About Time-Variant Filtering of Speech Signals with Time-Frequency Distributions for Hands-Free Telephone Systems

Srdjan Stanković

Abstract— **Joint time-frequency distributions represent the energy or intensity of signal simultaneously in time and frequency. In this paper we introduce a smoothed version of the method denoted as S-method (SM) [18]. It is shown that it is an effective tool for representation of speech signals disturbed by noise. As a consequence, it can be efficiently applied to time-variant filter problems, as they occur in hands-free telephones. The optimal time-variant filter is derived in terms of the smoothed SM, and is illustrated by examples.**

I. Introduction

Speech signals are highly nonstationary, with a wide dynamic range of multiple frequency components in the short-time spectra [2], [6], [26]. Time-frequency distributions have been introduced for analyzing the frequency components as a function of time of nonstationary signals. The most common time-frequency representation, the spectrogram, is characterized by a trade-off between time and frequency resolution. Consequently, the development of other quadratic time-frequency distributions for representation and processing of nonstationary signals is an interesting challenge. One possible application is e.g. the Wiener optimum filtering of speech signals corrupted by noise. The more accurate the desired signal spectrum can be estimated, the better the noise components can be filtered.

There are few time-frequency distributions which are used in practical applications. They belong to the general class of quadratic time-frequency distributions, the Cohen class [5].

Signal Processing, Vol.80, No.9, 2000.

Among all distributions of this class, the best auto-terms concentration is obtained with the Wigner distribution [5], [21]. However, for multicomponent signals, such as speech signals, the Wigner distribution is useless because of the great number of strong cross-terms [8], [19]. The SM has been introduced with the aim to avoid the cross-terms while keeping the auto-terms concentration of the Wigner distribution [18], [20]. Thereby, it is possible to combine the advantage of the spectrogram (absence of cross-terms) with those of the Wigner distribution (high time and frequency resolution). When comparing the SM with other distributions, it can be seen that the cross-terms are not reduced at the expense of the auto-terms concentration [21].

Because of the nonstationary nature of speech signals, statistically optimum filtering requires time-variant filtering methods. Filtering in the time-frequency domain could be advantageous compared to separate filtering in the time or frequency domain. Since there exists no unique definition of time-frequency spectra, many approaches for time-variant filtering have been proposed. Zadeh [27] suggested to use the Rihaczek distribution [17]. However, this time-frequency spectrum is complex valued and badly concentrated in time. Therefore, filtering in the time-frequency domain has been redefined using the Wigner distribution with the drawbacks mentioned above [3], [9], [13], [23]. In this paper we will consider this approach by using the Weyl correspondence [9], [11], [12], [13]. The time-variant transfer function has been defined as the Weyl symbol mapping of the impulse response into the time-frequency plane. The Wigner spectrum is used in order to average out the cross-terms. Many different realizations of the same random process, at a given instant, are necessary to obtain the mean value of the Wigner distribution, i.e. the Wigner spectrum. However, since our processing will be based on a single noisy speech signal realization, the original definitions, which use the mean value of the Wigner distribution, are not applicable since the cross-terms are not averaged out. By using the SM for statistically optimum filtering in the time-frequency domain cross-terms are suppressed and a more accurate time-frequency spectrum is obtained [23]. Here, a smoothed version of the SM is introduced. It improves the Wiener optimum filtering. Experiments with speech signals disturbed by car noise illustrate the advantages of the presented approach.

The paper is organized as follows. The SM and its application on time-frequency representation of speech signals are presented in Section II. Theory and definition of time-variant filtering are given in Section III. In Section IV time-frequency representation and time-variant filtering of speech signals corrupted by car noise are illustrated by some experiments.

