

EDGE-PRESERVING DENOISING BASED ON DYNAMIC PROGRAMMING ON THE FULL SET OF ADJACENCY GRAPHS

Pham Cong Thang¹, Andrei V. Kopylov², Sergei D. Dvoenko²

¹The University of Da Nang - University of Science and Technology, 54 Nguyen Luong Bang, Da Nang, Viet Nam
pcthang@dut.udn.vn

²Tula State University, 92 Lenin Ave., Tula, Russia
And.Kopylov@gmail.com, dsd@tsu.tula.ru

Commission II, WG II/10

KEYWORDS: Full set of adjacency graphs, Dynamic programming, Bayesian framework, Edge-preserving smoothing

ABSTRACT:

The ability of a denoising procedure to preserve fine image structures when suppressing unwanted noise has crucial importance for an accurate and effective medical diagnosis. We introduce here a new procedure of edge-preserving denoising for medical images, that combines the flexibility in prior assumptions, and computational effectiveness of parametric multi-quadratic dynamic programming with the increased accuracy of a tree-like representation of a discrete lattice based on the full set of possible adjacency graphs of image elements. Proposed procedure can effectively remove an additive white Gaussian noise with high quality. We provide experimental results in image denoising as well as comparison with related methods.

1. INTRODUCTION

Medical images such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI) or other tomographic modalities, like SPECT, PET, or ultrasound, is an essential source of non-invasive information both for diagnosis and treatment planning. However, the image acquisition and transmission process inevitably causes noise in the image due to Imaging plate nonuniformity, noise in the electronics chains, source power fluctuations, quantization noise in the analog-to-digital conversion process, and so on. But in the field of medical imaging, a precise representation of the fine local image structure is extremely important since the detection of small objects or a tissue type is often an objective. Thereby, the edge preserving properties of the denoising procedures become especially significant.

Despite of many present powerful methods, described in the literature, like nonlinear total variation (Rudin et al., 1992; Wang Y., et al., 2011), fourth order PDEs (You Y., Kaveh M., 2000), and nonlinear anisotropic diffusion (Perona, P. and Malik, J., 1990, Gerig G., et al., 1992; Yuanquan Wang, et al., 2013), none of them can simultaneously achieves both sufficient accuracy to provide a highly reliable data, and computational speed to process super-resolution or dynamic images at practice-relevant time.

In this paper, we proposed a parametric procedure for edge preserving image denoising using Bayesian framework as one of the most popular approaches to image processing. In this approach, the image analysis problem can be described as the problem of estimation of a hidden Markov component of a two-component random field, where the analyzed image plays the role of the observed component. An equivalent representation of Markov random fields in the form of Gibbs random fields, according to the Hammersley-Clifford theorem, can be used to define a priori

probability properties of a hidden Markov field by means of so called Gibbs potentials for cliques.

In the case of a so called singular loss function, the Bayesian estimation of the hidden component can be found as a maximum a posteriori probability (MAP) estimation, which leads us to the problem of minimization of the objective function, often called the Gibbs energy function (Besag J.E., 1974).

A new non-convex type of pairwise Gibbs potentials was proposed in the papers (Pham C.T. and Kopylov A.V., 2015, 2016), with the ability to flexibly define a priori preferences, using separate penalties for various ranges of differences between values of an image adjacent elements. A special version of the parametric dynamic programming procedure was elaborated for optimization of the objective function, based on the tree-like approximation of the lattice neighborhood graph. Experiments show that the proposed procedure can effectively manage heterogeneities and discontinuities in the source data.

The tree-like approximation method (Mottl V., et al., 1998) decomposes the original lattice-like adjacency graph into several tree-like ones. Each of these graphs covers all elements of the pixel grid. The final decision for each variable is based on the separate partial objective function with the tree-like adjacency of variables instead of the overall objective function. Thus, some relations between goal variables are eliminated and are not taken into account, which leads to decreasing in accuracy.

