ACADEMIC PRESS

Appl. Comput. Harmon. Anal. 13 (2002) 1–21

Applied and Computational Harmonic Analysis

www.academicpress.com

Adaptive Gabor transforms

Ingrid Daubechies and Fabrice Planchon *,1

Program in Applied and Computational Mathematics, Princeton University, Princeton NJ 08544-1000, USA

Received 2 September 2000; revised 28 February 2002

Communicated by P. Flandrin

Abstract

We aim to provide time-frequency representations of a one-dimensional signal where the window is locally adapted to the signal, thus providing a better readability of the representation.

Keywords: Gabor transform; Wigner–Ville transform; Adaptive time-frequency analysis

Introduction

All the linear transforms which allow to depict a signal in phase space (time-frequency plane) have a blurring effect, because they typically introduce an auxiliary function which can be chosen arbitrarily and which serves as a "template:" the window for a Gabor transform, the wavelet for a wavelet transform. Various attempts have been made to correct such blurring artifacts while retaining the interesting properties of such transforms. Let us cite reallocation methods [11], or squeezing [6]. On the other hand, bilinear transform such as the Wigner transform don't introduce extraneous "templates," and have less blurring for some classes of signals. In particular, the Wigner transform is well-localized for linear chirps. However, for more complicated signals, interferences appear due to the quadratic nature of the transform; often these are difficult to separate from the interesting components in the representation. Adaptive methods exist in this context to correct such problems [4, 10], but they usually suffer from a high-computational cost when compared to linear transforms. We aim at providing various adaptive Gabor representations that behave well on simple signals, and can be computed at reasonable cost. Techniques of this type have also been developed in [1], where they introduce the so-called warped-Gabor representation. Other adaptive methods are presented in [13], which are based on an adaptive Wigner–Ville transform (using different windows). This later approach however does not solve the problem of cross-term effects.

1. Wigner–Ville transform and Gabor transform

We recall that, given a one-dimensional signal $f(x)$, its Wigner–Ville transform is given by

$$W(f)(x,\xi) = \int e^{-i\xi u} f\left(x + \frac{u}{2}\right) f\left(x - \frac{u}{2}\right) \mathrm{d}u. \tag{1}$$

* Corresponding author.
E-mail address: fab@ann.jussieu.fr (F. Planchon).
[1] Current address: Laboratoire d'Analyse Numérique, URA CNRS 189, Université Pierre et Marie Curie, 4 place Jussieu BP 187, 75 252 Paris Cedex, France.

1063-5203/02/$ – see front matter
PII: S1063-5203(02)00003-9

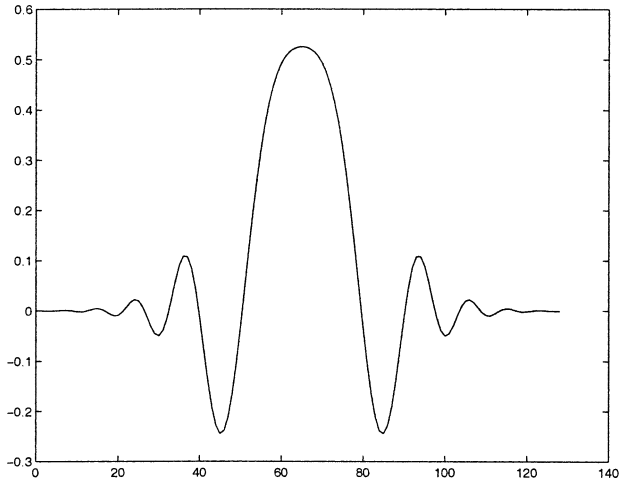

Fig. 1. Real part of signal g^{β_0,δ_0}.

Let's define a family of Gaussian windows,

$$g^{\beta,\delta}(x) = \delta^{\frac{1}{4}} e^{i(\beta/2)x^2 - \delta x^2}, \tag{2}$$

and the collection

$$g^{\beta,\delta}_{p,q}(x) = e^{ipx - pq/2} g^{\beta,\delta}(x-q). \tag{3}$$

Then, for fixed β and δ, the Gabor transform w.r.t. $g^{\beta,\delta}$ reads

$$G^{\beta,\delta}(f)(p,q) = \int g^{\bar{\beta},\delta}_{p,q}(x) f(x)\,\mathrm{d}x. \tag{4}$$

We will call such a transform the (β,δ)-Gabor transform; for the $(0,1)$-Gabor transform we often drop the label $(0,1)$, calling it simply the Gabor transform. The Wigner–Ville transform of such a window reads

$$W\big(g^{\beta,\delta}\big)(x,\xi) = Ce^{-(2\delta+\beta^2/2\delta)x^2 - (1/2\delta)\xi^2 + (\beta/\delta)x\xi}. \tag{5}$$

It is well known that we can obtain the energy of the Gabor transform by convolving the Wigner–Ville transform of f with the Gaussian $W(g^{\beta,\delta})$ [8]

$$\big|G^{\beta,\delta}(f)(p,q)\big|^2 = W\big(g^{\beta,\delta}\big) * W(f). \tag{6}$$

This explains why $|G^{\beta,\delta}(f)|^2$ shows blurring when compared to $W(f)$. On very simple signals like linear chirps, the Wigner–Ville transform is perfectly localized as a Dirac along the instantaneous frequency, while the "sharpness" of the Gabor transform depends a lot on the choice of windows. Let us illustrate this on a simple signal, say one of our (β,δ) windows, i.e., take $f = g^{\beta_0,\delta_0}$, with $\beta_0 \neq 0$ and $\delta_0 > 1$. The (real part of) the signal appears in Fig. 1 (notice that we chose a non-zero β value, therefore we get a chirped Gaussian) and its Wigner–Ville transform and $(0,1)$-Gabor transform (for which we plot the modulus) are in Figs. 2 and 3, respectively. However, other $G^{\beta,\delta}$ transforms of this f can look much better or worse than $G^{0,1}$. For instance, choosing $\beta = \beta_0$, $\delta = \delta_0$ leads to Fig. 4, while the choice $\beta = -\beta_0$, $\delta = \delta_0$ gives Fig. 5 which clearly looks different. To understand how these differences are caused, let us look more specifically at the role played by β and δ. For any (β,δ), the Wigner transform $W(g^{\beta,\delta})$ is an elliptically concentrated Gaussian (5). To this ellipsoid we can associate a couple (θ,λ), where λ is the eccentricity of the ellipse and θ is the angle of the largest axis with the time axis. Switching the sign of β means switching the sign of θ. This explains why in our previous example changing the choice of β had such a marked effect on the Gabor transform representation, and why the choice $\beta = -\beta_0$ gave misleading results, caused by the mismatch between f and the template $g^{\beta,\delta}$. Because the choice of (β,δ) that will bring out features best will depend on the time-frequency content of the signal itself, it is therefore natural to adapt the choice of parameters (β,δ) to

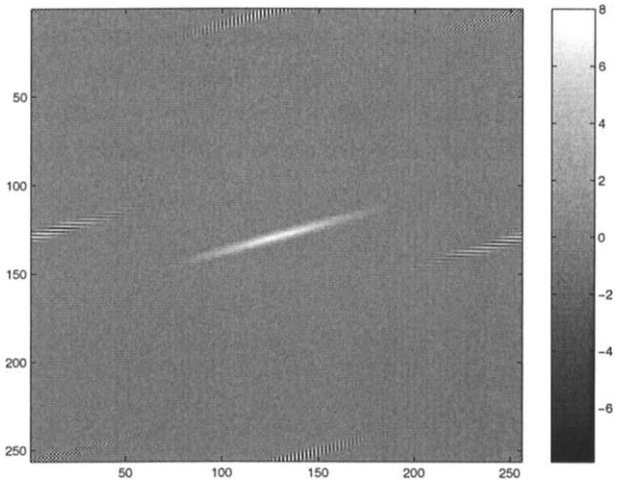

Fig. 2. Wigner–Ville transform of g^{β_0,δ_0}. Note that this plot was obtained via a numerical implementation that uses the FFT to speed up computations, which effectively periodizes in the same time-frequency plane. The "ghost" terms noticeable near the edges result from interference between the signal at the center and the periodized copies.

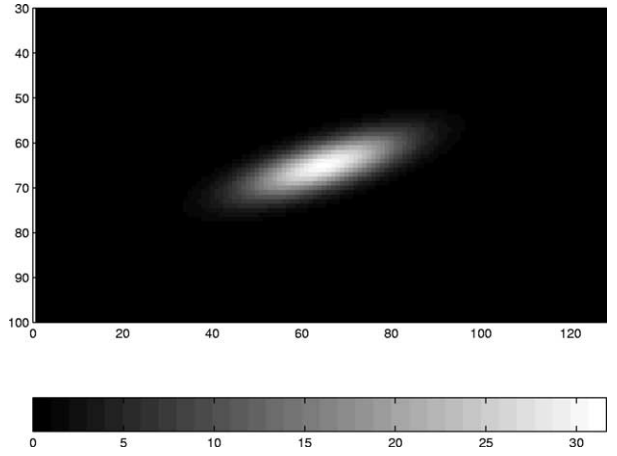

Fig. 3. Gabor transform of g^{β_0,δ_0}.

