NOISE REDUCTION AND VOICE SEPARATION ALGORITHMS APPLIED TO WOLF POPULATION COUNTING

B. Dugnol, C. Fernández, G. Galiano and J. Velasco
Dpto. de Matemticas, Universidad de Oviedo, c/ Calvo Sotelo s/n, 33007 Oviedo, Spain

Keywords: Time-frequency distribution, instantaneous frequency, signal separation, noise reduction, chirplet transform, partial differential equation, population counting.

Abstract: We use signal and image theory based algorithms to produce estimations of the number of wolves emitting howls or barks in a given field recording as an individuals counting alternative to the traditional trace collecting methodologies. We proceed in two steps. Firstly, we clean and enhance the signal by using PDE based image processing algorithms applied to the signal spectrogram. Secondly, assuming that the wolves chorus may be modelled as an addition of nonlinear chirps, we use the quadratic energy distribution corresponding to the Chirplet Transform of the signal to produce estimates of the corresponding instantaneous frequencies, chirp-rates and amplitudes at each instant of the recording. We finally establish suitable criteria to decide how such estimates are connected in time.

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 (Skonhoft, 2006). 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. The use of a Chirplet type transform (S. Mann, 1995; L. Angrisani, 2002) is investigated in this article, although an equivalent formulation in terms of the Fourier fractional transform (H. M. Ozaktas, 2001) could be employed as well. 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

Dugnol B., Fernández C., Galiano G. and Velasco J. (2008).
NOISE REDUCTION AND VOICE SEPARATION ALGORITHMS APPLIED TO WOLF POPULATION COUNTING.
In *Proceedings of the First International Conference on Bio-inspired Systems and Signal Processing*, pages 77-84
DOI: 10.5220/0001057000770084

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, and which should be treated in advance. We accomplish this task by using PDE-based techniques which transforms the image of the signal spectrogram into a smoothed and enhanced approximation to the reassigned spectrogram introduced in (K. Kodera, 1978; F. Auger, 1995) as a spectrogram readability improving method.

2 SIGNAL ENHANCEMENT

In previous works (B. Dugnol, 2007a; B. Dugnol, 2007d), we investigated the noise reduction and edge (IF lines) enhancement on the spectrogram image by a PDE-based image processing algorithm. For a clean signal, the method allows to produce an approximation to the reassigned spectrogram through a process referred to as *differential reassignment*, and for a noisy signal this process is modified by the introduction of a nonlinear operator which induces isotropic diffusion (noise smoothing) in regions with low gradient values, and anisotropic diffusion (edge-IF enhancement) in regions with high gradient values.

Let $x \in L^2(\mathbb{R})$ denote an audio signal and consider the Short Time Fourier transform (STFT)

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

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

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

which may be expressed also as (Mallat, 1998)

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

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 (Mallat, 1998). However, it is well known that it presents high amplitude sign-varying cross-terms for multi-component signals which makes its interpretation difficult. 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 was suggested in (K. Kodera, 1978) 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. As deduced in (F. Auger, 1995), the gravity center may be computed by the following formulas

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

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. 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 (Auger, 1991). 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 (E. Chassandre-Mottin, 1997),

$$\mathbf{r}(t,\omega) = \frac{1}{2}\nabla\log(S_\varphi(x;t,\omega)), \qquad (4)$$

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 $(\hat{t},\hat{\omega})$. Rewriting 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* (E. Chassandre-Mottin, 1997), defined as the transformation given by the dynamical system corresponding

to such velocity field,

$$\begin{cases} \dfrac{d\chi}{d\tau}(t,\omega,\tau) = \mathbf{r}(\chi(t,\omega,\tau)), \\ \chi(t,\omega,0) = (t,\omega), \end{cases} \tag{5}$$

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

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

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 for the differential reassignment is obtained by solving the following problem for $u(t,\omega,\tau)$ and $\tau > 0$,

$$\frac{\partial u}{\partial \tau} + \operatorname{div}(u\mathbf{r}) = 0, \tag{6}$$

$$u(\cdot,\cdot,0) = u_0, \tag{7}$$

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 (6)-(7) to hold in a bounded time-frequency domain, Ω, in which we assume non energy flow conditions on the solution and the data

$$\nabla u \cdot \mathbf{n} = 0, \quad \mathbf{r}\cdot\mathbf{n} = 0 \quad \text{on } \partial\Omega \times \mathbb{R}_+, \tag{8}$$

being $\mathbf{n}$ the unitary outwards normal to $\partial\Omega$. Finally, observe that the positivity of the spectrogram (Mallat, 1998) 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 (6)-(8) admits a unique smooth solution.

