The signal flow graph for the fast recursive implementation method of OC with the GSE of size two is presented in Fig. 4. The comparison of Fig. 3 with Fig. 4 shows that the fast recursive structure in Fig. 4 requires significantly fewer computations for an output of OC.

V. CONCLUSION

Efficient real-time implementation methods for the FP morphological operators were presented by extending our previous work [5], [6]. It was shown that the proposed recursive algorithms can improve the computational efficiency of the basis matrix implementation method by avoiding the redundant steps in computing overlapping min/max operations. It was also shown that, with the proposed recursive algorithms, both opening and closing can be determined in real time by $2N - 2$ additions and $2N - 2$ comparisons, and both OC and CO by $4N - 4$ additions and $4N - 4$ comparisons when the size of the GSE is equal to N. Moreover, the proposed recursive algorithms can reduce the memory requirement further than the basis matrix representations.

REFERENCES

[1] G. Matheron, *Random Sets and Image Geometry*. New York: Wiley, 1975.

[2] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.

[3] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[4] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 523–550, July 1987.

[5] S. J. Ko, A. Morales, and K. H. Lee, "Matrix representation of composite morphological function processing systems," in *Proc. 36th Midwest Symp. Circuits Syst.*, Aug. 1993.

[6] ———, "Block basis matrix implementation of the morphological open-closing and clos-opening," *IEEE Signal Processing Lett.*, vol. 2, pp. 7–9, Jan. 1995.

[7] P. Maragos and R. W. Schafer, "Morphological filters–Part I: Their set-theoretic analysis and relations to linear shift-invariant filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1153–1169, Aug. 1987.

[8] P. Maragos, "A representation theory for morphological image and signal processing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 586–599, June 1989.

Motion Estimation Using Higher Order Statistics

Elisa Sayrol, Antoni Gasull, and Javier R. Fonollosa

Abstract—The objective of this paper is to introduce a fourth-order cost function of the displaced frame difference (DFD) capable of estimating motion even for small regions or blocks. Using higher than second-order statistics is appropriate in case the image sequence is severely corrupted by additive Gaussian noise. Some results are presented and compared to those obtained from the mean kurtosis and the mean square error of the DFD.

I. INTRODUCTION

There is a growing interest in applications involving the estimation of 2-D motion or velocity field between consecutive image frames [14]. There are some situations where motion has to be estimated in the presence of noise. These include motion estimation or motion compensation applications, as in images from surveillance cameras or medical images such as echographics with speckle noise. In such cases, most existing methods do not work properly, and more robust techniques are necessary. On the other hand, noise can be realistically described as a colored Gaussian process. In such circumstances, higher order statistics (HOS) may offer some advantages since cumulants of Gaussian processes are asymptotically zero.

HOS-based methods have already begun to be used in motion estimation. In [7] the displacement vector is obtained by maximizing a third-order statistics criterion. In [2] several algorithms are developed based on a parametric cumulant method, a cumulant-matching method, and a mean kurtosis error criterion. The latter is an extension of the quadratic pixel-recursive method by Netravali and Robbins [10]. Other improved extensions to this algorithm were given by Walker and Rao [18] and Biemond et al. [3]. On the other side, iterative solutions that involve additional constraints to compute optical flow have been studied by Horn and Shunk [5] and latter by Nagel [9] using, in both cases, smoothness constraints. See [17] for a review on motion estimation techniques.

In this correspondence, we propose an alternative criterion that exploits HOS [12], [13]. However, our goal is to obtain a low-variance cost function to reduce the problems associated with the estimation of HOS for small blocks of data. Our method is based on an adaptive algorithm that was proposed in [1] for the estimation of fourth-order cumulants. The motivation behind this approach is to use previous frames and previously estimated displacements.

This work is organized as follows. The problem formulation is introduced in Section II. In Sections III and IV, cost functions based on the variance and the kurtosis of the DFD are revised. In Section V, a new class of HOS-based cost functions is derived. In Section VI, a recursive version of the new cost function is presented. Simulation results are provided in Section VII and, finally, Section VIII is devoted to conclusions and final remarks.

II. PROBLEM FORMULATION

The problem of motion estimation can be stated as follows: "Given an image sequence, compute a representation of the motion field that best aligns pixels in one frame of the sequence with those in the next" [10]. This is formulated as

$$g_{k-1}(\boldsymbol{m}) = f_{k-1}(\boldsymbol{m}) + n_{k-1}(\boldsymbol{m})$$
$$g_k(\boldsymbol{m}) = f_k(\boldsymbol{m}) + n_k(\boldsymbol{m}) = f_{k-1}(\boldsymbol{m} - \boldsymbol{d}_k^o(\boldsymbol{m})) + n_k(\boldsymbol{m}) \quad (1)$$

where $\boldsymbol{m} = (m, n)$ denotes spatial image position of a point; $g_k(\boldsymbol{m})$ and $g_{k-1}(\boldsymbol{m})$ are observed image intensities at instant k and $k - 1$, respectively; $f_k(\boldsymbol{m})$ and $f_{k-1}(\boldsymbol{m})$ are noise-free frames; $n_k(\boldsymbol{m})$ and $n_{k-1}(\boldsymbol{m})$ are assumed to be spatially and temporally stationary, zero-mean image Gaussian noise sequences with unknown covariance; and $\boldsymbol{d}_k^o(\boldsymbol{m})$ is the displacement vector of the object during the time interval $[k - 1, k]$. The noise-free signals are assumed to be zero-mean non-Gaussian random fields that are statistically independent of the noise. In this formulation, the basic assumption is intensity constancy, as follows:

$$f_k(\boldsymbol{m}) = f_{k-1}(\boldsymbol{m} - \boldsymbol{d}_k^o(\boldsymbol{m})). \quad (2)$$
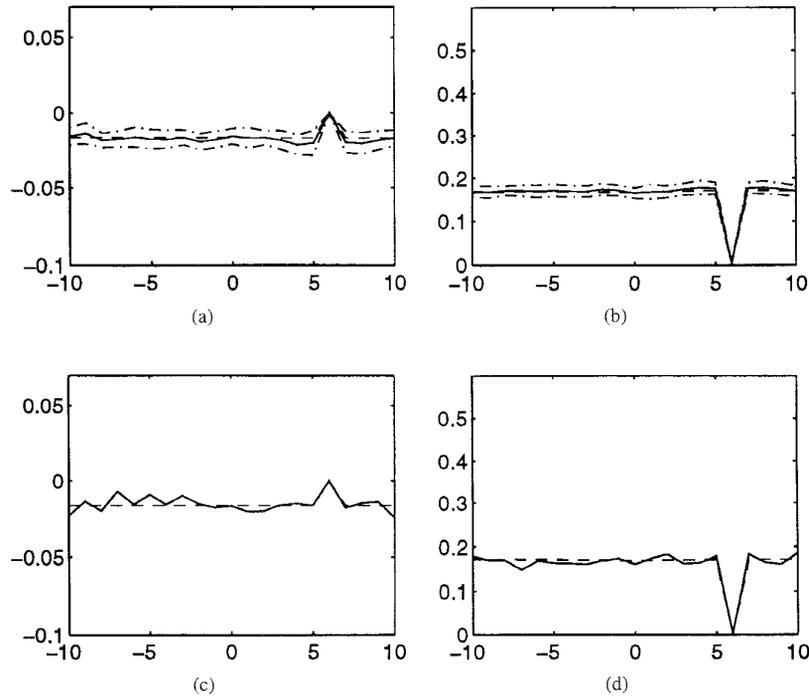
Fig. 1. Cost functions for $SNR = 15$ dB of a 256 length object where $d_k^0 = 6$. (a) $J_{41k}(d)$. (b) $J_{2k}(d)$. (c) single realization of $J_{41k}(d)$. (d) single realization of $J_{2k}(d)$.

The problem is to estimate $d_k^o(m)$ from the observation of $g_k(m)$ and $g_{k-1}(m)$. The $DFD_k(d)$ is defined [10] as follows:

$$DFD_k(d) = g_k(m) - g_{k-1}(m - d) \qquad (3)$$

or equivalently

$$DFD_k(d) = [f_{k-1}(m - d_k^o) - f_{k-1}(m - d)] \\ + [n_k(m) - n_{k-1}(m - d)] \qquad (4)$$

where we omit the space dependency of the displacement to simplify notation. In our model we consider that pixels of regions are visible over the entire frame; that is, no occlusions occur, pixels do not move in or out. Given the previous assumptions, we propose a motion estimation scheme that is divided into two steps; we concentrate our efforts on the second step.

Segmentation: We may work with motion estimation based on blocks or on a segmentation approach. The latter aims to adapt the segmentation to the scene such that each region uniquely corresponds to one continuously moving 2-D object [15]. Several problems that are inherent to block-oriented approaches are avoided; i.e., blocking artifacts are drastically reduced and small region sizes are not imposed. However, the method becomes increasingly complex as the number of regions undergoing different displacements increases. In this case, we can also apply block-oriented methods without loss of generality.

Motion Estimation: For every moving region or block, we estimate motion using a HOS-based cost function that is maximized or minimized for the desired displacement.

Next, we introduce different criteria to obtain the displacement vector based on second- and fourth-order statistics of the $DFD_k(d)$. We analyze under which conditions it is more appropriate to utilize each of the cost functions.

III. VARIANCE OF THE DFD

The classical solution to obtain the displacement vector from the $DFD_k(d)$ is the minimum square error [10] as follows:

$$J_{2k}(d) = E\{DFD_k^2(d)\}. \qquad (5)$$

An estimation of this cost function is given by the sample averaging

$$\hat{J}_{2k}(d) = \frac{1}{N} \sum_{m \in \Omega_m} DFD_k^2(d) \qquad (6)$$

where Ω_m denotes the spatial domain that contains the pixels from a region, and N the number of such pixels.

Equation (5) can also be expressed as a function of signal and noise covariances, substituting (4) in (5) as follows:

$$J_{2k}(d) = 2\sigma_f^2 - 2E\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\} \\ + 2\sigma_n^2 - 2E\{n_k(m)n_{k-1}(m - d)\} \qquad (7)$$

where σ_n^2 and σ_f^2 are the signal and noise variances at time k and $k - 1$. In case images are affected by white noise, the above cost function can be utilized to detect the correct displacement, since the noise covariance term cancels out and its only contribution in (7) is the variance term, which is assumed constant. Nevertheless, we are interested in studying the case when noise is colored; consequently, its covariance contribution is different from zero. In this situation, the cost function at the correct displacement may not be a minimum. The characterization of the signal and noise covariances may allow a deeper knowledge of the effects of the degradation. The signal covariance term in (7) has a maximum at the correct displacement for any statistics of the region. Analogously, the noise covariance term will also show a maximum for displacement zero. Thus, $J_{2k}(d)$ will show two local minima, one at displacement zero and the other at the correct displacement.

To obtain the correct displacement, we require the global minimum of the function to be located at the desired displacement. Developing
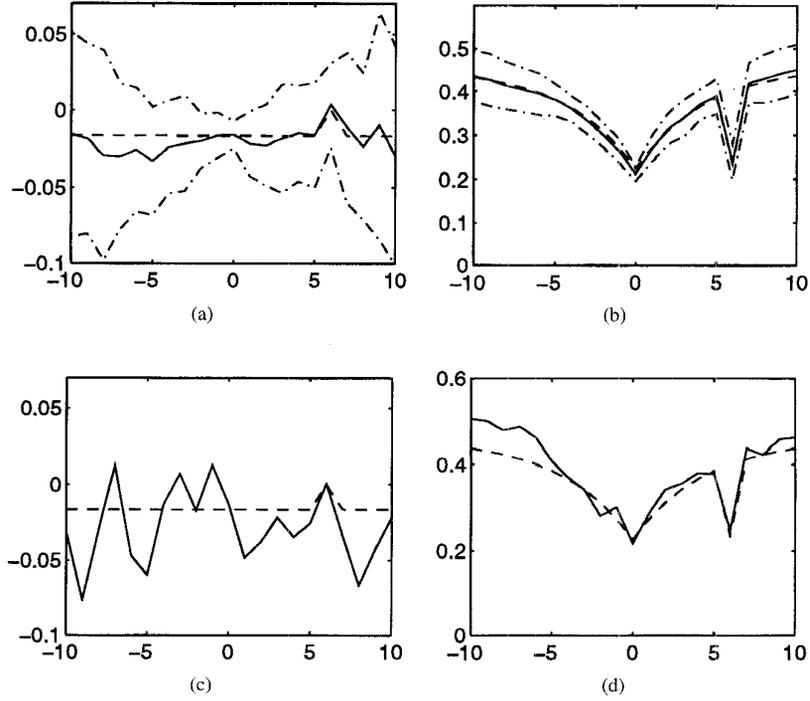
Fig. 2. Cost functions for $SNR = -2.5$ dB of a 256 length object where $d_k^o = 6$. (a) $J_{41k}(d)$. (b) $J_{2k}(d)$. (c) Single realization of $J_{41k}(d)$. (d) Single realization of $J_{2k}(d)$.