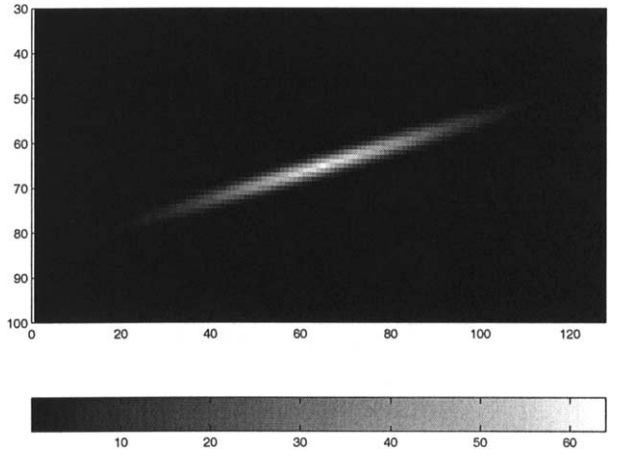

Fig. 4. (β_0, δ_0)-Gabor transform of g^{β_0,δ_0}.

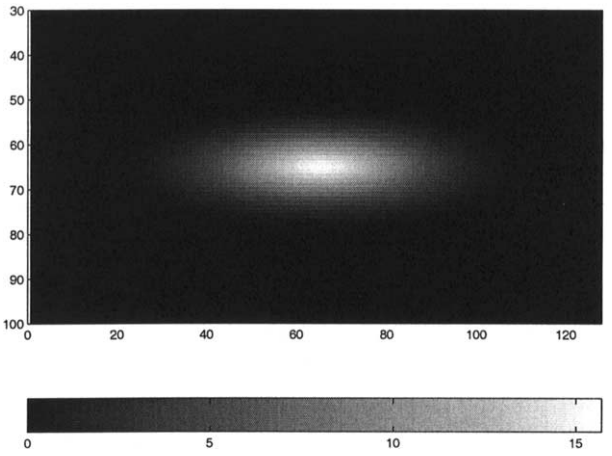

Fig. 5. $(-\beta_0, \delta_0)$-Gabor transform of g^{β_0,δ_0}.

the phase point (p, q). However, there are several difficulties with this approach. First of all, it is not clear what would be good criteria for choosing appropriate (β, δ). Adaptivity also means that, unlike for the linear transforms, there is no a priori way to recover the signal from the transform. While reconstruction is not always a priority, it certainly is useful in some circumstances, since it enables one to extract separate components from complex signals. Computational costs are also a problem, since the complexity for linear representations is proportional to (if, for the sake of simplicity, we look only at squared representations) $N^2 \log N$, while a naive way to implement the aforementioned adaptive transform would cost N^3. Note that several approaches in order to "improve" the Wigner transform exist in the literature, for example, [2] and the references given in the book by Flandrin [8] which presents a comprehensible survey of time-frequency representations.

2. Picking the right window

In this section we present different methods to optimize the window choice, depending on the location in the time-frequency plane. We remark that in order to keep the computational cost reasonable, we should not compute too many Gabor transforms with different parameters. In effect, in [13], only two Wigner transforms are computed and then a choice is performed. On the other hand, the second method presented in [1], called the energy method, can be compared to the curvature approach we present in Section 2.2. Though different by the criterion which is used to perform the window choice, both methods use a local grid around a given point in the time-frequency plane, arising from a set of precomputed transforms. By contrast, the so-called masking method from [1] is computationally more expensive.

2.1. Maximizing the modulus of the Gabor transform

It is apparent on the chirp example above that choosing a window with phase orientation in the same direction as the chirp will give a better localized representation. This is because convolving with an elongated Gaussian which points in a different direction causes more broadening than if its direction is "matched" to that of the signal. On the other hand, in regions of the time-frequency domain where the signal has little energy, convolving with a wide range of Gaussian will consistently give negligible results.

Consider

$$s(j) = \exp\left(2\pi i j \left(2 + 4\cos\left(2\pi \frac{j}{n}\right)/n\right)\right),$$

where $1 \leqslant j \leqslant 128$. Then its Gabor transform appears in Fig. 6. By maximizing pointwise over a set of twelve Gabor transforms corresponding to all pairings of six different β

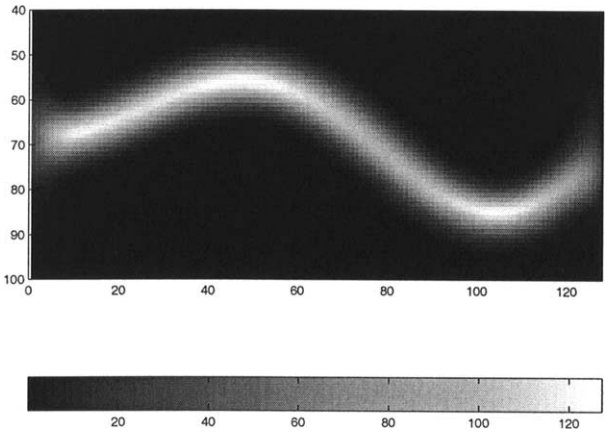

Fig. 6. Gabor transform of $s(j)$.

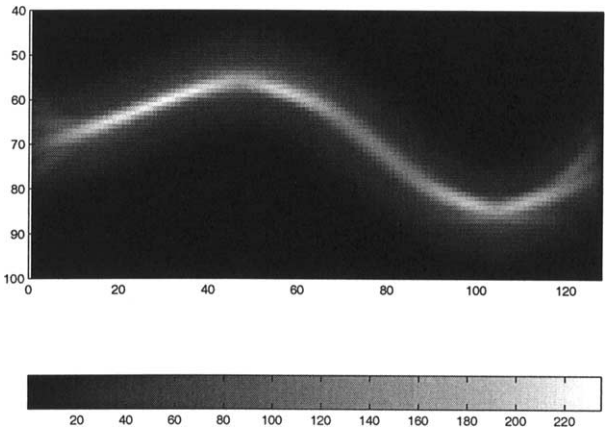

Fig. 7. Max–Gabor transform of $s(j)$.

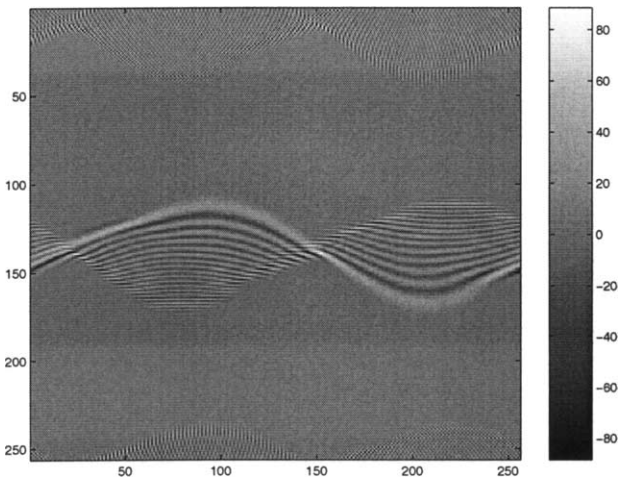

Fig. 8. Wigner–Ville transform of $s(j)$. The interference terms at top and bottom are caused by the numerical implementation, unlike the interference at the center.

values with two possible λ values, we obtain Fig. 7. For the same signal, the Wigner–Ville transform is given by Fig. 8 (we used the numerical implementation proposed in [12]).

