

# A Novel Euler's Elastica based Segmentation Approach for Noisy Images using the Progressive Hedging Algorithm

Lu Tan,\* Ling Li<sup>†</sup>, Wanquan Liu<sup>‡</sup>, Jie Sun<sup>§</sup>, Min Zhang<sup>¶</sup>

August 13, 2019

## Abstract

Euler's elastica based unsupervised segmentation models have strong capability of completing the missing boundaries for existing objects in a clean image, but they are not working well for noisy images. This paper aims to establish a Euler's elastica based approach that can properly deal with the random noises to improve the segmentation performance for noisy images. The corresponding formulation of stochastic optimization is solved via the progressive hedging algorithm (PHA) and the description of each individual scenario is obtained by the alternating direction method of multipliers (ADMM). Technically, all the sub problems derived from the framework of PHA can be solved by using the curvature weighted approach and the convex relaxation method. Then an alternating optimization strategy is applied by using some powerful accelerating techniques including the fast Fourier transform (FFT) and generalized soft threshold formulas. Extensive experiments have been conducted on both synthetic and real images, which displayed significant gains of the proposed segmentation models and demonstrated the advantages of the developed algorithms.

**Keywords** Euler's elastic energy, stochastic noises, progressive hedging algorithm (PHA), alternating direction method of multipliers (ADMM), curvature weighted approach.

---

\*School of EECMS, Curtin University, Perth, Australia. Email: lu.tan1@postgrad.curtin.edu.au

<sup>†</sup>School of EECMS, Curtin University, Perth, Australia. Email: l.li@curtin.edu.au

<sup>‡</sup>School of EECMS, Curtin University, Perth, Australia. Email: w.liu@curtin.edu.au

<sup>§</sup>School of EECMS, Curtin University, Perth, Australia. Email: jie.sun@curtin.edu.au

<sup>¶</sup>State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China; University of Chinese Academy of Sciences, Beijing, China. Email: zhangmin1206@ms.xjb.ac.cn

# 1 Introduction

In image segmentation, the Mumford-Shah model [1] is regarded as one of the most significant region-based models and has been applied to many applications. In 2001, the two-phase Chan-Vese (CV) model [2] was proposed to detect objects in a given image. With the increasing complexity of images, the multiphase segmentation models [3, 4] were proposed and these models mainly represent different regions by using the level set functions [5, 12, 14, 16]. In order to reduce the number of level set functions, Chan et al. proposed a multiphase segmentation model [6], which is a generalization of CV model.

Some specific segmentation models [7, 8, 9] were also established subsequently according to different noise distribution. They obtained the characteristic information contained in images by estimating corresponding parameters. When dealing noisy images, it is known that many segmentation problems need a suitable noise model, e.g., synthetic aperture radar, positron emission tomography, electron micrograph or medical ultrasound imaging, etc. Especially when the data were collected with poor statistics, it is necessary to consider the influence of the noise probability distribution in segmentation implementation.

Recently, authors in [10, 11, 12] made some progress in achieving illusory contour recovery while doing segmentation, which can identify absent boundaries or missing shapes successfully without necessary region features. In detail, [10] employed the fitting terms of two-phase CV model and the Euler's elastica term [11] as the regularization. Its major contribution was that the missing boundaries were interpolated automatically without specifying the regions. [12] improved the segmentation with depth problem [13, 14] and achieved acceleration via the strategies of model simplification and constraint projection. Its significant performance enhancements included shape reconstruction of occluded objects and determination of their ordering relation in a specific scene based on only one single image. Many other works illustrated that the curvature-related terms have played crucial roles in the boundary reconstruction [15, 16] and image restoration [17, 18, 19] with the capacity of producing excellent edge and corner preservation results. All of these researches show the significant potential for curvature-based methods.

However, the current segmentation models mentioned above cannot be directly applied for noisy images when the type of noise is unknown or there are more than one type of noise in the image. The reason is that in these models there exists a one-to-one mapping relationship between the parameters to be evaluated and the noisy images with some probability density distribution. Besides, the curvature-related terms will bring extra computational complexity due to the existence of nonlinear higher-order derivatives. This issue also appears in other variational models such as the non-texture image inpainting [20] and image denoising [21] with features (edge, corner, smoothness, contrast, etc.) preservation. Hence it is essential to take some mathematical optimization techniques, for example with global solution, stability guarantee and calculation acceleration, into consideration in the process of algorithm design.

For the difficulty of dealing with stochastic noises which are inherently generated

by the acquisition procedure of imaging due to various issues, [22, 23] have given us a great source of inspiration. The authors extended the range of applications for progressive hedging algorithm (PHA) in multistage stochastic variational inequality problems and explored stochastic complementarity problems in a two-stage formulation. Convergence of the algorithm was proved in detail and how the PHA performs was validated through numerical experiments as well as its stability and practicability. One of our major motivations in this paper is to embed stochastic property into segmentation energy functional in images with unknown noises or arbitrary damages and then implement PHA to solve it. Furthermore, Euler's elastica term will be also considered as the regularization in our variational formulations design since it has better properties in dealing with image feature information. Last but not least, in order to improve the computational efficiency and solve problems caused by the non-convex, non-differentiable, nonlinear and higher order terms involved in Euler's elastica related functional, fast algorithms of alternating direction method of multipliers (ADMM) [12, 16, 18, 26] and curvature weighted approach [20, 27, 28] will be systematically designed in the PHA algorithm framework as a fusion for energy minimization problems. The novel introduction of Euler's elastica term along with its analytical study will be another main contribution of our research.

Our contributions can be summarized in the following aspects:

- (i) We propose novel formulations for image segmentation with stochastic noises for various applications by transforming the original minimization problems into the optimization framework of stochastic programming.
- (ii) Euler's elastica term is employed to achieve completion of meaningful missing boundaries, reconstruction of occluded structures of objects and realization of segmentation simultaneously, which further enhances the segmentation performance.
- (iii) Our novel variational formulations will be applied in the problems of two-phase Euler's elastica based segmentation and segmentation with depth in gray and color spaces respectively.
- (iv) A general numerical algorithm based on PHA is proposed. Fusion of ADMM and curvature weighted approach (ADMM-C) is designed for the minimization of the Euler's elastica energy related sub variational problems. The minimization problems derived from ADMM-C will be then efficiently solved by Fast Fourier transform (FFT) [28, 36] and analytical soft threshold formulas [16, 18].

The rest of this paper is structured as follows. Section 2 briefly reviews the related approaches in this field. Our proposed approach and algorithm framework are presented in Section 3. The experiments conducted with performance evaluation and comparison are described in Section 4 followed by the conclusion in Section 5.

## 2 Research background

For the purpose of clarifying the motivations in this paper, some related works will be briefly reviewed before presenting our contributions clearly in Sections 3.

### 2.1 Euler’s elastica based segmentation

Illusory contour capture and shape reconstruction [29, 30] are a challenging problem which aims to complete the missing boundaries or fuzzy areas for existing objects in an image. It is a very common phenomenon in human vision. Part of illusory contours consists of objects’ actual boundaries and the other part is made of missing perceptual edges. However, current computer technique can only deal with closed boundaries. It is extremely difficult for computers to identify illusory contours automatically. Euler’s elastica based segmentation techniques play an important role due to their crucial roles in boundary reconstruction and image restoration.

**Two-phase Euler’s elastica based segmentation:** Zhu, Tai, and Chan [10] proposed the Chan-Vese-Euler (CVE) model designed for the foreground shape recovery problem by combining the CV model [2] and Euler’s elastica regularizer [11]. This model could recover the illusory contours and form a complete meaningful object, even without requiring initialization of fixed points. According to their work, the energy functional is defined as

$$E(\phi, c) = \alpha_1 \int_{\Omega} (f - c_1)^2 \phi dx + \alpha_2 \int_{\Omega} (f - c_2)^2 (1 - \phi) dx + \int_{\Omega} (\alpha + \beta \kappa^2) |\nabla \phi| dx, \quad (1)$$

where  $\mu, \alpha, \beta$  are positive penalty parameters,  $\phi$  is a binary level set representation supposed to take on either 0 or 1. The last term of this functional is the classic Euler’s elastica term.  $\kappa$  denotes the curvature represented as  $\kappa = \nabla \cdot (\nabla \phi / |\nabla \phi|)$ .

**Segmentation with depth information:** In [13], Nitzberg, Mumford and Shiota defined the problem of segmentation with depth information as a problem of recovering occluded shapes and their ordering relations based on a 2D image. The variables defined in this problem are in three folds: 1) the shapes of the regions  $R_1, R_2, \dots, R_n$  to which different objects belong; 2) the ordering relations among objects; 3) the pixel intensities of objects. Without loss of generality, one can assume that the objects  $R_1, R_2, \dots, R_n$  in an image are in ascending order, i. e.,  $R_1$  is the nearest object to the observer while  $R_n$  is the farthest one (i.e. background). Let  $R'_i$  be defined as the visible part of  $R_i$ , i.e.,  $R'_1 = R_1, R'_i = R_i - \bigcup_{j < i} R_j, (i = 2, \dots, n)$ . In addition,  $R'_{n+1} = \Omega - \bigcup_{j < n+1} R_j$  is defined as the visible background. Based on the above assumptions and definitions, the level set based energy functional is formulated as in [14]

$$E(\varphi, c) = \sum_{i=1}^n \int_{\Omega} (\alpha + \beta |\kappa_i|) |\nabla \varphi_i| \delta(\varphi_i) dx + \int_{\Omega} (f - c_{n+1})^2 \prod_{j=1}^n (1 - H(\varphi_j)) dx + \sum_{i=1}^n \left( \int_{\Omega} (f - c_i)^2 H(\varphi_i) \prod_{j=1}^{i-1} (1 - H(\varphi_j)) dx \right), \quad (2)$$

where  $\alpha, \beta$  are two positive penalty parameters,  $c_i \in R_i$  is the pixel intensity of the  $i$ -th object, and  $f$  is the image to be processed.  $\kappa_i$  denotes the curvature of boundary for region  $R_i$ . Here  $|\kappa|$  is chosen to substitute the square in Euler's elastica term with the reason that the object corners can be preserved when  $|\kappa|$  becomes large. In contrast, the model containing curvature related terms in quadratic form would inevitably smear those corners. The difference about the function of curvature selection between  $\kappa^2$  and  $|\kappa|$  for elastica term reflecting in segmentation performance were also indicated in [12, 14]. The level set function  $\varphi$  is represented by a continuous signed distance function.  $H(x)$  and  $\delta(x)$  are Heaviside function and Dirac delta function described in detail in [2, 6].

