Application of Backward Stochastic Differential Equations to Reconstruction of Vector-Valued Images

Dariusz Borkowski[1] and Katarzyna Jańczak-Borkowska[2]

[1] Faculty of Mathematics and Computer Science,
Nicolaus Copernicus University Chopina 12/18, 87-100 Toruń, Poland
dbor@mat.umk.pl
[2] Institute of Mathematics and Physics, University of Technology and Life Sciences
al. prof. S. Kaliskiego 7, 85-789 Bydgoszcz, Poland
kaja@utp.edu.pl

Abstract. In this paper we explore the problem of reconstruction of vector-valued images with additive Gaussian noise. In order to solve this problem we use backward stochastic differential equations. Our numerical experiments show that the new approach gives very good results and compares favourably with deterministic partial differential equation methods.

1 Introduction

Let D be a bounded, convex domain in $\mathbf{R}^2$, $u : \overline{D} \to \mathbf{R}^3$ be an original colour image and $u_0 : \overline{D} \to \mathbf{R}^3$ be a noisy observation of the form: $u_0 = u + \eta$, where η means white Gaussian noise. Having u_0 we have to reconstruct an original image u. Presented problem is a classic example of an inverse problem [3].

Problem of denoising colour images using fully automatic and reliable methods is one of the most important issues of digital image processing and computer vision. Efficient and effective reconstruction of images is an essential element of most image processing and recognizing algorithms. Algorithms of reconstruction allow us to make initial treatment of data for further analysis, which is very important especially in astronomy, biology and medicine.

The four most popular methods in reconstruction of images are statistic methods, linear filtration, methods based on partial differential equations and stochastic methods. Stochastic methods of denoising images mostly base on theory of random Markov fields. Backward stochastic differential equations give us a new approach to stochastic image processing. In bibliography one can only find theoretical bases of usage of backward stochastic differential equations to image reconstruction [1] and some practical results in the case of grey images [4].

A novel look on the reconstruction problem of gray images with using backward stochastic differential equations was fruitful and gave results that are usually better than existing methods. The idea of this paper is to generalize these results to images with values in $\mathbf{R}^n$, in particular to colour images.

L. Bolc et al. (Eds.): ICCVG 2012, LNCS 7594, pp. 38–47, 2012.

The paper is organized as follows. Section 2 contains definitions and fundamental facts of stochastic analysis. In Section 3 we recall basic ideas from [4]. Section 4 provides new results to reconstruction of colour images. Section 5 is devoted to presenting experimental results and comparing with PDE methods.

2 Mathematical Preliminaries

Let $D \subset \mathbf{R}^n$ be a domain with closure $\overline{D}$ and boundary ∂D. Let $T > 0$ and by $\mathbf{C}([0,T];\mathbf{R}^n)$ denote the set of continuous functions $f : [0,T] \to \mathbf{R}^n$.

Definition 1. *Let $y \in \mathbf{C}([0,T];\mathbf{R}^n)$, $y_0 \in \overline{D}$. A pair $(x,k) \in \mathbf{C}([0,T];\mathbf{R}^{2n})$ is called a solution to the Skorokhod problem associated with y and D if*
1. $x_t = y_t + k_t, \quad t \in [0,T]$,
2. $x_t \in \overline{D}, \quad t \in [0,T]$,
3. k is a function with bounded variation $|k|$ on $[0,T]$, $k_0 = 0$ and

$$k_t = \int_0^t n_s \, d|k|_s, \ t \in [0,T], \ |k|_t = \int_0^t 1_{\{x_s \in \partial D\}} \, d|k|_s, \ t \in [0,T],$$

where $n_s = n(x_s)$ is an inward normal unit vector at $x_s \in \partial D$.

It is known that if D is a convex set, then there exists a unique solution to the Skorokhod problem [11].

Definition 2. *Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space.*
1. An n-dimensional stochastic process $X = \{X_t; t \in [0,T]\}$ is a parametrised collection of random variables defined on $(\Omega, \mathcal{F}, \mathcal{P})$ with values in $\mathbf{R}^n$.
2. For each fixed $\omega \in \Omega$ the function $X_t(\omega)$, $t \in [0,T]$ is called a trajectory of X and is denoted by $X(\omega)$.
3. A filtration $(\mathcal{F}_t) = \{\mathcal{F}_t; t \in [0,T]\}$ is a nondecreasing family of sub-σ-fields of $\mathcal{F}$ i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \le s < t \le T$.
4. A process X is $(\mathcal{F}_t)$ adapted if for each $t \in [0,T]$, X_t is $\mathcal{F}_t$ - measurable random variable.