We note how the numerical Wigner–Ville transform, while well-localized on the chirp, shows not only the interference terms expected in a quadratic transform, but suffers moreover from aliasing (due to the use of the FFT in the numerical implementation),

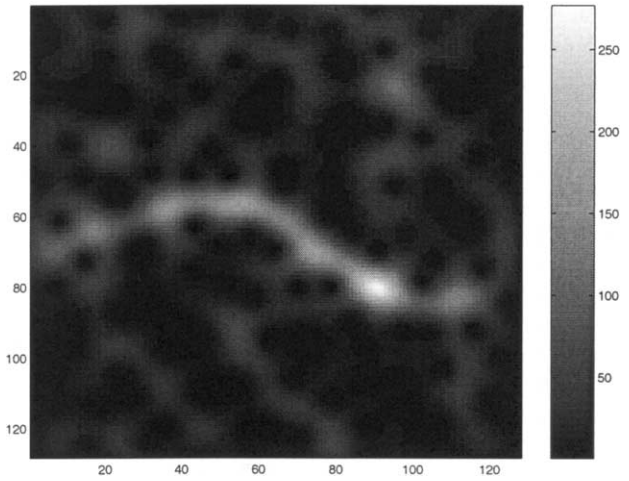

Fig. 9. Gabor transform of noisy $s(j)$.

which is non-existent in the two other transforms (even though they use the FFT as well). Maximizing the modulus of the Gabor transform adaptively has the following benefits:

1. As we shall see below, it gives good localization for several types of simple signals rather well, such as chirps or signals with an instantaneous frequency law with slow variation (note that only very simple signals have a perfectly localized Wigner transform, see [9]);
2. It is reasonably fast to compute, since we can compute several Gabor transforms and take the maximum;
3. It seems to be resistant to noise. For instance, if we add white noise to our previous example, with a signal-to-noise (SNR) ratio of -3 dB then the usual Gabor transform of the noisy signal (in Fig. 9) is much less "clear" than its Max–Gabor transform (in Fig. 10). This can be explained heuristically by the observation that the phase space representation of the noise doesn't have a lot of coherence, so that averaging it in any direction will tend to give a relatively small result; the maximum of all these smoothed out versions will be significantly smaller than the maximum of the Gabor transforms of a true "component" of the signal.

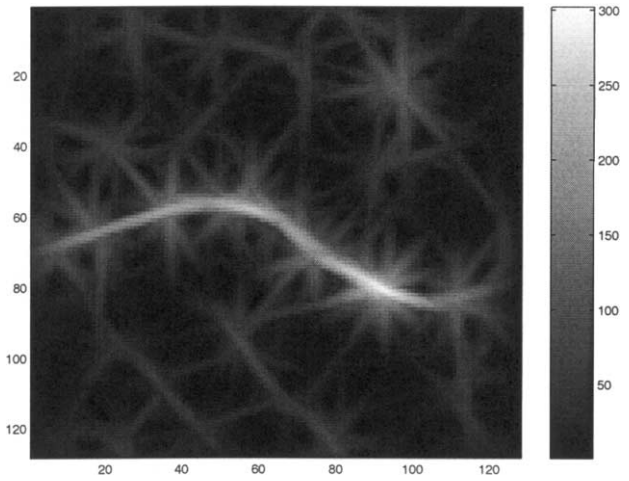

Fig. 10. Max–Gabor transform of noisy $s(j)$. This should be compared with Fig. 9; note that the dynamic range (indicated by the scaling on the gray level bar) is more extended here than in Fig. 9.

The method also has drawbacks, however:

1. There is no way to recover the signal;
2. It tends to raise the values of the representation everywhere (which is not seen on the pictures here because we renormalized the displays);
3. It has artifacts, when components cross or when the "instantaneous" frequency varies too fast (as can be observed in both Figs. 7 and 10, or, below, in Fig. 18, showing the artifacts produced by this method for a clean signal with crossing components, and in Fig. 19, showing similar artifacts for a noisy version of the same signal).

Several variations on this method can be considered. For instance, one could argue that it would be useful to use the *minimum*, instead of the maximum, of the different local Gabor transforms, for the following reason. Next to a "ridge" in the signal, the maximal Gabor transform uses an elliptical orientation that will be orthogonal to the ridge, thereby capturing as much as possible of the ridge contribution. This will lead to some broadening of the representation. Choosing the minimal Gabor contribution in such a point would lead to lining up the ellipse parallel to the ridge, avoiding "contamination" by the ridge; the resulting representation would be better focused. On the other hand, a similar argument holds if we place ourselves on the ridge: the maximal Gabor transform will line up with the ridge; the minimal Gabor transform will select an ellipse orthogonal to the ridge. The contribution of the minimal Gabor transform will therefore be much smaller. Whether one or the other representation is more useful depends on a comparison of the dynamical range of the two proposals on and near components; in practice it turns out that this is much better for the maximal Gabor transform than for the minimal Gabor transform.

One can also set thresholds, to eliminate artifacts. Figures 11 and 12 give thresholded versions of the standard Gabor transform, and the Max–Gabor transform, respectively, of our earlier noisy signal. (The thresholding is done rather brutally: only coefficients that exceed 0.75 times the maximum of the representation are kept in each case.) It is

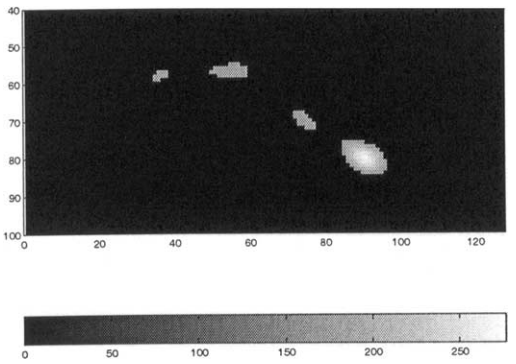

Fig. 11. Thresholding of the Gabor transform.

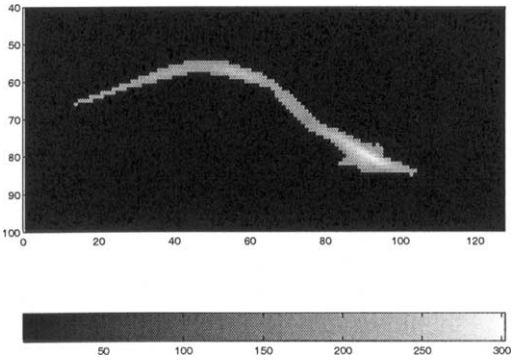

Fig. 12. Thresholding of the Max–Gabor transform.

clear that this works better for the sharper Max–Gabor representation than for the standard Gabor representation; again, this expresses the heuristic that the Max–Gabor representation provides a better "focusing" on the signal, without amplifying the noise.

Although the Max–Gabor transform does therefore a reasonably good job of presenting a sharper picture for the signal than standard Gabor transforms, it is, to say the least, a very ad-hoc procedure. In the next section we propose a different approach that has a more "physical" intuition.

2.2. Curvature and Wigner–Ville

In this section, we carry out an adaptive smoothing, still by computing windowed Fourier transforms with different chirped Gaussians at different phase space locations, but we propose to select the parameters for this adaptive smoothing based on the information in the Wigner–Ville transform Wf of the signal f. It is tempting to suggest that the parameter choices at (p,q) for the best ellipsoidal localization of the convolving Gaussian at that point should be linked to the curvature at (p,q) of the graph of Wf; intuitively, this would ensure the convolution does "not spread out things too much." If we admit that the Wigner–Ville transform localizes very well the kind of signals in which we are interested (modulo interferences), then this suggests that on any ridge, we use its Hessian to compute adapted (β,δ). If we use this strategy where interference terms dominate, then the sign changes near the interference terms will lead to near cancellation after convolution, so that these terms get attenuated. Practically, we can separate two components as long as their own width (as measured by their separate Wigner–Ville representations) does not exceed the distance between them.

The orientation θ of the ellipsoidal localization of the adaptive $G^{\beta,\delta}$ will be chosen to line up with the orientation of the principal axis of the Hessian of $W(f)$ at (p,q). It then remains to determine the eccentricity λ; this will be given by the ratio of long and short axes. It is easy to see how to do this in a coordinate system where the two axis are simply vertical and horizontal. A simple computation shows that to optimize

$$\int e^{-\mu x^2/2-\nu\xi^2/2-\lambda^2x^2-\lambda^2x^2-\xi^2/\lambda^2}\,\mathrm{d}x\,\mathrm{d}\xi,$$

where the first two factors of the integrand stand for the local behavior of the Wigner transform (where μ and ν can be computed from the local Hessian of Wf), and the second two factors give the eccentricity of the elliptical convolution, we have to choose $\lambda^2=\sqrt{\mu/\nu}$. In practice, this may not always be the best choice for λ; we shall come back to variations on this theme below.

In theory, we would therefore proceed as follows:

1. We compute the Wigner–Ville transform;
2. We compute the Hessian at point (p,q). This is achieved numerically by taking the best approximation by a paraboloid on a grid which can be $9\otimes 9$ or bigger;
3. From the Hessian we get the two eigenvalues and eigenvectors, from which we get the couple (β,δ) corresponding to orienting the Gaussian so that its curvature matches the Hessian of the Wigner–Ville transform at (p,q) as well as possible.
4. We compute $G^{\beta,\delta}(p,q)$.

Note that this last step is the most computationally intensive. There are several alternatives that would speed up the algorithm: we can compute $G^{0,1}$ and get $G^{\beta,\delta}(p,q)$ from $G^{0,1}(p',q')$ for (p',q') close to (p,q). Or we can compute several $G^{\beta,\delta}$ for a set of (β,δ) and then choose the closest one, or interpolate.

