

INSTANTANEOUS FREQUENCY ESTIMATION BY GROUP DELAY ATTRACTORS AND INSTANTANEOUS FREQUENCY ATTRACTORS

Soo-Chang Pei and Shih-Gu Huang

Graduate Institute of Communication Engineering, National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., 10617, Taipei, Taiwan, ROC
Email: pei@cc.ee.ntu.edu.tw, d98942023@ntu.edu.tw

ABSTRACT

Instantaneous frequency attractors (IFAs), obtained from the phase of a time-frequency representation, have been introduced for instantaneous frequency (IF) estimation. In this paper, another kind of attractors called group delay attractors (GDAs) are proposed to improve the IFA-based method. The GDAs can reveal IFs which cannot be estimated from the IFAs. Simulation results show that the IF estimation method based on both the GDAs and IFAs outperforms the well-known estimation method, i.e. ridge detection. Also, it is shown that the proposed method creates much less spurious IFs than the IFA-based method in noisy environments.

Index Terms— Instantaneous frequency estimation, local group delay, local instantaneous frequency, time-frequency representation

1. INTRODUCTION

Instantaneous frequency (IF) is an important characteristic used to describe the mechanisms for nonlinear and nonstationary processes. It arises in various areas such as communications, acoustics, speech, geophysics and biomedicine. Time-frequency (TF) analysis is one of the most popular and efficient techniques in IF estimation, especially for multicomponent signals. An outstanding time-frequency representation (TFR)-based IF estimation method contains a signal-dependent TFR with high energy concentration and a well-designed IF estimation method. The focus of this paper is on the latter no matter the given TFR is satisfactory or not.

Since TFRs can concentrate the energy of a signal along its IFs, the IFs are commonly estimated from the positions of the local maxima on the TF energy distribution [1], also known as ridge detection. Besides the envelope of the TFR, the phase part can also be used for IF estimation. In [2], Abe and Honda proposed an estimation method based on IF attractors (IFAs). The IFAs are local IFs, calculated from the phase of the TFR, satisfying two particular conditions. If the TFR has high enough energy concentration, the IFAs can exactly reflect the actual IFs of the signal. It has been shown that the IF curves estimated from the IFAs are more smooth and accurate, while ridge detection yields stepwise IF contours.

In this paper, the IFA-based method is improved by introducing another kind of attractors, called group delay attractors (GDAs). Similar to the IFAs, the GDAs are obtained from the local group delays (GDs), which are also calculated from the phase of the TFR. Since the IFs can also be estimated from the GDAs, an IF estimation method based on both the IFAs and GDAs is proposed. Simulation results show that the GDAs can help reveal the IFs that cannot be estimated from the IFAs due to the unsatisfactory energy concentration of the TFR. Plus, it is shown that the proposed method provides more accurate IF estimation with less artifacts, compared with ridge detection on S-method (SM) [3] and ridge detection on re-assigned SM [4]. In noisy environments, besides the accuracy of the estimated IFs, the number of spurious IFs caused by noise is also a critical concern, especially when the actual number of IFs is unknown. Accordingly, another estimation strategy that uses the IFAs, GDAs and two TFRs with different window widths is presented. Statistical results demonstrate that the proposed method creates much less spurious IFs than the IFA-based method.

The concept of local IFs and local GDs has been used in various reassignment methods [5, 6]. However, in most of these methods, the local IFs and the local GDs are used to produce a more concentrated TFR. In [7], the authors estimate the IFs based on the similar notion as the IFAs, but the GDAs are not considered. Although the work in [8] estimates the IFs directly from the local IFs and/or local GDs, there are more spurious IFs because the determination of whether or not the local IFAs/GDAs are the desired IFs is less rigorous.