the inequality $J_{2k}(d_k^o) < J_{2k}(0)$ as a function of signal and noise covariances,we obtain a minimum SNR from which the correct displacement can be found. In lower SNR scenarios, we should seek an alternative measure to obtain displacement. We are interested in choosing a cost function based on HOS to reduce the presence of colored Gaussian noise. Recall that HOS methods are blind to Gaussian processes and, thus, information due to deviations from Gaussianity may be extracted [8], [11].

IV. KURTOSIS OF THE DFD

HOS-based cost functions can be built from different criteria. The approach in [2] is based on a fourth-order statistics cost function that utilizes the kurtosis of the $DFD_k(d)$, which is asymptotically unaffected by correlated Gaussian noise. It is defined as

$$J_{41k}(d) = K(DFD_k(d)) \tag{8}$$

where the kurtosis is by definition [11], as follows:

$$K(DFD_k(d)) = E\{DFD_k^4(d)\} - 3[E\{DFD_k^2(d)\}]^2 \tag{9}$$

The correct displacement is found by minimizing $J_{41k}(d)$ when the region has a positive kurtosis or maximizing it if the kurtosis is negative. Tugnait [16] was the first to propose this criterion to estimate the time delay between two signals as an extension to the performance index $J_{2k}(d)$. Later, Anderson and Giannakis [2] used the above cost function to recursively estimate the displacement of pixels between two images.

The corresponding consistent estimation of $J_{41k}(d)$ is given by

$$\hat{J}_{41k}(d) = \frac{1}{N} \sum_{m \varepsilon \Omega_m} DFD_k^4(d) - 3\left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} DFD_k^2(d)\right]^2 \tag{10}$$

Asymptotically, the presence of Gaussian noise does not degrade the detection process and the cost function will depend only on the

signal second- and fourth-order moments. Substituting (4) in (9) and developing the two expectation terms, the cost function becomes

$$\begin{aligned}
J_{41k}(d) = {} & 2m_{f4} + 6E\{f_{k-1}^2(m - d_k^o)f_{k-1}^2(m - d)\} \\
& - 4E\{f_{k-1}(m - d_k^o)f_{k-1}^3(m - d)\} \\
& - 4E\{f_{k-1}^3(m - d_k^o)f_{k-1}(m - d)\} \\
& - 12\sigma_f^4 - 12E^2\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\} \\
& + 24\sigma_f^2 E\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\} \tag{11}
\end{aligned}$$

where $m_{f4} = E\{f_{k-1}^4(m - d_k^o)\} = E\{f_{k-1}^4(m - d)$. Observe that all noise terms have been canceled.

As important as the theoretical result is the study of the variance in estimating each of the cost functions. In the following example we consider AR models; hence, the noise covariance may be given by [6, p. 214]

$$E\{n_k(m)n_{k-1}(m - d)\} = \sigma_n^2 a_m^{|d_m|} a_n^{|d_n|} a_k \tag{12}$$

where $d = (d_m, d_n)$ are the two displacement components and a_m, a_n, a_k are the AR coefficients. The signal covariance term within the region is also given by

$$E\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\} = \sigma_f^2 b_m^{|d-d_m|} b_n^{|d-d_n|} \tag{13}$$

where b_m and b_n are the AR coefficients. Similar expressions can be derived for the fourth-order moments in (11); although we avoid them here for the sake of simplicity, they depend on the same parameters.

