# Space/Spatial-Frequency Analysis Based Filtering

LJubiša Stanković, *Senior Member, IEEE*, Srdjan Stanković, *Member, IEEE*, and Igor Djurović, *Student Member, IEEE*

***Abstract*—Space-invariant filtering of signals that overlap with noise in both space and frequency can be inefficient. However, the signal and noise may be well separated in the joint space/spatial-frequency domain. Then, it is possible to benefit from the application of space/spatial frequency approaches. Processing based on these approaches can outperform space or frequency invariant-based methods. To this aim the concept of nonstationary space-varying filtering is introduced in this paper as an extension of the time-varying filtering concept. The filtering definitions are based on statistical averages, although the filtering should commonly be applied knowing only a single noisy signal realization. The procedures that can produce good estimates of quantities crucial for efficient filtering, based on a single noisy signal realization, are considered. Special attention has been paid to the region of support estimation and cross-term effects removal. The efficiency of the proposed space/spatial-frequency filtering concept is tested on the signal forms inspired by the interferograms in optics, including real images as disturbances. Examples demonstrate the superiority of the proposed filtering over the space-invariant one for the considered type of signals and noise.**

***Index Terms*—Estimation, filtering, noisy signals, space varying filtering, time–frequency analysis, Wigner distribution.**

## I. INTRODUCTION

WHEN a two-dimensional (2-D) signal and noise do not occupy the same frequency range, then efficient filtering can be performed using space-invariant filters. However, in the cases when the signal and noise overlap in a significant part of the space and frequency domain, then the stationary filtering may be difficult and inefficient. We will consider noisy signals that occupy the same space and frequency range when their separation may be done in the joint space/spatial-frequency domain. For these kinds of noisy signals, a concept of space-varying filters is presented. It is an extension of the one-dimensional (1-D) time–varying filtering approach, [4], [22], [23], [25], [29], [33] to the 2-D problems. Since the concept of time-varying filtering is based on the time–frequency distributions, we will use joint space/spatial-frequency distributions [7], [8], [16], [17], [27], [37], [39], [42]–[44] in order to define and implement space-varying filtering. The 2-D Wigner distribution, along with the 2-D extension of the Weyl operator [18], [23], [25], [28], is used as a basic distribution. In order to produce undistorted version of the frequency modulated signal passing through the filter whose region of support is ideally concentrated along the local frequency (or group shift), a slight modification of the existing time–frequency filtering relation is proposed and justified. *The implementation is performed using a single noisy signal realization.* The algorithm for the Wigner distribution estimation by using only one noisy signal observation that would be close to the optimal one, with respect to the bias and variance, is presented [11], [20], [35], [36]. This is important since the Wigner distribution plays a crucial role in the region of support estimation and, consequently, determines filtering efficiency. In order to extend the presented forms for the application on multicomponent signals, the multidimensional S-method is used [30], [37]. The results presented in this paper are applied to the space-varying filtering of signals with a high amount of noise, including real image as a disturbance.

The paper is organized as follows. The concept of space-varying filtering is presented in Section II. A slight modification with respect to the common time-varying form is introduced and justified in the Appendix. The pseudo and discrete forms of the filtering relations are given. Efficient region of support estimation, as a crucial part of good filtering, is considered in detail in Section III. Since the Wigner distribution turns out to be the basic form for the region of support estimation, its variance and bias are considered. Based on a specific statistical approach of comparing the bias and the variance, an algorithm for the optimal Wigner distribution estimation, i.e., the filtering region of support estimation with minimal mean square error, is presented in this section. Efficiency of the proposed space-varying filtering is demonstrated in Section IV by examples.

## II. THEORY

Consider a 2-D noisy signal

$$s(x, y) = f(x, y) + \nu(x, y) \tag{1}$$

where $f(x, y)$ is the signal, whereas $\nu(x, y)$ denotes the noise. The above relation may be written in a vector notation as

$$s(\vec{r}) = f(\vec{r}) + \nu(\vec{r}) \text{ where } \vec{r} = (x, y). \tag{2}$$

The nonstationary 2-D filter relation will be defined in analogy with the 1-D time-varying filtering [4], [22], [23], [25], [33], [40]

$$(Hs)(\vec{r}) = \int_{\vec{v}} h(\vec{r} + \vec{v}/2, \vec{r} - \vec{v}/2) s(\vec{r} + \vec{v}) \, d\vec{v} \tag{3}$$

where $h(\vec{r}, \vec{v})$ is an impulse response of the space-varying 2-D filter.

This definition is slightly modified with respect to the existing 1-D definitions [4], [22], [23], [25], [40]. The modification has been introduced in order to provide that for $f(\vec{r}) = Ae^{j\phi(\vec{r})}$, we get $(Hf)(\vec{r}) = cf(\vec{r})$ if the filter in space/spatial-frequency

Manuscript received February 26, 1999; revised March 27, 2000. The work of LJ. Stanković and S. Stanković was supported by the Alexander von Humboldt Foundation. The associate editor coordinating the review of this paper and approving it for publication was Prof. Gregori Vazquez.

LJ. Stanković and I. Djurović are with the Elektrotehnicki Fakultet, University of Montenegro, Podgorica, Montenegro, Yugoslavia (e-mail: l.stankovic@ieee.org).

S. Stanković is with the Institute of Communications Technology, Darmstadt University of Technology, Germany, on leave from the Elektrotehnicki Fakultet, University of Montenegro, Podgorica, Montenegro, Yugoslavia.

Publisher Item Identifier S 1053-587X(00)05975-4.

domain is defined as a delta function $\delta(\vec{\omega} - \nabla\phi(\vec{r}))$ along the local frequency $\nabla\phi(\vec{r})$ for signals that satisfy stationary phase method conditions [2], [6], [26] (see the Appendix).

The optimal transfer function derivation will be done by analogy with the Wiener filter derivation in the stationary signal cases [22], [26], [40]. The error $f(\vec{r}) - (Hs)(\vec{r})$ is orthogonal to the data $s^*(\vec{r} + \vec{\alpha})$ when the mean square error $e^2 = E\{|f(\vec{r}) - (Hs)(\vec{r})|^2\}$ reaches its minimum [26], [40]

$$E\left\{\left[f(\vec{r}) - \int_{\vec{v}} h\left(\vec{r} + \frac{\vec{v}}{2}, \vec{r} - \frac{\vec{v}}{2}\right) s(\vec{r} + \vec{v})\, d\vec{v}\right] s^* (\vec{r} + \vec{\alpha})\right\} = 0. \tag{4}$$

Denote the expected value of the ambiguity function $AF_{ss}(\vec{\theta}, \vec{v})$, [15], [24] as

$$\overline{AF}_{ss}(\vec{\theta}, \vec{v}) = \int_{\vec{r}} E\{s(\vec{r} + \vec{v}/2)s^*(\vec{r} - \vec{v}/2)\} e^{-j\vec{\theta}\vec{r}}\, d\vec{r} \tag{5}$$

and the Fourier transform of $h(\vec{r} + \vec{v}/2, \vec{r} - \vec{v}/2)$ over $\vec{r}$ by

$$A_H(\vec{\theta}, \vec{v}) = \int_{\vec{r}} h(\vec{r} + \vec{v}/2, \vec{r} - \vec{v}/2) e^{-j\vec{\theta}\vec{r}}\, d\vec{r}. \tag{6}$$

From (4), it then follows that

$$\overline{AF}_{fs}(\vec{\theta}, \vec{\alpha}) = \int_{\vec{v}} \int_{\vec{u}} A_H(\vec{u}, -\vec{v}) \overline{AF}_{ss}(\vec{\theta} - \vec{u}, \vec{\alpha} - \vec{v})\, d\vec{v} d\vec{u} \tag{7}$$

for the processes that are mainly concentrated around the ambiguity domain origin (being different from zero only for small $|\vec{\theta}|$, $|\vec{u}|$, $|\vec{\alpha}|$, $|\vec{v}|$) when $e^{j(\vec{\theta}\vec{v} - \vec{u}\vec{\alpha} - \vec{u}\vec{r})/2} \cong 1$. In the 1-D case, these processes are referred to as the *underspread processes* [15], [22], [24].

*1) Note:* Relation (7) directly follows from (4) without any additional assumption if the considered processes are quasistationary [26], [33], [40], i.e., if $E\{f(\vec{r})s^*(\vec{r} + \vec{\alpha})\} = E\{f(\vec{r} - \vec{\alpha}/2)s^*(\vec{r} + \vec{\alpha}/2)\}$, and $E\{s(\vec{r} + \vec{v})s^*(\vec{r} + \vec{\alpha})\} = E\{s(\vec{r} + (\vec{v} - \vec{\alpha})/2)s^*(\vec{r} - (\vec{v} - \vec{\alpha})/2)\}$. It provides an additional physical motivation for modification (3).