In the paper (Dvoenko S.D., 2012) another way for the tree-like approximation of a lattice based on the full set of acyclic adjacency graphs was proposed. Let a hypothetical covering set of all spanning acyclic graphs (the full set) be given. For the finite set of image elements, the number of such graphs is also finite. Let us

assume that all elements of the data array be roots for several unknown for us acyclic adjacency graphs from the full set. Expanding step-by-step vicinities of descendants for each element simultaneously, we can obtain its maximal vicinity including the element itself, and thus obtain the final decision for that element as a combination of decisions based on acyclic adjacency graphs from the full set. The paper describes the general probabilistic framework for dependent objects recognition.

In this work, we propose a new edge-preserving procedure for medical images, that combines the flexibility in prior assumptions, and computational effectiveness of parametric dynamic programming shown in (Pham C.T. and Kopylov A.V., 2015, 2016) with the increased accuracy of the tree-like representation of a lattice on the basis of the full set of adjacency graphs, described in (Dvoenko S.D., 2012). It should be noticed, in this case the only forward move of the dynamic programming procedure in needed to find the final solution. According to it, we sufficiently simplify the recalculation of the intermediate Bellman functions.

In experimental studies, we compare the performance of image denoising algorithms by using well-known criteria (Bovik A. C., Wang Z., 2006) like the Mean Structure SIMilarity Index (MSSIM) and the Peak to Signal Noise Ratio (PSNR) for Gaussian denoising. We provide experimental results in medical image denoising as well, as comparison with other related methods.

2. THE PRIOR MODEL FOR EDGE-PRESERVING DENOISING

Within the Bayesian framework to image processing, denoising problem can be formulated as the problem of estimation of a hidden Markov component $\mathbf{X} = (x_t, t \in T)$ on the basis of observation $\mathbf{Y} = (y_t, t \in T)$, where $T = \{\mathbf{t} = (t_1, t_2)\}$, $t_1 = 1 \dots N_1$, $t_2 = 1 \dots N_2$ is a discrete image lattice. Hidden component \mathbf{X} represents the “true” image that we would like to recover, and \mathbf{Y} is observed intensity function of a noisy image.

An equivalent representation of Markov random fields in the form of Gibbs random fields, according to the Hammersley-Clifford theorem, can be used to define a priori probability properties of a hidden Markov field by means of so called Gibbs potentials for cliques of adjacency graph $G \subset T \times T$ of image elements (Figure 3). The clique number of a lattice-like graph is equal to two. Therefore, the pairwise Gibbs potentials, or so-called edge functions are the primary means to specify the statistical relations between image elements and, in turn, to define an edge-preserving properties of estimation procedures. Edge functions $\gamma_{t',t''}(x_{t'}, x_{t''})$ is defined over $\mathbf{X} \times \mathbf{X}$ for each pair $(\mathbf{t}', \mathbf{t}'') \in G$ of neighboring pixels and takes the greater value the greater is the discrepancy between the respective hidden values.

Different pairwise potentials such as Huber (Stevenson R., Stevenson D.E., 1990), semi-Huber function Fleury G., De la Rosa J. I., 2004), generalized gaussian function (Bouman, C., Sauer K., 1993), Besag function (Besag J., 1986), Green function (Green P. J. 1990) and others, can be found in the literature (Nikolova M., et al., 2010). Nonconvex functions offer the best possible quality of image reconstruction with neat and exact edges. One of the main problems in these approaches is high computational complexity of corresponding estimation procedures which can hardly be applied to high-resolution images.

A new non-convex type of pairwise Gibbs potentials (Figure 1):

$$\gamma_{\mathbf{t}}(x_{\mathbf{t}}, x_{\mathbf{t}-1}) = \min \left[\gamma_{\mathbf{t}}^{(1)}(x_{\mathbf{t}}, x_{\mathbf{t}-1}), \dots, \gamma_{\mathbf{t}}^{(L)}(x_{\mathbf{t}}, x_{\mathbf{t}-1}) \right], \quad (1)$$

where $\gamma_{\mathbf{t}}^{(i)}(x_{\mathbf{t}}, x_{\mathbf{t}-1}) = \lambda^{(i)}(x_{\mathbf{t}} - x_{\mathbf{t}-1})^2 + d^{(i)}$ – are quadratic functions with parameters $\lambda^{(i)}$ и $d^{(i)}$, $i = 1, \dots, L$, was proposed in the papers (Pham C.T. and Kopylov A.V., 2015, 2016).

