

Fast Global Image Denoising Algorithm on the Basis of Nonstationary Gamma-Normal Statistical Model

Inessa Gracheva^(✉), Andrey Kopylov, and Olga Krasotkina

Tula State University, Tula, Russian Federation
gia1509@mail.ru, And.Kopylov@gmail.com, ko180177@yandex.ru

Abstract. We consider here a Bayesian framework and the respective global algorithm for adaptive image denoising which preserves essential local peculiarities in basically smooth changing of intensity of reconstructed image. The algorithm is based on the special nonstationary gamma-normal statistical model and can handle both Gaussian noise, which is an ubiquitous model in the context of statistical image restoration, and Poissonian noise, which is the most common model for low-intensity imaging used in biomedical imaging. The algorithm being proposed is simple in tuning and has linear computation complexity with respect to the number of image elements so as to be able to process large data sets in a minimal time.

Keywords: Image denoising · Bayesian framework · Nonstationary gamma-normal statistical model

1 Introduction

Image analysis and information extraction are often used in many imaging systems. General degradation of image quality during acquisition and transmission process in many cases is described by an additive Gaussian noise model. At the same time, variations in low-intensity levels in quantum noise caused by fluctuations in either the number of detected photons or the inherent limitation of the discrete nature for photon detection can be described by a Poissonian noise model, which is the most popular in biomicroscopy. The primary aim of the image denoising technique is to remove the noisy observations while reconstructing a satisfying estimation of the original image, thereby enhancing the performance of these applications in imaging systems.

The most popular methods are good in case when we have only one type of noise. For example, for denoising of Gaussian used the Steins unbiased risk estimate from the concept of linear expansion of thresholds (SURE-LET) [1], where directly the denoising process is parametrized as a sum of elementary nonlinear processes with unknown weights; block-matching 3-D (BM3D) algorithm [2] based on enhanced sparse representation in transform domain; fast bilateral filter (FBF) [3] was first termed by Tomasi and Manduchi [4] based on

the work [5, 6], and later modified and improved in [7]. A widespread alternative to the direct handling of Poisson statistics is to apply variance-stabilizing transforms (VSTs) with the underlying idea of exploiting the broad class of denoising methods which are based on a Gaussian noise model [8], such as the HaarFisz transform [9]. Platelet approach [10], which stands among the state-of-the-art algorithms for Poisson intensity estimation. Poisson unbiased risk estimate from the concept of linear expansion of thresholds (PURE-LET) [11] is based on the minimization of an unbiased estimate of the MSE for Poisson noise, a linear parametrization of the denoising process and the preservation of Poisson statistics across scales within the Haar VST.

The purpose of this article is to present an algorithm that would satisfy compromise between restoration quality and computational complexity. Firstly, we want a method which is able to effectively remove Gaussian noise as well as Poissonian noise. In the second place, the method should satisfy strict constraints in terms of computational cost and memory requirements, so as to be able to process large data sets. To achieve these purpose a nonstationary gamma-normal noise model has been proposed in the framework of Bayesian approach to the problem of image processing. This model allows us to develop a fast global algorithm on the basis of Gauss-Seudel procedure and Kalman filter-interpolator.

2 Gamma-Normal Model of Hidden Random Field

The task of image reconstruction within the Bayesian approach can be expressed as the problem of estimating a hidden Markov component $X = (x_t, t = 1, \dots, N)$ $T = \{t = (t_1, t_2) : t_1 = 1, \dots, N_1, t_2 = 1, \dots, N_2\}$ of a two-component random field, where observed component $Y = (y_t, t \in T)$ is the analyzed image.

Probabilistic properties of a two-component random field (X, Y) are completely determined by the joint conditional probability density $\Phi(Y|X, \delta)$ of original functions $Y = (y_t, t \in T)$ with respect to the secondary data $X = (x_t, t \in T)$, and the a prior joint distribution $\Psi(X|\Lambda, \delta)$ of hidden component $X = (x_t, t \in T)$.

