INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 6, Number 3, Pages 389–401

FINITE ELEMENT APPROXIMATION OF THE GRADIENT FLOW FOR A CLASS OF LINEAR GROWTH ENERGIES WITH APPLICATIONS TO COLOR IMAGE DENOISING

XIAOBING FENG AND MIUN YOON

This paper is dedicated to Professor Roland Glowinski on the occasion of his seventieth birthday

Abstract. This paper concerns with the finite element approximation of a nonlinear second order parabolic system which describes the L^2-gradient flow for a class of linear growth energy functionals. Besides their appeals in differential geometry and calculus of variations, linear growth energy functionals and their gradient flows also arise naturally from emerging applications of image processing such as color image denoising. In this paper, we introduce a family of variational models for color image denoising which minimize linear growth energy functionals of maps into the unit sphere in $\mathbf{R}^3$. These models generalize the popular 1-harmonic map model which has been studied intensively in recent years. To compute the solutions of the variational models, we first derive their L^2-gradient flow equations and then introduce some fully discrete implicit finite element method for the gradient flow equations. It is proved that the proposed finite element method is uniquely solvable and absolutely stable, and the finite element solution converges to the PDE solution as the mesh sizes tend to zero. Numerical experiments are presented to demonstrate the effectiveness of the proposed variational models for color image denoising and to show the efficiency of the proposed finite element method. A numerical comparison of the proposed models with the channel-by-channel model is also presented.

Key Words. Linear growth energy functionals, gradient flow, p-harmonic maps, BV functions, color image denoising, finite element methods

1. Introduction

Image denoising and restoration are two most basic tasks in low level image processing. Tremendous progresses have been made in this area, particularly for gray images, in the past two decades. In addition to the further development in the traditional methods, techniques and algorithms, there has been an explosive development and growth of image processors based on partial differential equations (PDEs) and variational approach (cf. [4, 23, 6, 21] and the references therein). Compared with the traditional approaches, PDE and variational methods have remarkable advantages in both theory and computation. It allows to directly handle and process visually important geometric features such as gradients, tangents and curvatures, and to model visually meaningful dynamic process such as linear

Received by the editors October 30, 2008 and, in revised form, January 5, 2009.
2000 *Mathematics Subject Classification.* 65M12, 65M60, 35K65, 58E20.
This work was partially supported by the NSF grants DMS-0410266 and DMS-0710831.

and nonlinear diffusion. Computationally, it can greatly benefit from the existing wealthy amount numerical methods for PDEs (cf. [4, 20, 19, 21]).

A color image is often represented by a vector-valued image function $\mathbf{I}(x) = (r(x), g(x), b(x))$, where $r(x), g(x)$, and $b(x)$ denote the intensity values of three primary colors (Red, Green, Blue (RGB)) at an image pixel x. Suppose that $\mathbf{I}(x)$ describes a given scene which contains some random noise, the goal of color image denoising is to remove the noise such that the recovered image is as close as possible to the true image. One of the classical approaches for color image denoising is the median filter which works especially well for enhancing edges [18]. Recently, the median filter has been generalized to denoising chromaticity features on the unit sphere [28, 29]. Another approach is to treat the RGB color system directly as a vector space and to denoise it channel-by-channel [22, 8]. Most recent studies have proposed the use of the chromaticity and brightness decomposition (CBD). Indeed, the CBD approach has received a lot attention lately since it is well suited for denoising, edge detection, and segmentation, see [11, 27, 26, 28] and the references therein.

Mathematically, the CBD is nothing but the polar decomposition of the vector-valued image function. That is, we write

$$\mathbf{I}(x) = \rho(x)\mathbf{g}, \tag{1}$$

where

$$\rho(x) := |\mathbf{I}(x)| = \sqrt{r(x)^2 + g(x)^2 + b(x)^2}, \qquad \mathbf{g} := \frac{\mathbf{I}(x)}{\rho(x)}. \tag{2}$$

Therefore, $\rho(x)$, called the *brightness* of the image, is the length of the RGB color vector, and $\mathbf{g}$, called the *chromaticity* of the image, denotes the direction of the color vector which must lie on the unit sphere $\mathbf{S}^2$. One advantage of CBD approach is that it allows one to denoise the chromaticity and the brightness separately by different methods. For instance, one can use the well-known total variation (TV) model of Rudin-Osher-Fatemi [20] (also see [15, 16]) to denoise the rightness $\rho(x)$ and use another method to denoise the chromaticity $\mathbf{g}$. A number of authors have addressed color image denoising using directional diffusion of $\mathbf{R}^n$ vectors [8, 22, 24]. All these works extended the well established scalar diffusion flows [3, 20] in different forms for the vector-valued image and do not separate the chromaticity and brightness. Blomgren and Chan [8] proposed a new definition of the total variation norm for vector-valued functions which is the extension of the scalar TV norm and applied this new TV norm to restore color images. Some authors have used p-harmonic map flows for chromaticity denoising [30, 26, 5]. Most of these works considered the case $1 < p < \infty$ [30, 26]. Barrett, Feng and Prohl [5] proved the existence of weak solutions for the whole spectrum $1 \leq p < \infty$. Feng [14] extended the 1-harmonic map results of [5] to gradient flows of linear growth functionals. Chan, Kang and Shen [10] applied the general framework of non-flat TV denoising model [11] to chromaticity denoising.