## 2.2 Segmentation models incorporating noise distributions

Studies [7, 8, 9] investigated noisy image segmentation problems by using specific parameter estimation based on different noise distributions. All the related parameters are calculated via the maximum a posteriori probability (MAP) estimation from the viewpoint of Bayesian probability models. For example, estimation of variance information is used for images degraded with Gaussian noise; The square of image intensity value with capacity of enhancing weak properties is incorporated in the Rayleigh model; Models with great segmentation performance of dealing with Poisson and Gamma noises are built on the standard deviation and average. The general variational functional is written as follows

$$E(\theta, \phi) = \alpha_1 \int_{\Omega} Q_1(x, \theta_1) \phi dx + \alpha_2 \int_{\Omega} Q_2(x, \theta_2) (1 - \phi) dx + \gamma \int_{\Omega} |\nabla \phi| dx, \quad (3)$$

where  $\alpha_1, \alpha_2, \gamma$  are positive penalty parameters,  $\phi$  is a binary level set as defined in functional (1). Specific representations of function  $Q$  derived from the maximum likelihood method and the computation of their related parameters are given in Table 1.  $\theta = (\mu, \sigma)$  refers to the corresponding parameters of function  $Q$  need to be estimated.

**TABLE 1** Potential functions of different noise distributions

Functions	Gaussian noise	Rayleigh noise
$Q_{i(i=1,2)}$	$\frac{1}{2} \log 2\pi + \log \sigma_i + \frac{(f-\mu_i)^2}{2\sigma_i^2}$	$2 \log \sigma_i - \log f + \frac{f^2}{2\sigma_i^2}$
Parameters $\theta_i = (\mu_i, \sigma_i)$	$\mu_i = \frac{\int_{\Omega} f \phi^{2-i} (1-\phi)^{i-1} dx}{\int_{\Omega} \phi^{2-i} (1-\phi)^{i-1} dx}$ $\sigma_i^2 = \frac{\int_{\Omega} (f-\mu_i)^2 \phi^{2-i} (1-\phi)^{i-1} dx}{\int_{\Omega} \phi^{2-i} (1-\phi)^{i-1} dx}$	$\sigma_i^2 = \frac{\int_{\Omega} f^2 \phi^{2-i} (1-\phi)^{i-1} dx}{2 \int_{\Omega} \phi^{2-i} (1-\phi)^{i-1} dx}$
Functions	Poisson noise	Gamma noise
$Q_{i(i=1,2)}$	$\sigma_i - f \log \sigma_i$	$\frac{f}{\mu_i} + \log \mu_i$
Parameters $\theta_i = (\mu_i, \sigma_i)$	$\sigma_i = \frac{\int_{\Omega} f \phi^{2-i} (1-\phi)^{i-1} dx}{\int_{\Omega} \phi^{2-i} (1-\phi)^{i-1} dx}$	$\mu_i = \frac{\int_{\Omega} f \phi^{2-i} (1-\phi)^{i-1} dx}{\int_{\Omega} \phi^{2-i} (1-\phi)^{i-1} dx}$

### 2.3 The progressive hedging algorithm for two-stage stochastic programming

The progressive hedging algorithm (PHA) was originally designed by Rockafellar and Wets [25] for multistage stochastic minimization problems and it has recently been extended [23] to monotone stochastic variational inequality (SVI) problems of the form

$$x(\cdot) \in \mathcal{N} \text{ and } w(\cdot) \in \mathcal{M} \text{ such that } -\nabla F(x(\cdot)) - w(\cdot) \in N_{\mathcal{C}}(x(\cdot)), \quad (4)$$

where  $F$  is a convex continuously differentiable functional,  $N_{\mathcal{C}}(x(\cdot))$  is the normal cone to  $\mathcal{C}$  at  $x(\cdot)$ ,  $\mathcal{N}$  is a linear subspace of the Hilbert space  $\mathcal{L}$  composed of all  $x(\cdot)$ , and  $\mathcal{M}$  is the complementary subspace of  $\mathcal{N}$ . The notation of  $x(\cdot)$  indicates that the solution of Problem (4) is a function and  $\mathcal{L}$  is therefore a functional space. This understanding of (4) is particularly suitable for multistage stochastic optimization, where one seeks the best response function  $x(\xi)$  for each realization of a random vector  $\xi$ . In such case Problem (4) represents the necessary condition for optimality and it is also sufficient if the stochastic programming problem is convex.

To clarify this point, let us present a brief introduction to the two-stage stochastic programming problem, which will be the model we shall use for the image segmentation problem. A two-stage stochastic programming is an optimization problem to make two decisions, one at each stage. The first-stage decision is made before a random vector is realized, and the second-stage decision is a response after the random vector is observed. Let  $\xi$  be the random vector and let  $\Xi$  be the set of all the possible realizations of  $\xi$  (assumed to be finite). Then, the response mapping  $x(\cdot)$  of a two-stage stochastic programming has the following form:

$$x(\cdot) : \xi \mapsto x(\xi) = (x_1(\xi), x_2(\xi)) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^n,$$

where  $x_1(\xi)$  is the first-stage decision variable and  $x_2(\xi)$  is the second-stage decision variable. Consider the Hilbert space  $\mathcal{L}$  consisted of all the mappings  $x(\cdot) \in \mathbb{R}^n$ . Since in our two-stage model the first decision is made before  $\xi$  is realized, the solution to the two-stage stochastic programming problem should satisfy the so-called nonanticipativity constraint  $x(\cdot) \in \mathcal{N}$ , where

$$\mathcal{N} := \{x(\cdot) = (x_1(\cdot), x_2(\cdot)) \in \mathcal{L} \mid x_1(\xi) \text{ does not depend on } \xi\}.$$

Clearly,  $\mathcal{N}$  is a subspace of  $\mathcal{L}$ . In addition, suppose every decision  $x(\cdot)$  must satisfy a set of constraints and the constraints generally depend on  $\xi$ . We write this fact in the form of

$$x(\cdot) \in \mathcal{C} \subset \mathcal{L}, \text{ which means } x(\xi) \in C(\xi) \forall \xi \in \Xi,$$

where each  $C(\xi)$  refers to a nonempty closed convex subset of  $\mathbb{R}^n$  and the set  $\mathcal{C}$  therefore denotes a nonempty closed convex subset of  $\mathcal{L}$ . Let the total cost of the decision  $x(\xi)$  in two stages be  $f(x(\xi), \xi)$ . Then the two-stage stochastic programming problem

can be formulated as

$$\begin{aligned} \min_{x(\cdot)} \quad & F(x(\cdot)) = \mathbb{E}_\xi[f(x(\xi), \xi)] := \sum_{\xi} p(\xi) f(x(\xi), \xi) \\ \text{s.t.} \quad & x(\cdot) \in \mathcal{C} \cap \mathcal{N}, \end{aligned} \quad (5)$$

where  $\mathbb{E}$  stands for the expectation,  $p(\xi) > 0$  represents the known probability of each scenario  $\xi$ , and the sum of these probabilities is 1. If  $f(\cdot, \xi)$  is a continuously differentiable convex function, then the the following Karush-Kuhn-Tucker (KKT) conditions will be both necessary and sufficient for the optimality of  $x(\cdot)$ :

$$-\nabla F(x(\cdot)) \in N_{\mathcal{C} \cap \mathcal{N}}(x(\cdot)). \quad (6)$$

When certain constraint qualifications such as  $\text{ri } \mathcal{C} \cap \mathcal{N} \neq \emptyset$  (or  $\mathcal{C} \cap \mathcal{N} \neq \emptyset$  if the sets  $\mathcal{C}(\xi)$  are all polyhedral) are met, the following equality holds.

$$N_{\mathcal{C} \cap \mathcal{N}}(x(\cdot)) = N_{\mathcal{C}}(x(\cdot)) + N_{\mathcal{N}}(x(\cdot)).$$

Since  $\mathcal{N}$  is a subspace, which yields that  $N_{\mathcal{N}}(x(\cdot)) = \mathcal{N}^\perp := \mathcal{M}$ . Based on this equality, the SVI (6) can be rewritten as the following equivalent form:

$$x(\cdot) \in \mathcal{N} \text{ and } w(\cdot) \in \mathcal{M} \text{ such that } -\nabla F(x(\cdot)) \in N_{\mathcal{C}}(x(\cdot)) + w(\cdot). \quad (7)$$

Subsequently, SVI (4) can be easily obtained from (7) just through moving the term  $w(\cdot)$  from the righthand side to the lefthand side. One can also regard SVI (4) as an equivalent alternative extension of (6). The focus of PHA is to transform solving the original problem stated in (6) into solving the problem

$$-\nabla F(x(\cdot)) - w(\cdot) \in N_{\mathcal{C}}(x(\cdot)) \iff -\nabla_x f(x(\xi), \xi) - w(\xi) \in N_{\mathcal{C}(\xi)}(x(\xi)), \forall \xi. \quad (8)$$

As described in [23], the powerful feature of PHA is its capability of decomposition in terms of solving separate ‘‘deterministic’’ problem for each scenario  $\xi$ , while keeping each iterate in the space  $\mathcal{N}$  by projection. In this regard, the vector  $w(\cdot)$  serves as the multiplier vector for the nonanticipativity constraint, each of the  $w$ -iterates is also kept in the space  $\mathcal{M}$ , see the following description of the algorithm.

---

### PHA in two-stage stochastic programming

---

Given  $x^\nu(\cdot) \in \mathcal{N}$  and  $w^\nu(\cdot) \in \mathcal{M} := \mathcal{N}^\perp$  with a fixed parameter  $r > 0$ ,

**Step 1.** For every  $\xi$ , solve

$$\hat{x}^\nu(\xi) = \arg \min_{x(\xi) \in \mathcal{C}(\xi)} \{f(x(\xi), \xi) + w^\nu(\xi)^T x(\xi) + \frac{r}{2} \|x(\xi) - x^\nu(\xi)\|^2\};$$

**Step 2.** Update  $x^{\nu+1}(\cdot) = P_{\mathcal{N}}(\hat{x}^\nu(\cdot))$ ,  $w^{\nu+1}(\cdot) = w^\nu(\cdot) + r(\hat{x}^\nu(\cdot) - x^{\nu+1}(\cdot))$ .

$\nu := \nu + 1$ , **repeat** until a stopping criterion is met.

---

Here the existence and uniqueness of the solution  $\hat{x}^\nu(\xi)$  will be guaranteed by the fact that the  $r$ -term makes the related minimization subproblems be strongly convex.  $P_{\mathcal{N}}$

is the projection mappings onto the linear subspaces  $\mathcal{N}$ , and actually has the following explicit form in the two-stage stochastic programming:

$$P_{\mathcal{N}}(x(\cdot)) := (z_1(\cdot), z_2(\cdot)) \text{ with } z_1(\xi) = \sum_{\xi} p(\xi)x_1(\xi) \text{ and } z_2(\xi) = x_2(\xi) \forall \xi. \quad (9)$$

If problem (5) is convex and there exists at least one solution, then the sequence generated by PHA converges to an optimal solution. More details in terms of theorems and proofs can be found in [23]. An alternative proof of its convergence is provided in [42] in terms of alternating direction method of multipliers (ADMM). The convergence properties of PHA for the two-stage stochastic programming problem can be summarized as follows.

