural Riemannian metric of the continuous dictionary induced by $d(g_{\gamma_1}, g_{\gamma_2}) = 1 - |\langle g_{\gamma_1}, g_{\gamma_2}\rangle|$. The same point of view leads to sampling the chirp rate as $c = la^{-2j}\Delta c$, $l \in \mathbb{Z}$. The discrete chirp dictionary $\mathcal{D}^+$ is thus the family of atoms $g_{(s,u,\xi,c)}$ such that

$$(s, u, \xi, c) \in \Gamma^+ = \left\{(a^j, na^j\Delta u, ka^{-j}\Delta\xi, la^{-2j}\Delta c), \; j, n, k, l \in \mathbb{Z}\right\}. \quad (2)$$

As the set of atoms at a given scale j and chirp rate l is a Weyl-Heisenberg frame, it can only span $\boldsymbol{L}^2(\mathbb{R})$ if $\Delta u \times \Delta\xi \leq 2\pi$ [10]. When $\Delta u \times \Delta\xi < 2\pi$, $\mathcal{D}$ is complete [22], thus $\mathcal{D}^+ \supset \mathcal{D}$ is also complete.

Size of the discrete chirp dictionary

The size of the discrete chirp dictionary $\mathcal{D}^+$ is a function of the sampling steps $a, \Delta u$, $\Delta\xi$ and Δc. When analyzing a discrete N point signal, one also has to consider the limitations of the sampling rate and the signal size. The scale a^j can thus only vary between 1 and N, which makes a total of $\mathcal{O}(\log N)$ scales. At each scale, there are $\mathcal{O}(N)$ sampled values of $(u, \xi) \in [0, N-1] \times [-\pi, \pi]$. Because of the Nyquist condition, the instantaneous frequency is constrained to $\xi + c(t-u) \in [-\pi, \pi]$, $\forall t \in [u - a^j/2, u + a^j/2]$, *i.e.* $|c|\, a^j/2 \leq \pi - |\xi|$. For given a^j and ξ, the chirp rate c can take $\mathcal{O}((\pi - |\xi|)a^j)$ values. On the average, at scale a^j, it thus takes $\mathcal{O}(a^j)$ distinct values. The total number of chirp atoms in the discrete chirp dictionary $\mathcal{D}^+$ is thus of the order of $\sum_{j=0}^{\log N} \mathcal{O}(N/a^j) \times \mathcal{O}(a^j) \times \mathcal{O}(a^j) = \mathcal{O}(N^2)$.

3 Standard Matching Pursuit with $\mathcal{D}^+$

The Matching Pursuit [22] is a greedy strategy to decompose a signal x into a linear combination of atoms chosen among a dictionary $\mathcal{D} = \{g_\gamma, \gamma \in \Gamma\}$, *i.e.* a redundant family of unit vectors in a Hilbert space $\boldsymbol{H}$. It iteratively defines an m^{th}-order residual $R^{m-1}x$ (starting with $R^0 x = x$) in the following way :

1. Compute $\left|\left\langle R^{m-1}x, g_\gamma\right\rangle\right|^2$ for all $\gamma \in \Gamma$.
2. Select the best atom of the dictionary
$$g_{\gamma_m} = \arg\max_{\gamma\in\Gamma} \left|\left\langle R^{m-1}x, g_\gamma\right\rangle\right|^2. \quad (3)$$
3. Compute the new residual by removing the component along the selected atom
$$R^m x = R^{m-1}x - \left\langle R^{m-1}x, g_{\gamma_m}\right\rangle g_{\gamma_m}. \quad (4)$$

After M iterations, one gets an M-term approximation $x_M = x - R^M x = \sum_{m=1}^{M} \left\langle R^{m-1}x, g_{\gamma_m}\right\rangle g_{\gamma_m}$. The energy is split among the selected components

as $\|x\|^2 = \sum_{m=1}^{M} \left|\left\langle R^{m-1}x, g_{\gamma_m}\right\rangle\right|^2 + \left\|R^M x\right\|^2$. The Matching Pursuit is very similar to the Projection Pursuit principle introduced in statistics by Huber [14], which strong convergence $\lim_{M\to\infty} \left\|R^M x\right\| = 0$ was proved by Jones [16] whenever the dictionary $\mathcal{D}$ is complete, *i.e.* $\overline{\text{span}}(\mathcal{D}) = \boldsymbol{H}$.

With the chirp dictionary $\mathcal{D}^+$ and a N-point signal, the computation of $\left\langle R^{m-1}x, g_\gamma\right\rangle, g_\gamma \in \mathcal{D}^+$ can be done with $\mathcal{O}(N^2 \log N)$ operations, using FFT-based algorithms with appropriate windows [5, 32]. The search for the "best" atom (3) costs $\mathcal{O}(N^2)$, and the update of the residual (4) only costs $\mathcal{O}(N)$, hence the total complexity $\mathcal{O}(MN^2 \log N)$ of M iterations of pursuit with the chirp dictionary. Such a "brute force" Chirp Matching Pursuit is thus limited to the analysis of small signals, with only few iterations.

4 *Ridge* Pursuit

Because of the large size of $\mathcal{D}^+$, one cannot afford to compute the correlation $\left\langle R^{m-1}x, g_{(s,u,\xi,c)}\right\rangle$ of the residual with *every* atom of $\mathcal{D}^+$. As a consequence, the choice of the "best" atom $g_{(s_m,u_m,\xi_m,c_m)}$ must be done in an *approximate* way. In other words, one needs to "guess" where a "good" chirp atom is located, without scanning the whole dictionary.