The 2-D Fourier transform of (7) results in

$$\overline{WD}_{fs}(\vec{r}, \vec{\omega}) = L_H(\vec{r}, \vec{\omega}) \overline{WD}_{ss}(\vec{r}, \vec{\omega}) \tag{8}$$

where $\overline{WD}_{ss}(\vec{r}, \vec{\omega})$ is the Wigner spectrum (the expected value of the Wigner distribution [9]) of signal $s(\vec{r})$

$$\overline{WD}_{ss}(\vec{r}, \vec{\omega}) = \int_{\vec{v}} E\{s(\vec{r} + \vec{v}/2)s^*(\vec{r} - \vec{v}/2)\} e^{-j\vec{\omega}\vec{v}}\, d\vec{v}$$

and $L_H(\vec{r}, \vec{\omega})$ is defined by

$$L_H(\vec{r}, \vec{\omega}) = \int_{\vec{v}} h\left(\vec{r} - \frac{\vec{v}}{2}, \vec{r} + \frac{\vec{v}}{2}\right) e^{-j\vec{\omega}\vec{v}}\, d\vec{v}. \tag{9}$$

If the signal and noise are not correlated, then

$$L_H(\vec{r}, \vec{\omega}) = \frac{\overline{WD}_{ff}(\vec{r}, \vec{\omega})}{\overline{WD}_{ff}(\vec{r}, \vec{\omega}) + \overline{WD}_{\nu\nu}(\vec{r}, \vec{\omega})}. \tag{10}$$

*2) Note:* For the stationary processes, with $\overline{WD}_{ff}(\vec{r}, \vec{\omega}) = S_{ff}(\vec{\omega})$, $h(\vec{r} - \vec{v}/2, \vec{r} + \vec{v}/2) = h(\vec{v})$, and $L_H(\vec{r}, \vec{\omega}) = H(\vec{\omega}) = FT\{h(\vec{v})\}$, (3), (9), and (10) reduce to the well-known stationary Wiener filter forms [26], [40].

When the Wigner spectrum of signal $f(\vec{r})$ lies inside a space/spatial-frequency region denoted by $R_f$, while the noise is dominantly spread outside this region, then a simple solution satisfying (10) is given by

$$L_H(\vec{r}, \vec{\omega}) = \begin{cases} 1, & \text{for } (\vec{r}, \vec{\omega}) \in R_f \\ 0, & \text{for } (\vec{r}, \vec{\omega}) \notin R_f \end{cases} \tag{11}$$

where $R_f$ is the region where $\overline{WD}_{ff}(\vec{r}, \vec{\omega}) \neq 0$. This is true, for example, for a wide class of frequency modulated (highly concentrated in the space/spatial-frequency plane) signals $f(\vec{r})$ corrupted with a white noise $\nu(\vec{r})$ widely spread in the space/spatial-frequency plane.

*3) Note:* Relation (11) could also be obtained in a semi-intuitive way, as in [23]. Then, some restrictions imposed in its derivation, like the one about underspreadness, would not be necessary either. More details about this derivation, in the 1-D case, may be found in [23].

In the numerical implementations, the pseudo (space limited) forms of the filter relations should be introduced. The pseudo form of operator (3) will be defined as

$$(Hs)(\vec{r}) = \int_{\vec{v}} h\left(\vec{r} + \frac{\vec{v}}{2}, \vec{r} - \frac{\vec{v}}{2}\right) w(\vec{v}) s(\vec{r} + \vec{v})\, d\vec{v}. \tag{12}$$

This relation enables one to use the space limited intervals. It may be shown that for the signals ideally concentrated along the local frequency in the space/spatial-frequency domain, the following important conclusion holds: The window $w(\vec{v})$ does not influence the filter output (12) as far as $w(\vec{0}) = 1$ (see the Appendix). Using the Parseval's theorem, (12) assumes the form

$$(Hs)(\vec{r}) = \frac{1}{4\pi^2} \int_{\vec{\omega}} L_H(\vec{r}, \vec{\omega}) STFT(\vec{r}, \vec{\omega})\, d\vec{\omega} \tag{13}$$

where $STFT(\vec{r}, \vec{\omega})$ is the "short space" Fourier transform defined by

$$STFT(\vec{r}, \vec{\omega}) = \int_{\vec{v}} s(\vec{r} + \vec{v}) w(\vec{v}) e^{-j\vec{\omega}\vec{v}}\, d\vec{v}. \tag{14}$$

Discrete form of (13), as it used in the implementation, is

$$(Hs)(\vec{n}) = \sum_{\vec{k}} L_H(\vec{n}, \vec{k}) STFT(\vec{n}, \vec{k}) \tag{15}$$

where $L_H(\vec{n}, \vec{k})$ assumes unity values where the Wigner distribution of signal is different from zero. Therefore, in order to perform a 2-D space-varying filtering, we should know the $STFT$ and $L_H$. Computation of the $STFT$ is simple, but the problem of $L_H$ determination still remains. Obviously, a precise determination of $L_H$ is directly related to precise $R_f$ determination, which further leads to *the estimation of the Wigner distribution of signal without noise using only one noisy signal observation with the mean square error as small as possible*.

Note that the support region determination based on the squared modulus of $STFT(\vec{r}, \vec{\omega})$, i.e., multidimensional spectrogram, would be appropriate only in the case when *the local frequency does not change over space or changes very slowly* so that it may be considered as a constant within the window $w(\vec{v})$. Otherwise, when the local frequency variations within the considered domain are significant, for low and high frequencies, then its estimation based on either spectrogram or scalogram is not reliable. In addition, the region of support determined by using these distributions would be very spread out and would result in inefficient filtering.

## III. Single Realization-Based Estimation of the Filter Region of Support

### A. Optimal Window Width in the Wigner Distribution

The analysis in this paper is focused on the one realization based filtering of noisy signals. Accurate determination of the filtering region of support, i.e., the Wigner distribution of the signal, is crucially important for the efficient filtering. Two parameters that determine the mean square error of the Wigner distribution estimation are the bias and variance. They will be analyzed for the pseudo Wigner distribution of discrete unknown deterministic noisy signal

$$\begin{aligned} &PWD_{ss}(\vec{n}, \vec{\omega}) \\ &\quad = \sum_{\vec{m}} w(\vec{m}) w^*(-\vec{m}) s(\vec{n}+\vec{m}) s^*(\vec{n}-\vec{m}) e^{-j2\vec{\omega}\vec{m}} \end{aligned} \tag{16}$$

where $s(\vec{n}) = f(n_1, n_2) + \nu(n_1, n_2) = f(\vec{n}) + \nu(\vec{n})$, with $f(\vec{n})$ denoting signal and $\nu(\vec{n})$ additive Gaussian white noise with independent real and imaginary parts and total variance $\sigma_\nu^2$. This form corresponds to the practical cases when only a single realization of the image is known. Therefore, we may treat the signal as deterministic. The noise autocorrelation function is $R_{\nu\nu}(2\vec{m}) = \sigma_\nu^2 \delta(2\vec{m})$.

The pseudo Wigner distribution mean value is [34]

$$\begin{aligned} &E\{PWD_{ss}(\vec{n}, \vec{\omega})\} \\ &\quad = WD_{ff}(\vec{n}, \vec{\omega}) ** FT_{ww}(2\vec{\omega}) + \sigma_\nu^2 ** FT_{ww}(2\vec{\omega}) \end{aligned} \tag{17}$$

where $FT_{ww}(2\vec{\omega})$ is the Fourier transform of the 2-D window $w(\vec{m})w(-\vec{m})$, whereas $**$ is a 2-D convolution along $\omega_1$ and $\omega_2$. Since the term $\sigma_\nu^2 ** FT_{ww}(2\vec{\omega}) = \sigma_\nu^2 w(0, 0) = \sigma_\nu^2$ is constant pedestal and does not depend on the window shape and the signal form, it will not be considered further. The first term can be written as $WD_{ff}(\vec{n}, \vec{\omega}) ** FT_{ww}(2\vec{\omega}) = WD_{ff}(\vec{n}, \vec{\omega}) + bias(\vec{n}, \vec{\omega})$. The bias can be approximated by

$$bias(\vec{n}, \vec{\omega}) = \frac{\partial^2 WD_{ff}(\vec{n},\vec{\omega})}{\partial \omega_1^2} \frac{m_{21}}{8} + \frac{\partial^2 WD_{ff}(\vec{n},\vec{\omega})}{\partial \omega_2^2} \frac{m_{22}}{8} \tag{18}$$

where $m_{2i} = (1/2\pi)^2 \int \omega_i^2 FT_{ww}(\vec{\omega})\, d\vec{\omega}$.

