# Variational Image Segmentation Model Coupled with Image Restoration Achievements

Xiaohao Cai

*Department of Plant Sciences, and Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, UK*
*xc274@cam.ac.uk*

**Abstract**

Image segmentation and image restoration are two important topics in image processing with a number of important applications. In this paper, we propose a new multiphase segmentation model by combining image restoration and image segmentation models. Utilizing aspects of image restoration, the proposed segmentation model can effectively and robustly tackle images with a high level of noise or blurriness, missing pixels or vector values. In particular, one of the most important segmentation models, the piecewise constant Mumford-Shah model, can be extended easily in this way to segment gray and vector-valued images corrupted, for example, by noise, blur or information loss after coupling a new data fidelity term which borrowed from the field of image restoration. It can be solved efficiently using the alternating minimization algorithm, and we prove the convergence of this algorithm with three variables under mild conditions. Experiments on many synthetic and real-world images demonstrate that our method gives better segmentation results in terms of quality and quantity in comparison to other state-of-the-art segmentation models, especially for blurry images and those with information loss.

*Keywords:*
image segmentation, image restoration, piecewise constant Mumford-Shah model

## 1. Introduction

Image segmentation and image restoration are two important concepts in image processing. Image segmentation consists in partitioning a given image into multiple segments to transfer the representation of the image into a more meaningful one which is easier to analyze. It is typically used to locate objects and boundaries within an image. Image restoration is the operation of estimating a desired clean image from its corrupted version. Corruption may come in many forms, such as blur, noise, camera misfocus, or information loss. Obviously, image segmentation can be used as preprocessing or postprocessing of image restoration. In other words, these two topics influence each other substantially.

Let $\Omega \subset \mathbb{R}^2$ be a bounded, open, connected set, and $f : \Omega \to \mathbb{R}$ a given image. Without loss of generality, we restrict the range of $f$ to [0,1]. Let $g : \Omega \to \mathbb{R}$ denote the desired clean image, then $f = g + n_f$, where $n_f$ is the additive noise. Many image restoration models can be written in the form

$$E(g) = \mu\Phi(f, g) + \phi(g), \tag{1}$$

where $\Phi(f, g)$ is the data fidelity term, $\phi(g)$ is the regularization term, and $\mu > 0$ is a regularization parameter balancing the trade-off between terms $\Phi(f, g)$ and $\phi(g)$. If set $\Phi(f, g) = \int_\Omega (f - g)^2 dx$ and $\phi(g) = \int_\Omega |\nabla g| dx$ (total variation term), then model (1) becomes the ROF model proposed by Rudin, Osher and Fatemi in 1992 [37], i.e.,

$$E(g) = \mu \int_\Omega (f - g)^2 dx + \int_\Omega |\nabla g| dx. \tag{2}$$

One important advantage of model (2) is that it preserves the edge information of $f$ very well, but it also introduces the staircase effect. Many previous attempts to remove the staircase effect are based on higher-order derivative terms,

see [5, 14, 29, 38, 45]. For example in [5], tight-frame technic was used in $\phi(g)$ to obtain more details of higher-order derivative information of $f$. The relationship between total variation and tight-frame can be found in [41]. As we know, the data fidelity term $\Phi(f,g) = \int_\Omega (f-g)^2 dx$ is especially effective for Gaussian noise [37]. For removing Poisson noise, $\Phi(f,g) = \int_\Omega (g - f\log g)dx$ is proposed [17], whilst $\Phi(f,g) = \int_\Omega |f-g|dx$ is proposed for removing impulsive noise [31]. Please refer to [1, 11, 12, 17, 25, 31, 27, 39, 40] and references therein for the details of Poisson and impulsive noise removal. Note that the image restoration model (1) can be extended to process blurry images after introducing a problem related linear operator $\mathcal{A}$ in front of $g$ [36].

Let $\Gamma \in \Omega$ represent the boundary of the objects within an image, and $\Omega_i$ be the parts of the segmented objects fulfilling $\Omega = \cup_i \Omega_i \cup \Gamma$. The Mumford-Shah model is one of the most important image segmentation models, and has been studied extensively in the last twenty years. More precisely, in [30], Mumford and Shah proposed an energy minimization problem which approximates the true solution by finding optimal piecewise smooth approximations. The energy minimization problem was formulated as

$$E(g,\Gamma) = \frac{\lambda}{2}\int_\Omega (f-g)^2 dx + \frac{\mu}{2}\int_{\Omega\setminus\Gamma} |\nabla g|^2 dx + \text{Length}(\Gamma), \tag{3}$$

where $\lambda$ and $\mu$ are positive parameters, and $g : \Omega \to \mathbb{R}$ is differentiable in $\Omega \setminus \Gamma$ but may be discontinuous across $\Gamma$. Because model (3) is nonconvex, it is very challenging to find or approximate its minimizer, see [8, 9, 22]. Many works [26, 44] concentrate on simplifying model (3) by restricting $g$ to be a piecewise constant function ($g = c_i$ in $\Omega_i$), i.e.,

$$E(g,\Gamma) = \frac{\lambda}{2}\sum_{i=1}^{K}\int_{\Omega_i} (f-c_i)^2 dx + \text{Length}(\Gamma), \tag{4}$$

where $K$ is the known number of phases of $f$. Using the coarea formula [20], and recognizing that the total variation of the characteristic function of a set is its perimeter [20, 21], model (4) can be rewritten as

$$\begin{aligned} E(c_i,u_i) &= \lambda\sum_{i=1}^{K}\int_\Omega (f-c_i)^2 u_i dx + \sum_{i=1}^{K}\int_\Omega |\nabla u_i| dx, \\ \text{s.t.} \quad &\sum_{i=1}^{K} u_i(x) = 1, u_i(x) \in \{0,1\}, \forall x \in \Omega. \end{aligned} \tag{5}$$

Moreover, model (5) with $K = 2$ is the Chan-Vese model [16], and with fixed $c_i$ is a special case of the Potts model [34]. Due to the nonconvex property of (5), the exact convex version of (5) was proposed when $K = 2$ and $c_i$ fixed [13]. For $K > 2$, recent authors have focused on relaxing $u_i$ and solving the following model

$$\begin{aligned} E(u_i,c_i) &= \lambda\sum_{i=1}^{K}\int_\Omega (f-c_i)^2 u_i dx + \sum_{i=1}^{K}\int_\Omega |\nabla u_i| dx, \\ \text{s.t.} \quad &\sum_{i=1}^{K} u_i(x) = 1, u_i(x) \geq 0, \forall x \in \Omega. \end{aligned} \tag{6}$$

Please refer to [2, 24, 28, 33, 46] and references therein for more details. One drawback of model (6) is that *it is not good at segmenting images corrupted by blur or information loss*, and this is one of the main problems to solve in this current paper.

In [32], a model of coupling image restoration and segmentation based on a statistical framework of generalized linear models was proposed, but the analysis and algorithm therein are only focused on the two-phase segmentation problem. In our previous work [6], a two-stage segmentation method which provides a better understanding of the link between image segmentation and image restoration was proposed. The method suggests that for segmentation, it is reasonable and practicable to extract different phases in $f$ by using image restoration methods first and thresholding second. Moreover, in our recent work [7], we proved that the solution of the Chan-Vese model [16] for certain $\lambda$ can actually be given by thresholding the minimizer of the ROF model (2) using a proper threshold, which clearly

provides one kind of relationship between image segmentation and image restoration.

In this paper, we propose starting with the extension of the piecewise constant Mumford-Shah model (4) to manage blurry image, and then applying a novel segmentation model by composing model (4) with a data fidelity term borrowed from the field of image restoration. Since we only add a new fidelity term, which usually possesses good properties such as differentiability, the solution of the proposed model is not more involved compared with solving model (4). It can be solved efficiently using the alternating minimization (AM) algorithm [18] with the ADMM or primal-dual algorithms [4, 10, 23]. We prove that under mild conditions, the AM algorithm converges for the proposed model. The proposed model can segment blurry images easily but model (4) cannot. Moreover, it can also deal with images with information loss and vector-valued images (e.g., color images). Due to the advantage of two data fidelity terms, one from image restoration and the other from image segmentation, our model is much more robust and stable. Experiments on many kinds of synthetic and real-world images demonstrate that our method gives better segmentation results in comparison with other state-of-the-art segmentation methods, especially in the case of blurry images and images with information loss.

**Contributions.** The main contributions of this paper are summarized as follows.

1) Coupling variational image segmentation models and image restoration models, which provides a new methodology for multiphase image segmentation.

2) Extending the piecewise constant Mumford-Shah model (4) by incorporating image restoration achievements so that the new constructed variation segmentation model can handle blurry images easily.

3) Thanks to the image restoration achievements, the new variation segmentation model has the potential to process many different types of noise, for example Gaussian, Poisson and impulsive noise.

4) Different kinds of vector-valued image, for example color images, and images with information loss, are also covered in the proposed variational segmentation model.

5) The convergence of the AM algorithm with three variables to the proposed variational model is proved.

The rest of this paper is organized as follows. In Section 2, we propose our new segmentation model and extend it so that it can deal with vector-valued images and images with some pixel values missing. An AM algorithm for our model is introduced in Section 3. The convergence of it will be proved in Section 4. In Section 5, we compare the performance of our proposed method with state-of-the-art multiphase segmentation methods on various synthetic and real-world images. Conclusions are drawn in Section 6.

## 2. The Proposed Variational Image Segmentation Model

We propose our image segmentation model by combining the piecewise constant Mumford-Shah model (5) with the fidelity term $\Phi(f, g)$, which comes from the image restoration model (1). More precisely, our proposed segmentation model aims to minimize the energy

$$\begin{aligned} E(u_i, c_i, g) &= \mu\Phi(f, \mathcal{A}g) + \lambda\Psi(g, u_i, c_i) + \sum_{i=1}^{K} \int_\Omega |\nabla u_i| dx, \\ \text{s.t.} \quad & \sum_{i=1}^{K} u_i(x) = 1, u_i(x) \in \{0, 1\}, \forall x \in \Omega, \end{aligned} \tag{7}$$

where $g \in L^2(\Omega)$ and $\mathcal{A}$ is the problem-related linear operator. For example, $\mathcal{A}$ can be the identity operator for a noisy observed image $f$ or a blurring operator if there are both noise and blur in $f$. The blurring operator $\mathcal{A}$, for example Gaussian bur or motion blur, can be estimated by using image deblurring methods (please refer to [35, 43, 47] and references therein). Therefore, researchers generally assume $\mathcal{A}$ is known in image segmentation or image restoration. The first term $\Phi(f, \mathcal{A}g)$ is a data fidelity term arising from the image restoration model (1). It controls $g$ to within a short distance of the given corrupted image $f$; in other words, it aims to deblur and denoise according to the types of

noise in $f$. Term $\Psi(g, u_i, c_i)$ is also a data fidelity term but comes from an image segmentation model, and aims to separate $g$ into $K$ specified segments. In this paper, we restrict ourselves to

$$\Psi(u_i, c_i, g) = \sum_{i=1}^{K} \int_\Omega (g - c_i)^2 u_i dx.$$

The last term in (7) is the regularization term, which controls the length of the boundaries of the segmented parts $u_i$. The type of data fidelity term $\Phi(f, \mathcal{A}g)$ changes according to different noise models, for example,

i. Gaussian noise: $\Phi(f, \mathcal{A}g) = \int_\Omega (f - \mathcal{A}g)^2 dx$;

ii. Poisson noise (I-divergence): $\Phi(f, \mathcal{A}g) = \int_\Omega (\mathcal{A}g - f \log(\mathcal{A}g)) dx$;

iii. Impulsive noise: $\Phi(f, \mathcal{A}g) = \int_\Omega |f - \mathcal{A}g| dx$.