The primary goals of this paper are to introduce a family of variational models for color image denoising and to develop some fully discrete finite element method for computing solutions of the proposed models. The proposed models use convex linear growth functionals instead of the p-energy functional (cf. Section 2). We recall that there are two constraints for a chromaticity denoising model, that is,

$$\int_\Omega (\mathbf{v}-\mathbf{g})\, dx\, dy \quad = 0, \tag{3}$$

$$\frac{1}{2}\int_\Omega (\mathbf{v}-\mathbf{g})^2\, dx\, dy \quad = \sigma^2 \tag{4}$$

where $\sigma > 0$ is the error level and usually given. The first constraint corresponds to the assumption that the noise has zero-mean, and the second constraint uses a *priori* information that the standard deviation of the noise $\eta(x,y) := \mathbf{u}-\mathbf{g}$ is σ. In the proposed models, we enforce the fidelity of an image using general L^q-norm $(1 \le q < \infty)$ instead of the usual L^2-norm.

The remainder of this paper is organized as follows. In Section 2, we introduce a family of variational models for color image denoising which minimize linear growth energy functionals of the maps into the unit sphere in $\mathbf{R}^3$. These models generalize the popular 1-harmonic map model which has been studied intensively in recent years. We then introduce the L^2-gradient flow, which is described by a system of nonlinear parabolic PDEs, for the energy functionals as a way to solve the variation problem, and present the mathematical framework and setting for the gradient flow. In Section 3, we propose a fully discrete finite element method for approximating the gradient flow and carry out its stability and convergence analysis. In Section 4, we present a few numerical experiments to show the good performance of the proposed models and numerical method. We conclude the paper by a few remarks in Section 5.

Standard function and space notation will be adopted in this paper, we refer to [9, 13] for their precise definitions.

2. Gradient flow for linear growth energy functionals with applications to color image denoising

Let $\mathbf{u},\mathbf{v} : \Omega \subset \mathbf{R}^m \to \mathbf{S}^{n-1} \subset \mathbf{R}^n$ be vector-valued functions, where Ω is a bounded domain with Lipschitz boundary $\partial\Omega$ and $\mathbf{S}^{n-1}$ denote the unit sphere in $\mathbf{R}^n$. For image processing applications, $m = 2$ and $n = 3$, but the framework presented here applies to the general case as well. Let $\mathbf{g} : \Omega \subset \mathbf{R}^m \to \mathbf{S}^{n-1} \subset \mathbf{R}^n$ be a given map satisfying $|\mathbf{g}| \le 1$ a.e. in Ω. $\mathbf{g}$ denotes a given noisy chromaticity of some (unknown) color image in applications. Following [5, 26, 30], we introduce the following variational models with linear growth energy functionals to "denoise" the given "chromaticity" $\mathbf{g}$: Find $\mathbf{u} \in W^{1,1}_N(\Omega, S^{n-1})$ such that

$$\mathbf{u} = \operatorname*{argmin}_{\mathbf{v}\in W^{1,1}_N(\Omega,\mathbf{S}^{n-1})} \mathcal{J}_{\beta,\lambda}(\mathbf{v}), \tag{5}$$

where

$$\mathcal{J}_{\beta,\lambda}(\mathbf{v}) := \beta\, \mathcal{E}_\varphi(\mathbf{v}) + \frac{\lambda}{q}\int_\Omega |\mathbf{v}-\mathbf{g}|^q\, dx \quad \text{for } \beta > 0,\ \lambda > 0,\ 1 \le q < \infty,$$

$$\mathcal{E}_\varphi(\mathbf{v}) := \int_\Omega \varphi(|\nabla \mathbf{v}|)\, dx,$$

$$W^{1,1}_N(\Omega, S^{n-1}) := \left\{ \mathbf{v} \in W^{1,1}(\Omega, S^{n-1});\ \frac{\partial \mathbf{v}(x)}{\partial \mathbf{n}} = 0 \text{ for a.e. } x \in \partial\Omega \right\},$$

λ is a Lagrangian multiplier which is induced by the constraint (4). The energy density function $\varphi : \mathbf{R}_+ \cup \{0\} \to \mathbf{R}_+ := (0,\infty)$ is assumed to be a real-valued, continuous, nondecreasing, convex, and linear growth function (cf. [14]). The last property implies that there exists positive constants c_i for $i = 1, 2, 3, 4$ such that

$$c_1 s - c_2 \leq \varphi(s) \leq c_3 s + c_4 \qquad \forall s \in \mathbf{R}_+ \cup \{0\}.$$

In addition, we assume that there exists $c_5 > 0$ such that

$$0 \leq \varphi'(s) \leq c_5 \qquad \forall s \in \mathbf{R}_+ \cup \{0\}.$$

The best known example of the density function φ is

$$\varphi(s) = s \qquad \forall\, s \in \mathbf{R}_+ \cup \{0\}. \tag{6}$$

In this case, the above model reduces to the so-called 1-harmonic map model studied in [5]. Another example is

$$\varphi(s) = \sqrt{s^2 + 1} \qquad \forall\, s \in \mathbf{R}_+ \cup \{0\}, \tag{7}$$

which is often called the minimal surface energy density.

We remark that the variational models (5) are slightly different from those proposed in [14] where the number q is set to be 2. Besides their appeal in differential geometry and calculus of variation, the rationale for using linear growth density function $\varphi(s)$ in (5) is similar to that for the TV model [20], that is, to allow jumps in image functions and to minimize the diffusion at the place where $|\nabla \mathbf{u}|$ is relatively large. Recall that the set where $|\nabla \mathbf{u}|$ is large contains the edges of the image which should be kept in the recovered image. We also note that since the exact Lagrangian multiplier λ is difficult to determine, it is often estimated or chosen a *priori* in practice [8, 11, 20].

