

TENSOR-BASED IMAGE DIFFUSIONS DERIVED FROM GENERALIZATIONS OF THE TOTAL VARIATION AND BELTRAMI FUNCTIONALS

Anastasios Roussos and Petros Maragos

School of E.C.E., National Technical University of Athens, Greece

ABSTRACT

We introduce a novel functional for vector-valued images that generalizes several variational methods, such as the Total Variation and Beltrami Functionals. This functional is based on the structure tensor that describes the geometry of image structures within the neighborhood of each point. We first generalize the Beltrami functional based on the image patches and using embeddings in high dimensional spaces. Proceeding to the most general form of the proposed functional, we prove that its minimization leads to a nonlinear anisotropic diffusion that is regularized, in the sense that its diffusion tensor contains convolutions with a kernel. Using this result we propose two novel diffusion methods, the *Generalized Beltrami Flow* and the *Tensor Total Variation*. These methods combine the advantages of the variational approaches with those of the tensor-based diffusion approaches.

1. INTRODUCTION

Multiscale image analysis has been proven very useful in many image and vision applications. The first scale-spaces were linear and generated using Gaussian convolutions, which can be modeled by the *linear diffusion* [1]. Afterwards, various nonlinear modifications of the heat diffusion have been introduced, so that the diffusion respects the semantically important image features (see [2, 3] for detailed reviews).

Perona and Malik [4] proposed to apply the following type of nonlinear diffusion to the input image:

$$\partial u / \partial t = \operatorname{div} (g(|\nabla u|^2) \nabla u) \quad (1)$$

where the diffusivity function g is decreasing. The motivation was to favor intraregion over interregion smoothing. Two problems with this model are the amplification of noise by the diffusion coefficient and the sensitivity to initial conditions. In order to overcome them, Catté et al. [5] regularized the model by replacing $g(|\nabla u|^2)$ with $g(|\nabla G * u|^2)$ where G is a 2D Gaussian kernel.

Various image diffusion methods can be also derived from a variational framework, by minimizing a suitable functional of the image [6, 7]. A well-studied functional is the *Total Variation* (TV):

$$E[u] = \int_{\Omega} |\nabla u| \, d\mathbf{x} \quad (2)$$

where Ω is the image domain. Minimizing this functional using the Euler-Lagrange equations and the steepest descent method leads to a Partial Differential Equation (PDE) that is a special case of (1), with $g(s^2) = 1/s$. The TV (2) does not penalize discontinuities, but only strong oscillations, therefore the noise can be removed without blurring the edges. On the other hand, this model oversmooths homogenous regions and creates staircase artifacts.

In some more recent approaches, the image diffusion is not only nonlinear but also anisotropic, i.e. it is driven by an image dependent anisotropic tensor [2, 8]. This offers improved enhancement properties, since the diffusion can be more flexibly controlled and oriented according to the image structures.

Comparing the variational with the rest diffusion methods, it can be noted that the existence of a link between a diffusion method and a functional minimization offers several advantages. First, it offers better interpretation and makes the method less heuristic. Also, it facilitates the usage of the diffusion method in various image and vision problems, such as image restoration, inpainting, interpolation, that can be formulated as functional minimizations with constraints imposed by the problem. Further, it is easier to derive methods with reduced parameters (e.g. the steepest descent of the TV (2) has no parameters). Finally, the variational interpretation of a method can offer efficient implementations based on optimization techniques. Nevertheless, for several interesting types of diffusion methods, like regularized nonlinear diffusion [5] or nonlinear anisotropic diffusion [2], the corresponding PDEs have been directly designed, without any known variational interpretation.

Having the above as motivation, we propose a novel functional for vector-valued images that generalizes the Total Variation, the Beltrami Functional and other variational methods. It is based on the structure tensor that describes the geometry of image structures within the neighborhood of each point. We begin with a generalization of the Beltrami functional by representing an image with embeddings in high dimensional spaces that contain the image patches. Afterwards, we provide the general form of the new functional and prove that its minimization leads to a nonlinear anisotropic diffusion whose tensor is formed by using convolutions with a kernel. Applying this theoretical result to special cases, we propose two novel diffusion methods, the *Generalized Beltrami Flow* and the *Tensor Total Variation*. In contrast to other variational methods, e.g. [6, 7], the diffusion in the proposed methods is more efficiently controlled, since the diffusion tensors incorporate information about the image variation in the local neighborhood of each point. Also, as compared to other tensor-based diffusion approaches [2, 8], these methods have the advantage that come from a functional minimization. In parallel, the conducted experiments demonstrate the potential of the proposed diffusions as applied to the problem of image denoising.