One can notice that the chirp dictionary $\mathcal{D}^+$ is only an extension of the Gabor dictionary $\mathcal{D}$. As $\mathcal{D}$ is complete, the set of inner products $\left\langle R^{m-1}x, g_\gamma\right\rangle$, $g_\gamma \in \mathcal{D}$ contains all the information available about $R^{m-1}x$. It is thus theoretically sufficient to compute these inner products to select the best chirp atom. We will actually show, with theorem 1, that the behavior of $\left\langle R^{m-1}x, g_\gamma\right\rangle$ in the neighborhood of the best Gabor atom contains enough information to select a "locally optimal" chirp atom. A "good" chirp atom $g_{(s_m,u_m,\xi_m,c_m)}$ is selected with a two-step pursuit. First, one selects the best Gabor atom

$$g_{(s_m^\star,u_m^\star,\xi_m^\star)} \stackrel{\Delta}{=} \arg \max_{g_{(s,u,\xi)}\in\mathcal{D}} \left|\left\langle R^{m-1}x, g_{(s,u,\xi)}\right\rangle\right|. \quad (5)$$

Then one explores its neighborhood in $\mathcal{D}^+$, so as to find a good chirp atom

$$g_{(s_m,u_m,\xi_m,c_m)} \stackrel{\Delta}{=} \arg \max_{g_{(s,u_m^\star,\xi_m^\star,c)}} \left|\left\langle R^{m-1}x, g_{(s,u_m^\star,\xi_m^\star,c)}\right\rangle\right| \quad (6)$$

by selecting locally optimal chirp rate and scale parameters s_m and c_m. The time and frequency parameters $u_m^\star$ and $\xi_m^\star$ are kept constant.

One can see that after selecting the best Gabor atom (5), the second step (6) implies an *exhaustive scanning* of the neighborhood of this atom. However this scanning is still very costly. We replace it by a *fast estimation* $\mathcal{O}(1)$ of s_m and c_m, using again theorem 1, which helps us extract the information we need from the local behavior of $(s,u,\xi,c) \mapsto \left\langle R^{m-1}x, g_{(s,u,\xi,c)}\right\rangle$ in the neighborhood of the best Gabor atom. We hereby define a *Ridge Pursuit*, which complexity $\mathcal{O}(MN \log^2 N)$ is identical to that of the standard Matching Pursuit with the Gabor dictionary $\mathcal{D}$.

Ridges of the Gaussian chirp dictionary

From now on, we consider the model of analytic signal $R^{m-1}x(t) = a(t)e^{i\phi(t)}$ where the instantaneous frequency $\xi(t) = \phi'(t)$ and instantaneous chirp $c(t) = \phi''(t)$ are defined. Our results can be extended to the case of a superposition of finitely many such analytic signals, provided a sufficient separation of their instantaneous frequencies is granted.

The goal of the following ridge theorem (which is proved in appendix A) is to show that, under certain regularity conditions, the residual $R^{m-1}x$, seen "through" a Gaussian chirp atom $g_{(s,u,\xi,c)}$, looks like another Gaussian chirp atom $g_{\gamma^+(u)} \in \mathcal{D}^+$, *i.e.* $\left\langle R^{m-1}x, g_{(s,u,\xi,c)}\right\rangle \approx A(u)e^{i\Phi(u)} \left\langle g_{\gamma^+(u)}, g_{(s,u,\xi,c)}\right\rangle$.

Theorem 1 *Let $R^{m-1}x(t) = a(t)e^{i\phi(t)}$. Suppose that $\|a\|_\infty < \infty$, $\|\phi'''\|_\infty < \infty$ and $\|\alpha'''\|_\infty < \infty$, with $\alpha(t) \stackrel{\Delta}{=} -\log a(t)$. Let u a time where $\alpha''(u) > 0$, and let $g_{(s,u,\xi,c)}$ a* Gaussian *chirp atom. Then*

$$\begin{aligned}\left\langle R^{m-1}x, g_{(s,u,\xi,c)}\right\rangle &= \frac{ae^{i\phi}}{(\alpha''/\pi)^{1/4}} e^{-i\phi'\frac{\alpha'}{\alpha''}+i\frac{\phi''}{2}\left(\frac{\alpha'}{\alpha''}\right)^2} \\ &\quad e^{\frac{(\alpha')^2}{2\alpha''}} \left(\left\langle g_{\gamma^+(u)}, g_{(s,u,\xi,c)}\right\rangle \right. \\ &\quad \left. +\epsilon(s,u,\xi,c)\right) \quad (7)\end{aligned}$$

where every function $(a, \alpha', \alpha'', \phi, \phi', \phi'')$ is evaluated at u,

$$\gamma^+(u) = \left(\frac{1}{\sqrt{\alpha''}}, u - \frac{\alpha'}{\alpha''}, \phi' - \phi''\frac{\alpha'}{\alpha''}, \phi''\right) \quad (8)$$

and $|\epsilon(s,u,\xi,c)|$ is bounded by

$$\begin{aligned}\epsilon_{max}(s,u) &= (\alpha'')^{1/4}\pi^{-1/2}s^{+1/2}\left(\frac{Ks^3\sigma_3^3}{6}e^{1/6}\right. \\ &\quad \left. +\frac{4\|a\|_\infty}{a(u)}(Ks^3)^{1/3}e^{-\frac{1}{2(Ks^3)^{2/3}}}\right) \quad (9)\end{aligned}$$

with $K \stackrel{\Delta}{=} \|\alpha'''\|_\infty + \|\phi'''\|_\infty$ and $\sigma_3^3 \stackrel{\Delta}{=} \int |t|^3 g(t)dt$.

The hypothesis $\alpha''(u) > 0$ simply corresponds to $a''(u)/a(u) < (a'(u)/a(u))^2$ which is true for instance when $a''(u) \le 0$, *i.e.* on any concave part of

the amplitude a, such as the neighborhood of smooth local maxima of $a(u)$.

