3rd Congress of the Alps Adria Acoustics Association
27–28 September 2007, Graz – Austria

CHIRP-RATE ESTIMATION FOR TIME-FREQUENCY REASSIGNED SPECTROGRAMS

Franz Zotter[1], Robert Höldrich[1]
[1)] Institute of Electronic Music and Acoustics, http://iem.at,
University of Music and Dramatic Arts, Graz, Austria
E-Mail: {zotter,hoeldrich}@iem.at

Abstract: *Spectrogram analysis frequently suffers from inadequate time or frequency resolution, respectively. However, time-frequency reassigned spectrograms offer an attractive extension regarding the resolution trade-off. Specifically, this method deduces an improved time-frequency localization from the derivatives of the spectrogram phase, with respect to time and frequency. For high-quality images of reassigned spectrograms, however, there is a requirement for inconveniently short hop-sizes. Recently, the information of the cross-derivatives was introduced as local indicator for sinusoidal vs. impulsive signal characteristics. In this contribution, we try to extend the reassigned spectrogram points to lines, in order to allow for bigger hop-sizes. To this end, the second-order derivatives give coarse intermediate estimates of the chirp-rate. Using the cross-derivative indicator, we combine these intermediate estimates to a more accurate chirp-rate estimator. The properties of the novel time-frequency distribution is demonstrated in several examples.*

Key words: Time-frequency reassigned spectrogram, chirp-rate estimation, spectral tracks

1. INTRODUCTION

The magnitude of the short-time Fourier transform (STFT) can easily be calculated for a quantization grid in time and frequency. Using the window size N and a constant time-interval N_{hop} for the periodical discrete Fourier-transforms (DFTs) within the STFT, we obtain an analsyis grid at discrete times $t[n] = n \cdot N_{hop}/f_s$ and frequencies $f[k] = f_s/N$. As described in literature, utilization of smaller N_{hop} or zero-padding the DFT for the visual interpretation of the spectrogram, can only bring moderate improvements (cf. Auger [1]). The t-f-resolution in a strict sense, is not enhanced by this attempt.

However, it is possible to refine time-frequency localization in the short-term Fourier-transform. We may assume that an STFT-bin contains either a rather sinusoidal, or rather impulse-like signal. Using the derivatives of the phase-spectrum, we indeed obtain refined t-f-localization for the STFT evaluated at $n = mN_{hop}$ and $k = 0 \ldots N/2$, cf. [1]:

$$\hat{\eta}\left(X\left[n,k\right]\right) = \frac{\partial \Phi_x\left[n,k\right]}{\partial \omega} - n \qquad (1)$$

$$\hat{\omega}\left(X\left[n,k\right]\right) = \frac{\partial \Phi_x\left[n,k\right]}{\partial n} \cdot \frac{N}{2\pi} - k. \qquad (2)$$

Basically, the above equations may be interpreted as estimates of the group-delay (GD) and the instantaneous frequency (IF) that refine the localization in the STFTs.

Furthermore, Nelson [4] and Folup [7] offer an interpretation of the mixed derivative:

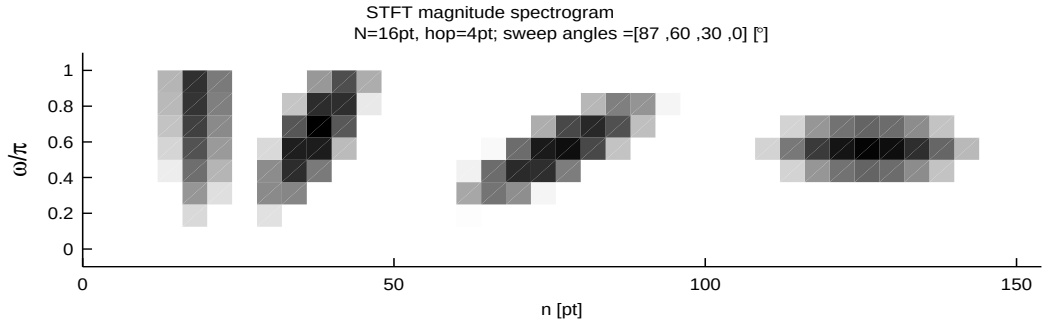
$$\frac{\partial^2 \Phi_x\left[n,k\right]}{\partial \omega \partial n} = \begin{cases} \to -1 & \text{stationary,} \\ \to 0 & \text{impulsive,} \end{cases} \qquad (3)$$

