A NEW APPROACH FOR THE REASSIGNMENT OF TIME-FREQUENCY REPRESENTATIONS

Ervin Sejdić[1], *Umut Ozertem*[2], *Igor Djurović*[3] *and Deniz Erdogmus*[4]

[1]Bloorview Research Institute, University of Toronto, Toronto, Ontario, Canada
[2] Yahoo! Inc., Santa Clara, CA, USA
[3]Elektrotehnički Fakultet, University of Montenegro, Podgorica, Montenegro
[4]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA
E-mails: esejdic@ieee.org, umut@yahoo-inc.com,
igordj@ac.me, erdogmus@ece.neu.edu

ABSTRACT

The reassignment method is a widespread approach for obtaining high resolution time-frequency representations. Nevertheless, its performance is not always optimal and can deteriorate for low signal-to-noise ratio (SNR) values. In order to overcome these obstacles, a novel method for obtaining high resolution time-frequency representations is proposed in this paper. The new method implements recently proposed nonparametric snakes in order to obtain accurate locations of the signal ridges in the time-frequency domain. The results of numerical analysis show that the proposed method is capable of achieving significantly higher concentration of signals in the time-frequency domain in comparison to the spectrogram and the traditional reassignment method. Furthermore, the new scheme also maintains good performance for low SNR values, while the performance of the other two considered methods significantly diminishes. It is clear from the results that the proposed method might be of significance in applications where accurate estimation of the signal components is required for low SNR values.

Index Terms— Time-frequency analysis, reassignment, nonparametric snakes.

1. INTRODUCTION

The main objectives of the various types of time-frequency representations (TFRs) and their modifications are to obtain a time-varying spectral density function with high resolution, and to overcome any potential interference [1]. Amongst various algorithms, the time-frequency reassignment is a widespread approach for obtaining high resolution representations. The reassignment method creates a modified version of a representation by moving its values away from where they are computed to produce a better localization of the signal components [2], [3]. However, a fundamental problem exists when the reassigned TFRs are used for the estimation of the instantaneous frequency. The reassignment method approaches the problem of the enhancement of energy concentration in the time-frequency domain by determining so called centers of gravities. Such centers are not necessarily located at the instantaneous frequencies of the signal components, except for chirps and impulses [3]. Hence, even though we might improve the energy concentration by reassigning the TFR of a signal, we often can diminish its accuracy since the peaks in the time-frequency domain can be moved away from the instantaneous frequencies. In order to maintain the accuracy of the instantaneous frequency estimation achieved with a TFR and to enhance its energy concentration, an idea from image processing is proposed as an alternative tool for the reassignment of the TFRs. The method proposed in this manuscript is based on the idea of recently proposed nonparametric snakes [4]. The method uses a nonparametric approach to detect ridges in the time-frequency domain. The detected ridges are then used for the reassignment of a TFR.

The proposed scheme has been tested using a set of synthetic signals and its performance is compared with the TFRs obtained by the spectrogram and the traditional reassignment method. The results of the numerical analysis showed that the proposed method significantly enhances the energy concentration of signals in the time-frequency domain. These results are confirmed for both, noise-free and noisy environments. As expected, the results deteriorate for signals in the noisy environment, however, the proposed scheme maintains the highest concentration amongst the considered representations.

This paper is organized as follows. In Section 2, concepts of the time-frequency analysis along with the reassignment of time-frequency representations are reviewed. The development of the proposed scheme is covered in Section 3. Section 4 evaluates the performance of the proposed scheme using test signals. Conclusions are drawn in Section 5.

2. TIME-FREQUENCY ANALYSIS

Time-frequency analysis (TFA) is of great interest when the time or the frequency domain descriptions of a signal alone cannot provide comprehensive information about a signal for further analysis. Therefore, the basic goal of the TFA is to determine the energy concentration along the frequency axis at a given time instant, i.e., to search for a joint time-frequency representation of the signal [5].

The TFRs can be classified according to the analysis approaches. In the first category, the signal, $x(t)$, is represented by time-frequency (TF) functions derived from translating, modulating and scaling a basis function having a definite time and frequency localization. The second category is based on Cohen's idea of time-frequency distributions (TFD). Short-time Fourier transform (STFT), wavelets, and matching pursuit algorithms are typical examples of signal decomposition based TFRs, while Wigner distribution, Choi-Williams distribution, and spectrogram are some of the methods commonly used for obtaining TFDs [5].

