MULTI-CHIRP SIGNAL SEPARATION

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

Keywords: Parametric model, Chirplet transform, Instantaneous frequency, Signal separation.

Abstract: Assuming that an specific audio signal, such as recordings of animal sounds, 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 and design an algorithm for tracking and separating the chirp components of the signal. We demonstrate the accuracy of our algorithm applying it to some synthetic and field recorded signals.

1 INTRODUCTION

For many audio recordings obtained from natural sources, such as emission from animals, it is a convenient approach to use a parametric model in terms of a super-position of chirps, $s(t) = \sum_n a_n(t) \exp[i\phi_n(t)]$, with a_n the analytic amplitude and ϕ_n the phase, whose derivative determines the chirp instantaneous frequency (IF).

There are several techniques focused in the amplitude and phase estimation of this kind of signals. McAulay and Quartieri (McAulay and Quatieri, 1986) propose algorithms based on the STFT, while for instance in (Maragos et al., 1993; Santhanam and Maragos, 2000) these estimates are obtained through the use of the Teager-Kaiser operator, $E = (x'(t))^2 - x(t)x''(t)$. The IF estimation via non-parametric models using Cohen distributions is also an important issue, see, for instance, (Stanković and Djurović, 2003; Kwok and Jones, 2000; Zhao et al., 2004; Boashash, 1992a; Boashash, 1992b).

Other kind of approaches, such as the empirical mode decomposition (EMC) (et al., 1998)) or the Gianfelici transform (Gianfelici et al., 2007) focus on the signal decomposition in terms of basis which depend on the given function.

In (Dugnol et al., 2008) we proposed a methodology based on the chirplet transform (Mann and Haykin, 1991) to identify and separate the different chirps conforming the given signal. To do this, not only amplitude and IF estimation was needed but also the so called chirp rate (CR, phase second derivative). We showed that the quadratic energy corresponding to the chirplet transform may be used for this task since for each time of a discrete time mesh the maxima points of the energy in the IF-CR plane correspond to *some* of the IF and CR of the existent chirps. However, some other maxima are *spurious* maxima, not corresponding to any chirp at all. Our previous algorithm, see (Dugnol et al., 2008), had some inefficiencies regarding the identification of these spurious maxima and, accordingly, to the matching process to construct the global chirp components of the signal. In this work we introduce important improvements to our previous algorithm.

2 CHIRPLET TRANSFORM ENERGY DISTRIBUTION

For a given signal $f \in L^2(\mathbb{R})$ we consider its chirplet transform given by

$$\Psi f(\tau,\xi,\mu;\lambda) = \int_{-\infty}^{\infty} f(t)\,\overline{\psi_{\mu,\lambda}(t-\tau)}\exp[-i\xi t]\,dt,$$

with τ,ξ and μ standing for time, IF and CR, respectively, parameter λ determining the window width and with $\psi_{\mu,\lambda}$ a complex window defined as

$$\psi_{\mu,\lambda}(s) = v_\lambda(s)\exp\left[i\left(\frac{\mu}{2}s^2\right)\right],$$

with $v_\lambda(s) = \lambda^{-1/2}v(s/\lambda)$ and v a non-negative, normalized, symmetric window. The quadratic energy corresponding to the chirplet transform is then given by

$$P_\psi f(\tau,\xi,\mu;\lambda) = |\Psi f(\tau,\xi,\mu;\lambda)|^2. \qquad (1)$$

Dugnol B., Fernández C., Galiano G. and Velasco J. (2009).
MULTI-CHIRP SIGNAL SEPARATION.
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pages 216-221
DOI: 10.5220/0001432802160221