Let the joint conditional probability density $\Phi(Y|X, \delta)$ be in the form of Gaussian distribution:

$$\Phi(Y|X, \delta) = \frac{1}{\delta^{N/2} (2\pi)^{N/2}} \exp\left(-\frac{1}{2\delta} \sum_{t \in T} (y_t - x_t)^2\right), \quad (1)$$

where $E(e_t^2) = \delta$ is the variance of the observation noise, which to be unknown.

Just as in [12, 13], we will express our prior knowledge about sought for estimates of a hidden component in the form of Markov random field. Let the conditional probability densities of hidden variables with respect to their neighbors be also Gaussian with some variance $E(\xi_t^2)$ and the conditional mathematical expectation equal to the value of the adjacent variable.

In the case of Poissonian noise the variance of corresponding random variables is determined by the intensity of photons emission and takes different values in different parts of an image. We do not use here VST to reforms the data so

that the noise approximately becomes Gaussian with a constant variance. On the contrary, we do not assume the common variance of hidden components to remain the same in an image plane, and suppose $E(\xi_t^2) = r_t$. The unknown variances ($r_t, t \in T$) are considered as proportional to the variance of the observation noise $r_t = \lambda_t \delta$ with the proportionality coefficients λ_t acting as factors of the unknown instantaneous volatility of hidden field $X = (x_t, t \in T)$, unknown as well.

Under this assumption, the a priori joint distribution of $X = (x_t, t \in T)$ is conditionally normal with respect to the field of the factors $\Lambda = (\lambda_t, t \in T)$. So, we come to the improper a priori density:

$$\begin{aligned} \Psi(X|\Lambda, \delta) &\propto \frac{1}{\left(\prod_{t \in T} \delta \lambda_t\right)^{1/2} (2\pi)^{n(N-1)/2}} \\ &\times \exp\left(-\frac{1}{2} \sum_{t', t'' \in V} \frac{1}{\delta \lambda_t} (x_{t'} - x_{t''})^2\right), \end{aligned} \quad (2)$$

where V is the neighborhood graphs of image elements having the form of a lattice.

Finally, we assume the inverse factors $1/\lambda_t$ to be a priori independent and identically gamma-distributed $\gamma(1/\lambda_t|\alpha, \vartheta) \propto (1/\lambda_t)^{\alpha-1} \exp(-\vartheta(1/\lambda_t))$ on the positive half-axis $\lambda_t \geq 0$. The mathematical expectation and variance of gamma-distribution are ratios α/ϑ , and α/ϑ^2 . The a priori distribution density of the entire field of the factors:

$$\begin{aligned} G(\Lambda|\alpha, \vartheta) &= \prod_{t \in T} \gamma(\lambda_t|\alpha, \vartheta) \propto \left(\prod_{t \in T} \frac{1}{\lambda_t}\right)^{\alpha-1} \exp\left(-\vartheta \sum_{t \in T} \frac{1}{\lambda_t}\right) \\ &= \exp\left[-(\alpha-1) \sum_{t \in T} \ln \lambda_t - \vartheta \sum_{t \in T} \frac{1}{\lambda_t}\right]. \end{aligned} \quad (3)$$

We redefine the parameters α and ϑ through new parameters μ and λ

$$\alpha = \frac{1}{2} \left[\frac{1}{\delta} \left(1 + \frac{1}{\mu}\right) + 1 \right], \vartheta = \frac{\lambda}{2\delta\mu}, \quad (4)$$

assuming thereby the parametric family of gamma distributions of inverse factors $1/\lambda_t$

$$\gamma(1/\lambda_t|\delta, \lambda, \mu) \propto (1/\lambda_t)^{\frac{2\mu+1}{2\delta\mu}} \exp\left(-\frac{\lambda}{2\delta\mu}(1/\lambda_t)\right), \quad (5)$$

with mathematical expectations and variances

$$E(1/\lambda_t) = \frac{(1+\delta)\mu+1}{\lambda}, \text{Var}(1/\lambda_t) = 2\delta\mu \frac{(1+\delta)\mu+1}{\lambda^2}.$$

In terms of this parameterization, the independent prior distribution of each instantaneous inverse factors $1/\lambda_t$ is almost completely concentrated around the mathematical expectation $1/\lambda$ if $\mu \rightarrow 0$. On the contrary, with $\mu \rightarrow \infty$ coefficient $1/\lambda$ have tends to the almost uniform distribution.