From this theorem, one can observe that if

$$\|\alpha'''\|_\infty \ll |\alpha''(u)|^{3/2} \text{ and } \|\phi'''\|_\infty \ll |\alpha''(u)|^{3/2} \tag{10}$$

then $\epsilon_{max}(1/\sqrt{\alpha''(u)}, u) \ll 1$, so that the best chirp atom at time u is close to $g_{\gamma^+(u)}$. The locally optimal parameters s, ξ, c can thus be obtained by estimating the index $\gamma^+(u)$. Let us now study how much information the location of the best Gabor atom gives about $\gamma^+(u)$.

Scale and frequency of the best Gabor atom

In the following we suppose that ϵ_{max} can be neglected. As the best Gabor atom (5) is the absolute maximum of $\left|\left\langle R^{m-1}x, g_{(s,u,\xi)}\right\rangle\right|$, it is a local maximum along s and ξ. If we additionally suppose that $\alpha'' \ll 1$, then the right hand side in (7) becomes $a(u)e^{i\phi(u)}\left\langle e^{i(\phi'(u)(t-u)+\phi''(u)(t-u)^2/2)}, g_{(s,u,\xi)}\right\rangle$. As the dictionary is Gaussian, the inner product that appears in this approximant is the Fourier transform $\widehat{g}_{(s,0,0,\phi''(u))}(\xi - \phi'(u))$ of a Gaussian chirp atom, which analytic expression is known [28]. For a given u, its maximum (or *ridge*) along s and ξ is located at $\xi = \phi'(u)$ and $s = 1/\sqrt{|\phi''(u)|}$. Thus one has

$$\xi_m^\star \approx \phi'(u_m^\star) \text{ and } s_m^\star \approx 1/\sqrt{|\phi''(u_m^\star)|}. \tag{11}$$

Bounds on the error of these estimates can be found in [13]. It is well known that the ridges of the wavelet transform or of the windowed Fourier transform give the instantaneous frequency [11, 20] : this result shows that the ridges of the Gabor dictionary additionally provide with the instantaneous chirp rate. Now, it is sufficient that

$$\|\alpha'''\|_\infty \ll |\phi''(u)|^{3/2} \text{ and } \|\phi'''\|_\infty \ll |\phi''(u)|^{3/2} \tag{12}$$

to get $\epsilon_{max}(1/\sqrt{\phi''(u)}, u) \ll 1$ and control the location (11) of the best Gabor atom, which gives information on the locally optimal chirp rate $|c_m| \approx |\phi''(u_m^\star)| \approx (s_m^\star)^{-2}$. Unfortunately the estimate $\widehat{c_m} = \pm(s_m^\star)^{-2}$ is far from the ideal one. First, because one has to determine its sign by computing the two inner products $\left\langle R^{m-1}x, g_{(s_m^\star, u_m^\star, \xi_m^\star, \pm(s_m^\star)^{-2})}\right\rangle$. But also mainly because it is a very poor estimate when, as usual, the scale $s = 2^j$ is coarsely quantized. Thus this estimate is not sufficient to avoid the costly $\mathcal{O}(N^2)$ "scanning" of the possible chirp atoms $g_{(s,u_m^\star,\xi_m^\star,c)}$.

Fast local estimation of the best chirp atom

The *local behavior* of $\xi \mapsto \left\langle R^{m-1}x, g_{(s_m^\star, u_m^\star, \xi)}\right\rangle$ in the neighborhood of $\xi_m^\star$ conveys much more information about the locally optimal chirp atom $g_{\gamma^+(u)}$ than the *location* $(s_m^\star, u_m^\star, \xi_m^\star)$ of the best Gabor atom does. Indeed, if $\epsilon_{max}(s_m^\star, u_m^\star) \ll 1$, then from theorem 1, $\left\langle R^{m-1}x, g_{(s_m^\star, u_m^\star, \xi)}\right\rangle \approx K\left\langle g_{\gamma^+(u_m^\star)}, g_{(s_m^\star, u_m^\star, \xi)}\right\rangle$, where $K \in \mathbb{C}$ is some constant independent on ξ. Using the analytic expression of the inner product between two Gaussian chirp atoms [28], one can get the following spectral estimation [25, 26] of the parameters of $g_{\gamma^+(u)}$, which is proved in [13].

Proposition 1 *If* $\epsilon_{max}(s_m^\star, u_m^\star) \ll 1$, *then* $\left\langle R^{m-1}x, g_{(s_m^\star, u_m^\star, \xi)}\right\rangle \approx Ae^{i\Phi}$ *where* $\xi \mapsto \log A(\xi)$ *and* $\xi \mapsto \Phi(\xi)$ *are second order polynomials in* ξ *with*

$$\phi''(u_m^\star) = -\frac{\Phi''(\xi)}{((\log A)''(\xi))^2 + (\Phi''(\xi))^2} \tag{13}$$

and

$$\alpha''(u_m^\star) + \frac{1}{(s_m^\star)^2} = \frac{-\log A''(\xi)}{((\log A'')(\xi))^2 + (\Phi''(\xi))^2}. \tag{14}$$

Moreover the following bounds hold

$$|\Phi''(\xi)| \leq \frac{(s_m^\star)^2}{2} \tag{15}$$

$$0 > \log A''(\xi) \geq -(s_m^\star)^2. \tag{16}$$

One can easily estimate $(\log A)''(\xi)$ and $\Phi''(\xi)$ (which are independent of ξ) using only the local behavior of $\xi \mapsto \left\langle R^{m-1}x, g_{(s_m^\star, u_m^\star, \xi)}\right\rangle$ around the best Gabor atom. Then (15) and (16) are used to *test* the validity of the approximation $\epsilon_{max} \ll 1$. Whenever the test is negative, the Ridge Pursuit is conservative : it does not try to find a better chirp atom than the best Gabor atom, but instead keeps it as its "good chirp atom" and steps forward to the next iteration. In the case of a positive test, we will *assume* that the model is valid. Thanks to (13) and (14), the estimates of $(\log A)''(\xi)$ and $\Phi''(\xi)$ provide with estimates of $\phi''(u_m^\star)$ and $\alpha''(u_m^\star)$, *i.e.* an estimate of $\gamma^+(u)$. This estimate is now obtained without costly "scanning".