It can be shown that the Euler-Lagrange equation of (5) is given by (cf. [14] for the case $q = 2$)

$$-\beta\, \mathrm{div} \mathcal{B} + \lambda |\mathbf{u} - \mathbf{g}|^{q-2} (\mathbf{u} - \mathbf{g}) = \mu_{\beta,\lambda} \mathbf{u} \qquad \text{in } \Omega, \tag{8}$$

$$|\mathbf{u}| = 1 \qquad \text{in } \Omega, \tag{9}$$

$$\mathcal{B}\mathbf{n} = 0 \qquad \text{on } \partial\Omega, \tag{10}$$

where

$$\begin{aligned} \mathcal{B} &:= \frac{\varphi'(|\nabla \mathbf{u}|)}{|\nabla \mathbf{u}|} \nabla \mathbf{u}, \\ \mu_{\beta,\lambda} &:= \beta \varphi'(|\nabla \mathbf{u}|) |\nabla \mathbf{u}| + \lambda |\mathbf{u} - \mathbf{g}|^{q-2} (1 - \mathbf{u} \cdot \mathbf{g}). \end{aligned}$$

The term on the right hand side of (8) is caused by the unit length constraint $|\mathbf{u}| = 1$ a.e. in Ω.

Let X be a Hilbert space and $\mathcal{F}$ be a (nonlinear) functional on X. Assume that $\mathcal{F}$ is differentiable with the Fréchet derivative $\mathcal{F}'$, the gradient flow of $\mathcal{F}$ is defined as (cf. [2]) seeking $u : (0, \infty) \to X$ such that

$$u'(t) = -\mathcal{F}'(u(t)).$$

If the above equation is interpreted in the dual space X' of X, testing the equation with $u'(t)$ immediately gives

$$\frac{d}{dt} \mathcal{F}(u(t)) = -\|u'(t)\|_{X'}^2 \leq 0,$$

where $\|\cdot\|_{X'}$ denotes the norm of the dual space X'. The above equation says that the "energy" $\mathcal{F}$ is decreasing along the flow $u(t)$. Under certain conditions on $\mathcal{F}$ (such as convexity) and on the initial data $u(0) = u_0$, it can be proved that $\lim_{t\to\infty} u(t)$ converges to a minimizer of the functional $\mathcal{F}$. As a result, the idea of gradient flow can be used to approximate solutions of variation problems such as

(5), although the identification of the Fréchet derivatives of the given functionals could be troublesome.

Nevertheless, it can be shown that (cf. [14] for the case $q = 2$) the L^2-gradient flow for the energy functional $\mathcal{J}_{\beta,\lambda}$ is given by

$$\mathbf{u}_t - \beta\,\mathrm{div}\mathcal{B} + \lambda|\mathbf{u}-\mathbf{g}|^{q-2}(\mathbf{u}-\mathbf{g}) = \mu_{\beta,\lambda}\mathbf{u} \qquad \text{in } \Omega_T := \Omega\times(0,T), \tag{11}$$

$$|\mathbf{u}| = 1 \qquad \text{in } \Omega_T, \tag{12}$$

$$\mathcal{B}\mathbf{n} = 0 \qquad \text{on } \partial\Omega_T := \partial\Omega\times(0,T), \tag{13}$$

$$\mathbf{u} = \mathbf{u}_0 \qquad \text{in } \Omega_0 = \Omega\times\{t=0\}, \tag{14}$$

where $\mathbf{u}_0$ is a given initial map. We remark that from the PDE point of view the idea of using linear growth density function $\varphi(s)$ in (5) is to minimize the diffusion produced by the first term in equation (8) and the second order term in equation (11) around the image edges where $|\nabla\mathbf{u}|$ is relatively large.

By inspecting the PDE system, one easily sees that three nonlinear terms appear in (11) which are expected to cause difficulties. The first one is the term involving $\mathcal{B}$ due to nonlinearity of φ, the second one is the fidelity term due to the nonlinearity of L^q-norm, the last one is the right-hand side due to the nonconvex constraint $|\mathbf{u}| = 1$. To handle the degeneracy of the leading term and the fidelity term in (11), we approximate the energy $\mathcal{E}_\varphi$ by the regularized energy $\mathcal{E}_\varphi^\varepsilon$ based on the approximation

$$|z| \approx |z|_\varepsilon := \sqrt{z^2+\varepsilon^2}.$$

To handle the nonconvex constraint $|\mathbf{u}| = 1$, we approximate it by the well known Ginzburg-Landau penalization [7], that is, instead of applying the exact constraint $|\mathbf{u}| = 1$, we enforce it approximately by adding a penalization term to the regularized energy $\mathcal{E}_\varphi^\varepsilon$. We then introduce the following regularized variation problem as an approximation to (5)

$$\mathbf{u}^{\varepsilon,\delta} = \underset{\mathbf{v}\in W_N^{1,1}(\Omega,\mathbf{R}^n)}{\mathrm{argmin}}\ \mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{v}) \qquad \text{for } \varepsilon,\ \delta > 0, \tag{15}$$

where

$$\mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{v}) := \beta\mathcal{E}_\varphi^{\varepsilon,\delta}(\mathbf{v}) + \frac{\lambda}{q}\int_\Omega |\mathbf{v}-\mathbf{g}|_\varepsilon^q\,dx, \tag{16}$$

$$\mathcal{E}_\varphi^{\varepsilon,\delta}(\mathbf{v}) := \frac{\varepsilon^\alpha}{2}\int_\Omega |\nabla\mathbf{v}|^2\ dx + \int_\Omega \varphi(|\nabla\mathbf{v}|_\varepsilon)\ dx + \frac{1}{\beta}\mathcal{L}^\delta(\mathbf{v}) \qquad (\alpha > 0), \tag{17}$$

$$\mathcal{L}^\delta(\mathbf{v}) := \frac{1}{\delta}\int_\Omega F(\mathbf{v})dx, \tag{18}$$

$$F(\mathbf{v}) := \frac{1}{4}(|\mathbf{v}|^2-1)^2. \tag{19}$$

Again, it can be shown that the gradient flow for the regularized energy functional $\mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}$ is given by (cf. [14] for the case $q = 2$)