As noted in (E. Chassandre-Mottin, 1997), 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 (B. Dugnol, 2007a; B. Dugnol, 2007d), 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

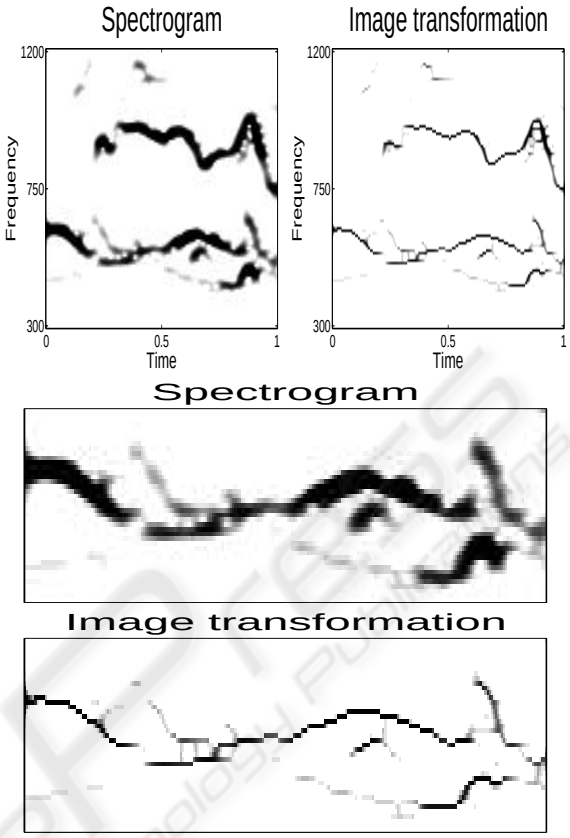

Figure 1: First row: Spectrogram and its transformation with the PDE model. Subsequent plots: detail of the howl contained within the range $300-750$ Hz. We observe the IF concentration and smoothing effect of the PDE algorithm.

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 modifying (6) by adding a diffusive term with the mentioned properties.

Let us make a final observation before writing the model we 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 (5) starting at (t_0,ω_0). Evaluating Eq. (6) along χ_0 we obtain

$$\frac{d}{d\tau}u = \frac{\partial u}{\partial \tau} + \mathbf{r}\cdot\nabla u = -u\operatorname{div}\mathbf{r}, \tag{9}$$

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 lies 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. (9), i.e., replacing Eq. (6) by the transport equation

$$\frac{\partial u}{\partial \tau} + \mathbf{r} \cdot \nabla u = 0. \qquad (10)$$

However, no energy conservation law will apply anymore (note that u is constant along the characteristics). The combination of the differential reassignment problem with the edge-detection image-smoothing algorithm (L. Álvarez, 1992) is written as

$$\frac{\partial u}{\partial \tau} + \frac{\varepsilon}{2}\nabla \log(u_0) \cdot \nabla u - g(|G_s * \nabla u|)A(u) = 0, \quad (11)$$

in $\Omega \times \mathbb{R}_+$, together with the boundary data (8) and the initial condition (7). Parameter $\varepsilon \geq 0$ allows us to play with different balances between transport and diffusion effects. The diffusion operator is given by

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

Let us briefly remind the properties and meaning of the diffusive term components in equation (11):

- 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

$$h(s) = \begin{cases} 0 & \text{for } s \leq h_0, \\ 1 & \text{for } s \geq 2h_0, \end{cases} \qquad (12)$$

being h_0 the *enhancement* parameter.

In Fig. 1 we show an example of the outcome of our algorithm for a signal composed by two howls. See (B. Dugnol, 2007d; B. Dugnol, 2007c) for more details and other numerical experiments.

3 HOWL TRACKING AND SEPARATION

A wolves chorus is composed, mainly, by howls and barks which, from the analytical point of view, may be regarded as chirp functions. The former has a long time support and a small frequency range variation, while the latter is almost punctually localized in time but posses a large frequency spectrum. It is convenient, therefore, adopting a parametric model to represent the wolves chorus as an addition of chirps given by the function $f : [0,T] \to \mathbb{C}$,

$$f(t) = \sum_{n=1}^{N} a_n(t)\exp[i\phi_n(t)], \qquad (13)$$

with T the length of the chorus emission, a_n and ϕ_n the chirps amplitude and phase, respectively, and with N, the number of chirps contained in the chorus. We notice that N is not necessarily the number of wolves since, for instance, harmonics of a given fundamental tone are counted separately.

