Implementation of a diffusive differential reassignment method for signal enhancement. An application to wolf population counting [1]

B. Dugnol, C. Fernández, G. Galiano [2] and J. Velasco

Dpto. de Matemáticas, Universidad de Oviedo, c/ Calvo Sotelo, 33007-Oviedo Spain

Abstract

We investigate the combination of two PDE-based image processing techniques applied to the image produced by time-frequency representations of one-dimensional signals, such as the spectrogram. Specifically, we consider the energy transport equation associated to the lagrangian coordinates corresponding to the spectrogram differential reassignment proposed by Chassandre-Mottin, Daubechies, Auger and Flandrin as a spectrogram readability improving method, together with the image restoration model proposed by Álvarez, Lions and Morel for noise reduction and edge enhancement. Our aim is to produce a transformation of the spectrogram in which the instantaneous frequency lines are easier to track, for using it as an input for a (wolves howls) counting algorithm.

After presenting the model derivation, we show some analytical properties of it, such as the existence of a unique solution and a comparison principle, and perform later a discretization to numerically investigate its performance for the cases of synthetic signals and field recorded wolves choruses. We finally compare our results with those obtained from well established techniques.

Key words: Spectrogram, differential reassignment, image restoration, partial differential equation, instantaneous frequency, population counting.

[1] The first three authors are supported by Project PC0448, Gobierno del Principado de Asturias, Spain. Third and fourth authors are supported by the Spanish DGI Project MTM2004-05417.

[2] Corresponding author. Phone:+34 985182299 Fax: +34 985103354

[3] dugnol@uniovi.es, carlos@uniovi.es, galiano@uniovi.es, julian@uniovi.es

1 Introduction

Wolf is a protected specie in many countries around the world. Due to their predator character and to their proximity to human settlements, wolves often kill cattle interfering in this way in farmers' economy. To smooth this interference, authorities reimburse the cost of these lost to farmers. Counting the population of wolves inhabiting a region is, therefore, not only a question of biological interest but also of economic interest, since authorities are willing to estimate the budget devoted to costs produced by wolf protection, see for instance [1]. However, estimating the population of wild species is not an easy task. In particular, for mammals, few and not very precise techniques are used, mainly based on the recuperation of field traces, such as steps, excrements and so on. Our investigation is centered in what it seems to be a new technique, based on signal and image theory methods, to estimate the population of species which fulfill two conditions: living in groups, for instance, packs of wolves, and emitting some characteristic sounds, howls and barks, for wolves. The basic initial idea is to produce, from a given recording, some time-frequency distribution which allows to identify the different howls corresponding to different individuals by estimating the instantaneous frequency (IF) lines of their howls.

Unfortunately, the real situation is somehow more involved due mainly to the following two factors. On one hand, since natural sounds, in particular wolf howling, are composed by a fundamental pitch and several harmonics, direct instantaneous frequency estimation of the multi-signal recording leads to an over-counting of individuals since various IF lines correspond to the same individual. Therefore, more sophisticated methods are indicated for the analysis of these signals, methods capable of extracting additional information such as the slope of the IF, which allows to a better identification of the harmonics of a given fundamental tone. Chirplet type transforms [2,3] or the use of the Fourier fractional transform [4] are possibilities which are under current development [5]. On the other hand, despite the quality of recording devices, field recordings are affected for a variety of undesirable signals which range from low amplitude broad spectrum long duration signals, like wind, to signals localized in time, like cattle bells, or localized in spectrum, like car engines. Clearly, the addition of all these signals generates an unstructured noise in the background of the wolves chorus which impedes the above mentioned methods to work properly.

In a previous work [6], we investigated the noise reduction and edge (IF lines) enhancement on the spectrogram image by a PDE-based image processing algorithm. Considering the spectrogram as the initial image, the algorithm induces isotropic diffusion (noise smoothing) in regions with low gradient values, and anisotropic diffusion (edge-IF enhancement) in regions with high gradient

values. We showed several numerical demonstrations applied to both synthetic and field recorded signals confirming a good performing of the algorithm. Moreover, in [7] we performed comparisons with several denoising techniques such as the nonlinear spectral substraction method and the 2D-stationary wavelet transform, showing that although the PDE method is more expensive in terms of time execution it is also more stable under threshold values perturbations in the denoised images. However, we found two main difficulties which motivated the present work. In one hand, the broadening effect of the spectrogram when compared to the WV distribution is still present if not slightly increased by this diffusive transformation. On the other hand, there is not a natural *stopping time* for the evolution of the algorithm, being this to be found by a trial and error mechanism, preventing therefore from getting a fully automatic algorithm.

In this article we address these difficulties partly by using the ideas introduced in [8] to improve the readability of the spectrogram through the so-called reassignment techniques, later developed in [9–11]. Let $x \in L^2(\mathbb{R})$ denote an audio signal and consider the Short Time Fourier transform (STFT)

$$\mathcal{G}_\varphi(x;t,\omega) = \int_{\mathbb{R}} x(s)\varphi(s-t)e^{-i\omega s}ds, \tag{1}$$

corresponding to the real, symmetric and normalized *window* $\varphi \in L^2(\mathbb{R})$. The energy density function or *spectrogram* of x corresponding to the window φ is given by

$$S_\varphi(x;t,\omega) = |\mathcal{G}_\varphi(x;t,\omega)|^2, \tag{2}$$

which may be expressed also as [12]

$$S_\varphi(x;t,\omega) = \int_{\mathbb{R}^2} WV(\varphi;\tilde{t},\tilde{\omega})WV(x;t-\tilde{t},\omega-\tilde{\omega})d\tilde{t}d\tilde{\omega}, \tag{3}$$

with $WV(y;\cdot,\cdot)$ denoting the Wigner-Ville distribution of $y \in L^2(\mathbb{R})$,

$$WV(y;t,\omega) = \int_{\mathbb{R}} y(t+\frac{s}{2})y(t-\frac{s}{2})e^{-iws}ds.$$