The definition of the Ridge Pursuit will be complete by showing how to efficiently estimate $(\log A)''(\xi)$ and $\Phi''(\xi)$.

Numerical estimation by linear interpolation

In order to get as local an estimation as possible, we estimate $(\log A)''(\xi)$ and $\Phi''(\xi)$ through a parabolic interpolation. We use three Gaussian Gabor atoms

$g_\varepsilon = g_{(s_m^\star, u_m^\star, \xi_m^\star + \varepsilon\Delta\xi/s_m^\star)}$, $\varepsilon \in \{-1, 0, +1\}$, of the discrete Gabor dictionary $\mathcal{D}$, and their inner products $\left\langle R^{m-1}x, g_\varepsilon \right\rangle = \tilde{A}_\varepsilon e^{i\tilde{\Phi}_\varepsilon}$. These inner products were already computed for the selection of the best Gabor atom.

The numerical parabolic interpolation of $\log \tilde{A}_\varepsilon$ (resp. $\tilde{\Phi}_\varepsilon$), taking into account the frequency bin size $\Delta\xi/s_m^\star$, leads to estimates

$$\Phi''(\xi) \approx \left(\tilde{\Phi}_{-1} - 2\tilde{\Phi}_0 + \tilde{\Phi}_1\right)\left(\frac{s_m^\star}{\Delta\xi}\right)^2 \quad (17)$$

$$(\log A)''(\xi) \approx \log\frac{\tilde{A}_{-1}\tilde{A}_1}{\tilde{A}_0^2}\left(\frac{s_m^\star}{\Delta\xi}\right)^2. \quad (18)$$

As $\tilde{\Phi}_\varepsilon$ is defined modulo 2π, the estimate of $\Phi''(\xi)$ is only defined modulo $2\pi(s_m^\star/\Delta\xi)^2$. However, thanks to condition (15), its only admissible value(s) lie within the interval $[-(s_m^\star)^2/2, +(s_m^\star)^2/2]$. In order to eliminate the ambiguity, it is necessary and sufficient to impose that the length of this interval is strictly less than $2\pi(s_m^\star/\Delta\xi)^2$, *i.e.* to choose $\Delta\xi < \sqrt{2\pi}$ in the definition of $\mathcal{D}$ and $\mathcal{D}^+$ (see (2)). Thus s_m and c_m are estimated at a cost $\mathcal{O}(1)$ from the inner products $\left\langle R^{m-1}x, g_\varepsilon \right\rangle$.

5 Fast Ridge Pursuit

For the analysis of real-valued signal, we do not make use of complex-valued atoms (1) but of real-valued ones. They are defined [22, 5] as

$$g_{(s,u,\xi,c,\phi)} = K_{(s,u,\xi,c,\phi)} g\left(\frac{t-u}{s}\right) \cos\left(\xi(t-u) + \frac{c}{2}(t-u)^2 + \phi\right) \quad (19)$$

with some normalizing constant $K_{(s,u,\xi,c,\phi)}$. Obviously $g_{(s,u,\xi,c,\phi)}$ lies in the two-dimensional subspace $\boldsymbol{V}_{(s,u,\xi,c)} = \mathrm{Span}\left\{g_{(s,u,\xi,c)}, \overline{g_{(s,u,\xi,c)}}\right\}$, and

$$\sup_\phi \left|\left\langle R^{m-1}x, g_{(s,u,\xi,c,\phi)}\right\rangle\right| = \left\| P_{\boldsymbol{V}_{(s,u,\xi,c)}} R^{m-1}x \right\|, \quad (20)$$

where $P_{\boldsymbol{V}}$ denotes the orthogonal projector onto the subspace $\boldsymbol{V}$. We show in appendix B that the right hand side of (20), as well as the corresponding optimal phase $\phi(s,u,\xi,c) = \arg\max_\phi \left|\left\langle R^{m-1}x, g_{(s,u,\xi,c,\phi)}\right\rangle\right|$, can be computed in $\mathcal{O}(1)$ from $\left\langle R^{m-1}x, g_{(s,u,\xi,c)}\right\rangle$.

Let us now summarize the Ridge Pursuit algorithm with real-valued Gaussian chirp atoms, and compute its numerical complexity. Each iteration is decomposed into a few steps.

Ridge Pursuit Algorithm

1. Compute $\left\langle R^{m-1}x, g_{(s,u,\xi)}\right\rangle$ for each complex Gaussian Gabor atom $[\mathcal{O}(N\log^2 N)]$.
2. Compute $\left\| P_{\boldsymbol{V}_{(s,u,\xi)}} R^{m-1}x \right\|$ and select the location $(s_m^\star, u_m^\star, \xi_m^\star)$ of the best real-valued Gaussian Gabor atom $[\mathcal{O}(N\log N)]$.
3. Estimate the locally optimal parameters s_m and c_m with a parabolic interpolation $[\mathcal{O}(1)]$.
4. Compute $\left\langle R^{m-1}x, g_{(s_m,u_m^\star,\xi_m^\star,c_m)}\right\rangle$, and determine the best real-valued chirp atom $g_{(s_m,u_m^\star,\xi_m^\star,c_m,\phi_m)}$ in $\boldsymbol{V}_{(s_m,u_m^\star,\xi_m^\star,c_m)}$ $[\mathcal{O}(N)]$.
5. Update the residual $[\mathcal{O}(N)]$.