II. Theoretical Background

The general class of quadratic time-frequency distributions, the Cohen class, is defined as [5]:

$$CD_{ff}(t,\omega)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}c(\tau,\theta)f(u+\frac{\tau}{2})f^*(u-\frac{\tau}{2})e^{-j\omega\tau-j\theta t+j\theta u}dud\tau d\theta, \quad (1)$$

where $f(t)$ is a signal, and $c(\tau,\theta)$ is a kernel function. Consider the kernel function in the form [21]:

$$c(\tau,\theta)=P(\frac{\theta}{2})*_\theta AF_{ww}(\tau,\theta) \quad (2)$$

where $P(\theta)$ is a rectangular window in the frequency domain, $AF_{ww}(\tau,\theta)=\int_{-\infty}^{\infty}w(u+\frac{\tau}{2})w(u-\frac{\tau}{2})e^{-j\theta u}du$ is the ambiguity function of the window $w(u+\frac{\tau}{2})w(u-\frac{\tau}{2})$, and $*_\theta$ denotes a convolution over θ.

The SM follows from (1) with the kernel function defined by (2):

$$SM_{ff}(t,\omega)=\frac{1}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}P(u)f(t-u+\frac{\tau}{2})f^*(t-u-\frac{\tau}{2})w(u+\frac{\tau}{2})w(u-\frac{\tau}{2})e^{-j\omega\tau}dud\tau. \quad (3)$$

Keeping in mind the definition of the short-time Fourier transform:

$$STFT(t,\omega)=\int_{-\infty}^{\infty}f(t+\tau)w(\tau)e^{-j\omega\tau}d\tau \quad (4)$$

the SM can be written as [18]:

$$SM_{ff}(t,\omega)=\int_{-\infty}^{\infty}P(\theta)STFT(t,\omega+\theta)STFT^*(t,\omega-\theta)d\theta. \quad (5)$$

Discretization of the SM (5) produces:

$$SM_{ff}(n,k)=\sum_{l=-L}^{L}P(l)STFT(n,k+l)STFT^*(n,k-l), \quad (6)$$

where $2L+1$ is the window length of $P(l)$, and n and k are the discrete time and frequency parameters, respectively. Noting that:

$$\begin{aligned}STFT(n,k+l)STFT^*(n,k-l)+\\STFT(n,k-l)STFT^*(n,k+l)=\\=2Re\{STFT(n,k+l)STFT^*(k-l)\}\end{aligned}$$

and taking a rectangular window for $P(l)$, we get:

$$SM_{ff}(n,k)=|STFT(n,k)|^2+2Re\{\sum_{l=1}^{L}STFT(n,k+l)STFT^*(n,k-l)\}, \quad (7)$$

where $|STFT(n,k)|^2$ is the spectrogram.

The pseudo Wigner distribution (as it is used in practical applications) is defined by:

$$WD_{ff}(t,\omega) = \int_{-\infty}^{\infty} f(t+\frac{\tau}{2})f^*(t-\frac{\tau}{2})\ x$$

$$w(\frac{\tau}{2})w^*(-\frac{\tau}{2})e^{-j\omega\tau}d\tau =$$

$$\frac{1}{\pi}\int_{-\infty}^{\infty} STFT(t,\omega+\theta)STFT^*(t,\omega-\theta). \quad (8)$$

In the case of multicomponent signals $f(t) = \sum_{i=1}^{N} f_i(t)$, (where $f_i(t)$ represents the $i-th$ component of the signal $f(t)$) the SM has the form [24]:

$$SM_{ff}(t,\omega) = \pi\sum_{i=1}^{N} WD_{f_i f_i}(t,\omega),$$

i.e., it is equal to the sum of the pseudo Wigner distributions of the auto components. Thus, the SM is the distribution that produces the cross-terms free pseudo Wigner distribution [24]. Factor $1/\pi$ could be included into the wondow $P(\theta)$.