Compared with model (5), model (7) has two advantages: it is able to segment blurry images, and its two data fidelity terms make it much more robust and stable in the processing of the given corrupted image $f$.

Obviously, for fixed $g$, model (7) is reduced to model (5). The following theorem 1 gives the uniqueness of $g$ when minimizingl (7) for fixed $c_i$ and $u_i$.

**Theorem 1.** *Assume $\Phi(f, \mathcal{A}g)$ in (7) is convex and continuous, then there exists one and only one $g$ which minimizes energy (7) for fixed $c_i$ and $u_i$.*

*Proof.* See the Appendix. □

In the following, we restrict ourselves to $\Phi(f, \mathcal{A}g) = \int_\Omega (f - \mathcal{A}g)^2 dx$ as one example to show how to extend model (7) so that it can handle images with missing information and vector-valued images. Let $\Omega'$ be the set containing the pixels whose pixel values are missing. Then model (7) can be extended to segment images with information loss as

$$E(u_i, c_i, g) = \mu \int_\Omega (f - \mathcal{A}g)^2 \omega dx + \lambda \sum_{i=1}^{K} \int_\Omega (g - c_i)^2 \omega u_i dx + \sum_{i=1}^{K} \int_\Omega |\nabla u_i| dx, \tag{8}$$

where

$$\sum_{i=1}^{K} u_i(x) = 1, u_i(x) \in \{0, 1\}, \ \omega(x) = \begin{cases} 1, & \text{if } x \in \Omega \setminus \Omega', \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

For the observed vector-valued image represented as $\boldsymbol{f} = (f_1, \cdots, f_N)$, let $\boldsymbol{g} = (g_1, \cdots, g_N)$ and $\boldsymbol{c}_i = (c_{i,1}, \cdots, c_{i,N})$, model (7) can be extended to segment vector-valued images with information loss as

$$E(u_i, \boldsymbol{c}_i, \boldsymbol{g}) = \mu \sum_{j=1}^{N} \int_\Omega (f_j - \mathcal{A}_j g_j)^2 \omega dx + \lambda \sum_{i=1}^{K} \sum_{j=1}^{N} \int_\Omega (g_j - c_{i,j})^2 \omega u_i dx + \sum_{i=1}^{K} \int_\Omega |\nabla u_i| dx, \tag{10}$$

where $u_i$ and $\omega$ are defined in (9).

## 3. The AM Algorithm

We first transfer (7) by relaxing $u_i$ to the following version

$$E(u_i, c_i, g) = \mu \Phi(f, \mathcal{A}g) + \lambda \sum_{i=1}^{K} \int_\Omega (g - c_i)^2 \omega u_i dx + \sum_{i=1}^{K} \int_\Omega |\nabla u_i| dx, \tag{11}$$

$$\text{s.t.} \quad \sum_{i=1}^{K} u_i(x) = 1, u_i(x) \geq 0, \forall x \in \Omega. \tag{12}$$

Using the AM algorithm [18], a partial minimizer $(g, c_i, u_i)$ of (11) can be computed alternatively as follows:

i. Find $g$ as minimizer of (11) for fixed $u_i$ and $c_i$. Obviously, $g$ is only contained in the first two terms of (11), and the second term can be regarded as one kind of Tikhonov regularizations when solving $g$, see [42]. The algorithm to find $g$ depends on the choice of $\Phi(f, \mathcal{A}g)$. For example, when $\Phi(f, \mathcal{A}g) = \int_\Omega (f - \mathcal{A}g)^2 \omega dx$, since it is differentiable, we have

$$g = (\mu \mathcal{A}^T \mathcal{A} + \lambda)^{-1} (\mu \mathcal{A}^T f + \lambda \sum_{i=1}^{K} c_i u_i)\omega. \tag{13}$$

For solving $g$ according to the choice of $\Phi(f, \mathcal{A}g)$ to Poisson or impulsive noise, we will pursue this in the future.

ii. Find $c_i$ as minimizer of (11) for fixed $u_i$ and $g$. Let $c = (c_1, \cdots, c_K)$. Clearly, $c_i$ is just related with the second term of (11), therefore

$$c_i = \frac{\int_\Omega g\omega u_i dx}{\int_\Omega \omega u_i dx}. \tag{14}$$

iii. Find $u_i$ as minimizer of (11) for fixed $g$ and $c_i$. The discussion of this is given in the following.

Note that when $g$ is fixed, the first term of model (11) is constant, hence the problem of finding $u_i$ is reduced to minimize model (6) with $\omega$. Therefore, there are many methods we can use, for example the ADMM method in [4, 23, 24] which will be given explicitly in the following. Alternatively, one can apply the primal-dual algorithm [10, 33] or the max-flow approach [46].

Let $u(j) = (u_i(j))_{i=1}^K$, $s = (s_i)_{i=1}^K = ((g - c_i)^2 \omega)_{i=1}^K$. Our problem can be transferred to be

$$\min_{v,u,d} \lambda \langle v, s\rangle + \|d\|_1 + \iota_S(u), \quad \text{s.t.} \quad \nabla v = d, v = u, \tag{15}$$

where $\iota_S(\cdot)$ is the indicator function defined as

$$\iota_S(y) := \begin{cases} 0, & \text{if } y \in S, \\ +\infty, & \text{otherwise,} \end{cases}$$

and $S := \{y \in \mathbb{R}^K | \sum_{i=1}^K y_i = 1, y \geq 0\}$. Then iterate the following steps until convergence

$$\begin{aligned} v^{k+1} &= \operatorname*{argmin}_v \left\{\lambda\langle v, s\rangle + \sigma(\|b_d^k + \nabla v - d^k\|^2 + \|b_u^k + v - u^k\|^2)\right\}, \\ d^{k+1} &= \operatorname*{argmin}_d \left\{\|d\|_1 + \sigma\|b_d^k + \nabla v^{k+1} - d\|^2\right\}, \\ u^{k+1} &= \operatorname*{argmin}_d \left\{\iota_S(u) + \sigma\|b_u^k + v^{k+1} - u\|^2\right\}, \\ b_d^{k+1} &= b_d^k + \nabla v^{k+1} - d^{k+1}, \\ b_u^{k+1} &= b_u^k + v^{k+1} - u^{k+1}. \end{aligned} \tag{16}$$

After $u = (u_i)_{i=1}^K$ is solved, each segment $\Omega_i$ can be obtained by

$$\Omega_i = \left\{x | u_i(x) = \max\{u_1(x), \cdots, u_K(x)\}, \forall x \in \Omega\right\}. \tag{17}$$

In summary, the AM algorithm to solve model (11) is given in Algorithm 1.

## 4. Convergence Analysis

In this section, we discuss the convergence property of Algorithm 1. We first give the very general conclusion to the AM algorithm for three variables. Let $X \subset \mathbb{R}^{m_1}, Y \subset \mathbb{R}^{m_2}$ and $Z \subset \mathbb{R}^{m_3}$ be closed sets, and the energy function

**Algorithm 1** The AM Algorithm to Model (11)

**Input:** Observed image $f$, number of phases $K$, $c^{(0)}$, and $u^{(0)}$.

1: **while** $\|c^{(k+1)} - c^k\| > \epsilon$ **do**
2: find $g^{(k+1)}$ as minimizer of (11) for fixed $u^{(k)}$, and $c^{(k)}$ using (13);
3: find $c^{(k+1)}$ as minimizer of (11) for fixed $g^{(k+1)}$ and $u^{(k)}$ using (14);
4: find $u^{(k+1)}$ as minimizer of (11) for fixed $g^{(k+1)}$ and $c^{(k+1)}$ using (16);
5: **end while**
6: **return** $\Omega_i, i = 1, \ldots, K$ using (17)

$E : X \times Y \times Z \to \mathbb{R}$ be continuous and bounded from below. To process the AM algorithm, we start with some initial guess $y^{(0)}, z^{(0)}$, then we successively obtain the alternating sequence (between $z, y$ and $x$) of conditional minimizers

$$z^{(0)}, y^{(0)} \to x^{(0)} \to z^{(1)} \to y^{(1)} \to x^{(1)} \to \cdots$$

from solving, for $k = 0, 1, \ldots$,

$$\begin{aligned} x^{(k)} &\in \operatorname*{argmin}_x E(x, y^{(k)}, z^{(k)}), \\ z^{(k+1)} &\in \operatorname*{argmin}_z E(x^{(k)}, y^{(k)}, z), \\ y^{(k+1)} &\in \operatorname*{argmin}_y E(x^{(k)}, y, z^{(k+1)}). \end{aligned} \tag{18}$$

**Theorem 2.** ***(Monotonicity of Alternating Minimization).*** *Let $X \subset \mathbb{R}^{m_1}, Y \subset \mathbb{R}^{m_2}$ and $Z \subset \mathbb{R}^{m_3}$ be closed sets, and $E : X \times Y \times Z \to \mathbb{R}$ be continuous and bounded from below. Then, for each $k \geq 0$, the following relations are satisfied*

$$\begin{aligned} E(x^{(k)}, y^{(k+1)}, z^{(k+1)}) &\leq E(x^{(k-1)}, y^{(k)}, z^{(k)}), \\ E(x^{(k)}, y^{(k)}, z^{(k+1)}) &\leq E(x^{(k-1)}, y^{(k-1)}, z^{(k)}), \\ E(x^{(k+1)}, y^{(k+1)}, z^{(k+1)}) &\leq E(x^{(k)}, y^{(k)}, z^{(k)}). \end{aligned}$$

*Hence, the sequence $\{E(x^{(k)}, y^{(k)}, z^{(k)})_{k\in\mathbb{N}}\}$ converges monotonically.*

*Proof.* See the Appendix. □

**Theorem 3.** *Let $X \subset \mathbb{R}^{m_1}, Y \subset \mathbb{R}^{m_2}$ and $Z \subset \mathbb{R}^{m_3}$ be closed sets, and $E : X \times Y \times Z \to \mathbb{R}$ be continuous and bounded from below. Then, for any convergent subsequence $(x^{(k_i)}, y^{(k_i)}, z^{(k_i)})_{i\in\mathbb{N}}$ of $(x^{(k)}, y^{(k)}, z^{(k)})_{k\in\mathbb{N}}$ generated from formula* (18) *with*

$$(x^{(k_i)}, y^{(k_i)}, z^{(k_i)}) \longrightarrow (x^*, y^*, z^*), \text{ as } i \to \infty,$$

*the following relations are satisfied:*

$$\begin{aligned} &E(x^*, y^*, z^*) \leq E(x, y^*, z^*) \quad \forall x \in X, \\ &E(x^*, y^*, z^*) \leq E(x^*, y, z^*) \quad \forall y \in Y, \\ &E(x^*, y^*, z^*) \leq E(x^*, y^*, z) \quad \forall z \in Z. \end{aligned} \tag{19}$$

*$(x^*, y^*, z^*)$ is called the the partial minimizer of $E(\cdot, \cdot, \cdot)$ if* (19) *is satisfied.*

*Proof.* See the Appendix □

Let $O$ be the set of all the partial minimizers of model (11) defined in Theorem 3. The following theorem 4 gives the convergence property of Algorithm 1.