The overall complexity of one iteration of real-valued Ridge Pursuit is $\mathcal{O}(N\log^2 N)$, hence the total cost $\mathcal{O}(MN\log^2 N)$ of M iterations. An accelerating technique was introduced by Bergeaud and Mallat [2, 3] for the Matching Pursuit analysis of images. It can be used to get a Fast Ridge Pursuit algorithm. The overall algorithm is described in full details in [13], and we give here its main features. We use *local maxima* of the Gabor dictionary $\mathcal{D}$, that is Gabor atoms $g_{(s,u,\xi)}$ where either $u \mapsto \left|\left\langle x, g_{(s,u,\xi)}\right\rangle\right|$ or $\xi \mapsto \left|\left\langle x, g_{(s,u,\xi)}\right\rangle\right|$ has a local maximum. A number P is fixed arbitrarily, and the following steps are done iteratively.

Fast Ridge Pursuit Algorithm

1. Build a sub-dictionary $\mathcal{D}_m$ of P local maxima of the Gabor dictionary $\mathcal{D}$.
2. For each atom in $\mathcal{D}_m$, use the fast local estimation procedure to get a good chirp atom. The collection of these chirp atoms is a sub-dictionary $\mathcal{D}_m^+$ of the chirp dictionary $\mathcal{D}^+$.
3. Run a "normal" pursuit in $\mathcal{D}_m^+$ until it is empty.

By choosing $P \propto \mathcal{O}(\sqrt{N}\log N)$, the overall complexity becomes $\mathcal{O}(MN)$ [13].

6 Applications

The Ridge Pursuit and Fast Ridge Pursuit algorithms were implemented using the Matching Pursuit Package of the LastWave program [18]. We used them to analyze a sound recording with sung voice and orchestra [8]. It is well known that a characteristic

of the sung voice is its vibrato [30], which the Gabor Matching Pursuit was not likely to decompose sparsely. The signal duration was approximately 2.5 seconds, at a sampling rate of $11,025$ Hertz, so the signal length was about $N \approx 30000$ samples. A Gabor Matching Pursuit and a Fast Ridge Pursuit were computed with $M = 5000$ iterations.

One needs first to realize how high the complexity of a "brute force" Matching Pursuit with the chirp dictionary would have been. With an (optimistic) average of 100 MFlops to 1 GFlops for todays computers, the $5000 \times (30000)^2 \log_2 30000 \approx 6 \times 10^{13}$ operations would have required 6×10^4 to 6×10^5 seconds of computation, that is between 16 and 160 hours of computation. This must be compared to the 2.5 seconds duration of the signal. On the other hand, the Fast Ridge Pursuit was run on a consumer PC. It only took 200 seconds to get the result.

Figure 2 displays the decrease, in decibels, of the energy of the residual. It is faster with the Fast Ridge Pursuit than with the standard Gabor Matching Pursuit. In order to get the same approximation quality, one thus needs fewer chirp atoms than Gabor atoms.

One can compare on figure 3 the time-frequency distributions [22, 5] associated to the Gabor Matching Pursuit and Fast Ridge Pursuit decompositions of the signal. The display is focused on a time-frequency area wherein the vibrato occurs. The Gabor Matching Pursuit decomposes the *vibrato* into several constant-frequency atoms, located on the "path" of the instantaneous frequency. On the contrary, the Fast Ridge Pursuit decomposes it into only a few chirp atoms, whose instantaneous frequency is alternatively increasing and decreasing.

7 Comments

We checked numerically that the fast estimate given by proposition 1 fails for non-Gaussian windows (even for B-spline windows, which in some sense are close to Gaussian windows). Even if an analogous of theorem 1 can be derived for such windows, the lack of analytic tools makes it difficult to derive an analogous of the fast and simple estimation procedure. It may be possible however to get fast estimates using *regression* [27] instead of linear interpolation to fit the local behavior of the spectrum around the best Gabor atom.

We do not cover in this article the theoretical question of the convergence of the Ridge Pursuit. One should notice that the convergence is in general not guaranteed by the fact that it is *stepwise more greedy* (the chosen chirp atom grabs more energy than the best Gabor atom) than the Gabor Matching Pursuit.

8 Conclusion

The Fast Ridge Pursuit algorithm iteratively decomposes an N-sample acoustic signal into M Gaussian chirp atoms with a computational cost $\mathcal{O}(MN)$. Thanks to its low computational complexity, the sparse structured representation of signals that it provides can become the basis for the implementation of a large variety of new processing tools.

Besides its evident use for signal compression, one of its most interesting features is its ability to decompose a signal into superimposed structures with different scale, frequency and chirp characteristics. Thanks to this decomposition property, it is possible to process separately the different parts (*e.g.* transients and steady parts) of a signal. Source separation can be achieved for sounds that have very different "chirp behavior", such as a singer (with a strong vibrato) and an orchestra. Additionally, considering time-stretching or pitch-shifting applications, it is possible to keep the fine structure of transients while processing the harmonic part of a sound. Because they respect the structure of the transients, and as the chirp parameter enables them to fit more finely the phase of the signal, such pitch shifting schemes will generate less "pipe noise" than standard windowed Fourier transform based techniques. Moreover their implementation using the chirplet decomposition is straightforward.