In Fig. 13 we try this strategy on our previous example; note that Fig. 13 shows how the interference terms that are particularly strong in the center of the corresponding "pure" Wigner–Ville transform in Fig. 8 lead to artifacts on our adaptive transform. Moreover, in other examples, we found that aliasing effects inherent to the numerical Wigner–Ville transform lead to artifacts as well.

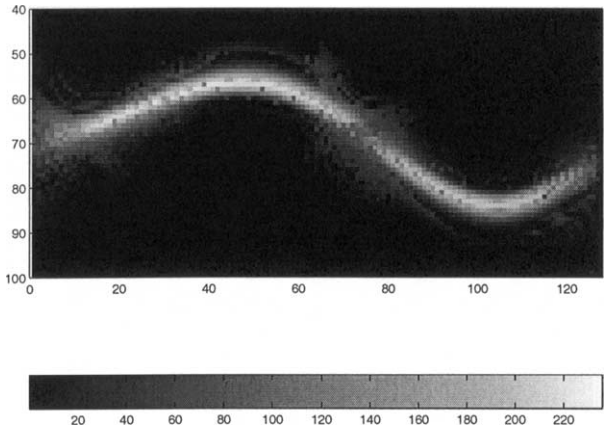

Fig. 13. Hessian–Gabor transform of $s(j)$.

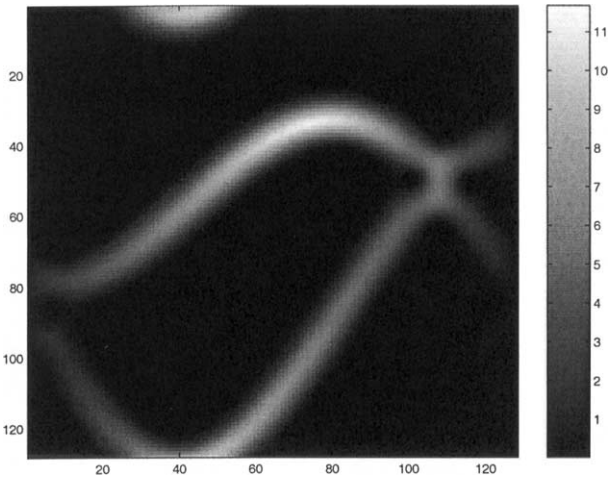

Fig. 14. Gabor transform of $s_2(j)$.

Let us consider a more complicated example,

$$s_2(j) = \exp\bigl(2\pi i\bigl(2(j-10)\bigr)\bigl(2+15\cos(\pi 1.2j/n)/n\bigr)\bigr) + \exp\bigl(2\pi i\bigl(2(j-n-60)\bigr)\bigl(2+15\cos(\pi 1.2j/n)/n\bigr)\bigr). \tag{7}$$

Figures 14, 15, 16 show the standard Gabor transform, the Wigner–Ville, and the adaptive transforms of s_2, respectively; Fig. 17 also shows the adaptive transform of the same signal to which white noise with SNR of 4 dB has been added. It turns out that this is very similar to the standard Gabor signal of this noisy signal, except for a slight increase in dynamic range.

Although the adaptive smoothing proposed here seems more "physical" than the simple selection of a the maximal Gabor transform of the last section, it leaves much to be desired, as shown by these examples. The main reason lies in the high-amplitude oscillation in the Wigner–Ville transform, due to the interference terms; in addition, there is also some instability due to the aliasing terms present in the numerical Wigner–Ville transform. As a result, our parameter computation is unreliable for noisy signals. We therefore present a more stable variation on this approach in the next subsection.

2.3. Variation on the curvature

We propose here to determine the parameters of the adaptive Gabor window from the curvature of a time-frequency representation of the signal that would be more stable than the discretized Wigner–Ville transform used above, which presents too many large local

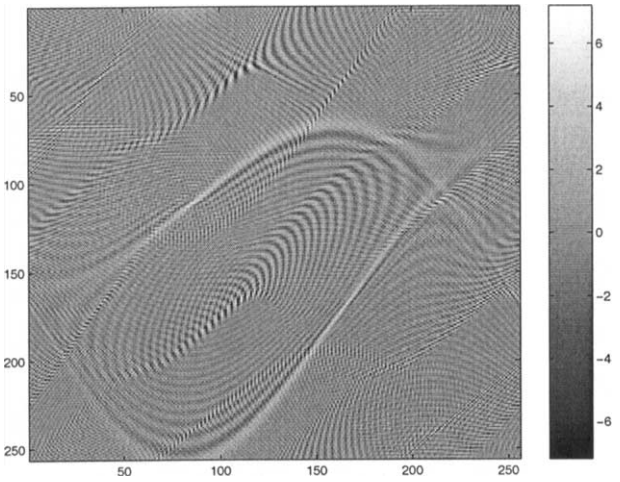

Fig. 15. Wigner–Ville transform of $s_2(j)$.

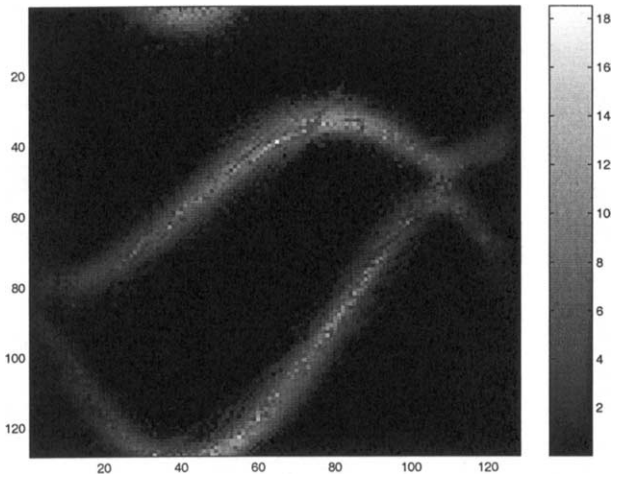

Fig. 16. Hessian–Gabor transform of $s_2(j)$.

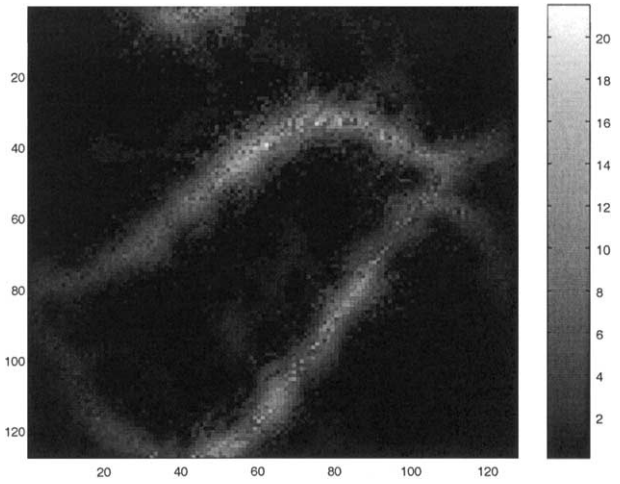

Fig. 17. Hessian–Gabor transform of noisy $s_2(j)$.

variations, already for clean signals, and even more so for noisy signals, as shown by the examples above. One easy choice for this more stable representation, used as a basis from which to compute the adaptive parameters, would be the most simple Gabor transform itself (using a regular Gaussian window). However, since we compute several Gabor transforms anyway (for reasons of computational cost: rather than computing different convolutions at different locations, we compute the different Gabor transforms globally, on the whole time-frequency plane, which allows us to use the FFT), we might as well take the maximum transform and use that as a basis for the choice of parameters. Another rationale for this approach is the following: in a sense, we are trying to focus on possible ridges in the representation. There exist several methods to detect such ridges in phase space representations (see [5]), not necessarily connected to the signal but fairly general to image processing. The difference between these approaches and ours is that our "focusing" method is inspired by the time-frequency nature of our "images." The Wigner–Ville based method of the last subsection breaks down for relatively low noise levels. We hope to achieve that the breaking point for our technique would occur only at much higher noise levels; as observed above, the maximum Gabor transform of the previous section resembles a Wigner transform with reduced interference effects that holds its "shape" better when noise is added. We thus propose to use the Max–Gabor transform: at every point (p, q) of interest, we shall determine its local curvature parameters, and then use these for an adaptive Gabor transform.