**Theorem 4.** *Assume the operator $\mathcal{A}$ in model* (11) *is a continuous mapping and $\Phi(f, \mathcal{A}g)$ is continuous and nonnegative. As $k \to \infty$, if $(u^{(k)}, g^{(k)}, c^{(k)}) \to (u^*, g^*, c^*)$, then $(u^*, g^*, c^*) \in O$. If $(u^{(k)}, g^{(k)}, c^{(k)})_{k\in\mathbb{N}}$ does not converge, it must contain a convergent subsequence and every convergent subsequence converges to a partial minimizer of model* (11).

*Proof.* See the Appendix. □

## 5. Experimental Results

We compare our segmentation model (11) with three state-of-the-art multiphase segmentation methods [6, 24, 46] on synthetic and real-world images of different phases corrupted by noise, blur and information loss. Methods [46] and [24] use max-flow approach and ADMM algorithm respectively to minimize model (6). The difference between them is that method [46] minimizes $u_i$ with fixed $c_i$, while method [24] minimizes both $u_i$ and $c_i$. Method [6] is a two-stage segmentation method, which first solves a convex variant of the Mumford-Shah model (3) and then applies a thresholding technique. Moreover, method [6] is very effective in segmenting general kinds of images including blurry images. All the codes of methods [6, 24, 46] are provided by their authors, and the parameters of them are chosen by trial and error to give the best results of the respective methods, see Tables 1 and 2. Note that methods [6, 24] only work for gray images, but it is straightforward to extend them to segment color images, i.e., using the strategy in (10). Indeed, the authors in [46] have used this strategy to segment color images. For the color images, we compare our method with the extended method [24].

The initial codebook $c_i$ for methods [24, 46] are computed by fuzzy C-means method [3] with 100 iteration steps, and the thresholds chosen in method [6] are based on the automatic strategy therein. The tolerance $\epsilon$ and the step size $\sigma$ in Algorithm 1 and (16) are fixed to be $10^{-4}$ and 2 respectively.The parameters $\lambda$ and $\mu$ in model (11) are chosen empirically. In practice, we set the value of $\lambda$ first, then choose the value of $\mu$ according to $\lambda$ to simplify the parameters selection process. The number of phases $K$ of each testing image is known in advance.

We apply MATLAB command `imnoise` with zero mean and variable variances to add Gaussian noise into the given images. In particular, the variance used for blurry images is fixed to be $10^{-4}$. The two most popular kinds of kernels, Gaussian kernel and motion kernel, are used in $\mathcal{A}$ to blur images. If there is no special explanation, the Gaussian kernel with size $15 \times 15$ and standard deviation 15 and the motion kernel with 15 pixels and an angle of 90 degrees are used. In addition to testing our proposed algorithm and [6] according to the true blur kernels of blurry images, we also test the performance of them on degraded blur kernels generated by removing 1 percentage magnitude from each entry of the true blur kernels. MATLAB command `rand` is employed to remove some information from images randomly, and the percentage of information loss is set to be 40% unless otherwise specified. The *segmentation accuracy* (SA) is defined as

$$\text{SA} = \frac{\#\text{correctly classified pixels}}{\#\text{all pixels}} \times 100,$$

and is used to evaluate the accuracy of segmentation results in detail if the ground truth of the testing image is known. Generally, it is easy to get the ground truth of synthetic images, but this is not the case with the ground truth of real-world images. One direct way is to treat the manual segmentation from experts as the ground truth; see Fig. 9 as an example. All the results were tested on a MacBook with 2.4 GHz processor and 4GB RAM.

### *5.1. Gray image segmentation*

*Example 1: two-phase synthetic images.* To illustrate the ability of our method on images with a high level of noise and images with information loss, we first test it on two two-phase synthetic images. One contains different shapes and the other is a 2D barcode image. The latter represents data relating to the object to which it is attached and which is the most frequently used type to scan with a smart phone, see Fig. 1. Fig. 1 (A1)–(A4) give the images corrupted by Gaussian noise with variance 0.2, and Fig. 1 (A2) and (A4) give the images with part of the information removed randomly. Columns two to four of Fig. 1 are the results of methods [46, 24, 6], respectively. The last column of Fig. 1 shows the results of our method. From rows one and three of Fig. 1, showing the results of segmenting the given noisy images, we see that all the methods can give very good results visually. After comparing the *segmentation accuracy* given in the brackets under each result, we see that our method gives the highest SA compared with the other three methods. This means that our model (11) is superior compared with model (6) in improving segmentation accuracy. From the second and the fourth rows of Fig. 1, showing the results of segmenting the given images with information loss, we can clearly see that only our method gives good results both in visual terms and *segmentation accuracy*. In particular, when scaning the barcode images using a barcode scan app on a smart phone (iPhone 5), the

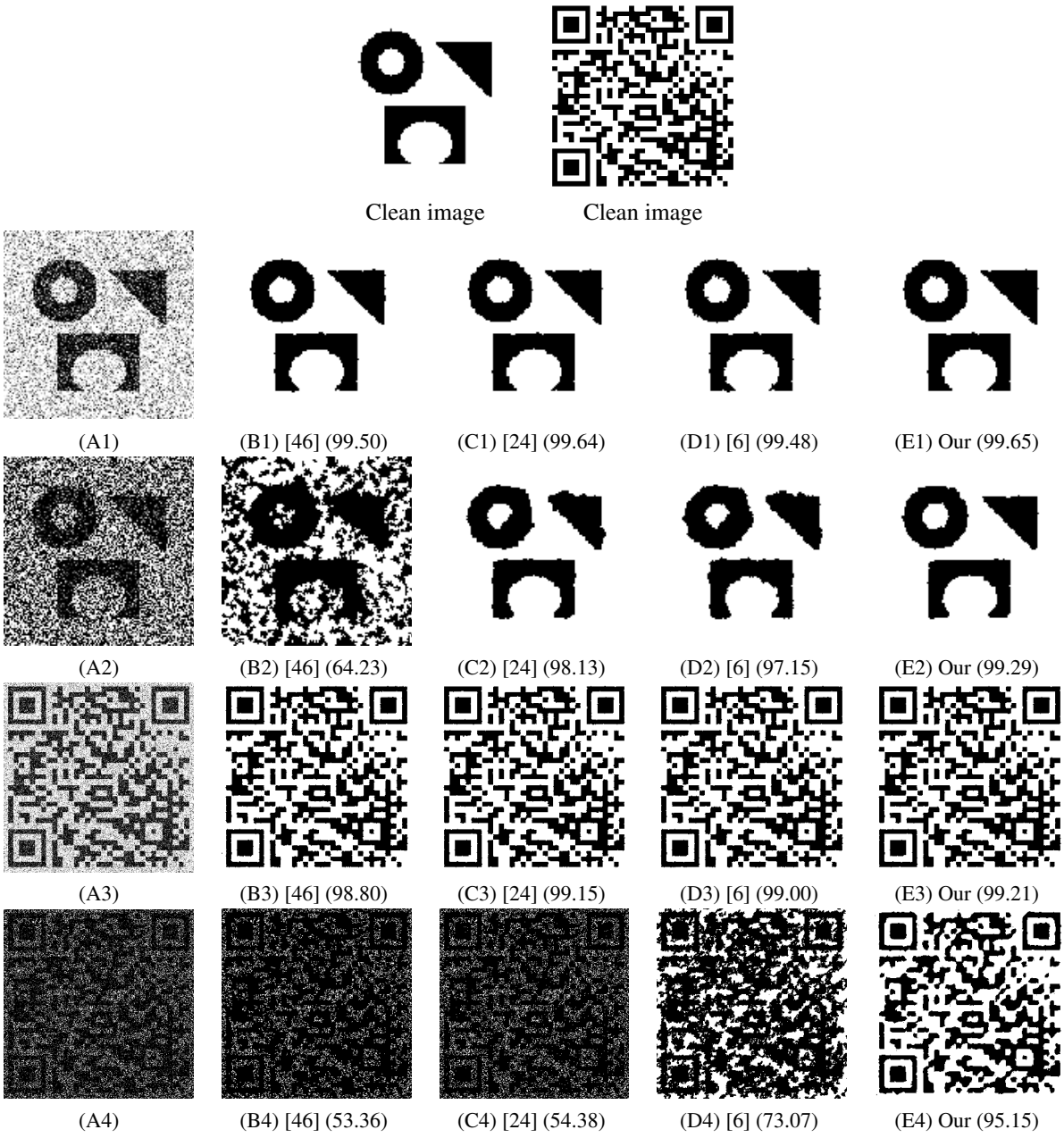


Figure 1: Segmentation of two-phase synthetic images (size: $128 \times 128$ and $333 \times 332$). (A1) and (A3): given noisy images; (A2) and (A4): given noisy images with 40% and 70% information loss, respectively; Columns two to five: results of methods [46, 24, 6] and our method, respectively. Numbers in brackets are the *segmentation accuracy*.

only image (row four of Fig. 1) that can be identified is our result (E4). The regularization parameters used in all methods are provided in Table 1.