Acknowledgments

The author would like to thank E. Bacry and S. Mallat, from Ecole Polytechnique, for their encouragement and all the interesting discussions they fed. The sound signal was kindly provided by Xavier Rodet, from IRCAM, whom I also wish to thank. All the numerical computations and figures were obtained using the GNU-licensed software LastWave [18].

A Proof of the ridge theorem

In this appendix, we give a proof of theorem 1. Building Taylor expansions of $\alpha(t)$ and $\phi(t)$ near $t = u$, one can find $\theta_1(t), \theta_2(t) \in [u, u+t]$ such that $a(u+t) = a(u) e^{-\alpha'(u)t - \frac{\alpha''(u)}{2}t^2 - \frac{\alpha'''(\theta_1(t))}{6}t^3}$ and $\phi(u+t) = \phi(u) + \phi'(u)t + \frac{\phi''(u)}{2}t^2 + \frac{\phi'''(\theta_2(t))}{6}t^3$. By

changing variables, and using the definition of the Gaussian window $g(t)$, we express $\left\langle x, g_{(s,u,\xi,c)} \right\rangle$ as

$$\frac{a(u)e^{i\phi(u)}}{(\alpha''(u)/\pi)^{1/4}} e^{\frac{(\alpha')^2}{2\alpha''}} \int_{-\infty}^{+\infty} g_{(1/\sqrt{\alpha''},-\alpha'/\alpha'',0,0)}(t) \cdot \overline{g_{(s,0,\xi-\phi',c-\phi'')}(t)} \cdot e^{\frac{t^3}{6}\left(-\alpha'''(\theta_1(t))+i\phi'''(\theta_2(t))\right)} dt.$$

The integral can be rewritten as

$$\left\langle g_{(1/\sqrt{\alpha''},-\alpha'/\alpha'',0,0)}, g_{(s,0,\xi-\phi',c-\phi'')} \right\rangle + \epsilon(s,u,\xi,c) \qquad (21)$$

where the error term $\epsilon(s,u,\xi,c)$ is

$$\int_{-\infty}^{+\infty} g_{(1/\sqrt{\alpha''},-\alpha'/\alpha'',0,0)}(t) \overline{g_{(s,0,\xi-\phi',c-\phi'')}(t)} \cdot \left(e^{\frac{t^3}{6}\left(-\alpha'''(\theta_1(t))+i\phi'''(\theta_2(t))\right)} - 1 \right) dt.$$

Let us now bound the error term, using again the expression of the Gaussian window, and splitting the integral with a parameter η :

$$\begin{aligned} |\epsilon(s,u,\xi,c)| &\leq (\alpha'')^{1/4}\pi^{-1/2}s^{-1/2} \cdot \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2s^2}-\frac{\alpha''}{2}\left(t+\alpha'/\alpha''\right)^2} \cdot \left| e^{\frac{t^3}{6}\left(-\alpha'''(\theta_1(t))+i\phi'''(\theta_2(t))\right)} - 1 \right| dt \\ &\leq (\alpha'')^{1/4}\pi^{-1/2}s^{-1/2} \cdot \left(\int_{|t|>\eta s} + \int_{|t|\leq \eta s} \right). \end{aligned}$$

The first part of the split integral is bounded by

$$\begin{aligned} & \int_{|t|>\eta s} e^{-\frac{t^2}{2s^2}} \left| e^{-\frac{\alpha''}{2}\left(t+\alpha'/\alpha''\right)^2} - e^{-\alpha' t - \frac{\alpha''}{2}t^2 - \frac{t^3}{6}\alpha'''(\theta_1(t)) - \frac{(\alpha')^2}{2\alpha''}} e^{i\frac{t^3}{6}\phi'''(\theta_2(t))} \right| dt \\ = & \int_{|t|>\eta s} e^{-\frac{t^2}{2s^2}} \left| e^{-\frac{\alpha''}{2}\left(t+\alpha'/\alpha''\right)^2} - \frac{a(u+t)}{a(u)} e^{-\frac{(\alpha')^2}{2\alpha''}} e^{i\frac{t^3}{6}\phi'''(\theta_2(t))} \right| dt \\ \leq & \int_{|t|>\eta s} e^{-\frac{t^2}{2s^2}} \left(\frac{\|a\|_\infty}{|a(u)|} + 1 \right) dt \\ \leq & \frac{2\|a\|_\infty}{|a(u)|} \int_{|t|>\eta s} e^{-\frac{t^2}{2s^2}} dt \\ \leq & s\frac{2\|a\|_\infty}{|a(u)|}\frac{2}{\eta}e^{-\eta^2/2} \leq s\frac{\|a\|_\infty}{|a(u)|}\frac{4}{\eta}e^{-\eta^2/2}. \end{aligned} \qquad (22)$$

Knowing that $\forall z \in \mathbb{C}$, $|e^z - 1| \leq e^{|z|} - 1 \leq |z|\, e^{|z|}$ we can bound the second part with

$$\begin{aligned} & \int_{|t|\leq \eta s} e^{-\frac{t^2}{2s^2}} \frac{|t|^3}{6} \left(|\alpha'''(\theta_1(t))| + |\phi'''(\theta_2(t))| \right) \cdot e^{\frac{|t|^3}{6}\left(|\alpha'''(\theta_1(t))|+|\phi'''(\theta_2(t))|\right)} dt \\ \leq & s\frac{\|\alpha'''\|_\infty + \|\phi'''\|_\infty}{6} \cdot e^{\frac{\eta^3 s^3}{6}\left(\|\alpha'''\|_\infty + \|\phi'''\|_\infty\right)} s^3\sigma_3^3. \end{aligned} \qquad (23)$$