In Figs. 1 and 2 we compare $J_{2k}(d)$ and $J_{41k}(d)$ for a rectangular 1-D object of length 256 whose kurtosis is negative. Each figure displays the following. The solid line is the mean of the cost function using 20 realizations of a sequence of two signals, where $d_k^o = 6$; the dashed line is the theoretical cost function; and finally, the dashed-dot line is the mean plus/minus the standard deviation. Fig. 1(a) and (b) shows the estimation for SNR = 15 dB, the pixels of the signal are uncorrelated, and the Gaussian noise follows an AR model with
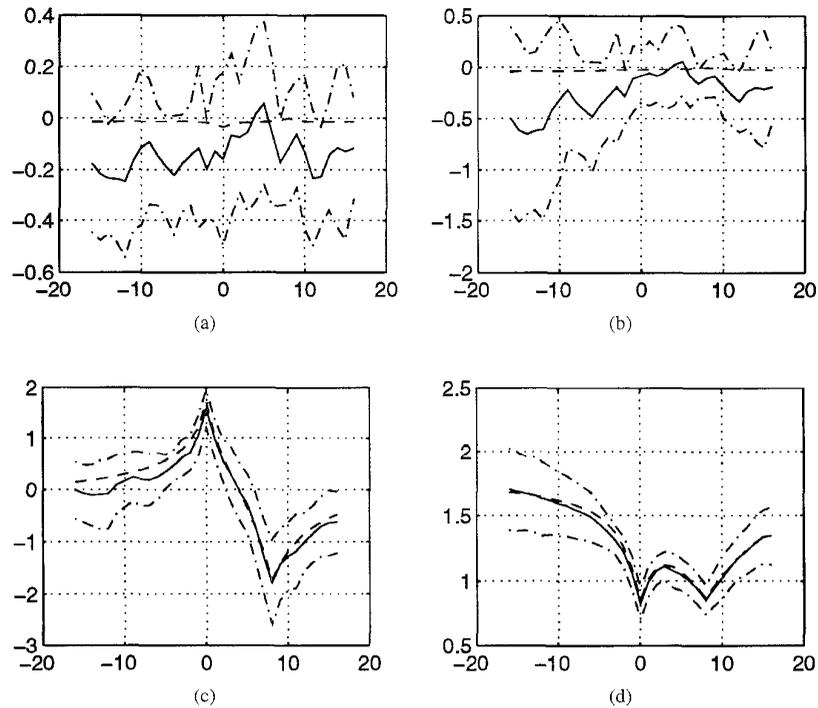
Fig. 3. Estimated cost functions when $d_{k-1}^0 = 0, d_k^0 = 8$ and $SNR = -2$ dB for a colored signal of length 256. (a) $J_{41k}(\boldsymbol{d})$ normalized. (b) $J_{41k}(\boldsymbol{d})$. (c) $J_{42k}(\boldsymbol{d})$. (d) $J_{2k}(\boldsymbol{d})$.
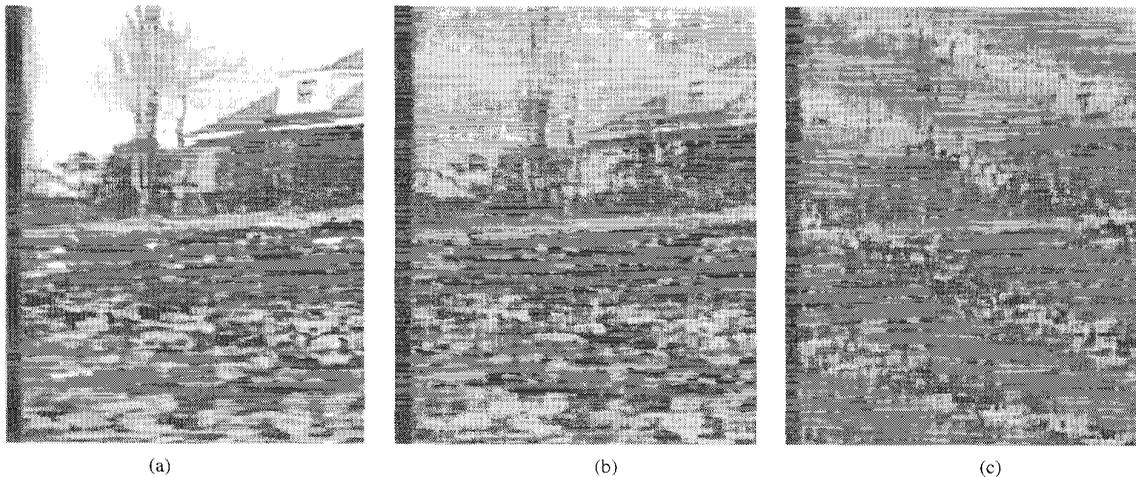
Fig. 4. (a) One of the original frames from the Flower Garden sequence. (b) One of the noisy frames where colored noise follows an AR model with $a_k = a_m = a_n = 0.6$ and the SNR is 10 dB. (c) One of the noisy frames when the SNR is 2 dB.

$a_m = a_k = 0.8$. The mean estimation is close to the theoretical expectations. Fig. 2(a) and (b) shows both cost functions when SNR = -2.5 dB. Clearly, although the mean estimation of the kurtosis tends to the theoretical result, the variance is very high. Similar behavior was observed for other AR parameters. These results suggest that the kurtosis-based cost function should not be used for low SNR unless dealing with longer signals, which would reduce the variance. We must be aware of the fact that we are going to use this cost function for image regions or blocks and the number of pixels may not be large.

V. MODIFIED KURTOSIS OF THE DFD

Image information is repeated along the sequence as it is established in (2). This redundancy may be used to obtain better

estimates of HOS to reduce the effect of additive noise. Amblard *et al.* [1], proposed an adaptive scheme for the estimation of fourth-order cumulants for transient detection that overcomes the problems of the classical adaptive scheme. It was proven for the case of i.i.d. random variables that the estimator is asymptotically unbiased. A convergence analysis was also carried out. The original expression was defined for a sequence of random variables. In this work, we have derived an expression as a function of the kurtosis of the DFD in a region or block. At time k this expression becomes [13], [1]

$$\hat{K}_k(\mathrm{DFD}(\boldsymbol{d})) = \hat{K}_{k-1}(\mathrm{DFD}(\boldsymbol{d})) + \gamma[\hat{k}(\mathrm{DFD}_k(\boldsymbol{d}))$$
$$- \hat{K}_{k-1}(\mathrm{DFD}(\boldsymbol{d}))] \tag{14}$$

TABLE I
PERCENTAGE OF ERROR WHEN TRACKING A MOVING OBJECT

size	32x32	16x16	32x32	16x16	32x32	16x16
SNR	0 dB	0 dB	-2 dB	-2 dB	-5 dB	-5 dB
$J_{42k}(d)$	0%	10%	10%	30%	80%	90%
$J_{2k}(d)$	0%	20%	40%	40%	100%	100%

where $\hat{k}(\mathrm{DFD}_k(d))$ is the "instantaneous kurtosis" given by

$$\hat{k}(\mathrm{DFD}_k(d)) = \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^4(d)) - 3$$
$$\cdot \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(d)\hat{E}_{k-1}\{\mathrm{DFD}^2(d)\} \right], \quad (15)$$

and

$$\hat{E}_k\{\mathrm{DFD}^2(d)\} = \hat{E}_{k-1}\{\mathrm{DFD}^2(d)\} + \mu \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(d) \right.$$
$$\left. - \hat{E}_{k-1}\{\mathrm{DFD}^2(d)\} \right] \quad (16)$$