We notice that Ψf may be rewritten as

$$\Psi f(\tau,\xi,\mu,\lambda) = \int_{-\infty}^{\infty} f'(x)\exp\left[-i\left(\xi x+\frac{\mu}{2}x^2\right)\right]dx,$$

with $f'(x) = f(x+\tau)v_\lambda(x)\exp[-i\xi\tau]$, implying that the Chirplet Transform provides a correlation measure of the linear chirp $\exp\left[-i\left(\xi x+\frac{\mu}{2}x^2\right)\right]$ and the portion of the signal centered at τ, . It is, then, clear that for a linear chirp

$$f(t) = a_1(t)\exp\left[i\left(\frac{\alpha_1}{2}(t-\tau)^2+\beta_1(t-\tau)+\gamma_1\right)\right] \quad (2)$$

and for fixed τ and λ, the quadratic energy distribution $P_\Psi f(\tau,\xi,\mu;\lambda)$ has a global maximum at $(\xi,\mu) = (\alpha_1,\beta_1)$, allowing us to determine the IF and CR of a given linear chirp at a fixed time by computing the maximum point of the energy distribution. Moreover, the following result holds, see (Dugnol et al., 2008): For

$$f(t) = a(t)\exp[i\phi(t)], \quad (3)$$

and for all $\varepsilon > 0$ and $\xi,\mu \in \mathbb{R}$ there exists $L > 0$ such that if $\lambda < L$ then

$$P_\Psi f(\tau,\xi,\mu;\lambda) \leqslant \lambda\varepsilon + P_\Psi f\left(\tau,\phi'(\tau),\phi''(\tau);\lambda\right). \quad (4)$$

In addition,

$$\lim_{\lambda\to 0}\frac{1}{\lambda}P_\Psi f\left(\tau,\phi'(\tau),\phi''(\tau);\lambda\right) = a(\tau)^2. \quad (5)$$

In other words, for a general mono-component chirp the energy distribution maximum provides an arbitrarily close approximation to the IF and CR of the signal. Moreover, its amplitude may also be estimated by shrinking the window time support at the maximum point. In fact, these estimates make also possible to establish criteria for deciding, in the case of a multi-chirp signal, when two energy maxima points in the IF-CR plane do belong to the same chirp or to different chirps, and in this way, it conforms a procedure for separating the chirp components (Dugnol et al., 2008).

2.1 Energy Distribution in the CR-IF Plane

It may be experimentally observed that the support of the energy distribution of a one-chirp signal like (3) for a fixed time τ has a geometric shape given by a double cone with vertex at $(\phi''(\tau),\phi'(\tau))$ in the CR-IF plane, see Figure 1.

Let us analyze the case of a linear chirp like (2) to get some analytical insight into this property. By using an appropriate window we may restrict the analysis to the time interval $(\tau-\lambda,\tau+\lambda)$. Fixing the ξ value in the energy distribution (1) we obtain a line

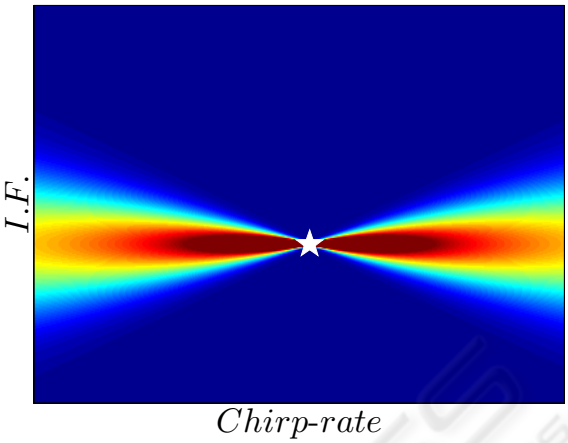

Figure 1: Energy distribution for one chirp signal at the CR-IF plane.

bundle at (τ,ξ) generated by the different values of the slope μ, see Fig. 2. For the energy distribution to be remarkable, the segment defined in the time-frequency plane by the chirp $\omega = \alpha_1(t-\tau)+\beta_1$ must intersect this bundle. Typically there exists a range of slopes for which this intersection is not empty, see Fig. 2. In particular, let $r_\xi(\mu) = \xi+\mu(t-\tau)$ be the line passing through (τ,ξ) with slope μ, and let

$$\mu_m := \min\left\{\alpha_1+\frac{\xi-\beta_1}{\lambda},\alpha_1-\frac{\xi-\beta_1}{\lambda}\right\}$$

and

$$\mu_M := \max\left\{\alpha_1+\frac{\xi-\beta_1}{\lambda},\alpha_1-\frac{\xi-\beta_1}{\lambda}\right\}.$$

Then if $\mu \in (\mu_m,\mu_M)$, the segment $r_\xi(\mu) = \xi+\mu(t-\tau)$ does not intersect the chirp and therefore the energy is negligible.

