

# ANISOTROPIC LOCAL APPROXIMATIONS FOR POINTWISE ADAPTIVE SIGNAL-DEPENDENT NOISE REMOVAL

Alessandro Foi<sup>a,c</sup>, Radu Bilcu<sup>b</sup>, Vladimir Katkovnik<sup>a\*</sup> and Karen Egiazarian<sup>a</sup>

<sup>a</sup>Institute of Signal Processing, Tampere University of Technology, Finland - firstname.lastname@tut.fi

<sup>b</sup>Nokia Research Center, Tampere, Finland - firstname.lastname@nokia.com

<sup>c</sup>Dipartimento di Matematica, Politecnico di Milano, Italy - firstname.lastname@polimi.it

## ABSTRACT

We present new methods for pointwise spatially-adaptive filtering of anisotropic multivariable signals. It is assumed that the observations are given by a broad class of models with a signal-dependent variance. The proposed methods are based on the local quasi-likelihood, incorporating the directional-windowed local polynomial approximations (*LPA*) of the signal. The intersection of confidence intervals (*ICI*) rule is used in order to determine the adaptive size of the directional windows. In this way we obtain multi-directional estimates which are spatially adaptive to unknown smoothness and anisotropy of the signal. Simulation experiments confirm the advanced performance of these new algorithms.

## 1. INTRODUCTION

In many applications the observed signal is corrupted by a signal-dependent noise. The most widely encountered models are Poisson, film-grain, multiplicative and speckle noise. Their common feature is that the variance of the noise is directly related to the signal. There are a number of adaptive approaches to this sort of observations based on local-statistics' calculation. In particular the filters by Lee [12, 13] and Kuan [11] are well known in the field.

Our approach is based on the *ICI* rule for pointwise adaptive estimation. Originally, the method has been developed for 1D signals [3, 7]. The idea was generalized for 2D image processing, where adaptive-size quadrant windows have been used [8]. Significant improvement of this approach was achieved on the base of anisotropic directional estimation [9, 2]. Multidirectional sectorial-neighborhood estimates are calculated for every point and the *ICI* rule is exploited for the adaptive selection of the size of each sector; as a result, the estimator is anisotropic and its support can have quite an exotic shape. In Figure 1 we show some examples of these anisotropic neighborhoods for the *Lena* and *Cameraman* images. The developed anisotropic estimates are highly sensitive with respect to change-points, and allow to reveal fine elements of images from noisy observations. In [10] we developed this idea for non-Gaussian observations using maximum likelihood (*ML*) estimates.

The contribution of this paper concerns the quasi-likelihood (*QL*) approach, which allows to extend our strategy to a much wider class of non-Gaussian models.

### 1.1 Quasi-likelihood

There are many practical circumstances in which even though the full likelihood is unknown, one can still specify the relationship between the mean and the variance. In this situation, estimation of the mean intensity (regression, local regression) can be achieved by replacing the conditional log-likelihood,  $\ln L(z, y)$ , by a quasi-likelihood function  $Q(z, y)$ . The *QL* has been proposed in [17] as a distribution-free method (note that contrary to it, the maximum likelihood (*ML*) is a distribution-based method).

#### General approach

Let us consider the observations  $z(x)$ ,  $x \in \mathbb{R}^d$ , with the expectations  $E\{z(x)\} = y(x)$ , where the errors  $\eta(x) = z(x) - y(x)$  are indepen-

\* The work of Dr. Katkovnik was supported in part by a *Visiting Fellow* grant from Nokia Foundation.

$\rho(y)$	$Q(z, y)$	Noise model
$\sigma^2$	$-(z - y)^2 / 2\sigma^2$	Gaussian
$y/\chi, y > 0$	$(z \ln y - y) \chi, z > 0$	Poissonian
$y(1 - y), 0 < y < 1$	$z \ln \frac{y}{1-y} + \ln(1 - y), 0 \leq z \leq 1$	Bernoulli
$y^2, y > 0$	$-z/y - \ln y, z > 0$	Gamma
$K^2 y^{2\alpha}, y > 0, \alpha \neq 1/2, 1$	$\frac{y^{(1-2\alpha)}}{K^2} (\frac{z}{1-2\alpha} - \frac{y}{2-2\alpha}), z \geq 0$	Film-grain