The Wigner-Ville (WV) distribution has received much attention for IF estimation due to its excellent concentration and many other desirable mathematical properties, see [12]. However, it is well known that it presents high amplitude sign-varying cross-terms for multi-component signals which makes its interpretation difficult, see Fig. 1. Expression (3) represents the spectrogram as the convolution of the WV distribution of the signal, x, with the smoothing kernel defined by the WV distribution of the window, φ, explaining the mechanism of attenuation of the cross-terms interferences in the spectrogram. However, an important drawback of the spectrogram with respect to the WV distribution is the broadening of the IF lines as a direct consequence of the smoothing convolution. To override this inconvenient, it is suggested in [8]

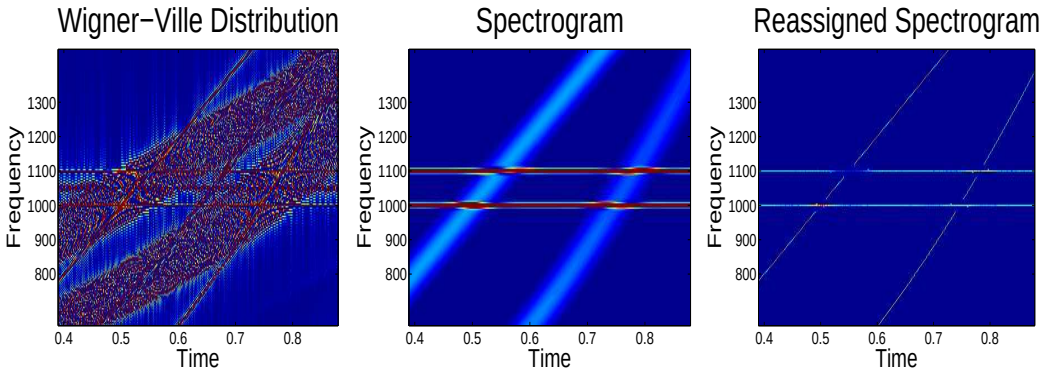

Fig. 1. Detail of the Wigner-Ville distribution, spectrogram and reassigned spectrogram corresponding to (clean) signal x_1 of Experiment 1.

that instead of assigning the averaged energy to the geometric center of the smoothing kernel, (t, ω), as it is done for the spectrogram, one assigns it to the *center of gravity* of these energy contributions, $(\hat{t}, \hat{\omega})$, which is certainly more representative of the local energy distribution of the signal, see Fig. 1. As deduced in [9], the gravity center may be computed by the following formulas

$$\hat{t}(x;t,\omega) = t - \Re\left\{\frac{\mathcal{G}_{T\varphi}(x;t,\omega)}{\mathcal{G}_{\varphi}(x;t,\omega)}\right\}, \quad \hat{\omega}(x;t,\omega) = \omega + \Im\left\{\frac{\mathcal{G}_{D\varphi}(x;t,\omega)}{\mathcal{G}_{\varphi}(x;t,\omega)}\right\},$$

where the STFT's windows in the numerators are $T\varphi(t) = t\varphi(t)$ and $D\varphi(t) = \varphi'(t)$. The reassigned spectrogram, $RS_\varphi(x;t,\omega)$, is then defined as the aggregation of the reassigned energies to their corresponding locations in the time-frequency domain

$$RS_\varphi(x;t,\omega) = \frac{1}{2\pi}\int_{\mathbb{R}^2} S_\varphi(x;t',\omega')\delta(t-\hat{t}(t',\omega'))\delta(\omega-\hat{\omega}(t',\omega'))dt'd\omega'. \quad (4)$$

Observe that energy is conserved through the reassignment process. Other desirable properties, among which non-negativity and perfect localization of linear chirps, are proven in [13] For our application, it is of special interest the fact that the reallocation vector, $\mathbf{r}(t,\omega) = (\hat{t}(t,\omega) - t, \hat{\omega}(t,\omega) - \omega)$, may be expressed through a potential related to the spectrogram [11],

$$\mathbf{r}(t,\omega) = \frac{1}{2}\nabla \log(S_\varphi(x;t,\omega)), \quad (5)$$

when φ is a Gaussian window of unit variance. Let $\tau \geq 0$ denote an artificial time and consider the dynamical expression of the reassignment given by $\Phi(t,\omega,\tau) = (t,\omega) + \tau\mathbf{r}(t,\omega)$ which, for $\tau = 0$ to $\tau = 1$, connects the initial point (t,ω) with its reassigned point $(\hat{t},\hat{\omega})$. Rewritting this expression as

$$\frac{1}{\tau}(\Phi(t,\omega,\tau) - \Phi(0,\omega,\tau)) = \mathbf{r}(t,\omega),$$

and taking the limit $\tau \to 0$, we may identify the displacement vector $\mathbf{r}$ as the velocity field of the transformation Φ. In close relation with this approach is

the process referred to as *differential reassignment* [11], defined as the transformation given by the (autonomous) dynamical system corresponding to such velocity field,

$$\begin{cases} \dfrac{d\chi}{d\tau}(t,\omega,\tau) = \mathbf{r}(\chi(t,\omega,\tau)), \\ \chi(t,\omega,0) = (t,\omega), \end{cases} \tag{6}$$

for $\tau > 0$. Observe that, in a first order approximation, we still have that χ connects (t,ω) with some point in a neighborhood of $(\hat{t},\hat{\omega})$, since

$$\chi(t,\omega,1) \approx \chi(t,\omega,0) + \mathbf{r}(\chi(t,\omega,0)) = (t,\omega) + \mathbf{r}(t,\omega) = (\hat{t},\hat{\omega}).$$

