An unconstrained multiphase thresholding approach for image segmentation

Benjamin Berkels

Institut für Numerische Simulation,
Rheinische Friedrich-Wilhelms-Universität Bonn,
Nussallee 15, 53115 Bonn, Germany
`benjamin.berkels@ins.uni-bonn.de`,
WWW home page: `http://numod.ins.uni-bonn.de/`

Abstract. In this paper we provide a method to find global minimizers of certain non-convex 2-phase image segmentation problems. This is achieved by formulating a convex minimization problem whose minimizers are also minimizers of the initial non-convex segmentation problem, similar to the approach proposed by Nikolova, Esedoḡlu and Chan. The key difference to the latter model is that the new model does not involve any constraint in the convex formulation that needs to be respected when minimizing the convex functional, neither explicitly nor by an artificial penalty term. This approach is related to recent results by Chambolle. Eliminating the constraint considerably simplifies the computational difficulties, and even a straightforward gradient descent scheme leads to a reliable computation of the global minimizer. Furthermore, the model is extended to multiphase segmentation along the lines of Vese and Chan. Numerical results of the model applied to the classical piecewise constant Mumford-Shah functional for two, four and eight phase segmentation are shown.

1 Introduction

Image segmentation is one of the fundamental research topics in the field of image processing. In particular, the Mumford-Shah model [1] is widely used in this context. One of the difficulties of this and many other variational image processing models is that the underlying energy functional has local, non-global minima. This is not only a theoretical problem, since the commonly used numerical minimization techniques often get stuck in local minima that differ considerably from a global minimum, hence possibly producing useless results. The goal of this paper is to introduce a method to obtain a global minimizer of the Mumford-Shah functional for 2-phase segmentation that only involves solving an unconstrained convex minimization problem. This method can be extended to multiphase segmentation by the ideas of Vese and Chan [2] in a canonical way.

1.1 Related work

The problem of minimizing the Mumford-Shah segmentation functional has been extensively studied in the last decade leading to a wide range of existing methods, each with its own shortcomings. One of the first numerical feasible methods to obtain (local) minimizers of the functional was proposed by Chan and Vese [3]. They build on the levelset methods of Osher and Sethian [4] and parameterize the unknown set by a levelset function.

Shen [5] developed a Γ-convergence formulation along with a simple implementation by the iterated integration of a linear Poisson equation. The unknown set is represented in a diffuse way by a phase field.

In [6], Esedoḡlu and Tsai tackle the minimization problem based on the threshold dynamics of Merriman, Bence and Osher [7] for evolving an interface by its mean curvature. Here the minimization is achieved by alternating the solution of a linear parabolic partial differential equation and simple thresholding.

Alvino and Yezzi [8] approximate Mumford-Shah segmentation using reduced image bases. According to them, the majority of the robustness of Mumford-Shah segmentation can be obtained without allowing each pixel to vary independently. Their approximative model has comparable performance to Mumford-Shah segmentations where each pixel is allowed to vary freely.

A way to obtain global minimizers was introduced by Nikolova, Esedoḡlu and Chan [9]. Here, a convex constrained minimization problem has to be solved followed by a simple thresholding of the latter minimizer. This method is closely related to the method we propose in this paper, the key difference is that [9] requires a constraint in the convex minimization while the model proposed in this paper does not involve any constraint in the convex formulation.

On the other hand there are methods to solve a certain class of minimal surface problems by unconstrained convex optimization, cf. the work of Chambolle and Darbon [10, 11]. The 2-phase Mumford-Shah functional belongs to this class, yet due to the best of our knowledge nobody seems to have tapped the potential offered by these general insights for Mumford-Shah based image segmentation so far.

2 Constrained global 2-phase minimization

First let us describe the general framework and revise the work of Nikolova et al. [9], the starting point for our model.