In this way, we come to the prior density:

$$G(\Lambda|\delta, \lambda, \mu) = \exp \left[-\frac{1}{2\delta\mu} \sum_{t \in T} \left(\lambda \frac{1}{\lambda_t} + \frac{1}{\lambda} \ln \lambda_t \right) \right], \quad (6)$$

and, so, have completely defined the joint prior normal gamma-distribution of both hidden fields $X = (x_t, t \in T)$ and $\Lambda = (\lambda_t, t \in T)$:

$$H(X, \Lambda|\delta, \lambda, \mu) = \Psi(X|\Lambda, \delta)G(\Lambda|\delta, \lambda, \mu).$$

Coupled with the conditional density of the observable field (2), it makes basis for Bayesian estimation of the field $X = (x_t, t \in T)$.

3 The Bayesian Estimate of the Hidden Random Field

Bayesian reasoning makes it possible to reduce a wide class of image analysis problems to the problem of maximum a posteriori (MAP) estimation.

The joint a posteriori distribution of hidden elements, namely, those of field $X = (x_t, t \in T)$ and its instantaneous factors $\Lambda = (\lambda_t, t \in T)$, is completely defined by (1) and (2) and, in terms of the original parameters (α, ϑ) , by (3):

$$P(X, \Lambda|Y, \delta, \alpha, \vartheta) = \frac{\Psi(X|\Lambda, \delta)G(\Lambda|\alpha, \vartheta)\Phi(Y|X, \delta)}{\int \int \Psi(X'|\Lambda', \delta)G(\Lambda'|\alpha, \vartheta)\Phi(Y|X', \delta)dX'd\Lambda'}.$$

The Bayesian estimate of (X, Λ) is the maximum point of the numerator

$$\left\{ \begin{aligned} &(\hat{X}, \hat{\Lambda}|\delta, \alpha, \vartheta) = \arg \max_{X, \Lambda} [\ln \Phi(Y|X, \delta) + \ln \Psi(X|\Lambda, \delta) + \ln G(\Lambda|\alpha, \vartheta)] \\ &= \arg \max_{X, \Lambda} \left(-\frac{1}{2\delta} \sum_{t \in T} (y_t - x_t)^2 - \frac{1}{2} \sum_{t \in T} \ln \delta \lambda_t - \frac{1}{2\delta} \sum_{t', t'' \in V} \frac{1}{\lambda_t} (x_{t'} - x_{t''})^2 \right. \\ &\left. - (\alpha - 1) \sum_{t \in T} \ln \lambda_t - \vartheta \sum_{t \in T} \frac{1}{\lambda_t} \right), \end{aligned} \right.$$

or, what is equivalent,

$$\left\{ \begin{aligned} &(\hat{X}, \hat{\Lambda}|\delta, \alpha, \vartheta) = \arg \min_{X, \Lambda} J(X, \Lambda|Y, \delta, \alpha, \vartheta), \\ &J(X, \Lambda|Y, \delta, \alpha, \vartheta) = \sum_{t \in T} (y_t - x_t)^2 \\ &\left\{ + \sum_{t', t'' \in V} \left\{ \frac{1}{\lambda_t} [(x_{t'} - x_{t''})^2 + 2\delta\vartheta] + \delta(2\alpha - 1) \ln \lambda_t \right\} \right\}. \end{aligned} \right.$$