i.e. this qualifier describes the *impulsiveness* of the signal. In the following sections, we demonstrate the performance of the other second order phase-derivatives as chirp-rate estimators. We introduce a new chirp-rate estimation, combining the advantage of both estimators. The estimator provides an improvement to specrogram analysis for visual inspection, and might prove useful in complex spectral modeling tasks.

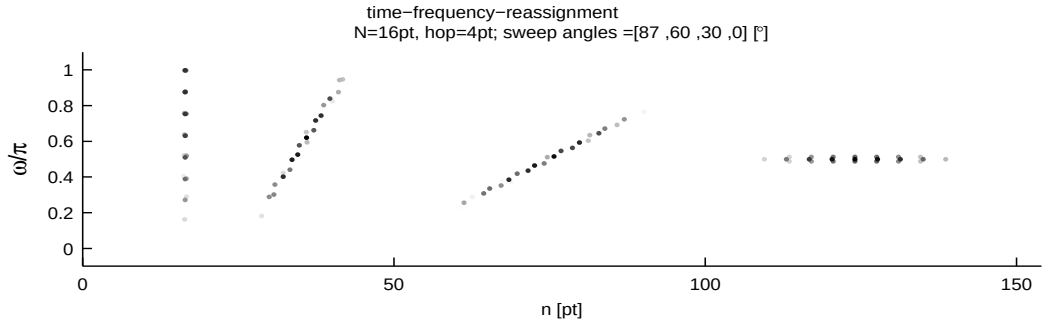
2. TIME-FREQUENCY REASSIGNMENT

2.1. First-Order *Fast Phase Derivatives*

There are several publications (Auger [1], Nelson [4]) describing the computation of the derivatives in frequency domain, using discrete Fourier transforms of differently windowed versions of the sigal $x[n]$. According to Fitz [6], we



(a) STFT-magnitude spectrogram



(b) Time-frequency reassigned spectrogram

Figure 1: These two figures show a comparison between the STFT spectrogram 1(a), and the time-frequency reassigned spectrogram 1(b).

may call these estimates *fast phase derivatives*. We need the following STFT-spectra:

$$X\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w[n]\right\} \quad (4)$$
$$X_t\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w_t[n]\right\} \quad (5)$$
$$X_d\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w_d[n]\right\}. \quad (6)$$

In the above equations, $w[n]$ is a symmetric window function (e.g. hanning) that determines the other time- and frequency derivative window functions:

$$w_t[n] = n\cdot w[n] \quad (7)$$
$$w_d[n] = \frac{d}{dn}w[n]. \quad (8)$$

The following formulae for the *fast* first-order *phase derivatives* are given in Fitz [6]:

$$\hat{\eta}[n,k] := \frac{\partial\Phi_x\left[n,k\right]}{\partial\omega} - n \quad (9)$$
$$= -\Re\left\{\frac{X_t\left[n,k\right]}{X\left[n,k\right]}\right\},$$
$$\hat{\omega}[n,k] := \frac{\partial\Phi_x\left[n,k\right]}{\partial n} - \frac{2\pi}{N}\cdot k \quad (10)$$
$$= \Im\left\{\frac{X_d\left[n,k\right]}{X\left[n,k\right]}\right\}. \quad (11)$$

2.2. Time-Frequency Reassignment

As stated above, the first and the second expression Eq. 9 and Eq. 10 have a very simple interpretation. $\hat{\eta}[n,k]$ is the refined location of the component $[n,k]$ in time (time-reassignment), whereas $\hat{\omega}[n,k]$ refines the localization in frequency (frequency-reassignment). An example thereof is illustrated in Fig. 1.