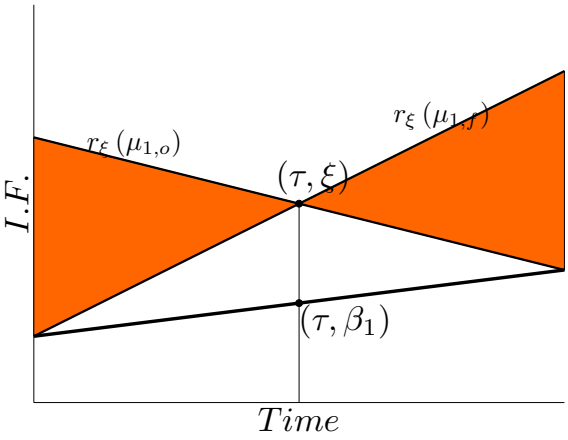

Figure 2: Shadow region corresponds to the chirplet negligible energy set.

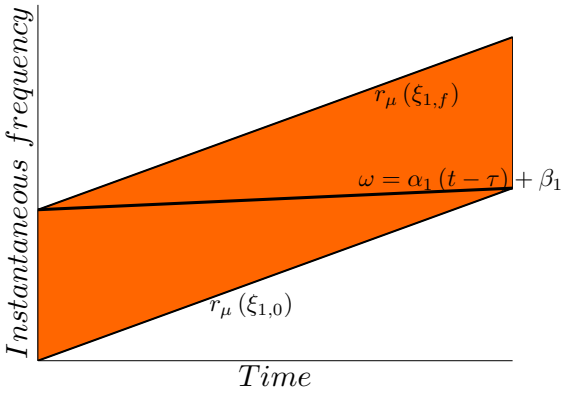

Figure 3: Shadow region corresponds to the chirplet concentration energy set.

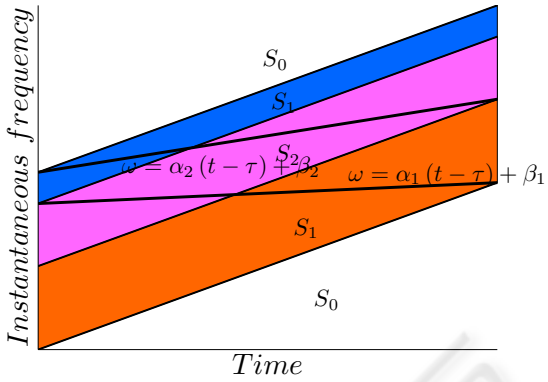

Figure 4: Two chirps energy regions classification.

In a similar way, if we keep fixed the CR value μ and move the IF value ξ we obtain that the energy is negligible outside an interval which depends on the slopes μ, α_1 y α_2, see Fig. 3. More precisely, writing $\xi_m = \min\{\beta_1+\lambda(\mu-\alpha_1), \beta_1-\lambda(\mu-\alpha_1)\}$ and $\xi_M = \max\{\beta_1+\lambda(\mu-\alpha_1), \beta_1-\lambda(\mu-\alpha_1)\}$, we have that the energy of the chirplet transform is concentrated in the interval (ξ_m, ξ_M).

Summarizing, for a one-chirp signal of the linear form (2) we have that, for fixed time, its energy is concentrated in a region of the CR-IF plane of the form

$$(-\infty, \mu_m(\xi)) \cup (\mu_M(\xi), +\infty) \times \mathbb{R} \equiv \mathbb{R} \times (\xi_m(\mu), \xi_M(\mu))$$

which illustrates de double cone shape in Fig. 1.