The variance is defined by

$$\sigma_{ss}^2 = E\left\{|PWD_{ss}(\vec{n}, \vec{\omega})|^2\right\} - |E\{PWD_{ss}(\vec{n}, \vec{\omega})\}|^2. \tag{19}$$

Using (19) we get, as in the 1-D case [1], [34], that

$$\sigma_{ss}^2 = \sigma_\nu^2 \sum_{\vec{m}} w^4(\vec{m}) \left[|f(\vec{n}+\vec{m})|^2 + |f(\vec{n}-\vec{m})|^2 + \sigma_\nu^2\right] \tag{20}$$

where a real and even window function is assumed.

For a signal with slow-varying amplitude, we get

$$\sigma_{ss}^2 \cong E_w \sigma_\nu^2 (2A^2(\vec{n}) + \sigma_\nu^2) \tag{21}$$

where $A(\vec{n})$ is the signal's amplitude, and $E_w$ is the window energy. For the 2-D separable window $w(m_1, m_2) = w(m_1)w(m_2)$, where $w(m)$ is the Hanning window, the energy is $E_w = 9N^2/64$, and $N$ is the window width. Similar results could be obtained for any other window shape. For example, for the separable Hamming window, the energy is $E_w = 0.197\,93N^2$, and for the rectangular window, it is $E_w = N^2$.

The mean square error is defined by $e^2(\vec{n}, \vec{\omega}) = bias^2(\vec{n}, \vec{\omega}) + \sigma_{ss}^2$. According to (18) and (21), it may be written as

$$e^2(\vec{n}, \vec{\omega}; N) = C(\vec{n}, \vec{\omega}; N)\frac{1}{N^4} + D(\vec{n})N^2 \tag{22}$$

where

$$\begin{aligned} &C(\vec{n}, \vec{\omega}; N) \\ &\quad = \left(\frac{\partial^2 WD_{ff}(\vec{n},\vec{\omega})}{\partial \omega_1^2} + \frac{\partial^2 WD_{ff}(\vec{n},\vec{\omega})}{\partial \omega_2^2}\right)^2 \frac{\pi^4}{16} \end{aligned} \tag{23}$$

$$D(\vec{n}) = 9(2A^2(\vec{n}) + \sigma_\nu^2)/64. \tag{24}$$

It is important to note that the variance is proportional to the square of the window width, whereas the bias decreases with the same rate. Optimal window width follows from $\partial e^2(\vec{n}, \vec{\omega}; N)/\partial N = 0$. From (22), we get

$$N_{opt} = \sqrt[6]{2C(\vec{n}, \vec{\omega}; N)/D(\vec{n})}. \tag{25}$$

Note that the optimal window width is space and frequency varying since it depends on $C(\vec{n}, \vec{\omega}; N)$. From (22), we can conclude that for the optimal width $N_{opt}$, the ratio of the bias and standard deviation

$$\frac{bias(\vec{n}, \vec{\omega})_{|N=N_{opt}}}{\sigma_{ss|N=N_{opt}}} = \frac{\sqrt{2}}{2} \tag{26}$$

is constant and depends neither on the problem parameters nor on the signal. The optimal window width is given by (25). However, it is not applicable to practical problems since it requires the knowledge of bias parameter $C(\vec{n}, \vec{\omega}; N)$, that is, a function of the distribution derivatives, and is therefore unknown in advance. An algorithm for the optimal window width determination without using the bias parameter $C(\vec{n}, \vec{\omega}; N)$ is described next.

*B. Algorithm*

In order to estimate the Wigner distribution of signal $f(\vec{n})$ and its region of support $R_f$ using the optimal window width for each space and frequency point, we will use a statistical approach [11], [20], [21], [29], [36]. It is based on the fact that the value of the pseudo Wigner distribution $PWD_s(\vec{n}, \vec{\omega}; N)$ of a noisy signal is a scalar random variable. It is spread around the exact Wigner distribution $WD_{ff}(\vec{n}, \vec{\omega}; N)$ with the bias denoted by $bias(\vec{n}, \vec{\omega}; N)$ (including constant pedestal $\sigma_\nu^2$) and the variance $\sigma_{ss}(N)$. As for any biased random variable, we may write the following inequality:

$$|WD_{ff}(\vec{n},\vec{\omega}; N) - [PWD_{ss}(\vec{n},\vec{\omega}; N) - bias(\vec{n}, \vec{\omega}; N)]| \leq \kappa\sigma_{ss}(N). \tag{27}$$

This inequality holds with probability $P(\kappa)$. For large $\kappa$, we have that (27) is satisfied with $P(\kappa) \to 1$ for any distribution law of random variable $PWD_{ss}(\vec{n}, \vec{\omega}; N)$. For example, with $\kappa = 3$ ("three sigma rule") for normal distribution of the random variable, $P(\kappa) = 0.997$.

For the cases of $N$ when the bias is small, i.e., $bias(\vec{n}, \vec{\omega}; N) \leq \triangle\kappa\sigma_{ss}(N)$, the above inequality may be written as

$$|WD_{ff}(\vec{n}, \vec{\omega}; N) - PWD_{ss}(\vec{n}, \vec{\omega}; N)| \leq (\kappa + \triangle\kappa)\sigma_{ss}(N). \tag{28}$$

According to (28), we may write the expressions for the lower and upper confidence interval limit $D(\vec{n}, \vec{\omega}; N) = [L_p, U_p]$ within which the value of the pseudo Wigner distribution $PWD_{ss}(\vec{n}, \vec{\omega}; N)$ is located with probability $P(\kappa)$

$$\begin{aligned} L_p &= WD_s(\vec{n},\vec{\omega}; N_p) - (\kappa + \triangle\kappa)\sigma_{ss}(N_p) \\ U_p &= WD_s(\vec{n}, \vec{\omega}; N_p) + (\kappa + \triangle\kappa)\sigma_{ss}(N_p). \end{aligned} \tag{29}$$

Index $p$ denotes an arbitrary window length $N_p$.

Consider now the successive values of the window lengths $N \in (N_1, N_2)$ such that $N_1 < N_2$. Consider two cases for $N$: a) small bias cases and b) small variance and large bias cases. Confidence intervals calculated with $N_1$, $N_2$ intersect in the case of small bias since, according to (28), the true value of the Wigner distribution belongs to both of these intervals [with probability $P(\kappa)$]. In the cases of very large bias and small variance, the confidence intervals do not intersect $D(\vec{n}, \vec{\omega}; N_1) \cap D(\vec{n}, \vec{\omega}; N_2) = \emptyset$. Values for $\triangle\kappa$, as well as interval for possible $\kappa$, will be determined from the condition that all confidence intervals for small bias such that $bias(\vec{n}, \vec{\omega}, N)/\sigma_{ss}(N) \leq \sqrt{2}/2$ intersect, whereas all confidence intervals for $bias(\vec{n}, \vec{\omega}, N)/\sigma_{ss}(N) > \sqrt{2}/2$ do not intersect. In this way, the examination of the intersection of two confidence intervals will be an indicator for the bias to standard deviation ratio. When critical value $bias(\vec{n}, \vec{\omega}, N)/\sigma_{ss}(N) = \sqrt{2}/2$ is reached, that means that we have the optimal window width for this space and frequency point (26).

To that aim, assume that the window widths have a dyadic scheme

$$\boldsymbol{N} = \{N_p| \quad N_p = 2^{-p}N_{opt}\} \tag{30}$$

with $p = \ldots, -2, -1, 0, 1, 2\ldots$, and $N_{opt}$ is unknown optimal width. It has been assumed that the optimal window width belonged to this set.[1] With (22) and (30), the relations for the bias and standard deviation for an arbitrary window width $N_p$ may be written as functions of the bias and standard deviation for optimal width $N_{opt}$

$$\begin{aligned} \sigma_{ss}(N_p) &= 2^{-p}\sigma_{ss}(N_{opt}) \\ bias(N_p) &= 2^{(2p-1/2)}\sigma_{ss}(N_{opt}). \end{aligned} \tag{31}$$