To identify the unknowns N, a_n and ϕ_n we proceed in two steps. Firstly, for a time discretization of the time interval $[0,T]$, say t_j, for $j = 0,\ldots,J$, we produce estimates of the amplitude $a_n(t_j)$ and the phase $\phi_n(t_j)$ of the chirps contained at such discrete times. Secondly, we establish criteria which allow us deciding if the computed estimates at adjacent times do belong to the same global chirp or do not.

For the first step we use the Chirplet transform defined by

$$\Psi f(t_o,\xi,\mu;\lambda) = \int_{-\infty}^{\infty} f(t)\,\overline{\psi_{t_o,\xi,\mu,\lambda}(t)}dt, \qquad (14)$$

with the complex window $\psi_{t_o,\xi,\mu,\lambda}$ given by

$$\psi_{t_o,\xi,\mu,\lambda}(t) = v_\lambda(t-t_o)\exp\left[i(\xi t + \frac{\mu}{2}(t-t_o)^2)\right]. \qquad (15)$$

Here, $v \in L^2(\mathbb{R})$ denotes a real window, $v_\lambda(\cdot) = v(\cdot/\lambda)$, with $\lambda > 0$, and the parameters $t_o,\xi,\mu \in \mathbb{R}$, stand for time, instantaneous frequency and chirp rate, respectively. The quadratic energy distribution corresponding to the chirplet transform (14) is given by

$$P_\Psi f(t_o,\xi,\mu;\lambda) = |\Psi f(t_o,\xi,\mu;\lambda)|^2. \qquad (16)$$

For a linear chirp of the form

$$f(t) = a(t)\exp[i(\frac{\alpha}{2}(t-t_o)^2 + \beta(t-t_o) + \gamma)],$$

it is straightforward to prove that the energy distribution (16) has a global maximum at (α, β), allowing us to determine the IF and chirp rate of a given linear chirp by localizing the maxima of the energy distribution. For more general forms of mono-component chirps we have the following localization result (B. Dugnol, 2007b)

Theorem 1 *Let* $f(t) = a(t)\exp[i\phi(t)]$, *with* $a \in L^2(\mathbb{R})$ *non-negative and* $\phi \in C^3(\mathbb{R})$. *For all* $\varepsilon > 0$ *and* $\xi, \mu \in \mathbb{R}$ *there exists* $L > 0$ *such that if* $\lambda < L$ *then*

$$P_\Psi f(t_o, \xi, \mu; \lambda) \leqslant \varepsilon + P_\Psi f\left(t_o, \phi'(t_o), \phi''(t_o); \lambda\right). \tag{17}$$

In addition,

$$\lim_{\lambda \to 0} P_\Psi f\left(t_o, \phi'(t_o), \phi''(t_o); \lambda\right) = a(t_o)^2. \tag{18}$$

In other words, for a general mono-component chirp the energy distribution maximum provides an arbitrarily close approximation to the IF and chirp rate of the signal. Moreover, its amplitude may also be estimated by shrinking the window time support at the maximum point.

Finally, for a multi-component chirp $f(t) = \sum_{n=1}^{N} a_n(t)\exp[i\phi_n(t)]$ the situation is somehow more involved since although the energy distribution still has maxima at $(\phi_n'(t_o), \phi_n''(t_o))$ for all n such that $a_n(t_o) \neq 0$, these are now of local nature, and in fact, spurious local maxima not corresponding to any chirp may appear due to the energy interaction among the actual chirps.