Table 1: Examples of quasi-likelihoods for some common variance functions  $\rho(y)$ .

dent and the variance of these observations is modeled as

$$\sigma_z^2(x) = \text{var}\{z(x)\} = \text{var}\{\eta(x)\} = \rho(y(x)), \quad (1)$$

$\rho$  being a given positive function of  $y$  called the *variance function*. For example, we have  $\rho(y) = (1 - y)y$  and  $\rho(y) = y$  for the Bernoulli and Poisson models respectively. The problem is to reconstruct the signal  $y$  from the noisy observations  $z$ .

The spatially-varying variance (1) gives the idea that the residuals in estimation should be weighted by these variances. The quasi-likelihood function  $Q(z, y)$  is defined by the relation [17]:

$$\partial_y Q(z, y) = \frac{z - y}{\rho(y)}. \quad (2)$$

Naturally,  $Q(z, y) = \int_{-\infty}^y \frac{z-u}{\rho(u)} du + f(z)$  (where  $f$  is a function of  $z$ , independent of  $y$ ), but practically, only the score function  $\partial_y Q(z, y)$  is used in estimation and we do not need to know the quasi-likelihood  $Q(z, y)$ .

Examples of  $Q(z, y)$  obtained according to (2) and associated with the variance function  $\rho$  in (1) are shown in Table 1. In general the *QL* is developed with connection to a class of exponential distributions and allows to estimate  $y$  when the variance of observations is a given linear or nonlinear function of  $y$ . One may note that the *QL* from the first four lines of Table 1 coincide with the log-likelihoods of the respective models. In fact, the *QL* and the likelihood are equivalent (i.e.  $\ln L(z, y) = Q(z, y)$ ) for the one-parameter exponential class of distributions, provided the corresponding definition of the variance function  $\rho(y)$  in (2). More details and further references on the quasi-likelihood can be found in [14, 17, 4, 5, 1].

#### Nonparametric quasi-likelihood

A local version of the quasi-likelihood is defined using the window function  $w$  as weights of the residuals. Then, the local polynomial approximation (*LPA*) [1] model  $y_h(x, x_s)$  of the estimated  $y$  is defined by the equation

$$\sum_s \partial_C Q(z(x_s), y_h(x, x_s)) w_h(x - x_s) = 0, \quad (3)$$

where  $y_h(x, x_s) = C^T \phi_h(x - x_s)$  is a *LPA* of  $y$  in a neighborhood of  $x$ ,  $\phi$  is a polynomial vector,  $\phi_h(x) = \phi(x/h)$ . The scaled window is  $w_h(x) = w(x/h)/h^d$  with the parameter  $h$  defining the size of the window and the scale of the estimate. When the estimate  $\hat{C}$  of  $C$  is found from (3), it is used for estimation of  $y$  at the point  $x$  as

$$\hat{y}_h(x) = \hat{C}^T \phi_h(0). \quad (4)$$

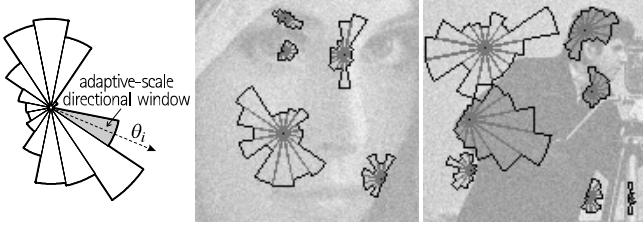


Figure 1: Anisotropic local approximations achieved by combining a number of adaptive-scale directional windows. The examples show some of these windows selected by the directional LPA-ICI for the noisy *Lena* and *Cameraman* images.

