Motion Parameters Estimation by New Propagation Approach and Time-frequency Representations

Igor Djurović, Srdjan Stanković, Akira Ohsumi and Hiroshi Ijima

Abstract— **Estimation of time-varying motion parameters of moving objects is considered. The motion parameters are mapped into the instantaneous frequency (IF) of a frequency modulated (FM) signal. The IF estimation is obtained using the Wigner distribution (WD). Generalization in the case of multiple objects and reduced interference distributions is given. A new propagation approach is proposed for joint estimation of velocity and position of objects in a video-sequence. This approach produces better accuracy than the one based on the variable μ-propagation-based one.**

I. INTRODUCTION

Motion estimation in video-sequences is a corner stone in numerous applications. Especially, one of the important application fields is the compression of video-signals. An overview of techniques for the motion estimation used in the video-signal compression is presented in [1]-[3]. Detailed analysis of the optical flow estimation techniques is given in [4]. Tracking and estimating motion parameters of several moving objects in the sequence are the other important issues. Recently, several techniques based on the spectral analysis methods have been proposed for this purpose.

In this paper, the time-frequency (TF) analysis tools will be applied to the motion parameters estimation. The spectral analysis methods are one of the research directions in the motion parameters and optical flow estimation. Details of this research are published in [5]-[16]. Numerous motion estimation algorithms use projections of video-frames to coordinate axes, in order to simplify the analysis of a 3D video-sequence (2 spatial and 1 time coordinate) [8]-[9]. Synthetic images, ob-

tained by projection of a video-sequence on the coordinate axes, are used to produce the frequency modulated (FM) signals [5], [6]. The motion parameters can be obtained by spectral analysis of these FM signals. The constant μ-propagation is proposed in [5], [6] for mapping the synthetic images to the FM signals. This propagation approach stems from a high-resolution estimation of direction-of-arrival of impinging signals on the linear sensors array [5]. The Fourier transform can be used for estimation in the case of constant velocities [6], while for the varying velocities the TF methods for extraction of object's velocities are applied in [13]. The variable μ-propagation is proposed in [6], [7] for joint estimation of velocity and object position. In the case of fast varying objects, this propagation approach produces the signal with highly nonlinear phase. The TF representation-based estimators of object parameters in this case exhibit significant bias and sensitivity to the spatio-temporal noise influence. In this paper, we will propose a new propagation approach that produces smaller bias in the estimate as well as higher robustness to the noise influence.

The paper is organized as follows. An overview of the instantaneous frequency (IF) estimation based on the TF representations, in particular the Wigner distribution (WD) and the S-method (SM), is presented in Section II. Transformation of the video-sequence to the synthetic images is analyzed in Section III. Motion parameters estimation by using constant μ-propagation, variable μ-propagation and new approach is given in Section IV. Numerical examples are presented in Section V. Conclusion is given in Section VI.

II. TF representations as IF estimator

TF representations are widely used for estimation of signal parameters [17]. The most important parameter is the IF. IF estimators based on the positions of the TF representation maxima are commonly used in practice [18]-[20]. In this paper the TF representations are applied to the estimation of motion parameters of moving objects in a video-sequence, where the motion parameters are embedded in the IF of the analyzed signal. The IF of the FM signal $z(t) = A_x(t)\exp(j\phi(t))$ is defined as the first derivative of signal phase: $\omega(t) = \phi'(t)$. Assume that the signal is transformed to the 2D TF plane, by using the TF method: $z(t) \rightarrow TF_z(t, \omega)$. The TF representations are usually highly concentrated on the IF. Therefore, the IF can be estimated by using the position of the TF representation maxima [17]: $\hat{\omega}(t) = \arg\max_\omega TF_z(t, \omega)$. For one-component signal, the WD is a commonly used TF representation (and IF estimator):

$$WD_z(t, \omega) =$$

$$= \sum_{\tau=-N/2}^{N/2-1} z(t + \tau)z^*(t - \tau)e^{-j2\omega\tau}. \quad (1)$$

The WD is ideally concentrated on the IF for linear FM signal with the IF $\omega(t) = at + b$. However, the WD exhibits bias for the nonlinear FM signal. The bias and noise influence in the WD-based IF estimator are analyzed in detail in [20]. A brief review of these error sources is given in Appendix.

The main problem in the WD application are interferences appearing in the case of multicomponent signals. For instance, assume that there is a two-component signal: $z(t) = z_1(t) + z_2(t) = A_1(t)e^{j\phi_1(t)} + A_2(t)e^{j\phi_2(t)}$. Then the WD is given as: $WD_{z_x}(t, \omega) = WD_{z_1}(t, \omega) + WD_{z_2}(t, \omega) + C(t, \omega)$, where $C(t, \omega)$ is the cross-term:

$$C(t, \omega) =$$

$$= 2\,\mathrm{Re}\left\{\sum_{\tau=-N/2}^{N/2-1} z_1(t + \tau)z_2^*(t - \tau)e^{-j2\omega\tau}\right\}. \quad (2)$$

The cross-term can be of higher magnitude than each of the auto-terms $WD_{z_i}(t, \omega)$, $i = 1, 2$, and it can mask useful information about the signal components. This is the reason for introducing distributions with reduced interference [23], [24]. The S-method (SM) [24]:

$$SM_z(t, \omega)$$

$$= \sum_{m=-L}^{L} STFT_z(t, \omega+m\Delta)STFT_z^*(t, \omega-m\Delta), \quad (3)$$

where $2L + 1$ is the frequency window width, can be used to reduce interferences. Here, $STFT_z(t, \omega)$ is the short-time Fourier transform given as:

$$STFT_z(t, \omega) = \sum_\tau z(t + \tau)w^*(\tau)e^{-j\omega\tau}, \quad (4)$$

$w(t)$ is the window function and Δ is the frequency resolution.

The procedure for the IF estimation of an M-component signal can be summarized as follows [25], [26]:

(a) Determination of the SM maximum that corresponds to the IF of the first component:

$$\hat{\omega}^{(1)}(t) = \arg\max_\omega SM_z(t, \omega). \quad (5)$$

(b) Obtain a new TF representation $SM_z^{(1)}(t, \omega)$ from $SM_z(t, \omega)$ by letting all values be zero in the region around the determined maximum $[\hat{\omega}^{(1)}(t) - \delta, \hat{\omega}^{(1)}(t) + \delta]$. This TF representation will be used for velocity estimation of the next component:

$$\hat{\omega}^{(2)}(t) = \arg\max_\omega SM_z^{(1)}(t, \omega). \quad (6)$$

This step is repeated for each component $l = 3, ..., M$. If the number of components is not known in advance, it can be obtained by comparing the maxima of the modified $SM_z^{(m)}(t, \omega)$ with a threshold value. Threshold determination is discussed in [26]. If $\max_\omega SM_z^{(m)}(t, \omega) < \beta$, where β is the determined threshold, we can assume that there are no more moving object in the sequence, and the procedure can be stopped.

In the next section, we will explain the transformation of a 3D video-sequence to the FM signal. Then, information about object motion are embedded in the parameters of the FM signal. Motion parameters can be extracted from obtained FM signal by using the TF representation-based IF estimators.