In the following, Ω denotes our computational domain, an arbitrary but fixed subset of $\mathbb{R}^n$. For given indicator functions $f_1, f_2 \in L^1(\Omega)$ such that $f_1, f_2 \geq 0$ a.e. we consider the prototype Mumford-Shah energy

$$E_{\mathrm{MS}}[\Sigma] := \int_\Sigma f_1 \mathrm{d}x + \int_{\Omega\setminus\Sigma} f_2 \mathrm{d}x + \nu \operatorname{Per}(\Sigma), \tag{1}$$

where $\operatorname{Per}(\Sigma)$ denotes the perimeter of the set $\Sigma \subset \Omega$ in Ω. If u_0 is an image, $c_1, c_2 \in \mathbb{R}$ are two grey values and $f_i(x) := (u_0(x) - c_i)^2$, this is the well known piecewise constant Mumford-Shah functional for 2-phase segmentation, i.e.

$$E[\Sigma, c_1, c_2] = \int_\Sigma (u_0 - c_1)^2 \mathrm{d}x + \int_{\Omega\setminus\Sigma} (u_0 - c_2)^2 \mathrm{d}x + \nu \operatorname{Per}(\Sigma). \tag{2}$$

Remark 1. Because of

$$E_{\mathrm{MS}}[\Sigma] = \underbrace{\int_\Sigma (f_1 - f_2)\mathrm{d}x + \nu \operatorname{Per}(\Sigma)}_{=:\hat{E}_{\mathrm{MS}}[\Sigma]} + \int_\Omega f_2 \mathrm{d}x,$$

E_{MS} and $\hat{E}_{\mathrm{MS}}$ share the same minimizers.

Remark 2. For $h(x) := e^{-|x|^2}$, we have

$$E_{\mathrm{MS}}[\Sigma] = \int_\Sigma (f_1 + h)\,\mathrm{d}x + \int_{\Omega\setminus\Sigma} (f_2 + h)\,\mathrm{d}x + \nu \operatorname{Per}(\Sigma) - \underbrace{\int_\Omega h \mathrm{d}x}_{=C<\infty},$$

i.e. replacing f_1 and f_2 by $f_1 + h$ and $f_2 + h$ does not affect the minimizers of E_{MS}. This, combined with $f_1, f_2 \geq 0$ a.e., means that we can assume $f_1, f_2 > 0$ a.e. in Ω without loss of generality.

To obtain (local) minimizers of the functional above, Chan and Vese [3] proposed to parametrize the unknown set Σ by a levelset function ϕ and get the energy

$$E_{\mathrm{CV}}[\phi] := \int_\Omega H(\phi) f_1 + (1 - H(\phi)) f_2 + \nu |\nabla(H(\phi))| \mathrm{d}x.$$

Here, $H(\cdot)$ denotes the Heaviside function, i.e. $H(s) = 1$ for $s > 0$ and $H(s) = 0$ else. A gradient descent will be used for minimization, therefore H is replaced by a smeared out Heaviside function, e.g. $H_\delta(x) := \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x}{\delta}\right)$, where $\delta > 0$. While the specific choice is not important, it is important to use a function whose derivative does not have compact support (cf. [3]). This gives the regularized energy

$$E_{\mathrm{CV},\delta}[\phi] := \int_\Omega H_\delta(\phi) f_1 + (1 - H_\delta(\phi))\, f_2 + \nu\, |\nabla(H_\delta(\phi))|\, \mathrm{d}x \tag{3}$$

and yields the gradient descent

$$\partial_t \phi = H_\delta'(\phi) \left[(f_2 - f_1) + \nu \operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|} \right) \right]. \tag{4}$$

One of the major drawbacks of the energy (3) is its non-convexity in ϕ. In [9], Nikolova et al. noted that the gradient descent (4) and

$$\partial_t \phi = \left[(f_2 - f_1) + \nu \operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|} \right) \right]$$

have the same stationary points, because $H_\delta'(\phi) > 0$. Obviously the latter is the gradient descent of the energy

$$E_{\mathrm{CE}}[\phi] := \int_\Omega (f_1 - f_2)\phi + \nu |\nabla\phi| \mathrm{d}x.$$

In general, $f_1 - f_2$ takes positive and negative values, therefore the energy is not bounded (neither from below nor from above). In other words, it does not necessarily have a minimizer. However, this is easily fixed by restricting the minimization to $0 \leq \phi(x) \leq 1$ for all $x \in \Omega$. Based on this, the following theorem holds:

Theorem 1. *For given indicator functions $f_1, f_2 \in L^1(\Omega)$ such that $f_1, f_2 \geq 0$ a.e., let*

$$u := \operatorname*{argmin}_{0 \leq \tilde{u} \leq 1} \int_\Omega (f_1 - f_2)\tilde{u} + \nu |\nabla \tilde{u}| \mathrm{d}x = \operatorname*{argmin}_{0 \leq \tilde{u} \leq 1} E_{\mathrm{CE}}[\tilde{u}]$$

and $\Sigma_c := \{x \in \Omega | u(x) > c\}$. Then Σ_c is a minimizer of the Mumford-Shah energy (1) for all $c \in [0, 1)$.