Now consider the SM in the case of a noisy signal. It has been shown that there is a trade-off in changing the window width L, [22]: (1) by increasing L, the auto-term concentration tends toward the one obtained by the Wigner distribution, (2) by decreasing L, smaller noise is added, while the auto-term concentration could be lower, as in the spectrogram. For reduction of the noise influence and keeping sufficient energy concentration, the smoothed SM is now introduced:

$$P(\theta) = \begin{cases} 1-\left|\frac{\theta}{a}\right| & \text{for} \quad -a \le \theta \le a \\ 0 & \text{elsewhere} \end{cases}. \quad (9)$$

The motivation for introducing this form of the SM, that is not based on the rectangular window $P(\theta)$, has been found in the analysis of the maximal auto-term value and the variance in the SM [1], [21], [22] (see Appendix).

The discrete version of the window (9) is:

$$P(l) = \begin{cases} 1-\left|\frac{l}{L+1}\right| & \text{for} \quad -L \le l \le L \\ 0 & \text{elsewhere} \end{cases}. \quad (10)$$

For example, the smoothed SM for the particular value of $L = 3$ can be written in the form:

$$SM_{ff}(n,k) = |STFT(n,k)|^2 +$$

$$2Re\{\frac{3}{4}STFT(n,k+1)STFT^*(n,k-1)+$$

$$\frac{1}{2}STFT(n,k+2)STFT^*(k-2)+$$

$$\frac{1}{4}STFT(n,k+3)STFT^*(n,k-3)\}. \quad (11)$$

This smoothed version of the SM turns out to be very simple and suitable for representation and processing of noisy speech signals.

III. Time-variant filtering of speech signals in hands-free telephone systems

Within the Wigner distribution framework, time-variant filtering of noisy signals $x(t) = f(t) + n(t)$ with the desired signal $f(t)$ and the noise signal $n(t)$ is defined by [11], [13], [23]:

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, t-\frac{\tau}{2})x(t+\tau)d\tau, \quad (12)$$

where $h(t+\frac{\tau}{2}, t-\frac{\tau}{2})$ is the impulse response of the time-variant filter. The optimal transfer function may be derived in analogy with the Wiener filter derivation for the case of stationary signals [13], [14]. When the mean square error reaches its minimum [14],

$$H_{opt} = \arg\{\min_{H} E\{|f(t)-(Hx)(t)|^2\}\}, \quad (13)$$

the error $e(t) = f(t) - (Hx)(t)$ is orthogonal to the signal $x^*(t+\tau+\alpha)$ [14]:

$$E\{[f(t) - \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, t-\frac{\tau}{2})\ x$$

$$x(t+\tau)d\tau]x^*(t+\tau+\alpha)\} = 0, \quad (14)$$

where α is an arbitrary constant. Let us define the expected value of the ambiguity function $AF_{xx}(\tau,\theta)$ as [12]:

$$\overline{AF}_{xx}(\tau,\theta) =$$

$$\int_{-\infty}^{\infty} E\{x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})\}e^{-j\theta t}dt, \quad (15)$$

and the Fourier transform of $h(t+\frac{\tau}{2}, t-\frac{\tau}{2})$, over t by:

$$AF_H(\tau,\theta) = \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, (t-\frac{\tau}{2})\}e^{-j\theta t}dt. \quad (16)$$

Assuming that the processes are mainly concentrated around the origin and the axes of the ambiguity plane (underspread processes [10], [12], [23]), it follows then from (14):

$$\overline{AF}_{fx}(\alpha,\theta) = \iint_{-\infty}^{\infty} AF_H(u,-\tau)\overline{AF}_{xx}(\theta-u,\alpha-\tau)dud\tau. \quad (17)$$

Note that relation (17) directly follows from (14), without additional assumptions, if the considered processes are quasistationary [14], i.e., if $E\{f(t)x^*(t+\alpha)\} = E\{f(t-\frac{\alpha}{2})x^*(t+\frac{\alpha}{2})\}$ and $E\{x(t+\tau)x^*(t+\alpha)\} = E\{x(t+(\tau-\alpha)/2)x^*(t-(\tau-\alpha)/2)\}$.