In either approach, poor energy concentration of a signal can

The work of Deniz Erdogmus is supported by NSF grants ECS-0524835 and ECS-0622239.

be achieved in the time-frequency domain for the following reasons: The signal decomposition based TFRs can have poor concentration due to localization of the basis functions, while TFDs usually suffer from cross-terms and inner interferences [5]. These can be removed by a kernel function, however, such a function deteriorates localization of the signal in the time-frequency domain. Several approaches have been proposed to improve the energy concentration in the time-frequency domain (e.g. [6], [7]) with the previously mentioned reassignment method being one of them.

2.1. Traditional Reassignment Approach

The first step in the reassignment approach is to calculate the center of gravity of the signal's energy for each point on the time-frequency plane. Mathematically, this is given by [2], [3]:

$$\widehat{t}(t,\omega) = t - \frac{\int\int uTFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega} \tag{1}$$

$$\widehat{\omega}(t,\omega) = \omega - \frac{\int\int \Omega TFR(t-u,\omega-\Omega)dud\Omega}{\int\int TFR(t-u,\omega-\Omega)dud\Omega} \tag{2}$$

Given these centers of gravities, the reassigned time-frequency representation is obtained by

$$RTFR(t,\omega) = \int\int TFR(\tau,\upsilon)\delta(t-\widehat{t}(\tau,\upsilon))\delta(\omega-\widehat{\omega}(\tau,\upsilon))d\tau d\upsilon \tag{3}$$

where $\delta(t)$ is a Dirac function.

Nevertheless, this approach has certain disadvantages. Since the centers of gravity are calculated for every point on the time-frequency plane, it is easy to imagine that for noisy signals some of the calculated centers are actually not part of the signal. Hence, the main challenge is to accurately determine these centers.

3. NONPARAMETRIC SNAKES FOR REASSIGNMENT OF TIME-FREQUENCY REPRESENTATIONS

A related problem to accurate estimation of the centers of gravities is that of the edge-based approaches for image segmentation [4]. These approaches detect the edges of the image and subsequently connect them to build object contours. However, if the edges cannot be computed, the application of these methods becomes limited. To resolve this issue, so-called snakes (or active contours) have been proposed [8], [9]. Snakes are based on the utilization of the shape priors with the gradient of the edge map of the image. As with most parametric methods, the usual way of seeking the desired result is to run the algorithm several times for a set of different parameter values until a satisfactory performance is obtained.

Image segmentation based on snakes was recently addressed using a nonparametric scheme [4]. The proposed approach translated the problem of seeking efficient values of the snake parameters into the problem of kernel density estimation (KDE) and derived an algorithm that exploits the underlying kernel density estimate of the edge image. Determining a suitable kernel function is the most significant step in KDE, and there is a wide breadth of literature about how to select the kernel function [10].

The application of nonparametric snakes to image segmentation requires KDE for each pixel. However, in our case the energy distribution $E_x(t,\omega) = |TF_x(t,\omega)|^2$ of a time-frequency representation of a signal, provides exactly that. In order to simplify the notation throughout this section, the vector $\mathbf{s} = [t\ \omega]^T$ denotes a location for each point on a time-frequency plane. Hence, the energy distribution of a signal can be rewritten as $E_x(\mathbf{s}) = E_x(t,\omega)$.

Given some samples of an active contour denoted by s_j^{snake} $j = 1, ..., N_{snake}$ and the energy distribution of a signal in the time-frequency domain, our aim becomes to find ridges that capture the structure of a signal. A solution for such a problem can be formulated through maximizing the inner product between the probability density function of the snake $p_{snake}(\mathbf{s})$ and the energy distribution of the signal:

$$\max_{\{\mathbf{s}^{snake}\}} J(\{\mathbf{s}^{snake}\}) = \max \int p_{snake}(\mathbf{s})E_x(\mathbf{s})d\mathbf{s} \tag{4}$$

where the probability density of the snake, $p_{snake}(\mathbf{s})$, is also evaluated as a KDE, using the samples of the snake

$$p_{snake}(\mathbf{s}) = \frac{1}{N_{snake}} \sum_{j=1}^{N_{snake}} K_\sigma(\mathbf{s} - \mathbf{s}_j^{snake}) \tag{5}$$

where N_{snake} is the number of points on the snake. Therefore, the cost function $J(\mathbf{s}^{snake})$ is given by

$$J(\mathbf{s}^{snake}) = \frac{1}{N_{snake}} \sum_{j=1}^{N_{snake}} \int K_\sigma(\mathbf{s} - \mathbf{s}_j^{snake})E_x(\mathbf{s})d\mathbf{s} \tag{6}$$

Note that this cost function is additive in terms of the samples of the snake. Therefore, near the optimal point along the ridge of the energy distribution, higher and lower sampling rates of the snake would lead to an accordingly denser or sparser evaluation of the optimality criterion along the ridge. The energy distribution of a signal contains all smoothness information needed by the snake, which can be simply extracted by sampling the snake at a higher rate. In practice, utilizing an isotropic fixed-bandwidth for p_{snake} is sufficient.