**Convergence properties of the PHA.** (Theorem 2 of [23]) As long as the (monotone) variational inequality (4) satisfies the constraint qualification and has at least one solution, the sequence  $\{x^\nu(\cdot), w^\nu(\cdot)\}$  generated by the progressive algorithm will converge to some pair  $(\bar{x}(\cdot), \bar{w}(\cdot))$  satisfying (7) and thus furnish  $\bar{x}(\cdot)$  as a solution to (4). In this the distance expressions

$$\|x^\nu(\cdot) - \bar{x}(\cdot)\|^2 + \frac{1}{r^2}\|w^\nu(\cdot) - \bar{w}(\cdot)\|^2$$

will form a decreasing sequence that goes to zero. The decrease will surely be at a linear rate if, in particular, the sets  $C(\xi)$  are polyhedral and the functions  $f(x(\xi), \xi)$  are convex quadratic (including linear as a special case).

The motivation of our work is to treat the noisy image segmentation process as a two-stage stochastic programming problem, in which the first-stage decision is to determine the curve of the object in the image and the randomness  $\xi$  is the distributions of noise, thus we adopt PHA framework to handle the curve under stochastic noises, which will be discussed in more details in the following.

### 3 Novel formulations for different segmentation problems incorporating stochastic noises via progressive hedging

Motivated by the research using PHA to solve the minimization problem of stochastic programming, we aim to propose novel formulations tackling different segmentation issues in consideration of the advantages of Euler's elastica term and the influence of unknown noises. In this way, we not only can fulfill general segmentation tasks with the classic model, but also can deal with some worse situations such as low quality images with large noises, absent boundaries, missing shapes or occlusion. Then we show how to implement PHA with developed ADMM-C algorithm to obtain the optimal solutions efficiently.

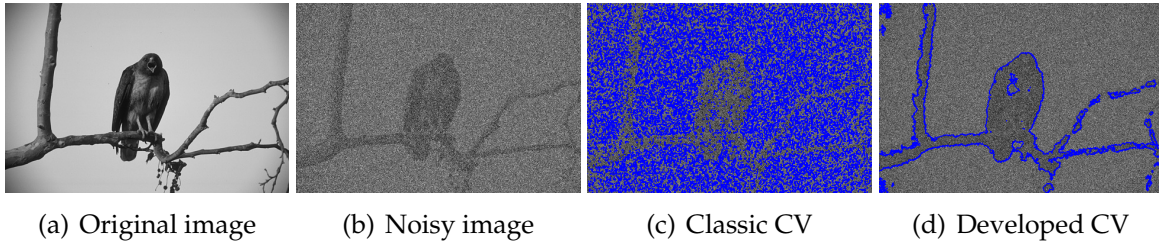
Before we get into our paper topic, we first illustrate how to formulate a new problem and apply progressive hedging algorithm in stochastic programming to improve



the segmentation performance. The obvious feature of this approach is not just robust to noise. It can also deal with stochastic noises, which means we don't need to know what the specific types of noises contained in the image. Figure 1 presents the comparison of results to further demonstrate this feature. The original image is given in (a). (b) presents the image with unknown noises. **Results without or with PHA in stochastic programming are shown in (c) and (d) respectively.** Here we take the classic CV model [2] as an example. And this application will be also a preprocessing procedure providing the initialization of binary level set functions for our segmentation with depth cases. **Specifically, the model for this example can be expressed as the minimization of the following energy functional**

$$E_{\text{SP-gray}}^{\text{CV}}(\theta_\xi, \phi_\xi) = \alpha_1 \int_{\Omega} Q_1(x, \theta_1(\xi))\phi(\xi)dx + \alpha_2 \int_{\Omega} Q_2(x, \theta_2(\xi))(1 - \phi(\xi))dx + \gamma \int_{\Omega} |\nabla\phi(\xi)|dx + \int_{\Omega} (v^k(\xi)\phi(\xi) + \frac{\tau}{2}(\phi(\xi) - \phi_\xi^k)^2)dx, \quad (10)$$

which is an PHA-extended variational model comes from model (3). Here each element  $\xi_i$  of set  $\xi = (\xi_1, \xi_2, \dots, \xi_N)$  is one specific noise type as summarized in Table 1. In accordance with the PHA framework in two-stage stochastic programming introduced in Section 2.3, binary level set functions  $\phi(\xi) = (\phi(\xi_1), \phi(\xi_2), \dots, \phi(\xi_N))$  are the variables of Stage 1 solved by Step 1. Likewise,  $\phi_\xi$  belongs to the variable of Stage 2 relying on the projection mappings implemented on the variables from Stage 1 in Step 2. And the multipliers  $v(\xi) = (v(\xi_1), v(\xi_2), \dots, v(\xi_N))$  correspond to the multipliers  $w(\xi)$ , which also need to be upgraded in Step 2. Now stochastic property of noises has been successfully formulated into classic CV segmentation model by taking advantage of PHA framework.



**Figure 1:** Example of PHA in stochastic programming for segmentation improvement. (a): original image; (b): noisy image; (c): results obtained by CV model; (d): results obtained by CV model with PHA in stochastic programming.

The region-based CV model is well known for capabilities of being insensitive to noises and capturing desired results even on weak boundaries. From Figure 1 (c), we can see that the performance of CV model may be influenced by stronger noises. And (d) validates that the classic CV model using PHA in stochastic programming can achieve improvements for standard segmentation.

We plan to design along this way and develop our work into Euler's elastica based segmentation. In this paper, our works possess the abilities of: 1) producing standard segmentation results as good as the classic CV model [2]; 2) achieving better performance on images with strong unknown noises, missing boundaries or fuzzy areas

(e.g. can complete the disconnected parts of the branch in Figure 1 (d), while CV-PHA functional (10) doesn't have this advantage); and 3) additionally, handling problem of recovering occluded shapes while multiphase segmentation models [3, 4, 6] can't do.

### 3.1 Two-phase segmentation based application

Based on two-phase Euler's elastica based segmentation formulated in (1) and CV-PHA model (10), we propose a novel segmentation model incorporating influence of different noises expressed as the following stochastic programming (SP) form. The reconstructed contour can be obtained by minimizing the following energy functional with respect to  $\phi(\xi)$ .

$$\begin{aligned} & \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-gray}}^{\text{Euler's elastica}}(\theta_\xi, \phi_\xi) \right. \\ &= \alpha_1 \int_{\Omega} Q_1(x, \theta_1(\xi)) \phi(\xi) dx + \alpha_2 \int_{\Omega} Q_2(x, \theta_2(\xi)) (1 - \phi(\xi)) dx \\ & \left. + \int_{\Omega} \left( \alpha + \beta \left| \nabla \cdot \frac{\nabla \phi(\xi)}{|\nabla \phi(\xi)|} \right| \right) |\nabla \phi(\xi)| dx + \int_{\Omega} (v^k(\xi) \phi(\xi) + \frac{\tau}{2} (\phi(\xi) - \phi_\xi^k)^2) dx \right\} \end{aligned} \quad (11)$$

$\xi = (\xi_1, \xi_2, \dots, \xi_N)$  represent different noise distributions and  $Q_i(x, \theta_i(x))$  contain the stochastic information need to be estimated.  $\phi(\xi)$  is the optimal solution of (11) under distribution  $\xi$ . Here we continue the definition of  $\phi(\xi)$  in [10] using the binary representation which can also be explained as a substitution  $\phi = H(\varphi)$ . As they described, this binary representation was originally used for finding the global minimizer. And it can also reduce the computational complexity to some extent such as avoiding the necessary calculation associated with level sets. One can be observed that (11) will reduce to the CV-PHA model (discussed in Functional (10)) when  $\beta$  equals 0. In this way, we can also regard CV based models as particular cases of our proposed segmentation models. **Functional (10) gives a good example that incorporating different noise distributions into one formulation based on PHA can enhance the ability of handling strong noise better than applying them separately and the improved performance has been validated in Figure 1 as well. In addition, Euler's elastica was studied for visual construction in [10, 13, 14, 15], of which the major advantages lied in the effect of reconstructing illusory contours or recovering occluded shapes. [12, 16] found that this good property of Euler's elastica term also reflected in robust against noises. According to the limitation of functional (10) discussed previously, we aim to employ Euler's elastica term as the regularization into our PHA-based segmentation framework with the purpose of merging the advantages of these techniques and make great progress in performance.**

According to the segmentation model for vector-valued images proposed in [31], the averages of the data terms over all channels are used for coupling. Let  $f = (f_1, f_2, \dots, f_m)$  be a original color image defined on a domain  $\Omega$ . Then function  $Q$  should be also in the multichannel form  $(Q_1, Q_2, \dots, Q_m)$ . In fact, our proposed model

used to solve color image segmentation can be stated as follows:

$$\begin{aligned}
& \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-color}}^{\text{Euler's elastica}}(\theta_\xi, \phi_\xi) \right. \\
&= \alpha_1 \int_{\Omega} \sum_{l=1}^m Q_{1l}(x, \theta_{1l}(\xi)) \phi(\xi) dx + \alpha_2 \int_{\Omega} \sum_{l=1}^m Q_{2l}(x, \theta_{2l}(\xi)) (1 - \phi(\xi)) dx \quad (12) \\
& \left. + \int_{\Omega} (\alpha + \beta \left| \nabla \cdot \frac{\nabla \phi(\xi)}{|\nabla \phi(\xi)|} \right|) |\nabla \phi(\xi)| dx + \int_{\Omega} (v^k(\xi) \phi(\xi) + \frac{\tau}{2} (\phi(\xi) - \phi_\xi^k)^2) dx \right\}
\end{aligned}$$

where  $l = 1, 2, \dots, m$  denotes the number of layers of a vector-valued image. Then novel Euler's elastica based formulations embedding stochastic noises for two-phase segmentation are obtained. In the following section we shall implement the calculation under PHA, which is one useful and effective tool for solving above multistage stochastic programming problem.

### 3.2 PHA with developed ADMM-C algorithm for two-phase segmentation application

In order to demonstrate the precise numerical procedure of PHA in solving (11) and (12), we focus on the general model integrating the gray space and color space cases. Detailed proofs for convergence of this algorithm are provided in [23]. The original minimization problems (11) and (12) are based on separate sub optimization problems,

$$\arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_\xi, \phi_\xi) \right\} \Rightarrow \begin{cases} \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_1}, \phi_{\xi_1}) \right\} \\ \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_2}, \phi_{\xi_2}) \right\} \\ \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_3}, \phi_{\xi_3}) \right\} \\ \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_4}, \phi_{\xi_4}) \right\} \end{cases} \quad (13)$$

where  $\xi = (\xi_1, \xi_2, \dots, \xi_N)$  refer to noise distributions. For each distribution  $\xi_i$ , it has a known probability  $p(\xi_i) > 0$  which can be set empirically through experiments, and the sum  $\sum_i p(\xi_i) = 1$  has to be guaranteed. According to PHA approach,  $\phi_\xi$  can be obtained by the following steps.