2. REVIEW OF INSTANTANEOUS FREQUENCY ATTRACTORS

In the sinusoidal model of [2], the IFs are estimated from the phase of a short-time Fourier transform (STFT). The STFT of a signal $x(t)$ with the kernel function $h(t)$ is defined as

$$X(t, \omega) = \int_{-\infty}^{\infty} x(\tau)h(\tau - t)e^{-j\omega\tau} d\tau. \quad (1)$$

With fixed ω , $F(t, \omega) = e^{j\omega t}X(t, \omega)$ can be deemed as the bandpass signal of $x(t)$ at frequency ω . The local IF is defined

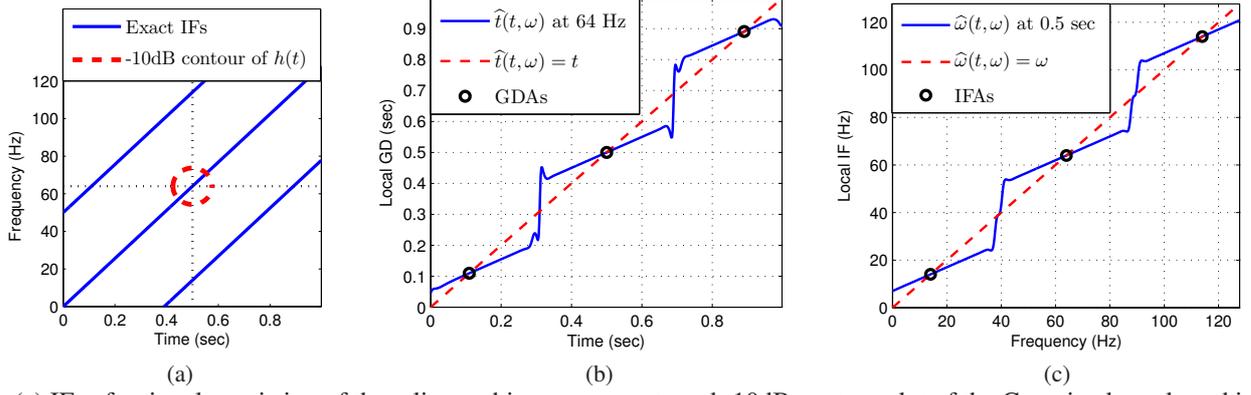


Fig. 1. (a) IFs of a signal consisting of three linear chirp components and -10dB contour plot of the Gaussian kernel used in the STFT. (b) Local GD $\hat{t}(t, \omega)$ versus time t (the solid line) and the GDAs (marked by small circles) at 64 Hz. (c) Local IF $\hat{\omega}(t, \omega)$ versus frequency ω (the solid line) and the IFAs (marked by small circles) at time instant $t = 0.5$ sec.

as the IF (i.e. first derivative of the phase) of $F(t, \omega)$:

$$\hat{\omega}(t, \omega) = \frac{\partial}{\partial t} \arg [F(t, \omega)] = \omega + \frac{\partial}{\partial t} \arg [X(t, \omega)]. \quad (2)$$

To avoid the discontinuity problem when calculating the phase, assume $X(t, \omega) = a + jb$ such that

$$\hat{\omega}(t, \omega) = \omega + \frac{a \frac{\partial b}{\partial t} - b \frac{\partial a}{\partial t}}{a^2 + b^2}, \quad (3)$$

where the time derivative of $X(t, \omega)$ can be alternatively calculated by replacing $h(t)$ in (1) by $\partial h(t)/\partial t$:

$$\frac{\partial X(t, \omega)}{\partial t} = - \int_{-\infty}^{\infty} x(\tau) \frac{\partial}{\partial t} h(\tau - t) e^{-j\omega\tau} d\tau. \quad (4)$$

Define $\mu(t, \omega) = \hat{\omega}(t, \omega) - \omega$. Then, at each time instant t , the IFAs are defined as the TF points satisfying

$$\mu(t, \omega) = 0 \quad \text{and} \quad \frac{\partial}{\partial \omega} \mu(t, \omega) < 0. \quad (5)$$

Due to the lack of a clear explanation for the above two conditions in [2], later together with the GDAs, a simple example will be given to illustrate why these two conditions make the IFAs reflecting the IFs.

3. PROPOSED INSTANTANEOUS FREQUENCY ESTIMATION METHOD

In this section, another kind of attractors called GDAs are introduced to improve the IFA-based method. Although the STFT is still adopted to explain the notion of the GDAs, in the end of this section, the extension to quadratic TFRs will be discussed.