In order to achieve faster convergence than gradient-like step based algorithms, a fixed-point approach will be preferred in optimization algorithm design. To derive a fixed-point iteration for the samples of the snake, we equate the gradient of the optimization criterion by using the fact that for any fixed point of the density inner product cost function, the gradient of the inner product with respect to $\mathbf{s}^{snake}$ should be equal to zero. This yields the following:

$$\frac{\partial J(\mathbf{s}^{snake})}{\partial \mathbf{s}} = 0. \tag{7}$$

Reorganizing the terms and solving for $\mathbf{s}^{snake}$, the fixed-point update rule can be written as

$$\mathbf{s}_j^{snake} \leftarrow \frac{\int \mathbf{s}K_\sigma(\mathbf{s} - \mathbf{s}_j^{snake})E_x(\mathbf{s})d\mathbf{s}}{\int K_\sigma(\mathbf{s} - \mathbf{s}_j^{snake})E_x(\mathbf{s})d\mathbf{s}}. \tag{8}$$

This iteration is convergent since it is similar to that of the mean-shift [11]. In practice, the iteration above converges at a rate proportional to the cube of the eigenvalue of the local Hessian in the vicinity of a local maximum. Consequently, along a relatively level ridge, one eigenvalue is significantly closer to zero (corresponding to the eigenvector pointing along the ridge) than the other. Due to the elimination of spurious edge maxima, the nonparametric snake does not suffer from poor capture range. The fixed-point iterations and the optimization criterion presented above can be utilized to increase the number of samples in the snake in order to densely populate the boundary once initial convergence is achieved by the original snake samples. The idea is to initialize multiple snake samples for each original sample around the corresponding convergence points and have the new samples converge to the boundary utilizing the same fixed-point iterations.

The determined snake points, $\mathbf{s}^{snake}$, denote the points located along the instantaneous frequency of each signal component. Hence, these points represent the centers of gravities:

$$\mathbf{s}^{snake} = \left[\widehat{t}(t,\omega), \widehat{\omega}(t,\omega)\right] \tag{9}$$

and can be used in the reassignment of the time-frequency representations as proposed by (3).

4. PERFORMANCE ANALYSIS

In this section, the performance of the proposed scheme is examined using a set of synthetic test signals. The goal is to examine the performance of the scheme in comparison to the standard reassignment approach. To illustrate the performance, the STFT is used as a time-frequency representation. Nevertheless, the principles presented here can be applied to other TFRs as well. As for the synthetic signals, the sampling period used in the simulations is $T_s = 1/256$ seconds. The Gaussian window is used as the window function for the calculation of STFT:

$$w_{STFT}(t) = \frac{1}{\sigma_{gw}\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma_{gw}^2}\right) \tag{10}$$

where σ_{gw} is the standard deviation of the window, which dictates the width of window. In addition, a kernel based on the Gaussian function is used for KDE. The bandwidth of the kernel is set to 1.5, and the same value is used for all synthetic signals considered here.

4.1. Example 1

The first test signal is defined as:

$$x_1(t) = \cos\left(60\pi t + \frac{3\pi}{2}\cos(9\pi t)\right) + \cos\left(238\pi t - 56\pi t^2\right) \tag{11}$$

where $0 \leq t < 1$. The signal consists of a sinusoidally FM modulated component and a linear FM component and it is depicted in Fig. 1(a). There are several trade-offs between good localization of the sinusoidally FM component and the linear FM component. In order to achieve good localization of the sinusoidally FM signal, the standard deviation of the Gaussian window has to be manipulated either by trial-and-error or a more sophisticated method such as one presented in [12]. For our purposes, we set $\sigma_{gw} = \frac{1}{20}$and the resulting time-frequency representation of the signal obtained by STFT is depicted in Fig. 1(b). Improvements can be achieved with the application of the traditional reassignment method as depicted in Fig. 1(c).