The substitution of the new parameters (4) makes the Bayesian estimate independent of the observation noise variance δ :

$$\left\{ \begin{array}{l} (\hat{X}, \hat{\Lambda}|\lambda, \mu) = \arg \min_{X, \Lambda} J(X, \Lambda|Y, \lambda, \mu), \\ J(X, \Lambda|Y, \lambda, \mu) = \sum_{t \in T} (y_t - x_t)^2 \\ + \sum_{t', t'' \in V} \left\{ \frac{1}{\lambda_t} [(x_{t'} - x_{t''})^2 + \lambda/\mu] + (1 + 1/\mu) \ln \lambda_t \right\}. \end{array} \right. \quad (7)$$

As it will be shown, the growing value of parameter μ endows this criterion with a pronounced tendency to keep the majority of estimated volatility factors $\hat{\lambda}_t$ close to the basic low value λ and to allow single large outliers, revealing thereby hidden events in the primarily smooth original field.

4 The Gauss-Seudel Procedure of Edge-Preserving Image Denoising

The conditionally optimal factors $\hat{\Lambda}(X, \lambda, \mu) = [\hat{\lambda}_t(X, \lambda, \mu), t \in T]$ are defined independently of each other:

$$\begin{aligned} \hat{\Lambda}(X, \lambda, \mu) &= \arg \min_{\Lambda} J(\Lambda|X, \lambda, \mu) : \\ \frac{\partial}{\partial \lambda_t} \left\{ \frac{1}{\lambda_t} [(x_{t'} - x_{t''})^2 + \lambda/\mu] + (1 + 1/\mu) \ln \lambda_t \right\} &= 0. \end{aligned} \quad (8)$$

The zero conditions for the derivatives, excluding the trivial solutions $\lambda_t \rightarrow \infty$, lead to the equalities

$$\frac{1}{\lambda_t} [(x_{t'} - x_{t''})^2 + \lambda/\mu] = (1 + 1/\mu),$$

and, hence,

$$\hat{\lambda}_t(X, \lambda, \mu) = \lambda \frac{(1/\lambda)(x_{t'} - x_{t''})^2 + 1/\mu}{1 + 1/\mu}, \quad (9)$$

Substitution of (9) into (7) gives the equivalent form, which avoids immediate finding the factors themselves:

$$\left\{ \begin{array}{l} (\hat{X}, \hat{\Lambda}|\lambda, \mu) = \arg \min_{X, \Lambda} J(X, \Lambda|Y, \lambda, \mu), \\ J(X, \Lambda|Y, \lambda, \mu) = \sum_{t \in T} (y_t - x_t)^2 \\ + \sum_{t', t'' \in V} \left\{ (1 + 1/\mu) \ln \frac{(1/\lambda)(x_{t'} - x_{t''})^2 + 1/\mu}{1 + 1/\mu} \right\}. \end{array} \right. \quad (10)$$

It is almost quadratic function in a vicinity of the zero point $(x_{t'} - x_{t''})^2 = 0$ and remains being so practically over the entire number axis if μ is small (Fig. 1).

But as μ grows, the originally quadratic penalty undergoes more and more marked effect of saturation at some distance from zero. This means that the

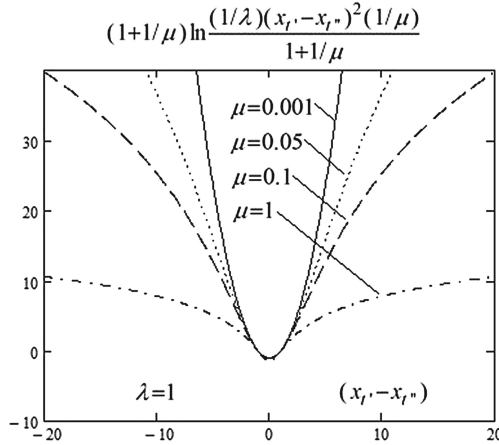


Fig. 1. The “saturation effect” of the unsmoothness penalty for sufficiently large values of the parameter μ and fixed value $\lambda=1$.

criterion strongly penalizes estimates of original function but becomes more and more indulgent to sharp discontinuities.