We denote by $K = \|\alpha'''\|_\infty + \|\phi'''\|_\infty$ and get from (22) and (23) that for all $\eta > 0$

$$|\epsilon(s,u,\xi,c)| \leq (\alpha'')^{1/4}\pi^{-1/2}s^{+1/2}\left(\frac{\|a\|_\infty}{a(u)}\frac{4e^{-\frac{\eta^2}{2}}}{\eta} + \frac{Ks^3\sigma_3^3}{6}e^{\frac{Ks^3}{6}\eta^3}\eta^3 \right). \qquad (24)$$

Choosing $\eta^3 = K^{-1}s^{-3}$ gives (9). To conclude the proof, we rewrite the first term of (21) as

$$e^{-i\phi'\frac{\alpha'}{\alpha''}+i\frac{\phi''}{2}\left(\frac{\alpha'}{\alpha''}\right)^2} \cdot \left\langle g_{\left(\frac{1}{\sqrt{\alpha''}},u-\frac{\alpha'}{\alpha''},\phi'-\phi''\frac{\alpha'}{\alpha''},\phi''\right)}, g_{(s,u,\xi,c)} \right\rangle.$$

B Real-valued atoms

Let $\tilde{g_\gamma} = \left(g_\gamma - \langle g_\gamma, \overline{g_\gamma}\rangle\, \overline{g_\gamma}\right) / \left(1 - |\langle \overline{g_\gamma}, g_\gamma\rangle|^2\right)$ and $\boldsymbol{V}_\gamma = \mathrm{Span}\left(g_\gamma, \overline{g_\gamma}\right)$. One can check that $\{\tilde{g_\gamma}, \overline{\tilde{g_\gamma}}\}$ is the dual basis of $\{g_\gamma, \overline{g_\gamma}\}$. Thus for all $x \in \boldsymbol{H}$

$$\begin{aligned} P_{\boldsymbol{V}_\gamma}x &= \langle x, \tilde{g_\gamma}\rangle\, g_\gamma + \langle x, \overline{\tilde{g_\gamma}}\rangle\, \overline{g_\gamma} \\ \left\| P_{\boldsymbol{V}_\gamma}x \right\|^2 &= \frac{2\left(|\langle x, g_\gamma\rangle|^2 - \Re(\langle g_\gamma, \overline{g_\gamma}\rangle\, \langle x, g_\gamma\rangle^2)\right)}{1 - |\langle \overline{g_\gamma}, g_\gamma\rangle|^2} \end{aligned}$$

where $\Re(z)$ denotes the real part of $z \in \mathbb{C}$. For x real-valued and $\gamma = (s,u,\xi,c)$, the first equality can be rewritten $P_{\boldsymbol{V}_{(s,u,\xi,c)}}x = \left\langle x, g_{(s,u,\xi,c,\phi)} \right\rangle g_{(s,u,\xi,c,\phi)}$ with $\left\langle x, g_{(s,u,\xi,c,\phi)} \right\rangle = \left\| P_{\boldsymbol{V}_\gamma}x \right\|$ and $e^{i\phi} = \frac{\langle x, \tilde{g_\gamma}\rangle}{|\langle x, \tilde{g_\gamma}\rangle|}$. The value $\left\langle \overline{g_{(s,u,\xi,c)}}, g_{(s,u,\xi,c)} \right\rangle$ can be computed up to an arbitrary precision with a cost $\mathcal{O}(1)$, thanks to an analytic expression [5] [28]. Once $\langle x, g_\gamma\rangle$ is known, so is its complex conjugate $\langle x, \overline{g_\gamma}\rangle$, thus $\phi(s,u,\xi,c)$ and $\left\langle x, g_{(s,u,\xi,c,\phi)} \right\rangle$ can be computed in $\mathcal{O}(1)$.

References

[1] R. G. Baraniuk and D. L. Jones. Shear madness : New orthonormal bases and frames using chirp functions. *IEEE Trans. Signal Process.*

Special Issue on Wavelets in Signal Processing, 41(12):3543–3548, December 1993.

[2] F. Bergeaud. *Représentations adaptatives d'images numériques,* Matching Pursuit. PhD thesis, Ecole Centrale Paris, 1995.

[3] F. Bergeaud and S. Mallat. Matching pursuit : adaptive representations of images and sounds. *Journal of Computational and Applied Mathematics*, 1996.

[4] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. *Commun. on Pure and Appl. Math.*, 44:141–183, 1991.

[5] A. Bultan. A four-parameter atomic decomposition of chirplets. *IEEE Trans. Signal Process.*, 47(3):731–745, March 1999.

[6] R. Coifman and M.V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, 38(2):713–718, March 1992.

[7] R. R. Coifman and Y. Meyer. Remarques sur l'analyse de Fourier à fenêtre. *Comptes-Rendus Acad. Sci. Paris(A)*, 312:259–261, 1991.

[8] M.-A. Dalbavie. "Marc-André Dalbavie". In *Compositeurs d'Aujourd'hui*. IRCAM, 19991-93.

[9] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Commun. on Pure and Appl. Math.*, 41:909–996, November 1988.

[10] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[11] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, Ph. Tchamitchian, and B. Torrésani. Asymptotic wavelet and Gabor analysis : Extraction of instantaneous frequency. *IEEE Trans. Inform. Theory*, 38(2):644–664, March 1992.

[12] P. Flandrin. *Temps-Fréquence*. Hermes, Paris, France, 1993.

[13] R. Gribonval. *Approximations non-linéaires pour l'analyse de signaux sonores*. PhD thesis, Université Paris IX Dauphine, September 1999.

[14] P. J. Huber. Projection pursuit. *The annals of statistics*, 13(2):435–475, 1985.

[15] S. Jaffard. Pointwise smoothness, two microlocalization and wavelet coefficients. *Publicacions Matemàtiques*, 35:155–168, 1991.