Let us illustrate this on the example we used before. Figures 18 and 19 show the Max–Gabor transform of the example (7), first without noise (Fig. 18) and then with noise (Fig. 19). Next, we show in Figs. 20 and 21 the adaptive transforms computed by using

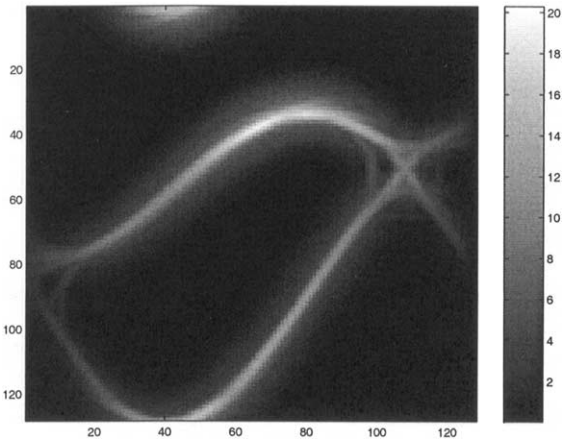

Fig. 18. Max–Gabor transform of $s_2(j)$.

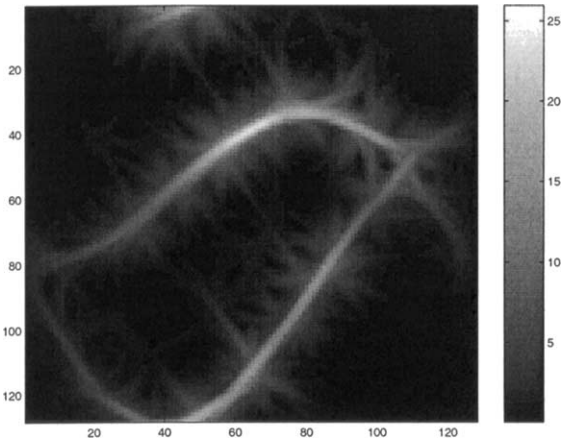

Fig. 19. Max–Gabor transform of noisy $s_2(j)$.

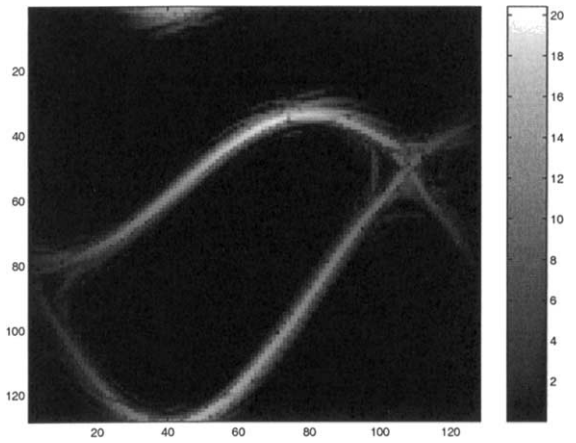

Fig. 20. Adaptive transform of $s_2(j)$ with adaptive parameters computed from the Max–Gabor transform.

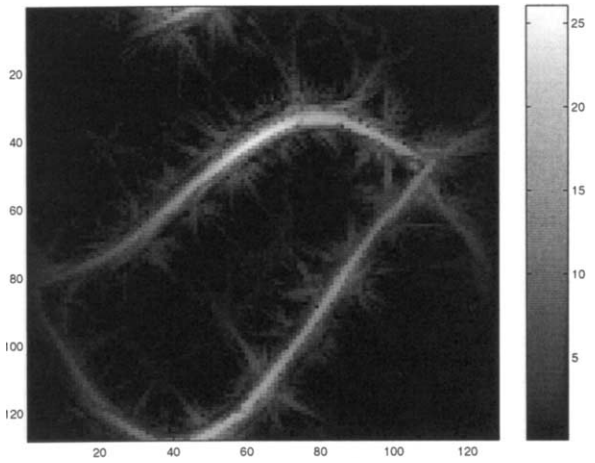

Fig. 21. Adaptive transform of noisy $s_2(j)$, with adaptive parameters computed from the Max–Gabor transform.

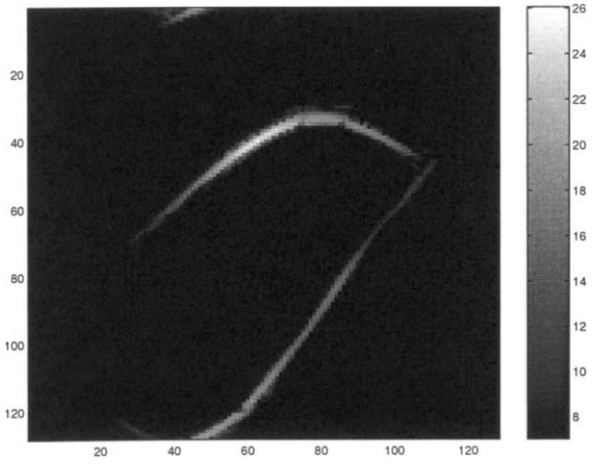

Fig. 22. Thresholded adaptive transform of noisy $s_2(j)$.

window parameters computed from the curvature of the Max–Gabor transforms, for the "clean" and the "noisy" versions of the signal, respectively.

At this point, we can further improve the readability of the method by performing various tricks. For instance, we can threshold the adaptive transform; an example is given in Fig. 22, which shows the effect of a simple threshold (we kept every coefficient larger than $\frac{1}{4}$ of the maximum) on the noisy adaptive transform of $s_2(j)$. One could also threshold the

maximum transform (or the $(1, 0)$-Gabor transform) from which the curvature parameters are computed, so that when there is little energy, we do not try to adjust parameters in parts of the time-frequency domain where there will be little variation anyway. This can be refined by using a grid of neighbors. All these tricks make the edges more apparent, and remove most of the "star-shaped" artifacts.

2.4. Reallocation and gradient methods

Reallocation methods have been developed by various authors: in [6], reallocation of coefficients of a continuous wavelet transform along "vertical" lines (i.e., to different p values, with same q) is proposed. This method provides good results, particularly for voice signals. One drawback is that reallocation is only performed along the scales, at fixed time. On the other hand, such a limited transform allows for a reconstruction formula. A more general method for reallocation in any direction of the phase space has been developed in [11]. It starts by computing a Gabor transform; next, it determines the logarithmic gradient of the transform; finally it "reallocates" the transform in the direction of this vector. There is no direct reconstruction formula. Such a method performs very well on a clean signal (without noise), that has several clearly separated components. It becomes less stable for noisy signals. The use of the gradient of a time-frequency transform in these reallocation methods inspired us to try out whether gradient information could be used instead of curvature information to determine the parameters of the adaptive Gabor transform; we explore this idea in this subsection. More specifically, we first compute the Gabor transform of the signal; then we compute its gradient at (p, q); next we use this to determine the (β, δ) parameters adapted to (p, q), in such a way that the corresponding ellipsoid's long axis is perpendicular to the gradient, and its eccentricity increases as the gradient does. Another variation would follow the same procedure, starting from the Max–Gabor transform rather than the Gabor transform. The next figures show the results of this approach, for the clean signal, with parameters computed from the $(1, 0)$ gradient in Fig. 23, and then from the Max–Gabor transform in Fig. 24. Notice that there is practically no difference between the two. Where the gradient is zero, however (on the ridge), the selected β, δ is the standard $(0, 1)$, leading to lower values than nearby, which seems counterproductive. However, all in all, if we work with a clean one component signal, things are going well, as illustrated on the previous figures. While this technique performs very well on clean signals, it turns out, however, that it performs poorly as soon as we add noise. Figures 25 and 26 show the same as Figs. 23 and 24 but for the noisy signal; Fig. 26 (using the Max–Gabor as starting point) is not quite as bad as Fig. 25, but it is still not as good as Fig. 21 or Fig. 22. Figures 26–29 show the same for representations (the standard Gabor transform, the adaptive Gabor transform with parameters computed from the gradient of the standard Gabor

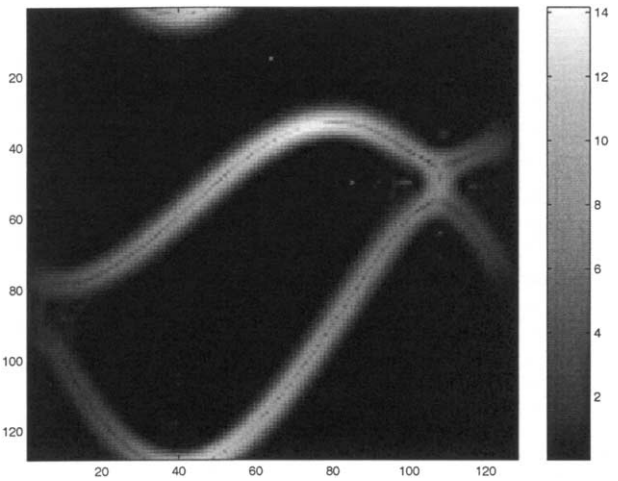

Fig. 23. Gradient $(0, 1)$-Gabor transform.