III. MOTION MODEL

Consider a video-sequence consisting of a background $m(x, y)$ and moving object $s(x, y)$:

$$f(x, y, t) = m(x, y)$$
$$+ s(x - \varphi_x(t), y - \varphi_y(t)), \qquad (7)$$

where $(\varphi_x(t), \varphi_y(t))$ is position of the moving object in frame t. In order to simplify our analysis, in the sequel we will consider the single moving object case. The starting object position is given as $(\varphi_x(0), \varphi_y(0))$, while its velocity in the considered frame is given as a vector with coordinates $(d\varphi_x(t)/dt, d\varphi_y(t)/dt)$. In order to get information about motion parameters, consider projections of the 3D video-sequence onto the coordinate axes:

$$P_x(x, t) =$$
$$= \sum_y [m(x, y) + s(x - \varphi_x(t), y - \varphi_y(t))]$$
$$= m_x(x) + s_x(x - \varphi_x(t)), \qquad (8)$$
$$P_y(y, t) =$$
$$= \sum_x [m(x, y) + s(x - \varphi_x(t), y - \varphi_y(t))]$$
$$= m_y(y) + s_y(y - \varphi_y(t)), \qquad (9)$$

where $m_x(x)$, $m_y(y)$, $s_x(x)$ and $s_y(y)$ are projections of the background and object on each coordinate axis. In the sequel we will consider the estimation of motion parameters along the x-axis, since estimation along the y-axis can be done in a similar way. In order to avoid the background influence, the difference between two consecutive projections is introduced

$$\Pi_x(x, t) = P_x(x, t + 1) - P_x(x, t)$$
$$= s_x(x - \varphi_x(t + 1)) - s_x(x - \varphi_x(t)). \quad (10)$$

It has been derived in [13] that $\Pi_x(x, t)$ can be approximately written as

$$\Pi_x(x, t) \approx s_x(x - \varphi_x(t))$$
$$- \left(\sum_{k=1}^{\infty} \frac{1}{k!} \varphi_x^{(k)}(t) \right) \frac{\partial s_x(x - \varphi_x(t))}{\partial x}$$
$$- s_x(x - \varphi_x(t))$$
$$= - \left(\sum_{k=1}^{\infty} \frac{1}{k!} \varphi_x^{(k)}(t) \right) \Gamma(x - \varphi_x(t)), \qquad (11)$$

with $\Gamma(x) = \partial s_x(x)/\partial x$, under the following assumptions: (a) $\varphi_x(t)$ has continuous and bounded derivatives; (b) $s_x(x)$ is continuous function around $x - \varphi_x(t)$; (c) object velocity is relatively small. In this way, we obtain synthetic images $\Pi_x(x, t)$ and $\Pi_y(y, t)$, containing information on the object motion. Due to the spatio-temporal noise, extraction of the motion parameters, performed directly through these images, will be inaccurate. Instead, these synthetic images are mapped to FM signals, and then the velocity parameters are extracted using spectral analysis techniques.

Note that this simple differentiation procedure (10) can introduce errors in our algorithm for noisy sequences. More sophisticated differentiation relationships are derived in the optical flow estimation where the displacement is calculated for each pixel in the sequence. An overview of these techniques can be found in [4]. Here, we will use a simple filter of the projection function $P_x(x, t)$,

$$\tilde{P}_x(x, t) =$$
$$= \frac{1}{\sum_\xi \sum_\theta a(\xi, \theta)}$$
$$\times \sum_\xi \sum_\theta a(\xi, \theta) P_x(x - \xi, y - \theta), \qquad (12)$$

where $a(\xi, \theta)$ are coefficients of the considered filter. After performing filtering (10), $\tilde{\Pi}_x(x, t) = \tilde{P}_x(x, t + 1) - \tilde{P}_x(x, t)$ can be used in the previously described procedure. In this paper we used the simple filter $a(\xi, \theta) = \frac{1}{9}$ for $|\xi| \leq 1$ and $|\theta| \leq 1$ and $a(\xi, \theta) = 0$ elsewhere. In this case, $\tilde{\Pi}_x(x, t)$ can be calculated

without evaluation $\tilde{P}_x(x, t)$, i.e., $\tilde{\Pi}_x(x, t) = \frac{2}{9} \sum_{\xi=x-1}^{\xi=x+1} [P_x(\xi, t+2) - P_x(\xi, t-1)]$. More sophisticated filtering forms can be found in [4].

The next step in our procedure is obtaining the FM signal from synthetic image $\tilde{\Pi}_x(x, t)$ (or for $\tilde{\Pi}_y(y, t)$). This will be done applying propagation approach to the image $\tilde{\Pi}_x(x, t)$. The term propagation stems from the direction-of-arrival estimation of impinging signals on a sensor array network where this algorithm is used to produce highly accurate estimate [5], [22]. Recently, this approach has been applied in [6] to the velocity estimation. In the next section, the existing propagation functions are reviewed and the new function, that can be used to get more accurate results compared to the ones obtained by the previously defined functions is introduced.

IV. PROPAGATION

A. Review of existing procedures

Transformation of the synthetic image $\Pi_x(x, t)$ to an FM signal can be obtained by using the constant μ-propagation [5], [6] that can be written in the generalized form as

$$z_x(t) = \sum_x \Pi_x(x, t) e^{j\mu(x, t)x}$$

$$= A_x(t) \sum_x \Gamma(x - \varphi_x(t)) e^{j\mu(x, t)x}, \quad (13)$$

where $A_x(t) = -\sum_{k=1}^{\infty} (1/k!) \varphi_x^{(k)}(t)$ and $\mu(x, t)$ is a propagation function. Assume that the object velocity is small compared with the object size. Then, the following approximation holds $\Gamma(x - \phi_x(t)) \approx \delta(x - \varphi_x(t))$, i.e., signal (13) can be written as:

$$z_x(t) = A_x(t) e^{j\mu(\varphi_x(t), t)\varphi_x(t)}. \quad (14)$$

For constant μ-propagation, the propagation function exhibits a form $\mu(x, t) = \mu$ and the IF function of (14) is $\omega_x(t) = \mu d\varphi_x(t)/dt$. Note that the IF is proportional to the object velocity, so that the WD- and SM-based estimator can be applied to estimate it. Performance of the velocity estimators are the same as performance of the IF estimator described in Appendix. However, we should keep in mind

that the signal amplitude $A_x(t)$, is now proportional to the object velocity. Therefore, the estimator variance for a small velocity could be very high. Fortunately, such a small velocity case is not of our practical interest. Besides, the high variance can be avoided by introducing a proper threshold. An alternative approach for reducing errors in the small velocity case is discussed in [13]. Note that the IF of signal (14) has no information about the object position. The variable μ-propagation is developed in [6], [7] with propagation function $\mu(x, t) = \mu x$. For instance, for the constant velocity case of $\varphi_x(t) = v_x t + \phi_x(0)$, this propagation form produces the signal $z_x(t) \approx A_x(t) e^{j\mu(v_x t + \varphi_x(0))^2}$ with the WD concentrated along the IF:

$$WD_{zx}(t, \omega) \approx$$

$$\approx 4\pi A_x^2(t)\delta(\omega - 2\mu v_x^2 t - 2\mu v_x \varphi_x(0)). \quad (15)$$

Then the IF line $\omega_x(t) = 2\mu v_x^2 t + 2\mu v_x \varphi_x(0)$ determines both parameters v_x and $\varphi_x(0)$.

These parameters can be generally extracted from the IF estimate based on some simple interpolation procedure. For the object whose position in the x-coordinate is described by a polynomial, $\varphi_x(t) = \sum_{l=0}^{M} a_l t^l$, the phase of the FM signal with the variable μ−propagation is given by: $\phi_x(t) = \mu \left(\sum_{l=0}^{M} a_l t^l \right)^2$, and the IF is

$$\omega_x(t) = 2\mu \sum_{l=0}^{M} \sum_{k=0}^{M} (l+k) a_l a_k t^{l+k-1}. \quad (16)$$

The highest order in the IF function (16) is $(2M - 1)$, and the bias in the WD-based IF estimation can be very high because of the high IF nonlinearity, as stated in Section II [20]. Therefore, the procedure of interpolation of the IF estimate with $(2M-1)$-order polynomial model becomes quite difficult to handle, because the coefficients in the expression for $\omega_x(t)$ are no longer linear.