*Example 2: multiphase synthetic images.* Two multiphase synthetic images are tested in this example: one is a fourphase image with different shapes inside and the other is a fivephase image including stars with different intensities. For Fig. 2 (A1)–(A4), the variances used to add noise on the fourphase and fivephase images are 0.05 and 0.01, respectively. Fig. 2 (A2) and (A4) are the noisy images with 20% pixels randomly removed. From the results

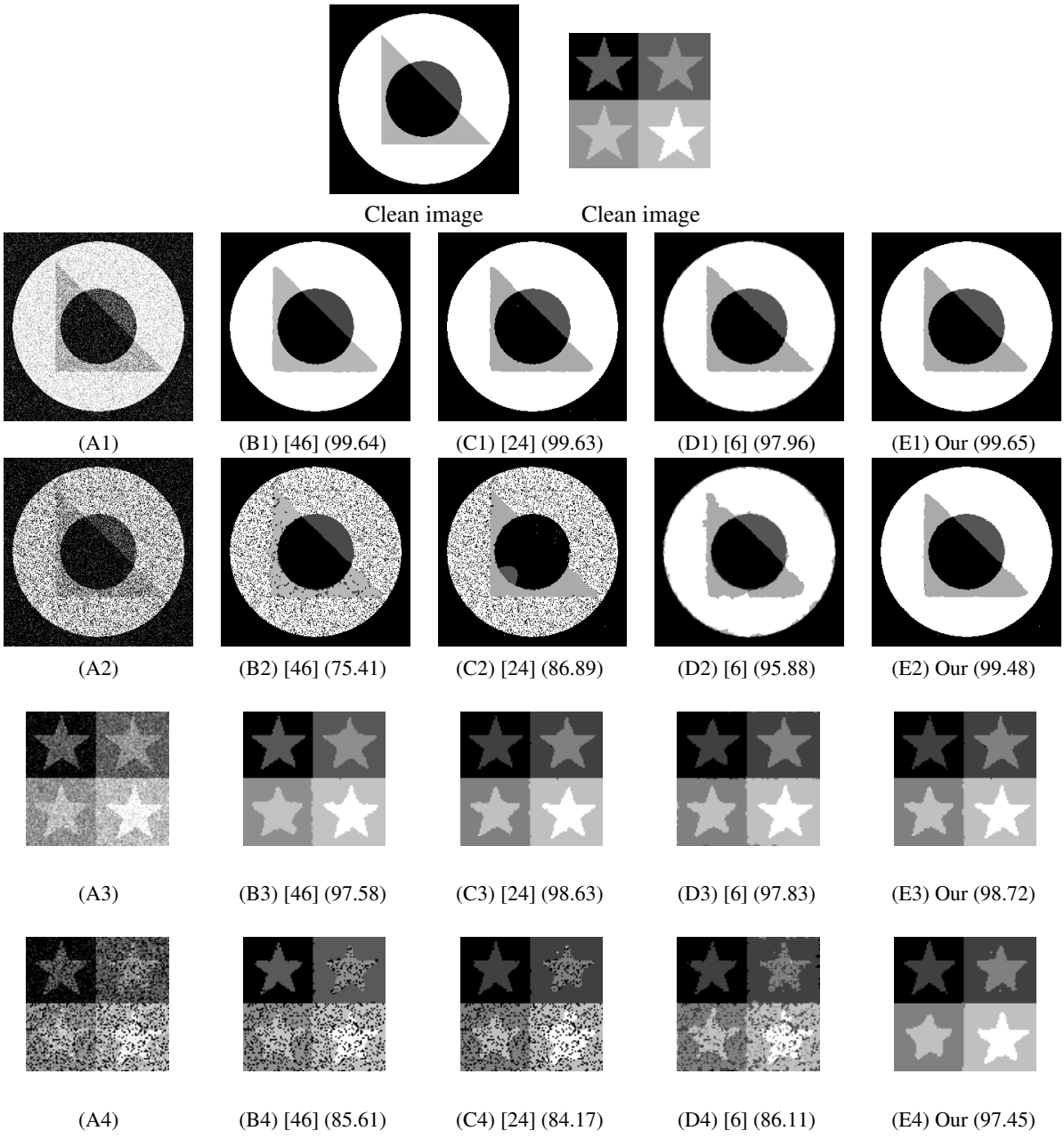


Figure 2: Segmentation of fourphase and fivephase synthetic images (size: $256 \times 256$ and $91 \times 91$). (A1) and (A3): given noisy images; (A2) and (A4): given noisy images with 20% information loss; Columns two to five: results of methods [46, 24, 6] and our method, respectively. Numbers in brackets are the *segmentation accuracy*.

in Fig. 2, we can get very similar conclusions to those obtained in example 1, i.e., all the methods give very good results visually with the noisy images but our method gives the highest *segmentation accuracy*. This illustrates that our model (11) is superior compared with model (6). Moreover, only our method can give good results when dealing with images with information loss.

To illustrate the effectiveness of our method in segmenting blurry images, we first test our method on two synthetic multiphase images used in Fig. 2 but with Gaussion blur and motion blur involved, see Fig. 3. Fig. 3 (A3) is blurred

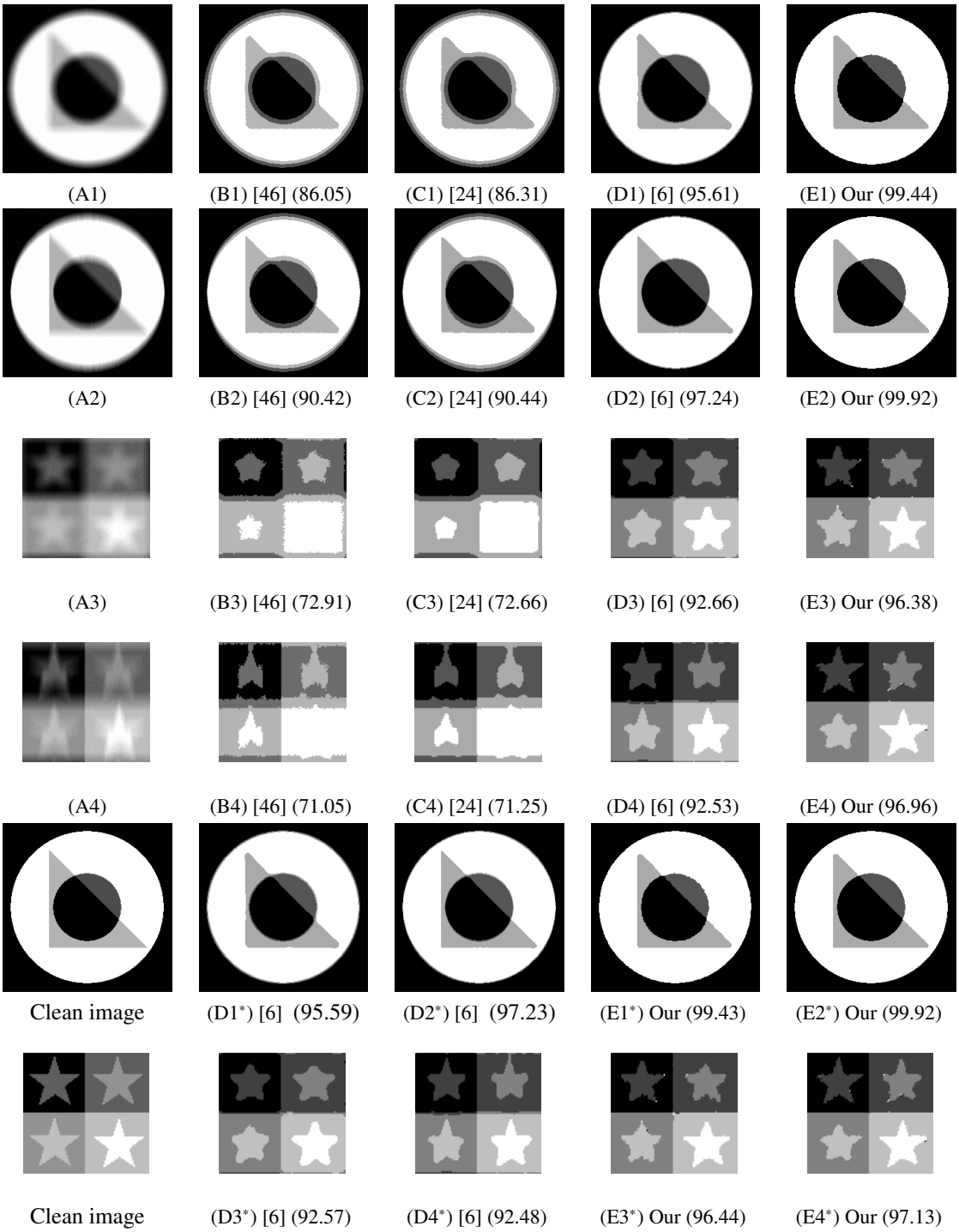


Figure 3: Segmentation of fourphase and fivephase synthetic blurry images (size: $256 \times 256$ and $91 \times 91$). (A1) and (A3): given images with Gaussian blur; (A2) and (A4): given images with motion blur; Columns two to five of rows one to four: results of methods [46, 24, 6] and our method, respectively. Rows five and six: results of method [6] and our method according to degraded operator $\mathcal{A}$, respectively. Numbers in brackets are the *segmentation accuracy*.

by using a Gaussian kernel with size $10 \times 10$ and standard deviation 10. Fig. 3 (D1)–(D4) and (E1)–(E4) are the results using true blur kernels and Fig. 3 (D1$^*$)–(D4$^*$) and (E1$^*$)–(E4$^*$) are the results using degraded blur kernels. From the comparison, we see that only method [6] and our method can produce good results visually. However, the *segmentation accuracy* tells us that our method gives a much higher SA than that of method [6]. Moreover, after comparing the results obtained by true blur kernels and degraded blur kernels, we can see that both method [6] and our method are all quite robust to blur kernels. Again, the *segmentation accuracy* reveals that our method is much better. Methods [46, 24] are not able to segment the blurry images accurately. More precisely, for the results of methods [46, 24] in Fig. 3 (A1) and (A2), the pixels around the boundaries are segmented incorrectly; their results in Fig. 3 (A3) and (A4) show that one star located in the right bottom corner is missed.

*Example 3: real-world images.* Two real-world images are tested: one cameraman and the other a MRI (magnetic resonance imaging) brain image, see Fig. 4. We first test our method on noisy images and images with information loss. In Fig. 4, the variance used for adding noise is 0.01. The conclusions we get are very close to those obtained from testing the methods on synthetic images in examples 1 and 2. From rows one and three of Fig. 4, we see that all methods give very good results in segmenting the two original real-world images. But for the images with information loss, especially Fig. 4 (A4), the results of methods [46, 24] are worse than the results of our method and that of [6], see Fig. 4 (B4), (C4), (D4) and (E4). Moreover, with the results of the latter two methods, we see that our result gives much more detail in the white matter, see Fig. 4 (D4) and (E4).

The ability of our method to segment real-world blurry images is given in Fig. 5. From the comparison, we can see that method [6] and our method give comparably good results, better than the results of methods [46, 24].