We do not need to know the  $QL$  function  $Q$  to solve (3), because

$$\begin{aligned} \partial_C Q(z(x_s), y_h(x, x_s)) &= \partial_{y_h} Q(z(x_s), y_h(x, x_s)) \partial_C y_h(x, x_s) = \\ &= \partial_{y_h} Q(z(x_s), y_h(x, x_s)) \phi_h(x - x_s), \end{aligned}$$

and  $\partial_y Q(z, y)$  is defined in (2). It gives the local  $QL$  equations as

$$\sum_S \frac{z(x_s) - y_h(x, x_s)}{\rho(y_h(x, x_s))} w_h(x - x_s) \phi_h(x - x_s) = 0, \quad (5)$$

or in the equivalent form

$$\sum_S \frac{w_h(x - x_s) \phi_h(x - x_s) \phi_h^T(x - x_s)}{\rho(y_h(x, x_s))} \hat{C} = \sum_S \frac{z(x_s) w_h(x - x_s) \phi_h(x - x_s)}{\rho(y_h(x, x_s))}. \quad (6)$$

This equation is non-linear on  $C$  as the weights  $1/\rho$  depend on it. Consequently, the estimate of  $C$ , as well as of the signal  $\hat{y}_h$ , are nonlinear functions of the observations. However, the equation (6) is linear provided a fixed value of the weights. This linearity is exploited in the iteratively reweighted least-squares method [16] - conventionally applied to this sort of estimates - as well as in our algorithms.

The local  $QL$  estimates (4)-(6) can be used with an  $LPA$  of any order and for any window  $w_h$  compatible with this order. Thus, it is quite a general construction applicable to observations characterized by a signal dependent variance.

A theoretical study of the accuracy of the  $QL$  estimates has been done. It is shown that the  $ICI$  rule is applicable as an adaptive-scale selector for the  $QL$  estimates.

## 2. $QL$ ANISOTROPIC LPA-ICI ALGORITHMS

We consider the  $ICI$  adaptive-scale estimates separately for the zero- and higher-order  $LPA$ . As it is shown in what follows, it corresponds to a solution of (5) that is, respectively, shift-invariant and shift-variant. The higher-order  $LPA$  obviously enables a better approximation of the signal, however, because of the increased complexity due to the shift-variant estimators, it may not be suitable for practical applications. On the other hand, the simple zero-order (locally constant)  $LPA$  is rather limiting when restoring signals that are not piecewise-constant. We propose an efficient "order-mixture" algorithm, which to some extent retains the advantages and minimizes drawbacks of the straightforward zero- and higher-order  $LPA$ .

The diagram in Figure 2 shows the layout of all three types of considered algorithms: zero order, higher order, and order-mixture.

### 2.1 Shift-invariant zero-order estimators: LPA-ICI-AV

The  $LPA-ICI-AV$  (adaptive variance) algorithm consists of the following four steps.

#### (I) CALCULATION OF THE ESTIMATE AND ITS VARIANCE

Let us use the quasi-likelihood equation (6) for zero-order  $LPA$ . It means that the estimate  $\hat{y}_h(x, x_s) = \hat{C}$  is constant. Then, the variance of the observations  $\rho(y_h(x, x_s))$  is also constant and the solution of (5) can be given in the explicit form as the weighted mean

$$\hat{y}_h(x) = \hat{C} = \sum_S z(x_s) g_h(x - x_s) = (\hat{y}_h \otimes g_h)(x), \quad g_h(\cdot) = \frac{w_h(\cdot)}{\sum_S w_h(x_s)}. \quad (7)$$

The above formulas define the conventional Nadaraya-Watson convolution estimate. Its variance is calculated, using  $\rho(\hat{y}_h(x))$  as

$$\sigma_{\hat{y}_h}^2(x) = \sum_S \rho(\hat{y}_h(x_s)) g_h^2(x - x_s) = (\rho(\hat{y}_h) \otimes g_h^2)(x). \quad (8)$$