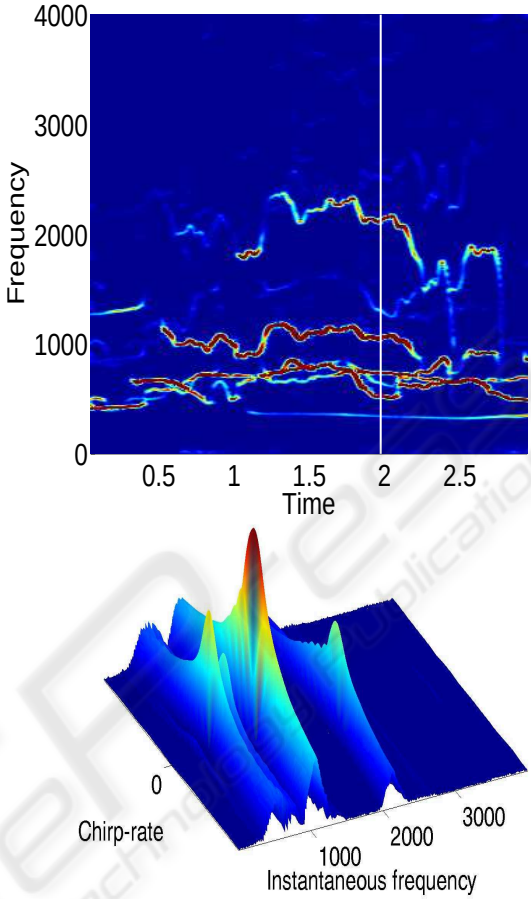

Figure 2: Above: STFT of a field recorded signal. Below: quadratic energy distribution of the chirplet transform at $t_0 = 2$. Maxima correspond to IF and chirp rate chirps locations. We observe the different behavior in the ξ and μ directions at these maxima.

4 NUMERICAL EXPERIMENTS

According to the recording quality, we start our algorithm enhancing the signal with the PDE algorithm explained in Section 2 or directly with the separation algorithm introduced in Section 3. For details about the implementation of the former, we refer the reader to (B. Dugnol, 2007c). Following, we briefly comment about the separation algorithm implementation. We start by computing the energy distribution, $P_\Psi f(\tau_m, \xi, \mu; \lambda)$ at a set of discrete times $\tau_m = m * \tau$ of constant time step, τ, and for a fixed window width λ. Next we compute the maxima of the energy at each of these times. When the signal is mono-component or the various components of the signal are far from each other relative to the window width, the maxima of P_Ψ correspond to some $(\phi_n'(\tau_m), \phi_n''(\tau_m))$ which are then identified as the IF and chirp-rate of a chirp candidate. However, when multi-component signals are close to each other or are crossing, some spurious local maxima are produced which do not correspond to any actual chirp. Therefore, some criterium must be used to select the correct local maxima at each τ_m. Although we lack of an analytical proof, there are evidences suggesting that maxima produced by chirps, i.e., at points of the type $(\phi_n'(\tau_m), \phi_n''(\tau_m))$, decrease much faster in the ξ direction than in the μ direction, see Fig. 2, a phenomenon that does not occur at spurious maxima. We use this fact to choose the candidates first by selecting ξ_k, for $k = 1, \ldots, K$, which are maxima for

$$\sup_\mu P_\Psi f(\tau_m, \xi, \mu; \lambda),$$

and, among them, selecting the maxima with respect to μ of $P_\Psi f(\tau_m, \xi_k, \mu; \lambda)$. We finally establish a threshold parameter to filter out possible local max-

ima located at points that do not correspond to any $\phi_n'(\tau_m)$ but which are close to two of them. We set this threshold such that the existence of two consecutive maxima is avoided.

In this way we obtain, for each τ_m, a set of points (μ_{i_m}, ξ_{i_m}), for $i_m = 1, \ldots, I_m$, which correspond to the IF's and chirp-rates of chirps with time support including τ_m.

The next step is the chirp separation. We note that if the time step $\tau = \tau_{m+1} - \tau_m$ is small enough, then

$$\xi_{j_{m+1}} - \tau \frac{\mu_{j_{m+1}}}{2} \approx \xi_{i_m} + \tau \frac{\mu_{i_m}}{2}.$$

Introducing a new parameter, ν, we test this property by imposing the condition

$$\frac{1}{\nu} < \frac{2\xi_{j_{m+1}} - \tau\mu_{j_{m+1}}}{2\xi_{i_m} + \tau\mu_{i_m}} < \nu, \qquad (19)$$

for two points to be in the same chirp. In the experiments we take $\nu = 2^{1/13} \approx 1.0548$.

Finally, in the case in which test (19) is satisfied by more than one point, i.e., when there exist points $\left(\xi_{j_{m+1}}, \mu_{j_{m+1}}\right)$ and $\left(\xi_{k_{m+1}}, \mu_{k_{m+1}}\right)$ such that (19) holds for the same (ξ_{i_m}, μ_{i_m}), we impose a regularity criterium and choose the point with a closer chirp-rate to that of (ξ_{i_m}, μ_{i_m}). This is a situation typically arising at chirps crossings points.