$$\begin{aligned}\mathbf{u}_t^{\varepsilon,\delta} - \beta\,\mathrm{div}\mathcal{B}^{\varepsilon,\delta} + \lambda|\mathbf{u}^{\varepsilon,\delta}-\mathbf{g}|_\varepsilon^{q-2}(\mathbf{u}^{\varepsilon,\delta}-\mathbf{g})&\\ +\frac{1}{\delta}(|\mathbf{u}^{\varepsilon,\delta}|^2-1)\mathbf{u}^{\varepsilon,\delta} = 0 &\qquad \text{in } \Omega_T,\end{aligned} \tag{20}$$

$$\mathcal{B}^{\varepsilon,\delta}\mathbf{n} = 0 \qquad \text{on } \partial\Omega_T, \tag{21}$$

$$\mathbf{u}^{\varepsilon,\delta} = \mathbf{u}_0 \qquad \text{in } \Omega_0, \tag{22}$$

where

$$\mathcal{B}^{\varepsilon,\delta} := \frac{\varphi'(|\nabla \mathbf{u}^{\varepsilon,\delta}|_\varepsilon)}{|\nabla \mathbf{u}^{\varepsilon,\delta}|_\varepsilon} \nabla \mathbf{u}^{\varepsilon,\delta}.$$

It is easy to see that (20)-(22) is an approximation to the gradient flow (11)-(14). In [14] the first author of this paper gave a comprehensive analysis for the above regularized flow in the case $q = 2$. It was proved that for each fixed pair of positive numbers (ε, δ) the flow has global weak solutions provided that $|\mathbf{u}_0| \leq 1$ and $|\mathbf{g}| \leq 1$ a.e. in Ω, and has global classical solutions if the domain and the data are sufficiently smooth. Moreover, it was shown that (11)-(14) possesses global (in time) weak solutions in $BV(\Omega)$, the space of functions of bounded variations (cf. [1]), which was done by developing a much involved compactness technique and passing to the limit in (20)-(22) as $\varepsilon,\ \delta \to 0$. It is expected that the techniques and results of [14] can be easily extended to the case $q \neq 2$. We leave the verification to the interested reader. We also note that the regularized flow (20)-(22) plays an important role not only for establishing existence result for the gradient flow (11)-(14) but also provide a practical and convenient formulation for computing the solution of the gradient flow. Indeed, in the next section we shall develop some fully discrete finite element method for computing the solution of the gradient flow (11)-(14) via the regularized flow (20)-(22), and shall also present a few numerical experiment results in §4.

We conclude this section by giving definitions of solutions to problem (20)-(22) and stating two well-posedness results for the problem without giving proofs. We refer the reader to [5, 14] for their proofs in some special cases.

Definition 2.1. *A map $\mathbf{u}^{\varepsilon,\delta} : \Omega_T \to \mathbf{R}^n$ is called a global* weak *solution to (20)-(22) if*

(i) $\mathbf{u}^{\varepsilon,\delta} \in L^\infty((0,T); H^1(\Omega, \mathbf{R}^n)) \cap H^1((0,T); L^2(\Omega, \mathbf{R}^n))$,
(ii) $|\mathbf{u}^{\varepsilon,\delta}| \leq 1$ *a.e. in* Ω_T,
(iii) $\mathbf{u}^{\varepsilon,\delta}$ *satisfies (20)-(22) in the distributional sense.*

Definition 2.2. *A weak solution $\mathbf{u}^{\varepsilon,\delta}$ to (20)-(22) is called a* strong *solution if $\mathbf{u}^{\varepsilon,\delta} \in L^2((0,T);\ H^2(\Omega, \mathbf{R}^n))$. It is called a* regular *solution if in addition $\mathbf{u}^{\varepsilon,\delta} \in H^1((0,T);\ H^1(\Omega, \mathbf{R}^n))$. It is called a* classical *solution if $\mathbf{u}^{\varepsilon,\delta} \in C^{2,1}(\overline{\Omega}_T)$ and satisfies (20)-(22) pointwise.*

Theorem 2.1. *Let $\Omega \subset \mathbf{R}^m$ be a bounded domain with smooth boundary. Suppose that $\mathbf{u}_0$ and $\mathbf{g}$ are sufficiently smooth functions (say, $\mathbf{u}_0, \mathbf{g} \in [C^3(\overline{\Omega})]^n$) with $|\mathbf{u}_0| \leq 1$ and $|\mathbf{g}| \leq 1$ in Ω. Then, the regularized flow (20)-(22) has a unique global classical solution $\mathbf{u}^{\varepsilon,\delta}$ for each fixed pair of positive numbers (ε, δ). Moreover, $\mathbf{u}^{\varepsilon,\delta}$ satisfies the following energy law*

$$\mathcal{J}^{\varepsilon,\delta}_{\beta,\lambda}(\mathbf{u}^{\varepsilon,\delta}(s)) + \int_0^s \|\, \mathbf{u}^{\varepsilon,\delta}_t(t) \,\|^2_{L^2}\, dt = \mathcal{J}^{\varepsilon,\delta}_{\beta,\lambda}(\mathbf{u}_0) \qquad \forall s \in [0,T]. \tag{23}$$

For less regular datum functions $\mathbf{u}_0$ and $\mathbf{g}$, and less regular domain Ω, we have the following weaker existence result.