#### (II) ADAPTIVE SCALE SELECTION

The estimates  $\hat{y}_h(x)$  in (7) and the corresponding variances  $\sigma_{\hat{y}_h}^2(x)$  in (8) are calculated for a set  $H = \{h_j\}_{j=1}^J$  of increasing scales. The  $ICI$  rule [3, 7] is exploited in order to obtain a pointwise adaptive estimate  $\hat{y}_{h^+}(x)$ , where for every point  $x$  an adaptive scale  $h^+(x) \in H$  is used. The  $ICI$  rule is as follows. Consider the intersection of confidence intervals  $\mathcal{I}_j = \bigcap_{i=1}^j \mathcal{D}_i$ , where  $\mathcal{D}_i = [\hat{y}_{h_i}(x) - \Gamma \sigma_{\hat{y}_{h_i}}(x), \hat{y}_{h_i}(x) + \Gamma \sigma_{\hat{y}_{h_i}}(x)]$  and  $\Gamma > 0$  is a threshold parameter, and let  $j^+$  be the largest of the indexes  $j$  for which  $\mathcal{I}_j$  is non-empty,  $\mathcal{I}_{j^+} \neq \emptyset$  and  $\mathcal{I}_{j^++1} = \emptyset$ . Then, the adaptive scale  $h^+$  is defined as  $h^+ = h_{j^+}$  and the adaptive estimate is  $\hat{y}_{h^+}(x)$ .

#### (III) FUSING OF MULTIPLE DIRECTIONAL ESTIMATES

The anisotropy of the  $QL$  estimates is enabled by the convex fusing [9, 2] of a number of adaptive estimates obtained, according to the above procedure, using windows oriented in different directions. Figure 1 shows the conical supports of such directional windows.

Let us use the notation  $\hat{y}_{\theta_k}^+(x)$  to indicate the adaptive estimate  $\hat{y}_{h^+}(x)$  selected by the  $ICI$  rule (step II) from a set  $\{\hat{y}_h(x)\}_{h \in H}$  of estimates obtained (step I) using directional windows oriented towards the direction  $\theta_k$ ,  $k = 1, \dots, N$ . Thus, we have  $N$  directional adaptive estimates,  $\{\hat{y}_{\theta_k}^+(x)\}_{k=1}^N$ . These directional adaptive estimates are fused together into the anisotropic estimate  $\hat{y}(x)$ :

$$\hat{y}(x) = \sum_k \lambda_k(x) \hat{y}_{\theta_k}^+(x), \quad \lambda_k(x) = \sigma_{\hat{y}_{\theta_k}^+}^{-2}(x) / \sum_i \sigma_{\hat{y}_{\theta_i}^+}^{-2}(x). \quad (9)$$

This fused estimate is a convex linear combination of the directional ones with data-driven adaptive weights (inverse-variances of the adaptive estimates). It is anisotropic because a pointwise adaptive scale  $h^+ = h^+(\theta_k)$  is used for each specified direction, as shown in Figure 1. If directional windows of distinct directions do not overlap, then the directional estimates are independent and the variance of the fused estimate (9) is  $\sigma_{\hat{y}}^2(x) = 1 / \sum_i \sigma_{\hat{y}_{\theta_i}^+}^{-2}(x)$ .

#### (IV) VARIANCE UPDATE

The anisotropic estimate  $\hat{y}(x)$  is used for updating the variances  $\sigma_{\hat{y}_h}^2$  of the individual estimates  $\hat{y}_h$  through the variance function  $\rho$ . Insert  $\hat{y}(x)$  in the variance formula (8),

$$\sigma_{\hat{y}_h}^2(x) = \sum_S \rho(\hat{y}(x_s)) g_h^2(x - x_s) = (\rho(\hat{y}) \otimes g_h^2)(x). \quad (10)$$

and repeat the  $ICI$  scale selection (II) and the fusing (III) with these adjusted values of the variance. This recursive-update procedure is iterated a few times, up to numerical convergence of the estimates.

In (8) we use the variances  $\rho(\hat{y}_h(x_s))$  obtained for the scale  $h$  whereas in (10) we use the variances  $\rho(\hat{y}(x_s))$  obtained from the fused estimate (9). Note that even the simple  $\rho(z(x_s))$  can be a satisfactory initialization for the recursive algorithm, because these approximative variances will eventually be updated.

This adaptation of the variance concerns essentially only the scale selection and the same set of the estimates  $\{\hat{y}_h(x)\}_{h \in H}$  (7) calculated at the first iteration, is used for the following iterations.