Proof. Nikolova et al. proved this theorem in [9] for a.e. $c \in [0, 1]$, we extend it here to hold not only for almost every, but for every $c \in [0, 1)$. First, we briefly sketch the prove given by Nikolova et al. for a.e. $c \in [0, 1]$.

Using $0 \leq u \leq 1$ and the coarea formula, one can show

$$E_{\mathrm{CE}}[u] = \int_0^1 E_{\mathrm{MS}}[\Sigma_c] \mathrm{d}c - C,$$

where C is a constant independent of u. Let $\Sigma_* \subset \Omega$ be a minimizer of E_{MS} (the existence of such minimizers using convergence in measure follows from standard arguments) and let $M := \{c \in [0, 1] | E_{\mathrm{MS}}[\Sigma_c] > E_{\mathrm{MS}}[\Sigma_*]\}$. Assuming $\mu(M) > 0$ leads to the contradiction $E_{\mathrm{CE}}[\chi_{\Sigma_*}] < E_{\mathrm{CE}}[u]$ therefore $\mu(M) = 0$ holds and the statement is proven for a.e. $c \in [0, 1]$. Here, χ_A denotes the characteristic function of the set A.

Now we extend the statement to all $c \in [0, 1)$, inspired by the proof of Lemma 4 (iii) in [12]: Again let u be a minimizer of E_{CE} under the constraint $0 \leq u \leq 1$ and denote its superlevelsets by Σ_c. Choose an arbitrary but fixed $\hat{c} \in [0, 1)$. The statement holds for a.e. $c \in [0, 1]$, so by Remark 1, there exists a sequence $(c_n) \in [0, 1]^{\mathbb{N}}$ with $c_n \downarrow \hat{c}$ such that

$$\Sigma_{c_n} \in \operatorname*{argmin}_{\Sigma \subset \Omega} \hat{E}_{\mathrm{MS}}[\Sigma].$$

Since the superlevelsets of a function are contained in each other, we have $\chi_{\Sigma_{c_n}} = \chi_{\bigcup_{k=1}^n \Sigma_{c_k}} \to \chi_{\Sigma^\cup}$ pointwise a.e., where $\Sigma^\cup := \bigcup_{n=1}^\infty \Sigma_{c_n}$. Setting $g := f_1 - f_2$ and using Lebesgue's dominated convergence theorem, we obtain

$$\int_{\Sigma^\cup} g \mathrm{d}x = \int_\Omega g \chi_{\Sigma^\cup} \mathrm{d}x = \lim_{n \to \infty} \int_\Omega g \chi_{\Sigma_{c_n}} \mathrm{d}x = \lim_{n \to \infty} \int_{\Sigma_{c_n}} g \mathrm{d}x.$$

Here we used $\left| g \chi_{\Sigma_{c_n}} \right| \leq |g| \leq |f_1| + |f_2|$ to provide the integrable upper bound. For each n and $\Sigma \subset \Omega$, we have

$$\int_{\Sigma_{c_n}} g \mathrm{d}x + \nu \operatorname{Per}(\Sigma_{c_n}) \leq \int_\Sigma g \mathrm{d}x + \nu \operatorname{Per}(\Sigma).$$

Using the continuity argument from above and the lower semicontinuity of the perimiter (cf. [13]), we get

$$\int_{\Sigma^{\cup}} g\mathrm{d}x + \nu\,\mathrm{Per}(\Sigma^{\cup}) \leq \int_{\Sigma} g\mathrm{d}x + \nu\,\mathrm{Per}(\Sigma),$$

i.e. $\Sigma^{\cup}$ is a minimizer of $E_{\mathrm{MS}}[\cdot, c_n]$. Combining this with

$$\Sigma_c = \{x \in \Omega | u(x) > c\} = \bigcup_{n=1}^{\infty} \{x \in \Omega | u(x) > c_n\} = \bigcup_{n=1}^{\infty} \Sigma_{c_n}$$

concludes the proof. □

Knowing that Theorem 1 holds true for all $c \in [0, 1)$ also remedies the last bit of "uncertainty" left in [9].

Remark 3. For any function u that fulfills the constraint, obviously $\{u > 1\} = \emptyset$. Therefore we cannot expect Theorem 1 to hold for $c = 1$.