In addition, for $\tau \to \infty$, each particle (t,ω) converges to some local extremum of the potential $\log(S_\varphi(x;\cdot,\cdot))$, among them the maxima and ridges of the original spectrogram. The conservative energy reassignation analogous to (4) for the differential reassignment is obtained by solving the following problem for $u(t,\omega,\tau)$ and $\tau > 0$,

$$\frac{\partial u}{\partial \tau} + \operatorname{div}(u\mathbf{r}) = 0, \tag{7}$$

$$u(\cdot,\cdot,0) = u_0, \tag{8}$$

where we introduced the notation $u_0 = S_\varphi(x;\cdot,\cdot)$ and, consequently, $\mathbf{r} = \frac{1}{2}\nabla \log(u_0)$. Since in applications both signal and spectrogram are defined in bounded domains, we assume (7)-(8) to hold in a bounded time-frequency domain, Ω, in which we assume non energy flow conditions on the solution and the data

$$\nabla u \cdot \mathbf{n} = 0, \quad \mathbf{r} \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega \times \mathbb{R}_+, \tag{9}$$

being $\mathbf{n}$ the unitary outwards normal to $\partial\Omega$. Finally, observe that the positivity of the spectrogram [12] and the fact that it is obtained from a convolution with a C^∞ kernel implies the regularity $u_0, \mathbf{r} \in C^\infty$ and, therefore, problem (7)-(9) admits a unique smooth solution.

As noted in [11], differential reassignment can be viewed as a PDE based processing of the spectrogram image in which the energy tends to concentrate on the initial image ridges (IF lines). As mentioned above, our aim is not only to concentrate the diffused IF lines of the spectrogram but also to attenuate the noise present in our recordings. It is clear that noise may distort the reassigned spectrogram due to the change of the energy distribution and therefore of the gravity centers of each time-frequency window. Although even a worse situation may happen to the differential reassignment, due to its convergence to spectrogram local extrema (noise picks among them) an intuitive way to correct this effect comes from its image processing interpretation. As shown in [6], when a strong noise is added to a clean signal better results are obtained for approximating the clean spectrogram if we use a noise reduction edge

enhancement PDE based algorithm than if we simply threshold the image spectrogram. This is due to the local application of gaussian filters in regions of small gradients (noise, among them) while anisotropic diffusion (in the orthogonal direction to the gradient) is applied in regions of large gradients (edges-IF lines). Therefore, a possible way to improve the image obtained by the differential reassigned spectrogram is to modify (7) by adding a diffusive term with the mentioned properties.

Let us make a final observation before writing the model we shall work with. In the derivation of both the reassigned and the differential reassigned spectrogram the property of energy conservation is imposed, implying that energy values on ridges increase. Indeed, let B be a neighborhood of a point of maximum for u_0, in which $\operatorname{div} \mathbf{r} = \Delta \log u_0 < 0$, and let $(t_0, \omega_0) \in B$. Let $\chi_0(t, \omega, \tau)$ denote the characteristic defined by (6) starting at (t_0, ω_0). Evaluating Eq. (7) along χ_0 we obtain

$$\frac{d}{d\tau} u = \frac{\partial u}{\partial \tau} + \mathbf{r} \cdot \nabla u = -u \operatorname{div} \mathbf{r}, \tag{10}$$

implying that u experiments exponential increase in B. For image processing, it is desirable the maximum principle to hold, i.e., that the bounds $\min u_0 \leq u \leq \max u_0$ hold for any $(t, \omega, \tau) \in \Omega \times \mathbb{R}_+$, ensuring that the processed image is within the range of image definition ($[0, 255]$, usually). A simple way, which we shall address, to ensure this property is by dropping the right hand side term of Eq. (10), i.e., replacing Eq. (7) by the transport equation

$$\frac{\partial u}{\partial \tau} + \mathbf{r} \cdot \nabla u = 0. \tag{11}$$

However, no energy conservation law will apply anymore (note that u is constant along the characteristics).

The combination of the differential reassignment problem (8), (9) and (11) with the edge-detection image-smoothing algorithm [14] is written as

$$\frac{\partial u}{\partial \tau} + \frac{\varepsilon}{2} \nabla \log(u_0) \cdot \nabla u - g(|G_s * \nabla u|) A(u) = 0 \quad \text{in } \Omega \times \mathbb{R}_+, \tag{12}$$

$$\nabla u \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega \times \mathbb{R}_+, \tag{13}$$

$$u(0, \cdot, \cdot) = u_0 \quad \text{in } \Omega, \tag{14}$$

where $\varepsilon \geq 0$, see Remark 1, and

$$A(u) = (1 - h(|\nabla u|))\Delta u + h(|\nabla u|) \sum_{j=1,\ldots,n} f_j\left(\frac{\nabla u}{|\nabla u|}\right) \frac{\partial^2 u}{\partial x_j^2}.$$

Let us briefly remind the properties and meaning of the diffusive term com-

ponents in equation (12):

- Function G_s is a Gaussian of variance s. The variance is a *scale parameter* which fixes the minimal size of the details to be kept in the processed image.
- Function g is non-increasing with $g(0) = 1$ and $g(\infty) = 0$. It is a *contrast* function, which allows to decide whether a detail is sharp enough to be kept.
- The composition of G_s and g on ∇u rules the speed of diffusion in the evolution of the image, controlling the *enhancement* of the edges and the noise smoothing.
- The diffusion operator A combines isotropic and anisotropic diffusion. The first smoothes the image by local averaging while the second enforces the diffusion only on the orthogonal direction to ∇u (along the edges). More precisely, for $\theta_j = (j-1) * \pi/n$, $j = 1, \ldots, n$ we define x_j as the orthogonal to the direction θ_j, i.e., $x_j = -t \sin\theta_j + \omega \cos\theta_j$. Then, smooth non-negative functions $f_j(\cos\theta, \sin\theta)$ are designed to be *active* only when θ is close to θ_j. Therefore, the anisotropic diffusion is taken in an approximated direction to the orthogonal of ∇u. The combination of isotropic and anisotropic diffusions is controlled by function $h(s)$, which is nondecreasing with

$$h(s) = \begin{cases} 0 \text{ for } s \leq \epsilon, \\ 1 \text{ for } s \geq 2\epsilon, \end{cases} \tag{15}$$

  being ϵ the *enhancement* parameter.