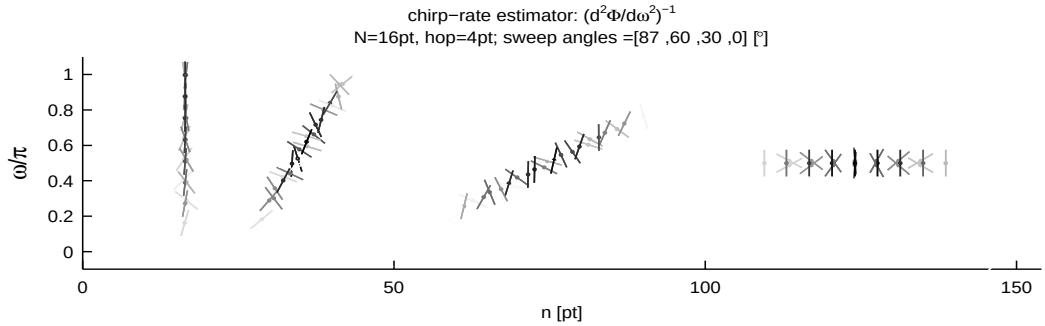
3. SECOND ORDER FAST PHASE DERIVATIVES

According to Fitz [6], there also exist fast phase derivative formulae for the pure and mixed second order derivatives, using:

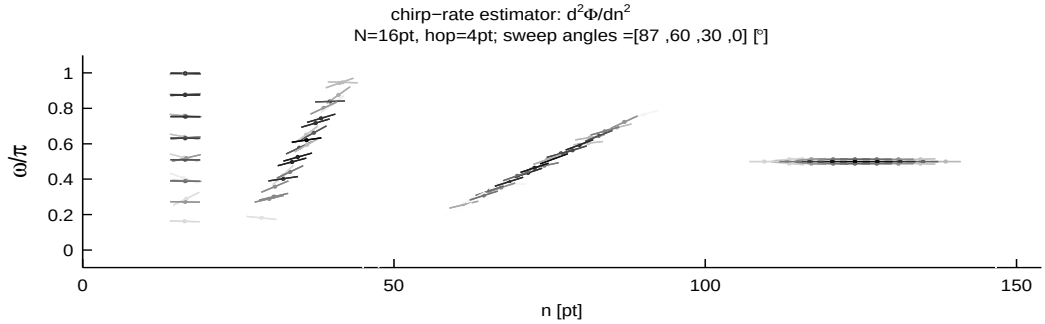
$$X_{dt}\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w_{dt}[n]\right\} \quad (12)$$
$$X_{tt}\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w_{tt}[n]\right\} \quad (13)$$
$$X_{dd}\left[m,k\right] = \mathcal{DFT}\left\{x[n+mN_{hop}]\cdot w_{dd}[n]\right\}, \quad (14)$$

and the windowing functions:

$$w_{dt}[n] = n\cdot\frac{d}{dn}w[n] \quad (15)$$
$$w_{tt}[n] = n^2\cdot w[n] \quad (16)$$
$$w_{dd}[n] = \frac{d^2}{dn^2}w[n]. \quad (17)$$



(a) Time-frequency reassigned spectrogram with the chirp-rate estimation $\gamma_{IF}[n,k]$.



(b) Time-frequency reassigned spectrogram with the chirp-rate estimation $\gamma_{GD}[n,k]$.

Figure 2: The chirp-rate estimation with.