2.2 Spurious Maxima

For the more interesting case of a signal composed by several chirps, the energy distribution will have maxima not only at ridge points but at the so called spurious maxima points. To understand this situation, we consider a signal composed by two linear chirps

$$f(t) = \sum_{k=1}^{2} a_k(t) \exp\left[i\left(\frac{\alpha_k}{2}(t-\tau)^2 + \beta_k(t-\tau) + \gamma_k\right)\right]$$

Considering again a window centered at τ and of width λ the time-frequency representation of the signal will be a pair of segments defined on the time interval $(\tau-\lambda, \tau+\lambda)$. If we keep constant the slope value μ and move the IF value ξ, we see that the energy distribution will be negligible outside an interval which depends on μ, α_1 y α_2. Since the signal is composed by two chirps, there arises a set S_2 in which the segment $r_\mu(\xi)$ intersects both chirps, a set S_1 in which the segment intersects only one chirp and a set S_0 with empty intersection, see Fig. 4, and therefore where the energy is negligible.

Since the instantaneous amplitudes are continuous there exists a maximum point in the interval $[\max\{\xi_{1,m}, \xi_{2,m}\}, \min\{\xi_{1,M}, \xi_{2,M}\}]$, which is not a ridge point. It is a *spurious maximum*. We note that these maxima do appear for large slope values and IF's localized between the IF's of the corresponding chirps at time τ.

3 CHIRP-TRACKING

Let us consider the multi-chirp signal

$$f(t) = \sum_{k=1}^{N} a_k(t) \exp[i\phi_k(t)]$$

with $\phi_k \in C^3$, i.e., such that there exists M such that $|\phi_k'''(t)| < M$ for $k = 1, \ldots, N$ and for all t.

Assume that we know that at $(\tau, \phi_k'(\tau))$ our signal has a ridge point in the time-frequency plane and that its CR is given by $\phi_k''(\tau)$. Then, for $h > 0$ small we have

$$\phi_k'(\tau+h) = \phi_k'(\tau) + h\phi_k''(\tau) + \frac{\phi_k'''(\nu)}{6}h^2$$

with $\nu \in (\tau, \tau+h)$. From the regularity assumed on f we deduce

$$\left|\phi_k'(\tau+h) - \phi_k'(\tau) - h\phi_k''(\tau)\right| \leqslant \frac{1}{6}Mh^2.$$

Therefore, once we know the IF and CR at a given time, we may estimate the range in which the CR will be at the next time step.

On the other hand, we have

$$\phi_k''(\tau+h) \approx \phi_k''(\tau) + h\phi_k'''(\tau)$$

i.e., $\phi_k''(\tau+h) \in (\phi_k''(\tau) - hM, \phi_k''(\tau) + hM)$. Therefore, we may also estimate the range of slopes for the CR at subsequent times.

Our algorithm is organized as follows:

1. We use a the signal spectrogram to choose the time components of points at which the spectrogram has maxima, $t = \tau_1, ..., \tau_L$. These are candidates to belong to the support of the α_k functions.
2. For each τ_l we compute the chirplet energy distribution corresponding to $\Psi f(\tau_l, \xi, \mu; \lambda)$ for a discrete set of slopes, $\mu_1, ..., \mu_r$, using the DFT.
3. We compute the maximum of the energy obtaining an estimate of the CR and IF $(\mu^l_{0,1}, \xi^l_{0,1})$ of one chirp at the given time.
4. We advance one step in time and compute the energy at $\tau_l + h$.
5. To avoid spurious maxima, we seek the maximum through the window estimated in the previous section, i.e.

$$\mu \in (\mu^l_{0,1} - hM, \mu^l_{0,1} + hM)$$

and

$$\xi \in \left(\xi^l_{0,1} + h\mu^l_{0,1} - \frac{1}{6}Mh^2, \xi^l_{0,1} + h\mu^l_{0,1} + \frac{1}{6}Mh^2\right)$$

This maximum provides us with an estimate for the CR and IF, $(\mu^l_{1,1}, \xi^l_{1,1})$.

6. We continue this process until the energy maximum in the selected window is lower than some fixed threshold. In this way, we obtain a set of ridge points corresponding to the same chirp.
7. We iterate steps 2-6 on every point computed at step 1, until all the chirps of signal f are identified.