The two-dimensional Fourier transform of (17) results in:

$$\overline{WD}_{fx}(t,\omega) = L_H(t,\omega)\overline{WD}_{xx}(t,\omega), \quad (18)$$

where

$$\overline{WD}_{xx}(t,\omega) = E\{WD_{xx}(t,\omega)\} = \int_{-\infty}^{\infty} E\{x(t+\tau/2)x^*(t-\tau/2)\}e^{-j\omega\tau}d\tau \quad (19)$$

is the mean value of the Wigner distribution $WD_{xx}(t,\omega)$ of the signal $x(t)$. It is referred to as the Wigner spectrum of the signal $x(t)$ [7].

The support function $L_H(t,\omega)$ is defined by:

$$L_H(t,\omega) = \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, t-\frac{\tau}{2})e^{-j\omega\tau}d\tau. \quad (20)$$

If the signal and noise are not correlated, we have:

$$L_H(t,\omega) = 1 - \frac{\overline{WD}_{nn}(t,\omega)}{\overline{WD}_{ff}(t,\omega)+\overline{WD}_{nn}(t,\omega)}. \quad (21)$$

In general, the mean value $E\{WD_{ff}(t,\omega)\} = \overline{WD}_{ff}(t,\omega)$ will eliminate uncorrelated cross-terms in the Wigner distribution, since $E\{f_i(t+\frac{\tau}{2})f_j^*(t-\frac{\tau}{2}\} = 0$ for $i \neq j$, as long as components $f_i(t)$ and $f_j(t)$ are uncorrelated [4], [7], [23]. Thus, the problem of cross-terms does not exist in the Wigner spectrum, if we are able to use a large number of realizations belonging to the same random process. Note that the filter is also signal dependent, since its region of support depends on the signal form, and it is signal adaptive.

In numerical implementations the pseudo (i.e., windowed) form of the filtering relation (12):

$$(Hx)(t) = \int_{-\infty}^{\infty} h(t+\frac{\tau}{2}, t-\frac{\tau}{2})w(\tau)x(t+\tau)d\tau \quad (22)$$

should be used. In this version a lag window $w(\tau)$ is introduced. In [23] it is shown that $w(\tau)$ does not influence the output signal $(Hx)(t)$ if $w(0) = 1$. By using Parseval's theorem, (22) can be written in the form:

$$(Hx)(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} L_H(t,\omega)STFT(t,\omega)d\omega. \quad (23)$$

Consider the determination of the support function $L_H(t,\omega)$, that plays a crucial role in time-variant filtering. In most practical applications filtering must be based on a single noisy signal realization. Therefore, it is not possible to calculate the mean value Wigner distribution to obtain the cross-terms free Wigner spectrum. By using the SM in (21) we obtain a cross-term free version of the Wigner distribution:

$$L_H(t,\omega) = 1 - \frac{SM_{nn}(t,\omega)}{SM_{xx}(t,\omega)}, \quad (24)$$

where $SM_{xx}(t,\omega)$ and $SM_{nn}(t,\omega)$ represent the SM of the noisy signal $x(t)$ and the SM of the noise $n(t)$, respectively.

A realization of the proposed time-variant Wiener optimum filter is illustrated in Fig.1.

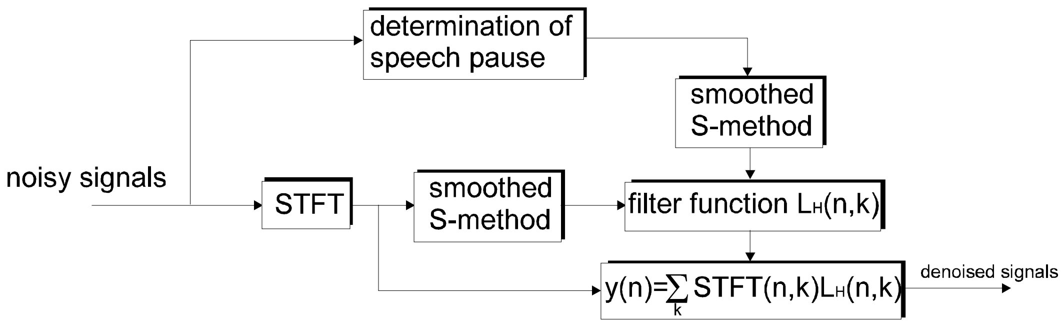

Fig. 1. The filter scheme for time-variant filtering of speech signals.