μ and γ are forgetting factors that adapt the estimation to changing conditions. The subindex k has been suppressed in the DFD's where previous frames are involved. We are interested in obtaining the displacement vector at time k and this only depends on the instantaneous kurtosis, that is, (15) and (16). It is clear that there is no information on the current displacement between two frames in previous images or in previous displacements (unless we introduce some prediction or motion model). Previous frames can be used to obtain information on the statistics of the regions and/or statistics of the noise. Our goal is to define a low variance cost function that should be, at the same time, asymptotically unaffected by correlated Gaussian noise. Thus, we choose (15) as the cost function to obtain the displacement. Observe that, comparing this equation with the kurtosis in (10), instead of having the instantaneous second-order moment estimate to the power of two, which may show a high variance, we have a term that is the product of the instantaneous second-order moment by a past estimation of the second-order moment using more than two frames. The new cost function is then given by

$$\hat{J}_{42k}(d) = \frac{1}{\hat{J}_{2k}^2(d)} \cdot \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^4(d) - 3\frac{1}{N} \right.$$
$$\left. \cdot \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_K^2(d)\hat{E}_{k-1}\{\mathrm{DFD}^2(d)\} \right] \quad (17)$$

where the normalization by the quadratic function improved the results obtained when considering AR models and over the examples using real sequences. The adjustment of the weight μ should be carried out. For small values of μ, past frames become more significant than the previous one. Hence, its value is chosen close to one when the scene in the image sequence changes rapidly (for $\mu = 1$ only the previous frame intervenes), and closer to zero otherwise.

An alternative estimation is necessary in case only two frames are available, or for the first two frames of a sequence. We propose the following approach:

$$\hat{J}_{43k}(d) = \frac{1}{\hat{J}_{2k}^2(d)} \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^4(d) - 3\frac{1}{N} \right.$$
$$\left. \cdot \sum_{m \varepsilon \Omega_m} [g_{k-1}(m) - g_{k-1}(m - d)]\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(d) \right] \quad (18)$$

where we have substituted the estimation of the second-order term by a quadratic difference term from a single frame. The resulting cost function displays a behavior similar to the one in (17), yielding to better estimates of the displacement. It has been successfully applied to the problem of time delay between two 1-D signals [12].

In terms of expectations, (17) is rewritten as

$$J_{42k}(d) = \frac{1}{[E\{\mathrm{DFD}_k^2(d)\}]^2} \cdot [E\{\mathrm{DFD}_k^4(d)\}$$
$$- 3E\{\mathrm{DFD}_k^2(d)\}E_{k-1}\{\mathrm{DFD}^2(d)\}]. \quad (19)$$

As was done for the previous cost functions, we can decompose the cost function in terms of second- and fourth-order moments of the signal and noise by substituting (4) in (19). We obtain for $\mu = 1$:

$$J_{42k}(d) = \frac{1}{J_{2k}^2(d)}[2m_{f4} + 6E\{f_{k-1}^2(m - d_k^o)f_{k-1}^2(m - d)\}$$
$$- 4E\{f_{k-1}^3(m - d_k^o)f_{k-1}(m - d)\}$$
$$- 4E\{f_{k-1}(m - d_k^o)f_{k-1}^3(m - d)\}$$
$$- 3[2\sigma_f^2 - 2E\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\}]$$
$$\cdot [2\sigma_f^2 - 2E\{f_{k-1}(m - d_{k-1}^o)f_{k-1}(m - d)\}]$$
$$- 3[2\sigma_n^2 - 2E\{n_k(m)n_{k-1}(m - d)\}]$$
$$\cdot [2E\{f_{k-1}(m - d_k^o)f_{k-1}(m - d)\}]$$
$$- 2E\{f_{k-1}(m - d_{k-1}^o)f_{k-1}(m - d)\}]. \quad (20)$$

In the case of nonuniform motion, this cost function shows a minimum at the correct displacement for regions characterized by positive or negative kurtosis. The noise term contributes with a maximum to $J_{42k}(d)$ and, thus, the minimum that provides the correct displacement is not mistaken as it happens to $J_{42k}(d)$ for low SNR. This behavior has been studied for AR models, but can be extended to other models whose covariance shows an absolute maximum at zero. In the case of uniform motion, that is, when the optimal displacements at time k and $k - 1$ are equal, we can deduce that $J_{42k}(d)$ for $\mu = 1$ is nothing else but the normalized kurtosis of the $\mathrm{DFD}_k(d)$. It displays a maximum at the correct displacement when the region has negative kurtosis and a minimum when the region kurtosis is positive. Hence, the newly defined cost function always exhibits an absolute minimum at the desired displacement for positive kurtosis regions. The same behavior is observed for negative kurtosis regions except when $d_{k-1}^o = d_k^o$ and $\mu = 1$, in which case the cost function becomes the normalized kurtosis and the desired displacement becomes a maximum. Thus, except for this case, in all other situations the correct displacement is derived from an absolute minimum. We need to introduce a mechanism to avoid searching for a wrong minimum in that case. We use $\mu < 1$; thus, the influence from previous frames is not restricted to frame at time $k - 1$. An additional step can be added in this case, which consists on introducing a shift of $-d_{k-1}^o$ to $\hat{E}_{k-1}\{\mathrm{DFD}^2(d)\}$ (that is available since it has been obtained when estimating the displacement between time instants $k - 1$ and $k - 2$). As a consequence, in the following iteration, $\hat{k}(\mathrm{DFD}_k(d))$ will have a maximum at $d = 0$ and a minimum at $d = d_k^o$.

In Section VII, we present some examples where we compare the cost functions defined up to this point. We will see the effects of the estimation and the advantages of using the newly defined cost function.

VI. RECURSIVE ESTIMATION OF THE DISPLACEMENT

In the previous section, the estimated displacement was taken from an exhaustive search of the displacement that provided the absolute minimum or maximum of a given cost function. Recursive

(a)   (b)   (c)

(d)   (e)   (f)