3.1. Definition of group delay attractors (GDAs)

For a linear chirp $x(t) = \exp(j\frac{a}{2}t^2)$, it is apparent that the IF is $\omega(t) = \frac{\partial}{\partial t} \arg [x(t)] = at$. Also, the IF can be determined from the group delay (GD) of the signal, defined as $t(\omega) =$

$-\frac{\partial}{\partial \omega} \arg [X(\omega)] = \frac{1}{a}\omega$. Applying this notion to the STFT in (1), for each time instant t , the local GD is obtained from the frequency function $X(t, \omega)$,

$$\hat{t}(t, \omega) = -\frac{\partial}{\partial \omega} \arg [X(t, \omega)]. \quad (6)$$

Similar to (3), assume $X(t, \omega) = a + jb$ such that the local GD can be calculated from

$$\hat{t}(t, \omega) = -\frac{a \frac{\partial b}{\partial \omega} - b \frac{\partial a}{\partial \omega}}{a^2 + b^2}. \quad (7)$$

The frequency derivative of $X(t, \omega)$ can be alternatively carried out from the following formula:

$$\begin{aligned} \frac{\partial}{\partial \omega} X(t, \omega) &= \int_{-\infty}^{\infty} x(\tau) (-j\tau \cdot h(\tau - t)) e^{-j\omega\tau} d\tau \\ &= -j \int_{-\infty}^{\infty} x(\tau) (\tau - t) h(\tau - t) e^{-j\omega\tau} d\tau - jtX(t, \omega), \end{aligned} \quad (8)$$

where the first term of (8) is equivalent to the STFT in which $h(t)$ is replaced by $th(t)$. Define $\nu(t, \omega) = \hat{t}(t, \omega) - t$. At each frequency ω , the GDAs are defined as TF points satisfying

$$\nu(t, \omega) = 0 \quad \text{and} \quad \frac{\partial}{\partial t} \nu(t, \omega) < 0. \quad (9)$$

An example is given in Fig. 1 to illustrate why the above two conditions make it possible to estimate the IFs from the GDAs. Fig. 1(a) displays the IFs of a signal comprising three linear chirp components. If the kernel of the STFT is a Gaussian function, its TF distribution is a 2D Gaussian mask with -10dB contour shown as the ellipse in Fig. 1(a). The local GD at TF point p can be deemed as the local mean time (weighted by the mask) with p as the center. As the Gaussian mask shifts right, it would contain the part of signal at higher time interval and lead to higher local GD. For example, for the TF points near (0.109 sec, 64 Hz), the leftmost component dominates the local GDs, and thus the local GDs are near 0.109 sec. Since there are three components at 64 Hz, the local GD

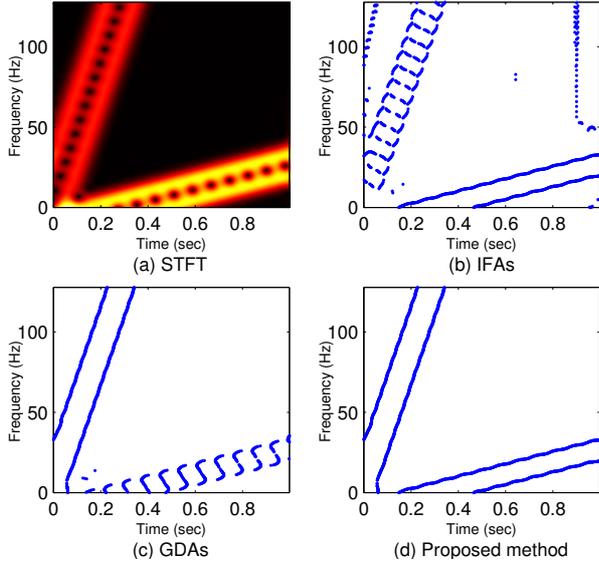


Fig. 2. (a) Spectrogram of a signal consisting of two smooth linear chirps and two sharp linear chirps, (b) IFAs, (c) GDAs, and (d) proposed IF estimation combining IFAs and GDAs (IF curves with too short length are eliminated).

