Nonparametric Algorithm for Local Frequency Estimation of Multidimensional Signals

Igor Djurović, LJubiša Stanković

Abstract— **Local frequency (LF) estimation of multidimensional (md) signals is considered. The md-Wigner distribution (WD) is used as the LF estimator. The LF is estimated based on the positions of the WD maxima. A nonparametric algorithm for the LF estimation is developed. It is based on the intersection of confidence intervals (ICI) rule. This algorithm produces an adaptive window size in the WD which gives almost minimal mean squared error (MSE) of the estimate. A simplified version of this algorithm is developed, with the starting estimate being produced with the WD of 1D signals. Theory is illustrated on examples.**

I. Introduction

Time-frequency representations (TF) have numerous applications in estimation of signal parameters [1]-[5]. Multidimensional (md) TF representations can also be used for the same purpose thanks to the methods for their efficient calculation and hardware implementations [6]-[7]. Sometimes, methods for estimation of md-signal parameters are just straightforward generalizations of the corresponding ones developed for the 1D cases. However, in many cases it is not possible to perform a direct generalization of the 1D methods. Some of the latest developments in the area of md-signal analysis can be found in [8]-[10].

The problem of local frequency (LF) estimation appears in optical microscopy, analysis of optical interferograms, SAR imaging, digital watermarking, etc. [11]. In this paper, the md-Wigner distribution (WD) is used as the LF estimator. Our goal is to determine the window size in the WD that will produce an estimate of the LF with the mean squared error (MSE) as close as possible to its minimal value. A nonparametric algorithm for the instantaneous frequency (IF) estimation based on the WD, in the case of 1D signals, has been presented in [12]. The algorithm is based on the intersection of confidence intervals (ICI) rule. The MSE minimum is achieved at the optimal bias-to-variance trade-off of the estimate. *The bias and variance of the LF estimate in one direction depend on window sizes along all coordinates.* This dependence is different for different directions. It means that an algorithm for the LF estimation needs a *multiparameter optimization.* To this aim, the algorithm based on the calculation of several WDs is developed. Since the calculation of an md-WD can be time-consuming, a simplified version of the algorithm is proposed, as well. The first step in this algorithm is application of the nonparametric algorithm to the estimates produced with the WDs along one direction. In the second step, the algorithm based on the md-WD calculation is performed only in the points where the accuracy of the simplified algorithm is not satisfactory.

The paper is organized as follows. The WD as an LF estimator is described in Section II. Algorithm for the LF estimation is presented in Section III. A simplified version of the algorithm is given in Section IV. A numerical example is presented in Section V. Details on the performance analysis of the WD as an LF estimator, including an overview of the nonparametric algorithm that produces bias-to-variance trade-off close to the optimal one, are given in Appendices.

IEEE Transactions on Image Processing, vol.12, 2003.

II. Local Frequency Estimation

Consider an md-FM signal $m(\vec{r}) = Ae^{j\phi(\vec{r})}$, $\vec{r} = (r_1, r_2, ..., r_Q)$, $\vec{r} \in \mathbb{R}^Q$. Our goal is to estimate the LF defined as a vector $\vec{\omega}(\vec{r}) = \{\partial\phi(\vec{r})/\partial r_i = \omega_i(\vec{r}), i = 1, 2, ..., Q\}$, based on discrete-time observations

$$s(\vec{n}\rho) = m(\vec{n}\rho) + \nu(\vec{n}\rho) = Ae^{j\phi(\vec{n}\rho)} + \nu(\vec{n}\rho). \quad (1)$$

In (1), ρ is the sampling interval along each coordinate, $\vec{n} = (n_1, n_2, ..., n_Q)$, $\vec{n} \in \mathbb{Z}^Q$, $\nu(\vec{n}\rho)$ is an additive, white, complex, zero-mean Gaussian noise (AWGN), $E\{\nu(\vec{n}\rho)\} = 0$, with covariance function $E\{\nu(\vec{n}\rho)\nu(\vec{m}\rho)\} = \sigma_\nu^2\delta(\vec{n} - \vec{m})$, where $\delta(\vec{n}) = \prod_{i=1}^{Q}\delta(n_i)$, and $\delta(n_i)$ is the unit impulse. We will assume that $\vec{\omega}(\vec{r})$ is an arbitrary smooth differentiable function of $\vec{r}$, with bounded partial derivatives $|\partial^q\vec{\omega}(\vec{r})/\prod_{k=1}^{q}\partial r_{p_k}| \leq M(\vec{r}) < \infty$, for $p_k \in [1, Q]$.

Since the WD is highly concentrated around the LF, i.e., the signal auto-term is concentrated around the LF, the WD can be used to define the LF estimator for nonstationary signals. The WD of md-signals is defined by:

$$WD_{\vec{h}}(\vec{r}, \vec{\omega}) =$$

$$\sum\nolimits_{\vec{n}} w_{\vec{h}}(\vec{n}\rho)s(\vec{r} + \vec{n}\rho)s^*(\vec{r} - \vec{n}\rho)e^{-j2\vec{\omega}\vec{n}\rho}, \quad (2)$$

where $\vec{\omega} = (\omega_1, \omega_2, ..., \omega_Q)$, and $\omega_i \in [-\pi/2\rho, \pi/2\rho)$, $i = 1, ..., Q$, while $w_{\vec{h}}(\vec{n}\rho)$ is a separable window function, limited along all coordinates by $\vec{h} = \{h_1, h_2, ..., h_Q\}$, $w_{\vec{h}}(\vec{n}\rho) = \prod_{i=1}^{Q}\rho w(n_i\rho/h_i)/h_i$, where $w(r)$ is a real-valued and even function that satisfies $w(r) = 0$ for $|r| > 1/2$. The LF can be estimated by using the WD as:

$$\widehat{\vec{\omega}}_{\vec{h}}(\vec{r}) = \arg[\max_{\vec{\omega}} WD_{\vec{h}}(\vec{r}, \vec{\omega})], \quad (3)$$

with the estimation error:

$$\Delta\widehat{\vec{\omega}}_{\vec{h}}(\vec{r}) = \vec{\omega}(\vec{r}) - \widehat{\vec{\omega}}_{\vec{h}}(\vec{r})$$

$$= \{\Delta\omega_{\vec{h}}^{(i)}(\vec{r}), i = 1, 2, ..., Q\}. \quad (4)$$

In general, the LF estimator (3) is biased. Approximative expressions for the bias and variance of $\Delta\widehat{\vec{\omega}}_{\vec{h}}(\vec{r})$ will be presented next. To simplify the notation, components of $\Delta\widehat{\vec{\omega}}_{\vec{h}}(\vec{r})$ will be denoted as $\{\Delta\omega_i, i = 1, 2, ..., Q\}$. Expression for the estimator variance can be derived by following the results from [12]. It is of the form:

$$var(\Delta\omega_i) = c \cdot h_i^{-3}\prod\nolimits_{j=1, j\neq i}^{Q} h_j^{-1}, \quad (5)$$

where

$$c = \frac{\sigma_\nu^2(2A^2 + \sigma_\nu^2)}{A^4}\frac{(F_{(0,2)}\rho)^{Q-1}(\rho F_{(2,2)})}{F_{(2,1)}^2 F_{(0,1)}^{2Q-2}}. \quad (6)$$

The constants $F_{(a,b)}$ depend on the chosen window shape only

$$F_{(a,b)} = \sum_n (n\rho/h_i)^a w^b(n\rho/h_i)\rho/h_i. \quad (7)$$

For a small ρ/h_i, the constants $F_{(a,b)}$ can be written in a very simple form $F_{(a,b)} = \int_{-1/2}^{1/2} r^a w^b(r)dr$.

Estimator bias can be approximately expressed as:

$$E\{\Delta\omega_i\} \approx \sum\nolimits_{j=1}^{Q} b_{j,i}h_j^2, \quad (8)$$

where

$$b_{j,i} = \begin{cases} \frac{\partial^3\phi(\vec{r})}{\partial r_i^3} \cdot \frac{F_{(0,1)}^{Q-1}F_{(4,1)}}{6F_{(2,1)}}, & j = i \\ \frac{\partial^3\phi(\vec{r})}{\partial r_i^2\partial r_j} \cdot \frac{F_{(0,1)}^{Q-2}F_{(2,1)}}{2}, & j \neq i. \end{cases} \quad (9)$$

Constants $b_{j,i}$ depend on the third order partial derivatives of the signal's phase, and on the chosen window function. Index i denotes estimation along the frequency coordinate ω_i. The approximative nature of (8) is due to the neglected fifth and higher order derivatives. We can easily conclude that the same window can produce different estimation accuracies along different frequency coordinates. More details on the variance and bias derivation are given in Appendix A.

The MSE of the LF estimation along ω_i coordinate can be approximated for each $\vec{r}$, according to expressions (5) and (8), as:

$$MSE_i = (E\{\Delta\omega_i\})^2 + var(\Delta\omega_i)$$

$$= (\sum\nolimits_{j=1}^{Q} b_{j,i}h_j^2)^2 + c \cdot h_i^{-3}\prod\nolimits_{j=1, j\neq i}^{Q} h_j^{-1}. \quad (10)$$

Optimal window size can be derived from the system of equations: $\partial MSE_i/\partial h_i|_{\vec{h}=\vec{h}_{opt}} = 0$,

$i = 1, ..., Q$. However, expressions for the optimal window size (Appendix A) contain unknown higher-order derivatives of the LF, and they cannot be used in practice. In order to solve this problem, a non-parametric algorithm for the LF determination is derived in the next section.

III. Basic Algorithm

The non-parametric algorithm based on the confidence intervals is derived in [12] for one-dimensional estimation problems. Details on the realization of this algorithm are given in Appendix B. Detailed analysis of the algorithm parameters is presented in [13], [14]. Since the problem considered here is multidimensional, we present the algorithm modification which can be used in the LF estimation. For the sake of notation simplicity, consider the case of a 2D signal. Let the signal coordinates be denoted by $\vec{r} = (x, y)$, frequency coordinates by $\vec{\omega} = (\omega_x, \omega_y)$, and the window size as $\vec{h} = (h_x, h_y)$. Consider the following set of the window sizes:

$$H^2 = H \times H,$$

$$H = \{h^{(s)} | h^{(s)} = 2h^{(s-1)}, s = 1, ..., J\}, \quad (11)$$

where $h^{(0)}$ is length of the narrowest window from H that is suitable for FFT algorithm application, for example $h^{(0)} = 2\rho$ or $h^{(0)} = 4\rho$. The WD obtained with a window of size $h_x \times h_y$ is denoted by $WD_{\{h_x,h_y\}}(\vec{r}, \vec{\omega})$, while the corresponding LF estimate components are denoted by $\hat{\omega}_{x\{h_x,h_y\}}(\vec{r})$ and $\hat{\omega}_{y\{h_x,h_y\}}(\vec{r})$, respectively. In the sequel we will consider the estimation of $\omega_x(x, y) = \partial\phi(x, y)/\partial x$.