### 2.2 Shift-variant higher-order estimators: LPA-ICI-AVW

A similar, but more sophisticated approach is performed by the  $LPA-ICI-AVW$  (adaptive variance and weights) algorithm. It differs from the previous one in the way the estimates of  $C$  are calculated. Here we do not assume that the observations' variance is constant within the window. Instead, an estimate of the signal is used in order to define - through the variance function - the variance of the observations. As result, the nonparametric estimate (6) penalizes those observations  $z(x_s)$  that have a higher variance. Since the  $QL$  denominator in (6) is varying, the resulting estimator is no longer shift-invariant. The  $LPA-ICI-AVW$  algorithm is as follows.

#### (I) CALCULATION OF THE ESTIMATE AND ITS VARIANCE

Let some estimates  $\hat{y}(x_s)$  of the signal be known, then  $\rho(y_h(x, x_s))$  in the  $QL$  equation (6) can be replaced by  $\rho(\hat{y}(x_s))$

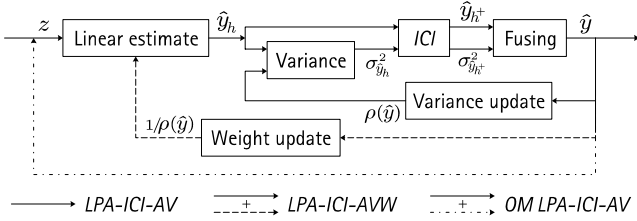


Figure 2: Layout of the *QL LPA-ICI* algorithms. The adaptive-variance algorithm (*LPA-ICI-AV*) is shown by the solid lines. The dashed lines correspond to the additional weights’ adaptation (*LPA-ICI-AVW*), whereas the dash-dot line is the recursive filtering in which the fused estimate is used as the observation data for the following iteration (order-mixture *LPA-ICI-AV* algorithm).

and the weights in these linear estimates are defined. We obtain the higher-order estimates in the kernel form

$$\hat{y}_h(x) = \sum_s z(x_s) g_{h,\rho}(x, x_s), \quad (11)$$

with the shift-variant  $g_{h,\rho}$  and the variance of  $\hat{y}_h(x)$  calculated as

$$g_{h,\rho}(x, x_s) = \phi_h^T(0) \Phi^{-1} \frac{w_h(x-x_s)}{\rho(\hat{y}(x_s))} \phi_h(x-x_s), \quad (12)$$

$$\Phi = \sum_s \frac{w_h(x-x_s)}{\rho(\hat{y}(x_s))} \phi_h(x-x_s) \phi_h^T(x-x_s),$$

$$\hat{\sigma}_{\hat{y}_h}^2(x) = \sum_s \rho(\hat{y}(x_s)) g_{h,\rho}^2(x, x_s). \quad (13)$$

Varying weights  $1/\rho(\hat{y}(x_s))$  in (12) may significantly improve the estimation. However, convolutions cannot be used anymore for calculations of (11), as the kernels  $g_{h,\rho}$  are not shift-invariant. The corresponding estimates have to be computed for each  $x$  in a point-wise adaptive manner. The computation becomes quite tedious.

The initial estimate  $\hat{y}(x_s)$  for  $\rho(\hat{y}(x_s))$  used above can be obtained either from the *LPA-ICI-AV* algorithm or by the iteratively reweighted least-squares method [16].

**(II)-(III) ARE EXACTLY AS IN THE *LPA-ICI-AV* ALGORITHM**  
**(IV) VARIANCE AND WEIGHTS UPDATE**

The fused estimate  $\hat{y}$  is used to update the variance  $\rho(\hat{y}(x_s))$ . This updated variance is recursively used in a number of places. First, it defines new weights for the kernel  $g_{h,\rho}$  (12), resulting in an update of the estimate  $\hat{y}_h(x)$  in (11). Then, the new  $\rho(\hat{y}(x_s))$  is used in (13) for the calculation of the variance of the estimate  $\hat{y}_h(x)$  from (11). The term *adaptive variance and weights (AVW)* is used for this recursive procedure as we use the fused improved estimate - obtained at the stage (III) of every iteration - to update the weights in (12) as well as the variance in (13). We remark again that, contrary to the *LPA-ICI-AV*, the kernels  $g_{h,\rho}$  used in the above estimates are not shift-invariant.

### 2.3 Order-mixture shift-invariant estimators: *OM LPA-ICI-AV*

With the intention to combine the improved accuracy of the higher-order *LPA* with the computational efficiency of the shift-invariant estimators, we developed a combined scheme.