Remark 1 *A re-parametrization $\tau \to \tau/\varepsilon$ transforms Eq. (11) to the equivalent form $\partial u/\partial\tau + \varepsilon \mathbf{r} \cdot \nabla u = 0$. Parameter $\varepsilon > 0$ allows us to play with different balances between transport and diffusion effects.*

2 Mathematical properties

For $\varepsilon = 0$, the following theorem is proven in [14].

Theorem 1 *Let $u_0 \in W^{1,\infty}(\Omega)$ and $\varepsilon = 0$. (i) Then, for any $T > 0$, there exists a unique (viscosity) solution, $u \in C([0,\infty) \times \Omega) \cap L^\infty(0, T; W^{1,\infty}(\Omega))$, of problem (12)-(14). Moreover,*

$$\inf_\Omega u_0 \leq u \leq \sup_\Omega u_0 \quad \text{in } \mathbb{R}_+ \times \Omega. \tag{16}$$

(ii) Let v be a solution of problem (12)-(14) corresponding to the initial data $v_0 \in L^\infty(\Omega)$. Then, for all $T \geq 0$, there exists a constant K which depends only on $\|u_0\|_{W^{1,\infty}}$ and $\|v_0\|_{L^\infty}$ such that

$$\sup_{0 \leq \tau \leq T} \|u(\tau, \cdot, \cdot) - v(\tau, \cdot, \cdot)\|_{L^\infty(\Omega)} \leq K \|u_0 - v_0\|_{L^\infty(\Omega)}. \tag{17}$$

A straightforward modification of the proof of Theorem 1 given in [14], affecting only to the constant K that now, in addition, depends on $\|u_0\|_{W^{2,\infty}}$, allows us to produce a similar proof for probem (12)-(14) with $\varepsilon > 0$. We, therefore, omit the proof. Let us note that the solution ensured by this theorem is not, in general, a classical solution. The notion of solution employed in [14] is that of *viscosity solution*, which coincides with the classical solution if it is regular enough. Since we will not enter in further discussions about regularity, we refer the reader to [15] for technical details about this notion of solution.

Part (ii) of Theorem 1 is specially useful to us for the following reasons. Spectrograms of a signal are computed relative to windows, i.e, for each window a different spectrogram (image) is got. Then, the time-frequency characteristics of the signal, like instantaneous frequency, look in a slight different way if two different windows are employed. It, therefore, arises the question of stability of the final images with respect to the windows, i.e., is it possible that starting from two spectrograms of the same signal for different windows the corresponding final images are very different from each other? Another stability question solved with the help of Theorem 1 is whether the transformed spectrograms of two close signals, for instance of a signal x and a signal $x + \delta n$, where n denotes a noise and $\delta > 0$ is small, are close or not. We have the following result.

Corollary 1 *(1) Let $\varphi, \psi \in W^{1,\infty}(\mathbb{R})$ be real, symmetric and normalized windows and denote by u_0 and v_0, respectively, the corresponding spectrograms of a given signal $x \in L^2(\mathbb{R})$. Let u and v be the solutions of problem (12)-(14) corresponding to the initial data u_0 and v_0, respectively. Then*

$$\sup_{0\leq\tau\leq T} \|u(\tau,\cdot,\cdot) - v(\tau,\cdot,\cdot)\|_{L^\infty(\Omega)} \leq c\|\varphi - \psi\|_{L^2(\mathbb{R})},$$

where the constant c depends only on $\|x\|_{L^\infty}$ and K.

(2) Let $x, y \in L^2(\mathbb{R})$ be two signals and $\varphi \in W^{1,\infty}(\mathbb{R})$ be a real, symmetric and normalized window. Let u_0 and v_0 be, respectively, their spectrograms. Finally, let u and v be the solutions of problem (12)-(14) corresponding to the initial data u_0 and v_0, respectively. Then

$$\sup_{0\leq\tau\leq T} \|u(\tau,\cdot,\cdot) - v(\tau,\cdot,\cdot)\|_{L^\infty(\Omega)} \leq c\|x - y\|_{L^2(\mathbb{R})},$$

where c depends only on K.

We omit the proof, which is a straightforward modification of that given for Corollaries 1 and 2 in [6].

3 Numerical Experiments

In this section we present numerical demonstrations of the diffusive differential reassignment method applied to spectrograms of synthetic and field signals, and some comparisons with other methods. Computation of the spectrogram is a standard operation performed by applying the discrete fast fourier transform (dfft) to the convolution of the signal with the window, being the latter imposed by relation (5) between the displacement vector and the spectrogram, which only holds for a unit variance Gaussian window. To facilitate the comparisons with other methods, we used the Matlab ToolBox [16] for computing both the spectrogram and the reassigned spectrogram. Once the spectrogram is produced, it is normalized in the usual digital image range $[0, 255]$, obtaining in this way the initial datum and transport term for problem (12)-(14). We use a time semi-implicit Euler scheme of the form

$$\frac{u^{k+1}-u^k}{\delta\tau} = -\frac{\varepsilon}{2}\nabla \log u_0 \cdot \nabla u^k + g(|G_s * \nabla u^k|)\Big(1 - h(|\nabla u^k|)\Delta u^{k+1} + h(|\nabla u^k|)\sum_{j=1,\dots,n} f_j(\frac{\nabla u^k}{|\nabla u^k|})\frac{\partial^2 u^{k+1}}{\partial x_j^2}\Big),$$