I. First the minimization problems need to be solved separately, which are given in the right side of (13).

$$\begin{aligned}
(\theta_{\xi_i}^{k+1}, \phi_{\xi_i}^{k+1}) &= \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_i}, \phi_{\xi_i}) \right. \\
&= \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1(\xi_i)) \phi(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2(\xi_i)) (1 - \phi(\xi_i)) dx \quad (14) \\
& \left. + \int_{\Omega} (\alpha + \beta |\kappa(\xi_i)|) |\nabla \phi(\xi_i)| dx + \int_{\Omega} (v^k(\xi_i) \phi(\xi_i) + \frac{\tau}{2} (\phi(\xi_i) - \phi_\xi^k)^2) dx \right\}
\end{aligned}$$

where  $\kappa = \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|}$ . For the gray space issue, there is only one layer of the image information to be calculated, which means  $Q^*(x, \theta(\xi_i)) = Q(x, \theta(\xi_i))$ . For the color space issue,  $Q^*(x, \theta(\xi_i))$  should be substituted with the coupling terms  $\sum_l Q_l(x, \theta_l(x; \xi_i))$ .  $Q^*(x, \theta(\xi_i))$

is the potential function for specific noise distribution in each channel of the image. Table 1 shows the representations of  $Q(x, \theta(\xi_i))$  and estimations of parameters  $\theta(\xi_i)$ .  $\phi_{\xi_i}$  is the optimal solution of (14) in distribution  $\xi_i$ .

**II.** Next, all of the obtained optimal solutions are utilized to gain the final optimum  $\phi_{\xi}$ .

$$\phi_{\xi}^{k+1} = \sum_{i=1}^N p(\xi_i) \phi_{\xi_i}^{k+1}, \quad (15)$$

**III.** At last, the sub problems' solutions  $\phi^k(\xi_i)$  and the Lagrangian multipliers  $v^k(\xi_i)$  need to be updated at the end of each iteration.

$$\begin{cases} \phi^{k+1}(\xi_1) = \phi_{\xi}^{k+1} \\ \phi^{k+1}(\xi_2) = \phi_{\xi}^{k+1} \\ \vdots \\ \phi^{k+1}(\xi_N) = \phi_{\xi}^{k+1} \end{cases}, \quad \begin{cases} v^{k+1}(\xi_1) = v^k(\xi_1) + \tau(\phi^{k+1}(\xi_1) - \phi_{\xi}^{k+1}) \\ v^{k+1}(\xi_2) = v^k(\xi_2) + \tau(\phi^{k+1}(\xi_2) - \phi_{\xi}^{k+1}) \\ \vdots \\ v^{k+1}(\xi_N) = v^k(\xi_N) + \tau(\phi^{k+1}(\xi_N) - \phi_{\xi}^{k+1}) \end{cases}. \quad (16)$$

Then updated  $\phi^{k+1}(\xi_i)$ ,  $v^{k+1}(\xi_i)$  and the parameters  $\theta^{k+1}(\xi_i)$  derived from (14) are passed to the next iteration from step **I**.

To solve the minimization problems (14) separately, both simplification and effectiveness of the algorithm should be considered. In fact, there are three main computational difficulties in the functional as listed below, followed by the corresponding algorithm design.

- (i) One main challenge of the Euler's Elastica based functional is due to the non-smoothness and non-convexity of  $g(\kappa) = \alpha + \beta|\kappa|$ . As described in [20, 27, 28], it is more efficient when the curvature term  $g(\kappa)$  is computed separately from the functional (14). Inspired by the concept of curvature weighted approach, we can rewrite functional (14) as the following simplified version

$$\begin{aligned} (\theta_{\xi_i}^{k+1}, \phi_{\xi_i}^{k+1}) &= \arg \min_{\theta, \phi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_i}, \phi_{\xi_i}) \right. \\ &= \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1(\xi_i)) \phi(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2(\xi_i)) (1 - \phi(\xi_i)) dx \\ &\quad \left. + \int_{\Omega} g(\kappa(\xi_i)) |\nabla \phi(\xi_i)| dx + \int_{\Omega} (v^k(\xi_i) \phi(\xi_i) + \frac{\tau}{2} (\phi(\xi_i) - \phi_{\xi}^k)^2) dx \right\} \end{aligned} \quad (17)$$

The proposed approach essentially reduces the minimization problem (17) to a total variation type [32]. Here the case of division by zero in  $g(\kappa)$  should be avoided. In practice, the denominator is often replaced by  $|\nabla \phi|_{\epsilon} = \max(\epsilon, |\nabla \phi|)$  ( $\epsilon$  is a arbitrarily small positive parameter). Then  $g(\kappa(\xi_i))$  is represented as  $\nabla \cdot (\nabla \phi(\xi_i)) / |\nabla \phi(\xi_i)|_{\epsilon}$ .

- (ii) Note that the binary constraint for  $\phi$  will also cause non-convexity in sub problems (17). [33] demonstrated that certain non-convex minimization problems

can be equivalent to the following convex minimization problems

$$\begin{aligned}
(\theta_{\xi_i}^{k+1}, \phi_{\xi_i}^{k+1}) &= \arg \min_{\theta, \phi \in [0,1]} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_i}, \phi_{\xi_i}) \right. \\
&= \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1(\xi_i)) \phi(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2(\xi_i)) (1 - \phi(\xi_i)) dx \\
&\left. + \int_{\Omega} g(\kappa(\xi_i)) |\nabla \phi(\xi_i)| dx + \int_{\Omega} (v^k(\xi_i) \phi(\xi_i) + \frac{\tau}{2} (\phi(\xi_i) - \phi_{\xi}^k)^2) dx \right\} \quad (18)
\end{aligned}$$

This convex minimization scheme could find global minimizers for (17) by thresholding the solution of (18), which was classified as a continuous min-cut algorithm. Together with its equivalent form, the continuous max-flow algorithm, the min-cut algorithm has been proved to be an exact convex relaxation of the original problem as discussed in [34, 35].

- (iii) Another critical issue for solving (18) is the inevitable high order derivatives in numerical implementation, which is proven to be tedious and prone to errors. Then a developed ADMM-C algorithm is designed to each sub problem by introducing auxiliary variables, Lagrangian multipliers and an alternating directional optimization strategy.

Here the detailed implementation on solving each sub-problem by the ADMM-C algorithm will be presented. Firstly, some auxiliary variables are introduced, i.e.,  $\vec{w}(\xi_i) = [w_1(\xi_i), w_2(\xi_i)]^T$  with property  $\vec{w} \approx \nabla \phi(\xi_i)$  and the Lagrangian multipliers  $\vec{\lambda}(\xi_i) = [\lambda_1(\xi_i), \lambda_2(\xi_i)]^T$ . Based on this observation, we can transform (18) into the following augmented Lagrangian functional

$$\begin{aligned}
(\theta_{\xi_i}^{k+1}, \phi_{\xi_i}^{k+1}, \vec{w}_{\xi_i}^{k+1}) &= \arg \min_{\theta, \phi \in [0,1]} \left\{ E_{\text{SP-general}}^{\text{Euler's elastica}}(\theta_{\xi_i}, \phi_{\xi_i}) \right. \\
&= \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1(\xi_i)) \phi(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2(\xi_i)) (1 - \phi(\xi_i)) dx \\
&+ \int_{\Omega} g(\kappa(\xi_i)) |\vec{w}(\xi_i)| dx + \int_{\Omega} (v^k(\xi_i) \phi(\xi_i) + \frac{\tau}{2} (\phi(\xi_i) - \phi_{\xi}^k)^2) dx \\
&\left. + \int_{\Omega} (\vec{\lambda}(\xi_i) \cdot (\vec{w}(\xi_i) - \nabla \phi(\xi_i)) + \frac{\mu}{2} (\vec{w}(\xi_i) - \nabla \phi(\xi_i))^2) dx \right\} \quad (19)
\end{aligned}$$

where  $\mu$  is a positive penalty parameter. It is worth noting that this kind of simple structure of (19) requires fewer variables and parameters compared with other works [11,18,19] using the ADMM to deal with the curvature term directly. After the initialization of  $\phi^0(\xi_i)$ ,  $\vec{w}^0(\xi_i)$  and  $\vec{\lambda}^0(\xi_i)$ , a minimization problem is carried out in each step with respect to one variable while keeping other variables fixed temporarily. When alternating optimization for all the variables is completed, the Lagrangian multipliers will be updated subsequently. This gives

$$\theta_{\xi_i}^{k+1} = \arg \min_{\theta} \left\{ \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1(\xi_i)) \phi^k(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2(\xi_i)) (1 - \phi^k(\xi_i)) dx \right\}, \quad (20)$$

$$\begin{aligned}
\phi_{\xi_i}^{k+1} &= \arg \min_{\phi \in [0,1]} \left\{ \alpha_1 \int_{\Omega} Q_1^*(x, \theta_1^{k+1}(\xi_i)) \phi(\xi_i) dx + \alpha_2 \int_{\Omega} Q_2^*(x, \theta_2^{k+1}(\xi_i)) (1 - \phi(\xi_i)) dx \right. \\
&+ \int_{\Omega} (v^k(\xi_i) \phi(\xi_i) + \frac{\tau}{2} (\phi(\xi_i) - \phi_{\xi}^k)^2) dx \\
&\left. + \int_{\Omega} (\vec{\lambda}^k(\xi_i) \cdot (\vec{w}^k(\xi_i) - \nabla \phi(\xi_i)) + \frac{\mu}{2} (\vec{w}^k(\xi_i) - \nabla \phi(\xi_i))^2) dx \right\} \quad (21)
\end{aligned}$$

$$\begin{aligned}
\vec{w}_{\xi_i}^{k+1} &= \operatorname{argmin}_{\vec{w}} \left\{ \int_{\Omega} g(\kappa^{k+1}(\xi_i)) |\vec{w}(\xi_i)| dx + \int_{\Omega} \vec{\lambda}^k(\xi_i) \cdot (\vec{w}(\xi_i) - \nabla \phi^{k+1}(\xi_i)) \right. \\
&\left. + \frac{\mu}{2} (\vec{w}(\xi_i) - \nabla \phi^{k+1}(\xi_i))^2 dx \right\}, \quad \text{where } g(\kappa^{k+1}(\xi_i)) = \nabla \cdot \frac{\nabla \phi_{\xi_i}^{k+1}}{|\nabla \phi_{\xi_i}^{k+1}|_{\epsilon}}, \quad (22)
\end{aligned}$$

$$\vec{\lambda}_{\xi_i}^{k+1} = \vec{\lambda}_{\xi_i}^k + \mu (\vec{w}^{k+1}(\xi_i) - \nabla \phi^{k+1}(\xi_i)). \quad (23)$$

**To obtain**  $\theta^{k+1} = (\mu, \sigma)$ : In the  $k + 1$  step of the proposed ADMM-C, the average image intensity values  $\mu_{\xi_i}$  as well as variances  $\sigma_{\xi_i}$  in the foreground and background can be obtained by using the standard variational method for (20). Table 1 gives all the solutions for distributions  $\xi = (\xi_1, \xi_2, \dots, \xi_N)$ .