Definition 3. *Let Y be an $(\mathcal{F}_t)$ adapted process with continuous trajectories, $Y_0 \in \overline{D}$. We say that a pair (X,K) of $(\mathcal{F}_t)$ adapted processes is a solution to the Skorokhod problem associated with Y and D, if for almost every $\omega \in \Omega$, $(X(\omega), K(\omega))$ is a solution to the Skorokhod problem associated with $Y(\omega)$ and D.*

Definition 4. *Assume that we are given $x_0 \in \overline{D}$ and $\sigma : \mathbf{R}^n \to \mathbf{R}^n \times \mathbf{R}^n$. Let Y be an $(\mathcal{F}_t)$ adapted process and by $W = \{W_t; t \in [0,T]\}$ denote an n-dimensional Wiener process. A pair $(X, K^{\overline{D}})$ of $(\mathcal{F}_t)$ adapted processes is called a solution to reflected stochastic differential equation (in short reflected SDE)*

$$X_t = x_0 + \int_0^t \sigma(s, X_s) \, dW_s + K_t^{\overline{D}}, \ t \in [0,T], \tag{1}$$

if $(X, K^{\overline{D}})$ is a solution to the Skorokhod problem associated with

$$V_t = x_0 + \int_0^t \sigma(s, X_s) \, dW_s, t \in [0,T] \text{ and } D.$$

The proof of existence and uniqueness of the solution to reflected SDEs can be found in [11]. The process X satisfying (1) is a diffusion process with values in domain $\overline{D}$. The process X is called the process with reflection.

Let $(\mathcal{F}_t^W)$ be a filtration generated by an l-dimensional Wiener process W, $\xi \in \mathbf{L}^2(\Omega, \mathcal{F}_T, P, \mathbf{R}^k)$ be a square integrable random variable and let $f : \Omega \times [0,T] \times \mathbf{R}^k \to \mathbf{R}^k$ be a Lipschitz continuous function in the space variable.

Definition 5. *A solution to the backward stochastic differential equation (BSDE) associated with ξ and f is a pair of $(\mathcal{F}_t^W)$ adapted processes (Y, Z) with values in $\mathbf{R}^k \times \mathbf{R}^{k\times l}$ satisfying $\mathbf{E}\left[\int_0^T \|Z_s\|^2\, ds\right] < \infty$ and*

$$Y_t = \xi + \int_t^T f(s, Y_s)ds - \int_t^T Z_s\, dW_s, \quad t \in [0,T].$$

See [7] for the proof of existence and uniqueness of the solution to BSDEs.

3 Reconstruction of Gray Levels Images

A general model of the image reconstruction is the following:

$$\begin{cases} X_t = x + \displaystyle\int_0^t \sigma(s, X_s)\, dW_s + K_t^{\overline{D}}, \\ Y_t = \xi + \displaystyle\int_t^T f(s, Y_s, X_s)ds - \int_t^T Z_s\, dW_s, \quad t \in [0,T], \end{cases}$$

where ξ depends on u_0 and the process X.

Note that, the process X has values in domain of the image $\overline{D}$ and is driven by a function σ, the process Y has values in codomain of the image and is driven by a function f. Moreover, the value of the process Y at time $t = 0$ is the reconstructed pixel $u(x)$.

3.1 Stochastic Representation of Solution to the Heat Equation

Before presenting a general method, we will illustrate our ideas by constructing a model which is equivalent to a commonly used filter, namely, the convolution of the noise image with two-dimensional Gaussian mask. We suppose for a while that the image is given by a function defined on the whole plane. Put

$$\begin{cases} X_t = W_t^x, & t \in [0,T], \\ Y_t = u_0(X_T) - \displaystyle\int_t^T Z_s\, dW_s, & t \in [0,T], \end{cases} \tag{2}$$

where W^x is two-dimensional Wiener process starting from $x \in \overline{D}$. From (2) we deduce

$$\begin{cases} X_t = W_t^x, \quad t \in [0,T], \\ Y_0 = u_0(X_T) - \displaystyle\int_0^T Z_s\, dW_s = \mathbf{E}\left(u_0(X_T)\right) = \int_{\mathbf{R}^2} \frac{1}{2\pi T} e^{-\frac{|x-y|^2}{2T}} u_0(y)\, dy. \end{cases} \tag{3}$$

A value of the process Y at time $t = 0$ is the reconstructed pixel $u(x)$. Therefore, by (3) the image is the convolution of the noisy image with two-dimensional Gaussian mask.