together with a finite differences discretization in space to find the numerical approximation of the solution, u. For the finite differences approximation of the diffusive term, we follow the discretization indicated in [14], considering four directions based in a nine nodes local grid to implement the anisotropic diffusion operator, and use these nodes to compute also the discretized laplacian and gradient, according to the formulas given in [14]. The convolution $G_s * \nabla u$ is done in a nine nodes local grid, with variance $s = 1$, and normalized to get a partition of the unity in the discrete grid, and the enhancement parameter, ϵ, shaping function h and controlling the combination of isotropic and anisotropic diffusion applied to the image, is taken as $\epsilon = 0.75 \max |\nabla u_0|$, implying that isotropic diffusion dominates over anisotropic diffusion, see (15). For the convective term we use an upwind scheme involving also the nine nodes of the local grid. Since the distance between time-frequency nodes (pixels) is set equal to one, we choose constant evolution steps $d\tau < 1/(\varepsilon\|\nabla \log u_0\|)$, for stability issues. Observe that since the scheme is semi-implicit in time, the discrete comparison principle holds, see (16), and therefore the range of the discrete solution is always within the range of image definition.

A important advantage of our algorithm is that it provides us with a clue for stopping the evolution of the spectrogram transformation. As mentioned in the Introduction, for $\tau = 1$ ($\tau = 1/\varepsilon$, after the re-parametrization, see Remark 1), the solution of the transport equation (7) must be close to the reassigned spectrogram (4), while for $\tau \to \infty$, the solution converges to the extrema of the spectrogram. Therefore, it seems reasonable to stop the algorithm for some

τ_{end} of the order of $1/\varepsilon$, for which a good concentration on ridges should be achieved without affecting to the main geometric features of the image. On the other hand, the parameter ε has also a clear physical meaning. For the situation of a signal free of noise, the effect of the diffusion operator, which is intended mostly to smoothen the noise, is not necessary and therefore $\varepsilon =$ "∞" (no diffusion) should be considered. When noise appears and increases, i. e., when the signal to noise ratio (SNR) decreases, more weight should be given to the diffusion operator, which means decreasing ε. Therefore, parameter ε may be regarded as an increasing function of the SNR.

Finally, to show more clearly the advantages of our technique, in the subsequent plots we used a simple algorithm to produce candidates to instantaneous frequency lines of the corresponding spectrograms. Let $\Omega = [0, T] \times [0, F]$ be the time-frequency domain of the image and $u : \Omega \to [0, 255]$ be the starting image. We consider its truncation $v(t, \omega) = u(t, \omega)$ if $u(t, \omega) \geq \beta$, and $v(t, \omega) = 0$ elsewhere, with $\beta = \text{Mean}_\Omega(u)$ in the experiments. For each $t \in [0, T]$ we consider the N connected components of the set $\{\omega \in (0, F) : v(t, \omega) > 0\}$, say $C_n(t)$, for $n = 1, \ldots, N(t)$, and define the function

$$\text{IF}(t, n) = \frac{\int_{C_n(t)} \omega v(t, \omega) d\omega}{\int_{C_n(t)} v(t, \omega) d\omega},$$

which is the frequency gravity center of the component $C_n(t)$. In this way, we shrink each connected component to one point to which we assign the average image intensity through the function $\text{INT}(t, n) = \text{Mean}_{C_n(t)}(v(t, \cdot))$. Finally, we plot function IF only for components with averaged intensity, INT, greater than a certain threshold, $i \in [0, 255]$. This final image does not seem to be very sensible under small perturbations of the parameters β and i.

Experiment 1. We use a one second 4KHz synthetic signal composed by the addition of two signals. The first is the addition of pure tones and chirps:

$$x_1(t) = c_1(\sin 2\pi 1000t + \sin 2\pi 1100t + \sin 2\pi 1000t^2 + \sin 2\pi 600t^3),$$

while the second, x_2, is a uniformly distributed real random variable. We normalize them to have $\|x_i\|_{L^2} = 1$ (so the constant c_1) and define the test signal as $x = x_1 + x_2$, i.e., with SNR = 0. After computing the spectrogram (SP) and the reassigned spectrogram (RSP), we transform the former with three algorithms: the one presented in this paper (convective-diffusive, CD), with $d\tau = 0.1$, $\varepsilon = 0.2$ and $\tau_{end} = 1/\varepsilon = 5$ (50 iterations), the corresponding only-diffusive (D) algorithm ($\varepsilon = 0$) for the same $d\tau$ and τ_{end}, and the 2D-Stationary Wavelet Transform (SWT) for the biorthogonal basis with five levels and hard thresholding [17]. The results are shown in Figure 2. First column corresponds to the images obtained with the different algorithms, second shows a detail of these images, and third the corresponding IF functions for threshold level INT = 10. In the first column we observe that noise is smoothen efficiently by

the three algorithms D, CD and SWT. However, only for the CD we obtain narrower instantaneous frequency lines which approach quite well those of the RSP, specially in the horizontal direction. The details of the pure tones confirm this observation. Finally, the third column is intended to show more clearly the noise elimination of these algorithms, which is evident upon comparison of the plots produced for the SP or RSP, and those produced for the D, CD, and SWT. Differences among the latter are almost imperceptible, partly due to the fact that our IF algorithm computes averages of the true IF lines, hiding therefore the shrinking property of the CD algorithm.