We finally notice that this new algorithm narrows the search area for maxima points in the CR-IF plane, implying a gain in computational time as well as a reduction in the spurious maxima processing.

4 NUMERICAL EXPERIMENTS

4.1 Experiment 1. A Synthetic Signal

We consider a synthetic signal composed by the addition of three chirps (see Fig.5) $f(t) = \sum_{i=1}^{3} a_i(t)\cos[\phi_i(t)]$ with $t \in [0, 10]$ and with amplitudes given by amplitudes

$$a_1(t) = \frac{1}{3}\exp\left(-\frac{(t-3.5)^2}{5}\right)X_{[1,6]}(t),$$

$$a_2(t) = \frac{1}{3}\exp\left(-\frac{(t-5)^2}{4}\right)X_{[2,8]}(t),$$

$$a_3(t) = \frac{1}{3}\exp\left(-\frac{(t-6)^2}{4}\right)X_{[3.5,9.5]}(t),$$

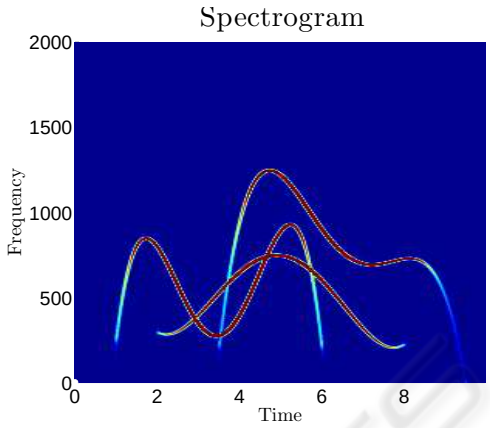

Figure 5: Spectrogram of synthetic signal of Experiment 1.

with $X_{[a,b]}(t)$ the characteristic function of the interval $[a, b]$, and with phases

$$\phi_1(t) = -13.07t^5 + 226.2t^4 - 1434t^3 + 4055t^2 - 4408t,$$

$$\phi_2(t) = 1.733t^5 - 42.92t^4 381.1t^3 - 1459t^2 + 2798t,$$

$$\phi_3(t) = -4.973t^5 + 166.3t^4 - 2172t^3 + 0.138t^2 + 0.411t.$$

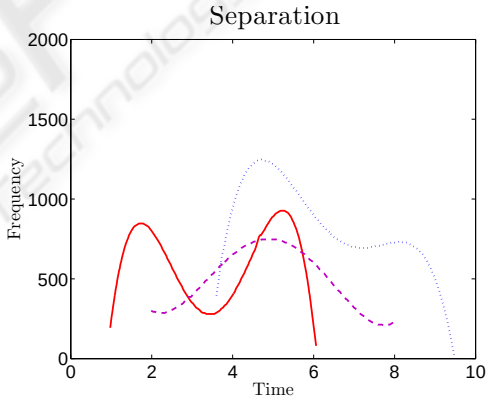

Figure 6: Separation of the three chirps of Experiment 1.

We used a variable time step $h \in (0.01, 0.05)$ sec. The initial step is 0.01 sec. and in subsequent iterations we use $h = 0.01c + 0.05(1-c)$ with $c = \min(|\mu|, 1000)/1000$, for μ the CR estimation of the previous step. Finally, we neglect maxima with amplitude lower than 0.001. Relative error results for amplitude, IF and CR are plotted in Figs. 9, 10 and 11. We observe that, specially the IF, is very accurately approximated, allowing the separation of the three chirps, see Fig. 6. In a further experiment not shown in this paper, we repeated the previous experiment adding a noise to the signal with SNR=0. In this case, the amplitude threshold was set to 0.025 and the relative error results were very similar to those exhibited by Figs. 9-11.