While discussing the above example, we assumed that the image is the function given on the whole plane. Since we want to consider the image as a function defined on the bounded, convex set, we have to introduce a new assumption for X. We assume that X is a stochastic process with reflection with values in $\overline{D}$. In this case X is a Wiener process with reflection, which we can write as

$$\begin{cases} X_t = W_t^x + K_t^{\overline{D}}, & t \in [0,T], \\ Y_t = u_0(X_T) - \displaystyle\int_t^T Z_s \, dW_s, & t \in [0,T]. \end{cases}$$

3.2 Anisotropic Diffusion

In the case of smoothing filters we will consider BSDEs associated with $\xi = u_0(X_T)$ and $f(t,y) = 0$, where X is a diffusion process with reflection. Following [13] we provide a construction of a model where process X diffuses along edges. This condition may be achieved by imposing

$$\begin{cases} X_t = x + \displaystyle\int_0^t \begin{bmatrix} -\frac{(G_\gamma * u_0)_{x_2}(X_s)}{|\nabla (G_\gamma * u_0)(X_s)|}, 0 \\ \frac{(G_\gamma * u_0)_{x_1}(X_s)}{|\nabla (G_\gamma * u_0)(X_s)|}, 0 \end{bmatrix} dW_s + K_t^{\overline{D}}, \\ Y_t = u_0(X_T) - \displaystyle\int_t^T Z_s \, dW_s, \qquad t \in [0,T] \end{cases} \tag{4}$$

where $u_{x_i}(y) = \frac{\partial u}{\partial x_i}(y)$. In particular $Y_0 = \mathbf{E}\,[u_0(X_T)]$.

To avoid false detections due to noise, u_0 is convolved with a Gaussian kernel $G_\gamma(x) = \frac{1}{2\pi\gamma^2} e^{-\frac{|x|^2}{2\gamma^2}}$ (in practice 3×3 Gaussian mask).

3.3 Backward Diffusion

In the case of enhancing images we will consider BSDEs associated with $\xi = u_0(x)$ and $f(t,y) = c(y - u_0(X_t))$, where X is a Wiener process with reflection and $c > 0$ is some constant. In accordace with Theorem 3 in [4] the reconstructed image $u(x)$

$$u(x) = Y_0^m = \sum_{k=0}^{m-1} a_k \mathbf{E}\left[u_0(W_{t_k}^x + K_{t_k}^{\overline{D}})\right], t_k = \frac{kT}{m}$$

is a combination of convolutions of the noisy image and two dimensional Gaussian mask with coefficients a_k where $a_0 > 0$, $a_k < 0$, for $k = 1, ..., m-1$ and $\sum_{k=0}^{m-1} a_k = 1$. This mean that the kernel of filtering is made up of positive weight for central pixel and negative weights for neighborhood pixels and finally is a model of enhancing filter. As it was shown in [4] if parameter c is greater than the result is more enhancing.

4 Reconstruction of Vector-Valued Images

Now we concentrate on images with values in $\mathbf{R}^3$. A very common idea to restore vector-valued images is to use scalar diffusion on each channel of a noisy image. But one fastly notices that this scheme is useless, since each image channel evolves independently with different smoothing geometries. To avoid this blending effect, the regularization process have to be driven in a common and coherent way for all the vector image channels. In order to execute that we use Di Zenzo geometry [5,6].

Let $u : D \to \mathbf{R}^3$ be a vector-valued image and $x \in D$ be fixed. Consider the function $F_x : V \to \mathbf{R}$, $F_x(v) = \left|\frac{\partial u}{\partial v}(x)\right|^2$, where $V = \{v \in \mathbf{R}^2; |v| = 1\}$. We are interested in finding the arguments $\theta_+(u,x), \theta_-(u,x)$ and corresponding values $\lambda_+(u,x) = F_x(\theta_+(u,x))$, $\lambda_-(u,x) = F_x(\theta_-(u,x))$ which maximize and minimize the function F_x, respectively.