[16] L.K. Jones. On a conjecture of Huber concerning the convergence of PP-regression. *The Annals of Statistics*, 15:880–882, 1987.

[17] H. K. Kwok and D. L. Jones. Improved FM demodulation in a fading environment. In *Proc. IEEE Conf. Time-Freq. and Time-Scale Anal. (TFTS'96)*, pages 9–12, Paris, June 1996.

[18] LastWave software. Freely available at http://wave.cmap.polytechnique.fr/soft/LastWave/.

[19] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, pages 674–693, July 1989.

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

[21] S. Mallat and W.L. Hwang. Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, 38(2):617, 1992.

[22] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, December 1993.

[23] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 40:2464–2482, July 1992.

[24] S. Mann and S. Haykin. The chirplet transform : Physical considerations. *IEEE Trans. Signal Process.*, 43(11):2745–2761, November 1995.

[25] J. S. Marques and L. B. Almeida. A background for sinusoid based representation of voiced speech. In *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP'86)*, pages 1233–1236, Tokyo, 1986.

[26] J. S. Marques and L. B. Almeida. Frequency-varying sinusoidal modeling of speech. *IEEE Trans. Speech and Audio Process.*, 37(5):763–765, May 1989.

[27] C.M. McIntyre and D.A. Dermott. A new fine-frequency estimation algorithm based on parabolic regression. In *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP'92)*, pages 541–544, 1992.

[28] A. Papoulis. *The Fourier integral and its applications*. Mcgraw-Hill Publisher Co., 1987.

[29] S. Qian and D. Chen. Signal representation using adaptive normalized Gaussian functions. *Signal Process.*, 36(1):1–11, 1994.

[30] X. Rodet. Time-domain formant-wave functions synthesis. In J.C. Simon, editor, *Spoken Language Generation and Understanding*, C: Mathematical and Physical Sciences, chapter 4-Speech Synthesis, pages 429–441. D. Reidel Publishing Company, 1980.

[31] B. Torrésani. Wavelets associated with representations of the affine Weyl-Heisenberg group. *J. Math. Phys.*, 32:1273–1279, May 1991.

[32] G.H. Watson and K. Gilholm. Signal and image feature extraction from local maxima of generalized correlation. *Pattern Recognition*, 31(11):1733–1745, 1998.

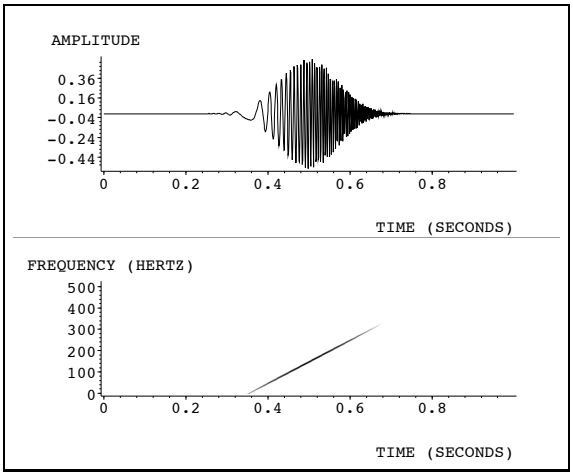

Figure 1: A Gaussian chirp atom and its Wigner-Ville transform.

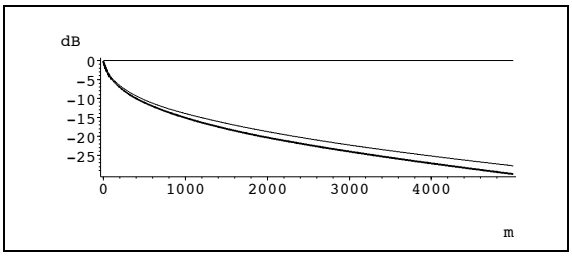

Figure 2: Decay (in decibels) of the relative energy $\|R^m x\|^2 / \|x\|^2$ of the residual with the number m of iterations. Plain : Gabor Matching Pursuit. Bold : Fast Ridge Pursuit with chirp dictionary. One needs fewer chirp atoms than Gabor atoms to get the same approximation quality.

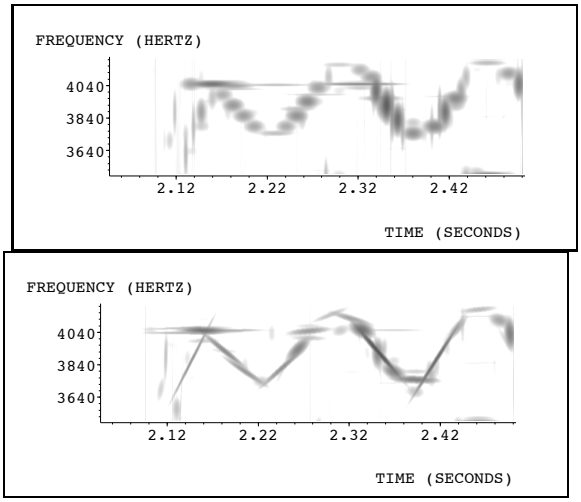

Figure 3: Time-frequency distributions of a sound signal of size $N \approx 30000$, with $M = 5000$ iterations of (a) Gabor Matching Pursuit (b) Fast Ridge Pursuit. The display is focused on a time-frequency region wherein the vibrato occurs. Vertical lines (*e.g.* at time $t = 2.1$) correspond to short scale atoms that represent transients. Horizontal lines, associated to large scale constant frequency atoms, represent the resonance of the notes of the instruments of the orchestra. The vibrato is decomposed into several constant frequency atoms by the Gabor Matching Pursuit. On the contrary the Fast Ridge Pursuit decomposes it into only a few chirp atoms.