2. PATCH-BASED GENERALIZATION OF THE BELTRAMI FUNCTIONAL

Sochen et. al. [7] interpret a vector-valued image \mathbf{u} with N channels as the 2D surface (\mathbf{x}, \mathbf{u}) embedded in a $N+2$ dimensional space. They propose to evolve the image according to the so-called *Beltrami flow*, which is an anisotropic diffusion flow towards the minimization of the area of this surface, using the induced metric. The *Beltrami functional* of this area offers an efficient and elegant way to couple the image channels and to extend in the vector-valued case

This work was supported in part by the EU projects FP6-IST-021324 and FP7-ICT-3-231135.

the properties of the Total Variation. The tensor field that controls the Beltrami flow combines the information from all channels but on each image point separately; it does not take into account the image variation in the neighborhood of the point. This lack of spatial extension has negative effects on the robustness to noise and on the enhancement of the image edges.

In order to overcome these limitations, we generalize the Beltrami functional by using higher dimensional mappings of the form $(\mathbf{x}, \mathcal{P}^u(\mathbf{x}))$ where the vector $\mathcal{P}^u(\mathbf{x})$ is an image patch [9] that contains weighted image values not only at point \mathbf{x} but also at points in a window around it. In this way, each point contributes to the area of the embedded surface by considering the image variation in its neighborhood. As it will be shown in Sec. 6.1, if we let the patches to sample the image with a step that tends to zero, the area of the embedded surface tends to the following functional:

$$A[\mathbf{u}] = \int_{\Omega} \sqrt{(\alpha^2 + \lambda_1)(\alpha^2 + \lambda_2)} d\mathbf{x} \quad (3)$$

where α is a constant, $\mathbf{u}(\mathbf{x}) = [u_1(\mathbf{x}) \cdots u_N(\mathbf{x})] : \Omega \rightarrow \mathbb{R}^N$ is the vector-valued image and $\lambda_i = \lambda_i(J_K(\nabla \mathbf{u}))$, $\lambda_1 \geq \lambda_2$, are the eigenvalues of the *structure tensor*:

$$J_K(\nabla \mathbf{u}) = K * \sum_{i=1}^N \nabla u_i \otimes \nabla u_i, \quad (4)$$

where K is a convolution kernel (e.g. a 2D gaussian). This tensor measures the geometry of image structures in the neighborhood of each point [2]. Note that the classic Beltrami functional corresponds to the case where the kernel K is the Dirac delta function $\delta(\mathbf{x})$. The generalized Beltrami functional (3) is a special case of the more general functional that will be introduced in the next section, where we provide a result regarding the functional minimization. In Sec. 4.2, we use this result to derive the *Generalized Beltrami Flow* method.

3. A GENERALIZED FUNCTIONAL BASED ON THE STRUCTURE TENSOR

Bearing in mind the analysis of the previous sections, we propose the following generalized functional:

$$E[\mathbf{u}] = \int_{\Omega} \psi(\lambda_1(J_K(\nabla \mathbf{u})), \lambda_2(J_K(\nabla \mathbf{u}))) d\mathbf{x} \quad (5)$$

where $\psi(\lambda_1, \lambda_2) : (\mathbb{R}^+)^2 \rightarrow \mathbb{R}^+$ is a cost function that is increasing w.r.t. both its arguments. The eigenvalues λ_i that are used in (5) measure the maximum and minimum vectorial variation of the image over a specific direction in the neighborhood of each point.

The difficulty in the theoretical analysis of the functional (5) is that, in contrast to most of the related variational methods (e.g. [6, 7, 8]), it integrates quantities that depend not directly on the image derivatives but on their convolutions with a kernel. Thus, the functional gradients that are needed for the minimization cannot be derived using the Euler-Lagrange equations. Nevertheless, using the definition of the functional gradient and some more elaborate calculus (see Sec. 6.2), we have proven the following theorem:

Theorem 1. *The functional gradient of $E[\mathbf{u}]$ (5) w.r.t. each image component $u_i(\mathbf{x})$ is:*

$$\delta E[\mathbf{u}]/\delta u_i = -\text{div}(D_K \nabla u_i), \quad i = 1, \dots, N, \quad (6)$$

$$\text{where: } D_K = K * \left(2 \frac{\partial \psi}{\partial \lambda_1} \boldsymbol{\theta}_1 \otimes \boldsymbol{\theta}_1 + 2 \frac{\partial \psi}{\partial \lambda_2} \boldsymbol{\theta}_2 \otimes \boldsymbol{\theta}_2 \right) \quad (7)$$

and $\boldsymbol{\theta}_{1,2}(\mathbf{x})$ are the eigenvectors of $J_K(\nabla \mathbf{u})$.