In particular, the derivatives are:

$$\frac{\partial^2\Phi_x[n,k]}{\partial\omega\partial n} = \Re\left\{\frac{X_{dt}[n,k]}{X[n,k]}\right\} - \Re\left\{\frac{X_t[n,k]\,X_d[m,k]}{X^2[n,k]}\right\}, \quad (18)$$

$$\frac{\partial^2\Phi_x[n,k]}{\partial\omega^2} = \Im\left\{\left(\frac{X_t[n,k]}{X[n,k]}\right)^2\right\} - \Im\left\{\frac{X_{tt}[n,k]}{X[n,k]}\right\}, \quad (19)$$

$$\frac{\partial^2\Phi_x[n,k]}{\partial n^2} = -\Im\left\{\left(\frac{X_d[n,k]}{X[n,k]}\right)^2\right\} + \Im\left\{\frac{X_{dd}[n,k]}{X[n,k]}\right\}. \quad (20)$$

3.1. CHIRP-RATE ESTIMATION

We can interpret the second-order phase derivatives as:

- $\partial^2\Phi/\partial\omega\partial n$: the *impulsiveness* of the signal (cf. Fulop [7], Eq. 3)
- $\partial^2\Phi/\partial n^2$: change rate of the instantaneous frequency for approximately stationary signals, i.e. when $\partial^2\Phi/\partial\omega\partial n \to -1$,
- $\partial^2\Phi/\partial\omega^2$: change rate of group-delay over frequency, i.e. dispersive characteristics of impulsive signals, when $\partial^2\Phi/\partial\omega\partial n \to 0$.

The following formulae show, how the two pure second order derivatives can be used as chirp-rate estimators:

$$\gamma_{IF}[n,k] = \frac{\partial^2\Phi_x[n,k]}{\partial n^2} \quad (21)$$

$$\gamma_{GD}[n,k] = \left(\frac{\partial^2\Phi_x[n,k]}{\partial n^2}\right)^{-1}. \quad (22)$$

These two estimates provide a linear extrapolation $\Omega[n+n_{obs},k]$ of the instantaneous frequency reassignment at time instants $n+n_{obs}$. Where $\hat{\eta}[n,k]$ is the time-frequency-reassignment in time, $\hat{\omega}[n,k]$ the reassignment in frequency, and n_{obs} the observed sample around the time index n of the analysis DFTs:

$$\Omega[n+n_{obs},k] = \hat{\omega}[n,k] + \gamma[n,k]\,(n_{obs} - \hat{\eta}[n,k]) \quad (23)$$

For graphical representations and some theoretical aspects, the description of the chirp rate as an angle within the discrete time-frequency grid is useful:

$$\alpha[n,k] = \arctan\left(\gamma[n,k]\cdot\frac{N}{2\pi}\right). \quad (24)$$

It can easily be shown that the two second order derivatives w.r.t. n and ω tend to be very imprecise if the underlying signals are either too impulsive, or too stationary. The signals were Gauss-windowed chirps, each centered at the bin-frequency $N/4+0.3$. An illustration thereof can be found in Fig. 2.

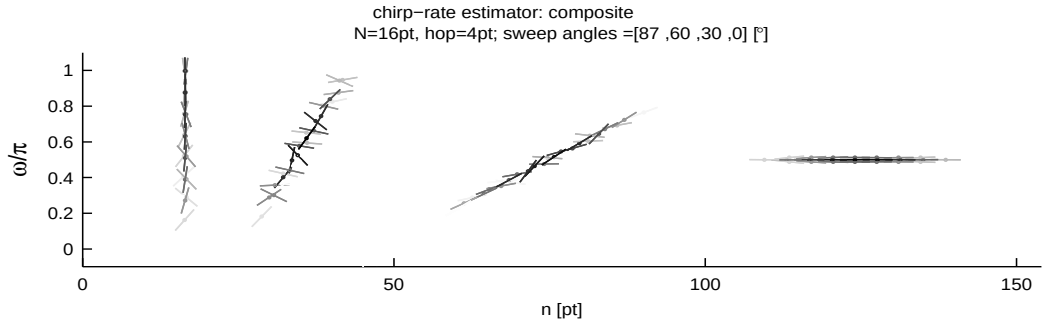


Figure 3: It can be clearly seen that the composite estimator combines some advantages of both estimators. On the other hand some error remains.