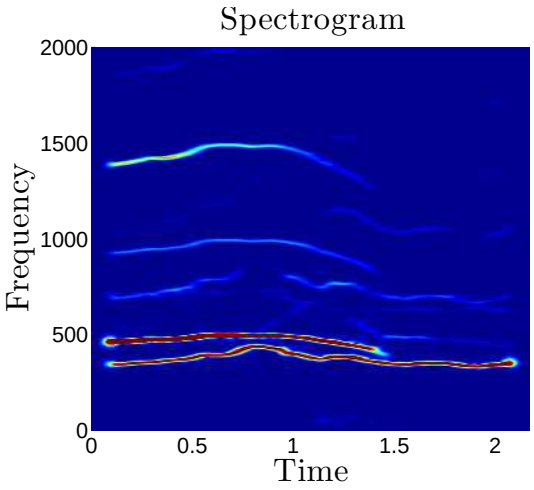

Figure 7: Signal spectrogram corresponding to Experiment 2.

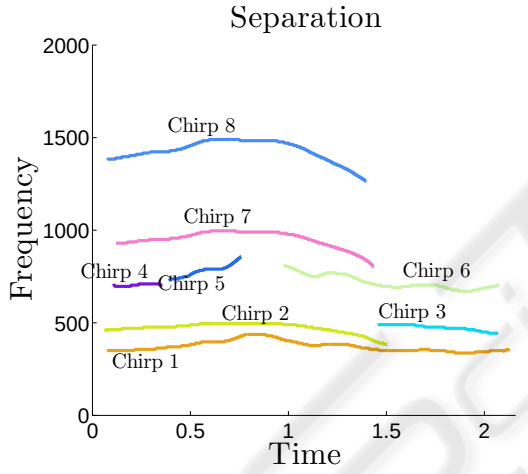

Figure 8: Separation result corresponding to Experiment 2.

4.2 Experiment 2. A Field Recorded Wolves Chorus

We use a wolf chorus field recorded signal. Parameters are set as follows: window width, $\lambda = 0.15$, time step $h \in (0.001, 0.005)$ sec. and the amplitude threshold to 0.006. Fig. 7 shows the signal spectrogram while Fig. 8 shows the separation result.

In Fig. 8 we observe the detection of eight chirps. Due to the threshold level election some chirps are split into several pieces due to the usual variation in the amplitude of wolf howls, e. g. chirps 4, 5 and 6 in Fig. 8. However, if we consider a lower amplitude threshold, the noise present in the recording induces the erroneous identification of some points as ridge

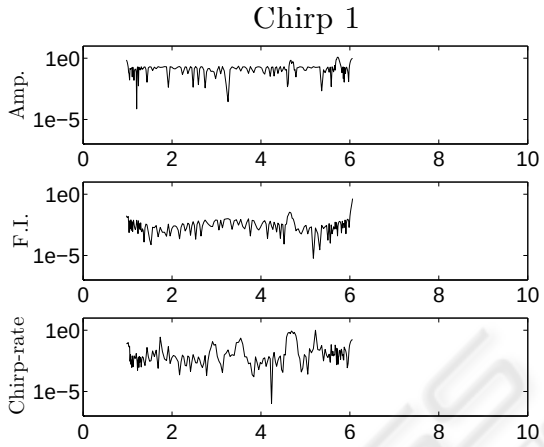

Figure 9: Relative errors for chirp 1 of Experiment 1.

points. Hence, a suitable threshold level must be set to balance these effects.

5 CONCLUSIONS

Many biological audio signals may be modelled as an addition of modulated chirps being a problem of interest the separation or decomposition of the signal in its fundamental components. For this separation, specially at chirps crossing points in the time-frequency plane, some additional information to instantaneous frequency is needed in order to track correctly the different components. In a previous work, we introduced an algorithm using the chirp-rate estimation based on the chirplet transform. Some drawbacks of this algorithm were the difficulties for defining an accurate matching algorithm for the chirp segments computed in each time step and the accuracy on the chirp-rate computation, specially in which respects to the detection of spurious maxima. In this paper we presented an improved algorithm which avoids the previous matching algorithm by means of a straight chirp tracking procedure, based on an improved methodology for energy maxima detection and spurious maxima identification.

ACKNOWLEDGEMENTS

All authors are supported by Project PC07-12, Gobierno del Principado de Asturias, Spain. Third and fourth authors are supported by the Spanish DGI Project MTM2007-65088.