IV. Experimental Illustrations

A. Example 1

First, we compare the spectrogram and the SM of a speech signal (see Fig.2). The SM is realized with a rectangular window $P(l)$ of length $2L+1=7$ and a Hanning window of length 1024 for the short-time Fourier transform. The spectrogram is implemented with a window length of 1024 for ensuring good frequency resolution. However, the SM frequency resolution is still higher, since the SM is able to concentrate the components with variations in the instantaneous frequency. Time resolution is also better as the SM is based on the Wigner distribution (see frequency components above $1\,kHz$). In summary, information about time-frequency characteristics from the SM are much more reliable than from the spectrogram.

B. Example 2

The spectrogram, the SM with $L=3$ and the smoothed SM realized by using the window function (10) with $L=4$ of a noisy speech signal are shown in Fig.3. The speech signal is recorded in a middle class car cruising along the highway. The noise results from motor hum and from noise produced by wind and wheels. The same window type and window width is used for the short-time Fourier transform calculation as in Example 1. Observe that the noise in the whole time-frequency plane, especially within the frequency range from $500\,Hz$ to $2\,kHz$ (Figs.3 a) and b)), is significantly reduced using the smoothed SM, Fig.3 c). This fact justifies the use of smoothed version of the SM in a time-variant filter realization (24).

C. Example 3

The above theoretical considerations are applied to the filtering of noisy speech signals. The signal has been recorded in the same way as in Example 2. Estimations of the spectrogram and of the SM of noise are performed during a speech pause. A window width of 256 samples (with zero padding up to 1024 samples) has been used for the calculation of the short-time Fourier transform. The same short-time Fourier transform has been used in (23). The calculation of $SM_{xx}(t,\omega)$ is performed by using the smoothed version (10). The time-variant filter has also been realized by using the spectrogram in (24). In both realizations, equation (24) has been slightly modified, by using the spectral floor [6], [26]:

$$L_H(t,\omega)=\max\left\{1-\frac{SM_{nn}(t,\omega)}{SM_{xx}(t,\omega)},\beta\right\} \quad (25)$$

and:

$$L_H(t,\omega)=\max\left\{1-\frac{SPEC_{nn}(t,\omega)}{SPEC_{xx}(t,\omega)},\beta\right\}. \quad (26)$$

The spectral floor is set to $\beta=0.12$.

Time-frequency representations of: a) a noisy signal, b) a denoised signal filtered by using the spectrogram (26), and c) a denoised signal filtered by using (25) are shown in Fig.4.

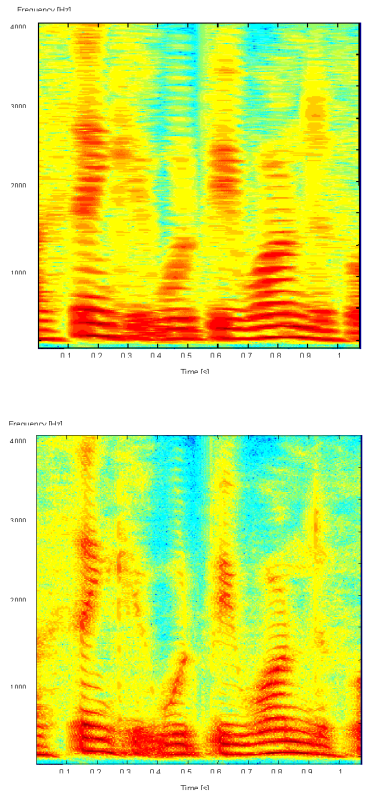

Fig. 2. Time-frequency representation of the speech signal without noise: a) Spectrogram; b) SM with L=3.