Summarizing, the chirp separation algorithm is implemented as follows:

- Each point (ξ_{i_1}, μ_{i_1}), for $i_1 = 1, \ldots, I_1$, is assumed to belong to a distinct chirp.
- For $k = 2, 3, \ldots$, we use the described criteria to decide if $\left(\xi_{i_k}, \mu_{i_k}\right)$, for : $i_k = 1, \ldots, I_k$, belongs to an already detected chirp. On the contrary, it is established as the starting point of a new chirp.
- When the above iteration is finished and to avoid artifacts due to numerical errors, we disregard chirps composed by a unique point.

Finally, once the chirps are separated, we use the following approximation, motivated by Theorem 1, to estimate the amplitude

$$a(\tau_m)^2 \approx \frac{1}{\lambda \left[\widehat{v}(0)\right]^2} P_\Psi f\left(t_o, \phi'(\tau_m), \phi''(\tau_m); \lambda\right).$$

Again, to avoid artifacts due to numerical discretization, we neglect portions of signals with an amplitude lower than certain relative threshold , $\varepsilon \in (0,1)$, of the maximum amplitude of the whole signal, considering that in this case no chirp is present.

4.1 Experiment 1. A Synthetic Signal

In this first experiment we test our algorithm with a synthetic signal, f, composed by the addition of three

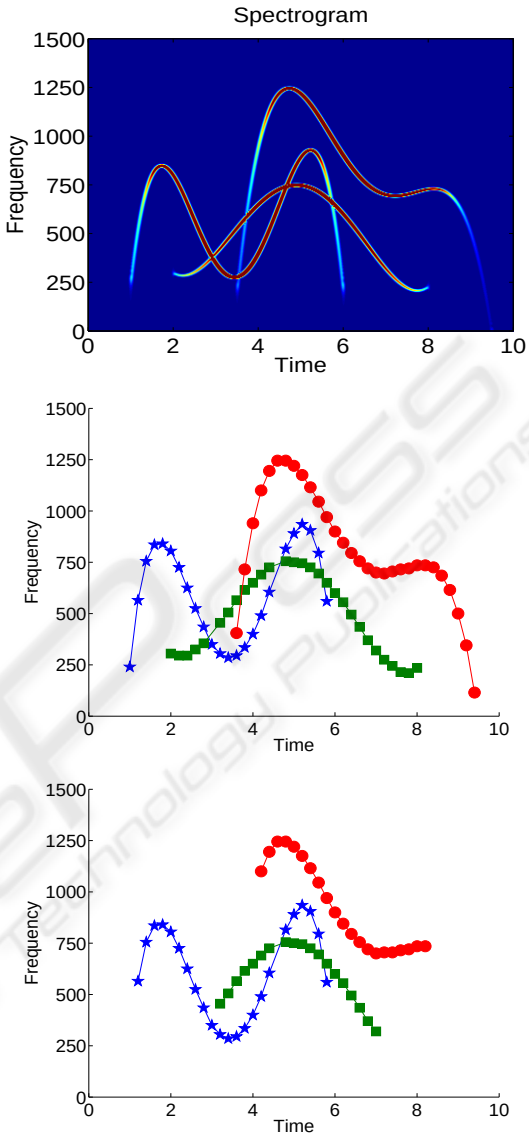

Figure 3: Spectrogram of the clean signal and results of the chirp localization and separation algorithm for clean and noisy signals, respectively.

nonlinear chirps (spectrogram shown in Fig. 3) and with the same signal corrupted with an additive noise of similar amplitude than that of f, i.e., with $SNR = 0$.

We used the same time step, $\tau = 0.2$ sec, and window width, $\lambda = 0.1$ sec, to process both signals, while we set the relative threshold amplitude level to $\varepsilon = 0.01$ for the clean signal and to $\varepsilon = 0.1$ for the noisy signal. The results of our algorithm of denoising, detection and separation is shown in Fig. 3. We observe that for the clean signal all chirps are captured with a high degree of accuracy even at crossing points. We also observe that the main effect of noise corruption is the lose of information at chirps

low amplitude range. However, the number of them is correctly computed.

In Fig. 4 we show the amplitude, IF and chirp-rate estimations of the chirp which is more affected by the noise corruption, for both clean and noisy signals. The main effect of noise corruption is observed in the amplitude computation and in the lose of information in the tails of the three quantities.