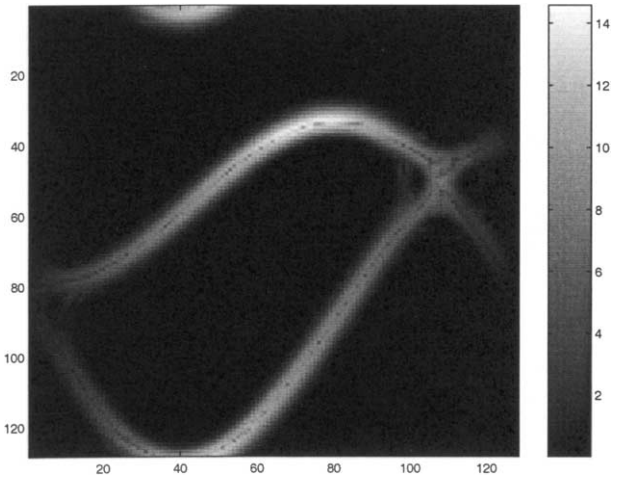

Fig. 24. Gradient Max–Gabor transform.

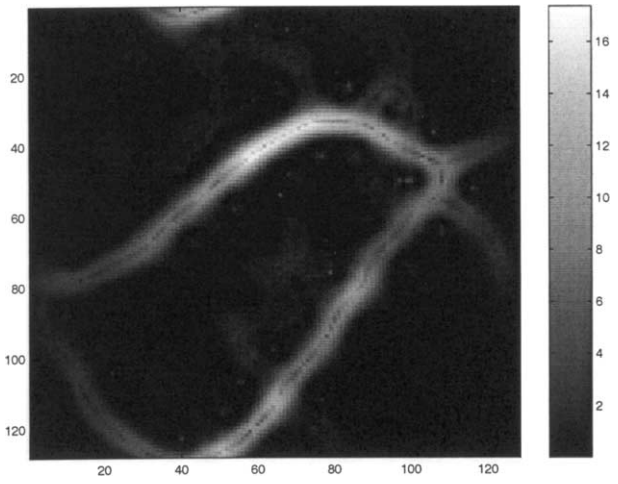

Fig. 25. Gradient $(0, 1)$-Gabor transform of noisy $s_2(j)$.

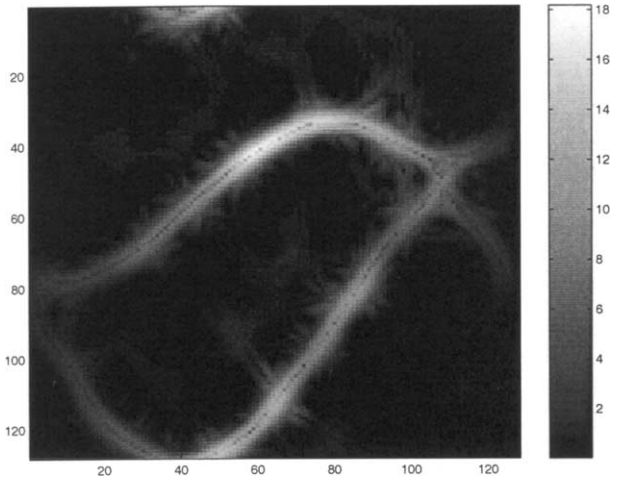

Fig. 26. Gradient Max–Gabor transform of noisy $s_2(j)$.

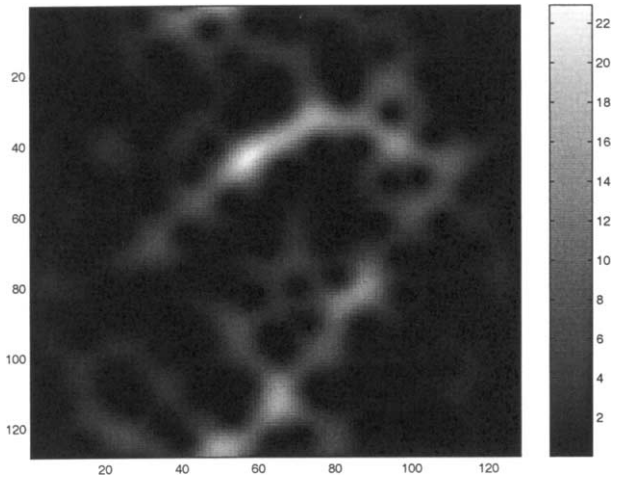

Fig. 27. Gabor transform of really noisy s_2.

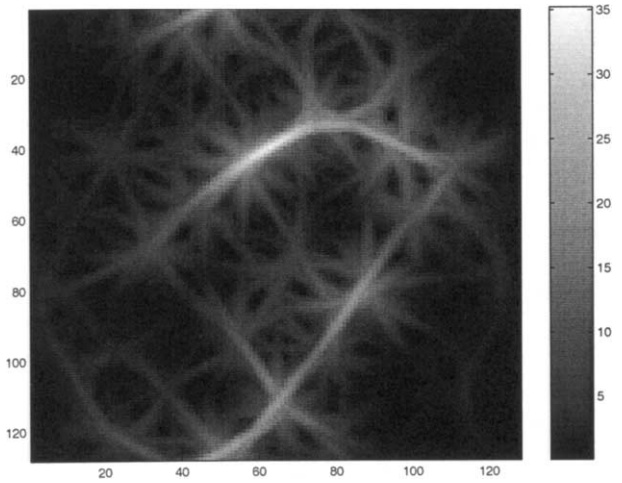

Fig. 28. Max–Gabor transform.

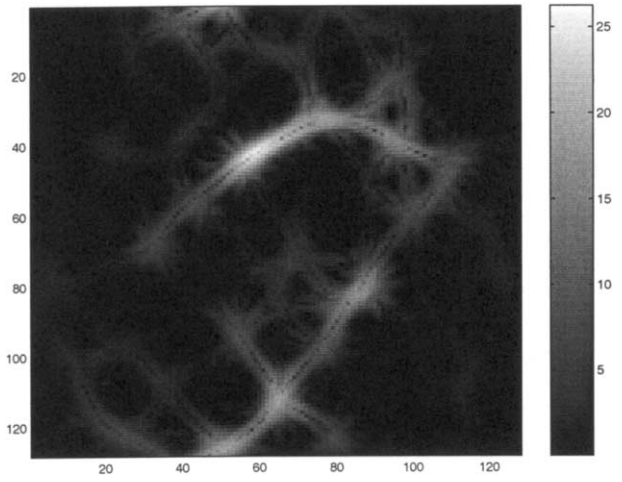

Fig. 29. Gradient Max–Gabor transform.

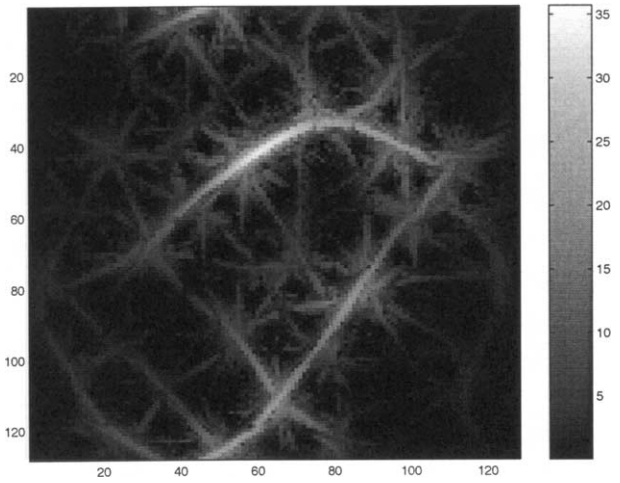

Fig. 30. Adaptive Gabor transform from Hessian of the Max–Gabor transform.

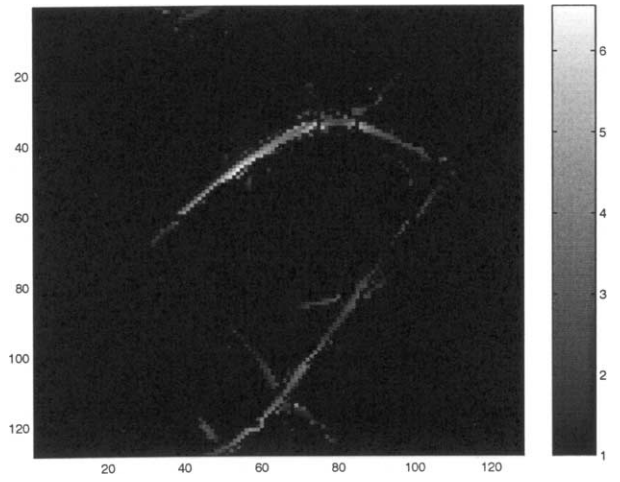

Fig. 31. Thresholded-adaptive Gabor transform.

transform, the Max–Gabor transform, and the adaptive Gabor transform computed from the gradient of the Max–Gabor transform) for the same signal, now with white noise added with SNR 4 dB.