The algorithm requires the specification of six parameters (κ', $\Delta\kappa'$, κ'', $\Delta\kappa''$, κ''', $\Delta\kappa'''$) as discussed in details in [13]. Appendix B gives the range of possible values of these parameters, as well.

Algorithm can be divided into three parts.

A. First we will perform optimization along the "diagonal" of the window set, i.e., $h_x = h_y$. The next window from the set H^2, larger than the optimal one, has an increased bias. However, it is not known which of the components, $\partial^3\phi(x, y)/\partial x^3$ or $\partial^3\phi(x, y)/\partial x\partial y^2$, increases the bias.

B. Thus, in the second part we will try to decrease the variance by increasing h_x, while keeping window width h_y the same.

C. In the third part of the algorithm we try to decrease the variance by increasing the window width along the y-coordinate.

The algorithm can be described, in more details, as follows.

A1. Consider a point $\vec{r}$. The initial guess is set to be the one produced with the smallest window from the set H^2 (with the smallest bias): $\hat{\omega}_x^0(\vec{r}) = \hat{\omega}_{x\{h^{(1)},h^{(1)}\}}(\vec{r})$.

A2. Calculate $\hat{\omega}_{x\{h^{(p)},h^{(p)}\}}$ with other window widths from H^2. The LF estimate is obtained with the largest window $h^{(p)} \times h^{(p)}$ that satisfies the following inequality:

$$|\hat{\omega}_{x\{h^{(p)},h^{(p)}\}}(\vec{r}) - \hat{\omega}_{x\{h^{(p+1)},h^{(p+1)}\}}(\vec{r})|$$
$$\leq (\kappa' + \Delta\kappa')[\hat{\sigma}(h^{(p)}, h^{(p)}) + \hat{\sigma}(h^{(p+1)}, h^{(p+1)})]. \quad (12)$$

A3. If inequality (12) is satisfied for all windows from the set H^2, and the widest window in the set $h^{(p+1)} = h^{(J)}$ is reached on the right-hand side of (12), then $\hat{\omega}_{x\{h^{(p)},h^{(p)}\}}(\vec{r})$ is the estimate of the LF component along the x-axis, $\hat{\omega}_x(\vec{r}) = \hat{\omega}_{x\{h^{(p)},h^{(p)}\}}(\vec{r})$. The remaining two parts of the algorithm can be skipped. If (12) is not satisfied for a window $h^{(p)}$ within H, then go to part B of the algorithm.

B. In the second part of the algorithm, the window size along the y-coordinate is held at the value $h^{(p)}$, while the x-coordinate size of the window is increased. The estimate of this LF component is produced with the largest window $h^{(q)} \times h^{(p)}$, $h^{(q)} > h^{(p)}$, that satisfies the following inequality:

$$|\hat{\omega}_{x\{h^{(q)},h^{(p)}\}}(\vec{r}) - \hat{\omega}_{x\{h^{(q+1)},h^{(p)}\}}(\vec{r})|$$
$$\leq (\kappa'' + \Delta\kappa'')[\hat{\sigma}(h^{(q)}, h^{(p)}) + \hat{\sigma}(h^{(q+1)}, h^{(p)})], \quad (13)$$

as $\hat{\omega}_x(\vec{r}) = \hat{\omega}_{x\{h^{(q)},h^{(p)}\}}(\vec{r})$. In this case the third step of the algorithm can be skipped. Otherwise, go to part C of the algorithm.

C. The LF component estimate is the one produced with the largest window $h^{(p)} \times h^{(t)}$, $h^{(t)} > h^{(p)}$, satisfying:

$$|\hat{\omega}_{x\{h^{(p)},h^{(t)}\}}(\vec{r}) - \hat{\omega}_{x\{h^{(p)},h^{(t+1)}\}}(\vec{r})|$$

$$\leq (\kappa''' + \Delta\kappa''')[\hat{\sigma}(h^{(p)}, h^{(t)}) + \hat{\sigma}(h^{(p)}, h^{(t+1)})]. \quad (14)$$

The estimate of this LF component is then $\hat{\omega}_x(\vec{r}) = \hat{\omega}_{x\{h^{(p)},h^{(t)}\}}(\vec{r})$. In the case that there is no window satisfying (14) and $h^{(t)} > h^{(p)}$, the LF component estimate is the output of part A of the algorithm.

Comments:

1. An alternative way for optimization could be: a) The initial estimate is produced with the narrowest window size; b) Optimization is performed by keeping the window along y-axis the same, and by increasing the window size along the x-axis; c) Optimize the y-coordinate width by holding the window size along the x-coordinate at the value obtained from step b). A similar form of the successive optimization can be applied to the LF estimation of md-signals, for $Q > 2$.

2. The MSE of the WD based LF estimation in the case of 2D signals can be written as, (10),

$$MSE_x = (E\{\Delta\omega_x\})^2 + var(\Delta\omega_x)$$

$$= \left(b_{xx}h_x^2 + b_{xy}h_y^2\right)^2 + c \cdot h_x^{-3}h_y^{-1}. \quad (15)$$

3. The standard deviation can be expressed as

$$\sigma(h_x, h_y) = \sqrt{\frac{\sigma_\nu^2(2A^2 + \sigma_\nu^2)\rho F_{(2,2)}F_{(0,2)}}{A^4 F_{(2,1)}^2 h_x^3 h_y}}.$$

For its estimation, one needs to estimate the variance of noise σ_ν^2 and the signal amplitude A^2. It can be done by a straightforward extension of 1D case presented in [12].