**To obtain**  $\phi^{k+1}$ : Optimal value of  $\phi_{\xi_i}$  for a certain distribution  $\xi_i$  is obtained by solving the minimization of (21) with respect to  $\phi(\xi_i)$ . We can get the update rule based on the corresponding Euler-Lagrange equations

$$(-\mu \Delta + v^k(\xi_i) + \tau) \phi(\xi_i) = \tau \phi_{\xi}^k - r_{\xi_i}(\theta_1^{k+1}, \theta_2^{k+1}) - \nabla \cdot \vec{\lambda}^k(\xi_i) - \mu \nabla \cdot \vec{w}^k(\xi_i), \quad (24)$$

where  $r_{\xi_i}(\theta_1^{k+1}, \theta_2^{k+1}) = \alpha_1 Q_1^*(x, \theta_1^{k+1}(\xi_i)) - \alpha_2 Q_2^*(x, \theta_2^{k+1}(\xi_i))$ . Like in [36], equation (24) is a screened Poisson equation for which the Fast Fourier transform (FFT) [28, 36] is a well-known solver with very low computational cost for imaging problems. Here FFT is applied for further improving the calculation efficiency. Equation (24) can be rewritten as

$$F^* L F \phi(\xi_i) = \tau \phi_{\xi}^k - r_{\xi_i}(\theta_1^{k+1}, \theta_2^{k+1}) - \nabla \cdot \vec{\lambda}^k(\xi_i) - \mu \nabla \cdot \vec{w}^k(\xi_i), \quad (25)$$

where  $L = -\mu F \Delta F^* + v^k(\xi_i) + \tau$  and  $F^*$  is the discrete inverse Fourier transform. Then we can obtain  $\phi_{\xi_i}^{k+1}$  as follows

$$\phi_{\xi_i}^{k+1} = F^*(L^{-1} F(\tau \phi_{\xi}^k - r_{\xi_i}(\theta_1^{k+1}, \theta_2^{k+1}) - \nabla \cdot \vec{\lambda}^k(\xi_i) - \mu \nabla \cdot \vec{w}^k(\xi_i))). \quad (26)$$

**To obtain**  $\vec{w}^{k+1}$ : The minimization problem (22) of  $\vec{w}$  can be solved via the generalized soft thresholding formula [16, 18], which is given by

$$\vec{w}_{\xi_i}^{k+1} = \max \left( \left| \nabla \phi_{\xi_i}^{k+1} - \frac{\vec{\lambda}^k(\xi_i)}{\mu} \right| - \frac{g(\kappa^{k+1}(\xi_i))}{\mu}, 0 \right) \frac{\nabla \phi_{\xi_i}^{k+1} - \frac{\vec{\lambda}^k(\xi_i)}{\mu}}{\left| \nabla \phi_{\xi_i}^{k+1} - \frac{\vec{\lambda}^k(\xi_i)}{\mu} \right|}. \quad (27)$$

For clarity, we present the overall algorithm for the two-phase Euler's elastica based segmentation in stochastic programming in a pseudo code format as follows.

---

**Algorithm 1 Computing framework for (11) and (12) via PHA**

---

**Input:**  $\phi^0(\xi_i), p(\xi_i), v^0(\xi_i), \alpha, \beta, \tau, \alpha_1, \alpha_2$

**for**  $k \geq 1$ , do the following steps recurrently

- 1: Obtain  $\phi_{\xi_i}^{k+1}$  via **Algorithm 2**
- 2: Update  $\phi_{\xi}^{k+1}$  via Equation (15)
- 3: Update  $\phi^{k+1}(\xi_i), v^{k+1}(\xi_i)$  via Equation (16)
- 4: **if** some stopping criteria (given in Section 3.5) are satisfied **break**

**Return** optimal value  $\phi_{\xi}^{k+1}$  after thresholding

---

---

**Algorithm 2 Detailed implementation for step 1 in Algorithm 1 via ADMM-C**

---

**If**  $k = 1$

input  $\vec{w}^0(\xi_i), \vec{\lambda}^0(\xi_i), \mu$

**else** solve the following problems alternatively

- 1: Update  $\theta_{\xi_i}^{k+1}$  according to distribution laws
  - 2: Update  $\phi_{\xi}^{k+1}$  via minimization problem (21)
  - 3: Update  $\vec{w}_{\xi_i}^{k+1}$  via minimization problem (22)
  - 4: Update  $\vec{\lambda}_{\xi_i}^{k+1}$  via (23) using gradient ascent method
- 

The idea of the ADMM-C method is intentionally applied for the main challenge of the non-convex, non-smooth and non-linear problems in Euler's elastics and it has attracted extensive attention. ADMM algorithm has been given analytical properties in [37, 38] and many other applications [39, 40] are provided in which many similar algorithms are developed and successfully used to achieve excellent performances via solving a variety of non-convex problems. Then a similar idea as in [20, 28] of designing a new algorithm to deal with the sub-problems derived from PHA framework is adopted.

### 3.3 Segmentation with depth based application

We intend to use similar stochastic programming skills as the ones applied in the two-phase issue in this part. Based on original segmentation with depth model for gray space reviewed in Section 2.1, we can establish the energy functional by introducing

random noise set as follows

$$\begin{aligned}
& \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-gray}}^{\text{Depth}}(\theta_\xi, \phi_\xi) \right. \\
& = \sum_{h=1}^n \int_{\Omega} (\alpha + \beta |\nabla \cdot \frac{\nabla \phi_h(\xi)}{|\nabla \phi_h(\xi)|}) |\nabla \phi_h(\xi)| dx + \sum_{h=1}^n \int_{\Omega} Q_h(x, \theta_h(\xi)) \phi_h(\xi) \prod_{j=1}^{h-1} (1 - \phi_j(\xi)) dx \\
& \left. + \int_{\Omega} Q_{n+1}(x, \theta_{n+1}(\xi)) \prod_{j=1}^n (1 - \phi_j(\xi)) dx + \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi) \cdot \phi_h(\xi) + \frac{\tau}{2} (\phi_h(\xi) - \phi_h^k(\xi))^2) dx \right\}
\end{aligned} \tag{28}$$

where  $h = 1, 2, \dots, n$  denote the number of objects in the image, and  $\phi_{n+1}(\xi) = 1$  is set only for consistency of description. We can still extend above segmentation with depth incorporating stochastic noises model to multichannel case. Analog to the coupling approach used in (12), the formulation is written as

$$\begin{aligned}
& \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-color}}^{\text{Depth}}(\theta_\xi, \phi_\xi) \right. \\
& = \sum_{h=1}^n \int_{\Omega} (\alpha + \beta |\nabla \cdot \frac{\nabla \phi_h(\xi)}{|\nabla \phi_h(\xi)|}) |\nabla \phi_h(\xi)| dx \\
& + \sum_{h=1}^n \int_{\Omega} \sum_{l=1}^m Q_{hl}(x, \theta_{hl}(\xi)) \phi_h(\xi) \prod_{j=1}^{h-1} (1 - \phi_j(\xi)) dx \\
& + \int_{\Omega} \sum_{l=1}^m Q_{(n+1)l}(x, \theta_{(n+1)l}(\xi)) \prod_{j=1}^n (1 - \phi_j(\xi)) dx \\
& \left. + \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi) \cdot \phi_h(\xi) + \frac{\tau}{2} (\phi_h(\xi) - \phi_h^k(\xi))^2) dx \right\}
\end{aligned} \tag{29}$$

In order to describe the calculation procedure explicitly, we plan to conduct on the general model integrating the gray space and color space cases. According to the curvature-weighted approach used in (17), the simplified version can be directly written as

$$\begin{aligned}
& \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Depth}}(\theta_\xi, \phi_\xi) \right. \\
& = \sum_{h=1}^n \int_{\Omega} g(\kappa_h(\xi)) |\nabla \phi_h(\xi)| dx + \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h(\xi)) \chi_h(\xi) dx \\
& \left. + \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}(\xi)) \chi_{n+1}(\xi) dx + \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi) \cdot \phi_h(\xi) + \frac{\tau}{2} (\phi_h(\xi) - \phi_h^k(\xi))^2) dx \right\}
\end{aligned} \tag{30}$$

where  $g(\kappa_h(\xi)) = \nabla \cdot (\nabla \phi_h(\xi)) / |\nabla \phi_h(\xi)|_\epsilon$  and the definition of  $|\nabla \phi_h(\xi)|_\epsilon$  is given in (17).  $Q_h^*(x, \theta_h(\xi))$  is  $Q_h(x, \theta_h(\xi))$  for gray space issue and  $\sum_l Q_{hl}(x, \theta_{hl}(\xi))$  for color space issue. And the characteristic function for the  $h$ -th region reads  $\chi_h(\xi) = \phi_h(x) \prod_{j=1:h-1} (1 - \phi_j(\xi))$ . Particularly,  $\chi_{n+1}(\xi) = \phi_{n+1}(\xi) \prod_{j=1:n} (1 - \phi_j(\xi)) = \prod_{j=1:n} (1 - \phi_j(\xi))$  representing the  $(n + 1)$ -th region background. Next section we shall calculate above general formulation under PHA with detailed implementation.