Consider now three regions $D_0$, $D_1$, $D_2$ and impose, according to the previous discussion, the condition that the regions $D_0$ and $D_1$ intersect and that the regions $D_1$ and $D_2$ do not intersect. Due to monotonicity of the bias and variance if $D_0$ and $D_1$ intersect, then all $D_{p-1}$ and $D_p$ for $p < 1$ will intersect. This means that the region $D_0$ is defined by the optimal window width. Assuming, without loss of generality, that the bias is positive, this condition may be written in the following way:

$$\min\{U_0\} \geq \max\{L_1\}; \quad \max\{U_1\} < \min\{L_2\} \tag{32}$$

where, for example, $\min\{U_0\}$ is the minimal possible value of the upper bound for $p = 0$. Based on (27), it follows that for the window of the width $N_p$, the Wigner distribution of $s(\vec{n})$ is within the interval

$$\begin{aligned} PWD_{ss}(\vec{n},\vec{\omega}; N_p) \in \\ [WD_{ff}(\vec{n},\vec{\omega}; N_p) + bias(N_p) - \kappa\sigma_{ss}(N_p) \\ WD_{ff}(\vec{n}, \vec{\omega}; N_p) + bias(N_p) + \kappa\sigma_{ss}(N_p)]. \end{aligned} \tag{33}$$

The lower and upper bounds of the confidence interval, for a given window width, according to (29), takes the values within

$$\begin{aligned} U_p \in [&WD_{ff}(\vec{n},\vec{\omega}; N_p) + bias(N_p) + \Delta\kappa\sigma_{ss}(N_p) \\ &WD_{ff}(\vec{n},\vec{\omega}; N_p) + bias(N_p) + (2\kappa + \Delta\kappa)\sigma_{ss}(N_p)] \\ L_p \in [&WD_{ff}(\vec{n},\vec{\omega}; N_p) + bias(N_p) - (2\kappa + \Delta\kappa)\sigma_{ss}(N_p) \\ &WDff(\vec{n}, \vec{\omega}; N_p) + bias(N_p) - \Delta\kappa\sigma_{ss}(N_p)]. \end{aligned} \tag{34}$$

Substituting (34) and (31) into (32), we get

$$\sqrt{2} \leq \triangle\kappa < 8\sqrt{2} - 2\kappa. \tag{35}$$

This means $\triangle\kappa > 0$ for $\kappa < 7\sqrt{2}/2 \approx 4.9$. For example, for $\kappa = 4$, we have $\sqrt{2} \leq \triangle\kappa < 8\sqrt{2} - 8 \approx 3.3$. The value of $\kappa$ determines the probability $P(\kappa)$ that (28) is satisfied. In that sense, it is desirable to take as large value of $\kappa$ as possible from the derived interval. For example, with $\kappa = 4$, we get $P(\kappa) > 0.999$. A further increase of $\kappa$ does not make sense since the probability is already close to 1 and other factors, like, for example, effects of discretization, become dominant. In this way, we have defined $\triangle\kappa$ and $\kappa$ as the key factors for the algorithm.

The above procedure may be significantly simplified using only two window widths in calculations [29]. Denote these two

[1]If $N_{opt}$ does not belong to the set $\boldsymbol{N}$, then the algorithm will produce the nearest value to this one. Since the MSE changes slowly around the stationary point, it will not significantly degrade performance.

window widths by $N_1 \times N_1$ and $N_2 \times N_2$, where $N_1 \ll N_2$. Window $N_1 \times N_1$ produces small variance, whereas $N_2 \times N_2$ has small bias. Therefore, when the confidence intervals for these two windows intersect, then it means that the bias is small, and we use $N_1 \times N_1$ in order to reduce the variance. Otherwise, the bias is large, and we use $N_2 \times N_2$ in order to reduce it. The resulting adaptive Wigner distribution is

$$PWD^e_{ss}(\vec{n}, \vec{\omega}) = \begin{cases} PWD_{ss}(\vec{n}, \vec{\omega}; N_1), & \Phi \text{ true} \\ PWD_{ss}(\vec{n}, \vec{\omega}; N_2), & \text{elsewhere} \end{cases} \tag{36}$$

where for $\Phi\ true$, it holds that

$$\begin{aligned} &|PWD_{ss}(\vec{n},\vec{\omega};\, N_1) - PWD_{ss}(\vec{n},\vec{\omega};\, N_2)| \\ &\quad \le (\kappa + \triangle\kappa)[\sigma_s(N_1) + \sigma_s(N_2)]. \end{aligned} \tag{37}$$

In this way, using only two window widths, we may expect significant improvements since the Wigner distribution is either slow-varying (bias very small) or highly concentrated along the local frequency (bias very large). If we used a multiwindow approach with many windows between $N_1 \times N_1$ and $N_2 \times N_2$, then the algorithm applied to the Wigner distribution would select one of these two extreme window widths almost everywhere.

Relation (36) reduces to the calculation of the Wigner distribution and its variance for two-window widths. One possible relation for the variance estimation is

$$\sigma^2_{ss}(N) \approx \frac{9N^2}{64}\left(\sum_{\vec{m}} |s(\vec{n} + \vec{m}|^2\right)^2. \tag{38}$$

It holds for the low signal-to-noise ratio. For cases of small noise, the variance estimation procedure is given in [21]. From (38), it follows that $\sigma^2_{ss}(N_1) = \sigma^2_{ss}(N_2)N_1^2/N_2^2$. Thus, we have defined the algorithm and all parameters for the Wigner distribution calculation, which is then used for the region of support $R_f$ determination and space varying filtering.

### *C. Multicomponent Signal Case*

For multicomponent signal $f(\vec{n}) = \sum_{i=1}^{M} f_i(\vec{n})$, the Wigner distribution contains $M$-signal terms (auto-terms) and $M(M-1)/2$ interference terms (cross-terms) with amplitude that could cover the signal terms

$$\begin{aligned} WD_{ff}(\vec{n}, \vec{\omega}) = &\sum_{i=1}^{M} WD_{ii}(\vec{n}, \vec{\omega}) \\ &+ 2Re\sum_{i=1}^{M}\sum_{j>i}^{M} WD_{ij}(\vec{n}, \vec{\omega}). \end{aligned} \tag{39}$$

This is the reason why the RID class of distributions [19] has been introduced

$$CD_{ff}(\vec{n}, \vec{\omega}) = WD_{ff}(\vec{n}, \vec{\omega}) *_{\vec{n}} *_{\vec{\omega}} \Pi(\vec{n}, \vec{\omega}) \tag{40}$$

where $\Pi(\vec{n}, \vec{\omega})$ [5] is the kernel in the space/spatial-frequency plane, whose Fourier transform $c(\vec{\theta}, \vec{\tau}) = FT\{\Pi(\vec{n}, \vec{\omega})\}$ (kernel in ambiguity plane) is a lowpass filter function. For multicomponent signals, we will use the S-method (SM) [30], [37]

$$\begin{aligned} &SM(\vec{n}, \vec{\omega}) \\ &\quad = \sum_{\vec{\theta}} P\left(\vec{\theta}\right) STFT\left(\vec{n}, \vec{\omega} + \vec{\theta}\right) STFT^*\left(\vec{n}, \vec{\omega} - \vec{\theta}\right) \end{aligned} \tag{41}$$

where $P(\vec{\theta})$ is rectangular window with $2L+1$ width in each direction. The SM belongs to general Cohen class of distributions [5]. When the components of a multicomponent signal do not overlap in the joint space/spatial-frequency plane, it is possible to determine the width of the frequency window $P(\vec{\theta})$ so that the SM is equal to a sum of the Wigner distributions of each signal component individually [30]

$$SM(\vec{n}, \vec{\omega}) = \sum_{i=1}^{M} WD_{ii}(\vec{n}, \vec{\omega}). \tag{42}$$

This interesting property has attracted the attention of some other researches to use the SM in their work [3], [10], [12]. Noise influence on the SM was considered in [31]. The variance for nonoverlapped multicomponent signals can here be expressed as (21)

$$\sigma^2_{ss}(\vec{n}, \vec{\omega}) = \begin{cases} \sigma^2_\nu \sum_{\vec{m}} w^4(\vec{m}) & \\ \quad \times [2A_i^2(\vec{n}) + \sigma^2_\nu], & \text{for } (\vec{n}, \vec{\omega}) \in R_{f_i} \\ \sigma^4_\nu \sum_{\vec{m}} w^4(\vec{m}), & \text{for } (\vec{n}, \vec{\omega}) \notin R_{f_i} \\ & i = 1, \ldots, M. \end{cases} \tag{43}$$