IV. Simplified Algorithm

If we consider relations (10) or (15), we can see that one window dimension dominantly influences the estimate variance of the LF for the considered component. Thus, by taking the window width along the y-axis equal to one sample, the bias term influenced by $\partial^3\phi(x,y)/\partial x^2\partial y$ can be neglected, while the variance decreases with a cube of the window width along the x-coordinate. The algorithm for $\omega_x(x,y)$ component of the LF, based on the 1D WD, is given in the sequel.

1. Consider the WD of a 1D signal:

$$WD_h(x, y; \omega_x) = \sum\nolimits_{n_x} w_h(n_x\rho)\times$$

$$s(x + n_x\rho, y)s^*(x - n_x\rho, y)e^{-j2\omega_x n_x\rho}. \quad (16)$$

2. Estimate $\hat{\omega}_x(x,y)$ as:

$$\hat{\omega}_{x\{h\}}(x,y) = \arg[\max_{\omega} WD_h(x, y; \omega_x)],\ h \in H. \quad (17)$$

3. The initial estimate is obtained by using the narrowest window from the set H, $\hat{\omega}_{x\{h^{(1)}\}}(x,y)$.

4. The optimal window is the widest $h^{(p)}$ from the set H that satisfies the inequality:

$$|\hat{\omega}_{x\{h^{(p)}\}}(x,y) - \hat{\omega}_{x\{h^{(p+1)}\}}(x,y)|$$

$$\leq (\kappa + \Delta\kappa)[\hat{\sigma}(h^{(p)}) + \hat{\sigma}(h^{(p+1)})]. \quad (18)$$

Values κ and $\Delta\kappa$ have to be chosen according to Appendix B, with $m = 4$ and $n = 3$. Estimate of the LF component is $\hat{\omega}_x(x,y) = \hat{\omega}_{x\{h^{(p)}\}}(x,y)$.

5. Estimation of the LF component along the y-axis, $\hat{\omega}_y(x,y)$, can be done similarly.

6. In this case the expected algorithm accuracy is lower than in the previous algorithm. Fortunately, we can use the property that $\partial\omega_x(x,y)/\partial y$ and $\partial\omega_y(x,y)/\partial x$ should be equal, since both of them reduce to $\partial^2\phi(x,y)/\partial y\partial x$. Therefore, if the following inequality:

$$|\partial\hat{\omega}_x(x,y)/\partial y - \partial\hat{\omega}_y(x,y)/\partial x| > \gamma, \quad (19)$$

holds for the partial derivatives, where γ is the accuracy threshold, then we can conclude that it is necessary to improve the accuracy for that point (x,y) by using the previous algorithm.

Accuracy of both of these algorithm forms is of the same order of magnitude for a wide range of the signal-to-noise (SNR) ratio, as it will be shown in the next section. The reason for this can be explained as follows. The simplified algorithm produces only slightly worse results after the first five steps than the previous one, since the basic algorithm can only slightly decrease variance by increasing the window size along the other coordinate. Also, the basic algorithm can include larger bias in the initial estimation than the simplified version. Points where accuracy can be improved are controlled by (19). Application of the 2D optimization algorithm, only in these points, produces the same order of accuracy as the algorithm with the 2D optimization for the entire plane. Note that calculation of the 1D WD needs approximately N times less multiplications and additions than the 2D WD.

V. Numerical Example

Consider the noisy signal:

$$s(x,y) = Ae^{j64\pi|x^2+y^2-0.1|} + \nu(x,y), \quad (20)$$

where $\nu(x,y)$, is the AWGN. The SNR was $A^2/\sigma^2 = 15$dB. Signal is considered within the interval $[x,y] \in [-1,1] \times [-1,1]$. Sampling intervals along both coordinates are $\rho = 1/256$. The true LF value components are:

$$\begin{aligned} &\omega_x(x,y) = \partial\phi(x,y)/\partial x \\ &= 128\pi x \mathrm{sgn}(x^2+y^2-0.1), \\ &\omega_y(x,y) = \partial\phi(x,y)/\partial y \\ &= 128\pi y \mathrm{sgn}(x^2+y^2-0.1). \end{aligned} \quad (21)$$

The windows used in the algorithm are rectangular, with size

$$H^2 = H \times H,$$

$$H = [4\rho, 8\rho, 16\rho, 32\rho, 64\rho, 128\rho]. \quad (22)$$

Several typical estimates of $\omega_x(x,y)$, for windows from set (22), as well as the estimation errors, are shown in Fig.1. The MSE of the LF component estimate obtained with the constant window size are shown in Table I. The smallest MSE is obtained for $4\rho \times 8\rho$, and it is equal to 114.0 for the component $\omega_x(x,y)$. For the component $\omega_y(x,y)$, the optimal window is $8\rho \times 4\rho$, with $MSE = 110.6$. The adaptive algorithm presented in Section III is applied for estimates obtained by using windows from set (22). Values of the parameters used in the algorithm are $\kappa' + \Delta\kappa' = 10$, $\kappa'' + \Delta\kappa'' = 4$, and $\kappa''' + \Delta\kappa''' = 2$. Estimate of the LF component obtained with the adaptive algorithm is shown in Fig.2. Algorithm produces the estimation errors 104.8 and 110.2. These values are smaller than those produced with any window with constant size, from the set H^2. In addition, the constant size window producing the best performance is not known in advance. The simplified algorithm from Section IV has been applied to this signal, as well. Windows are chosen from the set H. The MSE values, obtained with 1D WD based LF estimator, are given in Table I, last row. In both cases, the smallest MSE is obtained by the window of length 8ρ. It is 132.6 for the component $\omega_x(x,y)$, while for the component $\omega_y(x,y)$ its value is 137.4. For the simplified algorithm we have used $\kappa + \Delta\kappa = 4$. The accuracy threshold in (19) was $\gamma = 2\pi$ (two frequency steps). The estimate of $\partial\hat{\omega}_x(x,y)/\partial y$ is obtained by using:

$$\begin{aligned} \frac{\partial\hat{\omega}_x(x,y)}{\partial y} \approx \frac{1}{4\rho}[\hat{\omega}_x(x,y+\rho) + \frac{1}{2}\hat{\omega}_x(x+\rho,y+\rho) \\ +\frac{1}{2}\hat{\omega}_x(x-\rho,y+\rho) - \hat{\omega}_x(x,y-\rho) \\ -\frac{1}{2}\hat{\omega}_x(x+\rho,y-\rho) - \frac{1}{2}\hat{\omega}_x(x-\rho,y-\rho)], \end{aligned} \quad (23)$$

while $\partial\hat{\omega}_y(x,y)/\partial x$ is estimated similarly. Accuracy was not satisfactory in 13.7% of the considered points (19). In these points the accuracy was improved by increasing the window size along the other direction, and by applying the algorithm described in Section III. The obtained MSEs were 106.2 and 111.0, i.e., of the same order of magnitude as in the complete 2D optimization algorithm. The estimate of the LF component with the simplified algorithm is shown in Fig.3. Comparison of the results obtained with these two algorithms is given in Table II for various SNRs (6dB, 9dB, 12dB, and 15dB). It can be seen that for moderate SNR values both algorithms are accurate. However, this is not the case for lower SNR. Then, the initial estimate produced with the narrowest window width for high noise can result in the WD maxima outside the signal auto-term. The assumptions ("small noise") under which the presented algorithms are developed are no longer satisfied. Consideration of this effect is outside the scope of this paper.

VI. Conclusion

Adaptive algorithm for the LF estimation based on the md-WD is proposed. This algo-

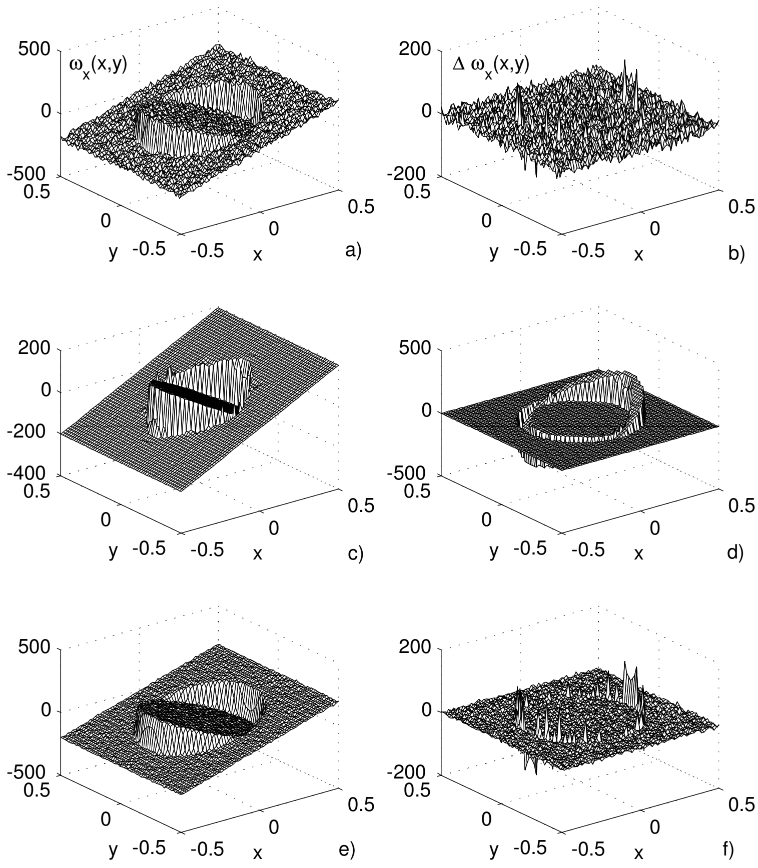

Fig. 1. Estimates of $\omega_x(t,\omega)$ obtained with the 2D WD and constant-size windows: a) Estimate for 4x4 window; b) Estimation error for 4x4 window; c) Estimate for 128x128 window; d) Estimation error for 128x128 window; e) Estimate for 4x8 window; f) Estimation error for 4x8 window.

rithm produces the MSE smaller than the LF estimate obtained with any constant window size. The values of the algorithm parameters are analyzed. The simplified algorithm version with the initial LF estimate obtained by using the 1D WD is presented. The proposed algorithms are compared with the nonadaptive LF estimation procedure based on the WD maxima. Numerical example confirms the theory.

Appendix A: Error Model, Variance and Bias

Consider the first WD partial derivatives around the exact LF value $\nabla WD_{\vec{h}}(\vec{r},\vec{\omega})$, whose components are:

$$\frac{\partial WD_{\vec{h}}(\vec{r},\vec{\omega})}{\partial \omega_i} = \sum_{\vec{n}} w_{\vec{h}}(\vec{n}\rho)\times$$

$$s(\vec{r}+\vec{n}\rho)s^*(\vec{r}-\vec{n}\rho)(-j2n_i\rho)e^{-j2\vec{\omega}\vec{n}\rho}. \quad (24)$$

In an ideal case, $\nabla WD_{\vec{h}}(\vec{r},\vec{\omega}) = 0$ on the LF. However, due to noise and to the fact that the WD ideally concentrates only the signal with a linear LF, this will not be true for a general signal, even without noise.