To solve the constrained optimization problem, Nikolova et al. show that the constrained problem has the same minimizers as the unconstrained problem if a penalty term of the form $\alpha \int p(u(x))$ is added with a sufficiently large coefficient α (cf. [9], Claim 1). Here p denotes $p(s) = \max\{0, 2\left|s - \frac{1}{2}\right| - 1\}$.

While this result already gives a method to find global minimizers of E_{MS} by solving a convex, unconstrained minimization problem, its practical relevance is limited. Most numerical minimizations methods rely on the gradient of the functional, but the proposed penalty term is not differentiable, making a regularization necessary. But any smooth regularization of the penalty term will stop the minimizers of the convex, constrained functional to coincide with those of the convex functional with penalty term. The stronger the regularization, the more the minimizers deviate.

Furthermore, the regularization imposes numerical difficulties. If an explicit gradient descent is used for the minimization (as proposed in [9]), a suitable timestep size control is needed to ensure convergence. The step sizes allowed by such methods, e.g. the Armijo rule [14], typically correspond to the size of the region in which the linearization of the functional properly approximates the functional. Due to the nature of the penalty term p, the linearization at 0 and 1 of a regularized version of it only approximates the regularization properly in a region that is of the size of the regularization parameter. So, as soon as the current iterate of the gradient descent takes values near 0 or 1, the timestep control only allows timestep sizes of the order of the regularization parameter, which, as mentioned above, cannot be chosen too big.

Instead of using a penalty term one could of course also approach the constrained convex optimization problem directly. This is done for example by Bresson et al. [15]. Their approach does not need a penalty term and gives an efficient algorithm to minimize E_{CE}, but has to introduce an additional unknown v and a regularization parameter θ and needs to minimize for u and v alternatingly. Furthermore, the key idea to apply Chambolle's TV minimization algorithm [16] can also be directly applied to our model to obtain a simpler and faster minimization algorithm: There is no need to introduce v, θ and the alternating minimization. Therefore it is worth to investigate whether it is possible to simplify the problem by getting rid of the constraint altogether.

3 Unconstrained global 2-phase minimization

Another alternative to Chan Vese is a phase field approach [6, 5] with a typical double well term:

$$E_{\mathrm{PH},\epsilon}[u] := \int_\Omega u^2 f_1 + (1-u)^2 f_2 + \nu \left(\frac{1}{\epsilon} u^2 (1-u)^2 + \epsilon \left|\nabla u\right|^2 \right) \mathrm{d}x.$$

A minimizer u_ϵ of this energy is a diffuse representation of the segmentation, i.e. $\{u_\epsilon = 0\}$ and $\{u_\epsilon = 1\}$ represent the two segments respectively with a smooth transition in between. $E_{\mathrm{PH},\epsilon}[u]$ is known to Γ-converge to E_{MS} [5], but unfortunately not convex and does not permit jumps in u for $\epsilon > 0$.

Knowing both E_{CE} and $E_{\mathrm{PH},\epsilon}$, the question arises whether it is possible to combine the advantages of both models while eliminating some of the disadvantages. Heuristically looking at both energies served as motivation to investigate the following energy:

$$E[u] := \int_\Omega u^2 f_1 + (1-u)^2 f_2 + \nu |\nabla u| \mathrm{d}x. \tag{5}$$

This energy is convex because it does not involve the non-convex double well term of $E_{\mathrm{PH},\epsilon}$, and can be minimized without imposing constraints because it does not have the indicator term from E_{CE} that is not bounded from below. Furthermore, it permits jumps in u.

Remark 4. Given a function u, obviously we have

$$E[\min\{\max\{0,u\},1\}] \leq E[u].$$

Therefore, a minimizer $u_{\min}$ fulfills $0 \leq u_{\min} \leq 1$.

While the proposed functional has some nice obvious properties, it is far from obvious whether there is a relation between its minimizer and minimizers of E_{MS}. Before we tackle this question, let us remark a link between E_{CE} and E:

Remark 5. There is a direct relationship between E_{CE} and E: A straightforward calculation shows

$$u^2 f_1 + (1-u)^2 f_2 = (f_1 - f_2)\, u + (u - \tfrac{1}{2})^2 (f_1 + f_2) - \tfrac{1}{4}(f_1 + f_2) + f_2.$$

Therefore

$$\begin{aligned} E[u] =& \int_\Omega (f_1 - f_2)\, u + (u - \tfrac{1}{2})^2 (f_1 + f_2) - \tfrac{1}{4}(f_1 + f_2) + f_2 + \nu |\nabla u| \mathrm{d}x \\ =& E_{\mathrm{CE}}[u] + \int_\Omega (f_1 + f_2)(u - \tfrac{1}{2})^2 \mathrm{d}x + C. \end{aligned}$$

In other words, E essentially equals E_{CE} plus an additional quadratic penalty energy. The constant C is clearly irrelevant for the minimizers.