Significant improvement in localization of the signal in time-frequency domain can be noticed in Fig. 1(d) in comparison to the TFR obtained by STFT. The proposed algorithm provides almost ideal TFR representation of the signal, where the ideal TFR is a theoretical model of a signal in time-frequency domain as described in [13]. In order to compare performance of the traditional reassignment method and the proposed method, a measure proposed in [14] is used. The measure is given as the ratio of the energy along the instantaneous frequencies and the energy outside these regions, and is defined as:

$$B = 10\log_{10}\left(\frac{\int\int_{(t,\omega)\in R} |TFR_x(t,\omega)|\, dtd\omega}{\int\int_{(t,\omega)\notin R} |TFR_x(t,\omega)|\, dtd\omega}\right) \tag{12}$$

where $TFR_x(t,\omega)$ is a TFR of the signal with region R corresponding to the instantaneous frequency lines of the signal components. From the definition of the measure, it can be concluded that a higher value of B signifies a more concentrated TFR. In addition, since the exact values of instantaneous frequencies are known, in the case of reassigned TFRs the measure produces higher value for the representation which reassigns the original TFR more accurately. For the considered representations, the values of B are given in Table 1. These results represent the analysis for noise-free and noisy environments. In order to simulate the noisy environment, the signal is contaminated with the additive white Gaussian noise whose variance is calculated based on the given SNR. The results clearly indicate that the proposed method achieves the highest concentration amongst considered methods. The reason behind significant improvement in the localization is the fact that the nonparametric snakes are capable of accurately searching for ridges, which are most of the time located at or around of the instantaneous frequencies of signal components [5]. In comparison, the traditional reassignment method determines the "centers of gravities", which are not necessarily located at the instantaneous frequencies.

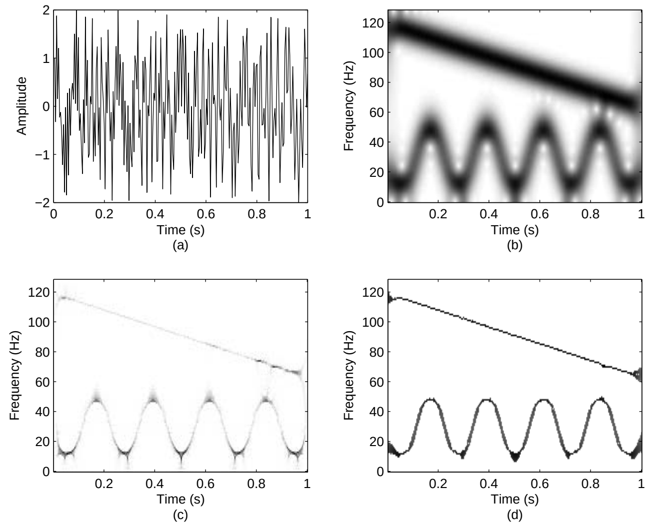

Fig. 1. Time-frequency analysis of $x_1(t)$: (a) time-domain representation of the signal; (b) the TFR of the signal obtained by STFT; (c) the TFR of the signal obtained by the traditional reassignment method; (d) the TFR of the signal obtained by the proposed method.

TFR	$\mathbf{B}$(noise-free)	$\mathbf{B}(SNR = 5$ dB)
STFT	4.87	4.40
TF reassignment	9.20	8.35
Proposed method	10.64	9.61

Table 1. The values of the measure for the three considered representations.

4.2. Example 2

A more complicated example, depicted in Fig. 2(a), can be defined as:

$$x_2(t) = \begin{cases} x_o(t) + \cos(60\pi t) & \text{for } 0.332 \leq t \leq 0.500 \\ x_o(t) & \text{elsewhere} \end{cases} \tag{13}$$

with

$$x_o(t) = \cos(176\pi t + 1.5\pi \cos(4\pi t)) + \cos\left(220\pi t - 106\pi t^2\right) \quad (14)$$

and where $x_2(t)$ exists only on the interval $0 \leq t < 1$. For this class of signals, similar conflicting issues occur as in the previous example; however, here exist additional constraints, i.e., the crossing components and a short-duration component. The TFR of the signal obtained with STFT is depicted in Fig. 2(b), and $\sigma_{gw} = \frac{1}{9}$ is used to obtain the TFR.