This type of edge functions has the ability to flexibly define a priori preferences, using separate penalties for various ranges of differences between values of an image adjacent elements, and leads to computationally effective procedure of maximum a posteriori probability (MAP) estimation of a hidden component of two component random field.

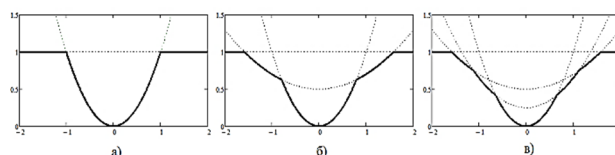


Figure 1. Examples of non-convex pairwise Gibbs potentials for a) $L = 2$; б) $L = 3$; в) $L = 4$

MAP estimation, can be formulated as a problem of minimization of the objective function of a special kind, called Gibbs energy function (Besag J.E., 1974).

$$J(\mathbf{X} | \mathbf{Y}) = \sum_{\mathbf{t} \in T} \psi_{\mathbf{t}}(x_{\mathbf{t}} | \mathbf{Y}) + \sum_{(\mathbf{t}, \mathbf{t}') \in G} \gamma_{\mathbf{t}, \mathbf{t}'}(x_{\mathbf{t}}, x_{\mathbf{t}'}). \quad (2)$$

where node function $\psi_{\mathbf{t}}(x_{\mathbf{t}} | \mathbf{Y})$, $x_{\mathbf{t}} \in \mathbf{X}$ represents unary Gibbs potentials and takes the greater value the more evident is the contradiction between the hypothesis that $x_{\mathbf{t}} \in \mathbf{X}$ is just the correct local value we are seeking and the current observable value $y_{\mathbf{t}} \in \mathbf{Y}$.

3. THE BASIC OPTIMIZATION PROCEDURE FOR A SEPARABLE FUNCTION SUPPORTED BY A TREE

If G is a tree on the set of nodes T (Figure 2.), there exists the highly effective global optimization procedure, based on a recurrent decomposition of the initial problem of minimizing multivariate function into a succession of partial problems, each of which consists in minimizing a function of only one variable. Such a procedure is nothing else than one of versions of the famous dynamic programming procedure.

Let the tree $G_{\mathbf{t}} \in G$ formed by a node \mathbf{t} and its descendants be called descendant tree of this node. The set of all nodes in the descendant tree of a node \mathbf{t} will be denoted by $T_{\mathbf{t}}$, symbol $T_{\mathbf{t}}(\mathbf{t})$ will mean the same set without this node itself. Analogously, symbols $\mathbf{X}_{\mathbf{t}} = (x_s, s \in T_{\mathbf{t}})$ and $\mathbf{X}_{\mathbf{t}}(\mathbf{t}) = (x_s, s \in T_{\mathbf{t}}(\mathbf{t}))$ will mean partial vectors of variables at the respective sets of nodes.

The principal idea of the optimization procedure for separable goal functions supported by trees is based, like the classical dynamic programming procedure for chains, on the notion of Bellman function. The fundamental property of the Bellman function (Mottl V., et al., 1998):

$$\tilde{J}_t(x_t) = \psi_t(x_t) + \sum_{s \in T(t)^0} \min_{x_s \in X} \{ \gamma_{t,s}(x_t, x_s) + \tilde{J}_s(x_s) \} \quad (3)$$

is called the upward recurrent relation. The inverted form of this relation

$$\tilde{x}_s(x_t) = \arg \min_{x_s \in X} \{ \gamma_{t,s}(x_t, x_s) + \tilde{J}_s(x_s) \}, \quad s \in T(t)^0, \quad (4)$$

is referred to the backward recurrent relation.