Table 1: Regularization parameters used in all methods for examples 1–3

| | | [46] | [24] | [6] | Our | |
|---|---|---|---|---|---|---|
| | | $\lambda$ | $\lambda$ | $\lambda$ | $\lambda$ | $\mu$ |
| Example 1 | Fig. 1 (A1) | 4 | 5 | 5 | 5 | $100 * \lambda$ |
| | Fig. 1 (A2) | 5 | 5 | 2 | 5 | $100 * \lambda$ |
| | Fig. 1 (A3) | 5 | 5 | 4 | 5 | $10 * \lambda$ |
| | Fig. 1 (A4) | 8 | 10 | 4 | 10 | $10 * \lambda$ |
| Example 2 | Fig. 2 (A1) | 6 | 10 | 5 | 10 | $8 * \lambda$ |
| | Fig. 2 (A2) | 10 | 10 | 2 | 11 | $100 * \lambda$ |
| | Fig. 2 (A3) | 40 | 50 | 10 | 60 | $30 * \lambda$ |
| | Fig. 2 (A4) | 30 | 30 | 10 | 50 | $100 * \lambda$ |
| | Fig. 3 (A1) | 50 | 50 | 50 | 10 | $1000 * \lambda$ |
| | Fig. 3 (A2) | 50 | 50 | 50 | 10 | $1000 * \lambda$ |
| | Fig. 3 (A3) | 100 | 200 | 70 | 40 | $1000 * \lambda$ |
| | Fig. 3 (A4) | 100 | 200 | 70 | 40 | $1000 * \lambda$ |
| Example 3 | Fig. 4 (A1) | 8 | 10 | 5 | 10 | $100 * \lambda$ |
| | Fig. 4 (A2) | 6 | 8 | 3 | 10 | $100 * \lambda$ |
| | Fig. 4 (A3) | 30 | 30 | 10 | 50 | $100 * \lambda$ |
| | Fig. 4 (A4) | 10 | 10 | 3 | 30 | $80 * \lambda$ |
| | Fig. 5 (A1) | 100 | 100 | 100 | 10 | $1000 * \lambda$ |
| | Fig. 5 (A2) | 100 | 100 | 100 | 10 | $1000 * \lambda$ |
| | Fig. 5 (A3) | 100 | 100 | 200 | 20 | $1000 * \lambda$ |
| | Fig. 5 (A4) | 100 | 100 | 50 | 20 | $1000 * \lambda$ |

### *5.2. Color image segmentation*

*Example 4: two-phase rose image.* Fig. 6 (A1)–(A4) give the original rose image, and the images corrupted by the random partial removal of information, Gaussian blur and motion blur, respectively. Columns two and three of Fig. 6 give the results of the extended method [24] and our method, respectively. From the comparison, we see that both of the two methods can give good results based on the original image (the first row of Fig. 6). The second row of Fig. 6

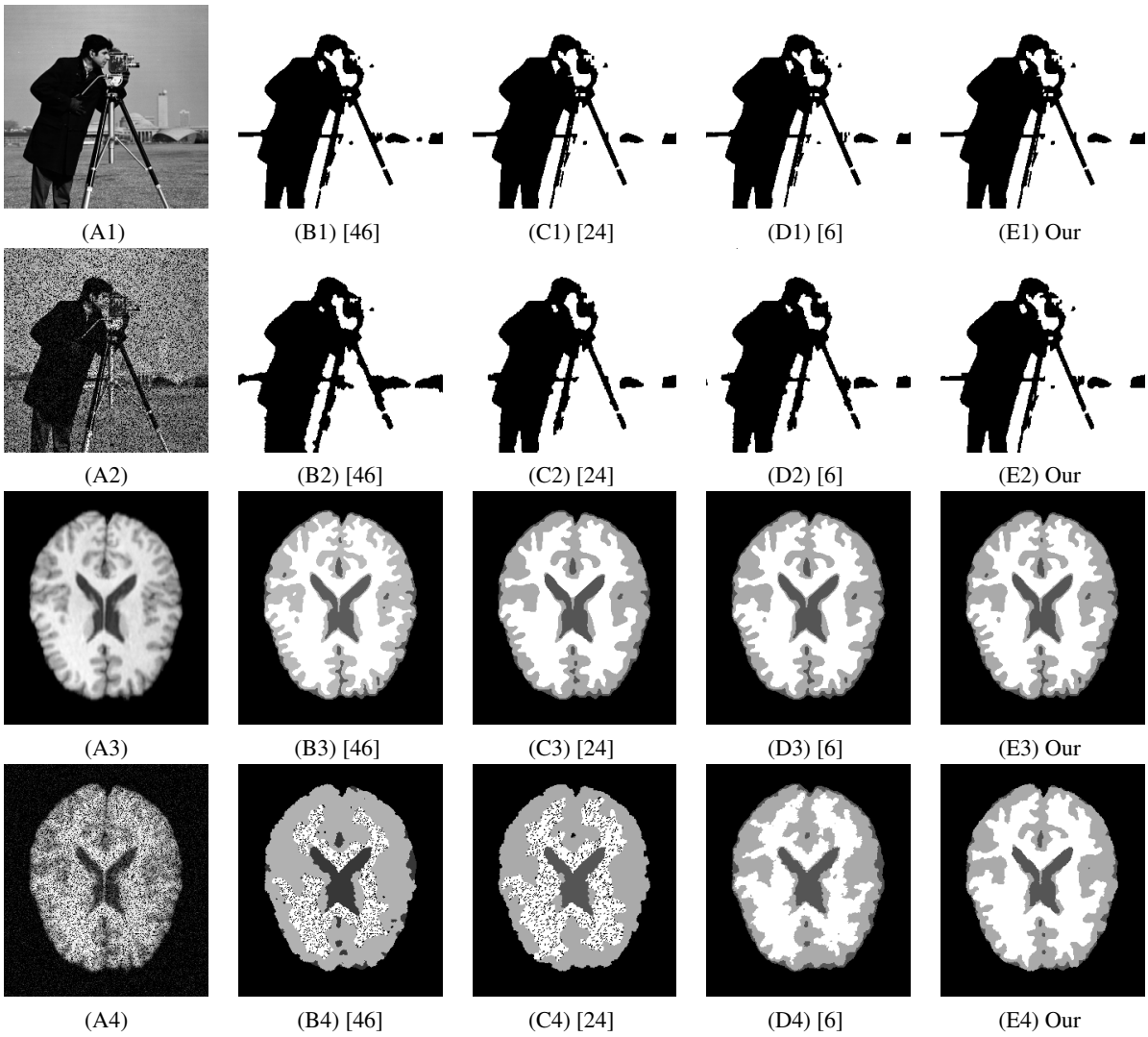


Figure 4: Segmentation of real-world images: cameraman and MRI brain (size: $256 \times 256$ and $319 \times 256$). (A1) and (A3): given images; (A2) and (A4): given noisy images with 20% information loss; Columns two to five: results of methods [46, 24, 6] and our method, respectively.

shows that the boundary of the result of the extended method [24] is coarse compared with our result; from rows three and four of Fig. 6, we can see that the results of the extended method [24] are over smoothed. Hence, our model can produce better results when segmenting blurry color images, while the extended method [24] cannot. Moreover, we can also see that the results of our method for the corrupted images are as good as the result for the original image, see the third column of Fig. 6. This demonstrates the ability of our method in segmenting images with information loss or blur. The regularization parameters used in the extended method [24] and our method are given in Table 2.

*Example 5: multiphase images.* Two more multiphase color images are tested, i.e. three phases crown image and four phases flowers image, see Fig. 7 and Fig. 8, respectively. Obviously, the extended method [24] fails to segment images with information loss, see Fig. 7 (B2) and Fig. 8 (B2). Moreover, from (C2) and (D2) of Fig. 7 and Fig. 8, we can see that the extended method [24] gives over smoothed results for blurry color images. On the contrary, from the third column of Fig. 7 and Fig. 8, we see that all our results are very good visually. In addition, the results of our method on the corrupted images are still as good as the results on the original crown and flowers images.

*Example 6: real-world color image with ground truth.* In order to test our method on real world images not only

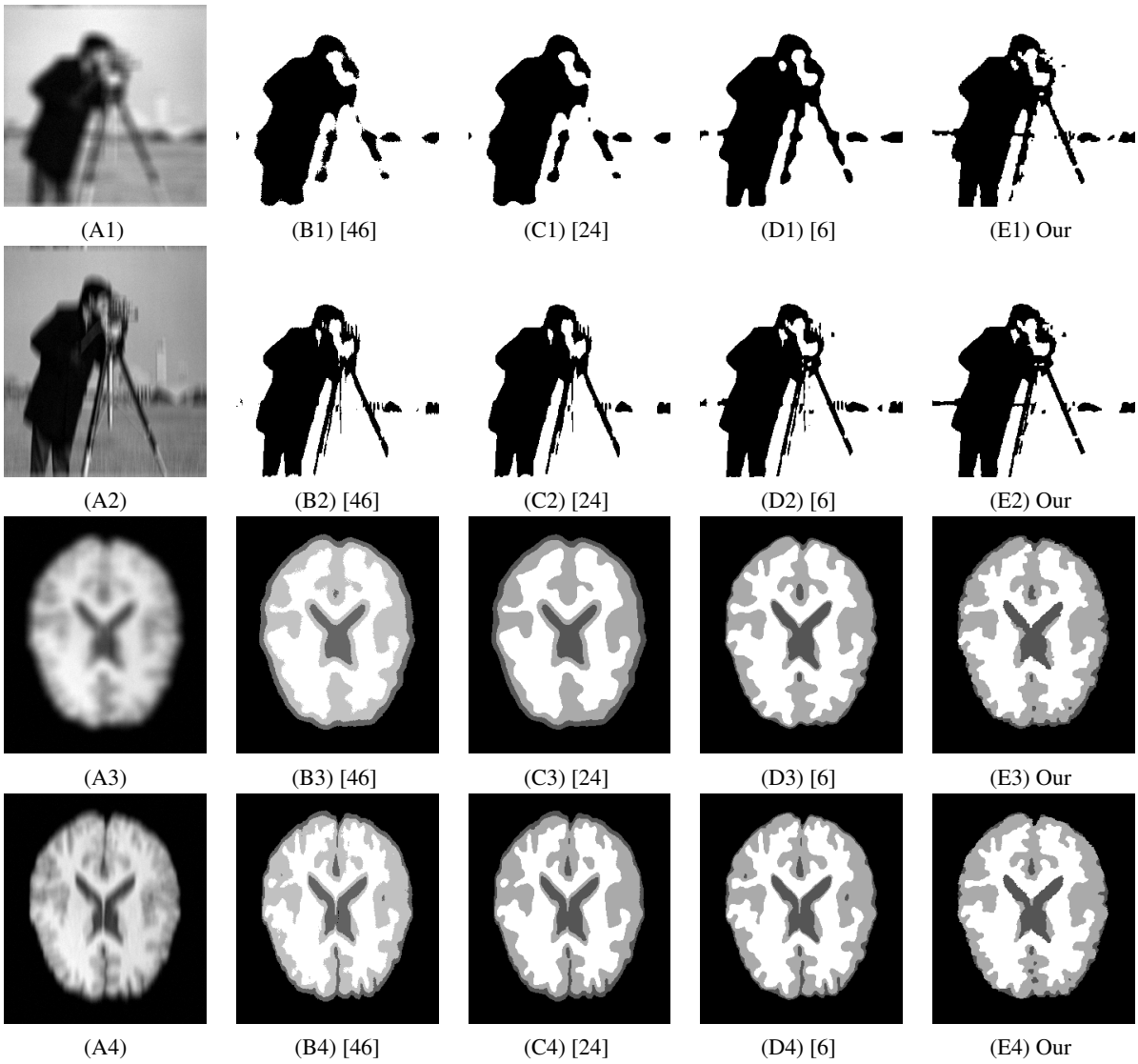


Figure 5: Segmentation of real-world blurry images: camera man and MRI brain (size: 256×256 and 319×256). (A1) and (A3): given images with Gaussian blur; (A2) and (A4): given images with motion blur; Columns two to five: results of methods [46, 24, 6] and our method, respectively.

qualitatively but also quantitatively, in this example the test image is chosen from the Berkeley Segmentation Dataset [48] with a manually segmented ground truth provided by an expert, see Fig. 9. The given degraded images are generated in the same way as in examples 5 and 6. In addition to giving very similar conclusions as those in examples 5 and 6, the *segmentation accuracy* also validates the better performance of our method quantitatively. Moreover, after comparing our results obtained by using true and degraded blurring kernels, we can see that our method is indeed robust to the blurring kernels, see Fig. 9 (C3)–(C4) and (C3$^*$)–(C4$^*$).