For finding the minimum point of the objective function with fixed structural parameters μ and λ , we apply the Gauss-Seidel iteration to both groups of variables $X = (x_t, t \in T)$ and $\Lambda = (\lambda_t, t \in T)$ starting with the initial values $\hat{\Lambda}^0 = (\hat{\lambda}_t^0 = \lambda, t \in T)$. At each iteration, current values of variables $\hat{\Lambda}^k = (\hat{\lambda}_t^k = 1, t \in T)$ according to (7) give calculate of the field $X = (x_t, t \in T)$, whose minimum point of the objective function gives the new approximation to the estimate of the mutually agreed field $\hat{X}^k = (\hat{x}_t^k, t \in T)$:

$$\begin{aligned} \hat{X}^k &= (\hat{x}^k, t \in T) = \arg \min_X J(X, \Lambda^k | Y, \lambda, \mu) \\ &= \arg \min_X \left\{ \sum_{t \in T} (y_t - x_t)^2 + \sum_{t', t'' \in V} \frac{1}{\lambda_t} (x_{t'} - x_{t''})^2 \right\}. \end{aligned} \quad (11)$$

It is clear that there is no way to replace the lattice-like neighborhood graph (Fig. 2a) by a tree-like one without loss of the crucial property to ensure smoothness of the hidden secondary data field in all directions from each point of the image plane.

To avoid this obstacle, for finding the values of the hidden field at each vertical row of the picture, we use a separate pixel neighborhood tree which is defined, nevertheless, on the whole pixel grid and has the same horizontal branches as the others (Fig. 2b). The resulting image processing procedure is aimed at finding optimal values only for the hidden variables at the stem nodes in each tree [14]. For this combination of partial pixel neighborhood trees, the algorithm of finding the optimal values of the stem node variables boils down into a combination of two usual Kalman filtration-interpolation procedures, each

dealing with single, respectively, horizontal and vertical image rows considered as signals on the one-dimensional argument axis.

First, such a one-dimensional procedure is applied to the horizontal rows $t_1 = 1, \dots, N_1$ independently for each $t_2 = 1, \dots, N_2$. The resulting marginal node functions $\hat{J}_{t_1, t_2}(x)$ should be stored in the memory. Then, the procedure is applied to the vertical rows $t_2 = 1, \dots, N_2$ independently for each $t_1 = 1, \dots, N_1$ with the only alteration: the respective marginal node functions $\hat{J}_t(x)$, obtained at the first step, are taken instead of the image-dependent node functions. In the case of real-valued variables x_t and quadratic pair-wise separable objective function, these elementary procedures applied to single horizontal and vertical rows are nothing else than Kalman filters-interpolators of special kind.

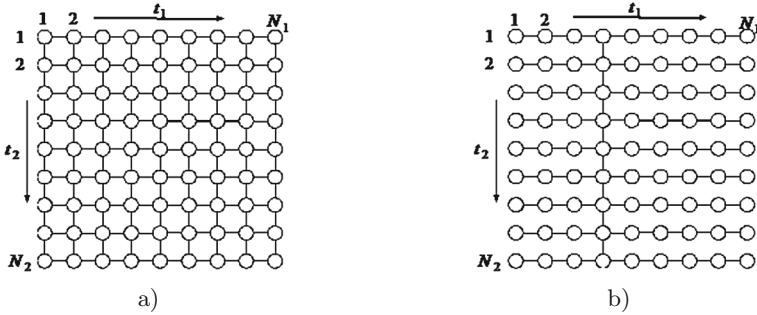


Fig. 2. Neighborhood graphs of image elements: (a) rectangular lattice; (b) simplest tree.