For the general motion model $\varphi_x(t)$ the phase function is given by $\phi_x(t) = \mu \varphi_x^2(t)$, so that the IF is obtained as $\omega_x(t) = 2\mu \varphi_x(t)[d\varphi_x(t)/dt]$. The information about the

object position can be extracted from estimates of the velocity and IF of signal obtained using the constant and variable μ-propagations, respectively.

B. New propagation approach

In our previous papers [13], [7] we used the variable μ-propagation for obtaining the joint information on the object position and velocity. However, in some applications this approach had produced unsatisfactory results due to the effects described in the previous subsection. Therefore, we investigated other possible propagation function should be used to handle this issue. The propagation function should satisfy the following properties: derivative of signal phase should contain information on both instantaneous velocity and object position relation between IF and motion parameters should be as simple as possible; order of polynomial in the signal phase should be as low as possible. We have found that all these requirements are satisfied by the time-varying propagation

$$\mu(x, t) = \mu t. \qquad (17)$$

Then, in the constant velocity case, the signal is expressed as:

$$z_x(t) \approx A_x(t)e^{j\mu(v_x t + \varphi_x(0))t}. \qquad (18)$$

The WD of (18) is concentrated ideally along the IF line:

$$WD_{zx}(t, \omega) \approx 4\pi A_x^2(t)\delta(\omega - 2\mu v_x t - \mu\varphi_x(0)). \qquad (19)$$

Thus, the motion parameters can be determined by the interpolation of this IF estimate. Consider now the case of the polynomial function. In this case, the time-varying propagation produces the FM signal with the phase function:

$$\phi_x(t) = \mu \sum_{l=0}^{M} a_l t^{l+1}. \qquad (20)$$

The IF of (20) is:

$$\omega_x(t) = \mu \sum_{l=0}^{M} a_l(l+1)t^l. \qquad (21)$$

Note here that the highest order is now M, so that in this case the bias can be reduced more than in the case of the variable μ-propagation. Owing to the linearity of the coefficients $\{a_l\}$ in (21), determination of these parameters through interpolation of $\omega_x(t)$ is rather easy, compared with the method with variable μ-propagation mentioned in the previous subsection. Furthermore, the lower order of polynomial in the signal's phase implies nothing but the robustness of the motion parameter estimation even though the noise influence is high. These facts clearly show that the proposed propagation approach can be used to produce more accurate motion parameter estimates than that preformed by the standard form of the variable μ-propagation.

In this case, for the general motion law the phase is given by $\phi_x(t) = \mu t\varphi_x(t)$, while the IF is $\omega_x(t) = \mu[\varphi_x(t) + td\varphi_x(t)/dt]$. The position can be estimated as $\hat{\varphi}_x(t) = \hat{\omega}_x(t)/\mu - t\hat{v}_x(t)$, where $\hat{\omega}_x(t)$ is the estimated IF of the signal produced by the new propagation approach, and $\hat{v}_x(t)$ is the velocity estimate obtained from the signal with the constant μ-propagation. It is interesting to discuss the choice of μ parameter in these propagations. It should be chosen as a trade-off according to two criteria: (a) larger μ means larger maximal possible detected value of the parameters; (b) larger μ means higher error caused by discretization. In order to explain these facts consider an example. Let the frame have $K \times K$ pixels and let maximal velocity of the object be $v$pixel/frame. Assume also that variations in the velocity are not extremely fast. Them the maximal value of the IFs obtained by the constant μ, variable μ and new propagation are $\mu_c v$, $\mu_v K v$ and $\mu_n[K/2 + Nv/2]$, where μ_c, μ_v and μ_n, are constants used in these propagations, respectively. Resolution in the estimation is, respectively, $\mu_c v/N$, $\mu_v K v/N$ and $\mu_n[K/2 + Nv/2]/N$. Assuming uniformly distributed error within one step, the mean squared error (MSE) caused by the discretization is $(\Delta\mu)^2/12$, where $\Delta\mu$ is resolution in the case of the corresponding propagation.. Therefore, maximal observed values of the parameters increase with the increase of propagation parameter μ but in the same time discretiza-
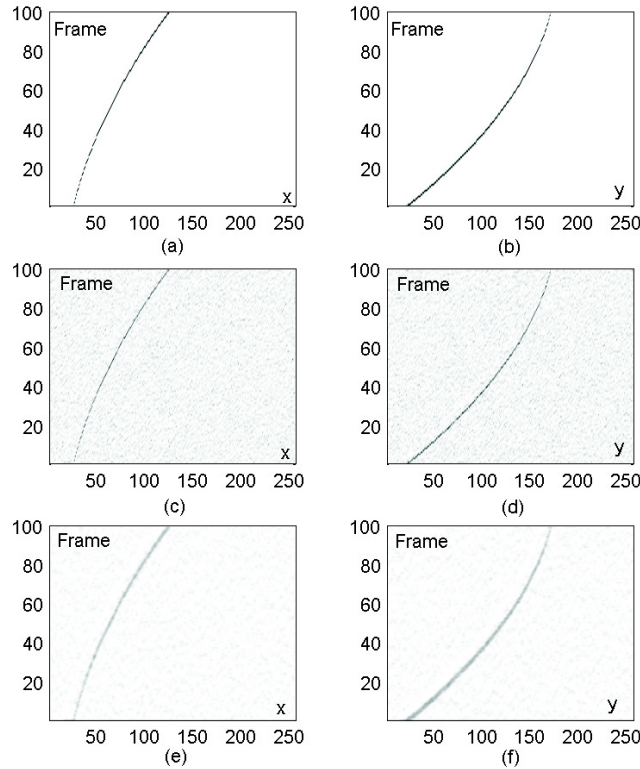
Fig. 1. Synthetic images $\Pi_x(x,t)$ (left column) and $\Pi_y(y,t)$ (right column): First row - Non-noisy sequence; Second row - Noisy sequence; Third row - Noisy sequence with smoothing of $P_x(x,t)$ and $P_y(y,t)$.

tion error increases. If maximal velocity is known in advance, the limits for propagation parameter can be easily set. Note that sampling rate in the signals obtained by propagation is $\Delta t = 1$ frame and that maximal frequency in the WD is $\pi/2$, while in the SM it is π. Therefore, in the case of the new propagation approach and the WD, the maximal value of the parameter μ is $\pi/[K + Nv_{\max}]$.

V. NUMERICAL EXAMPLES

Example 1: An artificial video-sequence with 100 frames of size 256×256 with white background is considered. The moving object is a black rectangle of size 10×10. Object positions can be described with:

$$\varphi_x(t) = \varphi_x(0) + v_x(0)t + a_xt^2/2, \qquad (22)$$

$$\varphi_y(t) = \varphi_y(0) + v_y(0)t + a_yt^2/2, \qquad (23)$$

where starting object position is $(\varphi_x(0), \varphi_y(0)) = (20, 15)$, while initial velocity is given as $(v_x(0), v_y(0)) = (0.5, 2.5)$. The object uniformly accelerates with $(a_x, a_y) = (0.01, -0.02)$. We considered two experiments: (a) Noiseless video-sequence; (b) Noisy video-sequence with a large white Gaussian noise $\nu(x, y, t)$ with standard deviation $\sigma = 40$. Note that the sequence has 255 grayscale levels where the white level is 255 and the black level is 0. Projections $\Pi_x(x, t)$ and $\Pi_y(y, t)$ for noise-free case are depicted in Figs. 1a,b, while for noisy case they are shown in Figs. 1c,d. From Figs. 1c,d one may notice a significant amount of noise. Synthetic images obtained after smoothing of projections by (12) are given in Figs. 1e,f. The used filtering func-
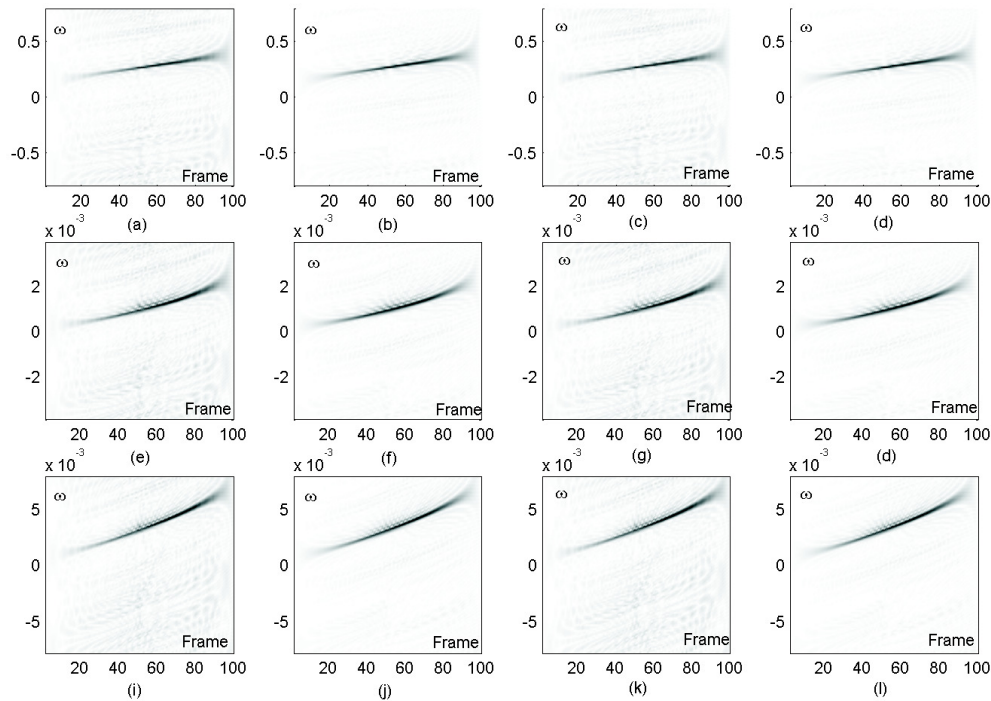
Fig. 2. WDs for non-noisy sequence: First row - WD obtained by using constant μ-propagation; Second row - WD obtained by variable μ-propagation; Third row - WD obtained by new propagation approach. First column - WD of signal $z_x(t)$ produced without filtering of projection; Second column - WD of signal $z_x(t)$ produced with filtering of projection; Third column - WD of signal $z_y(t)$ produced without filtering of projection; Fourth column - WD of signal $z_y(t)$ produced with filtering of projection.

tion was $a(\xi, \theta) = 1$ for $|\xi| \leq 1$ and $|\theta| \leq 1$ and 0 elsewhere. From Figs. 1e,f we know that not only the noise influence is attenuated but also that the pattern that representing motion becomes smoothed, as well.

The following propagation functions are applied to the synthetic images $\Pi_x(x, t)$ and $\Pi_y(y, t)$: the constant μ-propagation with $\mu = 0.5$, the variable μ-propagation with $\mu = \frac{1}{400}$, and the new propagation approach with $\mu = \frac{1}{200}$. The WDs are depicted in Figs. 2 and 3 for noise-free and noisy sequences, respectively. In both figures each row (from top to bottom) represents the WD of the signal obtained through the constant, variable μ-propagation, or by the new propagation; the first two columns depict $WD_{z_x}(t, \omega)$ with and without smoothing of $P_x(x, t)$ and the other two are

$WD_{z_y}(t, \omega)$. Corresponding IF estimates are depicted in Fig. 4. The results obtained by using smoothed projection are marked with thick lines while those produced with non-smoothed projection are shown with dashed lines. It can be seen that in the noise-free case (first two columns) all these estimates perform similarly. However, there are some errors at boundaries due to limited amount of data. In the noisy sequence case (last two columns) better behavior is shown by the smoothed projection. Also, it can be seen that the variable μ-propagation is more sensitive to the noise influence than the new approach. The reason for that is the inner interferences appearing in the non-linear FM signal case increase the bias in the estimation and in the same time decrease the maxima of the WD. Then, the errors caused by

TABLE I

MSE IN dB FOR VELOCITY ESTIMATION FOR VARIOUS
NOISE AMOUNTS. INDECES DENOTE: x, y -
CORRESPONDING COORDINATE; 1 - WITHOUT FILTERING
OF PROJECTION; 2 -WITH FILTERING OF PROJECTION.

σ	MSE_{x1}	MSE_{x2}	MSE_{y1}	MSE_{y2}
0	-36.02	-35.90	-37.69	-37.22
5.1	-26.33	-35.92	-37.35	-37.27
10.2	-15.12	-35.95	-23.57	-37.18
15.3	-12.35	-35.76	-16.11	-37.22
20.4	-8.60	-35.79	-10.93	-37.09
25.5	-5.72	-24.58	-8.40	-37.23
30.6	-4.43	-24.25	-6.91	-30.17
35.7	-3.98	-15.86	-5.23	-22.75
40.8	-3.53	-12.53	-4.17	-20.70
45.9	-3.64	-10.37	-3.20	-14.10
51	-3.48	-10.08	-2.43	-11.37

the high noise become larger in the case of the variable μ-propagation than in the case of the new propagation approach. This is an important advantage of the proposed approach. The MSE for the velocity estimation with respect to the different noise amounts is presented in Table I. Note that the first and the last ten samples are avoided in the calculation. The numerical results of the motion parameter estimation via variable μ- and new propagation methods are presented in Tables II and III. Results are obtained by using the Monte Carlo simulation with 100 trials. Again the first and the last ten frames are omitted. It can be seen that the new propagation approach outperforms the variable μ-propagation.

Example 2: The real video-sequence is considered, representing a man who moving from the upper right corner toward the left end of scene. The sequence has 220 frames of the size 240×320. Frames $24k$, $k = 1, 2, ..., 9$, are depicted in Fig. 5. We considered only the motion along the x-coordinate. The distributions obtained from the SM with constant μ-propagation with $\mu = 0.5$ and new propagation approach with $\mu = \frac{1}{800}$ are illustrated in Fig. 6 as well as the corresponding IF estimates. The position estimate is marked by thick line in Fig. 7. Object position is estimated in the following way. The constant

μ-propagation is used to produce estimate of the velocity $\hat{v}_x(t)$, while the time-varying propagation produces estimate of $\xi_x(t) = tv_x(t) + \phi_x(t)$. The object position can be estimated as $\hat{\phi}_x(t) = \hat{\xi}_x(t) - t\hat{v}_x(t)$. The 'true' value of the object position was calculated manually point by point, as the x-coordinate of the object's head. It can be seen that the accuracy is high, except in the region of the small x-component of velocity (first 75 frames).