We see that the noise suppression for the whole time-frequency plane is better when the smoothed SM is used. In the noise components of about $1\,kHz$ are efficiently filtered. The energy ratio for the noise present after filtering by using the spectrogram and the SM, during the speech pause, is 2.3 dB. This ratio for the spectrogram based filtered signal and the smoothed SM based one is 8 dB. In addition, from Fig.4, we can conclude that better time resolution is achieved by filtering based on the smoothed SM (Fig.4 b)) than by using the spectrogram (Fig.4 a)). These results are expected since the Wigner distribution is introduced in order to improve time and frequency resolution in time-frequency analysis.

It is important to note that the SM has a form very suitable for simple hardware realization. This property is attractive for on-line applications in processing of signals. Note that the hardware realization of the smoothed version of SM would be a straightforward extension of the realization presented in [15], [25].

V. Conclusion

Time-variant filtering of speech signals disturbed by car noise is presented. It has been shown that by using the smoothed version of the SM, as a basic distribution in the filtering definition, noise reduction is better than in the case when the spectrogram is used. Important properties of the proposed filter scheme are its efficiency and suitability for hardware realization.

VI. Acknowledgment

The author is grateful to the anonymous reviewers for comments which lead to a significant improvement of the paper. The author is thankful to Prof. E. Hänsler and J. Tilp for

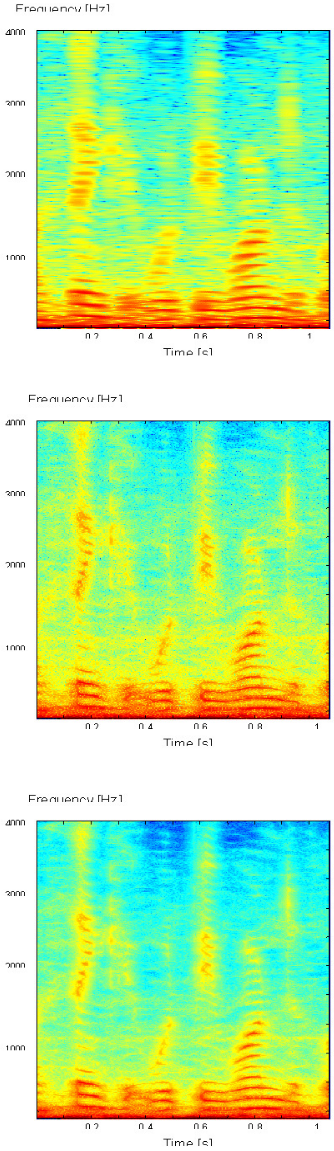

Fig. 3. Time-frequency representation of the noisy speech signal: a) Spectrogram; b) SM with L=3; c) Smoothed SM with L=4.

very helpful suggestions and discussions during the work on this paper. This research has been supported by the Alexander von Humboldt foundation.

VII. Appendix

According to the auto-term analysis [21], the maximal squared auto-term value in the SM (5), for linear FM chirp signals, is proportional to $\int_{-a}^{a} P(\theta)d\theta$. The value of a should lie within the auto-term width. The SM variance is approximately proportional, for white noise, to the window energy $\int_{-a}^{a} P^2(\theta)d\theta$, [1]. One can easily calculate the squared-amplitude-to-variance ratio, $R = \left|\int_{-a}^{a} P(\theta)d\theta\right|^2 / \int_{-a}^{a} P^2(\theta)d\theta$, for various commonly used windows: rectangular, Hanning, Hamming, and triangular ones. For each