(g)   (h)   (i)
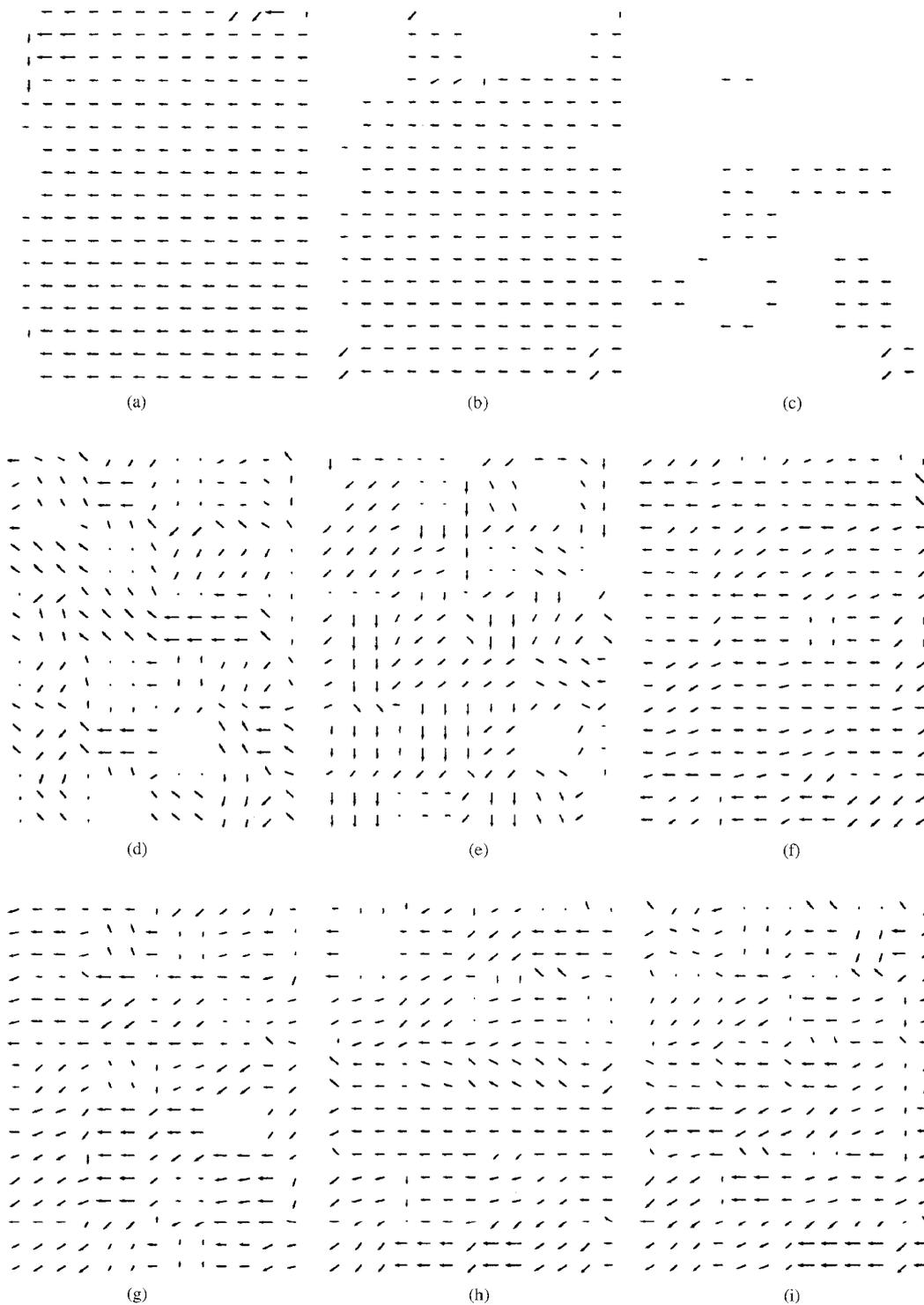
Fig. 5. (a) Noise-free real displacement map given blocks of 16 × 16 pixels obtained from $J_{2k}(d)$. (b) displacement map for noisy frames when using $J_{2k}(d)$, SNR = 10 dB. (c) $J_{2k}(d)$ for SNR = 2 dB. (d) $J_{41k}(d)$for SNR = 10 dB. (e) $J_{41k}(d)$for SNR = 2 dB. (f) $J_{42k}(d)$for SNR = 10 dB. (g) $J_{42k}(d)$)for SNR = 2 dB. (h) $J_{43k}(d)$for SNR = 10 dB. (i) $J_{43k}(d)$for SNR = 2 dB.

estimation algorithms aim to reduce the computational load, specially for subpixel displacements, since they use *a priori* information on the location of regions. Thus, given the ith estimate of the displacement, we obtain the $(i+1)$th estimate such that the value of the cost function resulting from the $(i+1)$th estimate is lower (or higher) than the one used in the ith. The gradient search procedure [10] is

$$d^i = d^{i-1} - \varepsilon \Delta_{d^{i-1}} J(d^{i-1}). \qquad (21)$$

The recursive estimation of $J_{42k}(\boldsymbol{d})$ for $\mu = 1$ substituting (17) in (21) becomes [13] the following:

$$
\boldsymbol{d}^i = \boldsymbol{d}^{i-1} - \varepsilon \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(\boldsymbol{d}^{i-1}) \right]^{-2}
$$

$$
\cdot \left[4 \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^3(\boldsymbol{d}^{i-1}) \nabla_{\boldsymbol{m}} g_{k-1}(\boldsymbol{m} + \boldsymbol{d}_k^o - \boldsymbol{d}^{i-1}) \right.
$$

$$
+ 6 \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_{k-1}^2(\boldsymbol{d}^{i-1}) \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k(\boldsymbol{d}^{i-1})
$$

$$
\cdot [\nabla_{\boldsymbol{m}} g(\boldsymbol{m} + \boldsymbol{d}_k^o - \boldsymbol{d}^{i-1})] - 6 \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(\boldsymbol{d}^{i-1})
$$

$$
\cdot \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_{k-1}(\boldsymbol{d}^{i-1}) \nabla_{\boldsymbol{m}} g_k(\boldsymbol{m} + \boldsymbol{d}_{k-1}^o - \boldsymbol{d}^{i-1})
$$

$$
- 4 \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k(\boldsymbol{d}^{i-1}) \nabla_{\boldsymbol{m}} g_k(\boldsymbol{m} + \boldsymbol{d}_{k-1}^o - \boldsymbol{d}^{i-1})
$$

$$
\left. \cdot \frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^4(\boldsymbol{d}^{i-1}) \left[\frac{1}{N} \sum_{m \varepsilon \Omega_m} \mathrm{DFD}_k^2(\boldsymbol{d}^{i-1}) \right]^{-1} \right].
$$

$$\tag{22}$$

where the gradients of the DFD's are expressed as a function of the gradient of the image intensity at time k considering the previous and current displacements of the region. An example comparing the recursive scheme of $J_{42k}(\boldsymbol{d})$ and $J_{2k}(\boldsymbol{d})$ was given in [12].