Example 3: Here, we consider an artificial video-sequence with two moving objects. Objects' coordinates can be described with:

$$\varphi_{xi}(t) = \varphi_{xi}(0) + v_{xi}(0)t + a_{xi}t^2/2, \quad (24)$$

$$\varphi_{yi}(t) = \varphi_{yi}(0) + v_{yi}(0)t + a_{yi}t^2/2, \quad i = 1, 2, \quad (25)$$

with the following parameters $(\varphi_{x1}(0), v_{x1}(0), a_{x1}, \varphi_{y1}(0), v_{y1}(0), a_{y1}) = (15, 2.5, -0.2, 20, 0.5, 0.01)$ and $(\varphi_{x2}(0), v_{x2}(0), a_{x2}, \varphi_{y2}(0), v_{y2}(0), a_{y2}) = (230, -3, 0.018, 220, 0.5, -0.03)$. Objects' size is 10×10. The sequence has 100 frames. We perform the constant μ-propagation of the considered sequence and the new propagation approach. The spectrograms and SMs of signals obtained by using these two propagation approaches are given in Fig. 8. It can be seen that the spectrogram produces components spread over the TF plane, while the SM gives highly concentrated components. The IF estimate is performed by using the SMs. The velocity estimates obtained from the signal produced by the constant μ-propagation are given in Figs. 9a and b. The IFs of signals obtained by using the new propagation approach contain information on both object velocity and position. Therefore, by combining information from the SMs of signal produced by the constant μ- and new propagation, we obtain the estimation of objects' position depicted in Figs. 9c and d. It can be seen that estimation is quite accurate, excluding the region of very small velocities. Note that, in this case, both propagation approaches produce TF components that do not overlap in the TF plane. Some sophisticated procedure for segmentation of the TF plane is required to produce separate estimation of particular object parameters in the case of overlapping TF components.
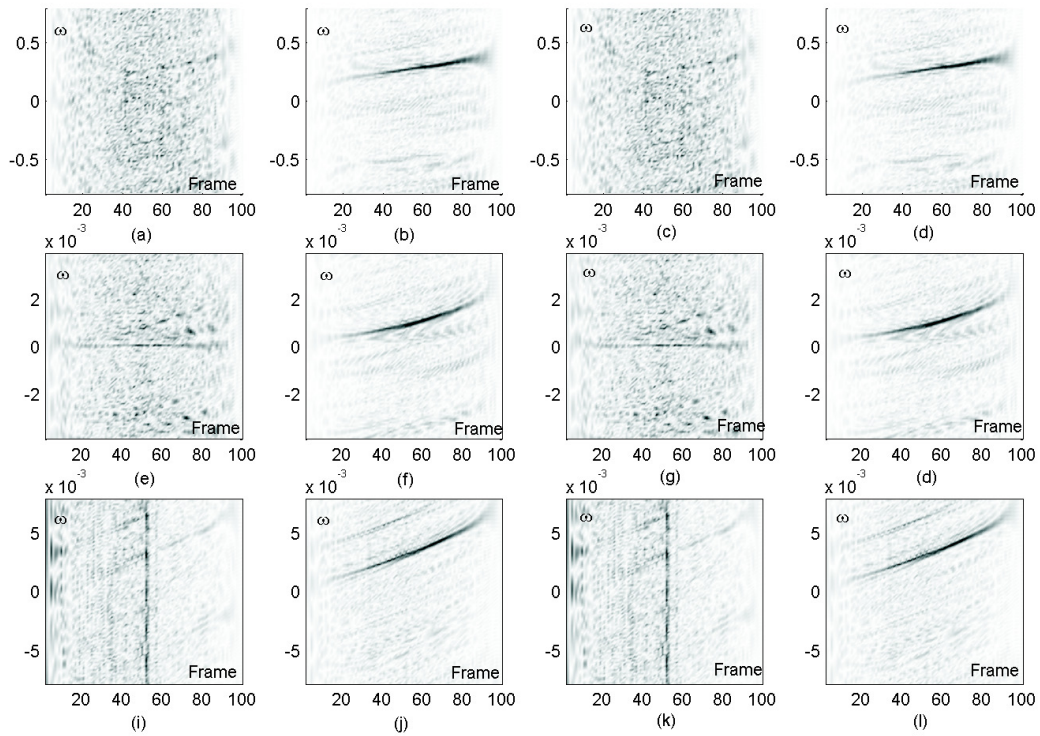
Fig. 3. WDs for noisy sequence: First row - WD obtained by using constant μ-propagation; Second row - WD obtained by variable μ-propagation; Third row - WD obtained by new propagation approach. First column - WD of signal $z_x(t)$ produced without filtering of projection; Second column - WD of signal $z_x(t)$ produced with filtering of projection; Third column - WD of signal $z_y(t)$ produced without filtering of projection; Fourth column - WD of signal $z_y(t)$ produced with filtering of projection.

TABLE II

Estimation of the motion parameters by using the variable μ-propagation. NF - without filtering of projection; F - with filtering.

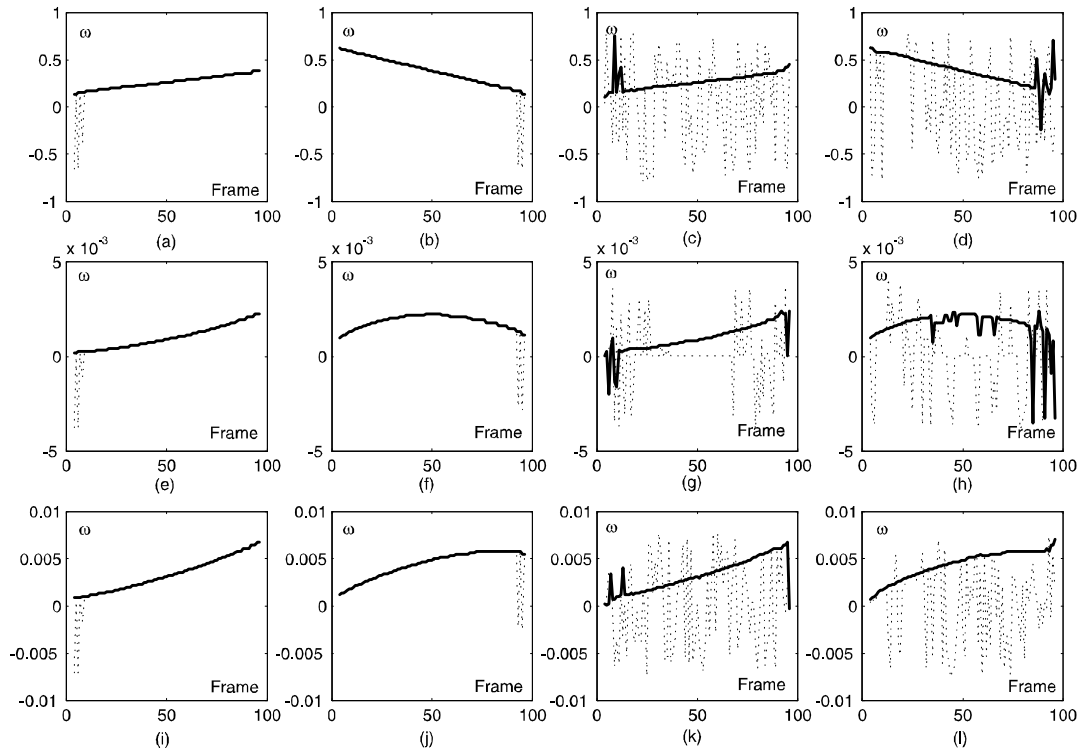|	true	$\sigma = 0, NF$	$\sigma = 0, F$	$\sigma = 40, NF$	$\sigma = 40, F$
a_x	0.0010	0.00073	0.00095	0.0021	0.0033
v_x	0.50	0.6928	0.3666	-2.1742	-1.8378
x_0	20	17.6741	32.8455	30.2183	28.7478
a_y	-0.0020	-0.0015	-0.0158	0.0025	0.0038
v_y	2.5	2.5193	2.5131	-1.2903	-0.3212
y_0	15	25.9683	25.4098	-199.5194	-313.3988

Fig. 4. IF estimation: First row - constant μ-propagation; Second row - variable μ-propagation; Third row - new propagation approach; First column - $z_x(t)$ for non-noisy sequence; Second column - $z_y(t)$ for non-noisy sequence; Third column - $z_x(t)$ for noisy sequence; Fourth column - $z_y(t)$ for noisy sequence; Thick line - with filtering; Dashed line - without filtering.