To investigate the relation between the minimizers of E and minimizers of E_{MS} we can make use of the theory derived in the context of the connection between minimal surface problems and total variation minimization.

The following general statement has been made by Chambolle [17], Chambolle and Darbon [11], in the continuous setting, its discrete counterpart is well known:

Theorem 2. *Let $\Psi : \Omega \times \mathbb{R} \to \mathbb{R}, (x,s) \mapsto \Psi(x,s)$ such that $\Psi(x,\cdot)$ is C^1 and uniformly convex for all $x \in \Omega$ and*

$$u := \operatorname*{argmin}_{\tilde{u}} \int_{\Omega} \Psi(x, \tilde{u}(x)) + \nu |\nabla \tilde{u}| \mathrm{d}x.$$

Then $\Sigma_c := \{x \in \Omega | u(x) > c\}$ for all $c \in \mathbb{R}$ is a minimizer of

$$\int_{\Sigma} \partial_s \Psi(x,c) \mathrm{d}x + \nu \operatorname{Per}(\Sigma).$$

Note that this general statement cannot be directly applied to the model of Nikolova et al. discussed in Section 2 because the integrand is neither uniformly (not even strictly) convex nor does the general statement incorporate the constraint.

As remarked in [11], the proof for a more specific statement given in [10] still applies to Theorem 2.

Theorem 3. *If u is a minimizer of (5), then $\{u > \frac{1}{2}\}$ minimizes*

$$E_{MS}[\Sigma] = \int_{\Sigma} f_1 \mathrm{d}x + \int_{\Omega \setminus \Sigma} f_2 \mathrm{d}x + \nu \operatorname{Per}(\Sigma).$$

Proof. Let $\Psi(x,s) := s^2 f_1(x) + (1-s)^2 f_2(x)$. Obviously $\Psi(x,\cdot)$ is C^2 for all $x \in \Omega$ and we have $\partial_s \Psi(x,s) = 2s f_1(x) + 2(s-1) f_2(x)$ and $\partial_s^2 \Psi(x,s) = 2(f_1(x) + f_2(x))$. From Remark 2, we know that $f_1, f_2 > 0$ a.e., therefore $\Psi(x,\cdot)$ is uniformly convex for a.e. $x \in \Omega$. Now just apply Theorem 2, noting $\partial_s \Psi(x, \frac{1}{2}) = f_1(x) - f_2(x)$ and Remark 1. $\square$

In this sense, our theorem is a corollary of Theorem 2.

The preceding theorem finally tells us how to find a global minimizer of $E_{\mathrm{MS}}[\cdot]$ given in (1): Minimize the convex energy (5) and threshold the minimizer to $\frac{1}{2}$. In case of the piecewise constant Mumford-Shah functional for 2-phase segmentation, we obtain a global minimizer of the Mumford-Shah energy (2) with respect to Σ for fixed gray values c_1, c_2. We do not necessarily find a global minimizer with respect to Σ, c_1 and c_2.

Another link between E_{CE} and $E_{\mathrm{PH},\epsilon}$ is the so-called piecewise constant levelset method [18] for 2-phase segmentation that constrains the levelset function to be piecewise constant. If this constraint is approximated with a penalty energy, the method equals the phase field approach. If the constraint is relaxed to a certain boundedness constraint, the method equals [9]. In both cases the fidelity term has to be altered accordingly, making use of the fact that this term is the same in E_{CE} and $E_{\mathrm{PH},\epsilon}$ if u only takes the values 0 and 1.

Since (5) is similar to the Rudin-Osher-Fatemi energy [19], there is a wide variety of established minimization schemes to choose from, ranging from a straightforward gradient descent scheme with a differentiable approximation of the BV term over primal thresholding methods [20] to sophisticated methods based on the dual formulation of the BV norm, e.g. [16, 11].