Note that F_x can be rewritten as $F_x(v) = F_x([v_1, v_2]^T) = v^T \mathbf{G}(x) v$, where

$$\mathbf{G}(x) = \begin{bmatrix} \sum_{i=1}^{3} \left(\frac{\partial u_i}{\partial x_1}(x)\right)^2, & \sum_{i=1}^{3} \frac{\partial u_i}{\partial x_1}(x)\frac{\partial u_i}{\partial x_2}(x) \\ \sum_{i=1}^{3} \frac{\partial u_i}{\partial x_1}(x)\frac{\partial u_i}{\partial x_2}(x), & \sum_{i=1}^{3} \left(\frac{\partial u_i}{\partial x_2}(x)\right)^2 \end{bmatrix}.$$

The interesting point about $\mathbf{G}(x)$ is that its positive eigenvalues $\lambda_+(u,x)$, $\lambda_-(u,x)$ are the maximum and the minimum of F_x while the orthogonal eigenvectors $\theta_+(u,x)$ and $\theta_-(u,x)$ are the corresponding variation orientations.

Three different choices of vector gradient norms $N(u,x)$ have been proposed in the literature $N(u,x) = \sqrt{\lambda_+(u,x)}$, $N(u,x) = \sqrt{\lambda_+(u,x) - \lambda_-(u,x)}$, $N(u,x) = \sqrt{\lambda(u,x) + \lambda_-(u,x)}$. In presented examples we have used $N(u,x) = \sqrt{\lambda_+(u,x)}$ as a natural extension of the scalar gradient norm viewed as the value of maximum variations.

4.1 Anisotropic Diffusion

Replacing in equation (4) $|\nabla(u,x)|$ and $[u_{x_1}(x), u_{x_2}(x)]^T$ respectively by $N(x,u)$ and $\theta_+(u,x) = [\theta_+^1(u,x), \theta_+^2(u,x)]^T$ we obtain the following model of anisotropic diffusion for vector-valued images:

$$\begin{cases} X_t = x + \displaystyle\int_0^t \begin{bmatrix} -\frac{\theta_+^1(G_\gamma * u_0, X_s)}{N((G_\gamma * u_0)(X_s))}, 0 \\ \frac{\theta_+^2(G_\gamma * u_0, X_s)}{N((G_\gamma * u_0)(X_s))}, 0 \end{bmatrix} dW_s + K_t^{\overline{D}}, \\ Y_t = u_0(X_T) - \displaystyle\int_t^T Z_s \, dW_s, \qquad t \in [0,T] \end{cases} \tag{5}$$

and in particular $u(x) = Y_0 = \mathbf{E}\,[u_0(X_T)]$.

4.2 Backward Diffusion

Theorem 1. *Let* $u_0 : \mathbf{R}^2 \to \mathbf{R}^n$, $x \in \mathbf{R}^2$, $f(t,y) = c(y - u_0(W_t^x))$, *where* W^x *is two-dimensional Wiener process starting from* x, $c > 0$. *If* (Y, Z) *is a solution to BSDE associated with* $\xi = u_0(x)$ *and* f *then*

$$\lim_{m\to+\infty} Y_0^m = \lim_{m\to+\infty} \sum_{k=0}^{m-1} a_k \mathbf{E}\left[u_0(W_{t_k}^x)\right] = \lim_{m\to+\infty} \sum_{k=0}^{m-1} a_k (G_{\sqrt{t}_k} * u_0)(x) = Y_0,$$

where $a_0 = \left(1 + \frac{cT}{m}\right)^m - \frac{cT}{m}$, $a_k = -\frac{cT}{m}\left(1 + \frac{cT}{m}\right)^k$, $t_k = \frac{kT}{m}$, $k = 0, 1, ..., m-1$ *($G_0 \equiv \delta$ is a Dirac function) and* $\sum_{k=0}^{m-1} a_k = 1$.

The above theorem is a generalization of results from [4] and its proof is silmilar to the proof of Theorem 3 in [4].