Therefore, the gradient descent flow for the minimization of the proposed functional (5), $\partial u_i/\partial t = -\delta E[\mathbf{u}]/\delta u_i$, is a novel general type of anisotropic diffusion of vector-valued images. The tensor D_K of this diffusion is formed using convolutions with the kernel K in two phases: **1**) in the formation of the structure tensor $J_K(\nabla \mathbf{u})$, from where its eigenvalues λ_i and eigenvectors $\boldsymbol{\theta}_i$ are computed and **2**) in the convolution of Eq. (7) that finally yields the diffusion tensor. The eigenvectors $\boldsymbol{\theta}_1(\mathbf{x})$ and $\boldsymbol{\theta}_2(\mathbf{x})$ correspond to the directions of maximum and minimum vectorial image variation, as this variation is viewed in a window around \mathbf{x} . In the next section, the above result is applied to some interesting special cases of the functional (5), resulting to novel diffusion methods.

4. NOVEL METHODS DERIVED AS SPECIAL CASES OF THE PROPOSED FUNCTIONAL

4.1. Tensor Total Variation

An effective choice for the cost function of the generic functional (5) is $\psi(\lambda_1, \lambda_2) = \sqrt{\lambda_1} + \sqrt{\lambda_2}$. Applying Theorem 1, we conclude that the steepest descent for the minimization of this functional is (for $i = 1, \dots, N$):

$$\frac{\partial u_i}{\partial t} = \text{div} \left(\left[K * \left(\frac{1}{\sqrt{\lambda_1}} \boldsymbol{\theta}_1 \otimes \boldsymbol{\theta}_1 + \frac{1}{\sqrt{\lambda_2}} \boldsymbol{\theta}_2 \otimes \boldsymbol{\theta}_2 \right) \right] \nabla u_i \right) \quad (8)$$

We call the derived method *Tensor Total Variation*, since it generalizes the TV and both the functional and the corresponding diffusion are based on tensors. The classic TV corresponds to the special case where $N=1$ (graylevel images) and $K = \delta(\mathbf{x})$. In this case, $\lambda_1 = |\nabla u|^2$ and $\lambda_2 = 0$, therefore $\psi(\lambda_1, \lambda_2) = |\nabla u|$.

The PDE (8) adaptively smooths a vector-valued image in the following way. The diffusion is strong and isotropic in the homogeneous image regions (small λ_1 , small λ_2), but weaker and mainly oriented by image structures in the vicinity of the edges (large λ_1 , small λ_2). Finally near image corners (large λ_1 , large λ_2), the diffusion is even weaker. The proposed *Tensor TV* combines the advantages of the TV method and the tensor-based diffusion methods: The diffusion comes from a variational formulation and is controlled by a tensor that flexibly adapts to the image structures.

4.2. Generalized Beltrami Flow

The Generalized Beltrami Functional (3) is derived from the functional (5) by setting $\psi(\lambda_1, \lambda_2) = \sqrt{(\alpha^2 + \lambda_1)(\alpha^2 + \lambda_2)}$. Theorem 1 implies that the gradient descent flow towards the minimization of this functional is given by (for $i = 1, \dots, N$):

$$\frac{\partial u_i}{\partial t} = \text{div} \left(\left[K * \left(\sqrt{\frac{\alpha^2 + \lambda_2}{\alpha^2 + \lambda_1}} \boldsymbol{\theta}_1 \otimes \boldsymbol{\theta}_1 + \sqrt{\frac{\alpha^2 + \lambda_1}{\alpha^2 + \lambda_2}} \boldsymbol{\theta}_2 \otimes \boldsymbol{\theta}_2 \right) \right] \nabla u_i \right)$$

Note that the Beltrami Flow of [7] is derived if we set in the above equation $K = \delta(\mathbf{x})$ but also add to its right-hand side (RHS) the multiplying term $1/\psi(\lambda_1, \lambda_2)$. This extra weighting is due to the fact that in [7] the gradient descent is applied in the space of embeddings instead of the space of images. But due to this weighting, the diffusion is too much reduced near image edges and cannot effectively enhance them.