## 6. Conclusions

In this paper, we proposed a new multiphase segmentation model by combining image restoration approaches with a variational image segmentation model. Utilizing aspects of image restoration, the proposed segmentation model is very effective and robust in tackling noisy images, blurry images, and images with information loss. In particular, the

Table 2: Regularization parameters used in the extended [24] and our method for examples 4–6

| | | Extended [24] | Our | |
|---|---|---|---|---|
| | | $\lambda$ | $\lambda$ | $\mu$ |
| Example 4 | Fig. 6 (A1) | 4 | 4 | $100 * \lambda$ |
| | Fig. 6 (A2) | 4 | 4 | $100 * \lambda$ |
| | Fig. 6 (A3) | 10 | 3 | $1000 * \lambda$ |
| | Fig. 6 (A4) | 6 | 2 | $100 * \lambda$ |
| Example 5 | Fig. 7 (A1) | 100 | 100 | $100 * \lambda$ |
| | Fig. 7 (A2) | 3 | 20 | $100 * \lambda$ |
| | Fig. 7 (A3) | 20 | 10 | $1000 * \lambda$ |
| | Fig. 7 (A4) | 20 | 20 | $1000 * \lambda$ |
| | Fig. 8 (A1) | 50 | 100 | $100 * \lambda$ |
| | Fig. 8 (A2) | 10 | 50 | $100 * \lambda$ |
| | Fig. 8 (A3) | 80 | 20 | $1000 * \lambda$ |
| | Fig. 8 (A4) | 80 | 10 | $1000 * \lambda$ |
| Example 6 | Fig. 9 (A1) | 2 | 2 | $100 * \lambda$ |
| | Fig. 9 (A2) | 5 | 4 | $100 * \lambda$ |
| | Fig. 9 (A3) | 2 | 1 | $1000 * \lambda$ |
| | Fig. 9 (A4) | 2 | 2 | $1000 * \lambda$ |

piecewise constant Mumford-Shah model was extended using our strategy so that it can process blurry images. Moreover, our model can also be extended to process vector-valued images, for example color images. It can be solved efficiently using the AM algorithm, and we prove its convergence property under mild conditions. Experiments on many kinds of synthetic and real-world images demonstrate that our method gives better segmentation results in terms of quality and quantity in comparison with other state-of-the-art segmentation methods, especially on blurry images and images with information loss. In future work, we will test our model on images corrupted by other types of noise, for example Poisson noise and impulsive noise. We will also use much more complicated and effective terms from image segmentation and image restoration in our method.

**Acknowledgements:** The author would like to thank professor Gabriele Steidl for fruitful discussions and invaluable comments. Thanks to Will Simonson for his instructive comments and proofreading on the paper, and we are grateful for the grant support from Wellcome Trust. The authors also would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper.

## Appendix

*Proof of Theorem 1*

*Proof.* Fix $c_i = c_i^*$ and $u_i = u_i^*$. Note that $L^2(\Omega)$ is a reflective Banach space (see Chapter II in [19]), and $E(g, c_i^*, u_i^*)$ in (7) is convex and lower semicontinuous. Using Proposition 1.2 in [19], for the existence of $g$, we just need to prove

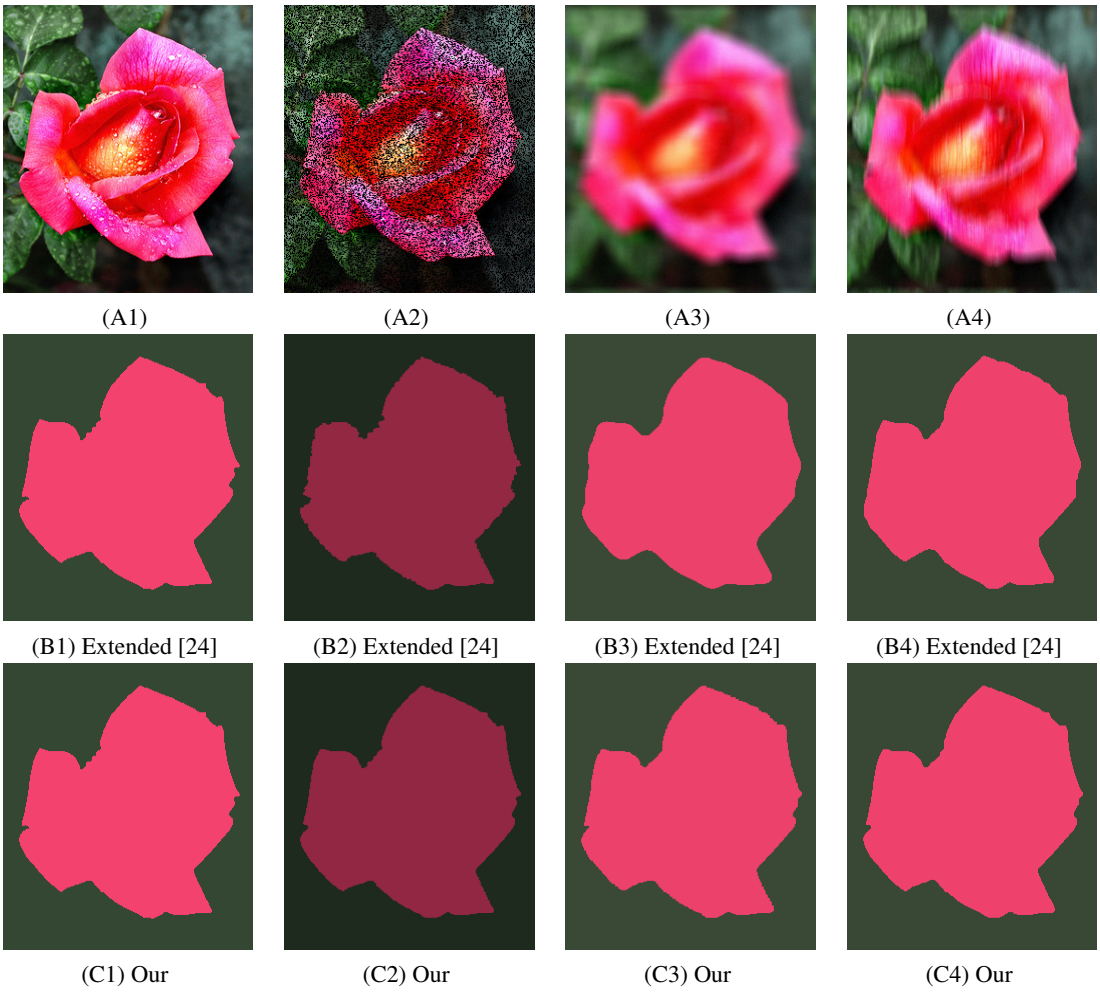


Figure 6: Segmentation of rose color image (size: $303 \times 250 \times 3$). Row one: given image, and given images corrupted by 40% information loss, Gaussian blur and motion blur, respectively. Rows two to three: results of the extended method [24] and our method, respectively.

that $E(g, c_i^*, u_i^*)$ is coercive over $L^2(\Omega)$. Note that $u_i(x) \in [0, 1]$, the coercive of $E(g, c_i^*, u_i^*)$ can be given by

$$\begin{aligned}
\|g\|_2 &= \|\sum_{i=1}^{K}(g - c_i)u_i + \sum_{i=1}^{K} c_i u_i\|_2 \\
&\leq \|\sum_{i=1}^{K}(g - c_i)u_i\|_2 + \|\sum_{i=1}^{K} c_i u_i\|_2 \\
&\leq \|\sum_{i=1}^{K}|g - c_i|u_i^{\frac{1}{2}}\|_2 + \|\sum_{i=1}^{K} c_i u_i\|_2 \\
&\leq \sqrt{K}\sqrt{\sum_{i=1}^{K}\int_\Omega (g - c_i)^2 u_i dx} + \|\sum_{i=1}^{K} c_i u_i\|_2 \\
&\leq \frac{\sqrt{K}}{\sqrt{\lambda}}\sqrt{E(g, c_i^*, u_i^*)} + \|\sum_{i=1}^{K} c_i u_i\|_2.
\end{aligned}$$

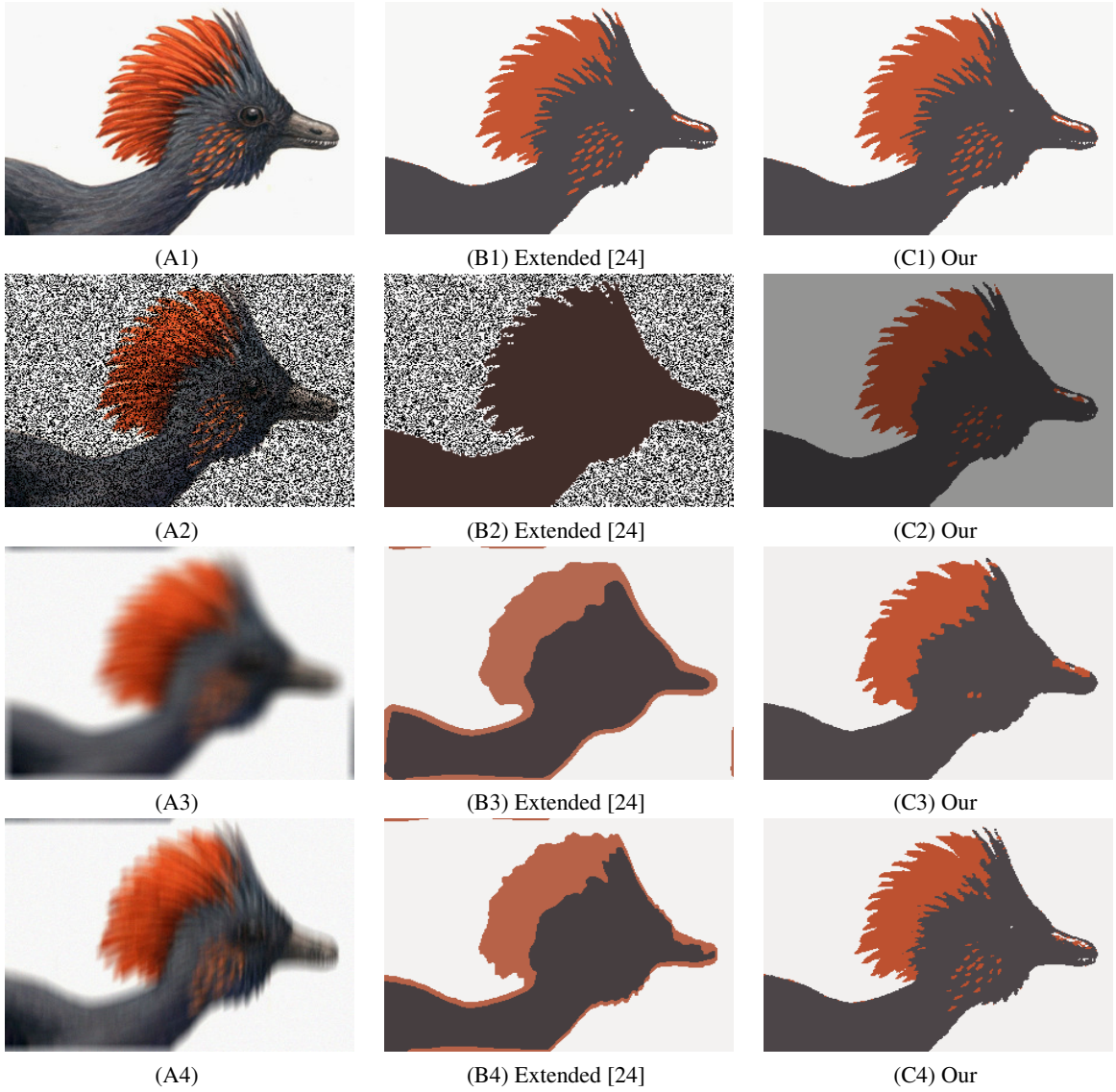


Figure 7: Segmentation of crown color image (size: $225 \times 300 \times 3$). Column one: given image, and given images corrupted by 40% information loss, Gaussian blur and motion blur, respectively. Columns two to three: results of the extended method [24] and our method, respectively.