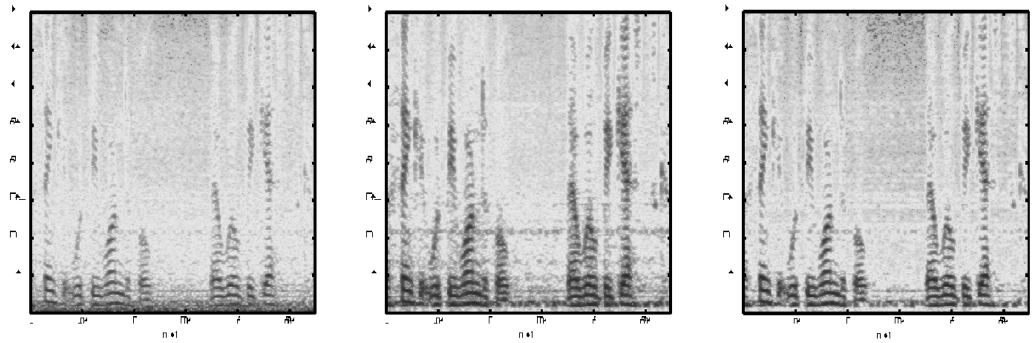

Fig. 4. Time-frequency representation of the noisy speech signal and denoised speech signals filtered by using time-variant filters: a) Spectrogram of noisy signal b) Spectrogram of the denoised signal, filtered by using the spectrogram c) Spectrogram of the denoised signal, filtered by using the smoothed SM with L=4.

considered window we assume the value of a such that the maximal value of the auto-term in the SM remains invariant. The values of R, normalized with the value of R for the rectangular window (the standard SM), are 1, 4/3, 1.36, 3/2, for rectangular, Hanning, Hamming and triangular windows, respectively. Thus, we conclude that the triangular window, as in (9), could improve results in the SM based time-frequency analysis of noisy signals. This is confirmed by examples and by the application of the proposed smoothed SM in time-variant filtering.

Application of the smoothed SM in time-variant filtering is one of the main objectives of this paper. Here, we have to estimate the mean value of the Wigner distribution of a noisy signal, $E\{WD_{xx}(t,\omega)\} = E\{WD_{ff}(t,\omega) + WD_{nn}(t,\omega)\}$. We have concluded that the smoothed SM, based on one signal realization, would produce a reliable estimate of the cross-term free Wigner spectrum. The term $E\{WD_{nn}(t,\omega)\}$ represents the noise spectral density function. Its estimation, obtained by using one signal realization and the smoothed SM, is biased. The bias is proportional to $\int_{-a}^{a} P(\theta)d\theta = const.$, while the estimation variance is proportional to the window $P(\theta)$ energy [1], [14]. Therefore, the estimation of noise spectral density will be the best for the window $P(\theta)$ having minimal energy, i.e. maximal ratio R.

Note that the harmonic shaped disturbances, which are stationary during the considered time interval, are also efficiently removed. For the points (t,ω) where $SM_{nn}(t,\omega)$ has significantly greater values than $SM_{ff}(t,\omega)$, we have $L_H(t,\omega) = 1 - SM_{nn}(t,\omega)/\left(SM_{ff}(t,\omega) + SM_{nn}(t,\omega)\right) \rightarrow 0$.

References

[1] M. G. Amin: "Minimum variance time-frequency distribution kernels for signals in additive noise", *IEEE Trans. Signal Processing*, vol. 44, no. 9, pp. 2352-2356, Sep. 1996.

[2] S. Boll: "Suppression of acoustic noise in speech using spectral substraction ", IEEE Transaction on Acoustics Speech and Signal Processing, vol.27, No.2, 1979, pp. 113-120.

[3] G. F. Boudreaux-Bartels, T. W. Parks: "Time-varying filtering and signal estimation using Wigner distribution", IEEE Transaction on Acoustic Speech and Signal Processing, vol.34. No.6, Jun 1986, pp. 442-451.

[4] J. F. Böhme, D. König: "Statistical processing of car engine signals for combustion diagnostic", Proc. IEEE 7th Workshop on statistical signal and array processing, Quebec 1994, pp. 369-374.

[5] L. Cohen: "Time-frequency distributions- a review ", Proc. IEEE, vol.77, No.7, July 1989, pp.941-981.

[6] P. Dreiseitel and H. Puder: "Speech enhancement for mobile telephony based on non-uniformly spaced frequency resolution", Proc. EUSIPCO Rhodes 1998, pp.965-968.