### 3.4 PHA with developed ADMM-C algorithm for segmentation with depth application

For the numerical part, we shall try to use similar ideas as used in the sophisticated two-phase imaging tasks. The main program loop for PHA is shown below

I. First we need to solve the sub minimization problems of (30) separately, which gives

$$\begin{aligned}
& (\theta_{h(\xi_i)}^{k+1}, \phi_{h(\xi_i)}^{k+1}) \\
= & \arg \min_{\theta_h, \phi_h \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Depth}}(\theta_{h(\xi_i)}, \phi_{h(\xi_i)}) \right. \\
= & \sum_{h=1}^n \int_{\Omega} g(\kappa_h(\xi_i)) |\nabla \phi_h(\xi_i)| dx + \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h(\xi_i)) \chi_h(\xi_i) dx \\
& \left. + \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}(\xi_i)) \chi_{n+1}(\xi_i) dx + \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi_i) \cdot \phi_h(\xi_i) + \frac{\tau}{2} (\phi_h(\xi_i) - \phi_{h(\xi)}^k)^2) dx \right\} \quad (31)
\end{aligned}$$

II. Next all of the obtained optimal solutions  $(\phi_{h(\xi_1)}, \phi_{h(\xi_2)}, \dots, \phi_{h(\xi_N)})$  are utilized to gain the final optimum  $\phi_{h(\xi)}$ .

$$\phi_{h(\xi)}^{k+1} = \sum_{i=1}^N p(\xi_i) \phi_{h(\xi_i)}^{k+1}. \quad (32)$$

III. At last, the sub problems' solutions  $\phi_h^k(\xi_i)$  and the Lagrangian multipliers  $v_h^k(\xi_i)$  need to be updated at the end of each iteration.

$$\left\{ \begin{array}{l} \phi_h^{k+1}(\xi_1) = \phi_{h(\xi)}^{k+1} \\ \phi_h^{k+1}(\xi_2) = \phi_{h(\xi)}^{k+1} \\ \vdots \\ \phi_h^{k+1}(\xi_N) = \phi_{h(\xi)}^{k+1} \end{array} \right\}, \left\{ \begin{array}{l} v_h^{k+1}(\xi_1) = v_h^k(\xi_1) + \tau(\phi_h^{k+1}(\xi_1) - \phi_{h(\xi)}^{k+1}) \\ v_h^{k+1}(\xi_2) = v_h^k(\xi_2) + \tau(\phi_h^{k+1}(\xi_2) - \phi_{h(\xi)}^{k+1}) \\ \vdots \\ v_h^{k+1}(\xi_N) = v_h^k(\xi_N) + \tau(\phi_h^{k+1}(\xi_N) - \phi_{h(\xi)}^{k+1}) \end{array} \right\}. \quad (33)$$

Then updated  $\phi_h^{k+1}(\xi_i)$ ,  $v_h^{k+1}(\xi_i)$  and the parameters  $\theta^{k+1}(\xi_i)$  derived from (31) are passed to the next iteration from step I.

In order to solve the minimization problems (31) in step I separately, we will show how to apply the designed ADMM-C algorithm. Note that there are  $n$  binary functions  $(\phi_1(\xi_i), \phi_2(\xi_i), \dots, \phi_n(\xi_i))$  need to be obtained for one specific noise distribution. Thus  $n$  auxiliary variables  $(\vec{w}_1(\xi_i), \vec{w}_2(\xi_i), \dots, \vec{w}_n(\xi_i))$  are introduced and each component is defined as  $\vec{w}_h(\xi_i) = [\vec{w}_{h1}(\xi_i), \vec{w}_{h2}(\xi_i)]^T$  with property  $\vec{w}_h \approx \nabla \phi_h(\xi_i)$ . Likewise,  $n$  Lagrangian multipliers  $(\vec{\lambda}_1(\xi_i), \vec{\lambda}_2(\xi_i), \dots, \vec{\lambda}_n(\xi_i))$  are also brought in with  $\vec{\lambda}_h(\xi_i) = [\vec{\lambda}_{h1}(\xi_i), \vec{\lambda}_{h2}(\xi_i)]^T$ . Based on the convex relaxation method, (31) is rewritten

into the following augmented Lagrangian functional

$$\begin{aligned}
& (\theta_{h(\xi_i)}^{k+1}, \phi_{h(\xi_i)}^{k+1}, \vec{w}_{h(\xi_i)}^{k+1}) \\
= & \arg \min_{\theta_h, \phi_h \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Depth}}(\theta_{h(\xi_i)}, \phi_{h(\xi_i)}) \right. \\
= & \sum_{h=1}^n \int_{\Omega} g(\kappa_h(\xi_i)) |\nabla \phi_h(\xi_i)| dx + \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h(\xi_i)) \chi_h(\xi_i) dx \\
+ & \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}(\xi_i)) \chi_{n+1}(\xi_i) dx + \sum_{h=1}^n \int_{\Omega} v_h^k(\xi_i) \cdot \phi_h(\xi_i) + \frac{\tau}{2} (\phi_h(\xi_i) - \phi_{h(\xi)}^k)^2 dx \\
+ & \left. \sum_{h=1}^n \int_{\Omega} \vec{\lambda}_h(\xi_i) \cdot (\vec{w}_h(\xi_i) - \nabla \phi_h(\xi_i)) + \frac{\mu}{2} (\vec{w}_h(\xi_i) - \nabla \phi_h(\xi_i))^2 dx \right\} \tag{34}
\end{aligned}$$

where  $\mu$  is a positive penalty parameter. Here  $h = 1, 2, \dots, n$  refer to the number of binary level set functions and  $i = 1, 2, \dots, N$  refer to the number of noise distributions. In order to solve (34) efficiently with the ADMM-C, we first initialize the unknown  $\phi_h^0(\xi_i)$ ,  $\vec{w}_h^0(\xi_i)$  and  $\vec{\lambda}_h^0(\xi_i)$  at the initial iterative step  $k = 0$ , then, we solve some minimization problems with respect to only one kind of unknowns while other ones are temporarily fixed at each step from  $k$ -th to  $(k + 1)$ -th until convergence is reached. With this alternating direction optimization strategy, we can divide the optimization problem (34) into three minimization problems in the iterative process from  $k$ -th to  $(k + 1)$ -th step:

$$\theta_{h(\xi_i)}^{k+1} = \arg \min_{\theta_h} \left\{ \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h(\xi_i)) \chi_h^k(\xi_i) dx + \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}(\xi_i)) \chi_{n+1}^k(\xi_i) dx \right\}, \tag{35}$$

$$\begin{aligned}
\phi_{h(\xi_i)}^{k+1} = & \arg \min_{\phi \in [0,1]} \left\{ \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h^{k+1}(\xi_i)) \chi_h(\xi_i) dx + \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}^{k+1}(\xi_i)) \chi_{n+1}(\xi_i) dx \right. \\
+ & \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi_i) \phi_h(\xi_i) + \frac{\tau}{2} (\phi_h(\xi_i) - \phi_{h(\xi)}^k)^2) dx \\
+ & \left. \sum_{h=1}^n \int_{\Omega} (\vec{\lambda}_h^k(\xi_i) \cdot (\vec{w}_h^k(\xi_i) - \nabla \phi_h(\xi_i)) + \frac{\mu}{2} (\vec{w}_h^k(\xi_i) - \nabla \phi_h(\xi_i))^2) dx \right\} \tag{36}
\end{aligned}$$

$$\begin{aligned}
\vec{w}_{h(\xi_i)}^{k+1} = & \arg \min_{\vec{w}_h} \left\{ \sum_{h=1}^n \int_{\Omega} g(\kappa_h^{k+1}(\xi_i)) |\vec{w}_h(\xi_i)| dx + \sum_{h=1}^n \int_{\Omega} \vec{\lambda}_h^k(\xi_i) \cdot (\vec{w}_h(\xi_i) - \nabla \phi_h^{k+1}(\xi_i)) \right. \\
+ & \left. \frac{\mu}{2} (\vec{w}_h(\xi_i) - \nabla \phi_h^{k+1}(\xi_i))^2 dx \right\}, \quad \text{where } g(\kappa_h^{k+1}(\xi_i)) = \nabla \cdot \frac{\nabla \phi_h^{k+1}(\xi_i)}{|\nabla \phi_h^{k+1}(\xi_i)|_{\epsilon}}, \tag{37}
\end{aligned}$$

$$\vec{\lambda}_{h(\xi_i)}^{k+1} = \vec{\lambda}_{h(\xi_i)}^k + \mu (\vec{w}_h^{k+1}(\xi_i) - \nabla \phi_h^{k+1}(\xi_i)). \tag{38}$$

**To obtain**  $\theta_h^{k+1} = (\mu_h, \sigma_h)$ : The average image intensity values  $\mu_h^{k+1}(\xi_i)$  as well as variances  $\sigma_h^{k+1}(\xi_i)$  in the foreground and background can be obtained by using the standard variational method based on (35), which are given by the following equations

**TABLE 2** Potential functions of different noise distributions for segmentation with depth application

Functions	Gaussian noise	Rayleigh noise
$Q_{h'}$ ( $h' = 1, \dots, n+1$ )	$\frac{1}{2} \log 2\pi + \log \sigma_{h'} + \frac{(f - \mu_{h'})^2}{2\sigma_{h'}^2}$	$2 \log \sigma_{h'} - \log f + \frac{f^2}{2\sigma_{h'}^2}$
<b>Parameters</b> $\theta_{h'} = (\mu_{h'}, \sigma_{h'})$	$\mu_{h'} = \frac{\int_{\Omega} f \chi_{h'} dx}{\int_{\Omega} \chi_{h'} dx}$ $\sigma_{h'}^2 = \frac{\int_{\Omega} (f - \mu_{h'})^2 \chi_{h'} dx}{\int_{\Omega} \chi_{h'} dx}$	$\sigma_{h'}^2 = \frac{\int_{\Omega} f^2 \chi_{h'} dx}{2 \int_{\Omega} \chi_{h'} dx}$
Functions	Poisson noise	Gamma noise
$Q_{h'}$ ( $h' = 1, \dots, n+1$ )	$\sigma_{h'} - f \log \sigma_{h'}$	$\frac{f}{\mu_{h'}} + \log \mu_{h'}$
<b>Parameters</b> $\theta_{h'} = (\mu_{h'}, \sigma_{h'})$	$\sigma_{h'} = \frac{\int_{\Omega} f \chi_{h'} dx}{\int_{\Omega} \chi_{h'} dx}$	$\mu_{h'} = \frac{\int_{\Omega} f \chi_{h'} dx}{\int_{\Omega} \chi_{h'} dx}$

**To obtain**  $\phi_h^{k+1}$ : For the minimization problem (36) with respect to the function  $\phi_h(\xi_i)$ , the corresponding Euler-Lagrange equation is given as

$$\begin{aligned}
& (-\mu\Delta + v_h^k(\xi_i) + \tau)\phi_h(\xi_i) \\
&= -Q_h^*(x, \theta_h^{k+1}(\xi_i)) \prod_{j=1}^{h-1} (1 - \phi_j(\xi_i)) + \tau\phi_h^k(\xi) - \nabla \cdot \vec{\lambda}_h^k(\xi_i) - \mu\nabla \cdot \vec{w}_h^k(\xi_i) \\
&+ \sum_{s=h+1}^{n+1} \{Q_s^*(x, \theta_s^{k+1}(\xi_i))\phi_s(\xi_i) \prod_{j=1}^{h-1} (1 - \phi_j(\xi_i)) \prod_{j=h+1}^{s-1} (1 - \phi_j(\xi_i))\} \quad (39)
\end{aligned}$$

Though above equation is more complicated, FFT can be still applied here for accelerating the calculation. Equation (39) can be rewritten as

$$F^* L F \phi_h(\xi_i) = \tau\phi_h^k(\xi) - \Lambda_h(\xi_i) - \nabla \cdot \vec{\lambda}_h^k(\xi_i) - \mu\nabla \cdot \vec{w}_h^k(\xi_i), \quad (40)$$

where  $L = -\mu F \Delta F^* + v^k(\xi_i) + \tau$  and  $F^*$  is the discrete inverse Fourier transform and

$$\begin{aligned}
\Lambda_h(\xi_i) &= Q_h^*(x, \theta_h^{k+1}(\xi_i)) \prod_{j=1}^{h-1} (1 - \phi_j(\xi_i)) \\
&- \sum_{s=h+1}^{n+1} \{Q_s^*(x, \theta_s^{k+1}(\xi_i))\phi_s(\xi_i) \prod_{j=1}^{h-1} (1 - \phi_j(\xi_i)) \prod_{j=h+1}^{s-1} (1 - \phi_j(\xi_i))\}.
\end{aligned}$$