4.3. Other interesting cases and relations to previous works

If in the functional (5) we choose $\psi(\lambda_1, \lambda_2) = \phi(\lambda_1 + \lambda_2)$ for some function ϕ , then the gradient descent is described by a diffusion that is nonlinear but isotropic:

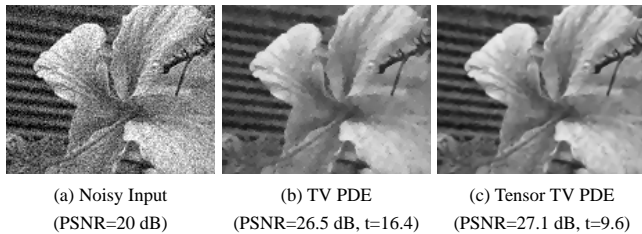


Fig. 1. graylevel image denoising using *TV* and *Tensor TV* methods.

$$\partial u_i / \partial t = \text{div} \left(2 \left[K * \varphi' \left(K * \|\nabla u\|^2 \right) \right] \nabla u_i \right) \quad (9)$$

where $\|\nabla u\|^2 = \sum_{i=1}^N |\nabla u_i|^2$. This case offers the following alternative to the regularization of the Perona-Malik model (1) by Catté et al. [5]: The diffusion coefficient $g(|\nabla u|^2)$ is replaced by $K * g(K * |\nabla u|^2)$ instead of $g(|\nabla K * u|^2)$. The advantage of this novel regularization is that it has a variational interpretation.

Note finally that a simpler form of the proposed functional that does not contain any convolution with a kernel has been already studied in [8]. In this case, the corresponding diffusion is anisotropic only if the number of image channels is $N \geq 2$. Also, its diffusion tensor does not incorporate information about the image variation in the vicinity of each image point.

5. EXPERIMENTAL RESULTS AND COMPARISONS

We conduct denoising experiments, adopting the following framework: Each time we use a noise-free reference image and we corrupt it with additive gaussian noise. The derived noisy image is used as input to the diffusion methods that are compared. Since the ground truth is available, we compute Peak Signal to Noise Ratios (PSNR) during the evolutions of the corresponding PDEs and we consider as output of each method the image from this flow that achieves the maximum PSNR. Of course, in practice the diffusion stopping time should be estimated, but the choice made here corresponds to the best-case scenario for each method. In all experiments, the kernel K that is used in the proposed methods was chosen to be a 2D gaussian with fixed standard deviation $\rho=0.5$ pixels.

In Fig. 1, the proposed Tensor TV and the classic TV are compared in a graylevel image experiment. We observe that the proposed method enhances the edges more effectively and generally yields a more plausible result. In addition, it achieves a higher maximum PSNR measure during the PDE evolution. This maximum corresponds to a smaller diffusion time t (less iterations), which reflects the fact that the Tensor TV exploits the convolutions with the kernel K in order to flexibly adapt to the image and to improve the robustness against noise.

Figure 2 demonstrates the application of the classic and the proposed generalized Beltrami Flows to a noisy color image. The proposed generalization seems to smooth the various image regions in a more balanced way and to remove the noise by introducing less artifacts. In addition, it yields an improved PSNR performance.

Finally, we followed the aforementioned experimental framework using 23 natural images¹. We run two series of denoising experiments, based on the graylevel and color versions of these images. Figure 3 demonstrates overall PSNR measures for each tested method and noise level of the input images. We observe that the proposed Tensor TV yields in all cases the best overall measures. The second proposed method, Generalized Beltrami Flow, yields worse

¹These images are from <http://www.cipr.ppi.edu/resource/stills/kodak.html>.

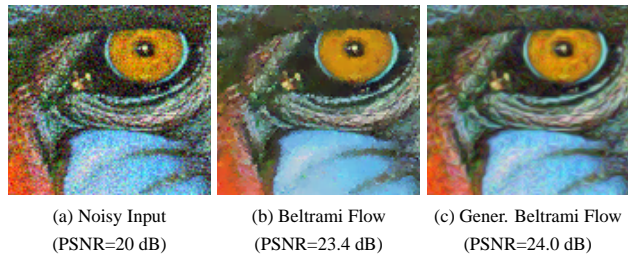


Fig. 2. Color image denoising example.