Theorem 2.2. *Let $\Omega \subset \mathbf{R}^m$ be a bounded Lipschitz domain, suppose that $\mathbf{u}_0 \in H^1(\Omega, \mathbf{R}^n)$, $|\mathbf{u}_0| \leq 1$ and $|\mathbf{g}| \leq 1$ a.e. in Ω. Then, for each fixed pair of positive numbers (ε, δ) the regularized flow (20)-(22) has a unique global weak solution $\mathbf{u}^{\varepsilon,\delta}$. Moreover, $\mathbf{u}^{\varepsilon,\delta}$ satisfies the following energy inequality*

$$\mathcal{J}^{\varepsilon,\delta}_{\beta,\lambda}(\mathbf{u}^{\varepsilon,\delta}(s)) + \int_0^s \|\, \mathbf{u}^{\varepsilon,\delta}_t(t) \,\|^2_{L^2}\, dt \leq \mathcal{J}^{\varepsilon,\delta}_{\beta,\lambda}(\mathbf{u}_0) \qquad \forall s \in [0,T]. \tag{24}$$

3. Finite element approximations

3.1. Formulation of fully discrete finite methods. We assume $m = 2$ and $n = 3$ in this section. Let $\mathcal{T}_h = \{K_1, \ldots, K_M\}$ be a quasi-uniform triangulation of Ω with mesh size $h \in (0,1)$ and $\overline{\Omega} = \bigcup_{K \in \mathcal{T}_h} \overline{K}$. Let $J_\tau := \{t_k\}_{k=0}^L$ be a quasi-uniform partition of $[0,T]$ with mesh size $\tau := \frac{T}{L}$, and

$$\partial_t \mathbf{v}^k := \frac{\mathbf{v}^k - \mathbf{v}^{k-1}}{\tau}.$$

We introduce the finite element space

$$\mathbf{V}_h := \{\mathbf{v} \in C^0(\overline{\Omega}, \mathbf{R}^n);\ \mathbf{v}_h|_K \in [P_r(K)]^n,\ \forall K \in \mathcal{T}_h\},$$

where $P_r(K)$ denotes the space of polynomials of degree less than or equal to r on K. We decompose the density function F, which is not convex, into the difference of two convex functions W_+ and W_-, that is,

$$F(\mathbf{v}) = W_+(\mathbf{v}) - W_-(\mathbf{v}).$$

One such example is $W_+ = \dfrac{|\mathbf{v}|^4}{4}$ and $W_- = \dfrac{|\mathbf{v}|^2}{2} - \dfrac{1}{4}$. Clearly, such a decomposition is not unique.

Our fully discrete finite element discretization for the initial boundary value problem (20)-(22) is defined as follows: Find $\mathbf{u}_h^k \in \mathbf{V}^h$ for $k = 1, \cdots, L$ such that

$$\begin{aligned}(\partial_t \mathbf{u}_h^k, \mathbf{v}_h) + \beta(\mathcal{B}_h^k, \nabla \mathbf{v}_h) + \lambda((\mathbf{u}_h^k - \mathbf{g})|\mathbf{u}_h^k - \mathbf{g}|_\varepsilon^{q-2}, \mathbf{v}_h)& \\ + \frac{1}{\delta}(\mathbf{W}'_+(\mathbf{u}_h^k), \mathbf{v}_h) = \frac{1}{\delta}(\mathbf{W}'_-(\mathbf{u}_h^{k-1}), \mathbf{v}_h) &\qquad \forall\, \mathbf{v}_h \in \mathbf{V}_h,\end{aligned} \tag{25}$$

where

$$\mathcal{B}_h^k = \frac{\varphi'(|\nabla \mathbf{u}_h^k|_\varepsilon)}{|\nabla \mathbf{u}_h^k|_\varepsilon} \nabla \mathbf{u}_h^k$$

with some starting value $\mathbf{u}_h^0 \in \mathbf{V}_h$. Note that we have omitted the indices ε and δ on $\mathbf{u}_h^k$ for notation brevity and clearly the above method is implicit in time.

3.2. Stability and convergence analysis. Since at each time step our finite element method (25) gives a nonlinear equation in $\mathbf{u}_h^k$, the well-posedness of the equation is not obvious and has to be addressed first. After that is done, we need to analyze the proposed finite element method by addressing the stability and convergence of the method, which we now describe.

First, the well-posedness and the stability of the method is ensured by the following theorem.

Theorem 3.1. *For each fixed $k \geq 1$, suppose that $\mathbf{u}_h^{k-1}$ is known, then there exists a unique solution $\mathbf{u}_h^k$ to (25). Moreover, $\{\mathbf{u}_h^k\}$ satisfies the following energy estimate*

$$\frac{\tau}{2}\sum_{k=1}^{\ell} \| \partial_t \mathbf{u}_h^k \|_{L^2}^2 + \mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{u}_h^k) \leq \mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{u}_h^0). \tag{26}$$

Here $\mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}$ is defined by (16).

Proof. For each fixed $k \geq 1$, we define the functional

$$\begin{aligned}G_k(\mathbf{v}) := \int_\Omega \Big\{ \frac{1}{2\tau}|\mathbf{v} - \mathbf{u}_h^{k-1}|^2 + \frac{\beta\varepsilon^\alpha}{2}|\nabla \mathbf{v}|^2 + \beta\varphi\big(|\nabla \mathbf{v}|_\varepsilon\big)& \\ + \frac{\lambda}{q}|\mathbf{v} - \mathbf{g}|_\varepsilon^q + \frac{\beta}{\delta}W_+(\mathbf{v}) \Big\}\, dx - \frac{\beta}{\delta}\int_\Omega W'_-(\mathbf{u}_h^{k-1}) \cdot \mathbf{v}\, dx.&\end{aligned} \tag{27}$$

It is easy to check that G_k is a convex, coercive and differentiable functional, then it has a unique minimizer $\mathbf{u}_h^k \in \mathbf{V}_r^h$ (cf. [25]). A direct calculation shows that (25) is exactly the weak form of the Euler-Lagrange equation of $G_k(\mathbf{v})$. Hence, (25) has a unique solution $\mathbf{u}_h^k$ which is the unique minimizer of $G_k(\mathbf{v})$.