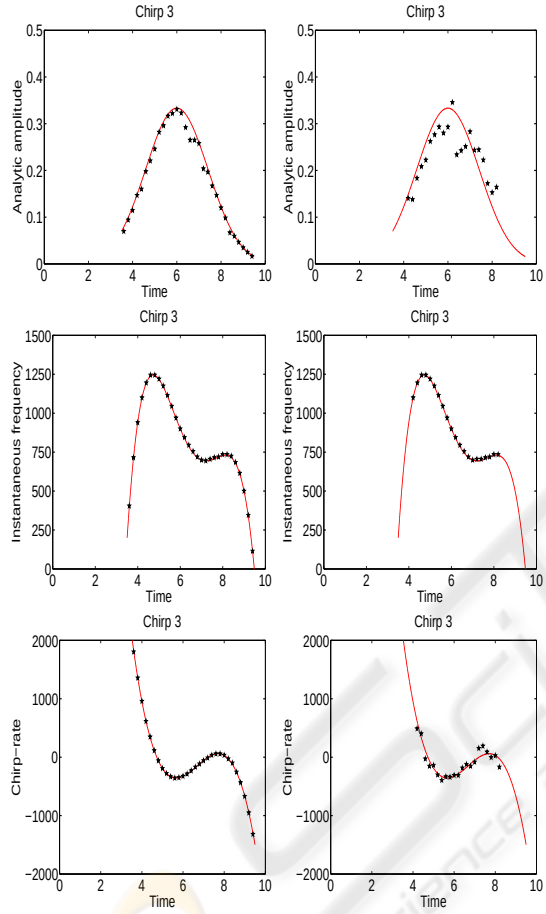

Figure 4: Amplitude, IF and chirp rate for the clean signal (left column) and noisy signal (right column). Solid lines correspond to exact values and crosses to computed values.

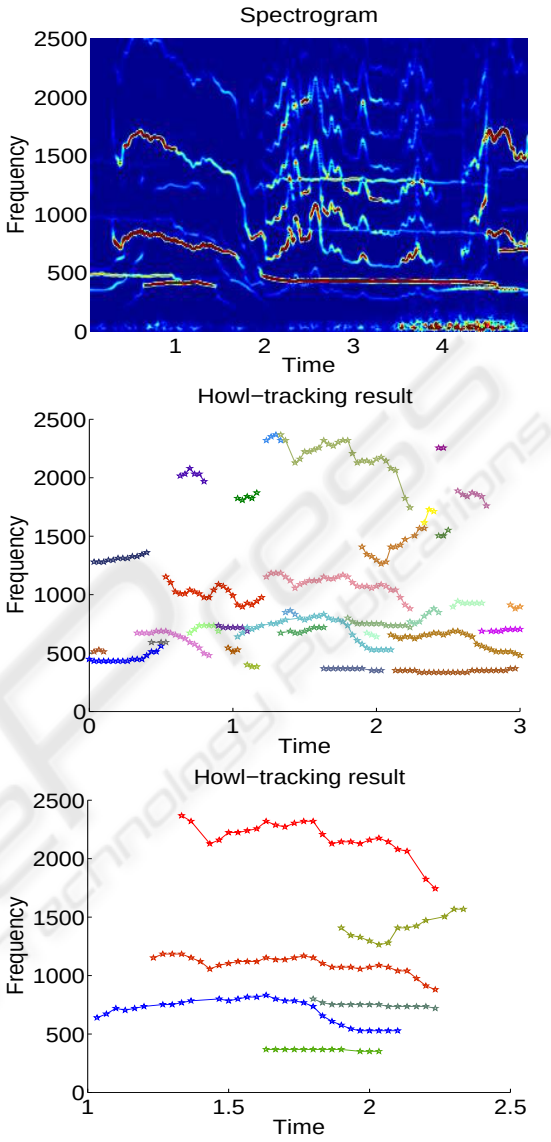

Figure 5: First row: spectrogram of the field recorded signal utilized in Experiment 2. Second row: result of the chirp localization and separation algorithm. Third row: a zoom of the previous plot showing six separated chirps corresponding to five wolves howls.

4.2 Experiment 2. A Field Recorded Wolves Chorus

In this experiment we analyze a rather complex signal obtained from field recordings of wolves choruses in wilderness, (L. Llaneza,). Due to the noise present in the recording, we first use the PDE algorithm to enhance the signal and reduce the noise, see (B. Dugnol, 2007c) for details. For the separation algorithm, we fixed the time step as $\tau = 0.03$ sec, the relative am-

plitude threshold as $\varepsilon = 0.01$ and the window width as $\lambda = 0.0625$ sec.