Then we can obtain optimal value of  $\phi_h(\xi_i)$  as follows

$$\phi_h^{k+1}(\xi_i) = F^*(L^{-1}F(\tau\phi_h^k(\xi) - \Lambda_h(\xi_i) - \nabla \cdot \vec{\lambda}_h^k(\xi_i) - \mu\nabla \cdot \vec{w}_h^k(\xi_i))). \quad (41)$$

**To obtain  $\vec{w}_h^{k+1}$ :** The calculation result of  $\vec{w}$  minimization problem (37) can be obtained via the generalized soft thresholding formula as

$$\vec{w}_{h(\xi_i)}^{k+1} = \max \left( \left| \nabla \phi_{h(\xi_i)}^{k+1} - \frac{\vec{\lambda}_h^k(\xi_i)}{\mu} \right| - \frac{g(\kappa_h^{k+1}(\xi_i))}{\mu}, 0 \right) \frac{\nabla \phi_{h(\xi_i)}^{k+1} - \frac{\vec{\lambda}_h^k(\xi_i)}{\mu}}{\left| \nabla \phi_{h(\xi_i)}^{k+1} - \frac{\vec{\lambda}_h^k(\xi_i)}{\mu} \right|}. \quad (42)$$

For clarity, the overall algorithm for the Euler's elastica based segmentation with depth in stochastic programming in a pseudo code format is presented as follows.

---

**Algorithm 3 Computing framework for (28) and (29) via PHA**

---

**Input:**  $\phi_h^0(\xi_i)$  ( $h = 1, \dots, n$ ),  $p(\xi_i)$ ,  $v_h^0(\xi_i)$ ,  $\alpha$ ,  $\beta$ ,  $\tau$

for  $k \geq 1$ , do the following steps in turn

- 1: Obtain  $\phi_{h(\xi_i)}^{k+1}$  via **Algorithm 4**
- 2: Update  $\phi_{h(\xi)}^{k+1}$  via Equation (32)
- 3: Update  $\phi_h^{k+1}(\xi_i)$ ,  $v_h^{k+1}(\xi_i)$  via Equation (33)
- 4: **if some stopping criteria (given in Section 3.5) are satisfied break**

**Return** optimal value  $(\phi_{1(\xi)}^{k+1}, \phi_{2(\xi)}^{k+1}, \dots, \phi_{n(\xi)}^{k+1})$  after thresholding

---



---

**Algorithm 4 Detailed implementation for step 2 in Algorithm 3 via ADMM-C**

---

**If**  $k = 1$

input  $\vec{w}_h^0(\xi_i)$ , and  $\vec{\lambda}_h^0(\xi_i)$  ( $h = 1, \dots, n$ ),  $\mu$

**else** solve the following problems alternatively

- 1: Update  $\theta_{h(\xi_i)}^{k+1}$  according to distribution laws
  - 2: Update  $\phi_{h(\xi_i)}^{k+1}$  via minimization problem (36)
  - 3: Update  $\vec{w}_{h(\xi_i)}^{k+1}$  via minimization problem (37)
  - 4: Update  $\vec{\lambda}_{h(\xi_i)}^{k+1}$  via (38) using gradient ascent method
- 

### 3.5 Termination criteria

The stopping criteria for our entire algorithm are described in this section. As described in [18, 41], the iterations need to be terminated when the following criteria are satisfied:

- **For major framework PHA in Algorithm 1 and Algorithm 3:** During iteration, the constraint errors of  $(\phi_{\xi_i} - \phi_{\xi})$ , the relative errors of Lagrange multipliers and the optimal solutions should be monitored. They should decrease to a suffi-

ciently small level

$$\text{Algorithm 1} \left\{ \begin{array}{l} R_{\tau}^k = \frac{\sum_{i=1}^N p(\xi_i) \|\phi_{\xi_i}^k - \phi_{\xi}^k\|_{L^1}}{\sum_{i=1}^N p(\xi_i) \|\phi_{\xi_i}^0 - \phi_{\xi}^0\|_{L^1}} \quad (43) \\ R_{v_{\xi}}^k = \frac{\|v_{\xi}^k - v_{\xi}^{k-1}\|_{L^1}}{\|v_{\xi}^{k-1}\|_{L^1}} \text{ with } v_{\xi}^k = \sum_{i=1}^N p(\xi_i) v^k(\xi_i) \quad (44) \\ R_{\phi_{\xi}}^k = \frac{\|\phi_{\xi}^k - \phi_{\xi}^{k-1}\|_{L^1}}{\|\phi_{\xi}^{k-1}\|_{L^1}} \quad (45) \end{array} \right.$$

$$\text{Algorithm 3} \left\{ \begin{array}{l} R_{\tau}^k = \frac{\sum_{h=1}^n \sum_{i=1}^N p(\xi_i) \|\phi_{h(\xi_i)}^k - \phi_{h(\xi)}^k\|_{L^1}}{\sum_{h=1}^n \sum_{i=1}^N p(\xi_i) \|\phi_{h(\xi_i)}^0 - \phi_{h(\xi)}^0\|_{L^1}} \quad (46) \\ R_{v_{h(\xi)}}^k = \frac{\|v_{h(\xi)}^k - v_{h(\xi)}^{k-1}\|_{L^1}}{\|v_{h(\xi)}^{k-1}\|_{L^1}} \text{ with } v_{h(\xi)}^k = \sum_{i=1}^N p(\xi_i) v_h^k(\xi_i) \quad (47) \\ R_{\phi_{h(\xi)}}^k = \frac{\|\phi_{h(\xi)}^k - \phi_{h(\xi)}^{k-1}\|_{L^1}}{\|\phi_{h(\xi)}^{k-1}\|_{L^1}} \quad (48) \end{array} \right.$$

where,  $\|\cdot\|_{L^1}$  denotes the  $L^1$  norm on image domain  $\Omega$ . All components are calculated in pixel wise. If  $R^k < l$  ( $l$  is a small enough parameter), the iteration process will be stopped. Note that Equation (44) and (47) can be quite small if the penalty parameters are large. This is due to their explicit dependence on the penalty parameters.

The relative energy error should also be considered, we can use the following form:

$$R_e^k = \|E^k - E^{k-1}\| / \|E^{k-1}\| \quad (49)$$

where  $E^k = \sum_{i=1}^N p(\xi_i) E^k(\xi_i)$ . The computation stops automatically when  $R_e^k$  is less than a predefined tolerance, which indicates that the energy approaches to its steady state.

- **For sub minimization problems using ADMM-C in Algorithm 2 and Algorithm 4:** The following constraint errors of  $(\vec{w}_{\xi_i} - \nabla \phi_{\xi_i})$  and the relative errors of its corresponding Lagrange multipliers in iterations need to be monitored

$$\text{Algorithm 2} \left\{ \begin{array}{l} R_{\vec{w}_{\xi}}^k = \frac{\sum_{i=1}^N p(\xi_i) \|\vec{w}_{\xi_i}^k - \nabla \phi_{\xi_i}^k\|_{L^1}}{\sum_{i=1}^N p(\xi_i) \|\vec{w}_{\xi_i}^0 - \nabla \phi_{\xi_i}^0\|_{L^1}} \quad (50) \end{array} \right.$$

$$R_{\vec{\lambda}_{\xi}}^k = \frac{\|\vec{\lambda}_{\xi}^k - \vec{\lambda}_{\xi}^{k-1}\|_{L^1}}{\|\vec{\lambda}_{\xi}^{k-1}\|_{L^1}} \text{ with } \vec{\lambda}_{\xi}^k = \sum_{i=1}^N p(\xi_i) \vec{\lambda}^k(\xi_i) \quad (51)$$

$$\text{Algorithm 4} \left\{ \begin{array}{l} R_{\vec{w}_{h(\xi)}}^k = \frac{\sum_{h=1}^n \sum_{i=1}^N p(\xi_i) \|\vec{w}_{h(\xi_i)}^k - \nabla \phi_{h(\xi_i)}^k\|_{L^1}}{\sum_{h=1}^n \sum_{i=1}^N p(\xi_i) \|\vec{w}_{h(\xi_i)}^0 - \nabla \phi_{h(\xi_i)}^0\|_{L^1}} \quad (52) \end{array} \right.$$

$$R_{\vec{\lambda}_{h(\xi)}}^k = \frac{\|\vec{\lambda}_{h(\xi)}^k - \vec{\lambda}_{h(\xi)}^{k-1}\|_{L^1}}{\|\vec{\lambda}_{h(\xi)}^{k-1}\|_{L^1}} \text{ with } \vec{\lambda}_{h(\xi)}^k = \sum_{i=1}^N p(\xi_i) \vec{\lambda}_h^k(\xi_i) \quad (53)$$

All numerical quantities are presented in log scale. Some specific methods are used to tune parameters in the implementation process of the proposed approach. The two parameters in  $g(\kappa) = \alpha + \beta|\kappa|$ ,  $\alpha$  and  $\beta$ , control the length and curvature of the segmentation boundary. The ratio between  $\alpha$  and  $\beta$  is related to the connectivity and smoothness of the level lines. As discussed in [10], the connection of disconnected level lines and smoothness of level lines can be guaranteed by a large parameter  $\beta$ . In addition, how to determine another two parameters:  $\tau$  and  $\mu$  associated with Lagrange multipliers will be illustrated. Numerical indicators give the basis of penalty parameter adjustment. One example of their values selection is given in Experiment 4.1.

Lastly, we turn to the discussion related to the convergency of PHA. PHA is applied to the general stochastic programming problems of two-phase image segmentation and segmentation with depth in gray and color images, which are formulated in (17) and (30). Then PHA is implemented relying on convex minimization for each scenario  $\xi_i$  (given in Algorithm 1 and Algorithm 3). In this way, the sequence of pairs  $(\phi_\xi^k, v_\xi^k)$  generated from PHA are hence sure to converge to a solution pair. The authors in [23] also came to a conclusion about the convergence of PHA theoretically with a clear proof procedure.

## 4 Experimental results

We apply the proposed segmentation formulations and developed algorithm extensively on various synthetic and real images for multiple purposes. Experimental results are used to validate the performance and efficiency of our proposed models and algorithm. All the experiments are implemented on the same platform (Matlab 8.2) on a PC (Intel (R), CPU: 2.80GHz, RAM: 16GB, cores number: 4, architecture: 64-bit).