Each iteration of this algorithm has linear computational complexity with respect to size of the original image. Once the estimates $\hat{X}^k = (\hat{X}_t^k = \hat{x}_t^k, t \in T)$ are found, the next approximation to the estimates of factors $\hat{\Lambda}^k = (\hat{\lambda}_t^k = \lambda, t \in T)$ is defined by the rule (9):

$$\hat{\lambda}_t^{k+1} = \frac{(\hat{x}_{t'}^k - \hat{x}_{t''}^k)^2 + \lambda/\mu}{1 + 1/\mu}, t', t'' \in T, \quad (12)$$

which, in accordance with (8), gives the solution of the conditional optimization problem

$$\begin{aligned} \hat{\Lambda}^{k+1} &= \arg \min_{\Lambda} J(\hat{X}^k, \Lambda | Y, \lambda, \mu) \\ &= \arg \min_{\Lambda} \left\{ \sum_{t \in T} (y_t - \hat{x}_t)^2 + \sum_{t', t'' \in V} \frac{1}{\lambda_t} (\hat{x}_{t'}^k - \hat{x}_{t''}^k)^2 \right\}. \end{aligned}$$

The very structure of the iterative procedure (11) and (12) provides objective function satisfying the inequality

$$J(\hat{X}^{k+1}, \hat{\Lambda}^{k+1} | Y, \lambda, \mu) < J(\hat{X}^k, \hat{\Lambda}^k | Y, \lambda, \mu),$$

which, as it is shown previously, the equality holds only at the stationary point.

It remains only to specify the way of choosing the values of the structural parameters μ and λ , which control, respectively, the basic average factors and the ability of instantaneous volatility factors to change along the image plane.

5 Experimental Results

We use here one of the common measures of image distortion, namely peak signal-to-noise ratio (PSNR). PSNR is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is most easily defined via the mean squared error (MSE). Given a noise-free mn monochrome image x and its noisy approximation \hat{x} , MSE is defined as:

$$PSNR = 10\log_{10} \left(\frac{\max(x^2)}{\langle |\hat{x} - x|^2 \rangle} \right)$$

We have tested the various denoising methods for a representative set of standard 8-bit grayscale images such as Barbara, Moon (size 512×512) and Peppers, Cameraman (size 256×256), corrupted by simulated additive Gaussian white noise and Poissonian noise at six different power levels, which corresponds to PSNR decibel values. The parameters of each method have been set according to the values given by their respective authors in the corresponding referred papers. Variations in output PSNRs are, thus, only due to the denoising techniques themselves.

Table 1 summarizes the PSNRs obtained by the various algorithms for denoising of Poisson. We can see that the PURE-based approach clearly outperforms

Table 1. Comparison of Poissonian noise removal algorithms

Images	Cameraman 256×256						Moon 512×512					
	120	60	30	20	10	5	120	60	30	20	10	5
Input PSNR	24.08	21.07	18.05	16.29	13.28	10.27	26.27	23.25	20.23	18.48	15.47	12.46
Haar-Fisz	28.49	26.40	24.36	23.69	22.39	20.93	29.03	26.89	25.01	24.33	23.55	22.78
Platelet	28.29	26.79	25.44	24.60	23.24	21.49	29.16	26.01	25.05	24.60	23.96	23.63
PURE-LET	30.07	28.28	26.54	25.55	23.94	22.42	29.62	27.97	26.56	25.87	24.92	24.23
Our algorithm	29.03	26.95	25.67	24.70	23.54	21.86	29.56	27.64	25.98	25.05	24.39	23.79

Table 2. Comparison of Gaussian noise removal algorithms

Images	Papers 256×256						Barbara 512×512					
	120	60	30	20	10	5	120	60	30	20	10	5
Input PSNR	8.13	14.15	18.59	22.11	28.13	34.15	8.34	14.35	19.23	22.48	28.47	34.46
FBF	21.39	24.39	27.57	29.60	32.62	35.72	22.66	24.21	26.43	28.32	32.92	36.51
SURE-LET	21.32	24.43	27.13	29.33	33.18	37.17	21.76	23.70	25.83	27.98	32.18	36.71
BM3D	21.49	24.40	27.70	30.16	34.68	37.93	23.03	24.59	26.99	28.72	33.98	37.78
Our algorithm	21.42	24.40	27.47	29.50	33.54	37.76	22.56	23.34	25.98	28.05	33.39	36.79