Experiment 2. We used a recording done in wilderness, see [18], from where we extracted a signal of 0.5 seconds which is affected by a strong background noise. Since field data recorders are set to 44.1 KHz meanwhile wolves signals are rarely out of the range $200 - 3000$ Hz, we start by filtering and downsampling the signal to speed up computations. After computing the spectrogram and the reassigned spectrogram, we again transform the former with the three algorithms mentioned in Experiment 1. In this case we take $\varepsilon = 2$ in the CD algorithm and $d\tau = 0.01$, $\tau_{end} = 2/\varepsilon = 1$ (100 iterations) in both the CD and D algorithms. The 2D-SWT is again computed for the biorthogonal basis with five levels and hard thresholding. The results are shown in Figure 3. The first column corresponds to the images obtained with the different algorithms, and the second to the corresponding IF functions for threshold level INT = 6. Although the results are not as neat as for the synthetic signal, we still appreciate the smoothing of the most disperse noise components and an enhancement of the strongest IF lines. In addition, the CD algorithm clearly shrinks these lines. In the images at the right hand side, the IF lines are much better identified in the CD processed image. However, some of them appear broken due to the fact that the intensity of noise and some portions of the IF lines are similar, and therefore, treated by the algorithm in the same way.

Experiment 3. We used a recording done in captivity, see [19], from where we extracted a rather clean signal of 1 second. We again filtered and downsampled the signal to the range $200-2000$ Hz. After computing the spectrogram and the reassigned spectrogram, the former was transformed with the CD algorithm to test the possible convergence towards the reassigned spectrogram. In this case we took $\varepsilon = 1000$, $d\tau = 0.00005$ and $\tau_{end} = 5/\varepsilon = 0.005$ (100 iterations). The results are shown in Figure 4. In the first row we show the complete spectrogram, reassigned spectrogram and the CD algorithm output. In the second to fourth columns we show a detailed plot of the howl contained within the range $200 - 500$ Hz. We observe that the CD algorithm gives a better concentration of the IF lines and, moreover, produces smoother lines than the RS algorithm. However, execution time plays against the CD algorithm: the RS takes about 2 sec. while the CD takes about 70 sec. in a standard laptop. Execution time is, anyway, not an important issue for our application since it does not requires real time processing of the data.

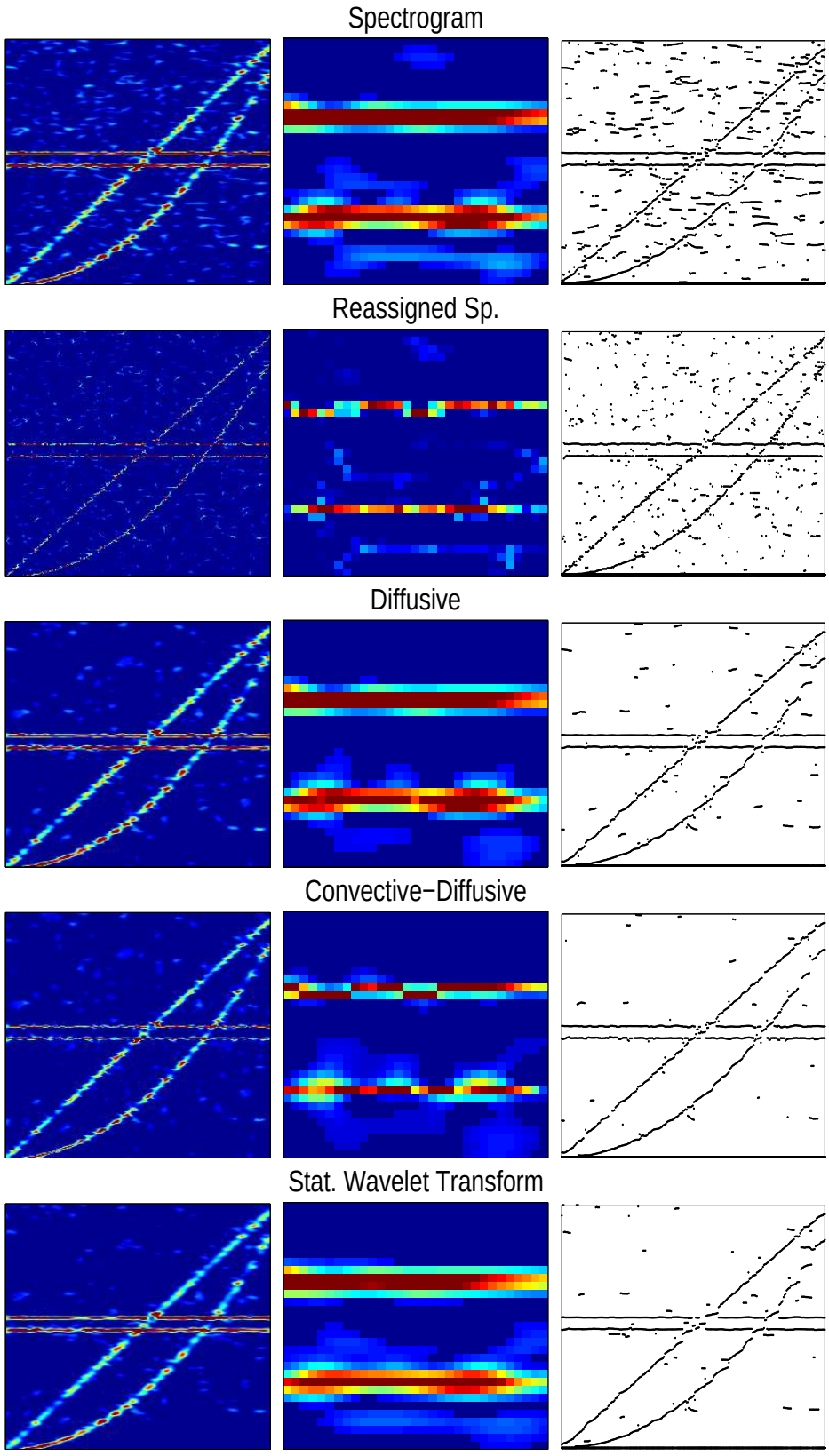

Fig. 2. Experiment 1. First column shows the images obtained with the different algorithms. Second column gives a detail of the pure tones in the previous image. Third column shows the corresponding transformation with the IF procedure.