TABLE III

ESTIMATION OF THE MOTION PARAMETERS BY USING NEW APPROACH. NF - WITHOUT FILTERING OF PROJECTION; F - WITH FILTERING.

	true	$\sigma = 0, NF$	$\sigma = 0, F$	$\sigma = 40, NF$	$\sigma = 40, F$
a_x	0.0010	0.00078	0.00083	-0.0045	0.0012
v_x	0.50	0.6925	0.6564	3.4750	0.3502
x_0	20	25.5096	26.9643	-142.6412	41.61
a_y	-0.0020	-0.0017	-0.0174	0.0014	-0.0019
v_y	2.5	2.2907	2.2925	-1.4083	2.3983
y_0	15	31.5317	31.4440	81.6503	25.1504

Example 4. Real video-sequence representing two walkers is considered. Walkers have opposite motion directions. Sequence has 220 frames. Walker 1 appear about frame 40 while walker 2 is present in the sequence from frame 1. The constant μ- and new propagation are performed with parameters $\mu = 0.2$ and $\frac{1}{1000}$ respectively. The SM calculated for sig-

nal produced with the constant μ-propagation is depicted in Fig. 10a, while the IF (velocity) estimations are given in Fig. 10b. Velocity estimates are marked with dotted and solid lines for walker 1 and 2, respectively. The SM obtained based on new approach is given in Fig. 10c. Position estimates obtained based on the results from constant μ- and new propagation
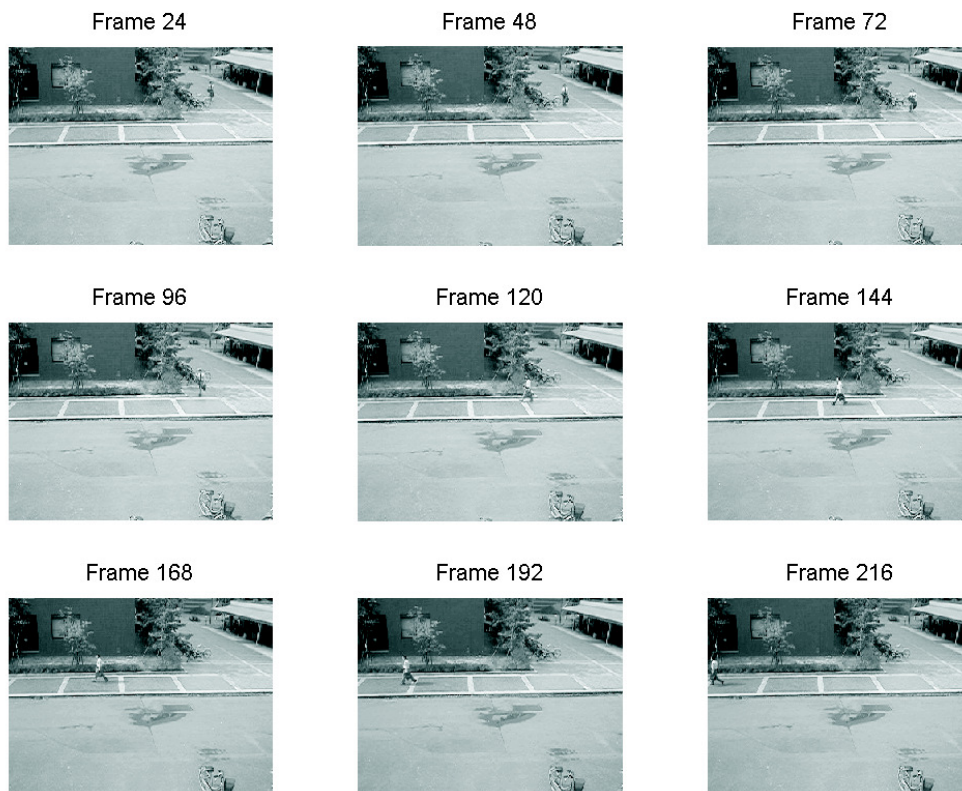
Frame 24   Frame 48   Frame 72

Frame 96   Frame 120   Frame 144

Frame 168   Frame 192   Frame 216

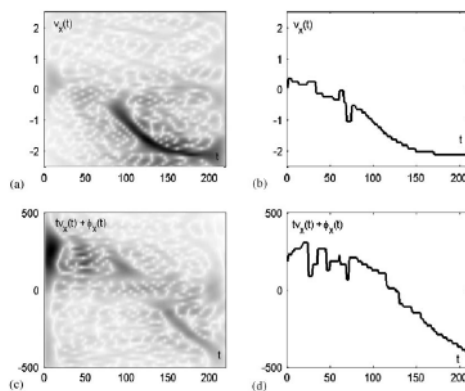Fig. 5. Frames in the real video-sequence.

Fig. 6. Position estimation by using constant μ-propagation and new approach. Thick line - estimate; Thin line - exact position.
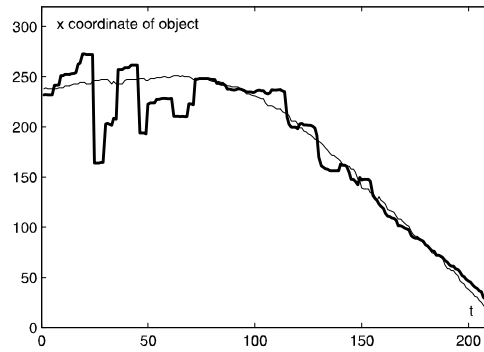
Fig. 7.   Position estimation by using constant μ-propagation and new approach. Thick line - estimate; Thin line - exact position.

are given in Fig. 10d (position estimate for walker 1 are given after frame 40 where this object appears in the scene). In order to visualize these results in a better way Fig. 11 is presented. Estimate of walker 1 position along x-coordinate are marked with sign '=' in bottom of frame, while position of the walker 2 are marked with sign '*' in top of frame. It can be seen that the proposed approach produces accurate indication of the objects' positions.

VI. Conclusion

A procedure for the velocity estimation using the TF representations has been presented. The WD was applied to a single moving object in the sequence, while for multiple objects and real-sequences the S-method can be used. The time-varying form of the variable μ-propagation is proposed in order to improve accuracy of the object position estimation. This propagation approach produces signal with smaller nonlinearity in the signal phase. It means that TF representation-based estimates of motion parameters have better accuracy due to the smaller bias. Furthermore, smaller bias implies better concentration along the IF and, at the same time, robustness to the noise influence. These effects have been proved in the numerical and real-life examples. The proposed approach can be used for tracking a small number of objects in video-sequences with highly non-stationary motion. Possible application areas can be traffic control and sports event monitoring. Note that current video-compression algorithms could be in-

accurate in the presence of objects with highly varying motion parameters. Therefore, application of the proposed algorithm to low-quality video-compression will be one of the further research directions.