$\hat{t}(t, \omega)$ versus time t at 64 Hz is a staircase-like function, as depicted in Fig. 1(b). If the second condition in (9) is not used, the TF points (0.309 sec, 64 Hz) and (0.69 sec, 64 Hz) will be misidentified as GDAs. Similarly, the local IF at TF point p can be deemed as the local mean frequency with p as the center. The local IF $\hat{\omega}(t, \omega)$ versus ω at $t = 0.5$ sec, as depicted in Fig. 1(c), is also a staircase-like function because there are three components at 0.5 sec. If the second condition in (5) is not used, the TF points (0.5 sec, 39 Hz) and (0.5 sec, 89 Hz) will be misidentified as IFAs.

3.2. IF estimation method based on GDAs and IFAs

Comparing (5) with (9), we can find out that the IFAs correspond to vertically estimating the IFs, while the GDAs correspond to horizontally estimating the IFs. Compared with vertical estimation, the horizontal estimation is more suitable to estimate rapidly varying IFs. Therefore, it can be expected that the GDAs may reveal the IFs which are not estimated from the IFAs, and it explains why the proposed IF estimation method takes both IFAs and GDAs into account. The algorithm of the proposed method is presented as follows:

1. The IFAs and GDAs are calculated from the given TFR.
2. The IFAs/GDAs are separated into several IF curves through curve tracing [9] on the TF plane.
3. The IF curves with too short length (depending on the length of the signal) are deemed as noise components or artifacts and discarded.
4. The remaining IF curves obtained from the IFAs and those from the GDAs are combined together by an OR operator.

If two TFRs with different window widths are available, the IFAs and GDAs can be obtained from different TFRs. In this

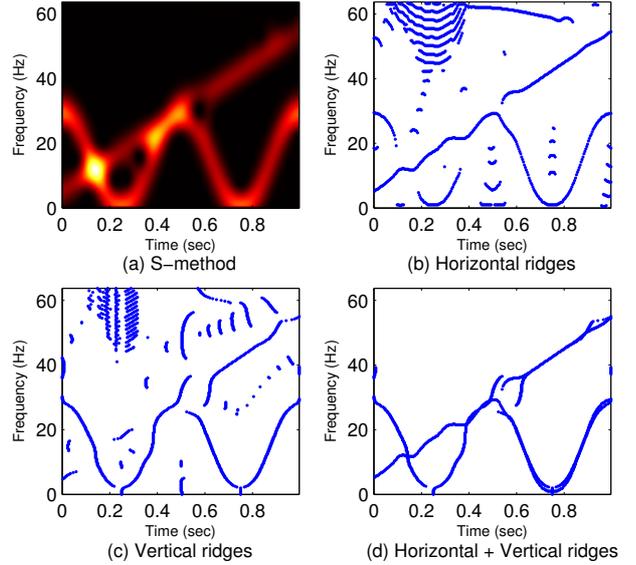


Fig. 3. (a) Magnitude of the S-method of a three-component signal, (b) horizontal ridges, (c) vertical ridges, and (d) combination of horizontal and vertical ridges after elimination of short-length ridges and energy thresholding.

case, the number of spurious IFs can be reduced, especially for noisy signals. Details will be illustrated directly by an example in Sec. 4.3

In [5], it has been indicated that the local IF and local GD are closely related to the TF reassignment; that is, (2) and (6) can also be expressed as

$$\hat{\omega}(t, \omega) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \eta W_x(\tau, \eta) W_h(t - \tau, \omega - \eta) d\tau d\eta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(\tau, \eta) W_h(t - \tau, \omega - \eta) d\tau d\eta}, \quad (10)$$

$$\hat{t}(t, \omega) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tau W_x(\tau, \eta) W_h(t - \tau, \omega - \eta) d\tau d\eta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x(\tau, \eta) W_h(t - \tau, \omega - \eta) d\tau d\eta}, \quad (11)$$

where $W_x(t, \omega)$ and $W_h(t, \omega)$ are the Wigner-Ville distributions of $x(t)$ and $h(t)$, respectively. Since $W_h(t, \omega)$ in (10) and (11) can be replaced by an arbitrary low-pass kernel, the proposed method can be extended to quadratic TFRs.