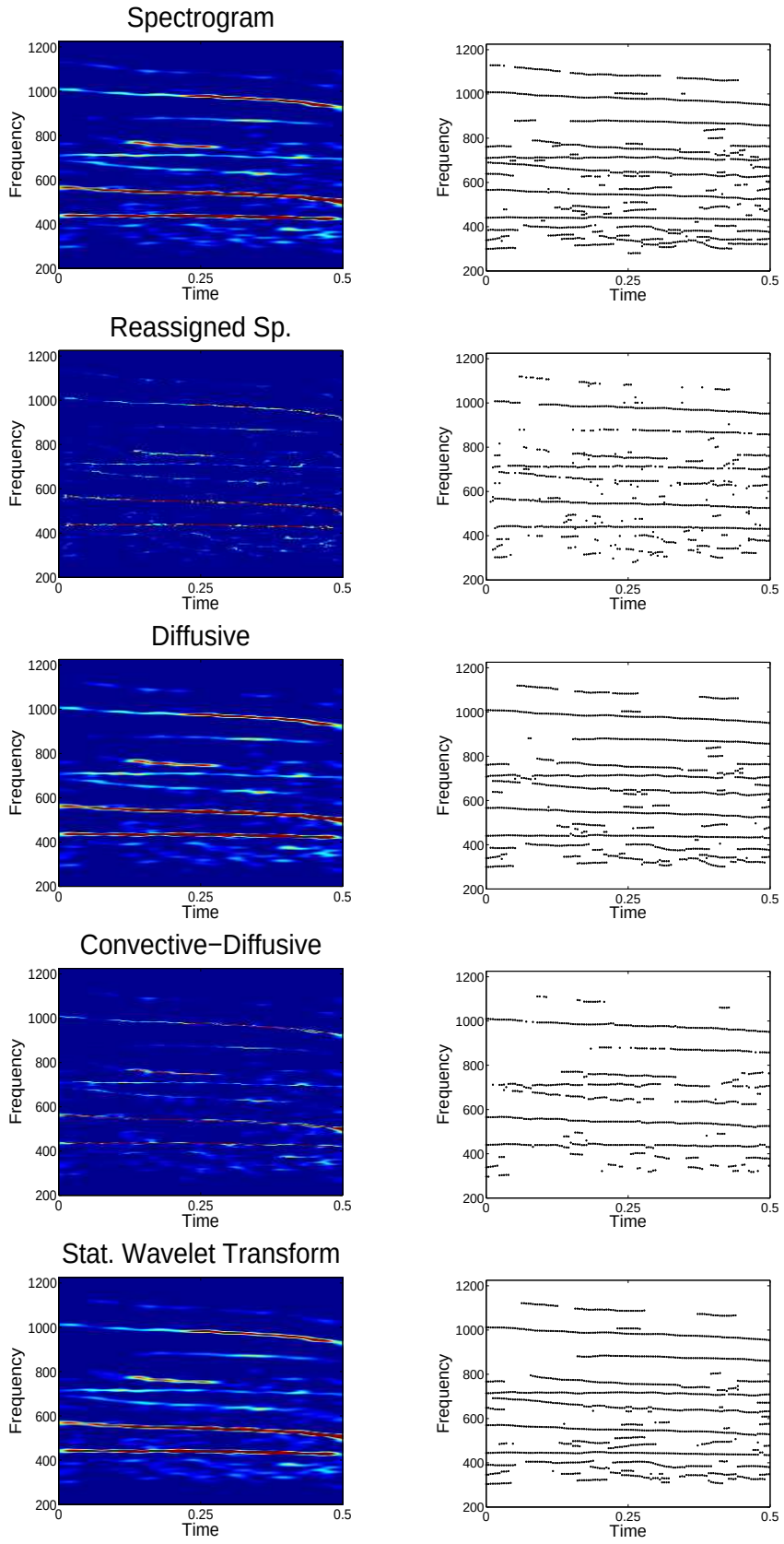

Fig. 3. Experiment 2. Images obtained with the different algorithms and their corresponding transformation with the IF routine.