Appendix A

The sources of error in the IF estimation are analyzed in [20]. The bias caused by the IF non-linearity is [20]:

$$bias(\hat{\omega}_x(t)) = \sum_{s=1}^{\infty} a_s \frac{d^{2s+1}\phi_x(t)}{dt^{2s+1}} N^{2s}, \quad (26)$$

where $\{a_s, \ s = 1, 2, ...\}$ are constants depending on the window width. This error increases as the window length N increases. Error caused by the stochastic influence asymptotically approaches:

$$var\{\omega_x(t) - \hat{\omega}_x(t)\}$$

$$= \frac{6\sigma_z^2}{|A_x(t)|^2} \left(1 + \frac{\sigma_z^2}{2|A_x(t)|^2}\right) \frac{1}{N^3}. \quad (27)$$
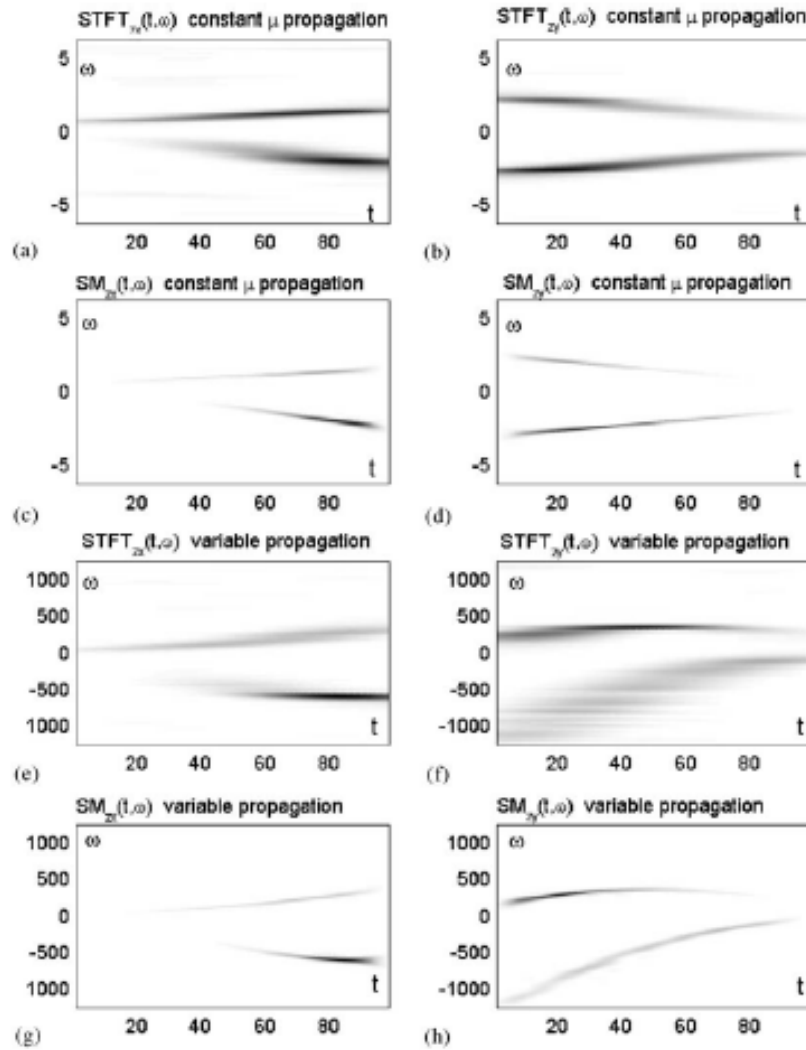
Fig. 8. Time-frequency representations: (a) $STFT_{zx}(t,\omega)$ - constant μ-propagation; (b) $STFT_{zy}(t,\omega)$ - constant μ-propagation; (c) $SM_{zx}(t,\omega)$ - constant μ-propagation; (d) $SM_{zy}(t,\omega)$ - constant μ-propagation; (e) $STFT_{zx}(t,\omega)$ - new propagation approach; (f) $STFT_{zy}(t,\omega)$ - new propagation approach; (g) $SM_{zx}(t,\omega)$ - new propagation approach; (h) $SM_{zy}(t,\omega)$ - new propagation approach.

Note that expressions (26)-(27) are derived under the assumption that the additive noise is of the moderate magnitude [20]. However, in the high noise environment the WD maxima can be outside the auto-term and causing the dominant error in the IF estimation [21]. This error becomes high in the case of signals with high non-linearity in the IF function.

This is the reason to present properties of the WD-based IF estimator in this section.

Note that the high noise has smaller impact to the WD-based IF estimator than to the higher-order TF representation-based ones. This is the reason why in this application we have used the WD instead of some higher-order TF representations, which can be very sensitive to the noise influence.
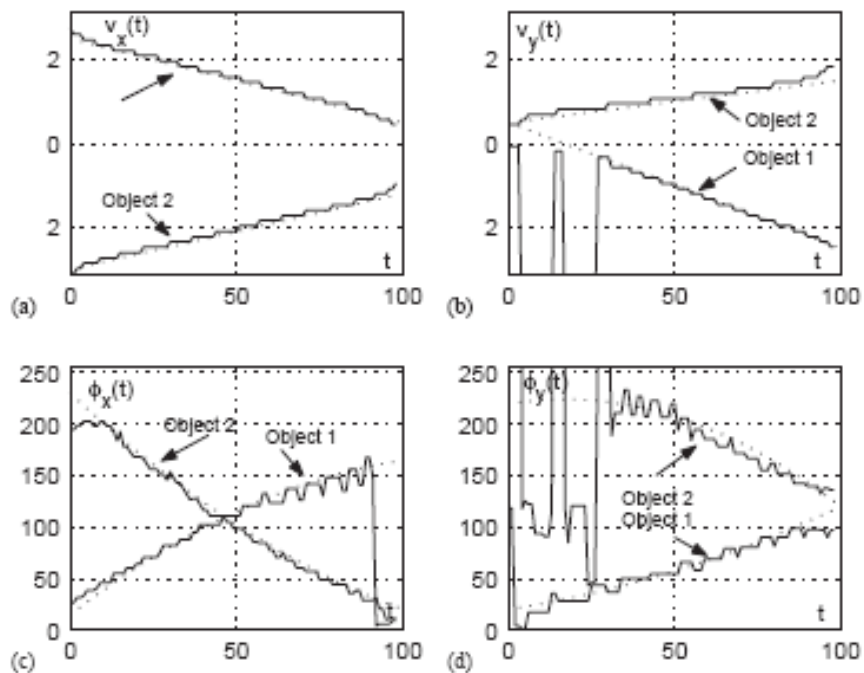
Fig. 9. Velocity and position estimation: (a) $\hat{v}_x(t)$; (b) $\hat{v}_y(t)$; (c) $\hat{\varphi}_x(t)$; (d) $\hat{\varphi}_y(t)$. Exact values - dashed lines; Estimates - solid lines.
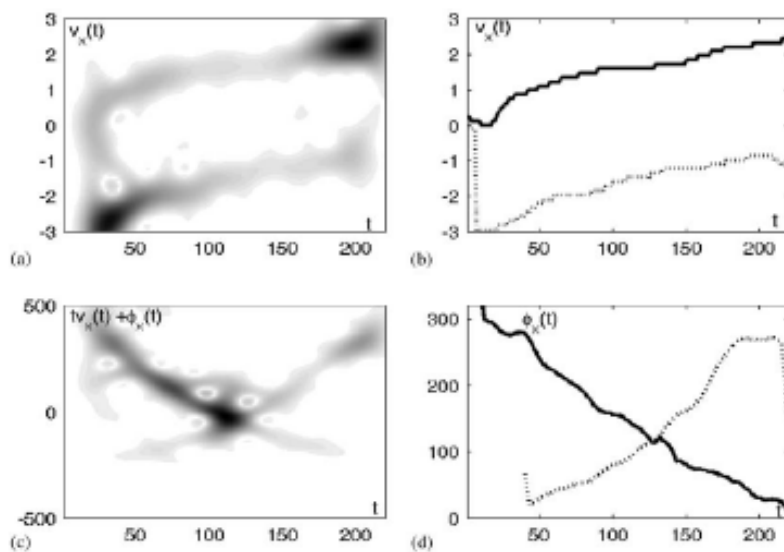