Moreover, since the middle term of energy (7) is quadratic with respect to $g$, hence energy (7) is strictly convex, it has unique minimizer $g$ for fixed $c_i$ and $u_i$. □

*Proof of Theorem 2*

*Proof.* From (18), we have

$$\begin{aligned} E(x^{(k)}, y^{(k)}, z^{(k)}) &\leq E(x^{(k-1)}, y^{(k)}, z^{(k)}), \\ E(x^{(k)}, y^{(k)}, z^{(k+1)}) &\leq E(x^{(k)}, y^{(k)}, z^{(k)}) \end{aligned}$$

and

$$E(x^{(k)}, y^{(k+1)}, z^{(k+1)}) \leq E(x^{(k)}, y^{(k)}, z^{(k+1)}).$$

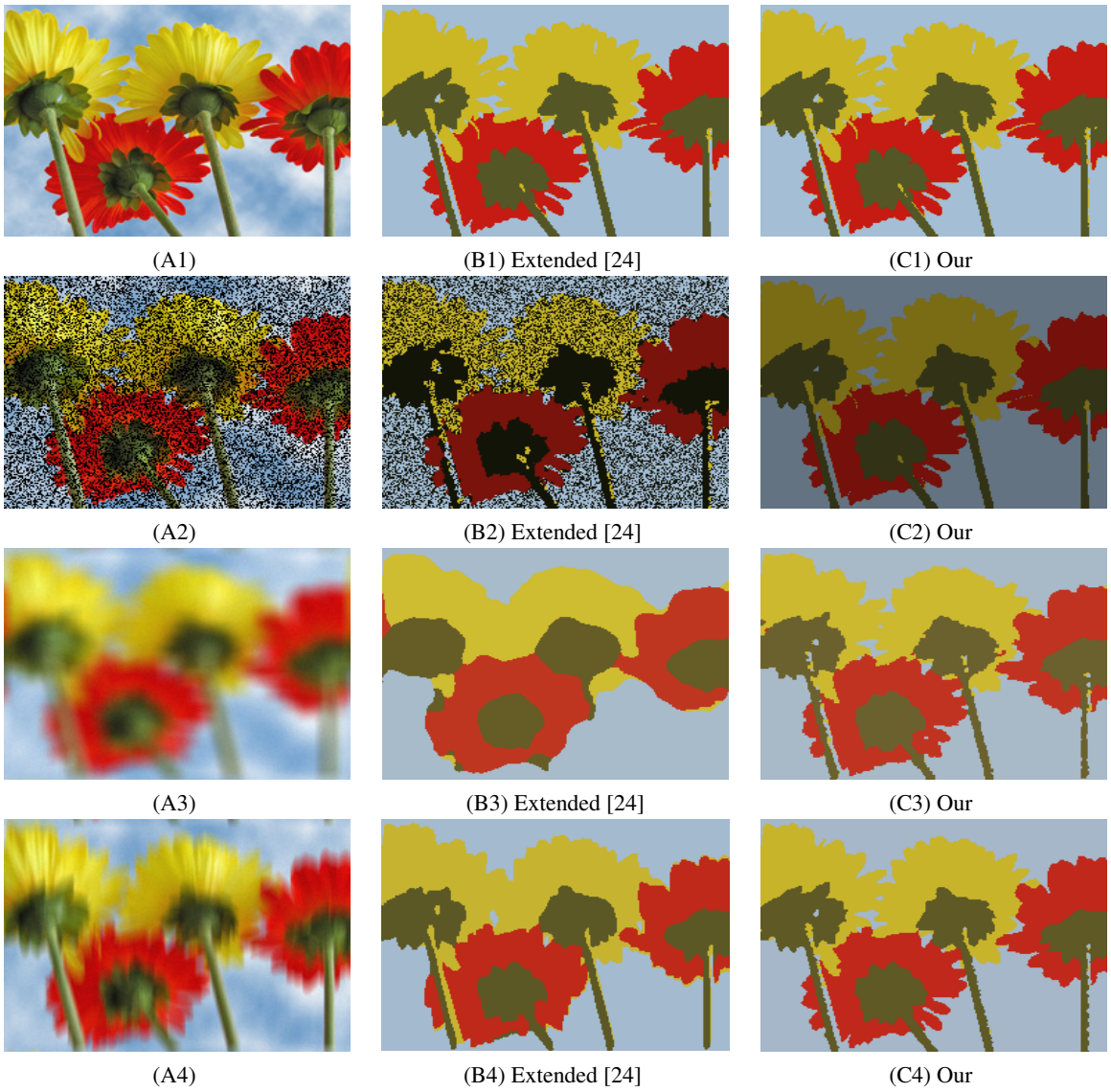


Figure 8: Segmentation of flowers color image (size: $188 \times 250 \times 3$). Column one: given image, and given images corrupted by 40% information loss, Gaussian blur and motion blur, respectively. Columns two to three: results of the extended method [24] and our method, respectively.

Then, it can be verified directly that

$$\begin{aligned} E(x^{(k+1)}, y^{(k+1)}, z^{(k+1)}) &\leq E(x^{(k)}, y^{(k+1)}, z^{(k+1)}) \\ &\leq E(x^{(k)}, y^{(k)}, z^{(k+1)}) \\ &\leq E(x^{(k)}, y^{(k)}, z^{(k)}) \\ &\leq E(x^{(k-1)}, y^{(k)}, z^{(k)}) \\ &\leq E(x^{(k-1)}, y^{(k-1)}, z^{(k)}). \end{aligned}$$

Hence, since $E(\cdot, \cdot, \cdot)$ is bounded from below, the sequence $\{E(x^{(k)}, y^{(k)}, z^{(k)})\}_{k \in \mathbb{N}}$ converges monotonically. □

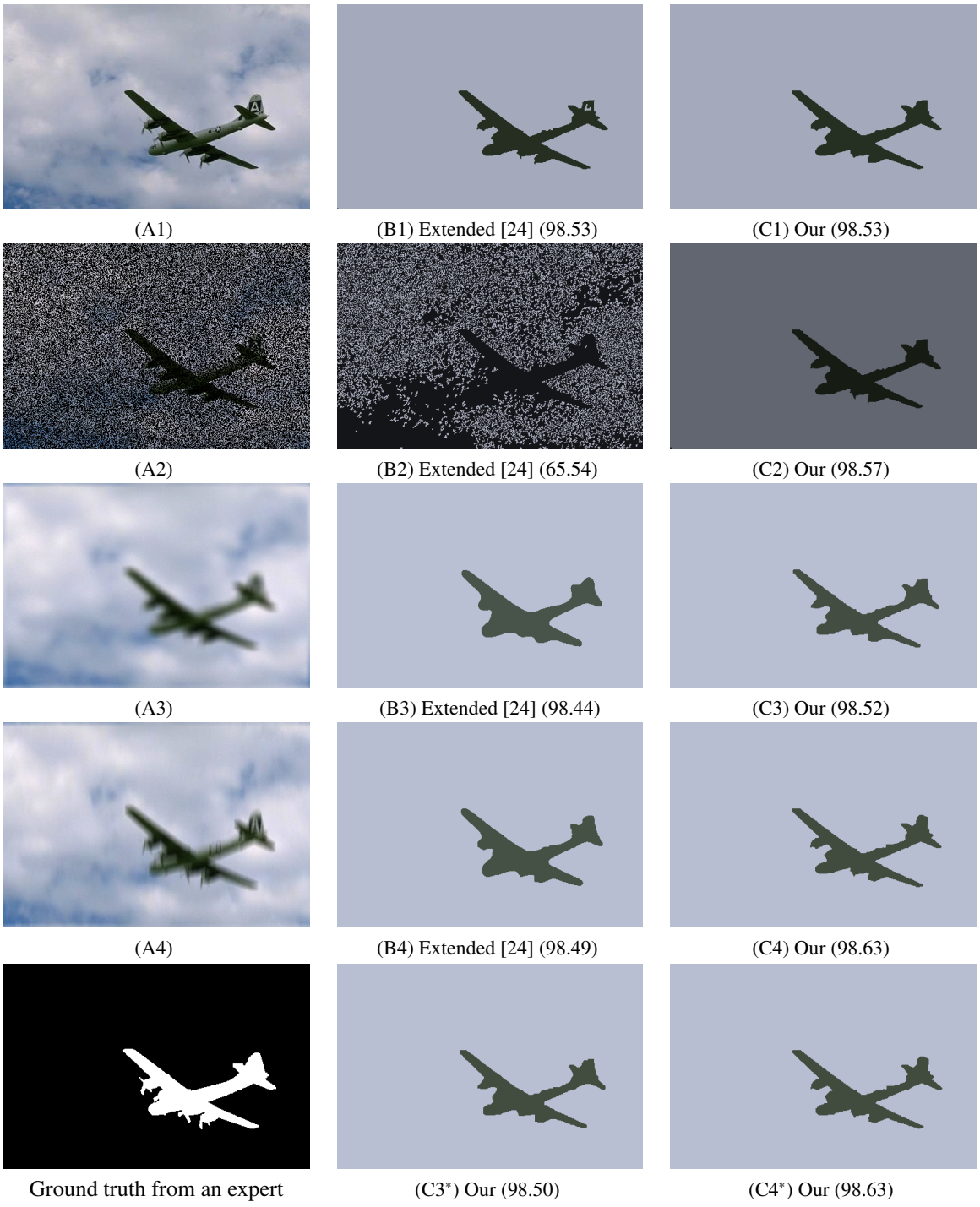


Figure 9: Segmentation of kangaroo color image (size: $321 \times 481 \times 3$). Column one: given image, and given images corrupted respectively by 60% information loss, Gaussian blur and motion blur, and ground truth. Columns two to three of rows one to four: results of the extended method [24] and our method, respectively. Row five: results of our method according to degraded operator $\mathcal{A}$. Numbers in brackets are the *segmentation accuracy*.