The algorithm uses convex mixtures  $g_h^\lambda$  of zero-order *LPA* kernels  $g_h^0$  (7) and first-order *LPA* kernels  $g_h^1$  (as in equation (12), but assuming constant  $\rho$ ):  $g_h^\lambda = (1-\lambda)g_h^0 + \lambda g_h^1$ . These mixtures allow to achieve a better fit of the data but, contrary to the pure higher-order polynomial kernels, they yield estimates with a smaller range of values. To be precise, the range of  $\hat{y}_h^\lambda$ ,  $\hat{y}_h^\lambda = z \otimes g_h^\lambda$ , is  $\lambda$ -times the range of  $\hat{y}_h^1$ . This allows to reasonably enable the assumption that the variance  $\rho(y_h)$  is constant within the support of the window  $w_h$ . Thus, similarly to what was done at the beginning of section 2.1, the non-parametric *QL* estimates are obtained by a shift-invariant kernel: it corresponds to the convolution  $\hat{y}_h^\lambda = z \otimes g_h^\lambda$ .

Because of these special kernels, we call this third algorithm an *order-mixture (OM) LPA-ICI-AV* algorithm. It follows the layout of the zero-order *LPA-ICI-AV* algorithm (shown as the solid line in Figure 2), using convolutional estimates (I), *ICI* adaptive-scale

	test image	noisy	Lee	<i>NURW</i>	<i>ANF</i>	<i>OM LPA-ICI-AV</i>
<i>Poisson</i>	<i>Lena</i> 512×512	1240	—	—	—	82 (11.8)
	<i>Peppers</i> 512×512	1197	—	—	—	79 (11.8)
	<i>Lena</i> 256×256	1239	200	177	151	120 (10.1)
<i>Film-grain</i>	<i>Peppers</i> 256×256	1206	184	160	145	120 (10.0)
	<i>Aerial</i> 256×256	766	231	252	179	179 (6.3)
	<i>Lena</i> 512×512	1343	—	—	—	83 (12.1)
<i>Speckle</i>	<i>Peppers</i> 512×512	1304	—	—	—	80 (12.1)
	<i>Lena</i> 256×256	1346	206	185	160	125 (10.3)
	<i>Peppers</i> 256×256	1311	199	169	150	120 (10.4)
	<i>Aerial</i> 256×256	828	242	267	188	185 (6.5)
	<i>Lena</i> 512×512	4375	—	—	—	196 (13.5)
	<i>Peppers</i> 512×512	4303	—	—	—	182 (13.7)
	<i>Lena</i> 256×256	4349	365	371	381	269 (12.1)
	<i>Peppers</i> 256×256	4304	370	372	378	269 (12.0)
	<i>Aerial</i> 256×256	1707	348	387	318	329 (7.1)

Table 2: *MSE* values for different images, noise models, and methods. In the last column, the value in parentheses is the *ISNR* (dB).

selection (II), fusing (III) and variance update recursion (IV). Further, to achieve an effect similar to the adaptive weighting (12), we rely (after numerical convergence of the variance) on an additional recursive adaptive filtering of the fused anisotropic estimates [2], shown as the dash-dot line in Figure 2. Indeed, the fusing formula (9) uses the variance of its addends as denominators, exactly as in (12). The only difference between the original variance-update iteration (IV) and this recursive filtering (which is performed after the numerical convergence of the variance-update iterations), is that, instead of the pair  $z$  and  $\sigma_z^2 = \rho(y)$ , the input variables for the following iteration’s stage (I) are the fused estimate  $\hat{y}$  and its variance  $\sigma_{\hat{y}}^2$ , respectively.