It appears that using the Hessian rather than the gradient (of any transform) to obtain parameters seems to be more noise resistant. In Figs. 27–30 we double the noise level (reaching SNR 0 dB) as compared with Figs. 25 and 26, illustrating the better stability of the Hessian method. Finally, Fig. 29 shows a modification of Fig. 30 where we have thresholded based on comparison between the adaptive transform and the Max–Gabor transform from which the parameters for the adaptive transform were obtained.

2.5. A simple real life example

We present an example of our analysis on a bat sonar signal, borrowed from the Time-Frequency Toolbox by Auger et al. [3], see also [7]. This signal was recorded with a 230.4 kHz sampling frequency, in the 8–80 kHz range. It fits nicely within the kind of signals that can be successfully studied using the adaptive methods we introduce. The first figure presents the clean signal, while the second figure is its Gabor transform with the standard Gaussian window, then we add noise with an SNR of 0 dB, and compute both the usual Gabor transform and the adaptive Gabor transform (based on the analysis of the Max–Gabor transform). This last two transforms have been thresholded to keep only the part relevant to the signal. We remark the adaptive transform is well-localized despite the relatively high noise level.

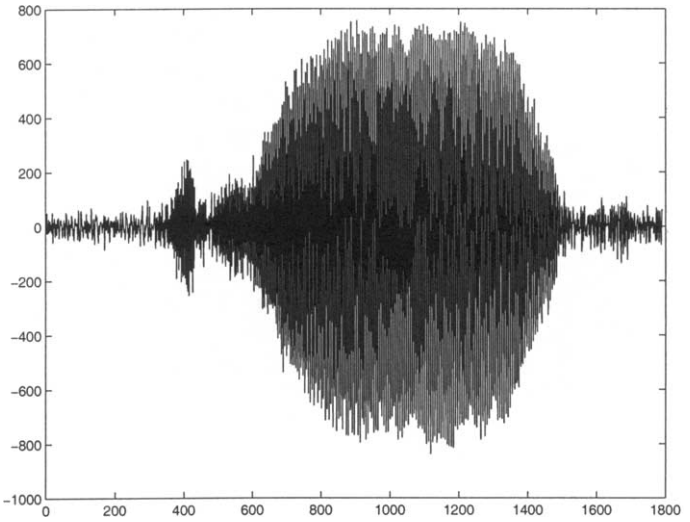

Fig. 32. Bat sonar signal.

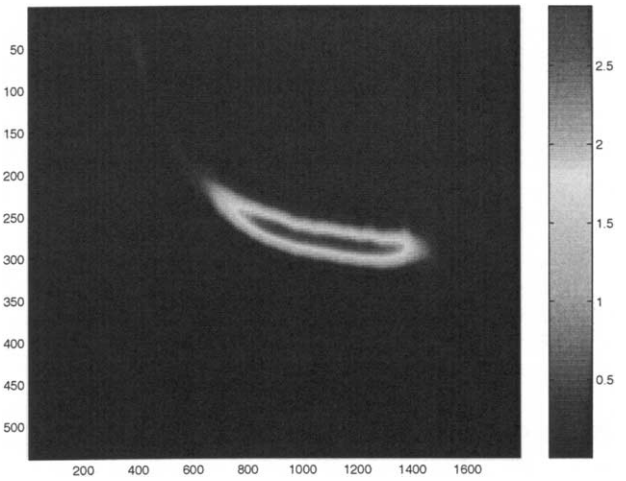

Fig. 33. Gabor transform for the bat sonar signal.

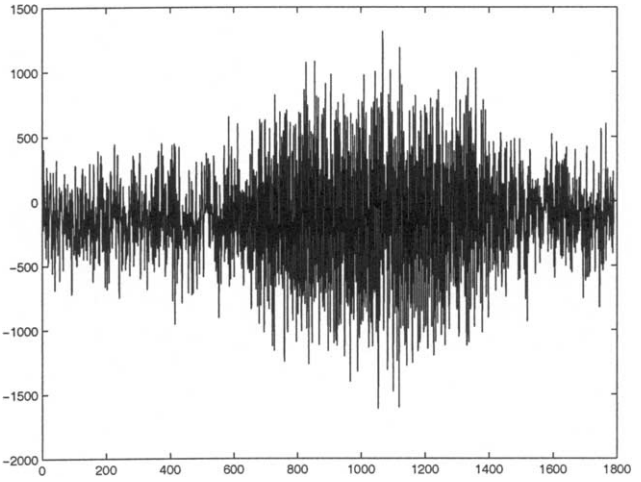

Fig. 34. Noisy bat sonar signal.

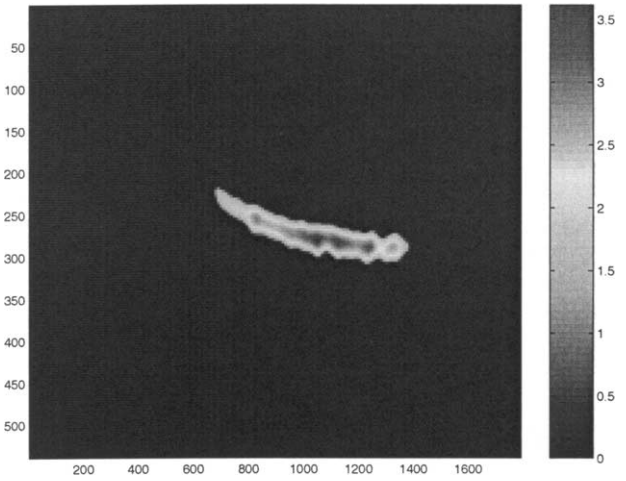

Fig. 35. (Thresholded) Gabor transform for the noisy bat sonar signal.

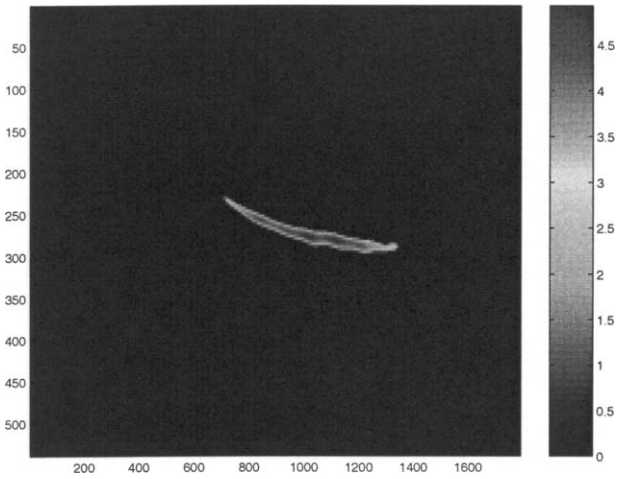

Fig. 36. (Thresholded) adaptive Gabor transform for the noisy bat sonar signal.

3. Reconstruction

It is well known that, for a fixed couple (β, δ), the Gabor transform defined by (4) can be inverted by the following formula:

$$f(x) = \int_{p,q} G^{\beta,\delta}(f)(p,q) g^{\tilde{\beta},\tilde{\delta}}_{p,q}(x)\, \mathrm{d}p\, \mathrm{d}q. \tag{8}$$

The choice of windows for the reconstruction being arbitrary, one can chose a Dirac window in order to integrate only over frequencies

$$f(x) = \int_{p} G^{\beta,\delta}(f)(p,x) e^{ipx - pq/2}\, \mathrm{d}p. \tag{9}$$

It seems very tempting to apply the same strategy to our adaptive Gabor transform. Theoretically, this should fail, except for a few particular cases. (One such particular case is when (β, δ) depend only on q. Then essentially the previous formula still apply.)

Another interesting remark has to do with formula (8). If we wish to use the same formula with (β, δ) dependent on (p, q), then one essentially obtains a pseudo-differential operator, which has the following Weyl symbol (up to constants):

$$\sigma(x,\xi) = \int_{p,q} e^{-(2\delta+\beta^2/2\delta)(x-q)^2 - (1/2\delta)(\xi-p)^2 + (\beta/\delta)(x-q)(\xi-p)}\, \mathrm{d}p\, \mathrm{d}q. \tag{10}$$

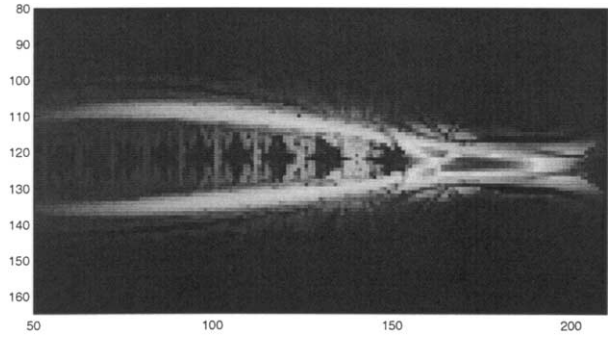

Fig. 37. Adaptive Gabor transform of signal s_3.