Let us call

$$\tilde{F}_t(x_t) = \sum_{s \in T(t)^0} \min_{x_s \in X} \{ \gamma_{t,s}(x_t, x_s) + \tilde{J}_s(x_s) \}, \quad (5)$$

a partial Bellman function, then $\tilde{J}_t(x_t) = \psi_t(x_t) + \tilde{F}_t(x_t)$.

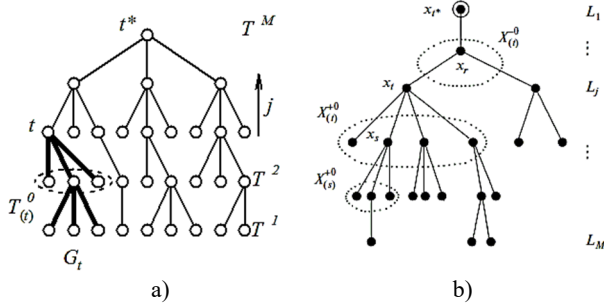


Figure 2: a) Structure of an arbitrary tree; b) Markov vicinity for the x_t variable divided into two arbitrary parts for treelike adjacency graph G

The procedure of dynamic programming searches for the minimum of the objective function (2) in two passes according to forward (3) and backward (4) recurrent relations.

Nevertheless, this procedure cannot be applied immediately to the image reconstruction tasks since discrete image lattice is not a tree. The tree-like approximation method (Mottl V., et al., 1998) decomposes the original lattice-like adjacency graph into several tree-like ones. Each of these graphs covers all elements of the pixel grid. The final decision for each variable is based on the separate partial objective function with the tree-like adjacency of variables instead of the overall objective function. Thus, some relations between goal variables are eliminated and are not taken into account, which leads to decreasing in accuracy.

As it was shown in (Pham C.T. and Kopylov A.V., 2015, 2016), in the case of a minimum of a finite set of quadratic functions of pairwise Gibbs potentials (1), and node functions are in quadratic form, the Bellman functions at each step of the dynamic programming are a minimum of a finite set of quadratic functions. The procedure breaks down at each step into several parallel procedures, according to the number of quadratic functions forming the intermediate optimization problems of one variable. The corresponding procedure is called a multi quadratic dynamic programming procedure (MQDP). A special version of the multi quadratic dynamic programming procedure was elaborated for optimization of the objective function, based on the tree-like approximation of the lattice neighborhood graph. Experiments

show that the proposed procedure can effectively manage heterogeneities and discontinuities in the source data.

The number of quadratic functions in a representation of Bellman functions grows during the forward move and a special technique on the basis of k-means clustering (Dvoenko S. D., 2009) was proposed to reduce their number.

We proposed here a new procedure on the basis of another type of tree-like representation of image lattice, which does not eliminate any pairwise connections between image elements and can sufficiently simplify the recalculation of the intermediate Bellman functions.

4. THE FULL SET OF ACYCLIC ADJACENCY GRAPHS

In the paper (Dvoenko S.D., 2012) another way for the tree-like approximation of a lattice based on the full set of acyclic adjacency graphs was proposed.

For an element $t \in T$, we have its vicinity relative to a certain acyclic adjacency graph G , which can be divided arbitrarily into two parts $T(t)^0 = T(t)^- \cup T(t)^+$ (Fig. 2.b).

Let us expand vicinity $T(t)^1 = T(t)^{+0}$ of descendants

$$\begin{aligned} T(t)^{\oplus 2} &= (T(t)^{+0})^{+0} = T(t)^+ \cup (\cup_s T(s)^+), \quad s \in T(t)^+, \\ T(t)^{\oplus 3} &= ((T(t)^{+0})^{+0})^{+0} = (T(t)^+ \cup (\cup_s T(s)^+)) \cup (\cup_v T(v)^+), \\ &s \in T(t)^+, \quad \cup_v T(v)^+, \dots, \\ T(t)^{\oplus i} &= T(t)^{\oplus(i-1)} \cup (\cup_v T(v)^+), \quad v \in T(t)^{\oplus(i-1)} \setminus \cup_v T(v)^{\oplus(i-2)}. \end{aligned}$$

This process is finished when the expanded descendant vicinity will contain all terminal elements as descendants of element $t \in T$.