### 3. NUMERICAL EXPERIMENTS

As an illustration of the potential of the proposed approach, we present some experimental results for three common types of signal-dependant noise: the “scaled” *Poisson* noise,  $z \sim \mathcal{P}(\chi y)/\chi$ ,  $\chi \in \mathbb{R}^+$ , the *film-grain* noise,  $z = y + Ky^\alpha \eta$ ,  $K, \alpha \in \mathbb{R}^+$  and  $\eta \sim \mathcal{N}(0, 1)$ , and the “multiple-look” *speckle* noise,  $z = L^{-1} \sum_{i=1}^L y \epsilon_i$ ,  $\epsilon_i \sim \mathcal{E}(\beta)$ ,  $\beta \in \mathbb{R}^+$ . The calligraphic letters  $\mathcal{P}$ ,  $\mathcal{N}$ , and  $\mathcal{E}$  denote, respectively, the Poisson, Gaussian, and exponential distributions. For the above observation models, the variance functions  $\rho(y) = \sigma_z^2$  are  $\rho(y) = y/\chi$ ,  $\rho(y) = K^2 y^{2\alpha}$ , and  $\rho(y) = y^2 \beta/L$ , respectively.

To enable an objective comparison with the many simulations presented in [15], we set  $\chi = 0.1$ ,  $K = 3.3$ ,  $\alpha = 0.5$ ,  $L = 4$ , and  $\beta = 1$ . The true signal  $y$  is assumed to have range [0,255]. Note that in the quasi-likelihood approach, the Poissonian and the film-grain observations with  $\alpha = 0.5$  are treated identically (up to a multiplicative factor). Nevertheless, even when  $K^2 = 1/\chi$  (i.e. when their corresponding variance functions coincide), their corresponding observations are quite different, because of the different distributions. In particular, Poissonian observations are always integer and positive, i.e.  $z \in \mathbb{N}/\lambda$ , whereas Gaussian distributed observations can take any real value.

In [15], where the main focus is on the adaptive-neighborhood filter (*ANF*) (a technique which - like ours - is based on anisotropic adaptation), are also considered the “refined” Lee filter (Lee) [13] and the noise-updating repeated Wiener filter (*NURW*) [6]. Table 2 includes the results from [15] and extends them with those obtained by the proposed *OM LPA-ICI-AV* method. Comparing the *MSE* values given in Table 2, we may note that for the *Lena* and *Pepper* images the developed algorithm gives essentially better results for all types of noise. For the *Aerial* image we obtain figures which are very close to the best, given by *ANF* algorithm. An illustration of some of these results, attesting the advanced filtering performance of our method, is given in Figure 3.

We also show some visual results obtained using real data ac-

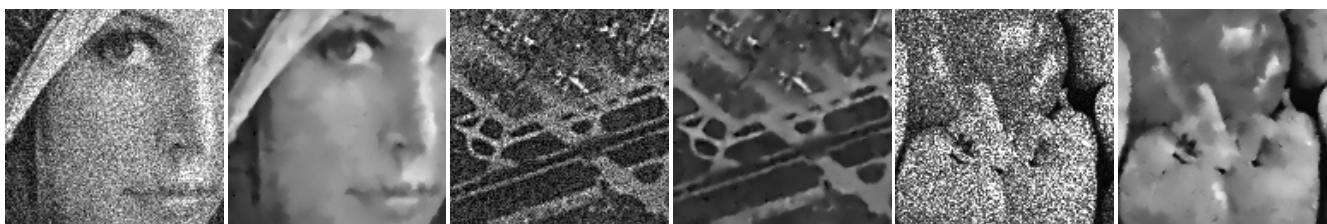


Figure 3: Fragments of the noisy and restored images: (from left to right) Poisson *Lena*, film-grain *Aerial*, and speckle *Peppers* 256×256.



Figure 4: Raw data from cameraphone's CMOS sensor, R channel, 1ms exposure (left), and reconstructed image using the *LPA-ICI* adaptive method with the estimated variance function.

quired using the CMOS sensor of a Nokia cameraphone. The statistical characteristics of the sensor's raw-data have been studied, and were found to follow very accurately the observation model (1). The corresponding variance function  $\rho(y)$  has been estimated and used in the algorithm. In extreme low-light conditions, or for extremely short exposure-times, the signal-to-noise ratio can be dramatically low. Figures 4(left) and 5(top) show, respectively, the raw data captured in dim light with an exposure time of 1ms, and the reconstructed color-image using the full image-processing chain (which includes white-balance, color correction, some denoising and the color-array interpolation). Figures 4(right) and 5(bottom) show the corresponding results obtained when the raw data is filtered by the adaptive *LPA-ICI* method using the estimated variance function  $\rho(y)$ : smooth areas are faithfully restored and finer details are accurately preserved.