With $\Psi(x,s) = \frac{1}{2}\left(s - (f_2(x) - f_1(x))\right)^2$, another immediate consequence of Theorem 2 is that the zero superlevelset of a minimizer of the ROF energy

$$E_{\mathrm{ROF}}[u] := \int_\Omega \frac{1}{2}\left(u - (f_2 - f_1)\right)^2 + \nu|\nabla u|\mathrm{d}x \tag{6}$$

is a global minimizer of $\hat{E}_{\mathrm{MS}}$ and therefore of E_{MS}. This is another way to obtain a global minimizer of E_{MS} by unconstrained convex optimization, but compared to (5) this method has a few shortcomings, cf. Sections 4 and 5. Furthermore, the boundedness mentioned in Remark 4 does not hold for minimizers of the ROF energy. Perhaps this is one of the reasons why nobody seems to have used the classical ROF function for Mumford-Shah based image segmentation so far.

4 Multiphase segmentation

Our functional can be extended to multiphase segmentation by the using the idea of Vese and Chan [2] in a straightforward manner. To keep notation at bay, we restrict the discussion to segmentation in 4 phases. The segmentation in 2^n phases works analogously. Let $f_1, f_2, f_3, f_4 \in L^1(\Omega)$ such that $f_i \geq 0$ a.e., then the multiphase functional is given by

$$\begin{aligned} E[u_1,u_2] := \int_\Omega & u_1^2 u_2^2 f_1 + (1-u_1)^2 u_2^2 f_2 \\ & + u_1^2(1-u_2)^2 f_3 + (1-u_1)^2(1-u_2)^2 f_4 \\ & + \nu\left(|\nabla u_1| + |\nabla u_2|\right)\mathrm{d}x. \end{aligned} \tag{7}$$

If we fix u_2, the reduced functional $E[\cdot, u_2]$ is the same as the 2-phase functional (5) with the indicator functions $\tilde{f}_1 = u_2^2 f_1 + (1-u_2)^2 f_3$ and $\tilde{f}_2 = u_2^2 f_2 + (1-u_2)^2 f_4$. As in the 2-phase case, we can assume $f_i > 0$ a.e. without loss of generality and because either $u_2^2 > 0$ or $(1-u_2)^2$ holds, we have $\tilde{f}_1, \tilde{f}_2 > 0$. Therefore, all statements proven for the 2-phase functional can be applied to $E[\cdot, u_2]$, i.e. we can compute the global minimum (for fixed u_2). The same applies for fixed u_1, so as an optimization strategy, we propose to minimize with respect to u_1 and u_2 alternatingly.

Even though it is easy to extend (5) to multiphase segmentation, the same does not apply to the ROF energy (6). There is no apparent extension in the sense of [2] to formulate the multiphase segmentation in a single functional.

5 Indicator parameters

In typical segmentation tasks, the indicator functions depend on unknown parameters, e.g. the grey values for each segment in case of the piecewise constant Mumford-Shah model. For the sake of simplicity, we discuss the latter model in its 2-phase formulation here, i.e. $f_i(x) := (u_0(x) - c_i)^2, i = 1, 2$, but this discussion applies to other indicator functions and multiphase segmentation as well.

During the minimization of (5) we have to minimize for c_1 and c_2 as well. This is typically done in an alternating fashion, but there are two apparent possibilities to update the grey values: Minimize (5) with respect to c_1 and c_2 or do so for the energy in the set formulation (1). The two possible updating formulae for c_1 two arising are

$$c_1 = \int_\Omega u^2 u_0 \mathrm{d}x \Big/ \int_\Omega u^2 \mathrm{d}x \text{ or } c_1 = \int_{\{u>\frac{1}{2}\}} u_0 \mathrm{d}x \Big/ \int_{\{u>\frac{1}{2}\}} \mathrm{d}x.$$

The two possibilities only coincide if u is binary. The first formula not only averages u_0 in $\{u > \frac{1}{2}\}$, instead it takes into account the values of u_0 everywhere, but weights the values according to u^2. To a certain degree this is similar to the effect of the regularization of the Heaviside function in the model of Chan and Vese. From our experiments, this reduces the chance of getting stuck in local minima that can still occur when minimizing over u and the indicator parameters. Particulary in the case of multiphase segmentation it turned out to be beneficial.

Due to the different way f_1 and f_2 are used in the ROF energy (6), it is not quadratic in c_1 and c_2. So this functional does not give a natural formula to update the grey values.