4. SIMULATION RESULTS

In all the simulations, STFT with Gaussian window function is adopted because the optimal window width can be easily determined from [10].

4.1. Comparisons between the proposed method and the IFA-based method

Consider a signal consisting of two smooth linear chirps and two sharp linear chirps, all of which have unit envelope. The STFT of the signal is shown in Fig. 2(a), and the used window width $\sigma^2 = 0.0013$ is not the optimal for either the smooth

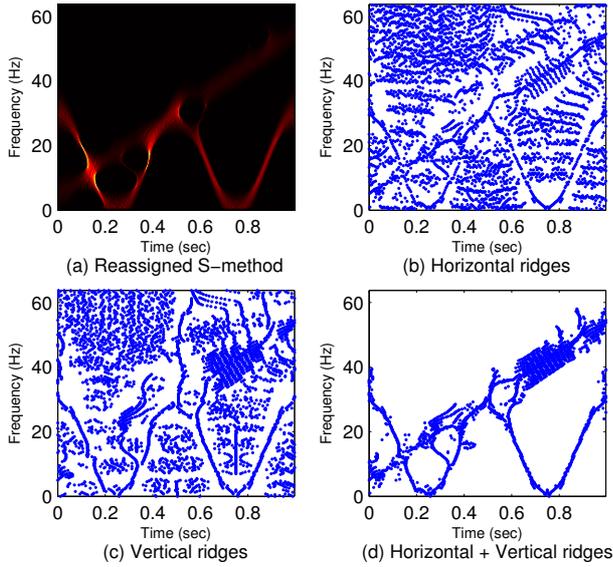


Fig. 4. (a) Magnitude of the reassigned S-method of a three-component signal, (b) horizontal ridges, (c) vertical ridges, and (d) combination of horizontal and vertical ridges after elimination of short-length ridges and energy thresholding.

chirps or the sharp ones. From the IFAs shown in Fig. 2(b), the IFs of the smooth chirps are estimated with high accuracy. However, the estimated IFs of the sharp chirps are discontinuous and there are a lot of artifacts between them. On the contrary, the GDAs shown in Fig. 2(c) yield good IF estimation for the sharp chirps but many artifacts for the smooth ones. This simulation verifies that the GDAs have the ability to reveal the IFs that cannot be estimated accurately from the IFAs. In the proposed method, those artifacts (i.e. IF curves with too short length) in the IFAs and GDAs are eliminated, and then good IF estimation result can be obtained by combining the IFAs and GDAs together, as depicted in Fig. 2(d).

4.2. Comparison between the proposed method and ridge detection

For a fairer comparison, the ridge detection algorithm used here combines the ridges detected along the frequency axis (i.e. horizontal ridges) with the ridges detected along the time axis (i.e. vertical ridges). The ridge detection algorithm is performed on the S-method (SM) [3] and the reassigned SM [4] because these two TFRs are based on the STFT, and thus the same window width is used for all the TFRs.

Fig. 3(a) depicts the SM of a signal consisting of three components. The first component has sinusoidal IF and is overlapped with the second one, a linear chirp. The third one (located on the upmost of the TFR) is also a linear chirp but has much lower energy, and thus it is hardly visible from the magnitude of the TFR. The horizontal ridges and vertical ridges, shown in Fig. 3(b) and (c) respectively, are combined after removing ridges with too short length. However, many long-length spurious ridges still remain. These ridges can be further removed by energy thresholding but at the cost of un-