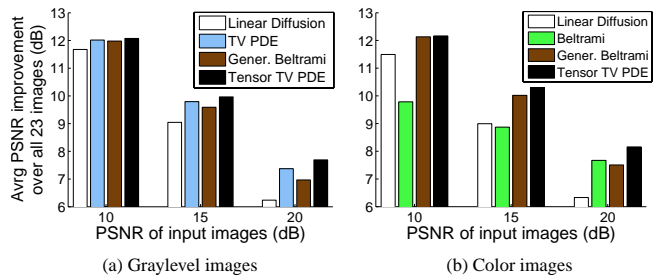


Fig. 3. Performance measures of different diffusion methods.

measures than the TV PDE in the experiments with graylevel images. On the other hand, as compared to the Beltrami flow in the case of color images, it demonstrates a significantly improved performance when the noise levels are relatively high (low input PSNR).

6. THEORETICAL DETAILS

6.1. Derivation of the Generalized Beltrami Functional

Here, we more formally derive the Generalized Beltrami Functional (3) as a limit of the area of the patch-based embeddings. For each point $x \in \Omega$, we define a sequence $\{\mathcal{P}_n^u(x)\}_{n=1,2,\dots}$ of image patches around x by **1)** sampling all the image channels inside the square $S_x = [x-L, x+L] \times [y-L, y+L]$, for some fixed $L > 0$, in the points of a regular grid with horizontal and vertical spacing $\epsilon_n = L/n$, **2)** weighting the samples with a window function $w(x)$ and **3)** forming a vector from these values. The function $w(x)$ is assumed to have a support that is a subset of the square $[-L, L]^2$ and to satisfy the symmetry property $w(x, y) = w(|x|, |y|)$. We are interested in the limit case $n \rightarrow \infty$, where $\epsilon_n \rightarrow 0$ and the patch tends to contain weighted versions of the image values at all the points of the square S_x . Based on the above sequence of image patches, we define the following sequence of embedded surfaces $F_n : \Omega \rightarrow \mathbb{R}^{D_n}$:

$$x \rightarrow F_n(x) = [\alpha x, \epsilon_n \mathcal{P}_n^u(x)], \quad (10)$$

where $D_n = 2 + N(2n + 1)^2$ and α is a positive constant that in practice is chosen to be small compared to the vectorial variations of the image. The weighting of the patches $\mathcal{P}_n^u(x)$ with ϵ_n in (10) serves as a compensation for the fact that their dimensionality grows with n . For every F_n , if we consider the Euclidean metric on the embedding space \mathbb{R}^{D_n} , the induced metric on Ω is given by:

$$\mathcal{G}_n(x) = \sum_{j=1}^{D_n} \nabla F_n^j \otimes \nabla F_n^j = \alpha^2 I_2 + \dots$$

$$\sum_{i=1}^N \sum_{k=-n}^n \sum_{\ell=-n}^n \epsilon_n^2 w^2(k\epsilon_n, \ell\epsilon_n) \nabla u_i \otimes \nabla u_i(x+k\epsilon_n, y+\ell\epsilon_n)$$

where F_n^i is the i -th component of $\mathbf{F}_n(\mathbf{x})$, I_2 is the 2×2 identity matrix and “ \otimes ” denotes tensor product. The double sum that corresponds to each image channel i at the above equation is a classic 2D Riemann sum. Given that specific conditions hold for \mathbf{u} and w (e.g. all their 1st derivatives are bounded and continuous almost everywhere), the sequence of these sums converges, as $n \rightarrow \infty$, to the convolution with the kernel $K(\mathbf{x}) \triangleq w^2(\mathbf{x})$. Thus, for every \mathbf{x} :

$$\mathcal{G}(\mathbf{x}) = \lim_{n \rightarrow \infty} \mathcal{G}_n(\mathbf{x}) = \alpha^2 I_2 + J_K(\nabla \mathbf{u}), \quad (11)$$

where $J_K(\nabla \mathbf{u})$ is the image’s structure tensor (4). We can choose the weight function $w(\mathbf{x})$ so that a desirable kernel K is used in $J_K(\nabla \mathbf{u})$ (e.g. K can be an isotropic 2D gaussian kernel that is truncated in the square $[-L, L]^2$). Using the metric \mathcal{G}_n , we can express the area of each embedded surface \mathbf{F}_n as $A_n = \int_{\Omega} \sqrt{\det(\mathcal{G}_n)}$ [7]. Therefore:

$$A[\mathbf{u}] = \lim_{n \rightarrow \infty} A_n = \int_{\Omega} \sqrt{\det(\mathcal{G})} \stackrel{(11)}{=} \int_{\Omega} \sqrt{(\alpha^2 + \lambda_1)(\alpha^2 + \lambda_2)},$$

since as it can be shown, the above limit commutes with the integration.