*6.1. Proof of Theorem 3*

*Proof.* Using (18) and the idea of [15, Theorem 5.5], for each $i$

$$E(x^{(k_i)}, y^{(k_i)}, z^{(k_i)}) \leq E(x, y^{(k_i)}, z^{(k_i)}) \quad \forall x \in X.$$

By the continuity of $E(\cdot,\cdot,\cdot)$, this gives, as $i \to \infty$,

$$E(x^*, y^*, z^*) \leq E(x, y^*, z^*) \quad \forall x \in X.$$

On the other hand, for each $i$, note that $k_{i-1} \leq k_i - 1$. From Theorem 2, we have for $\forall y \in Y$,

$$\begin{aligned} E(x^{(k_i)}, y^{(k_i)}, z^{(k_i)}) &\leq E(x^{(k_i-1)}, y^{(k_i)}, z^{(k_i)}) \\ &\leq E(x^{(k_{i-1})}, y^{(k_{i-1}+1)}, z^{(k_{i-1}+1)}) \\ &\leq E(x^{(k_{i-1})}, y, z^{(k_{i-1}+1)}), \end{aligned}$$

and $\forall z \in Z$,

$$\begin{aligned} E(x^{(k_i)}, y^{(k_i)}, z^{(k_i)}) &\leq E(x^{(k_i-1)}, y^{(k_i)}, z^{(k_i)}) \\ &\leq E(x^{(k_i-1)}, y^{(k_i-1)}, z^{(k_i)}) \\ &\leq E(x^{(k_{i-1})}, y^{(k_{i-1})}, z^{(k_{i-1}+1)}) \\ &\leq E(x^{(k_{i-1})}, y^{(k_{i-1})}, z). \end{aligned}$$

Coupled with the continuity of $E(\cdot,\cdot,\cdot)$, as $i \to \infty$, we have,

$$E(x^*, y^*, z^*) \leq E(x^*, y, z^*) \quad \text{and} \quad E(x^*, y^*, z^*) \leq E(x^*, y^*, z)$$

for $\forall y \in Y$ and $\forall z \in Z$ respectively. □

*Proof of Theorem 4*

*Proof.* For model (11), because all of its three terms are continuous and nonnegative, $E(\cdot,\cdot,\cdot)$ is continuous and nonnegative. If $(u^{(k)}, g^{(k)}, c^{(k)}) \to (u^*, g^*, c^*)$, as $k \to \infty$, using Theorem 3, we have $(u^*, g^*, c^*) \in \mathcal{O}$. Obviously, all the components of $u^{(k)}$ are in $[0, 1]$, hence $u^{(k)}$ is bounded. From (14), $c^{(k)}$ is bounded since it is just the convex combination of $g$. From (13), we have

$$\begin{aligned} \|g\|_2 &\leq \|(\mu\mathcal{A}^T\mathcal{A} + \lambda\mathcal{I})^{-1}\|_2 \|(\mu\mathcal{A}^T f + \lambda \sum_{i=1}^{K} c_i u_i)\omega\|_2 \\ &\leq (\mu\|\mathcal{A}^T\|_2\|f\|_2 + \lambda \sum_{i=1}^{K} c_i\|u_i\|_2)\|\omega\|_2/\lambda, \end{aligned}$$

hence $g^{(k)}$ is also bounded. Therefore $(u^{(k)}, g^{(k)}, c^{(k)})_{k\in\mathbb{N}}$ must contain convergent subsequence. Using Theorem 3, we can get that any of these subsequences converges to a partial minimizer of model (11). □

## References


[1] G. Aubert and J. Aujol, "A variational approach to removing multiplicative noise," *SIAM J. Appl. Math.*, vol. 68, no. 4, pp. 925–946, Jan. 2008.

[2] L. Bar, T. Chan, G. Chung, M. Jung, N. Kiryati, R. Mohieddine, N. Sochen, and L.A. Vese, "Mumford and Shah model and its applications to image segmentation and image restoration," *Handbook of Mathematical Imaging, Springer*, pp. 1095–1157, 2011.

[3] J. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy $c$-means clustering algorithm," *Comput. Geosci.*, vol. 10, no. 2-3, pp. 191–203, 1984.

[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. and Trends Mach. Learning*, vol. 3, no. 1, pp. 1–122, Jan. 2010.

[5] J. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Appl. Comput. Harmon. A.*, vol. 24, no. 2, pp. 131–149, Mar. 2008.
[6] X. Cai, R. Chan, and T. Zeng, "A two-stage image segmentation method using a convex variant of the Mumford-Shah model and thresholding," *SIAM J. Imaging Sci.*, vol. 6, no. 1, pp. 368–390, Feb. 2013.
[7] X. Cai and G. Steidl, "Multiclass segmentation by iterated ROF thresholding," *EMMCVPR, LNCS, Springer*, pp. 237–250, 2013.
[8] A. Chambolle, "Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations," *SIAM J. Appl. Math.*, vol. 55, no. 3, pp. 827–863, 1995.
[9] A. Chambolle, "Finite differences discretization of the Mumford-Shah functional," *RAIRO Math. Model. Numer. Anal.*, vol. 33, no. 2, pp. 261–288, 1999.
[10] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. math. Imaging Vis.*, vol. 40, no. 1, pp. 120–145, May 2011.
[11] R. Chan, C. Hu, and M. Nikolova, "An iterative procedure for removing random-valued impulse noise," *IEEE Signal Process. Lett.*, vol. 11, no. 12, pp. 921–924, Dec. 2004.
[12] R. Chan, C. Hu, and M. Nikolova, "Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1479–1485, Oct. 2005.
[13] T. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM J. Appl. Math*, vol. 66, no. 5, pp. 1632–1648, Sep. 2006.
[14] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 22, no. 2, pp. 503–516, 2000.
[15] T. Chan and J. Shen, "Image processing and analysis: variational, PDE, wavelet, and stochastic methods," *SIAM*, 2005.
[16] T. Chan and L.A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
[17] I. Csiszár, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, vol. 19, no. 4, pp. 2032–2066, 1991.
[18] I. Csiszár and G. Tusnady, "Information geometry and alternating minimization procedures," *Statistics & Decisions, Supplement Issu*, vol. 1, pp. 205–237, 1984.
[19] I. Ekeland and R. Temam, "Convex analysis and variational problems," *SIAM, Philadelphia*, 1999.
[20] W. Fleming, W. Rishel, and R. Rishel, "An integral formula for total gradient variation," *Arch. Math.*, vol. 11, no. 1, pp. 218–222, 1960.
[21] E. Giusti, "Minimal Surfaces and Functions of Bounded Variation," *Monogr. Math.*, vol. 80, 1984.
[22] M. Gobbino, "Finite difference approximation of the Mumford-Shah functional," *Comm. Pure Appl. Math.*, vol. 51, no. 2, pp. 197–228, Feb. 1998.
[23] T. Goldstein and S. Osher, "The split Bregman algorithm for L1 regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–343, Apr. 2009.
[24] Y. He, B. Shafei, M. Y. Hussaini, J. Ma, and G. Steidl, "A new fuzzy c-means method with total variation regularization for segmentation of images with noisy and incomplete data," *Pattern Recognition*, vol. 45, no. 9, pp. 3436-3471, Sep. 2012.
[25] Y. Huang, M. Ng, and Y. Wen, "Fast image restoration methods for impulse and Gaussian noise removal," *IEEE Signal Process. Lett.*, vol. 16, no. 6, pp. 457–460, Jun. 2009.
[26] G. Koepfler, C. Lopez, and J. Morel, "A Multiscale Algorithm for Image Segmentation by Variational Approach," *SIAM J. Numer. Anal.*, vol. 31, no. 1, pp. 282–299, Feb. 1994.
[27] T. Le, R. Chartrand and T. Asaki, "A variational approach to reconstructing images corrupted by Poisson noise," *J. Math. Imaging Vis.*, vol. 27, no. 3, pp. 257–263, Apr. 2007.
[28] F. Li, M. Ng, T. Zeng, and C. Shen, "A multiphase image segmentation method based on fuzzy region competition," *SIAM J. Imaging Sci.*, vol. 3, no. 3, pp. 277–299, Jul. 2010.
[29] M. Lysaker, A. Lundervold, and X. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Trans. on Image Process.*, vol. 12, no. 12, pp. 1579–1590, Dec. 2003.
[30] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Comm. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, Jul. 1989.
[31] M. Nikolova, "A variational approach to remove outliers and impulse noise," *J. Math. Imag. Vis.*, vol. 20, no. 1-2, pp. 99–120, Jan. 2004.
[32] G. Paul, J. Cardinale, and I. Sbalzarini, "Coupling image restoration and segmentation: a generalized linear model/bregman perspective," *Int. J. Comput. Vis.*, vol. 104, no. 1, pp. 69–93, Aug. 2013.
[33] T. Pock, D. Cremers, A. Chambolle, and H. Bischof, "A convex relaxation approach for computing minimal partitions," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 810–817, 2009.
[34] R. Potts, "Some generalized order-disorder transformations," *Proceedings of the Cambridge Philosophical Society 48*, pp. 106–109, 1952.
[35] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *SIGGRAPH*, 2008.
[36] L. Rudin and S. Osher, "Total variation based image restoration with free local constraints," *Proc. 1st IEEE ICIP*, vol. 1, pp. 31–35, 1994.
[37] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D.*, vol. 60, no. 1-4, pp. 259–268, Nov. 1992.
[38] S. Setzer and G. Steidl, "Variational methods with higher order derivatives in image processing," *Approximation XII, Nashboro Press, Brentwood*, pp.360–386, 2008.
[39] S. Setzer, G. Steidl, and T. Teuber, "Deblurring Poissonian images by split Bregman techniques," *J. Vis. Commun. Image R.*, vol. 21, no. 3, pp. 193–199, Apr. 2010.
[40] G. Steidl and T. Teuber, "Removing multiplicative noise by Douglas-Rachford splitting methods," *J. Math. Imaging and Vis.*, vol. 36, no. 2, pp. 168–184, Feb. 2010.
[41] G. Steidl, J. Weickert, T. Brox, P. Mrazek, and M. Welk, "On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularization, and SIDEs," *SIAM J. Numer. Anal.*, vol. 42, no. 2, pp. 686–713, May 2004.
[42] A. Tikhonov, "Regularization of ill-posed problems," *Dokl. Akad. Nauk SSSR 153*, vol. 1, pp. 49–52, 1963.

[43] M. Vairy and Y. Venkatesh, "Deblurring Gaussian blur using a wavelet array transform," *Pattern Recognition*, vol. 28, no. 7, pp. 965–976, 1995.
[44] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *Int. J. of Computer Vision*, vol. 50, no. 3, pp. 271–293, Dec. 2002.
[45] Y. You and M. Kaveh, "Fourth-order partial differential equation for noise removal," *IEEE Trans. on Image Process.*, vol. 9, no. 10, pp. 1723–1730, Oct. 2000.
[46] J. Yuan, E. Bae, X. Tai, and Y. Boycov, "A continuous max-flow approach to potts model," *ECCV*, 2010.
[47] S. Yun and H. Woo, "Linearized proximal alternating minimization algorithm for motion deblurring by nonlocal regularization," *Pattern Recognition*, vol. 44, no. 6, pp. 1312–1326, 2011.
[48] http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/, *The Berkeley Segmentation Dataset and Benchmark*.