To derive the desired energy inequality, on noting that $\mathbf{u}_h^k$ is the minimizer of G_k in $\mathbf{V}_r^h$, we have

$$G_k(\mathbf{u}_h^k) \le G_k(\mathbf{u}_h^{k-1}). \tag{28}$$

The convexity of W_- implies that

$$W'_-(\mathbf{u}_h^{k-1})\big(\mathbf{u}_h^k - \mathbf{u}_h^{k-1}\big) \le W_-(\mathbf{u}_h^k) - W_-(\mathbf{u}_h^{k-1}).$$

Combining the above inequality with (28) we get

$$\frac{\tau}{2}\| \partial_t \mathbf{u}_h^k \|_{L^2}^2 + \mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{u}_h^k) - \mathcal{J}_{\beta,\lambda}^{\varepsilon,\delta}(\mathbf{u}_h^{k-1}) \le 0.$$

The estimate (26) then follows from applying the summation operator $\sum_{k=1}^{\ell}$ $(1 \le \ell \le L)$ to the last inequality. The proof is complete. □

Corollary 3.1. *The finite element method* (25) *is stable for any $h, \tau > 0$, hence, it is absolutely stable.*

Next, we analyze the convergence of the method (25). Two approaches are often used to address the convergence of a numerical method. The first approach is to derive error estimates for the numerical solution, which then infer not only the convergence but also rates of convergence. The second approach directly proves convergence of the numerical solution in some (function) norm. Clearly, an error estimate is more desirable because it is a stronger result. On the other hand, error estimates are only possible if the exact (PDE) solution has certain regularity. In the case when the exact solution has very low regularities, convergence is usually the best one could hope and ask for.

It was pointed out in Section 2 that the solutions of the limiting problem (11)-(14) are expected to belong to $L^2((0,T);(BV(\Omega))^d)$. Recall that spatial gradient $\nabla\mathbf{v}$ of a BV function $\mathbf{v}$ is only a Radon measure (cf. [1]), that means $\nabla\mathbf{v}$ is not even an integrable function. Because this very low regularity of solutions to problem (11)-(14), it is natural for us to adopt the second approach described above to address the convergence of the method (25). To this end, we define the linear interpolation in t of the finite element solution $\{\mathbf{u}_h^k\}$

$$\mathbf{U}^{\varepsilon,\delta,h,\tau}(\cdot,t) := \frac{t - t_{k-1}}{\tau}\mathbf{u}_h^k(\cdot) + \frac{t_k - t}{\tau}\mathbf{u}_h^{k-1}(\cdot) \tag{29}$$

for any $t \in [t_{k-1}, t_k]$ and $1 \le k \le L$. Clearly, $\mathbf{U}^{\varepsilon,\delta,h,\tau}$ is continuous in both x and t.

We now are ready to state our convergence theorem for the finite element method (25).

Theorem 3.2. *Suppose that $\mathbf{u}_0 \in H^1(\Omega, \mathbf{R}^n)$, $|\mathbf{u}_0| = 1$ and $|\mathbf{g}| \le 1$ a.e. in Ω. For each pair of positive numbers (ε,δ), let $\mathbf{U}^{\varepsilon,\delta,h,\tau}$ be defined by* (29). *Then, there exists $\mathbf{u}^{\varepsilon,\delta} \in L^\infty(\Omega_T)$ such that*

$$\lim_{h,\tau\to 0} \| \mathbf{U}^{\varepsilon,\delta,h,\tau} - \mathbf{u}^{\varepsilon,\delta} \|_{L^\ell(\Omega_T)} = 0 \qquad \forall \ell \in [1,\infty), \tag{30}$$

provided that

$$\lim_{h\to 0} \| \mathbf{u}_h^0 - \mathbf{u}^0 \|_{H^1} = 0.$$

Moreover, $\mathbf{u}^{\varepsilon,\delta}$ solves (20)-(22) in the sense of Definition 2.1.

Proof. Since the proof is similar to that of Theorem 1.5 of [12], where the convergence of a general Galerkin approximation of the p-harmonic map heat flow for $p \geq 2$ was proved, and to that of Theorem 7.2 of [5], where the convergence of a fully discrete finite element method for the p-harmonic map heat flow with $1 \leq p < \infty$ was presented, we shall only outline the main idea and key steps of the proof in the following.

The main idea of the proof is to use the method of compactness. The proof can be divided into three steps.

Step 1: Energy estimates. The goal of this step is to derive uniform (in h, τ, and possibly in ε) energy estimates for the finite element solution $\{\mathbf{U}^{\varepsilon,\delta,h,\tau}\}$ (or $\{\mathbf{u}_h^k\}$). On noting that $\partial_t \mathbf{u}_h^k = \mathbf{U}_t^{\varepsilon,\delta,h,\tau}$ and $\tau \sum_{k=1}^L \| \partial_t \mathbf{u}_h^k \|_{L^2}^2 = \| \mathbf{U}_t^{\varepsilon,\delta,h,\tau} \|_{L^2(L^2)}$, the discrete energy estimate (26) then immediately infer the uniform (in h and τ) boundedness of $\{\mathbf{U}^{\varepsilon,\delta,h,\tau}\}$ in $L^\infty((0,T); H^1(\Omega, \mathbf{R}^n))$. We note that this bound is not uniform in ε. On the ohter hand, uniform in ε estimate can be proved in $L^\infty((0,T); W^{1,1}(\Omega, \mathbf{R}^n))$.