Estimation of the autoterms in the Wigner distribution, based on the SM, could be done as

$$SM^e(\vec{n}, \vec{\omega}) = \begin{cases} SM(\vec{n}, \vec{\omega}; N_1,\, L=0), & \Phi \text{ true} \\ SM(\vec{n}, \vec{\omega};\, N_2,\, L), & \text{elsewhere} \end{cases} \tag{44}$$

where for $\Phi\ true$, it holds that

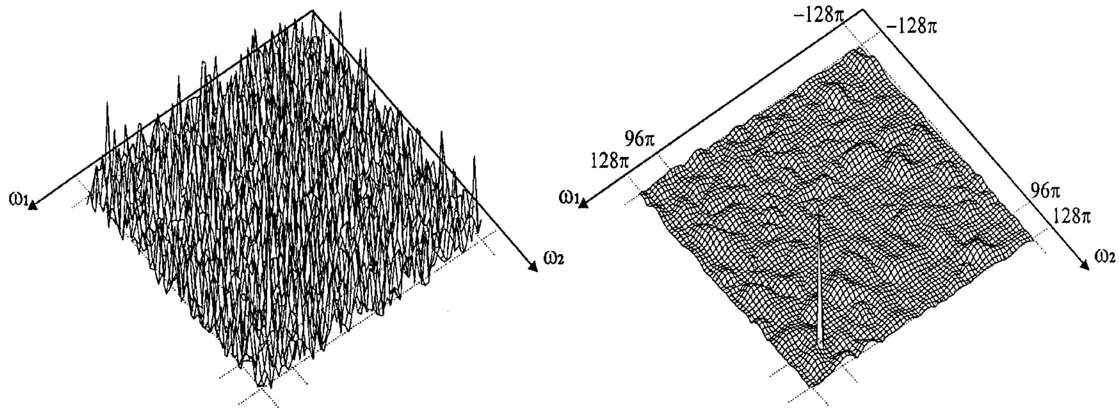
$$\begin{aligned} &|SM(\vec{n},\vec{\omega};\, N_1, L=0) - SM(\vec{n},\vec{\omega};\, N_2, L)| \\ &\quad \le (\kappa + \triangle\kappa)[\sigma_{ss}(N_1,\, L=0) + \sigma_{ss}(N_2,\, L)]. \end{aligned} \tag{45}$$

Note that for $L=0$, the spectrogram (square magnitude of the STFT) is obtained as a special case of the SM. Region of signal's support $R_f$ could be obtained from $SM^e(\vec{n}, \vec{\omega})$ as

$$L_H(\vec{n}, \vec{\omega}) = \frac{1}{2} + \frac{1}{2} sign(SM^e(\vec{n}, \vec{\omega}) - Lev(\vec{n})) \tag{46}$$

where $Lev$ is a given reference level. We used $Lev(\vec{n}) = 0.1\max_{\vec{\omega}}(SM^e(\vec{n}, \vec{\omega}))$.

The procedures and theory presented in the paper will be now summarized in the




(a) (b)

Fig. 1. Determination of the region of support $R_f$ for the space-varying filter at the point $(x, y) = (0.5, 0.5)$. (a) Wigner distribution of noisy signal calculated using the window $N_2 \times N_2 = 128 \times 128$. (b) Wigner distribution of the noisy signal calculated using the proposed algorithm. Distribution from (b) is used for $R_f$ determination at the point $(x, y) = (0.5, 0.5)$.

*Algorithm for Numerical Implementation:*

1) Calculate $STFT(\vec{n}, \vec{\omega})$ for a given noisy signal. In highly nonstationary cases, application of narrower windows is preferred.
2) Calculate the Wigner distribution either by definition (16) or by using (41), with two window widths and values of $L$. Details about (41), which has been used in numerical examples, including its realization without interpolation and oversampling, may be found in [37].
3) Find the "optimal" distribution, for each point, according to (44).
4) Determine the filter region of support by using (46). For one-component signals, determination of $L_H(\vec{n}, \vec{\omega})$ may be simplified just by taking $L_H(\vec{n}, \vec{\omega}) = 1$ at the point $(\vec{n}, \vec{\omega})$ where the maximum of $PWD_{ss}(\vec{n}, \vec{\omega})$ or $SM^e(\vec{n}, \vec{\omega})$ is detected for a given $\vec{n}$. The same procedure may be used for multicomponent signals with a known number of components [33]. Otherwise, the general form (46) must be used.
5) Compute the filtered signal according to (15).

## IV. Numerical Examples

*Example 1:* The presented theory will be illustrated on the numerical example with the signal

$$f(x, y) = 0.5e^{j(96\pi x^2 + 96\pi y^2)} e^{-1.5x^2 - 1.5y^2} \qquad (47)$$

corrupted with a high amount of additive noise with variance $\sigma_{\nu\nu}^2 = 1$, meaning that $10 \log(A^2/\sigma_{\nu\nu}^2) \leq -6$ [dB]. The signal form (47) is inspired with the interferograms in optics.

Fig. 1 demonstrates the algorithm for the region $R_f$ determination at the point $(x, y) = (0.5, 0.5)$. The Wigner distribution calculated using the window width $(N_2 \times N_2) = (128 \times 128)$ is shown in Fig. 1(a). Fig. 1(b) presents the Wigner distribution calculated using the two-window algorithm presented in Section IV with $(N_1 \times N_1) = (16 \times 16)$ and $(N_2 \times N_2) = (128 \times 128)$. The algorithm used the lower variance distribution calculated with $(N_1 \times N_1) = (16 \times 16)$ everywhere, except at the point where the signal energy is concentrated when the lower bias window $(N_2 \times N_2) = (128 \times 128)$ is used.

Noisy signal (47) is considered within the interval $[0 \leq x < 0.5, 0 \leq y < 0.5]$. Signal frequency changes along both axis from 0 to $(3/4)f_{\max}$, where $f_{\max}$ is the maximal frequency for a given sampling interval. In (37), we assumed $\kappa + \triangle\kappa = 3.5$.

The original signal without noise is shown in Fig. 2(a), whereas the noisy signal is given in Fig. 2(b). Signal filtered using stationary filters with the cutoff frequency in both directions $f_c = (3/4)f_{\max}$ is given in Fig. 2(c). Filtering with lower cutoff frequencies reduces the noise but also degrades the signal in Fig. 2(d) and (e) with $f_c = f_{\max}/2$ and $f_c = f_{\max}/4$, respectively. Signal obtained from the noisy signal in Fig. 2(b) using the space-varying filtering presented in this paper, by (3), (13), and the algorithm in Section III, is shown in Fig. 2(f). The advantage of the proposed concept with respect to the stationary filtering is evident.

*Example 2:* For the multicomponent case we will, as an example, consider the signal

$$\begin{aligned} f(x,y) = &\exp(j(-104\pi(x^2+x) - 104\pi y^2)) \\ &+ \exp(j(-104\pi x^2 - 104\pi(y^2+y))) \end{aligned} \qquad (48)$$

corrupted with a high amount of additive noise with variance $\sigma_{\nu\nu}^2 = 1.7$. In the spectrogram calculation, the Hanning window with $(N_1 \times N_1) = (16 \times 16)$ is used, whereas for the SM calculation, the Hanning window with $(N_2 \times N_2) = (128 \times 128)$ and $L = 2$ are taken. The original signal is shown in Fig. 3(a). The noisy signal is given in Fig. 3(b). The signal filtered with the proposed algorithm is presented in Fig. 3(c).

*Example 3:* Finally, we will consider a separation of linear frequency modulated signals from real image. This would correspond to a generalization of the notch filtering in the stationary cases. Here, the local frequency is varying, and the position of the notch frequency changes for each point. Its detection based on the spectrogram would not be precise what would cause an imprecise and wide support region, resulting in unsatisfactory separation. The spectrogram-based region of support determination would be efficient only in the cases of constant (or slow-varying) local frequency. The Wigner distribution approach gives very precise determination of the local frequency when it changes over the image. It results in an efficient separation using the proposed procedure, as demonstrated