Let a hypothetical covering set of all spanning acyclic graphs (the full set) be given.

For the finite set of elements for array T , the number of such graphs is also finite. In the general case for all feasible data arrays T , the full set of acyclic adjacency graphs is countable.

It is evident that in the general case the acyclic graphs from the full set are unknown. Let us accept that all elements $t \in T$ of the data array appear to be roots for several unknown for us acyclic adjacency graphs from the full set.

Let us expand vicinities of descendants for each element $t \in T$ simultaneously step-by-step. Since we use 4-connectivity of image elements, only four descendants must be considered simultaneously (Figure 3). Thus, when we obtain the maximal vicinity for each element, including the element itself, we will have the whole collection of the Bellman functions $\tilde{J}_t(x_t)$ (3), $t \in T$ already as a result of combining the acyclic adjacency graphs from the full set.

We use here non-convex type of pairwise potential functions (1) in the form:

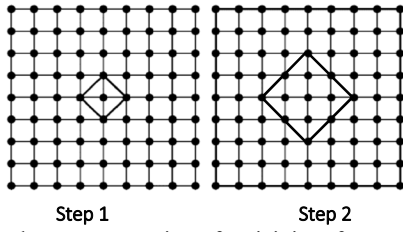


Figure 3. Step-by-step expansion of a vicinity of a central point

$$\gamma_{\mathbf{t}}(x_{\mathbf{t}}, x_{\mathbf{t}-1}) = u \min[(x_{\mathbf{t}} - x_{\mathbf{t}-1})^2, \lambda_1(x_{\mathbf{t}} - x_{\mathbf{t}-1})^2 + d_1, \Delta^2], \quad (6)$$

where u - smoothing parameters;

$\lambda_1(x_{\mathbf{t}} - x_{\mathbf{t}-1})^2 + d_1$ -quadratic functions with parameters λ_1 и d_1 .

Node functions are selected in the quadratic form:

$$\psi_{\mathbf{t}}(x_{\mathbf{t}}) = (x_{\mathbf{t}} - y_{\mathbf{t}})^2.$$

For each element $\mathbf{t} = \{t_1, t_2\}$ $\mathbf{t} \in T$ we calculate partial Bellman function (5):

$$\tilde{F}_{\mathbf{t}}(x_{\mathbf{t}}) = \begin{bmatrix} \min_{x_{\eta_1, \eta_2-1}} \left\{ \gamma_{t_1, t_2-1}(x_{t_1, t_2}, x_{t_1, t_2-1}) + \tilde{J}_{t_1, t_2-1}(x_{t_1, t_2-1}) \right\} \\ + \min_{x_{\eta_1, \eta_2+1}} \left\{ \gamma_{t_1, t_2+1}(x_{t_1, t_2}, x_{t_1, t_2+1}) + \tilde{J}_{t_1, t_2+1}(x_{t_1, t_2+1}) \right\} \\ + \min_{x_{\eta_1+1, \eta_2}} \left\{ \gamma_{t_1+1, t_2}(x_{t_1, t_2}, x_{t_1+1, t_2}) + \tilde{J}_{t_1+1, t_2}(x_{t_1+1, t_2}) \right\} \\ + \min_{x_{\eta_1-1, \eta_2}} \left\{ \gamma_{t_1-1, t_2}(x_{t_1, t_2}, x_{t_1-1, t_2}) + \tilde{J}_{t_1-1, t_2}(x_{t_1-1, t_2}) \right\} \end{bmatrix}. \quad (7)$$

Partial Bellman functions $\tilde{J}(x_{\mathbf{t}})$ in (7) are taken instead of the node functions $\psi_{\mathbf{t}}(x_{\mathbf{t}})$. Then we can calculate Bellman functions for each element $\mathbf{t} = \{t_1, t_2\}$:

$$\tilde{J}_{t_1, t_2}(x_{t_1, t_2}) = \psi_{t_1, t_2}(x_{t_1, t_2}) + \tilde{F}_{t_1, t_2}(x_{t_1, t_2}). \quad (8)$$

The number of quadratic functions in a representation of a current Bellman function can be naturally reduced by the following algorithm.