6.2. Sketch of the Proof of Theorem 1

Let $\mathbf{v}(\mathbf{x}) = [v_1, \dots, v_N]$ be an arbitrary vector-valued test function that vanishes on $\partial\Omega$, $V(\mathbf{x}) = \text{diag}(\mathbf{v})$ and $\varepsilon \in \mathbb{R}^N$. If we define $\Phi(\varepsilon) = E[\mathbf{u} + V\varepsilon]$, then the functional gradient of $E[u]$ w.r.t. each $u_i(\mathbf{x})$ must by definition satisfy:

$$\left. \frac{\partial \Phi(\varepsilon)}{\partial \varepsilon_i} \right|_{\varepsilon=0} = \int_{\Omega} v_i(\mathbf{x}) \frac{\delta E[\mathbf{u}]}{\delta u_i} d\mathbf{x} \quad (12)$$

for any test function $\mathbf{v}(\mathbf{x})$. On the other hand, this first variation of $E[u]$ can be written also as (using Eq. (5)):

$$\left. \frac{\partial \Phi(\varepsilon)}{\partial \varepsilon_i} \right|_{\varepsilon=0} = \int_{\Omega} d\mathbf{x} \sum_{j=1}^2 \frac{\partial \psi}{\partial \lambda_j} \frac{\partial \lambda_j(J_K(\mathbf{u} + V\varepsilon))}{\partial \varepsilon_i} \Big|_{\varepsilon=0} \quad (13)$$

again for any test function $\mathbf{v}(\mathbf{x})$. If $a_{k\ell}(\mathbf{u})$ are the elements of the structure tensor $J_K(\mathbf{u})$ (4), its eigenvalues are given from $\lambda_j(J_K(\mathbf{u})) = (a_{11} + a_{22} + C_j \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2})/2$, $j=1,2$, where $C_1 = 1$ and $C_2 = -1$ [2]. Using the definition (4), we can show that (for $k, \ell=1,2$ and $i=1, \dots, N$):

$$\left. \frac{\partial a_{k\ell}(\mathbf{u} + V\varepsilon)}{\partial \varepsilon_i} \right|_{\varepsilon=0} = K * \left(\frac{\partial u_i}{\partial x_k} \frac{\partial v_i}{\partial x_\ell} + \frac{\partial u_i}{\partial x_\ell} \frac{\partial v_i}{\partial x_k} \right). \quad (14)$$

We thus obtain the following (for $j=1,2$):

$$\left. \frac{\partial \lambda_j(J_K(\mathbf{u} + V\varepsilon))}{\partial \varepsilon_i} \right|_{\varepsilon=0} = [(a_{11} - a_{22})K * \langle M_1 \nabla u_i, \nabla v_i \rangle + 2a_{12}K * \langle M_2 \nabla u_i, \nabla v_i \rangle] \frac{C_j}{\lambda_1 - \lambda_2} + K * \langle \nabla u_i, \nabla v_i \rangle \quad (15)$$

where $M_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Based on the last equation, we observe that the RHS of Eq. (13) can be expressed as a sum of 6 integrals (3 for each j) that all have the form:

$$I = \int_{\Omega} f(\mathbf{x}) K * \langle M \nabla u_i, \nabla v_i \rangle d\mathbf{x}, \quad (16)$$

for some appropriate function $f(\mathbf{x})$ and 2×2 matrix M . At this point, we need a formal definition of the convolutions with K . Referring to the definition of the structure tensor (4), where these convolutions are initially used, we first extend the image \mathbf{u} to \mathbb{R}^2 by