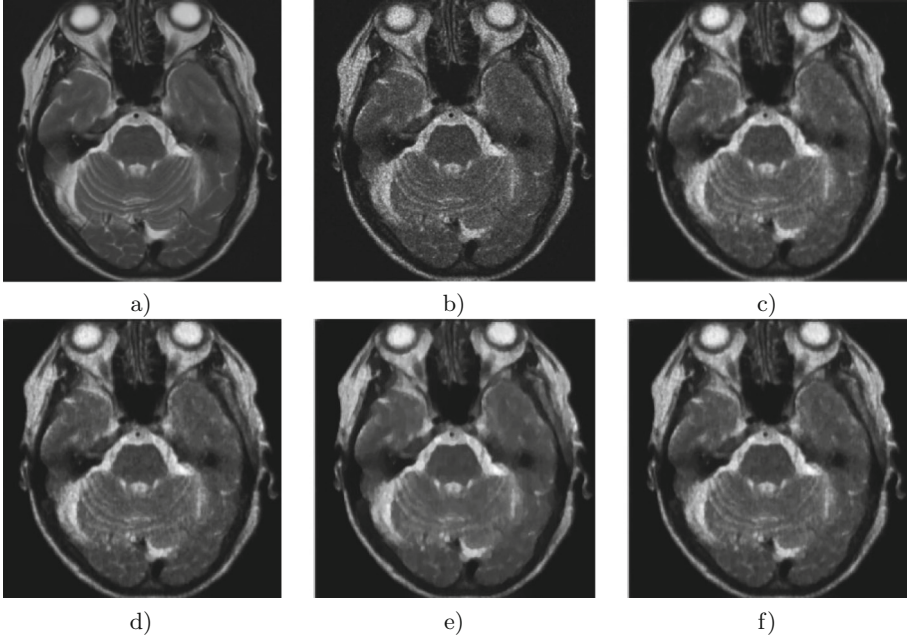


Fig. 3. (a) Part of the original MRI slice. (b) Noisy version with simulated Poissonian noise: PSNR = 22.31 dB. (c) Denoised with our algorithm ($\lambda = 10$, $\mu = 0.5$): PSNR = 29.03 dB. (d) Denoised with our algorithm ($\lambda = 0.1$, $\mu = 10$): PSNR = 29.54 dB. (e) Denoised with our algorithm ($\lambda = 0.0001$, $\mu = 0.1$): PSNR = 29.32 dB. (f) Denoised with our algorithm ($\lambda = 0.1$, $\mu = 0.5$): PSNR = 29.89 dB.

the standard VST-based wavelet denoiser applied in an orthonormal wavelet basis. Our solution also gives significantly better PSNRs than the non-redundant version of the Platelet approach and Haar-Fisz algorithm.

Table 2 summarizes the results obtained for denoising of Gaussian. Our results are already competitive with the best techniques available such as SURE-LET, FBF and BM3D.

Figures 3 and 4 present the results of our algorithm for different images. The values of the parameters (the basic average factors) and (the ability of instantaneous volatility factors to change in time) are chosen different that show their influence on the PSNRs value.

It is also interesting to evaluate the various denoising methods from a practical point of view: the computation time. Indeed, the results achieved by algorithms used for comparison in this paper are superiorly than many other algorithms, but their weakness is the time they require (on a Core i3 workstation with 1.6 GHz for 256×256 and 512×512 images to obtain the redundant results reported in Table 3. With our method, the whole denoising process lasts approximately 0.16 s for 256×256 images (0.53 s for 512×512 images), using a similar workstation. To compare with, FBF lasts approximately 1.8 s for



Fig. 4. (a) The original Barbara image. (b) Noisy version with simulated Gaussian noise: PSNR = 19.323 dB. (c) Denoised with our algorithm ($\lambda = 0.1$, $\mu = 1$): PSNR = 21.34 dB. (d) Denoised with our algorithm ($\lambda = 0.1$, $\mu = 0.5$): PSNR = 23.65 dB. (e) Denoised with our algorithm ($\lambda = 0.0001$, $\mu = 0.05$): PSNR = 22.59 dB. (f) Denoised with our algorithm ($\lambda = 0.01$, $\mu = 0.8$): PSNR = 25.98 dB.