Fig. 10.   Parameters estimation for video-sequence with two moving objects: (a) S-method obtained with constant μ-propagation; (b) Velocity estimates; (c) S-method obtained with new propagation approach; (d) Position estimates. Dotted line - walker 1; Solid line- walker 2.
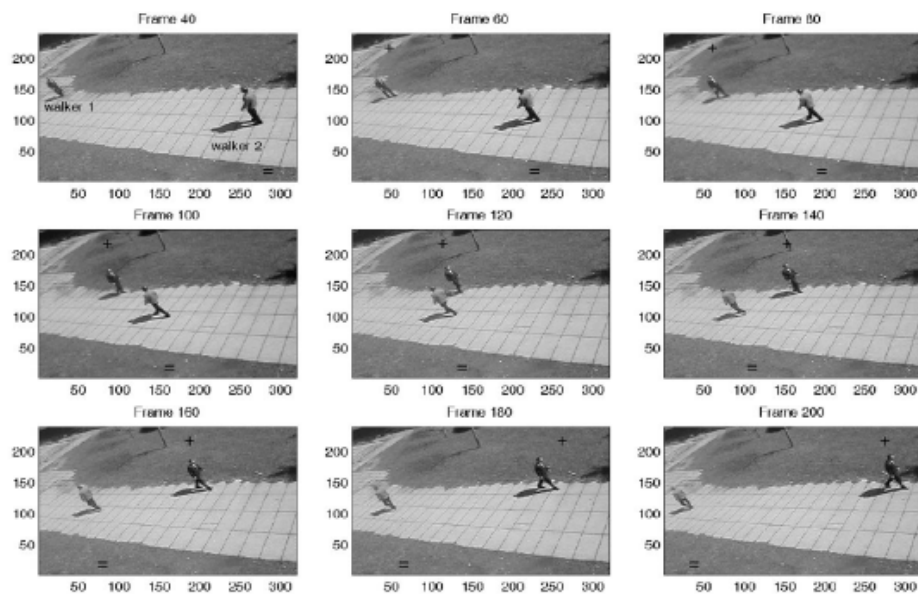
Fig. 11.　Frames from the real video-sequence. x-coordinate of the walker 1 are depicted with '+' in top of frames while for walker 2 it is marked with '=' in bottom of frames.

References

[1] C. Stiller and J. Konrad, "Estimating motion in image sequances," *IEEE Sig. Proc. Mag.*, Vol. 16, No.4, Apr. 1999, pp. 70-91.

[2] D. Vernon, *Machine vision*, Prentice-Hall, Englwood Cliffs, NJ. 1991.

[3] D. Vernon, M. Tistarelli, "Using camra motion to estimate range for robotic parts manipulation, *IEEE Tans. Robotics Automation*, Vol. 6, No. 5, Oct. 1990, pp. 509-521.

[4] J. L. Barron, D. J. Fleet and S. Beauchemin, "Performance of optical flow techniques," *Int. Jour. Comp. Vis.*, Vol. 12, No.1, 1994, pp. 43-77.

[5] H. K. Aghajan and T. Kailath, "SLIDE: Subspace-based line detection," *IEEE Trans. PAMI*, Vol. 16, No.11, Nov. 1994, pp. 1057-1073.

[6] H. K. Aghajan, B. H. Khalaj and T. Kailath, "Estimation of multiple 2-D uniform motions by SLIDE: Subspace-based line detection", *IEEE Trans. Im. Proc.*, Vol. 8, No.4, Apr. 1998, pp. 517-526.

[7] S. Stanković and I. Djurović, "Motion parameters estimation by using time-frequency representations," *El. Let.*, Vol. 37, No.24, Nov. 2001, pp. 1446-1448.

[8] P. Milanfar, "Two-dimensional matched filtering for motion estimation," *IEEE Trans. Image Process.*, Vol. 8, No. 3, Mar. 1999, pp. 438-444.

[9] P. Milanfar, "A model of the effects of the image motion in the Radon transformation domain," *IEEE Trans. Im. Proc.*, Vol. 8, No.3, Mar. 1999, pp. 438-444.

[10] A. Scaglione and S. Barbarossa, "Estimating motion parameters using parametric modeling based time-frequency representations," Radar 97, 1997, pp. 280-284.

[11] P. Milanfar, "Two-dimensional matched filtering for motion estimation," *IEEE Trans. Im. Proc.*, Vol. 8, No.9, Sep. 1999, pp. 1276-1281.

[12] K. Abed-Meraim and A. Beghdadi, "Multiline fitting using polynomial phase transforms and downsampling," in *Proc. of IEEE ICASSP'2001*, Salt Lake City, Vol. 3, pp. 1701-1704.

[13] I. Djurović and S. Stanković, "Estimation of time-varying velocities of moving objects in video-sequences by using time-frequency representations," *IEEE Trans. Image Proces.*, Vol. 12, No.5, Sep. 2003, pp. 550-562.

[14] C. E. Erdem, G. Z. Karabulut, E. Yanmaz and E. Anarim, "Motion estimation in the frequency domain using fuzzy c-planes clustering," *IEEE Trans. Im. Proc.*, Vol. 10, No.12, Dec. 2001, pp. 1873-1879.

[15] W.-G. Chen, G. B. Giannakis and N. Nadhakumar, "A harmonic retrivial framework for discontinous motion estimation," *IEEE Trans. Im. Proc.*, Vol. 7, No.9, Sep. 1998, pp. 1242-1257.

[16] L. Jacobson and H. Wechsler, "Derivation of optical flow using spatiotemporal-frequency approach," *Comp. Vis. Grap. Im. Proc.*, Vol. 38, 1987, pp. 29–65.

[17] *Proceedings of IEEE*, Special issue on time-frequency analysis, Vol. 84, No.9, Sep. 1996.

[18] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal - Part I," *Proc. IEEE*, Vol. 80, No.4, Apr. 1992, pp. 521-538.

[19] P. Rao and F. J. Taylor, "Estimation of the instantaneous frequency using discrete Wigner distribution," *El. Let.*, Vol. 26, 1990, pp. 246-248.

[20] LJ. Stanković and V. Katkovnik, "Instantaneous

frequency estimation using higher order distributions with adaptive order and window length," *IEEE Trans. Inf. Th.*, Vol. 46, No.1, Jan. 2000, pp. 302-311.

[21] I. Djurović and LJ. Stanković, "Influence of high noise on the instantaneous frequency estimation using time-frequency distributions," *IEEE Sig. Proc. Let.*, Vol. 7, No.11, Nov. 2000, pp. 317-319.

[22] R. Roy and T. Kailath, "ESPRIT: Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. ASSP*, Vol. 37, No.7, July 1989, pp. 984-995.

[23] J. Jeong and W. J. Williams, "Kernel design for reduced interference distributions," *IEEE Trans. Sig. Proc.*, Vol. 40, No.10, Oct. 1992, pp. 2498-2517.

[24] LJ. Stanković, "A method for time-frequency signal analysis," *IEEE Trans. Sig. Proc.*, Vol. 42, No.1, Jan. 1994, pp. 225-229.

[25] B. Barkat, B. Boashash, "A high-resolution quadratic time-frequency distribution for multi-component signals analysis," *IEEE Trans. Signal Process.*, Vol. 49, No. 10, Oct. 2001, pp. 2232-2239.

[26] A. Francos and M. Porat, "Analysis and synthesis of multicomponent signals using positive time-frequency distributions," *IEEE Trans. Sig. Proc.*, Vol. 47, No.2, Feb. 1999, pp. 493-504.