Assumptions:

1. By expanding the signal phase into the Taylor series, assuming limited partial derivatives $\phi(\vec{r}+\vec{n}\rho) = \sum_{i=0}^{\infty}(\vec{n}\rho\nabla)^i\phi(\vec{r})/i!$, the deterministic component in (24) can be written as $m(\vec{r}+\vec{n}\rho)m^*(\vec{r}-\vec{n}\rho) = |A|^2\exp(j\Phi(\vec{r},\vec{n}\rho))$, where $\Phi(\vec{r},\vec{n}\rho) = \sum_{j=1}^{Q} 2n_j\rho\partial\phi(\vec{r})/\partial r_j + \Delta\phi(\vec{r},\vec{n}\rho)$. Beside the first component, $\Phi(\vec{r},\vec{n}\rho)$ contains a component influenced by higher or-

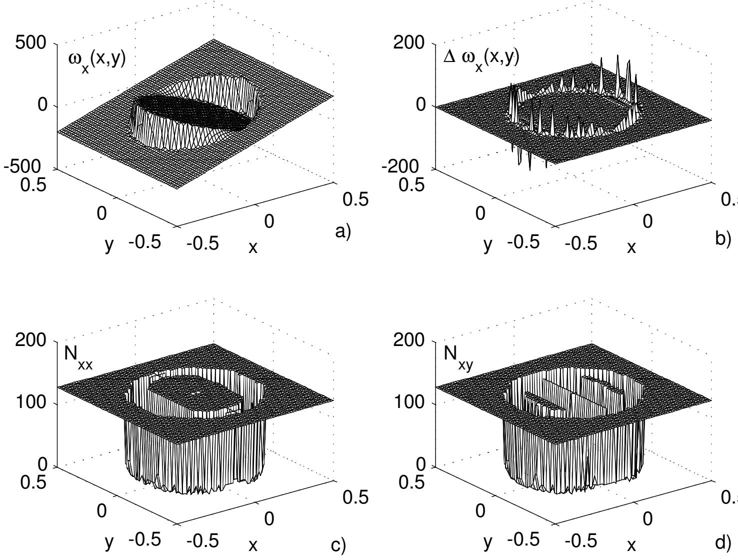

Fig. 2. Estimates of $\omega_x(x,y)$ obtained with the adaptive algorithm: a) Estimate of $\omega_x(x,y)$; b) Estimation error; c) Optimal window size along x−coordinate $\hat{h}_x(x,y)$; d) Optimal window size along y−coordinate $\hat{h}_y(x,y)$.

TABLE I
MSE OF THE LF ESTIMATE OF THE X AND Y COMPONENT BY USING CONSTANT-SIZE WINDOWS.

MSE_x / MSE_y	4ρ	8ρ	16ρ	32ρ	64ρ	128ρ
4ρ	236.9	114.0	194.7	698.9	1601.3	3320.2
	226.7	155.8	182.5	350.8	783.5	1441.3
8ρ	160.7	122.9	200.8	711.0	1614.6	3303.9
	110.6	120.7	175.2	347.6	772.3	1510.7
16ρ	186.3	176.0	231.4	712.3	1603.7	3337.7
	194.7	200.7	231.5	349.0	736.3	1559.4
32ρ	353.0	351.0	341.4	725.5	1533.1	3039.1
	706.3	707.7	690.4	723.0	779.3	1543.0
64ρ	781.9	773.4	746.4	782.5	1129.7	1892.9
	1600.9	1615.2	1599.8	1536.5	1119.6	1289.7
128ρ	1408.5	1511.6	1560.3	1546.7	1302.6	2645.5
	3259.6	3306.6	3341.0	3048.6	1904.4	2630.1
1D WD	573.6	132.6	196.6	701.4	1590.5	3171.0
	605.9	137.4	197.3	679.0	1592.8	3164.6

der phase derivatives:

$$\Delta\phi(\vec{r},\vec{n}\rho) = 2\sum_{i=1}^{\infty}\frac{(\vec{n}\rho\nabla)^{2i+1}}{(2i+1)!}\phi(\vec{r}) = \frac{(\vec{n}\rho\nabla)^3}{6}[\phi(\vec{r}+\vec{r}_1)+\phi(\vec{r}-\vec{r}_1)], \quad (25)$$

where components of $\vec{r}_1$ are within the interval $[0,\rho]$. In our analysis we assumed that higher