Step 2: Extracting convergent subsequence. The uniform energy estimates obtained in Step 1 allow us to extract a convergent subsequence of $\{\mathbf{U}^{\varepsilon,\delta,h,\tau}\}$ which converges weakly in $L^\ell((0,T); H^1(\Omega))$ for $1 < \ell < \infty$ and strongly in $L^\ell(\Omega_T)$ (by Sobolev embedding) to some function $\mathbf{u}^{\varepsilon,\delta} \in L^\infty(\Omega_T)$ as $h, \tau \to 0$, which shows (30) holds for the subsequence.

Step 3: Passing to the limit as $h, \tau \to 0$. The final step of the proof is to pass to the limit in the finite element method (25). Clearly, this can be done easily in the first term on the left-hand side because it is a linear term. Passing to the limit in all low order nonlinear terms (i.e., the third and fourth term on the left-hand side and the term on the right-hand side) is not hard either because the strong convergence of the subsequence in L^ℓ-norm for all $\ell \in [1, \infty)$. So the only term remains is the second term on the left-hand side which has strong nonlinearity. To pass to the limit in this term, we need to appeal to a compactness result given by Lemma 13 of [14]. Using this result we can pass to the limit in this nonlinear term and show that the corresponding subsequence of $\mathcal{B}_h^k$ converges to $\mathcal{B}^{\varepsilon,\delta}$ weakly in L^2. Consequently, we prove that $\mathbf{u}^{\varepsilon,\delta}$ is a weak solution of (20)-(22).

Finally, we note that the uniqueness of $\mathbf{u}^{\varepsilon,\delta}$ (cf. Theorem 2.2) ensures that (30) in fact holds for the whole sequence $\{\mathbf{U}^{\varepsilon,\delta,h,\tau}\}$. $\square$

Remark 3.1. *Several practical choices of $\mathbf{u}_h^0$ are possible. For instance, both the H^1-projection of $\mathbf{u}_0$ and the Clemént finite element interpolation of $\mathbf{u}_0$ into $\mathbf{V}_r^h$ (cf.* [13]*) are qualified candidates for $\mathbf{u}_h^0$.*

An immediate consequence of Theorem 3.2 and the convergence of $\mathbf{u}^{\varepsilon,\delta}$ (see the discussion in the paragraph before Definition 2.1) is the following convergence result.

Theorem 3.3. *Let $\mathbf{U}^{\varepsilon,\delta,h,\tau}$ be defined by* (29)*, assume the assumptions of Theorems 3.2 hold. Then, there exists a subsequence of $\{\mathbf{U}^{\varepsilon,\delta,h,\tau}\}$ (still denote by the same notation) and a weak solution $\mathbf{u}$ of (11)-(14) such that*

$$\lim_{\varepsilon,\delta\to 0,} \lim_{h,\tau\to 0,} \| \mathbf{u} - \mathbf{U}^{\varepsilon,\delta,h,\tau} \|_{L^\ell(\Omega_T)} = 0 \qquad \forall \ell \in [1, \infty). \tag{31}$$

4. Numerical Experiments

In this section we present some numerical experiment results for the finite element method proposed in Section 3. These numerical results not only demonstrate the efficiency of the proposed numerical method but also show the effectiveness of the proposed chromaticity denoising models. Since different density function φ and parameter q can be used in (5), here we consider the following four specific models:

- Model 1: $\varphi(s) = s$ and $q = 2$.
- Model 2: $\varphi(s) = s$ and $q = 1$.
- Model 3: $\varphi(s) = s^{\frac{3}{2}}$ and $q = 2$.
- Model 4: $\varphi(s) = \sqrt{s^2 + 1}$ and $q = 1$.

Note that Model 1 is the 1-harmonic map model and Model 3 is the $\frac{3}{2}$-harmonic map model both considered in [5]. It should be noted that although we use CBD approach, we only consider the chromaticity denoising in our numerical experiments. That is, we assume no noise in the brightness component. As explained in Section 1, if there is a noise in the brightness component, one can use any gray image denoising model, such as the Total Variation (TV) model [20], to denoise the brightness component.

Figure 1 displays the test results of Models 1 and 2 on a knee image. The left image is a noisy knee image with 20% Gaussian noise in the chromaticity, the middle image is the denoised image using Model 2 and the right image is the denoised image using Model 1. The SNR (signal to noise ratio) of the three images are $7.4499, 16.5151, 6.9723$, respectively.

Figure 2 displays the test results of Models 3 and 4 on the knee image. The left image is a noisy knee image with 20% Gaussian noise in the chromaticity, the middle image is the denoised image using Model 3 and the right image is the denoised image using Model 4. The SNR of the three images are $7.4499, 15.8193, 6.9245$, respectively.

FIGURE 1. Test results of Models 1 and 2 on a knee image

FIGURE 2. Test results of Models 3 and 4 on a knee image

Figure 3 displays the test result of Models 1 and 2 on a pepper image. The left image is a noisy pepper image with 20% Gaussian noise in the chromaticity, the middle image is the denoised image using Model 2 and the right image is the denoised image using Model 1. The SNR of the three images are $2.4936, 2.7659, 2.7659$, respectively.

FIGURE 3. Test results of Models 1 and 2 on a pepper image

Figure 4 displays the test result of Models 3 and 4 on the pepper image. The left image is a noisy pepper image with 20% Gaussian noise in the chromaticity, the middle image is the denoised image using Model 3 and the right image is the denoised image using Model 4. The SNR of the three images are $2.4936, 2.7495, 2.7641$, respectively.

FIGURE 4. Test results of Models 3 and 4 on a pepper image

As a comparison, we also present a simulation result of the channel-by-channel approach using the TV model to denoise each chromaticity component in Figure 5, where the left image is the same noisy pepper image as above and the right image is the denoised image. The result clearly shows that the channel-by-channel approach does a worse job than the brightness and chromaticity decomposition approach. However, it should be noted that the channel-by-channel approach takes less computer time to run the test.