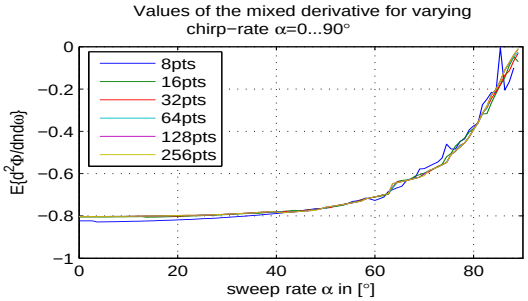


Figure 4: Average values of the mixed derivative w.r.t. varying chirp-rates at different DFT sizes.

3.2. Mixed Derivative Indicator

We illustrate the behaviour of the mixed second order derivative in Fig. 4, which is in fact similar to the description in Eq. 3. The chirp-rate angle proved to be very useful as a common parameter to several DFT sizes.

4. COMPOSITE CHIRP-RATE ESTIMATION

Combining the two second order derivatives from the previous section using a weighted geometric mean, we create an new composite chirp-rate estimate $\gamma\,[n,k]$ that shows a significantly improved performance, as shown in Fig 3:

$$\gamma\,[n,k] \;=\; \Re\left\{\gamma_{IF}[n,k]^{-g\left\{\frac{\partial^2\Phi_x[n,k]}{\partial\omega\partial n}\right\}}\times \gamma_{GD}[n,k]^{1+g\left\{\frac{\partial^2\Phi_x[n,k]}{\partial\omega\partial n}\right\}}\right\}. \qquad (25)$$

We chose a distortion function $g\{x\} = x$. Further investigation on different kinds of distortions seem to be useful, in particular regarding the results in Fig. 4.

4.1. Example

The properties of the new chirp rate estimator are illustrated in Fig. 3. It is obvious that it outperforms the other variants, cf. Fig. 2. Remaining errors can also be clearly identified.

5. CONCLUSION

We have shown an extension to the time-frequency reassignment technique applying chirp-rate estimation. In addition to time-frequency localisation, this method provides information about the slope of the regarded spectrogram component. Several examples of the different estimators could be given, and a comparison to the new estimator has been made. Prospective applications might not only be found in spectrogram illustration, but also in computer assisted interpretation of specral data. Nevertheless, furhter improvements of this technique are required, and will pose a future topic in our research.

6. ACKNOWLEDGEMENTS

We gratefully thank the Zukunftsfonds Steiermark (Prj. 3027) for supporting our research.

REFERENCES

[1] F. Auger, P. Flandrin: Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment-Method. IEEE Transactions on Signal Processing, vol. 43, no. 5, pp. 1068–1089, 1995.

[2] S.W. Hainsworth, M.D. Macelod: Time Frequency Reassignment: A Review and Analysis. Tech. Rep. CUED/FINFENG/TR.459, Camebridge University, Eng. Dep., 2003.

[3] : D.J. Nelson: Cross-Spectral Methods for Processing Speech. Journal of the Acoust. Soc. of Am., 110, pp. 2575–2592, 2001.

[4] : D.J. Nelson: Instantaneous Higher Order Phase Derivatives. Digital Signal Processing 12, pp. 416–428, 2002.

[5] K.R. Fitz, L. Haken: On the Use of Time-Frequency Reassignment in Additive Sound Modeling. Journal of the AES, vol. 50(11), pp. 879–893, November, 2002.

[6] K.R. Fitz, S.A. Fulop: A Unified Theory of Time-Frequency Reassignment. Signal Processing, Elsevier, Sept., 2005.

[7] S.A. Fulop K.R. Fitz: Separation of Components from Impulses in Reassigned Spectrocgrams. Journal of the Acoust. Soc. of Am., 121(3), pp. 1510–1518, 2007.