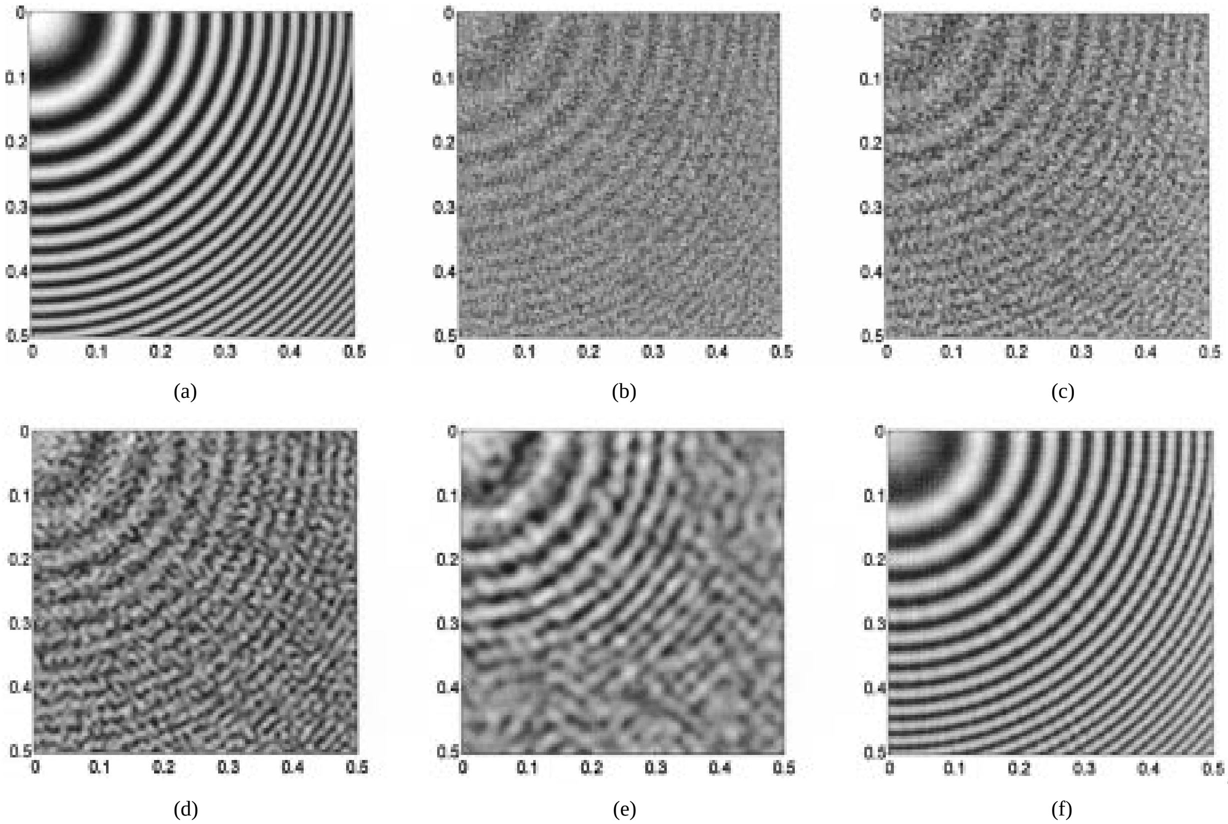



Fig. 2. Filtering of a 2-D signal. (a) Original signal without noise. (b) Noisy signal. (c) Noisy signal filtered using the stationary filter with a cutoff frequency equal to the maximal signal frequency. (d) Noisy signal filtered using the stationary filter with a cutoff frequency equal to a half of the maximal frequency. (e) Noisy signal filtered using the stationary filter with a cutoff frequency equal to a fourth of the maximal frequency. (f) Noisy signal filtered using the space-varying filter.

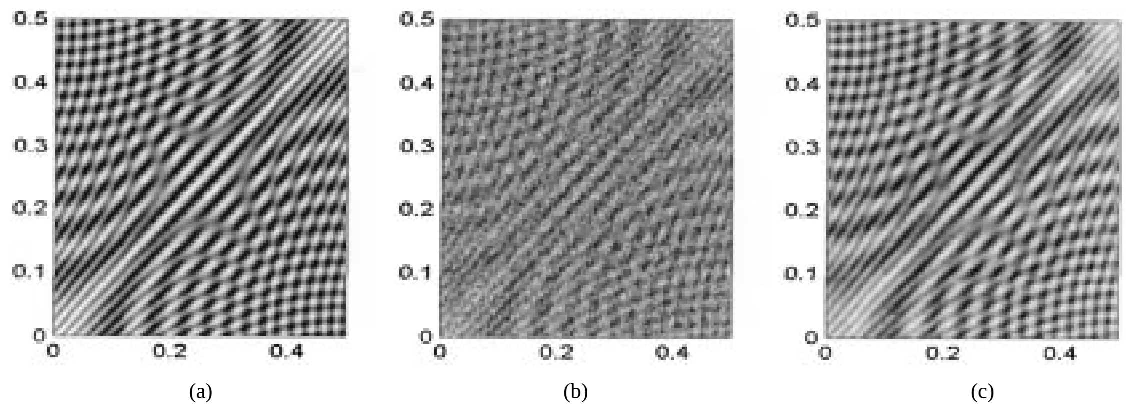



Fig. 3. (a) Two-dimensional multicomponent signal without noise. (b) Noisy 2-D multicomponent signal. (c) Filtered noisy signal.

in Fig. 4. Fig. 4(a) shows original image. A part of the frequency-modulated signal added to the image Fig. 4(a) is shown in Fig. 4(b). The image corrupted with the frequency modulated signal is shown in Fig. 4(c), whereas the frequency-modulated signal extracted from Fig. 4(c), by using the proposed procedure, is shown in Fig. 4(d). The reconstruction is very good, although the ratio of the original image maximal value to the frequency-modulated signal maximal value was extremely high

$$20\log_{10}\frac{\max\{\nu(\vec{r})\}}{\max\{f(\vec{r})\}} = 18.5[\text{dB}]$$

where $\nu(\vec{r})$ is the image, and $f(\vec{r})$ denotes frequency-modulated signal added to image.

## V. Conclusion

The concept of space-varying filtering of multidimensional signals is presented. It has been shown that it can outperform the space invariant approaches when the signal and noise overlap in both space and frequency domains separately but not in the joint space/spatial-frequency domain. Special attention has been paid to the realization of space/spatial-frequency-based filtering using only a single noisy signal realization. In order to produce accurate region of support estimation, the optimization of the Wigner distribution parameters is considered. This optimization produces space/spatial-frequency varying parameters based on a specific statistical approach of comparing the bias and variance of distribution. Further modifications are done in order to apply the algorithm on the multicomponent signals. Efficiency of

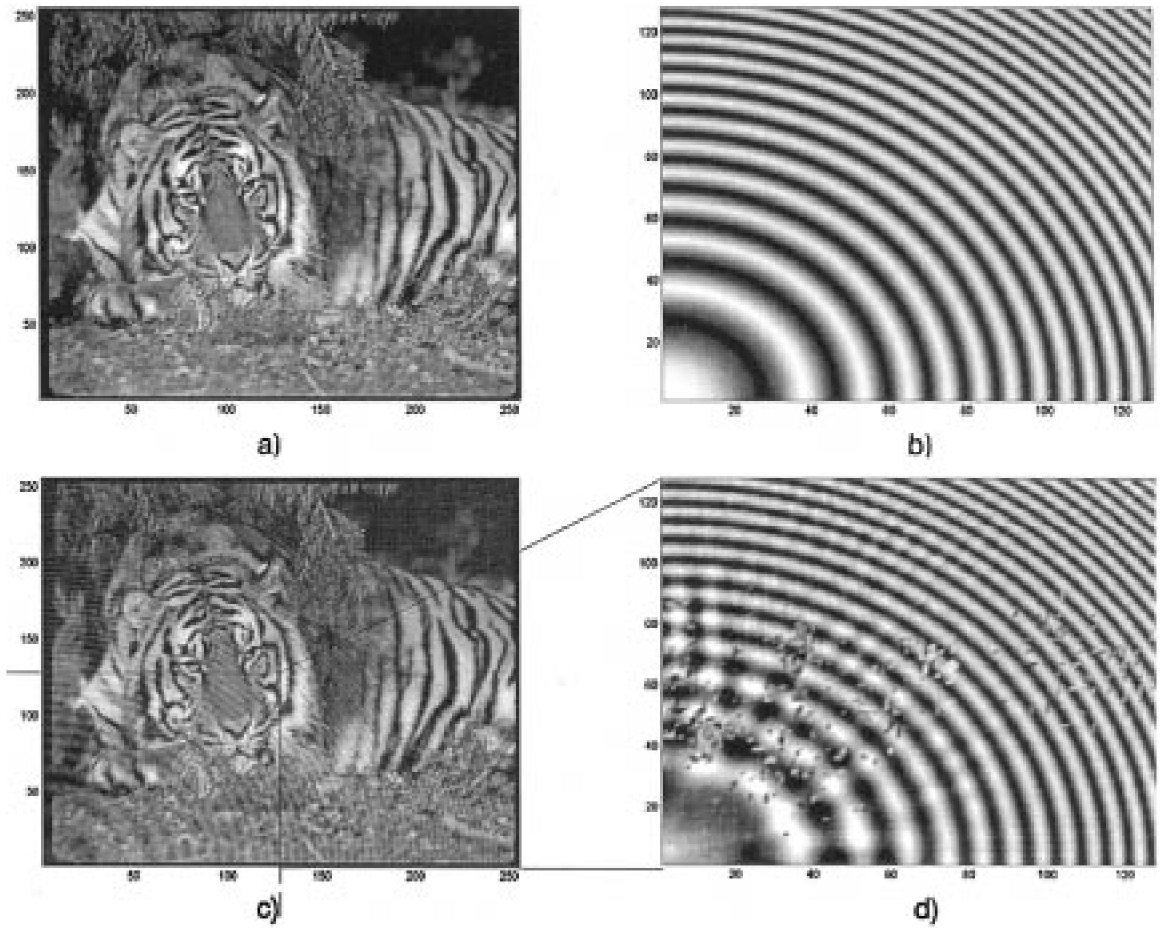


Fig. 4. (a) Original image. (b) Part of the linear frequency-modulated signal added to the image. (c) Sum of the previous two signals. (d) Reconstructed linear frequency modulated signal from (c).

the proposed filtering, and the algorithm for its support determination, has been illustrated on noisy signals whose form is inspired by the interferograms in optics, as well as on the signal corrupted with a real image. The second example is inspired by the watermarking in the space/spatial-frequency domain. From the presented theory and examples, we may conclude that in the cases when the signal and noise do not significantly overlap in the joint space/spatial-frequency domain, the proposed filtering may produce better results than the space or frequency-invariant one. Beside the application in filtering of images, this paper offers an interesting application possibility in watermarking in the space/spatial-frequency domain. This is a topic of current research [38].