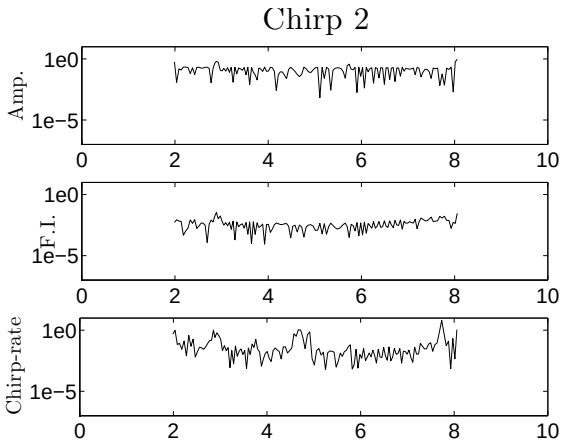

Figure 10: Relative errors for chirp 2 of Experiment 1.

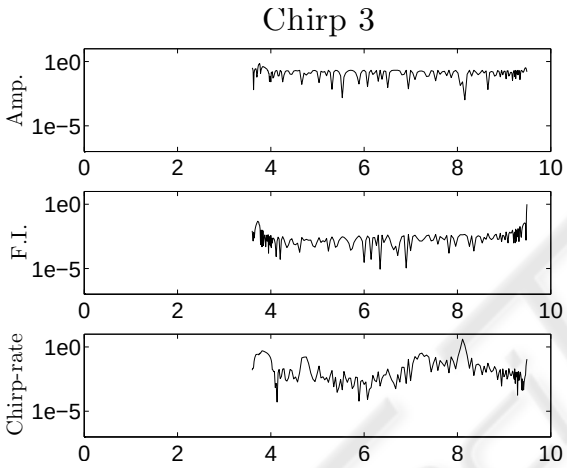

Figure 11: Relative errors for chirp 3 of Experiment 1.

REFERENCES

Boashash, B. (1992a). Estimating and interpreting the instantaneous frequency of a signal. i. fundamentals. *Proceedings of the IEEE*, 80(4):520–538.

Boashash, B. (1992b). Estimating and interpreting the instantaneous frequency of asignal. ii. algorithms and applications. *Proceedings of the IEEE*, 80(4):540–568.

Dugnol, B., Fernández, C., Galiano, G., and Velasco, J. (2008). On a chirplet transform based method applied to separating and counting wolf howls. *Signal Proc.*, 88(7):1817–1826.

et al., N. E. H. (1998). The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London A*, 454(1971):903–995.

Gianfelici, F., Biagetti, G., Crippa, P., and Turchetti, C. (2007). Multicomponent am-fm representations: An asymptotically exact approach. *IEEE Trans. Audio Speech Lang. Process.*, 15(3):823–837.

Kwok, H. and Jones, D. (2000). Improved instantaneous frequency estimation using an adaptive short-time fourier transform. *IEEE Trans. Signal Process.*, 48(10):2964–2972.

Mann, S. and Haykin, S. (1991). The chirplet transform: A generalization of gabor's logon transform. In *Proc. Vision Interface 1991*, pages 205–212.

Maragos, P., Kaiser, J. F., and Quatieri, T. F. (1993). Energy separation in signal modulations with applications to speech analysis. *IEEE Trans. Speech Audio Process.*, 41(10):3024–3051.

McAulay, R. J. and Quatieri, T. F. (1986). Speech analysis/synthesis based on a sinusoidal representation. *IEEE Trans. Acoustics Speech Signal Process.*, 34(4):744–754.

Santhanam, B. and Maragos, P. (2000). Multicomponent am-fm demodulation via periodicity-based algebraic separation and energy-based demodulation. *IEEE Trans. Commun.*, 48(3):473–490.

Stanković, L. and Djurović, I. (2003). Instantaneous frequency estimation by using the wigner distribution and linear interpolation. *Signal Process.*, 83(3):483–491.

Zhao, Z., Pan, M., and Chen, Y. (2004). Instantaneous frequency estimate for non-stationary signal. In *Fifth World Congress on Intelligent Control and Automation, 2004. WCICA 2004.*, volume 4, pages 3641–3643.