Since the image is a function defined on the bounded set we have to consider Wiener process with reflection with values in domain $\overline{D}$. Finally we have the following model of backward diffusion for vector-valued images:

$$u(x) = Y_0^m = \sum_{k=0}^{m-1} a_k \mathbf{E}\left[u_0(W_{t_k}^x + K_{t_k}^{\overline{D}})\right], \tag{6}$$

where Y^m is process with values in $\mathbf{R}^n$.

5 Experimental Results

Some results from our evaluation experiments regarding BSDE method and classic PDE methods: total variation [9] (in short TV) and Perona-Malik [8] (in short PM) for colour images [5,10] are presented in: Table 1, Table 2 and Fig. 1, Fig. 2. The results refer to RGB images: *Lenna*, *house* and *peppers* corrupted

Table 1. SSIM

Image Method\Standard deviation	*Lenna* $\rho = 30$	*Lenna* $\rho = 40$	*house* $\rho = 30$	*house* $\rho = 40$	*peppers* $\rho = 30$	*peppers* $\rho = 40$
PM	0.8422	0.8203	0.7528	0.7291	0.8373	0.8073
TV	0.8723	0.8333	0.7884	0.7550	0.8692	0.8249
BSDE (5)	**0.8823**	**0.8449**	**0.7896**	**0.7566**	**0.8705**	**0.8342**

Table 2. PSNR

Image Method\Standard deviation	*Lenna* $\rho = 30$	*Lenna* $\rho = 40$	*house* $\rho = 30$	*house* $\rho = 40$	*peppers* $\rho = 30$	*peppers* $\rho = 40$
PM	29.0295	27.6992	28.7392	27.4461	28.6412	27.3377
TV	**29.4952**	28.2040	29.2119	27.7876	**28.8832**	27.5218
BSDE (5)	29.3037	**28.2205**	**29.5891**	**28.2195**	28.7124	**27.5597**