## APPENDIX
## ON THE MODIFICATION (3)

The property of the space/spatial frequency distributions that they are well concentrated around the local frequency $\vec{\omega}_l(\vec{r}) = \nabla\phi(\vec{r})$ is one of the basic points for their introduction and application. The Wigner distribution can achieve complete concentration for the signals whose local frequency variations could be considered as linear within the considered domains, while, in order to improve concentration for nonlinear local frequency forms, various modifications have been defined [2], [3], [5], [20], [32].

Let us consider a 2-D FM signal $f(\vec{r}) = Ae^{j\phi(\vec{r})}$. Assume that the signal satisfies the condition that its Fourier transform may be obtained by using a 2-D form of the stationary phase method [2], [6], [26], which is given as

$$F(\vec{\omega}) = \int_{\vec{r}} Ae^{j\phi(\vec{r})-j\vec{r}\vec{\omega}}\, d\vec{r} = Ae^{j\phi(\vec{r}_0)-j\vec{r}_0\vec{\omega}} \sqrt{\frac{(2\pi j)^2}{\dfrac{\partial^2\phi(\vec{r}_0)}{\partial x^2}\dfrac{\partial^2\phi(\vec{r}_0)}{\partial y^2}}}$$

where $\vec{r}_0$ is the point where $\vec{\omega} = \nabla\phi(\vec{r}_0)$. Assume that we have achieved ideally concentrated space/spatial-frequency representation, which is given by $\delta(\vec{\omega} - \nabla\phi(\vec{r}))$, for $f(\vec{r})$. When there is no input noise $s(\vec{r}) = f(\vec{r})$, one expects from a filtering relation that it can produce undistorted signal at the output, at least in this ideal case.

According to Parseval's theorem, from (3), we get

$$\begin{aligned}(Hs)(\vec{r}) &= \int_{\vec{v}} h\left(\vec{r}+\frac{\vec{v}}{2}, \vec{r}-\frac{\vec{v}}{2}\right) s(\vec{r}+\vec{v})\, d\vec{v} \\ &= \frac{1}{4\pi^2}\int_{\vec{\omega}} L_H(\vec{r}, \vec{\omega})F(\vec{\omega})e^{j\vec{r}\vec{\omega}}\, d\vec{\omega} \\ &= \frac{A}{2\pi}\int_{\vec{\omega}} \delta(\vec{\omega}-\nabla\phi(\vec{r}))\frac{e^{j\phi(\vec{r}_0)-j\vec{r}_0\vec{\omega}+j\vec{r}\vec{\omega}}}{\sqrt{-\dfrac{\partial^2\phi(\vec{r}_0)}{\partial x^2}\dfrac{\partial^2\phi(\vec{r}_0)}{\partial y^2}}}\, d\vec{\omega} \\ &= cAe^{j\phi(\vec{r})}\end{aligned}$$

since, when the conditions for the stationary phase method applications are satisfied, then from $\vec{\omega} = \nabla\phi(\vec{r}_0) = \nabla\phi(\vec{r})$ follows $\vec{r}_0 = \vec{r}$, and $c = 1/(2\pi)\sqrt{-1/[(\partial^2\phi(\vec{r})/\partial x^2)(\partial^2\phi(\vec{r})/\partial y^2)]}$. Thus, at the output, we have obtained the original signal with an amplitude variation depending on the variations of local frequency. The same holds for (12) with $w(0) = 1$. In numerical realizations, the delta function assumes unity values, along the local frequency plane, so that it satisfies (11).

The commonly used definition for time-varying filtering, corresponding to $(Hs)(\vec{r}) = \int_{\vec{v}} h(\vec{r}+\vec{v}/2, \vec{r}-\vec{v}/2)s(\vec{v})d\vec{v}$, would in this case produce $(Hs)(\vec{r}) = cAe^{j\phi(\vec{r})}e^{-j\vec{r}\nabla\phi(\vec{r})}$. *This is not a desired output.* Beside amplitude variation signal has significant and complex phase deviation. This illustrates our motivation for a slight modification of the space/spatial-frequency filtering definition. By the way, the commonly used form of filtering was not able to recover signals in our numerical examples.

From the theoretical point of view, it is interesting to mention the case when the support function is not $\delta(\vec{\omega} - \nabla\phi(\vec{r}))$ but the geometrical mean of the ideally concentrated distributions along the local frequency $\delta(\vec{\omega} - \nabla\phi(\vec{r}))$ and the group shift $\delta(\vec{r} - \vec{r}_g(\vec{\omega}))$. For asymptotic signals, $\vec{r}_g(\vec{\omega}) = \nabla^{-1}\phi(\vec{\omega})$ holds [2]. Then, for

$$\begin{aligned} L_H(\vec{r},\vec{\omega}) &= 2\pi j\sqrt{\delta(\vec{\omega} - \nabla\phi(\vec{r}))\delta(\vec{r} - \nabla^{-1}\phi(\vec{\omega}))} \\ &= 2\pi\sqrt{-(\partial^2\phi(\vec{r})/\partial x^2)(\partial^2\phi(\vec{r})/\partial y^2)}\delta(\vec{\omega} - \nabla\phi(\vec{r})) \end{aligned}$$

the output signal would be the same as the input signal, without additional amplitude variations $(Hs)(\vec{r}) = Ae^{j\phi(\vec{r})}$.

## REFERENCES


[1] M. G. Amin, "Minimum variance time-frequency distribution kernels for signals in additive noise," *IEEE Trans. Signal Processing*, vol. 44, pp. 2352–2356, Sept. 1996.

[2] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal—Part 1: Fundamentals," *Proc. IEEE*, vol. 80, pp. 519–538, Apr. 1992.

[3] B. Boashash and B. Ristić, "Polynomial time-frequency distributions and time-varying higher order spectra: Applications to analysis of multicomponent FM signals and to treatment of multiplicative noise," *Signal Process.*, vol. 67, pp. 1–23, May 1998.

[4] G. F. Boudreaux-Bartels and T. W. Parks, "Time-varying filtering and signal estimation using the Wigner distribution," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 442–451, June 1986.

[5] L. Cohen, *Time-frequency Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[6] E. T. Copson, *Asymptotic Expansions*. New York: Cambridge Univ. Press, 1967.

[7] G. Cristobal, J. Bescos, and J. Santamaria, "Image analysis through the Wigner distribution function," *Appl. Opt.*, vol. 28, no. 2, pp. 262–271, 1989.

[8] G. Cristobal, C. Gonzalo, and J. Bescos, "Image filtering and analysis through the Wigner distribution function," in *Advances in Electronics and Electron Physics*, P. W. Haekes, Ed. Boston, MA: Academic, 1991.

[9] P. Flandrin and W. Martin, "The Wigner-Ville spectrum of nonstationary processes," in *The Wigner Distribution: Theory and Applications in Signal Processing*, W. Mecklenbrauker and F. Hlawatsch, Eds. Amsterdam, The Netherlands: Elsevier, 1997.

[10] B. Friedlander and L. L. Scharf, "On the structure of time-frequency spectrum estimators," IEEE Trans. Signal Processing, to be published.

[11] A. Goldenshluger and A. Nemirovski, "On spatial adaptive estimation of nonparametric regression," Technion, Haifa, Israel, Res. Rep. 5/94, Nov. 1994.

[12] P. Goncalves and R. G. Baraniuk, "Pseudo affine Wigner distributions: Definition and kernel formulation," *IEEE Trans. Signal Processing*, vol. 46, pp. 1505–1517, June 1998.

[13] S. B. Hearon and M. G. Amin, "Minimum-variance time-frequency distributions kernels," *IEEE Trans. Signal Processing*, vol. 43, pp. 1258–1262, May 1995.