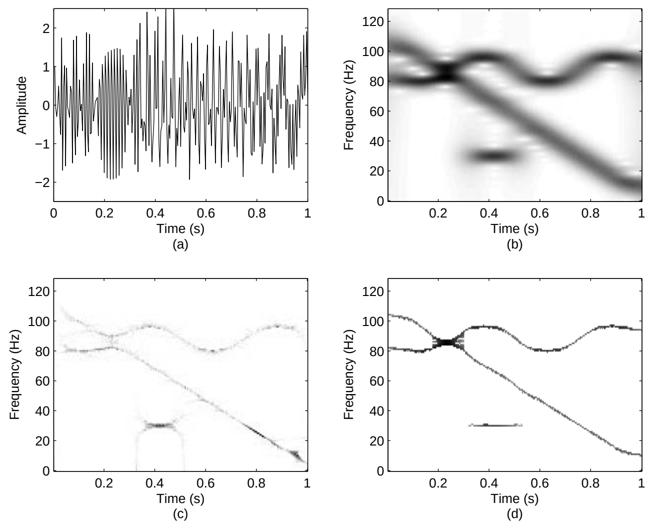

Fig. 2. Time-frequency analysis of $x_2(t)$: (a) time-domain representation of the signal; (b) the TFR of the signal obtained by STFT; (c) the TFR of the signal obtained by the traditional reassignment method; (d) the TFR of the signal obtained by the proposed method.

Improvements can be noticed with the traditional reassignment method depicted in Fig. 2(c). However, the best localization amongst the considered representations is obtained with the proposed reassignment method as shown in Fig. 2(d). To verify our claims, the measure given by (12) is calculated for the three considered representations. Similarly, as in the previous example, noise-free and noisy environments are considered. The proposed method achieves the best concentration again as shown in Table 2. As expected, the performance for all three considered methods deteriorates with an increased level of noise, however, the proposed method maintains overall good performance.

TFR	$\mathbf{B}$(noise-free)	$\mathbf{B}(SNR = 5$ dB)
STFT	6.40	5.91
TF reassignment	10.19	8.95
Proposed method	11.60	11.18

Table 2. The values of the measure for the three considered representations.

5. CONCLUSION

In this paper, a novel scheme for improving the energy concentration of signals in the time-frequency domain has been developed. The scheme is based on the idea of reassigning the time-frequency representations of signals in order to achieve sharper representations. Nevertheless, the important difference between the existing methods and the proposed approach is that here we implemented a novel concept from image processing as a tool aiding us in the reassignment of the time-frequency representations. The approach developed throughout this paper is based on the idea of nonparametric snakes which are used to accurately determine the instantaneous frequencies of signal components. In order to verify our claims, the performance of the proposed scheme has been evaluated and compared with the spectrogram and the traditional reassignment method by using a set of synthetic test signals. The performance was examined in noise-free and noisy environments. The results have shown that the scheme can achieve higher energy concentration of signals in the time-frequency domain in comparison with these traditional methods. As expected, the performance deteriorates in the noisy environment for all considered signal representations. Nevertheless, the proposed scheme maintains the best performance.

6. REFERENCES

[1] I. Djurović and LJ. Stanković, "A virtual instrument for time-frequency analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 48, no. 6, pp. 1086–1092, Dec. 1999.

[2] K. Kodera and R. G. C. De-Villedary, "Analysis of time-varying signals with small BT values," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-26, no. 11, pp. 64–76, Feb. 1978.

[3] F. Auger and 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, May 1995.

[4] U. Ozertem and D. Erdogmus, "Nonparametric snakes," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2361–2368, Sep. 2007.

[5] L. Cohen, *Time-Frequency Analysis*. Englewood Cliffs, N.J: Prentice Hall PTR, 1995.

[6] R. G. Baraniuk and D. L. Jones, "Signal-dependent time-frequency analysis using a radially Gaussian kernel," *Signal Processing*, vol. 32, no. 3, pp. 263–284, Jun. 1993.

[7] LJ. Stanković, "Highly concentrated time-frequency distributions: Pseudo quantum signal representation," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 543–551, Mar. 1997.

[8] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, Jan. 1988.

[9] C. Xu and J. L. Prince, "Snakes, shapes and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[10] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.

[11] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 8, pp. 790–799, Aug. 1995.

[12] LJ. Stanković, "A measure of some time-frequency distributions concentration," *Signal Processing*, vol. 81, no. 3, pp. 621–631, Mar. 2001.

[13] ——, "An analysis of some time-frequency and time-scale distributions," *Annales des Telecommunications*, vol. 49, no. 9/10, pp. 505–517, Sep./Oct. 1994.

[14] I. Djurović and LJ. Stanković, "Time-frequency representation based on the reassigned S-method," *Signal Processing*, vol. 77, no. 1, pp. 115–120, Aug. 1999.