[7] P. Flandrin, W. Martin: "The Wigner-Ville spectrum of nonstationary processes", *Wigner distribution: Theory and applications in signal processing*, eds. W. Mecklenbrauker, F. Hlawatsch, Elsevier, 1997.

[8] F. Hlawatsch, G. F. Boudreaux-Bartels: "Linear and quadratic time-frequency representations", IEEE Signal Processing Magazine, April 1992, pp.21-61.

[9] A. J. E. M. Jannssen: "Wigner weight functions

and Weyl symbols of nonnegative definite linear operators ",Phillips J. Res., vol.44, 1989, pp.7-42.

[10] H. Kirchauer, F. Hlawatsch, W. Kozek: "Time-frequency formulation and design of nonstationary Wiener filter", Proc. Int. Conf. ASSP, 1995, pp.1549-1552.

[11] W. Kozek: "Time-frequency processing based on the Wigner-Weyl framework", Signal Processing, vol.29, No.1, Oct. 1992, pp.77-92.

[12] W. Kozek, F. Hlawatsch, H. Kirchauer, U. Trautwein: "Correlative time-frequency analysis and classification of non stationary random process", Proc. IEEE Int. Symp. TFTSA, Philadelphia, PA, Oct. 1994, pp.417-420.

[13] G. Matz, F. Hlawatsch, W. Kozek: "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes", IEEE Transaction on Signal Processing, vol.45, No.6, Jun 1997, pp.1520-1534.

[14] A. Papoulis: "Signal Analysis", McGrow-Hill Book Company, 1977.

[15] D. Petranović, S. Stanković, LJ. Stanković: "Special purpose hardware for time-frequency analysis", Electronics Letters, vol.33, No.6, March 1997, pp.464-466.

[16] H. Puder: "Single channel noise reduction using time-frequency dependent voice activity detection ", Proc. 6th Workshop on Acoustic Echo and Noise Control, Pocono Manor, USA, Sept. 1999, pp.68-71.

[17] A. W. Richaczek: "Signal energy distribution in time and frequency", IEEE Transaction on Information Theory, vol.14, 1968, pp.369-374.

[18] LJ. Stanković: "A method for time-frequency analysis", IEEE Transaction on Signal Processing, vol.42, No.1, Jan. 1994, pp.225-229.

[19] LJ. Stanković: "An analysis of some time-frequency and time-scale distributions", Annales des Telecommunications, vol.49, No.9/10, Sep/Oct 1994, pp.505-517.

[20] LJ. Stanković: "A method for improvement energy concentration in the time-frequency signal analysis using L-Wigner distribution", IEEE Transaction on Signal Processing, vol.43, No.5, May 1995 pp.1262-1268.

[21] LJ. Stanković: "Auto-terms representation by the reduced interference distributions; Procedure for kernel design", IEEE Transaction on Signal Processing, vol.44, No.6, June 1996, pp.1557-1564.

[22] LJ. Stanković, V. Ivanović, Z. Petrović: "Unified approach to the noise analysis in the Wigner distribution and spectrogram using the S-method", Annales des Telecommunications, No.11/12, Nov/Dec 1996, pp.585-594.

[23] LJ. Stanković: "On the time-frequency analysis based filtering", Annales des Telecommunications, vol.54, No.5/6, May/June, 2000.

[24] LJ. Stanković, J. Böhme: "Time-frequency analysis of multiple resonances in combustion engine signals", Signal Processing, vol.79, Nov.1999, pp.15-28.

[25] S. Stanković, LJ. Stanković: "An architecture for the realization of a system for time-frequency analysis", IEEE Transaction on Circuits and Systems", vol.44, No.2, Feb. 1997, pp.600-604.

[26] N. Virag: "Single channel speech enhancement based on the masking properties of the human auditory system ", IEEE Transaction on Speech and Audio Processing, vol.7, no.2, March 1999, pp.126-137.

[27] L. A. Zadeh: "Frequency analysis of variable networks", Proc. IRE, vol.67, March 1950, pp.291-297.