Algorithm 1: Reducing a number of quadratic functions

Input: $\tilde{F}_{\mathbf{t}}^{(i)}(x_{\mathbf{t}}), i = 1, \dots, \tilde{K}_{\mathbf{t}}$.

Output: $\tilde{F}_{\mathbf{t}}^{(i)}(x_{\mathbf{t}}), i = 1, \dots, \tilde{K}'_{\mathbf{t}}, \tilde{K}'_{\mathbf{t}} < \tilde{K}_{\mathbf{t}}$

1. At the beginning, sort by ascending values $\tilde{d}_{\mathbf{t}}^{(i)}$ of array $\tilde{F}_{\mathbf{t}}^{(i)}(x_{\mathbf{t}})$.
 2. In case of the presence of a parameter Δ , at each step, we look for the minimum constant $D_{\mathbf{t}} = \tilde{d}_{\mathbf{t}}^{(1)} + u \cdot \Delta^2$ and reject all other constants.
 3. Discard all functions than have minimum greater than or equal to this constant $\tilde{d}_{\mathbf{t}}^{(i)} \geq D_{\mathbf{t}}, i = 2 \dots \tilde{K}_{\mathbf{t}}$.
-

4. Among the remaining functions, select the function with the smallest minimum and leave it.
 5. Find and check necessary and sufficient condition of intersection of quadratic functions. Discard all the functions which have no intersection.
 6. Among the remaining functions, select the function with the smallest minimum and leave it.
 7. Repeat until there will be functions for which have no decision on acceptance or discarding.
-

The denoising algorithm for the full set of acyclic graphs is described as following:

Algorithm 2: Image denoising algorithm for the full set of acyclic graphs

1. **Initialize:** $\tilde{J}_{\mathbf{t}}^{(1)}(x_{\mathbf{t}})$, **MaxIter**
 2. **While** k=2 to **MaxIter** **do**
 3. Compute Bellman functions $\tilde{J}_{\mathbf{t}}^{(k)}(x_{\mathbf{t}})$
 Using (7-8) with applying algorithm 1 for reducing number of quadratic functions.
 4. **Repeat** step (3)
 5. $x_{\mathbf{t}} = \arg \min \tilde{J}_{\mathbf{t}}^{(k)}(x_{\mathbf{t}})$
 6. **Return** $\tilde{\mathbf{X}}_{\mathbf{t}} = \{x_{\mathbf{t}}\}$
-

We assign that node functions $\psi_{\mathbf{t}}(x_{\mathbf{t}})$ are taken instead of Bellman functions $\tilde{J}_{\mathbf{t}}(x_{\mathbf{t}})$ in equation (7) for first step of algorithm 2.

In the next section, we performed studying experiments to test our model for medical image denoising.

5. EXPERIMENTAL RESULTS

We perform experiments to test our model in case of a 8 bit medical images from the BrainWeb database (<http://brainweb.bic.mni.mcgill.ca/brainweb/>). The experimental performance is measured by PSNR,

$$PSNR = 10 \log_{10} \frac{255^2}{\|\hat{\mathbf{X}} - \mathbf{X}\| / (N_1 N_2)},$$

and a mean value of Structure SIMilarity (SSIM) index:

$$SSIM(P, Q) = \frac{(2\mu_P \mu_Q + c_1)(2\sigma_{P, Q} + c_2)}{(\mu_P^2 + \mu_Q^2 + c_1)(\sigma_P^2 + \sigma_Q^2 + c_2)},$$

where μ_P, μ_Q - means of images;

σ_P, σ_Q - standard deviations (the square root of variance) of images;

$\sigma_{P, Q}$ - covariance of images;

$P(i, j)$ and $Q(i, j)$;

$c_1 = 6.5025, c_2 = 58.5225$;

$P(i, j)$ and $Q(i, j)$ denote pixel values of the original image and the reconstructed or noisy image accordingly.