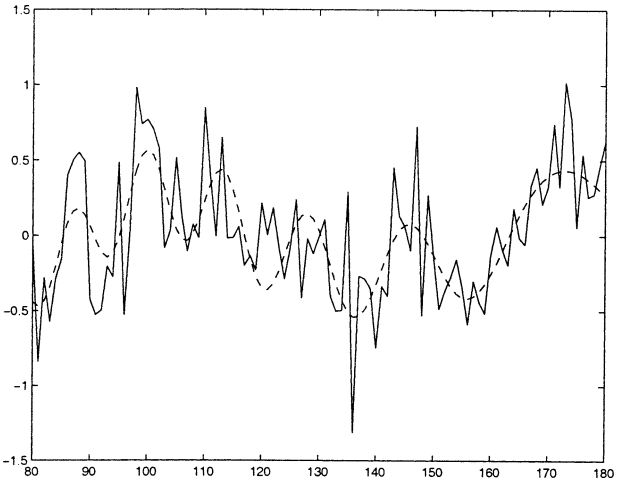

Fig. 38. (Real part of) signal s_3.

From this formula it is clear why a fixed choice of (β, δ) leads to the identity for the associated operator. If (β, δ) varies with (p, q), then the pseudo-differential operator can be viewed as a perturbation of the identity, but unfortunately for our needs this won't be sufficient.

For some signals however, which consist of well-defined components that are fairly well-separated in the time-frequency plane, we can argue heuristically that for each component we can apply a reconstruction formula with $(\beta, 0)$ depending only on q; because the components are well-localized in the time-frequency plane, we can do this simultaneously for the different components, each with their own q-varying (β, δ) localization. This suggests a reconstruction formula that naively applies (9) even for our adaptive (β, δ) depending on both p and q. Let us try this out in an example. We consider a signal s_3, which is the sum of two simple slowly varying chirps, as illustrated on Fig. 37, which is its adaptive Gabor transform.

The (real part of the) signal we are considering appears on Fig. 38, along with the same signal to which we added noise (with SNR of 6 dB).

Next, on Fig. 39 we have plotted the reconstructed signal from a thresholded Gabor transform: we kept 3% of the coefficients. This is to be compared with Fig. 40, on which we have the reconstructed signal from the adaptive Gabor transform, with the same threshold for the coefficients. Notice how we still can see the generic shape of the signal. Of course performing such a reconstruction on a clean signal, unlike for the true Gabor transform, will not reconstruct perfectly, as illustrated on Fig. 41. Various scaling effects observed on the reconstructions from adaptive transforms have to do with the non-conservation of the energy by the adaptive transform. Still, it is striking that we still get close to the signal, especially with noise.

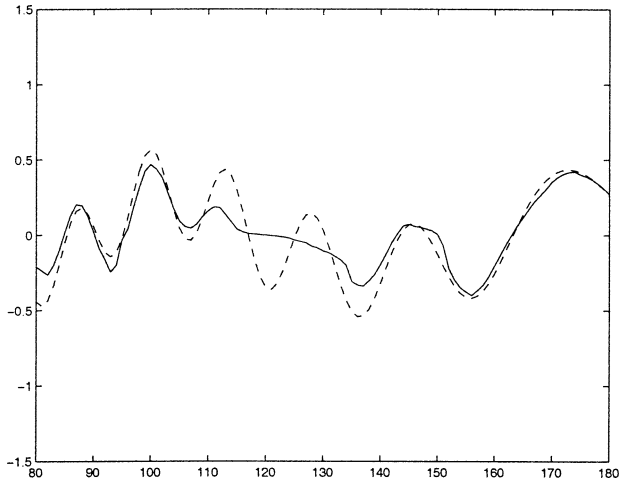

Fig. 39. Reconstructed s_3 from noisy Gabor transform.

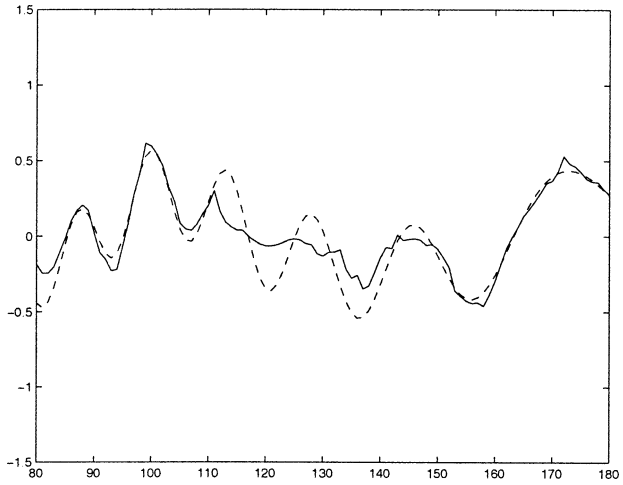

Fig. 40. Reconstructed s_3 from noisy adaptive transform.

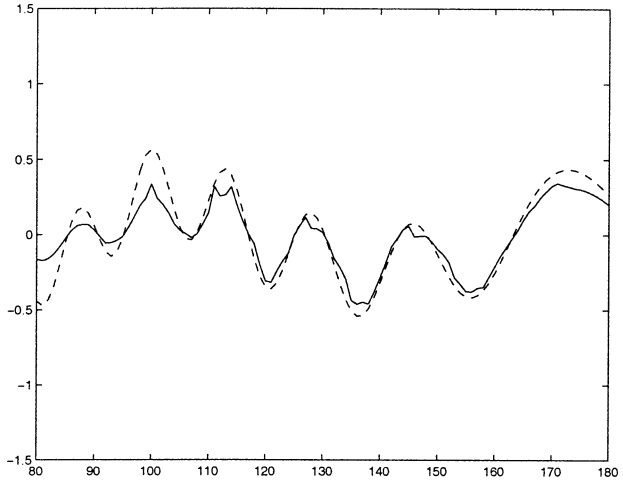

Fig. 41. Reconstructed s_3 from clean adaptive transform.

Acknowledgments

Most of this work was conducted while the second author was visiting the Program in Applied and Computational Mathematics at Princeton University, supported by ONR grant N00014-96-1-0367 and AFOSR grant F49620-98-1-0044.

References

[1] A. Akan, L. Chaparro, Evolutionary spectral analysis using a warped Gabor expansion, in: IEEE ICASSP96, Atlanta, 1996.
[2] J.C. Andrieux,, et al., Optimum smoothing of the Wigner–Ville distribution, IEEE Trans. Signal Process. ASSP-35 (6) (1987) 764–769.
[3] F. Auger, P. Flandrin, P. Gonçalvès, O. Lemoine, Time-Frequency Toolbox, http://www.isis.enst.fr/Applications/tftb/iutsn.univ-nantes.fr/auger/tftb.html.
[4] R.G. Baraniuk, D.L. Jones, A signal-dependent time-frequency representation: Optimal kernel design, IEEE Trans. Signal Process. SP-41 (4) (1993) 1589–1602.
[5] R. Carmona, W.-L. Hwang, B. Torrésani, Practical Time-Frequency Analysis, Academic Press, San Diego, CA, 1998, Gabor and wavelet transforms with an implementation in S.
[6] I. Daubechies, S. Maes, A non-linear squeezing of the continuous wavelet transform based on auditory nerve models, in: Wavelets in Medicine and Biology, CRC Press, 1996, Chap. 20.
[7] P. Flandrin, On detection-estimation procedures in the time-frequency plane, in: International Conference on Acoustics, Speech, and Signal Processing, 1886, pp. 43.5.1–43.5.4.
[8] P. Flandrin, Temps-Fréquence, Hermes, 1993.
[9] P. Flandrin, B. Escudié, Sur la localisation des représentations conjointes dans le plan temps-fréquence, C. R. Acad. Sci. Paris 295 (1982) 475–478.
[10] D. Jones, T. Parks, A high reslution data-adaptive time-frequency representation, IEEE Trans. Signal Process. ASSP-38 (1990) 2127–2135.
[11] P. Auger, P. Flandrin, Improving the readibility of time-scale and time-frequency representations by using the reassignement method, IEEE Trans. Signal Process. (1995).
[12] F. Peyrin, R. Prost, A unified definition for the discrete-time, descrete-frequency and discrete-time/frequency Wigner distributions, IEEE Trans. Acoust. Speech. Signal Process. ASSP-34 (4) (1986) 858–867.
[13] L.J. Stanković, V. Katkovnik, The Wigner–Ville distribution of noisy signals with adaptive time-frequency varying window, IEEE Trans. Signal Process. 47 (4) (1999) 1099–1108.