The algorithm output is composed by 32 chirps which should correspond to the howls and barks (with all their harmonics) emitted by the wolves along the duration of the recording (about five sec). The result is shown in Fig. 5. Since our aim is giving an estimate of how many individuals are emitting in a recording, we plot a zoom of the separating algorithm result for the time interval $(1, 2.5)$. Here, the number of chirps reduces to six. However, it seems that one couple of them are harmonics, the couples formed by the chirp around 1000 Hz and the highest IF chirp. Therefore, we may conclude that at least five wolves are emitting in this interval of time. A similar analysis is carried out with other time subintervals until all the recorded signal is analyzed.

5 CONCLUSIONS

A combined algorithm for signal enhancement and voice separation is utilized for wolf population counting. Although field recorded wolf chorus signals posses a complex structure due to noise corruption and nonlinear multi-component character, the outcome of our algorithm provides us with accurate estimates of the number of individuals emitting in a given recording. Thus, the algorithm seems to be a good complement or, even, an alternative to existent methodologies, mainly based in wolf traces collection or in the intrusive attaching of electronic devices to the animals. Clearly, our algorithm is not limited to wolves emissions but to any signal which may reasonably be modelled as an addition of chirps, opening its utilization to other applications. Drawbacks of the algorithm are related to the expert dependent election of some parameters, such as the amplitude threshold, or to the execution time when denoising and separation of long duration signals must be accomplished. We are currently working in the improvement of these aspects as well as in the recognition of components corresponding to the same emissor, such as harmonics of a fundamental chirp, pursuing the full automatization of the counting algorithm.

ACKNOWLEDGEMENTS

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.

REFERENCES

Auger, F. (1991). *Représentation temps-fréquence des signaux non-stationnaires: Syntheèse et contributions*. thèse de doctorat, Ecole Centrale de Nantes.

B. Dugnol, C. Fernández, G. G. (2007a). Wolves counting by spectrogram image processing. *Appl. Math. Comput.*, 186:820–830.

B. Dugnol, C. Fernández, G. G. J. V. (2007b). Implementation of a chirplet transform method for separating and counting wolf howls. Preprint 1, Dpt. Mathematics, Univ. of Oviedo, Spain.

B. Dugnol, C. Fernández, G. G. J. V. (2007c). Implementation of a diffusive differential reassignment method for signal enhancement. an application to wolf population counting. *Appl. Math. Comput.* To appear. E-version: doi:10.1016/j.amc.2007.03.086.

B. Dugnol, C. Fernández, G. G. J. V. (2007d). On pde-based spectrogram image restoration. application to wolf chorus noise reduction and comparison with other algorithms. In E. Damiani, A. Dipanda, K. Y. L. L. P. S., editor, *Signal processing for image enhancement and multimedia processing*, volume 34 of *Multimedia systems and applications*. Springer Verlag. To appear. (e-version in http://www.u-bourgogne.fr/SITIS/06/Proceedings/index.htm).

E. Chassandre-Mottin, I. Daubechies, F. A. P. F. (1997). Differential reassignment. *IEEE Signal Processing Letters*, 4(10):293–294.

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

H. M. Ozaktas, Z. Zalevsky, M. A. K. (2001). *The Fractional Fourier Transform with Applications in Optics and Signal Processing*. Wiley, Chichester.

K. Kodera, R. Gendrin, C. d. V. (1978). Analysis of time-varying signals with small bt values. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(1):64–76.

L. Álvarez, P. L. Lions, J. M. M. (1992). Image selective smoothing and edge detection by nonlinear diffusion. ii. *SIAM J. Numer. Anal.*, 29(3):845–866.

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

L. Llaneza, V. P. Field recordings obtained in wilderness in asturias (spain) in the 2003 campaign. Asesores en Recursos Naturales, S.L.

Mallat, S. (1998). *A wavelet tour of signal processing*. Academic Press, London.

S. Mann, S. H. (1995). The chirplet transform: Physical considerations. *IEEE Trans. Signal Processing*, 43(11):2745–2761.

Skonhoft, A. (2006). The costs and benefits of animal predation: An analysis of scandinavian wolf recolonization. *Ecological Economics*, 58(4):830–841.