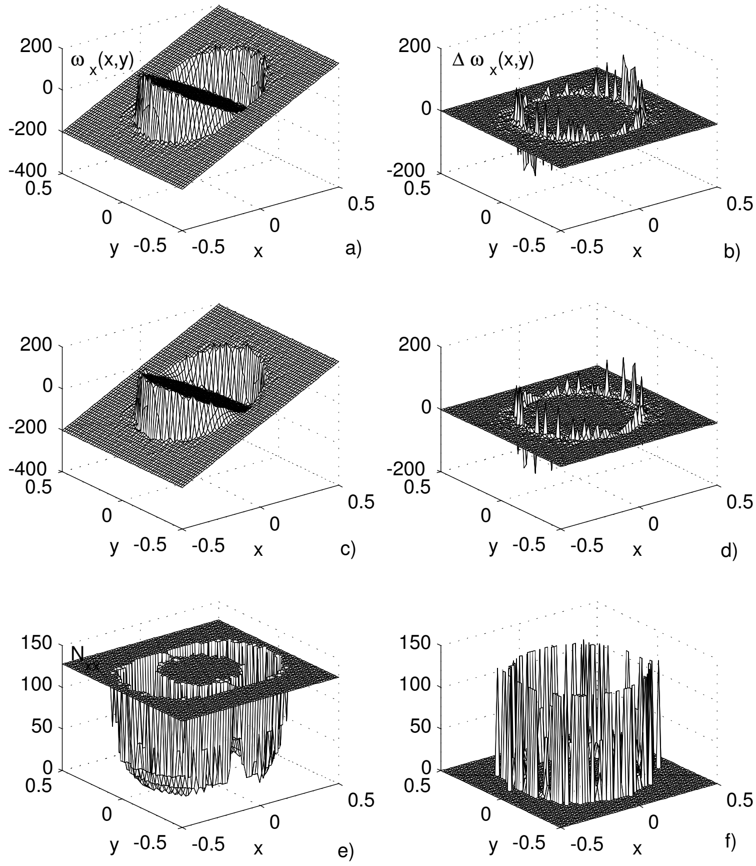

Fig. 3. Estimates of $\omega_x(x,y)$ obtained with the simplified algorithm: a) Estimate with the optimization of h_x window only; b) Estimation error after optimization of h_x window only; c) Estimate of $\omega_x(x,y)$ with the accuracy improvement; d) Estimation error; e) Optimal window size along $x-$coordinate $\hat{h}_x(x,y)$; f) Optimal window size along $y-$coordinate $\hat{h}_y(x,y)$.

TABLE II

MSE for various SNRs: Simp - simplified algorithm, Basic - basic algorithm.

SNR		6dB	9dB	12dB	15dB
MSE$_x$	Simp	675.7	130.9	114.3	106.2
	Basic	735.8	137.4	114.1	104.8
MSE$_y$	Simp	800.5	119.4	113.1	111.0
	Basic	821.5	124.2	112.8	110.2

order derivatives in $\Delta\phi(\vec{r},\vec{n}\rho)$ can be neglected within the considered window, and that it can be written as $\Delta\phi(\vec{r},\vec{n}\rho) = (\vec{n}\rho\nabla)^3\phi(\vec{r})/3$. This assumption is crucial for the analytic derivation of the bias expression, but not for the algorithm application itself, since it does not require the explicit knowledge about the bias.

2. The variance is derived under the assumption that the noise is small. "Small noise assumption" can be verbally defined as the noise level whose influence is not so strong that the points completely outside the auto-term can be detected as the WD maxima.

By using a linear approximation of $\partial WD_{\vec{h}}(\vec{r},\vec{\omega})/\partial\omega_i$ around the true LF value and following the derivations from [12], we get the variance and the bias of the LF estimator. The variance assumes the form:

$$var(\Delta\omega_i) = \frac{\sigma_\nu^2(2A^2+\sigma_\nu^2)\sum_{\vec{n}} n_i^2\rho^2 w_{\vec{h}}^2(\vec{n}\rho)}{A^4[\sum_{\vec{n}} n_i^2\rho^2 w_{\vec{h}}(\vec{n}\rho)]^2}. \qquad (26)$$

From (26) follows that the estimation variance is of the form (5), [12]. The estimation bias is:

$$E\{\Delta\omega_i\} = \frac{2\sum_{\vec{n}} w_{\vec{h}}(\vec{n}\rho)(2n_i\rho)\sum_{j=1}^{\infty}\frac{(\vec{n}\rho\nabla)^{2j+1}}{(2j+1)!}\phi(\vec{r})}{\sum_{\vec{n}} w_{\vec{h}}(\vec{n}\rho)(2n_i\rho)^2}. \qquad (27)$$

According to *Assumption 1,* and following derivations from [12], expression for bias (8) follows (27).

The optimal window size $\vec{h}_{opt}$ can be determined from the system of equations produced from the partial derivatives of the MSE_i (10). Its coordinates read:

$$h_l = h_i\sqrt{\frac{b_{i,i}}{3b_{l,i}}}|_{\vec{h}=\vec{h}_{opt}}, l \neq i,$$

$$\text{with } h_i = \sqrt[6+Q]{c\frac{3^{(Q+3)/2}}{4(Q+2)}\frac{\prod_{j\neq i}^{Q} b_{j,i}^{1/2}}{b_{i,i}^{(Q+3)/2}}}. \qquad (28)$$

However, since coefficients $b_{j,i}$, given by (9), contain unknown third order derivatives of the phase function, this optimal window size cannot be directly used in the LF estimation.

Appendix B: Algorithm Based on Intersection of Confidence Intervals

An adaptive algorithm for the IF estimation based on the ICI rule has been developed in [12]. It can be used when the MSE, in terms of parameter h, is of the form $MSE = A/h^m + B(t)h^n$, where $\sigma^2(h) = A/h^m$ is the estimation variance, while the second term represents the squared bias, $bias^2(t,h) = B(t)h^n$. In the IF estimation case, the parameter h was the window length. For implementation of the FFT algorithms, it is suitable that the windows have dyadic lengths, $H = \{h^{(s)}|h^{(s)} = 2h^{(s-1)}, s = 1,2,...,J\}$. The minimal MSE is produced with $h_{opt}(t) = (mA/nB(t))^{1/(m+n)}$. Its value is:

$$MSE(h_{opt}) = A\,(mA/(nB(t)))^{-m/(m+n)}\,\frac{n+m}{n}. \qquad (29)$$

In the worst case, the optimal window width is the geometrical mean of two adjacent window lengths $h_{opt}(t) = \sqrt{2}h^{(s)}$. Then the MSE, for the window width $h^{(s)}$ that can be selected as optimal in the algorithm, is $MSE(h^{(s)}) = MSE(h_{opt})(n2^{m/2} + m2^{-n/2})/(n+m)$. Therefore, even in the worst case, error produced by a reduced set H is of the same order of magnitude as the one produced with the optimal window size. It means that satisfactory results can be obtained by considering a set with relatively small number of the parameter h values.