Table 3. Relative computation time of various denoising techniques (seconds)

Methods	Image size	
	256 × 256	512 × 512
Platelet	856	1112
PURE-LET	4.6	10.2
BM3D	4.1	9.2
Haar-Fisz	1.3	3.2
FBF	0.5	1.8
SURE-LET	0.4	1.6
Our algorithm	0.16	0.53

512×512 images, PURE-LET lasts approximately 10.2s for 512×512 images. Besides giving competitive results, our method is also much faster.

6 Conclusion

The gamma-normal model of the image and the expected result of processing, proposed in this paper in a combination with computationally effective Kalman filtration-interpolation procedure, allows us to develop fast global image denoising algorithm which is able to preserve substantial local image features, in particular edges of objects.

The comparison of the denoising results, obtained with our algorithm with respect to the other methods, demonstrates the efficiency of our approach for most of the images. The visual quality of our denoised images is moreover characterized by fewer artifacts than the other methods.

However, the most important advantage of the algorithm is low computation time in comparison with other algorithms. With this algorithm you can handle large images in a relatively short time.

Acknowledgements. This research is funded by RFBR, grant #13-07-00529.

References

1. Luisier, F., Blu, T., Unser, M.: A new SURE approach to image denoising: inter-scale orthonormal wavelet thresholding. *IEEE Trans. Image Process.* **16**(3), 593–606 (2007)
2. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-D transform-domain collaborative ltering. *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007)
3. Yang Q., Tan K.H., Ahuja N.: Real-time $O(1)$ bilateral filtering. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 557–564, Miami (2009)
4. Tomasi C., Manduchi R.: Bilateral filtering for gray and color images. In: *6th International Conference on Computer Vision*, pp. 839–846, Bombay (1998)
5. Aurich V., Weule J.: Non-linear gaussian filters performing edge preserving diffusion. In: *DAGM Symposium*, pp. 538–545, Bielefeld (1995)
6. Smith, S.M., Brady, J.M.: SUSANA new approach to low level image processing. *Int. J. Comput. Vis.* **23**(1), 45–78 (1997)
7. Elad, M.: On the origin of the bilateral filter and ways to improve it. *IEEE Trans. Image Process.* **11**(10), 1141–1151 (2002)
8. Donoho D.L.: Nonlinear wavelet methods for recovery of signals densities and spectra from indirect and noisy data. In: Daubechies, I. (ed.) *Different Perspectives on Wavelets, Proceedings of Symposia in Applied Mathematics*, vol. 47, pp. 173–205. American Mathematical Society, Providence (1993)
9. Fryzlewicz, P., Nason, G.P.: A HaarFisz algorithm for poisson intensity estimation. *J. Computat. Graph. Stat.* **13**(3), 621–638 (2004)
10. Willett, R.M., Nowak, R.D.: Multiscale poisson intensity and density estimation. *IEEE Trans. Inf. Theory* **53**(9), 3171–3187 (2007)

11. Luisier, F., Vonesch, C., Blu, T., Unser, M.: Fast interscale wavelet denoising of poisson-corrupted images. *J. Sig. Process.* **90**, 415–427 (2010)
12. Gracheva I., Kopylov A.: Adaptivnyj parametriceskij algoritm sglazhivaniya izobrazhenij. *Izvestija TulGU*, ser. “Tehnicheskie nauki”, Tula: Izd-vo TulGU **9**(2), 61–67 (2013) (in Russian)
13. Markov, M., Mottl, V., Muchnik, I.: Principles of nonstationary regression estimation: a new approach to dynamic multi-factor models in finance. DIMACS Technical report, Rutgers University, USA (2004)
14. Mottl, V., Blinov, A., Kopylov, A., Kostin, A.: Optimization techniques on pixel neighborhood graphs for image processing. In: Jolion, J.-M., Kropatsch, W.G. (eds.) *Graph-Based Representations in Pattern Recognition*. Computing Supplement, vol. 12, pp. 135–145. Springer, Wien (1998)