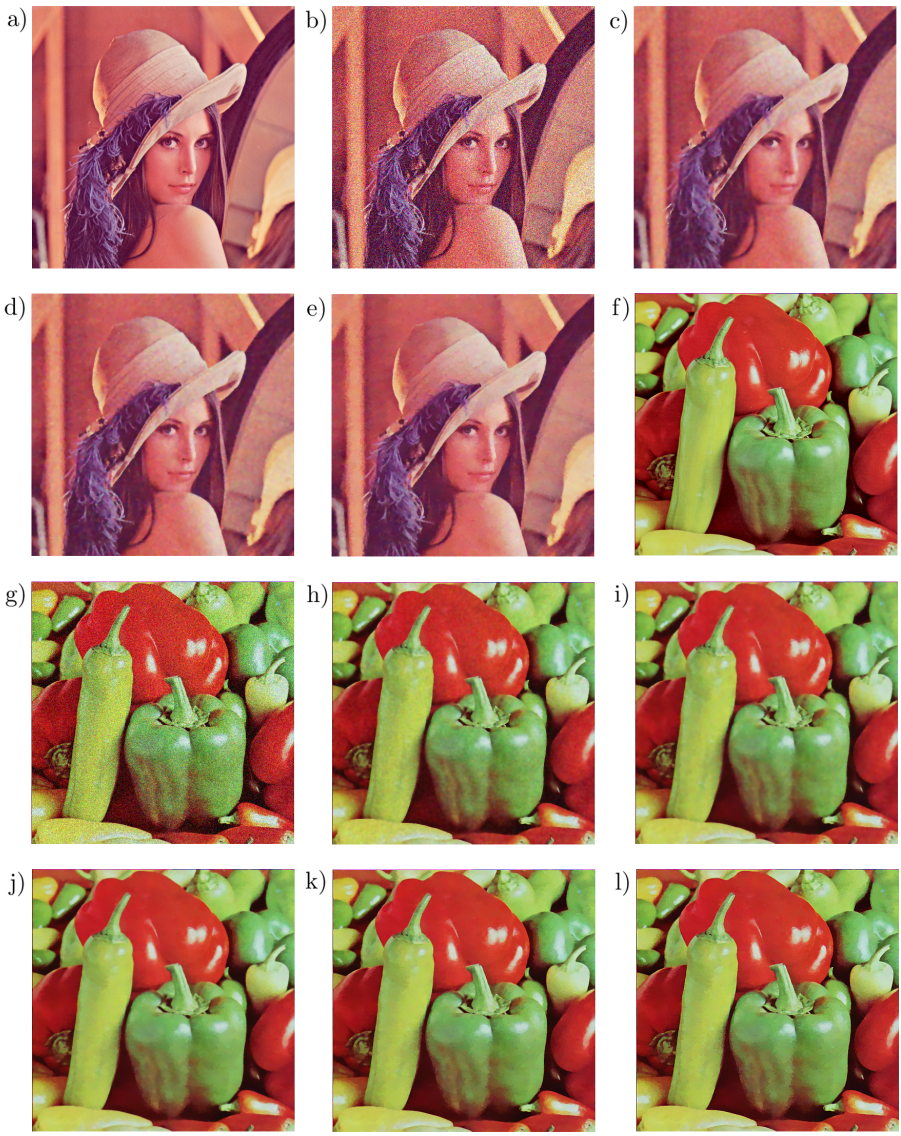

Fig. 1. a) Original *Lenna* image b) Noisy image: $\rho = 40$ c) PM d) TV e) BSDE (5) f) Original *peppers* image g) Noisy image: $\rho = 30$ h) PM i) TV j) BSDE (5) k) BSDE (6): $c = 0.6$, SSIM=0.8641, PSNR=28.2813 l) BSDE (6): $c = 0.8$, SSIM=0.8561, PSNR=27.5979

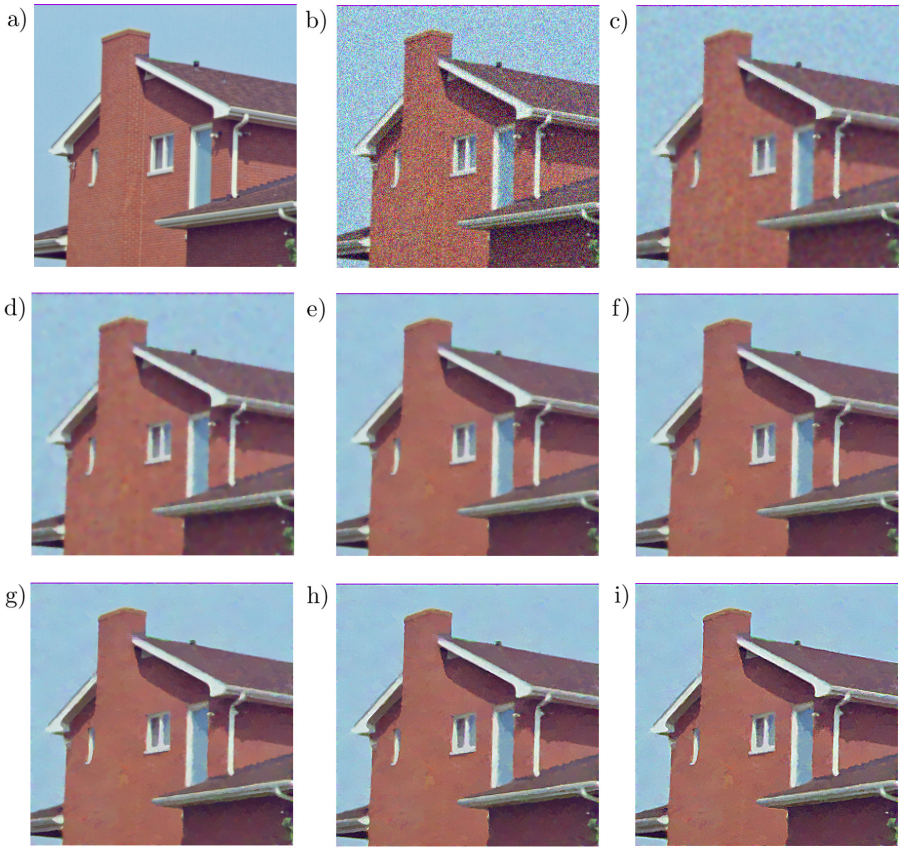

Fig. 2. a) Original *house* image b) $\rho = 30$, c) PM d) TV e) BSDE (5) f) BSDE (6): $c = 0.4$: SSIM=0.7762, PSNR=29.5289 g) BSDE (6): $c = 0.6$, SSIM=0.07589, PSNR=29.1452 h) BSDE (6): $c = 0.8$, SSIM=0.7270, PSNR=28.2127 i) BSDE (6): $c = 1$, SSIM=0.6760, PSNR=26.8126