[14] F. Hlawatsch and G. F. Broudreaux-Bartels, "Linear and quadratic time-frequency signal representation," *IEEE Signal Processing Mag.*, pp. 21–67, Apr. 1992.

[15] F. Hlawatsch, G. Matz, H. Kirchauer, and W. Kozek, "Time-frequency formulation and design of time-varying optimal filters," Vienna Univ. Technol., Vienna, Austria, Tech. Rep. 97-1, Jan. 1997.

[16] J. Hormigo and G. Cristobal, "High resolution spectral analysis of images using the pseudo-Wigner distribution," *IEEE Trans. Signal Processing*, vol. 46, pp. 1757–1763, June 1998.

[17] L. Jacobson and H. Wechsler, "Joint spatial/spatial-frequency representation," *Signal Process.*, vol. 14, pp. 37–68, 1988.

[18] A. J. E. M. Janssen, "Wigner weight functions and Weyl symbols of nonnegative definite linear operators," *Philips J. Res.*, vol. 44, pp. 7–42, 1989.

[19] J. Jeong and W. J. Williams, "Kernel desing for reduced interference distributions," *IEEE Trans. Signal Processing*, vol. ASSP-34, pp. 402–412, June 1986.

[20] V. Katkovnik, "Adaptive local polynomial periodogram for time-varying frequency estimation," in *Proc. IEEE-SP TFTS Anal.*, Paris, France, June 1996.

[21] V. Katkovnik and LJ. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data-driven window length," *IEEE Trans. Signal Processing*, vol. 46, pp. 2315–2325, Sept. 1998.

[22] H. Kirchauer, F. Hlawatsch, and W. Kozek, "Time-frequency formulation and design of nonstationary Wiener filters," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 1995, pp. 1549–1552.

[23] W. Kozek, "Time-frequency processing based on the Wigner-Weyl framework," *Signal Process.*, vol. 29, no. 1, pp. 77–92, Oct. 1992.

[24] W. Kozek, F. Hlawatsch, H. Kirchauer, and U. Trautwein, "Correlative time-frequency analysis and classification of nonstationary random processes," in *Poc. IEEE Int. Symp. TFTSA*, Philadelphia, PA, Oct. 1994, pp. 417–420.

[25] G. Matz, F. Hlawatsch, and W. Kozek, "Generalized evolutionary spectral analysis and the Weyl spectrum of nonstationary random processes," *IEEE Trans. Signal Processing*, vol. 45, pp. 1520–1534, June 1997.

[26] A. Papoulis, *Signal Analysis*. New York: McGraw-Hill, 1977.

[27] T. Reed and H. Wechsler, "Segmentation of textured images and Gestalt organization using spatial/spatial-frequency representations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 1–12, Jan. 1990.

[28] R. G. Shenoy and T. W. Parks, "The Weyl correspondence and time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 318–331, Feb. 1994.

[29] LJ. Stanković, "Algorithm for the Wigner distribution of noisy signals realization," *Electron. Lett.*, vol. 34, no. 7, pp. 622–624, Apr. 2, 1998.

[30] ——, "A method for time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 42, pp. 225–229, Jan. 1994.

[31] LJ. Stanković, V. Ivanović, and Z. Petrović, "Unified approach to the noise analysis in the Spectrogram and Wigner distribution," *Analles des Telecomm.*, vol. 51, no. 11/12, pp. 585–594, Nov./Dec. 1996.

[32] LJ. Stanković, "Highly concentrated time-frequency distributions: Pseudo quantum signal representation," *IEEE Trans. Signal Processing*, vol. 45, pp. 543–552, Mar. 1997.

[33] ——, "On the time-frequency analysis based filtering," *Annales des Telecommun.*, vol. 55, no. 5/6, May/June 2000.

[34] LJ. Stanković and S. Stanković, "On the Wigner distribution of discrete-time noisy signals with application to the study of quantization effects," *IEEE Trans. Signal Processing*, vol. 42, pp. 1863–1867, July 1994.

[35] LJ. Stanković and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with variable window width," *IEEE Signal Processing Lett.*, vol. 5, pp. 224–227, Sept. 1998.

[36] ——, "The Wigner distribution of noisy signals with adaptive time-frequency varying window," *IEEE Trans. Signal Processing*, vol. 47, pp. 1099–1108, Apr. 1999.

[37] S. Stanković, LJ. Stanković, and Z. Uskoković, "On the local frequency, group shift and cross-terms in some multidimensional time-frequency distributions: A method for multidimensional time-frequency analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1719–1725, July 1995.

[38] S. Stanković, I. Djurović, and I. Pitas, "Watermarking in the space/spatial-frequency domain using two-dimensional Radon-Wigner distribution," *IEEE Trans. Signal Processing*, to be published.

[39] H. Suzuki and F. Kobayashi, "A method of two-dimensional spectral analysis using the Wigner distribution," *Electron. Commun. Jpn.*, vol. 75, no. 1, pp. 1006–1013, 1992.

[40] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 1968.

[41] D. Wu and J. M. Morris, "Discrete Cohen's class of distributions," in *Proc. IEEE-SP TFTS Anal.*, Philadelphia, PA, Oct. 1994, pp. 532–535.

[42] Y. M. Zhu, F. Peyrin, and R. Goutte, "Transformation de Wigner Ville: Description d'un nouvel outil de traitment du signal et des images," *Ann. Telecommun.*, vol. 42, no. 3–4, pp. 105–117, 1987.

[43] Y. M. Zhu, R. Goutte, and F. Peyrin, "The use of a two-dimensional Wigner-Ville distribution for texture segmentation," *Signal Process.*, vol. 19, no. 3, pp. 205–222, 1990.

[44] Y. M. Zhu, R. Goutte, and M. Amiel, "On the use of a two-dimensional Wigner-Ville distribution for texture segmentation," *Signal Process.*, vol. 30, pp. 329–354, 1993.


**LJubiša Stanković** (M'91–SM'96) was born in Montenegro, Yugoslavia, on June 1, 1960. He received the B.S. degree in electrical enginering from the University of Montenegro, Podgorica, Yugoslavia, in 1982 with the honor of "the best student at the University," the M.S. degree in electrical enginering from the University of Belgrade, Belgrade, Yugoslavia, in 1984, and the Ph.D. degree in electrical enginering from the University of Montenegro in 1988.

As a Fulbright grantee, he spent 1984 to 1985 at the Worcester Polytechnic Institute, Worcester, MA. He was also active in politics, as a Vice President of the Republic of Montenegro from 1989 to 1991 and then the leader of democratic (anti-war) opposition in Montenegro from 1991 to 1993. Since 1982, he has been on the Faculty of the University of Montenegro, where he presently holds the position of a Full Professor. From 1997 to 1999, he was on leave at the Ruhr University Bochum, Bochum, Germany, with Signal Theory Group, which was supported by the Alexander von Humboldt foundation. His current interests are in the signal processing and electromagnetic field theory. He has published about 200 technical papers, 47 of them in the leading international journals, mainly IEEE editions. He has published several textbooks on signal processing (in Serbo-Croatian) and the monograph *Time–Frequency Signal Analysis* (in English).

Dr. Stanković was awarded the highest state award of the Republic of Montenegro for scientific achievements in 1997. He is a member of the National Academy of Science and Art of Montenegro (CANU).

**Srdjan Stanković** (M'94) was born in Montenegro, Yugoslavia, on May 9, 1964. He received the B.S. degree in electrical engineering (with honors) in 1988 from the University of Montenegro, Podgorica, Yugoslavia, the M.S. degree in electrical engineering from the University of Zagreb, Zagreb, Croatia, in 1991, and the Ph.D. in electrical engineering from the University of Montenegro in 1993.

From 1988 to 1992, he was with the Aluminum Plant of Podgorica, as a Research Assistant. In 1992, he joined the faculty of the Electrical Engineering Department, University of Montenegro, where he is currently an Associate Professor. His interests are in signal processing and digital electronics. Several of his papers have appeared in leading journals. He published a textbook on electronic devices in Serbo-Croatian and coauthored a monograph on time–frequency signal analysis. He is now on leave at the Institute of Communications Technology, Darmstadt University of Technology, Darmstadt, Germany.

Prof. Stanković was awarded the biannual young researcher prize for 1995 by the Montenegrin Academy of Science and Art.

**Igor Djurović** (S'99) was born in Montenegro in 1971. He received the B.S. and M.S. degrees, both in electrical engineering, from the University of Montenegro, Podgorica, Yugoslavia, in 1994 and 1996, respectively. Currently, he is a Teaching Assistant and is pursuing the Ph.D. degree in electrical engineering in the area of time–frequency signal analysis.

His current research interests include application of virtual instruments, time–frequency analysis-based methods for signal estimation and filtering, fractional Fourier transform applications, image processing, and digital watermarking.