All experiments presented in this paper were produced using the same algorithm parameters. In particular, in our implementation, the variance update is performed three times and the recursive adaptive filtering is repeated twice. A set  $H$  of seven scales is used, and the anisotropic estimates are obtained by fusing eight directional adaptive estimates.

## REFERENCES

- [1] Fan J. and I. Gijbels, *Local polynomial modelling and its application*. London, Chapman and Hall, 1996.
- [2] Foi, A., V. Katkovnik, K. Egiazarian, and J. Astola, "A novel anisotropic local polynomial estimator based on directional multiscale optimizations", *Proc. 6th IMA Int. Conf. Math. in Signal Processing*, Cirencester (UK), pp.79-82, 2004.
- [3] Goldenshluger, A., and A. Nemirovski, "On spatial adaptive estimation of nonparametric regression", *Math. Meth. Statistics*, vol.6, pp.135-170, 1997.
- [4] Hastie T.J. and R.J. Tibshirani. *Generalized linear models*, London, Chapman and Hall, 1990.
- [5] Heyde, C.C., *Quasi-likelihood and its applications*, New York, Springer, 1997.
- [6] Jiang S.S., and A.A. Sawchuk, "Noise updating repeated Wiener filter and other adaptive noise smoothing filters using local image statistics", *Appl. Opt.*, vol.25, pp.2326-2337, 1986.
- [7] Katkovnik V., "A new method for varying adaptive bandwidth selection", *IEEE Trans. on Signal Proc.*, vol.47, no.9, pp.2567-2571, 1999.
- [8] Katkovnik V., K. Egiazarian, and J. Astola, "Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule", *J. of Math. Imaging and Vision*, vol.16, no.3, pp.223-235, 2002.
- [9] Katkovnik V., A. Foi, K. Egiazarian, and J. Astola, "Directional varying scale approximations for anisotropic signal processing", *Proc. of EUSIPCO 2004*, pp.101-104, 2004.
- [10] Katkovnik V., A. Foi, K. Egiazarian, and J. Astola, "Anisotropic local likelihood approximations", *Proc. of Electronic Imaging 2005*, 5672-19, 2005.
- [11] Kuan D.T., A.A. Sawchuk, T.C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal dependent noise", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.7, pp.165-177, 1985.
- [12] Lee J.S., "Digital image enhancement and noise filtering by using local statistics", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.2, 1980.
- [13] Lee J.S., "Refined filtering of image noise using local statistics", *Comput. Graph. Image Proc.* vol.15, pp.380-389, 1981.
- [14] McCullagh P. and J.A. Nelder, *Generalized linear models*, (2nd ed.), London, Chapman and Hall, 1999.
- [15] Rangarayanan R.M., M. Ciuc, and F. Faghieh, "Adaptive-neighborhood filtering of images corrupted by signal-dependent noise", *Appl. Opt.*, vol.37, pp.4477-4487, 1998.
- [16] Seber G.A. and C.J. Wild, *Nonlinear regression*, New York, Wiley, 1989.
- [17] Wedderburn R.W.M., "Quasilielihood functions, generalized linear models and the Gauss-Newton method", *Biometrika*, vol.61, pp.439-447, 1974.

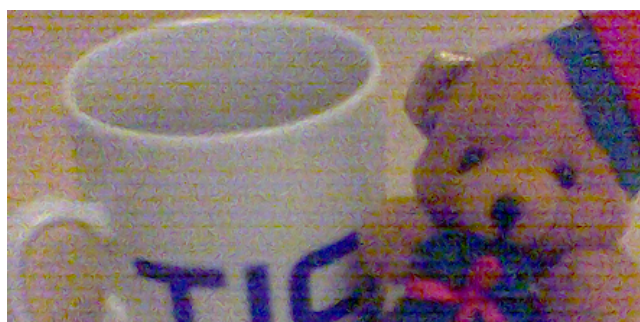


Figure 5: Color image reconstructed, using the standard imaging chain, from the noisy raw data (top) and from the *LPA-ICI*-filtered raw data (bottom).