(independent all channels) with the Gaussian noise with standard deviation ρ. Noisy images have been reconstructed with using vector analysis in RGB space. The maximum values of SSIM and PSNR are given in tables. PSNR is defined by the following formula:

$$\mathrm{PSNR}(U, \hat{U}) = 10 \log_{10} \left(\frac{255^2}{\mathrm{MSE}(U, \hat{U})} \right),$$

$$\mathrm{MSE}(U, \hat{U}) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} \|U(i,j) - \hat{U}(i,j)\|^2}{3 \cdot \mathrm{N} \cdot \mathrm{M}}, \quad \|(r,g,b)\| = \sqrt{r^2 + g^2 + b^2},$$

where M, N are the image dimensions, $U(i,j)$ and $\hat{U}(i,j)$ denote the original and the restored RGB vector, respectively. Definition of SSIM error to gray scale

can be found in [12]. In order to count SSIM to RGB color space, we apply SSIM measure to each individual color component and next we average the result [2]. Parameters of SSIM were set to the default values as recommended by [12].

Fig. 1 c,d,e,h,i,j, and Fig. 2 c,d,e, show images with maximum value of SSIM. In Fig. 1 k,l, and Fig. 2 f,g,h,i, we can see results obtained after applying backward diffusion (6) to Fig. 1 j, and Fig. 2 e, respectively. It is clear that if parameter c is greater than the result is more enhancing.

6 Conclusion

In this paper we have introduced a new method of colour image reconstruction. The idea presented here is the alternative to PDE vector-valued models and provide a new methodology based on advanced tools of stochastic analysis. Comparing figures one can observe that images created by the stochastic methods are visually more pleasant. The reason for it is that PDE methods show clear evidence of a block image, but this stair-case effect is reduced in our algorithm. Moreover analysing the measuring of image quality shows that BSDE methods perform better (for SSIM test) or are comparable to results of TV method (for PSNR test).

References

1. Abraham, R., Riviere, O.: Forward-backward stochastic differential equations and PDE with gradient dependent second order coefficients. ESAIM P&S 10, 184–205 (2006)
2. Åström, F., Felsberg, M., Lenz, R.: Color Persistent Anisotropic Diffusion of Images. In: Heyden, A., Kahl, F. (eds.) SCIA 2011. LNCS, vol. 6688, pp. 262–272. Springer, Heidelberg (2011)
3. Aubert, G., Kornprobst, P.: Mathematical problems in image processing, 2nd edn. Applied Mathematical Sciences, vol. 147. Springer (2006)
4. Borkowski, D.: Smoothing, Enhancing Filters in Terms of Backward Stochastic Differential Equations. In: Bolc, L., Tadeusiewicz, R., Chmielewski, L.J., Wojciechowski, K. (eds.) ICCVG 2010, Part I. LNCS, vol. 6374, pp. 233–240. Springer, Heidelberg (2010)
5. Deriche, R., Tschumperlé, D.: Diffusion PDE's on vector-valued images: local approach and geometric viewpoint. IEEE Signal Processing Magazine 19(5), 16–25 (2002)
6. Di Zenzo, S.: A note on the gradient of a multi-image. Comput. Vis. Graph. Image Process. 33(1), 116–125 (1986)
7. Pardoux, É.: Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In: Stochastic Analysis and Related Topics, VI, Geilo, vol. 42, pp. 79–127 (1998)
8. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12(7), 629–639 (1990)
9. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D 60(1–4), 259–268 (1992)

10. Sapiro, G., Ringach, D.L.: Anisotropic diffusion of multivalued images with applications to color filtering. IEEE Trans. Image Process. 5(11), 1582–1585 (1996)
11. Tanaka, H.: Stochastic differential equations with reflecting boundary condition in convex regions. Hiroshima Math. J. 9(1), 163–177 (1979)
12. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Process. 13(4), 600–612 (2004)
13. Weickert, J.: Theoretical Foundations Of Anisotropic Diffusion In Image Processing. Computing Suppement 11, 221–236 (1996)