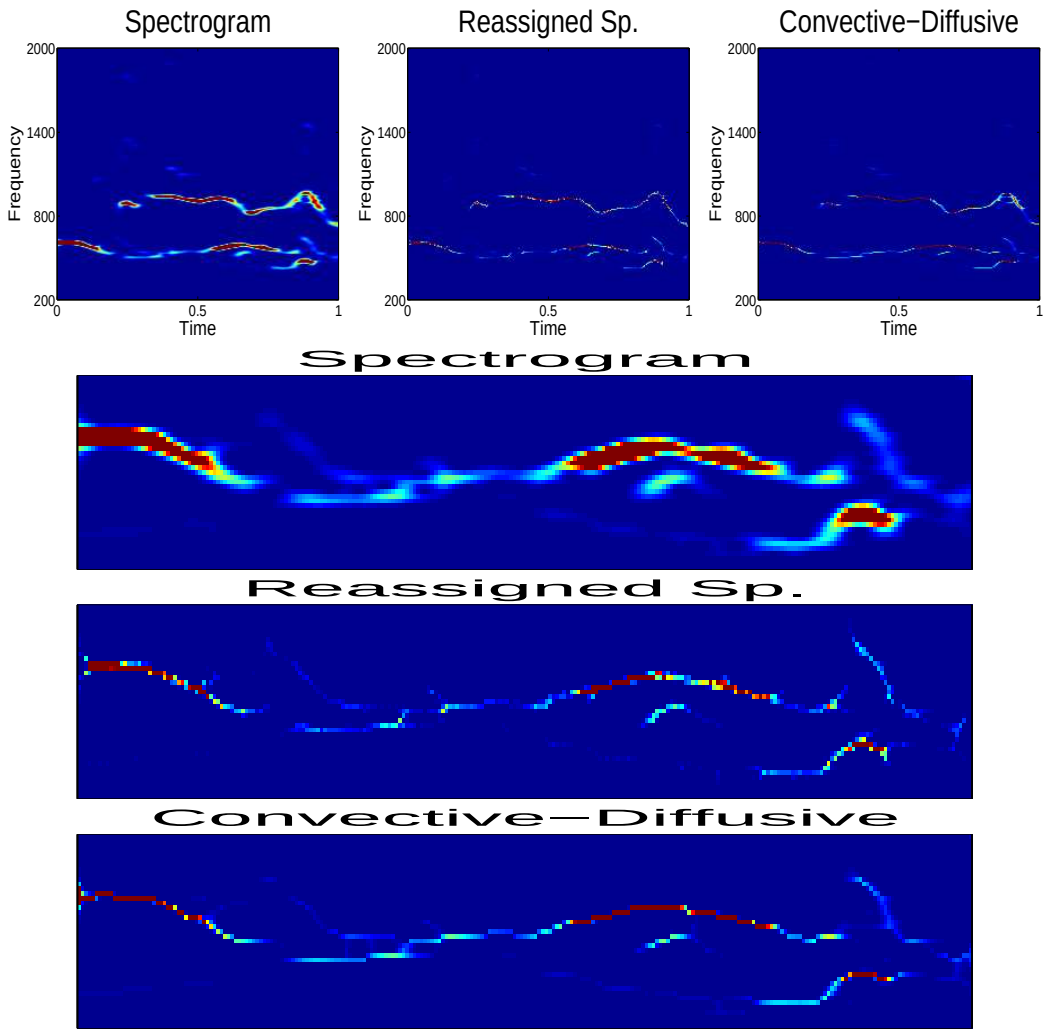

Fig. 4. Experiment 3. First row: Spectrogram, reassigned spectrogram and CD algorithm output. Subsequent plots: detail of the howl contained within the range $200 - 500$ Hz. The CD algorithm seems to concentrate and smoothen the IF lines better than the reassigned spectrogram.

References

[1] A. Skonhoft, The costs and benefits of animal predation: An analysis of Scandinavian wolf re-colonization, Ecological Economics 58 4 (2006) 830–841.

[2] S. Mann, S. Haykin, The Chirplet Transform: Physical Considerations, IEEE Trans. Signal Processing, 43 11 (1995) 2745–2761.

[3] L. Angrisani, M. D'Arco, A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation, IEEE Trans. Instrum. Meas., 51 (2002), 704–711.

[4] H. M. Ozaktas, Z. Zalevsky, M. A. Kutay, The Fractional Fourier Transform with Applications in Optics and Signal Processing, Wiley, Chichester, 2001.

[5] B. Dugnol, C. Fernández, G. Galiano, Implementation of a Chirplet Transform method for separating and counting wolf howls, Preprint of Dpt. Mathematics, Univ. of Oviedo, Spain, 2007.

[6] B. Dugnol, C. Fernández, G. Galiano, Wolves counting by spectrogram image processing, Appl. Math. Comput. (2006), doi:1016/j.amc.2006.08.173.

[7] B. Dugnol, C. Fernández, G. Galiano, Wolves chorus noise reduction by spectrogram image processing. To appear in The 2006 Int. Conf. Signal-Image Tech. & Internet-Based Systems (SITIS'06), IEEE, e-version in http://www.u-bourgogne.fr/SITIS/06/Proceedings/index.htm.

[8] K. Kodera, R. Gendrin, C. de Villedary, Analysis of time-varying signals with small BT values, IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-26 1 (1978) 64–76.

[9] F. Auger, P. Flandrin, Improving the readability of time-frequency and time-scale representations by the method of reassignment, IEEE Trans. Signal Processing, 43 5 (1995) 1068–1089.

[10] E. Chassandre-Mottin, F. Auger, P. Flandrin, On the statistics of spectrogramm reassignment, Multidimensional Systems and Signal Processing, 9 (1998) 355–362.

[11] E. Chassandre-Mottin, I. Daubechies, F. Auger, P. Flandrin, Differential Reassignment, IEEE Signal Processing Letters, 4 10 (1997) 293–294.

[12] S. Mallat, A wavelet tour of signal processing, Academic Press, London, 1998.

[13] F. Auger, Représentation temps-fréquence des signaux non-stationnaires: Sybtheèse et contributions, thèse de doctorat, Ecole Centrale de Nantes, 1991.

[14] L. Álvarez, P. L. Lions, J. M. Morel, Image selective smoothing and edge detection by nonlinear diffusion. II, SIAM J. Numer. Anal, 29 3 (1992), 845-866.

[15] M. G. Crandall, H. Ishii, and P. L. Lions, User's guide to viscosity solutions of second order partial differential equations, Bull. Amer. Math. Soc. (N.S.), 27 1 (1992) 1–67.

[16] F. Auger, P- Flandrin, P- Gonsalvès, O. Lemoine, Time-frequency toolbox for use with Matlab, CNRS (France) - Rice University (USA), 1996.

[17] Wavelet toolbox v. 3.0, Matlab v. 7.0, The MathWorks, Inc. 2004.

[18] L. LLaneza, V. Palacios, Asesores en Recursos Naturales, S.L., Field recordings obtained in wilderness in Asturias (Spain) in the 2003 campaign.

[19] L. LLaneza, V. Palacios, Asesores en Recursos Naturales, S.L., Field recordings obtained in captivity in Spain and Portugal in the 2005 campaign.