VII. EXAMPLES

Example 1: In this example we compare the cost functions for a 1-D object of negative kurtosis. Its length is 256 and it follows a first order AR model whose parameter is $b_m = 0.8$, which is moving along three frames where $\boldsymbol{d}_{k-1}^o = 0$ and $\boldsymbol{d}_k^o = 8$. Noise is simulated using a first order AR model with $a_m = a_k = 0.6$ and SNR $= -2$ dB. We generate the sequence 20 times and obtain the mean behavior of the cost function at time k (The dashed line represents the theoretical curve; the solid line, the mean-estimation; and the dashed-dotted line, the mean $\pm$ the standard deviation of the estimation). We can observe in Fig. 3(b) that the kurtosis does not show the maximum at the correct displacement, and its behavior is not improved after normalization by $J_{2k}^2(\boldsymbol{d})$ (Fig. 3(a)). Fig. 3(c) shows $J_{42k}(\boldsymbol{d})$ using the estimation given by (17), for $\mu = 1$. For the modified kurtosis, the standard deviation is much lower and follows the theoretical results. For $J_{2k}(\boldsymbol{d})$ in Fig. 3(d) the variance is low, but for this SNR the minimum is not absolute.

Example 2: We compare the modified kurtosis with $J_{2k}(\boldsymbol{d})$ in case of uniform motion. For this purpose, we utilize the procedure of shifting $\hat{E}_{k-1}\{\mathrm{DFD}^2(\boldsymbol{d})\}$ by $-\boldsymbol{d}_{k-1}^o$ and $\mu = 0.89$. We are given ten realizations of a sequence of seven synthetic noise-free images containing a 2-D rectangular object. The object was previously segmented and was moving $(3,1)$ pixels per frame. Colored Gaussian noise was generated from a first order AR model with $a_n = a_m = a_k = 0.6$. For this size and signal distribution, the cost function $J_{41k}(\boldsymbol{d})$ failed. Table I shows the percentage errors to reach the final position of the object for different sizes and SNR's. An error was generated when a wrong displacement was estimated between two frames, and this misdetection was not compensated for in the following frames.

Example 3: This example demonstrates the performance of the cost functions $J_{41k}(\boldsymbol{d}), J_{42k}(\boldsymbol{d}), J_{43k}(\boldsymbol{d})$ and $J_{2k}(\boldsymbol{d})$ when block-matching is applied to obtain the displacement between consecutive real frames. Fig. 4(a) shows a portion of one of the original images taken from the Flower Garden sequence (frames 112 to 116). Fig. 4(b) is the same image when colored Gaussian noise generated from an

AR model with $a_k = a_m = a_n = 0.6$ has been added to the sequence and the SNR is 10 dB. In Fig. 4(c), the SNR is 2 dB. The results obtained from the second-order cost function between frames 115 and 116 when no noise is added are illustrated in Fig 5(a). We use these results to compare them with those obtained when colored Gaussian noise is added (it is not our goal to study the noise-free case). Fig. 5(b) and (c) are the maps obtained for $J_{2k}(\boldsymbol{d})$ when colored noise is added to the sequence for SNR $= 10$ and 2 dB, respectively. Analogously, we represent the results for $J_{41k}(\boldsymbol{d})$ in Fig. 5(d) and (e). Fig. 5(f) and (g) shows $J_{42k}(\boldsymbol{d})$, while Fig. 5(h) and (i) illustrate $J_{43k}(\boldsymbol{d})$ for the two different SNR's. Missing arrows in the displacement map indicate that zero displacement has been estimated. The cost functions $J_{42k}(\boldsymbol{d})$ and $J_{43k}(\boldsymbol{d})$ clearly outperform the second-order cost function for SNR $= 2$ dB. In the case of SNR $= 10$ dB, the results given by $J_{42k}(\boldsymbol{d})$ seem to be better in the upper part of the image, whereas $J_{2k}(\boldsymbol{d})$ is preferable in other blocks. The results obtained using the kurtosis were not satisfactory.

VIII. CONCLUSIONS

There are some situations where motion between frames has to be estimated in the presence of noise. In such circumstances, HOS may offer some advantages, since cumulants of Gaussian processes are zero. Cost functions of the DFD based on second- and fourth-order statistics were examined and compared. Examples employing moderate-size regions and low SNR were given. The second-order cost function of the DFD is capable of detecting the correct displacement from an absolute minimum up to certain SNR. On the other hand, the mean behavior of the kurtosis followed the predicted function; however, its standard deviation was so high that detection of the correct displacement was difficult in most of the cases. This measure should be utilized for long data records that imply moderate variance. Finally, the modified kurtosis of the DFD has shown its usefulness for moderate size regions in a range of SNR where the second-order cost function is biased and the kurtosis shows a high variance.

The results of this paper suggest that low-variance cost functions based on higher order statistics may be defined and can outperform not only the second-order cost function but other existing HOS-based cost functions.

REFERENCES

[1] P. O. Amblard and J. M. Brossier," Adaptive estimation of the fourth-order cumulant of a white stochastic process," *Signal Processing*, vol. 42, no. 1, pp. 37–44, Feb. 1995.

[2] J. M. Anderson and G. B. Giannakis, "Image motion estimation algorithms using cumulants", *IEEE Trans. Image Processing*, vol. 4, pp. 346–357, Mar. 1995.

[3] J. Biemond, L. Looijenkga, and D. Boekee, " A pel-recursive Wiener-based displacement estimation algorithm," *Signal Processing*, vol. 13, pp. 399–412, Dec. 1987.

[4] D. R. Brillinger and M. Rosenblatt, "Computation and interpretation of kth-order spectra," in *Spectral Analysis of Time Series*, B. Harris, Ed. New York: Wiley, 1967, pp. 189–232.

[5] B. Horn and B. Shunk, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.

[6] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[7] R. P. Kleihorst, R. L. Lagendijk, and J. Biemond, "Noise reduction of severely corrupted images," in *Proc ICASSP '93*, Minneapolis, MN, pp. v293–96.