reflection followed by periodization and afterwards use the classic continuous 2D convolution. It can be shown that this process is equivalent to an integral transform with integration domain just the image domain Ω and using an integral kernel $\hat{K}(\boldsymbol{\tau}; \mathbf{x})$ that is derived from $K(\mathbf{x} - \boldsymbol{\tau})$, after appropriate warpings and replications. By assuming that the kernel K has the symmetric property $K(x, y) = K(|x|, |y|)$, it is not hard to show that $\hat{K}(\boldsymbol{\tau}; \mathbf{x}) = \hat{K}(\mathbf{x}; \boldsymbol{\tau})$. Having these in mind and using integration by parts and the fact that v_i vanishes on $\partial\Omega$, we can write the integral I as follows:

$$I = \int_{\Omega} f(\mathbf{x}) \left\{ \int_{\Omega} \langle M \nabla u_i(\boldsymbol{\tau}), \nabla v_i(\boldsymbol{\tau}) \rangle \hat{K}(\boldsymbol{\tau}; \mathbf{x}) d\boldsymbol{\tau} \right\} d\mathbf{x} \\ = - \int_{\Omega} v_i(\mathbf{x}) \text{div}((K * f) M \nabla u_i) d\mathbf{x}$$

Using Eqs. (12),(13),(15), the above result and the fundamental lemma of calculus of variations, we conclude that (for $i = 1, \dots, N$):

$$\delta E[\mathbf{u}]/\delta u_i = -\text{div}(D \nabla u_i), \quad \text{where:} \quad (17)$$

$$D = \sum_{j=1}^2 \left\{ (K * \frac{\partial \psi}{\partial \lambda_j} \frac{(a_{11} - a_{22})C_j}{\lambda_1 - \lambda_2}) M_1 + \right. \\ \left. (K * \frac{\partial \psi}{\partial \lambda_j} \frac{2a_{12}C_j}{\lambda_1 - \lambda_2}) M_2 + (K * \frac{\partial \psi}{\partial \lambda_j}) I_2 \right\} \\ = K * \sum_{j=1}^2 \frac{\partial \psi}{\partial \lambda_j} \underbrace{\left(I_2 + \frac{a_{11} - a_{22}}{\lambda_1 - \lambda_2} C_j M_1 + \frac{2a_{12}}{\lambda_1 - \lambda_2} C_j M_2 \right)}_{B_j}$$

After some algebra, we obtain that the eigenvalues of the matrix B_j , $j=1,2$, are equal to 2 and 0 and the unit eigenvector that corresponds to the nonzero eigenvalue is $\boldsymbol{\theta}_j$. Therefore $B_j = 2\boldsymbol{\theta}_j \otimes \boldsymbol{\theta}_j$. This completes the proof.

7. CONCLUSIONS

We introduced a generic functional for vector-valued images and proved that its minimization yields a tensor-based diffusion process. Using this result, we proposed two novel anisotropic diffusion methods. The presented experiments demonstrated the effectiveness of these methods when applied to image denoising. Thanks to the variational formulation of the adopted framework, such regularized tensor-based diffusions can be readily applied to various other image processing problems, such as image restoration, inpainting and interpolation. This could be a direction for future research.

8. REFERENCES

- [1] J.J. Koenderink, “The structure of images,” *Biological Cybernetics*, vol. 50, pp. 363–370, 1984.
- [2] J. Weickert, *Anisotropic Diffusion in Image Processing*, Teubner Stuttgart, 1998.
- [3] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of *Applied Math. Sciences*, Springer-Verlag, 2002.
- [4] P. Perona and J. Malik, “Scale space and edge detection using anisotropic diffusion,” *IEEE Trans. PAMI*, vol. 12, no. 7, pp. 629–639, July 1990.
- [5] F. Catté, P.L. Lions, J.M. Morel, and T. Coll, “Image selective smoothing and edge detection by nonlinear diffusion,” *SIAM J. Numer. Anal.*, vol. 29, no. 1, pp. 182–193, 1992.
- [6] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, pp. 259–268, 1992.
- [7] N. Sochen, R. Kimmel, and R. Malladi, “A general framework for low level vision,” *IEEE Trans. Im. Proc.*, vol. 7, pp. 310–338, 1998.
- [8] D. Tschumperlé and R. Deriche, “Vector-valued image regularization with PDE’s: A common framework for different applications,” *IEEE T-PAMI*, vol. 27, no. 4, 2005.
- [9] D. Tschumperlé and L. Brun, “Non-local image smoothing by applying anisotropic diffusion pde’s in the space of patches,” in *ICIP*, 2009.