Note that all the PSNR results (in dB) and MSSIM reported in Fig. 5 have been averaged over 10 noise realizations.

The figure 4 shows PSNR Variations on Loops at level of additive white Gaussian noise with a standard deviation $\sigma = 15$. Figure 5 shows result images of proposed and compared methods: Multi quadratic dynamic programming (MQDP), nonlinear anisotropic diffusion (AD P-M), modified Perona-Malik (MP-M), fourth order PDEs, nonlinear total variation (TV). For MQDP we use edge functions (5) with smoothing parameters of edge function with fixed values $\lambda = 0.2$, $d = 0.5\Delta^2$.

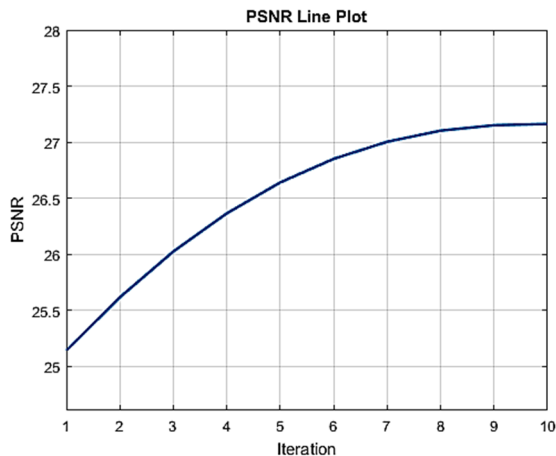


Figure 4. PSNR variations during the vicinity expansion steps

The experimental results show that our image denoising method allows image denoising as well as other related method.

6. CONCLUSION

Edge preserving image denoising has become an urgent step in imaging to remove noise and to preserve local image features for improving the quality of further analysis. In this paper, we proposed approaches to achieve these aims. Proposed edge-preserving procedure for medical images allows, sufficiently simplify the recalculation of the intermediate Bellman functions to find the final solution. The experimental results show that proposed algorithm allows get high results of dynamic programming for image processing.

Numerical results show that our proposed algorithms are efficient and allows get result as well as existed method.

ACKNOWLEDGEMENTS

This work was supported by Russian Foundation for Basic Research (RFBR), projects № 16-07-01039 and 16-57-52042.

REFERENCES

Besag J., 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, Vol. 36, pp. 192–236.

Besag J., 1986. On the Statistical Analysis of Dirty Pictures. *Journal of the Royal Statistical Society (Series B)*, Vol. 48, pp. 259–302.

Bovik A. C., Wang Z. (2006). *Modern Image Quality Assessment, Synthesis Lectures on Image, Video, and Multimedia Processing*. Morgan & Claypool Publishers, 156 pages.

Bouman, C., Sauer K., 1993. A Generalized Gaussian Image Model for Edge-Preserving Map Estimation. *IEEE Transactions on Image Processing*, Vol. 2 (3), pp. 296 - 310.

Dvoenko S.D., 2012. Recognition of dependent objects based on acyclic Markov models. *Journal of Pattern Recognition and Image Analysis*, Vol 22 (1), pp. 28-38.

Fleury G., De la Rosa J. I., (2004). Bootstrap Methods for a Measurement Estimation Problem. *IEEE Transactions on Instrumentation and Measurement*, Vol. 55(3). P. 820-827.

Gerig G., et al., 1992. Nonlinear anisotropic filtering of MRI data. *IEEE Transaction on Medical Imaging*, Vol. 11 (2), pp. 221–232.

Green P. J. (1990) Bayesian reconstruction from emission tomography data using a modified EM algorithm. *IEEE Trans. on Medical Imaging*, Vol. 9(1), pp.84–93.

Mottl V.V., et al., 1998. Optimization Techniques on Pixel Neighborhood Graphs for Image Processing. In: *Jolion JM., Kropatsch W.G. (eds) Graph Based Representations in Pattern Recognition. Supplement 12*. Springer–Verlag/Wien, pp. 135-145.