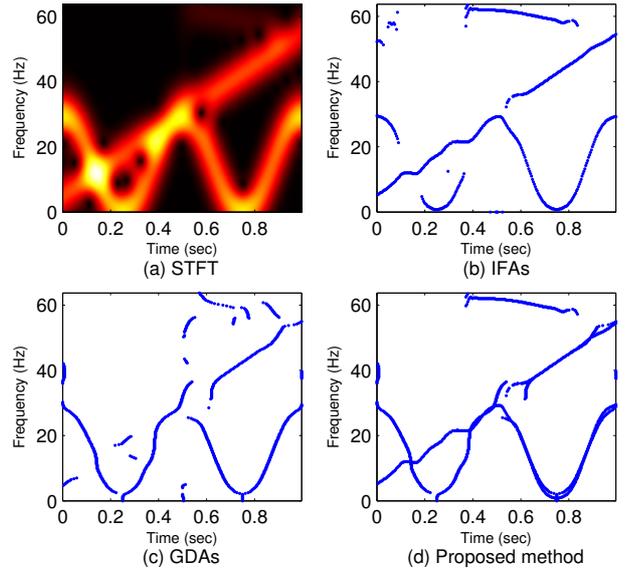


Fig. 5. (a) Spectrogram of a three-component signal, (b) IFAs, (c) GDAs, and (d) proposed IF estimation combining IFAs and GDAs (IF curves with too short length are eliminated).

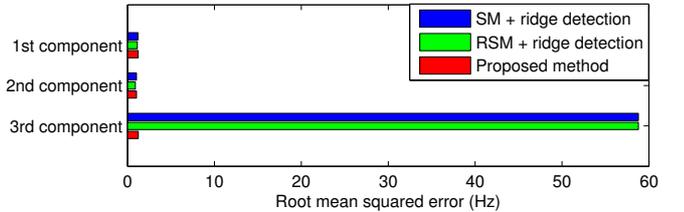


Fig. 6. Root mean squared errors of the estimated IFs shown in Fig. 3(d), Fig. 4(d) and Fig. 5(d): 1st component (sinusoidal IF), 2nd component (high-energy linear chirp), and 3rd component (lower-energy linear chirp).

necessarily eliminating the ridge of the low-energy component, as shown in Fig. 3(d). The ridge detection on the reassigned SM is illustrated in Fig. 4. It can be found that the reassigned SM provides higher energy concentration, but there are more remaining spurious ridges and the ridge of the low-energy component is also eliminated. Finally, the proposed method is depicted in Fig. 5. By comparison, much less spurious IFs are generated by the IFAs and GDAs. Since energy thresholding is unnecessary, the IFs of the three components can all be estimated. The root mean squared errors (RMSEs) of the estimated IFs shown in Fig. 3(d), Fig. 4(d) and Fig. 5(d) are depicted in Fig. 6, assuming that the remaining spurious ridges can be clearly eliminated. The proposed method has similar RMSEs for all the components, while the ridge detection has very poor performance for the 3rd component which has much lower energy.

4.3. Proposed method in noisy environments

For noisy signals, if two STFTs with different window widths are provided, the number of spurious IFs can be reduced by the proposed method. The signal in Sec. 4.1 is used again

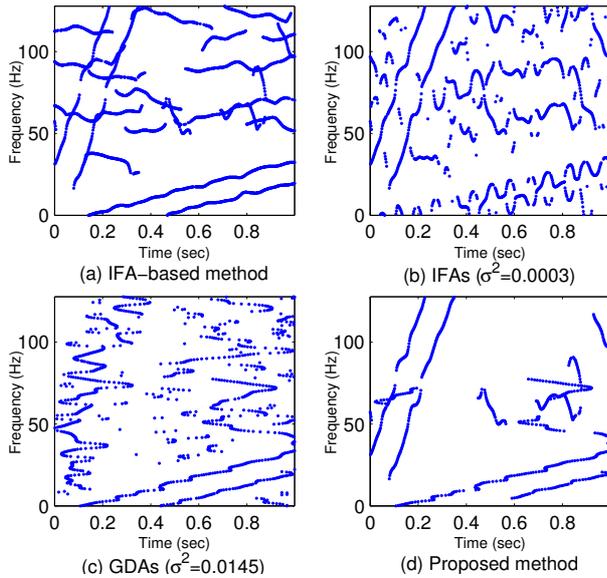


Fig. 7. IF estimation of a noisy signal with SNR being 1.5dB: (a) IFA-based method using the IFAs of STFT with $\sigma^2 = 0.0003$ and the IFAs of STFT with $\sigma^2 = 0.0013$, (b) IFAs of the STFT with $\sigma^2 = 0.0003$, (c) GDAs of the STFT with $\sigma^2 = 0.0145$, (d) proposed method based on the IFAs and the GDAs (IF curves with too short length have been eliminated.)

with SNR being 1.5dB. In the IFA-based method, two sets of IFAs obtained from the two STFTs are combined together after eliminating ridges with too short length. For minimal number of spurious IFs, the optimal choices of the window widths are $\sigma^2 = 0.0003$ and $\sigma^2 = 0.0013$, and the result is shown in Fig. 7(a). The remaining spurious IF curves are so long that they are misidentified as the IFs of the signal. Thus, the key for less spurious IF curves is reducing their length.