Let $f(t)$ be a true value of the estimated variable. The estimates performed with different parameters from the set H are denoted by $\hat{f}_{h^{(s)}}(t)$, $s = 1,2,...,J$. For the estimate $\hat{f}_{h^{(s)}}(t)$, as for any other random variable, the following inequality can be written:

$$|f(t)-(\hat{f}_{h^{(s)}}(t)-bias(t,h^{(s)}))| \leq \kappa\sigma(h^{(s)}), \qquad (30)$$

with probability $P(\kappa)$. For the algorithm it is essential that κ is such that $P(\kappa)$ is close to 1.

The confidence interval around the estimate is:

$$D_s \in [\hat{f}_{h^{(s)}}(t) - (\kappa+\Delta\kappa)\sigma(h^{(s)}), \hat{f}_{h^{(s)}}(t) + (\kappa+\Delta\kappa)\sigma(h^{(s)})]. \qquad (31)$$

The values of κ and $\Delta\kappa$ should satisfy [13]:

$$\Delta\kappa = \sqrt{\frac{m}{n}2^{m/2}\frac{2^{n/2}-1}{2^{m/2}+1}}, \qquad (32)$$

$$\kappa < \sqrt{\frac{m}{n}} 2^{m/2-1} \frac{2^{n/2}-1}{2^{m/2}+1} (2^{(m+n)/2} - 1). \quad (33)$$

The adaptive (close to optimal) value of h is determined as the largest $h^{(s)} \in H$ where two consecutive confidence intervals intersect: $|\hat{f}_{h^{(s)}}(t) - \hat{f}_{h^{(s+1)}}(t)| \leq (\kappa + \Delta\kappa)[\sigma(h^{(s)}) + \sigma(h^{(s+1)})]$.

According to (32) and (33), the range of possible values of the basic algorithm parameters (Section III) is $\Delta\kappa' = 2.4$, $\kappa' < 18$ (for $m = 4$, $n = 4$), $\Delta\kappa'' = 1.9194$, $\kappa'' < 9.8993$ (for $m = 3$, $n = 4$), $\Delta\kappa''' = 0.8787$, and $\kappa''' < 2.0459$ (for $m = 1$, $n = 4$).

VII. Acknowledgment

This work is supported by the Volkswagen Stiftung, Federal Republic of Germany.

References

[1] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - Part I: Fundamentals," *Proc. IEEE*, Vol.80, pp.521-538, Apr. 1992.

[2] S. Barbarossa and O. Lemoine, "Analysis of nonlinear FM signals by pattern recognition of their time-frequency representations," *IEEE Sig. Proc. Let.*, Vol.3, pp.112-115, Mar. 1996.

[3] B. Barkat, "Instantaneous frequency estimation of nonlinear frequency-modulated signals in the presence of multiplicative and additive noise," *IEEE Trans. Sig. Proc.*, Vol.49, pp.2214-2222, Oct. 2001.

[4] H.K. Kwok and D.L. Jones, "Improved instantaneous frequency estimation using an adaptive short-time Fourier transform," *IEEE Trans. Sig. Proc.*, Vol. 48, pp. 2964-2972, Oct. 2000.

[5] F. Cakrak and P.J. Loughlin, "Multiwindow time-varying spectrum with instantaneous bandwidth and frequency constraints," *IEEE Trans. Sig. Proc.*, Vol.49, pp.1656-1666, Aug. 2001.

[6] Y.M. Zhu, R. Goutte and M.Amiel, "On the use two-dimensional Wigner-Ville distribution for texture segmentation," *Sig. Proc.*, Vol.30, pp.329-354, 1993.

[7] J. Hormigo and G. Cristobal, "High resolution spectral analysis of images using the pseudo Wigner distribution," *IEEE Trans. Sig. Proc.*, Vol.46, pp.1757-1763, June 1998.

[8] A. Jakobsson, S.L. Marple Jr. and P. Stoica, "Computationally efficient two-dimensional Capon spectrum analysis," *IEEE Trans. Sig. Proc.*, Vol.48, pp.2651-2661, Sept. 2000.

[9] B. McGuffin and B. Liu, "An efficient algorithm for two-dimensional autoregressive spectrum estimation," *IEEE Trans. ASSP*, Vol.37, pp.106-117, Jan. 1989.

[10] B. Friedlander and J.M. Francos, "An estimation algorithm for 2D polynomial-phase signals," *IEEE Trans. Im. Proc.*, Vol.5, pp.1084-1087, June 1996.

[11] S.R. DeGraaf, "SAR imaging via modern 2-D spectral estimation methods," *IEEE Trans. Im. Proc.*, Vol.7, pp.729-761, May 1998.

[12] V. Katkovnik and LJ. Stanković, "Instantaneous frequency estimation using the Wigner distribution with varying and data driven window length," *IEEE Trans. Sig. Proc.*, Vol.46, pp.2315-2325, Sept. 1998.

[13] LJ. Stanković and V. Katkovnik, "Algorithm for the instantaneous frequency estimation using time-frequency distributions with variable window width," *IEEE Sig. Proc. Let.*, Vol.5, pp.224-228, Sept. 1998.

[14] LJ. Stanković, "Adaptive instantaneous frequency estimation using TFDs," in *Time-frequency signal analysis and processing*, ed. B. Boashash, Elsevier, 2003, in print.