Nikolova M., Michael K., and Tam C.P., 2010. Fast Nonconvex Nonsmooth Minimization Methods for Image Restoration and Reconstruction. *IEEE Transactions on Image Processing*, Vol. 19 (12), pp. 3073-3088.

Perona, P. and Malik, J., 1990. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12. pp. 629–639.

Pham Cong Thang and Kopylov A. V., 2015. Multi-quadratic dynamic programming procedure of edge-preserving denoising for medical images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-5/W6, pp. 101-106.

Pham Cong Thang, Kopylov A.V, 2016. Parametric Procedures For Image Denoising With Flexible Prior Model. *The Seventh International Symposium on Information and Communication Technology (SoICT 2016)*, pp: 294-301.

Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. *Physica. D*, Vol. 60, pp. 259–268.

Stevenson R., Stevenson D.E., 1990. Fitting curves with discontinuities. *Proc. of the first international workshop on robust computer vision*, pp. 127–136.

Wang Y., et al., 2011. MTV: modified total variation model for image noise removal. *IEE Electronics Letters*, Vol. 47 (10), pp.592-594.

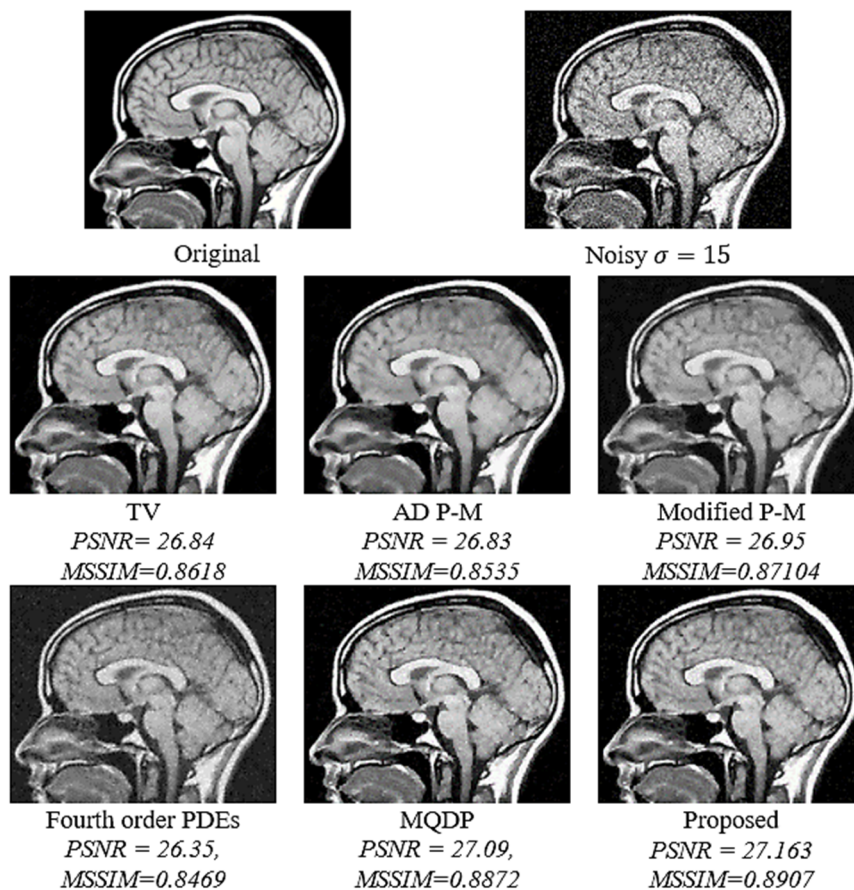


Figure 5. Results of processing for proposed and compared methods

You Y., Kaveh M., 2000. Fourth Order Partial Differential Equations for Noise Removal. *IEEE Trans. Image Processing*, Vol. 9 (10), pp. 1723-1730

Yuanquan Wang, J.C. Guo, W.F. Chen and Wenxue Zhang, 2013. Image denoising using modified Perona-Malik model based on directional Laplacian. *Signal Processing*, Vol. 93(9), September 2013, pp. 2548-2558.