It can be expected that the STFT with shorter window width yields longer vertical ridges, while the STFT with longer window width leads to longer horizontal ridges. Accordingly, in the proposed method, the IFAs are obtained from the STFT with $\sigma^2 = 0.0003$ (Fig. 7(b)), while GDAs are calculated from the STFT with $\sigma^2 = 0.0145$ (Fig. 7(c)). Then, most of the spurious IF curves are short enough to be eliminated, and combining the IFAs and the GDAs leads to Fig. 7(d). This experiment also shows that even if non-optimal window width is used, the proposed method has much less spurious IFs with acceptable IF estimation accuracy. A statistical result in Fig. 8 shows the average number of the remaining spurious IFs at any time instant. It can be found that the proposed method greatly outperforms the IFA-based method.

5. CONCLUSION

In the sinusoidal model of [2], the IFs are estimated from the IFAs, which are defined based on the phase of the TFR. In order to improve this IF estimation method, the GDAs, which are also obtained from the phase of the TFR, have been pro-

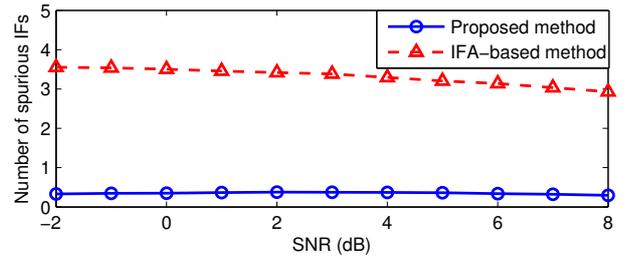


Fig. 8. Comparison of the proposed method and the IFA-based method on the average number of the remaining spurious IFs at any time instant.

posed in this paper. Since the GDAs can reveal IFs which are not estimated from the IFAs, the proposed method (i.e. the IF estimation based on GDAs and IFAs) outperforms the IFA-based method and creates less spurious IFs in noisy environments. Also, it has been shown that the proposed method provides better IF estimation results than ridge detection.

REFERENCES

- [1] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. II. Algorithms and applications," *Proc. IEEE*, vol. 80, no. 4, pp. 540–568, 1992.
- [2] T. Abe and M. Honda, "Sinusoidal model based on instantaneous frequency attractors," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1292–1300, 2006.
- [3] L. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, 2001.
- [4] I. Djurović and L. Stanković, "Time-frequency representation based on the reassigned S-method," *Signal Processing*, vol. 77, no. 1, pp. 115–120, 1999.
- [5] B. Boashash, *Time frequency signal analysis and processing: a comprehensive reference*, Elsevier Science Limited, 2003.
- [6] E. Sejdić, I. Djurović, and J. Jiang, "Time-frequency feature representation using energy concentration: an overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009.
- [7] K. R. Fitz and S. A. Fulop, "A unified theory of time-frequency reassignment," *ArXiv preprint arXiv:0903.3080*, 2009.
- [8] P. Guillemin and R. Kronland-Martinet, "Horizontal and vertical ridges associated to continuous wavelet transforms," in *Proc. IEEE-SP Int. Symp. on Time-Frequency and Time-Scale Analysis*, 1992, pp. 63–66.
- [9] K. Raghupathy and T. W. Parks, "Improved curve tracing in images," in *ICASSP 2004*, pp. 581–584.
- [10] S. C. Pei and S.-G. Huang, "STFT with adaptive window width based on the chirp rate," *IEEE Trans. on Signal Processing*, vol. 60, no. 8, pp. 4065–4080, 2012.