[8] J. M. Mendel, "Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proc. IEEE*, vol. 79, pp. 278–305, Mar. 1991.

[9] H.-H. Nagel, "On the estimation of optical flow: Relation between different approaches and some new results," *Artif. Intell.*, vol. 33, pp. 299–324, 1987.

[10] A. N. Netravali and J. D. Robbins, "Motion-compensated television: Part I," *Bell Syst. Tech. J.*, vol 58, pp. 613–670, Mar. 1979.

[11] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework.* Englewood Cliffs, NJ: Prentice-Hall, 1993.

[12] E. Sayrol, T. Gasull and J. R. Fonollosa, "Cost function for motion estimation based on higher-order statistics," in *Proc. EUSIPCO'94*, Edinburgh, UK, pp. 117–1120.

[13] E. Sayrol, "Higher-order statistics applications in image sequence processing," Ph.D. dissertation, Univ. Politecnica de Catalunya, Oct. 1994.

[14] M. I. Sezan and R. L. Lagendijk, Eds., *Motion Analysis and Image Sequence Processing.* Boston, MA: Kluwer, 1993.

[15] C. Stiller, "A statistical image model for motion estimation," in *Proc. ICASSP '93*, Minneapolis, MN, pp. v193–196.

[16] J. K. Tugnait, "Time delay estimation in unknown spatially correlated Gaussian noise using higher-order statistics," in *Proc. 23rd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, 1989, pp. 211–215.

[17] G. Tziritas and C. Labit, "Motion analysis for image sequence coding," in *Advances in Image Communication*, vol. 4. New York: Elsevier, 1994.

[18] D. Walker and K. Rao, " Improved pel-recursive motion compensation," *IEEE Trans. Commun.*, vol. COM-32, pp. 1128–1134, Oct. 1984.

An Estimation Algorithm for 2-D Polynomial Phase Signals

Benjamin Friedlander and Joseph M. Francos

Abstract—We consider nonhomogeneous 2-D signals that can be represented by a constant modulus polynomial-phase model. A novel 2-D phase differencing operator is introduced and used to develop a computationally efficient estimation algorithm for the parameters of this model. The operation of the algorithm is illustrated using an example.

I. INTRODUCTION

A fundamental problem in 2-D signal processing and in many image processing applications is the modeling and analysis of nonhomogeneous 2-D signals. For example, in almost any image taken by a camera, perspective exists, and hence, the acquired 2-D signal is nonhomogeneous, even if the original scene was homogeneous. Conventional approaches to the problems of perspective and camera orientation estimation usually involve *local* analysis of the image by means of edge detection algorithms [6]. Recently, a nonparametric method for estimating, and then canceling, the effects of perspective was suggested in [7], using the Chirplet transform. In this method, a 1-D cross section of the image is expanded onto a set of modulated and warped versions of one "mother-waveform" in order to later compute an unwarped representation of the original image.

Parametric models, when used in image processing, generally assume the observed image to be homogeneous or piecewise homogeneous. In this correspondence, we consider a parametric model that is *nonhomogeneous* and attempts to perform global (or at least less localized) image analysis. More specifically, the proposed model is aimed at modeling images that result from continuous coordinate transformations of homogeneous images. Since 2-D continuous functions can be approximated by 2-D polynomials, we will study a

B. Friedlander is with the Department of Electrical and Computer Engineering, University of California at Davis, Davis, CA 95616 USA.

J. M. Francos is with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel.

model consisting of a sine (or cosine) of a polynomial function of the image coordinates. In the special case of a first-order polynomial, this reduces to a homogeneous model—a simple 2-D sinusoid. When the polynomial order is higher, the model is no longer homogeneous: The spatial frequencies are now a function of location. This type of model arises, for example, when a homogeneous image consisting of a periodic structure undergoes distortion due to perspective. Using the 2-D Wold decomposition, it is shown in [3] that an approximate model for homogeneous textures is a sum of harmonic components in additive noise. Hence, in general, the deterministic component [3] of a homogeneous texture undergoing nonlinear continuous warping can be approximated by a sum of 2-D sinusoids of a polynomial function of the image coordinates.

The proposed model belongs to the general class of AM–FM signals that has been recently investigated both for 1-D and 2-D signals using nonparametric methods [8]–[11]. It is shown that the Teager–Kaiser energy operator can be used to approximately estimate the amplitude envelope of the AM component as well as the instantaneous frequency of the FM component. However, using this method, an approximation error exists even when no observation noise is present. The estimation algorithm of 2-D multicomponent AM–FM signals [11] initially uses multiband bank of Gabor wavelets to isolate the different components, thus avoiding the interference between the various components and increasing the effective SNR. The estimation of the AM and FM parts of each component follows in the next stage.

For reasons that will become clear later, it is more convenient to work with a complex valued model in which the sinusoidal function is replaced by a complex exponential. In some applications, such as synthetic aperture radar imaging, the 2-D signal is complex valued to begin with. In other applications, the 2-D signal is real but can be converted subject to some restrictive conditions into complex form through the Hilbert Transform [2].

Throughout this paper, we consider 2-D signals that can be represented by a constant amplitude complex exponential whose phase is a polynomial function of the coordinates. Having defined the model, we study the problem of estimating its parameters given observations on the 2-D signal. In the presence of additive white Gaussian noise, a straightforward but computationally prohibitive approach is to develop a maximum likelihood estimator for the polynomial phase parameters. This estimator involves a multidimensional search in the parameter space and is not practical except for very low order models. Here, we present a suboptimal but computationally efficient algorithm for estimating the parameters of 2-D constant amplitude polynomial phase signals. This algorithm is an extension of the so-called polynomial phase transform, which was introduced in [1]. The algorithm is based on the properties of a 2-D polynomial phase difference operator, which is defined in the next section.

The paper is organized as follows. In Section II, we define the parametric model of the observed signal, define the 2-D polynomial phase difference operator, and present some properties of the operator. In Section III, we present the proposed parameter estimation algorithm that is based on the 2-D polynomial phase difference operator and its properties. We then illustrate the algorithm operation using a numerical example.

II. THE PHASE DIFFERENCE OPERATOR

In this section, we define the phase difference operator and present some of its basic properties. We start with a description of the type of signal for which the operator was designed.