6 Numerical examples

To conclude, we show the practical usability of the proposed model by applying it to the classical piecewise constant Mumford-Shah functional, see equation (2). As minimization method we use an explicit gradient descent scheme with the Armijo rule [14] as timestep size control. The absolute value is regularized by $|z|_\epsilon = \sqrt{z^2 + \epsilon^2}$ (in all examples presented here, $\epsilon = 0.1$ is used). For the spatial discretization, we use bilinear finite elements on a regular quadrilateral grid, i.e. each pixel of the input image u_0 corresponds to a node of the finite element mesh. The grey values c_1 and c_2 are initialized with 0 and 1 respectively and updated occasionally during the gradient descent.

Figure 1 shows results of our method and of the one proposed by Nikolova et al. [9] on one artificial image and one digital photo. In both examples, the minimizer u from our model is far from being binary, but this is nothing to be expected from the theory presented in this paper. The 0.5-superlevelset gives an accurate segmentation that is not influenced by the presence of heavy noise (top row) and works on non-binary input images (bottom row). The minimizers u

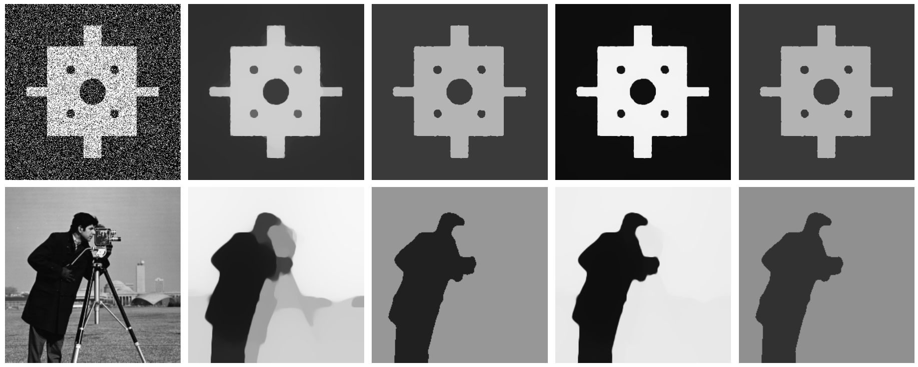

Fig. 1. Segmentation of an artificial noisy structure ($\nu = 2 \cdot 10^{-3}$, top row) and the well-known Matlab cameraman image ($\nu = 4 \cdot 10^{-3}$, bottom row): Input image u_0 (left), segmentation function u and 0.5-superlevelset of u colored with the average grey values c_1, c_2 obtained by our model (middle) and by using E_{CE} (right). The slight difference of the grey values is attributed to the employed update formula, cf. Section 5.

of the Nikolova et al. model look very different, but the segmentation obtained from the 0.5-superlevelsets is almost identical.

Upon closer inspection, the minimizer u of our model from the top row of Figure 1 looks very much like as obtained by minimizing the ROF energy with u_0 as input image. This is not surprising due to the following observation: If u_0 is binary, i.e. $u_0 = \chi_A$ for a set $A \subset \Omega$ and $c_1 = 0$, $c_2 = 1$ we have $f_1 = (\chi_A - 0)^2 = \chi_A$ and $f_2 = (\chi_A - 1)^2 = \chi_{\Omega \setminus A}$ and therefore

$$E[u] = \int_\Omega (u - \chi_{\Omega \setminus A})^2 + \nu |\nabla u| \mathrm{d}x,$$

i.e. E equals the ROF energy in this special case. This is not the case if u_0 is non-binary which can be seen from the bottom row of Figure 1.

Figure 2 shows 4-phase segmentation results. Those indicate the tendency of the segmentation functions to become binary for small values of ν.

Finally, Figure 3 illustrates the behavior of the method for different numbers of segments and Figure 4 shows three timesteps of the 8-phase segmentation.

References

1. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Communications on Pure Applied Mathematics **42** (1989) 577–685
2. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the Mumford and Shah model. International Journal of Computer Vision **50**(3) (2002) 271–293
3. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Transactions on Image Processing **10**(2) (2001) 266–277