FIGURE 5. Test results of a channel-by-channel model on a pepper image

5. Summary and concluding remarks

In this paper we present a class of variational models with linear growth energy functionals for color image denoising based on the chromaticity and brightness decomposition approach. To compute the solution of the variational models, we consider the L^2-gradient flow of regularized energy functionals and the nonconvex constraint $\mathbf{u} = 1$, which is hard to handle numerically, is enforced approximately using the Ginzburg-Landau approximation. The gradient flows are discretized by a fully discrete implicit finite element method. It is proved that the proposed finite element method is absolutely stable and enjoys a discrete energy law. Strong convergence in L^ℓ-norm for $\ell \in [1, \infty)$ is established for the numerical solution using the method of compactness. Numerical experiments are presented to show good performance of the proposed finite element scheme and the color image denoising models.

Our numerical experiment results indicate that all variational models of linear growth functionals are effective for chromaticity denoising (or directional diffusion), although the minimal surface model is easier to implement and faster to converge compared to the popular 1-harmonic map model. In particular, all models of this family do a very good job on preserving image edges. Mathematically, this is expected since the solutions of these models are only BV functions instead of Sobolev functions (which is the case for the p-harmonic map model for $p > 1$) (see [5, 14]). It is well-known that BV functions allow jumps while Sobolev functions do not. It would be interesting to investigate that if the results of this paper can be extended to nonconvex functionals.

References

[1] L. Ambrosio, N. Fusco, D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford University Press, New York, 2000.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows*. Birkhäuser, Basel, 2005.

[3] L. Alvarez, P. L. Lion, and J. M. Morel. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29:845, 1992.

[4] G. Aubert and P. Kornprobst. *Mathematical problems in image processing*, volume 147 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2006.

[5] J. W. Barrett, X. Feng, and A. Prohl. p-harmonic map heat flow for $1 \leq p < \infty$ and their finite element approximations with applications to image processing. *SIAM J. Math. Anal.* (to appear).

[6] M. Bertalmío, V. Caselles, G. Haro, and G. Sapiro. PDE-based image and surface inpainting. In *Handbook of mathematical models in computer vision*, pages 33–61. Springer, New York, 2006.

[7] F. Bethuel, H. Brezis, and F. Helein. *Ginzburg-Landau Vertices*. Birkhauser, New York, 1994.

[8] P. V. Blomgren and T. F. Chan. Color tv: Total variation methods for restoration of vector valued images. *IEEE Trans. Image Processing*, 7:304, 1998.

[9] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. 3rd edition, Springer, 2008.

[10] T. F. Chan, S. H. Kang, and J. Shen. Total variation denoising and enhancement of color images based on the cb and hsv color models. *Journal of Visual Communucation and Image Representation*, 12:422, 2001.

[11] T. F. Chan and J. Shen. Variational restoration of non-flat image features: Models and algorithms. *SIAM Journal of Applied Mathematics*, 61:1338, 2000.

[12] Y. Chen, M.-H. Hong, N. Hungerbühler. Heat flow of p-harmonic maps with values into spheres. *Math. Z.*, 215:25–35 (1994).

[13] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.

[14] X. Feng. Divergence-L^q and divergence-measure tensor fields and gradient flows for linear growth functionals of maps into the unit sphere. *Calc. Var. and PDEs*, (submitted).

[15] X. Feng and A. Prohl. Analysis of total variation flow and its finite element approximations. *Math. Mod. Num. Anal.*, 37:533, 2003.

[16] X. Feng, M. von Oehsen, and A. Prohl. Rate of convergence of regularization procedures and finite element approximations for the total variation flow. *Numer. Math.*, 100:441, 2005.

[17] Y. Giga, Y. Kashima, N. Yamazaki. Local solvability of a constrained gradient system of total variation, *Abstr. Appl. Anal.*, 8:651-682 (2004).
[18] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* Addison-Wesley, 1992.
[19] J.-M. Morel and S. Solimini. *Variational methods in image segmentation.* Progress in Nonlinear Differential Equations and their Applications, vol 14, Birkhäuser Boston Inc., Boston, 1995.
[20] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D.*, 60:259, 1992.
[21] G. Sapiro. *Geometric partial differential equations and image analysis.* Cambridge University Press, Cambridge, 2001.
[22] G. Sapiro and D. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Trans. Image Processing*, 5:1582, 1996.
[23] W. E. Snyder and H. Qi. *Machine Vision.* Cambridge University Press, Cambridge, 2004.
[24] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Trans. Image Processing*, 7:310, 1998.
[25] M. Struwe. *Variational Methods.* Springer, New York, 1990.
[26] B. Tang, G. Sapiro, and V. Caselles. Diffusion of general data on non-flat manifolds via harmonic maps theory: the direction diffusion case. *Int. J. Comp. Vision*, 36:149, 2000.
[27] B. Tang, G. Sapiro, and V. Caselles. Color image enhancement via chromaticity diffusion. *IEEE Trans. Image Processing*, 10:701, 2001.
[28] P. E. Trahanias, D. Karako, and A. N. Venetsanopoulos. Directional processing of color images: theory and experimental results. *IEEE Trans. Image Processing*, 5:868, 1996.
[29] P. E. Trahanias and A. N. Venetsanopoulos. Vector directional filters-a new class of multichannel image processing filters. *IEEE Trans. Image Processing*, 2:528, 1993.
[30] L. Vese and S. Osher. Numerical methods for p-harmonic flows and applications to image processing. *SIAM J. Numer. Anal.*, 40:2085, 2002.

Department of Mathematics, the University of Tennessee, Knoxville, TN 37996, U.S.A.
E-mail: `xfeng@math.utk.edu and yoon@math.utk.edu`
URL: `http://www.math.utk.edu/~xfeng/ and http://www.math.utk.edu/~yoon/`