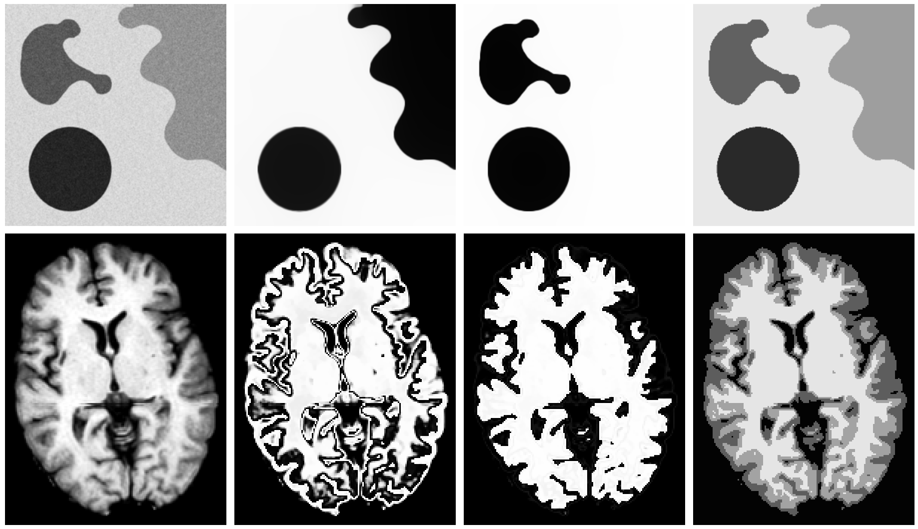

Fig. 2. 4-phase segmentation of an artificial noisy image (top row) and a MRI image (bottom row) ($\nu = 6 \cdot 10^{-4}$): Input image u_0 (left), segmentation functions u_1 and u_2 (middle), segmentation colored with the average grey values $c_1, ..., c_4$ (right).

4. Osher, S.J., Sethian, J.A.: Fronts propagating with curvature dependent speed: Algorithms based on Hamilton–Jacobi formulations. Journal of Computational Physics **79** (1988) 12–49
5. Shen, J.: Γ-convergence approximation to piecewise constant Mumford-Shah segmentation. In: 7th International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS'05). Volume 3708 of LNCS. (2005) 499–506
6. Esedoḡlu, S., Tsai, Y.H.R.: Threshold dynamics for the piecewise constant Mumford-Shah functional. Journal of Computational Physics **211**(1) (2006) 367–384
7. Merriman, B., Bence, J.K., Osher, S.J.: Diffusion generated motion by mean curvature. CAM Report 92-18, UCLA (1992)
8. Alvino, C.V., Yezzi, A.J.: Fast Mumford-Shah segmentation using image scale space bases. In: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. Volume 6498. (2007)
9. Nikolova, M., Esedoḡlu, S., Chan, T.F.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM Journal on Applied Mathematics **66**(5) (2006) 1632–1648
10. Chambolle, A.: An algorithm for mean curvature motion. Interfaces and free Boundaries **6** (2004) 195–218
11. Chambolle, A., Darbon, J.: On total variation minimization and surface evolution using parametric maximum flows. CAM Report 08-19, UCLA (2008)
12. Alter, F., Caselles, V., Chambolle, A.: A characterization of convex calibrable sets in $\mathbb{R}^N$. Mathematische Annalen **332**(2) (2005) 329–366
13. Ambrosio, L., Fusco, N., Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford Mathematical Monographs. Oxford University Press, New York (2000)

Fig. 3. Segmentation of a digital photo ($\nu = 2 \cdot 10^{-5}$). Input image u_0 (left), segmentation in four (middle) and eight (right) segments colored with the average grey values of the segments. Original image © bigmama / PIXELIO.

Fig. 4. Intermediate results of the segmentation in eight segments shown in Figure 3 after 50 (left), 250 (middle) and 700 (right) gradient descent steps.

14. Kosmol, P.: Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben. 2. edn. Teubner, Stuttgart (1993)
15. Bresson, X., Esedoḡlu, S., Vandergheynst, P., Thiran, J., Osher, S.: Fast global minimization of the active contour/snake model. Journal of Mathematical Imaging and Vision **28**(2) (2007) 151–167
16. Chambolle, A.: An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision **20**(1-2) (2004) 89–97
17. Chambolle, A.: Total variation minimization and a class of binary MRF models. In: Energy Minimization Methods in Computer Vision and Pattern Recognition. Volume 3757 of LNCS. (2005) 136–152
18. Lie, J., Lysaker, M., Tai, X.C.: A binary level set model and some applications to Mumford-Shah image segmentation. IEEE Transactions on Image Processing **15**(5) (2006) 1171–1181
19. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60** (1992) 259–268
20. Daubechies, I., Defrise, M., de Mol, C.: An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. Communications on Pure and Applied Mathematics **57**(11) (2004) 1413–1457