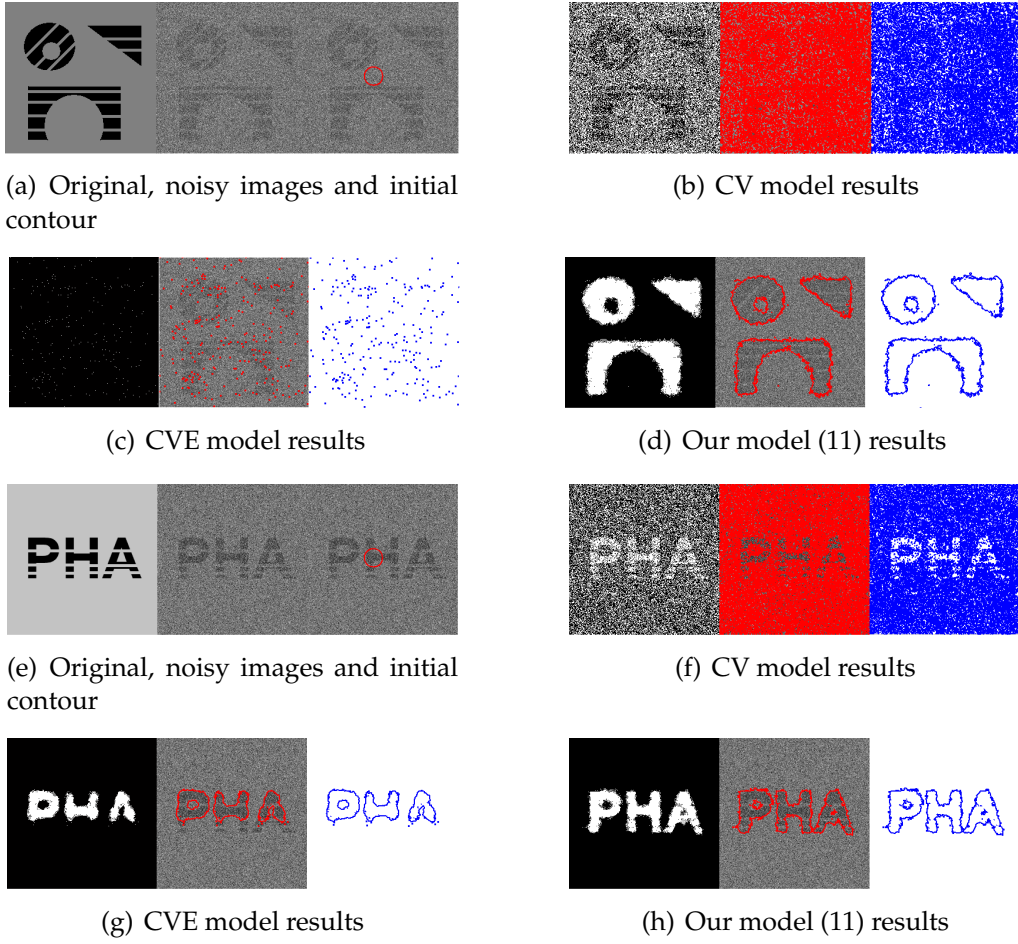
### 4.1 Experiments for two phase cases on Synthetic Images

Some Synthetic images of size  $256 \times 256$  pixels are used as the test images. In these experiments, two-phase CV model [2] and the CVE model [10] are used for comparison in order to show the performance of our proposed model. The reason of choosing [2] and [10] for comparison is that they are both classic models in variational segmentation with excellent results, which brought us a lot of inspiration. First, we set  $v^0(\xi_i) = 0$ ,  $w^0(\xi_i) = \vec{0}$ ,  $\vec{\lambda}^0(\xi_i) = \vec{0}$  and all the Lagrange multipliers are initially set to be 0 for all the numerical experiments in this section. The same initialization of variables in each experiment are used in order to have a relatively fair comparison. In Figure 2, some results of the CV model, CVE model and our proposed model are first presented respectively. The original images, noisy images with stochastic noises including the Gaussian noise, Rayleigh noise, Poisson noise and Gamma noise, and initialization for  $\phi^0$  are shown in (a) and (e). In addition, the pepper & salt noise is additionally contained in (e). And results obtained by the CV model are presented in (b) and (f). (c) and (g) give the final results obtained by CVE model, results from our model (11)

are presented in (d) and (h) separately. From left to right in (b)-(d) and (f)-(h), we start with the optimal solution  $\phi_\xi$ , followed by final curves plotted on noisy images (red lines) and final curves plotted individually (blue lines). It helps to distinguish the detailed differences among results obtained from different models by presenting the final results in blue separately. It can be clearly seen that the results obtained by our model (11) are much better than those two models. The results obtained from CV model are totally different since it is driven by the mean level of the target region resulting in the fact that it can not recognize whether one particular pixel belongs to big noises or objects. The CVE model's performance is unsatisfactory when the desired object has similar density as the background. With the increase of the homogeneity degree, this kind of drawback becomes more obvious.

One thing should be noted is that all the noises are randomly contained in the images for our experiment section, which means we don't need to set specific rate for each type of noise. What proportion of one noise is randomly chosen, so the percentage is also possible to be 0%. From the observation of all experiments, CV model can be robust against noises. But when the noise level is getting higher, CV model fails to obtain the desired results. And CV model can not complete disconnected curves or missing shapes. CVE model has better properties in dealing with segmentation problems in noisy images as well as capturing meaningful structures. While it is also able to omit relatively tiny ones of the objects, thus some small and homogeneous part can not be captured. The parameters for CV model, CVE model and our proposed model are given as follows

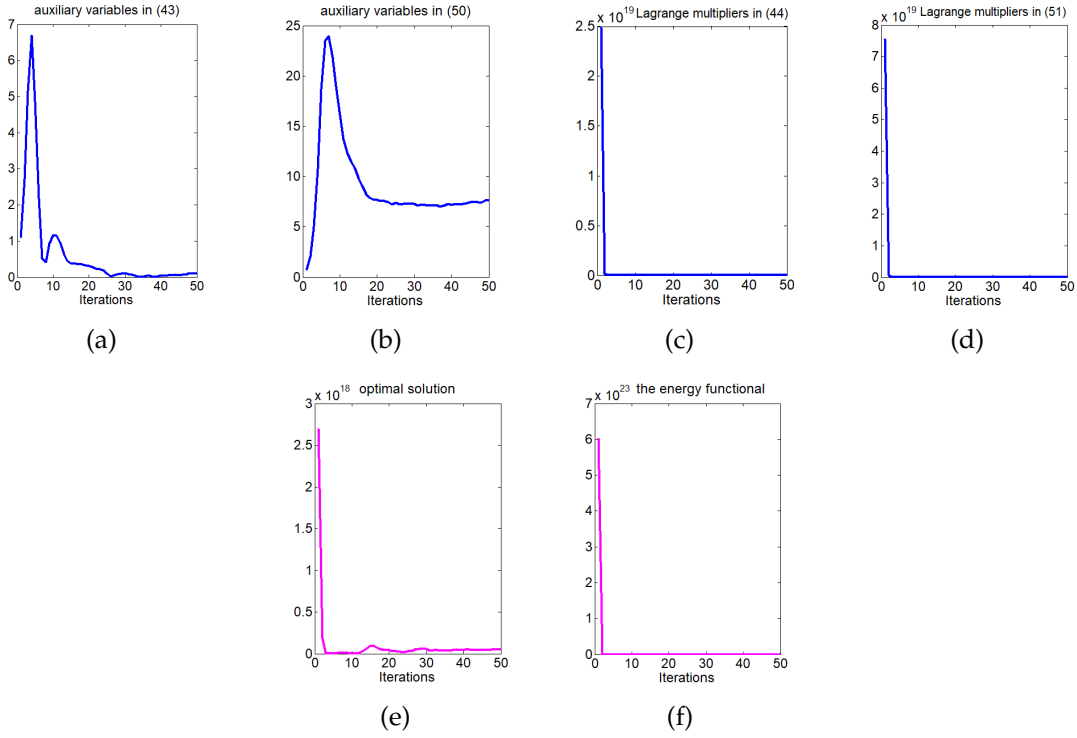
CV model [2]		CVE model [10]	
Figure 2 (b):	$\mu = 20, \gamma = 3$ $\alpha_1 = 10, \alpha_2 = 10$	Figure 2 (c):	$\alpha = 3, \beta = 15, \mu = 20$ $\alpha_1 = 8, \alpha_2 = 8$
Figure 2 (f):	$\mu = 20, \gamma = 3$ $\alpha_1 = 7, \alpha_2 = 7$	Figure 2 (g):	$\alpha = 3, \beta = 10, \mu = 20$ $\alpha_1 = 8, \alpha_2 = 8$
Our proposed model (11) via PHA with ADMM-C			
Figure 2 (d):	$\alpha = 3, \beta = 25, \tau = 5, \mu = 20$ $\alpha_1 = 10, \alpha_2 = 10, p(\xi) = (0.4, 0.1, 0.3, 0.2)$		
Figure 2 (h):	$\alpha = 3, \beta = 25, \tau = 5, \mu = 20$ $\alpha_1 = 7, \alpha_2 = 7, p(\xi) = (0.4, 0.1, 0.3, 0.2)$		



**Figure 2:** Two phase segmentation for synthetic noisy images with incomplete shapes and letters. (a) and (e): original images, noisy images and the same initial  $\phi^0$ ; (b) and (f): results obtained by CV model; (c) and (g): results obtained by CVE model; (d) and (h): final results from our proposed model (11).

Figure 3 gives an example that illustrates the convergence of the relative residuals (Eqn. 43 and 50), the relative errors of Lagrange multipliers (Eqn. 44 and 51), the relative error of  $\phi_\xi^k$  (Eqn. 45) and the energy curve (Eqn. 49) in our model respectively. They are obtained for the image in Figure 2 (d). It is clearly shown that the proposed algorithm has converged well before 100 iterations. They also give an important clue on how to choose the penalty parameters  $t$  and  $m$ . In order to ensure convergence as well as achieving a high speed of convergence, the errors  $R_\tau^k$ ,  $R_{\bar{w}}^k$ ,  $R_v^k$  and  $R_\lambda^k$  should converge steadily with nearly the same speed. If  $R_\tau^k$ ,  $R_{\bar{w}}^k$  go to zero faster than the others,  $\tau$  and  $\mu$  can be decreased and vice versa.  $R_\tau^k$ ,  $R_{\bar{w}}^k$  will converge to zero with the same speed as the iteration proceeds and the energy will decrease to a steady constant value when  $\tau$  and  $\mu$  are chosen properly.





**Figure 3:** The relative residual plots. (a): auxiliary variables in (43); (b): auxiliary variables in (50); (c): Lagrange multipliers in (44); (d): Lagrange multipliers in (51); (e): optimal solution  $\phi_\xi^k$ ; (f): the energy functional.

## 4.2 Experiments for two phase cases on real Images

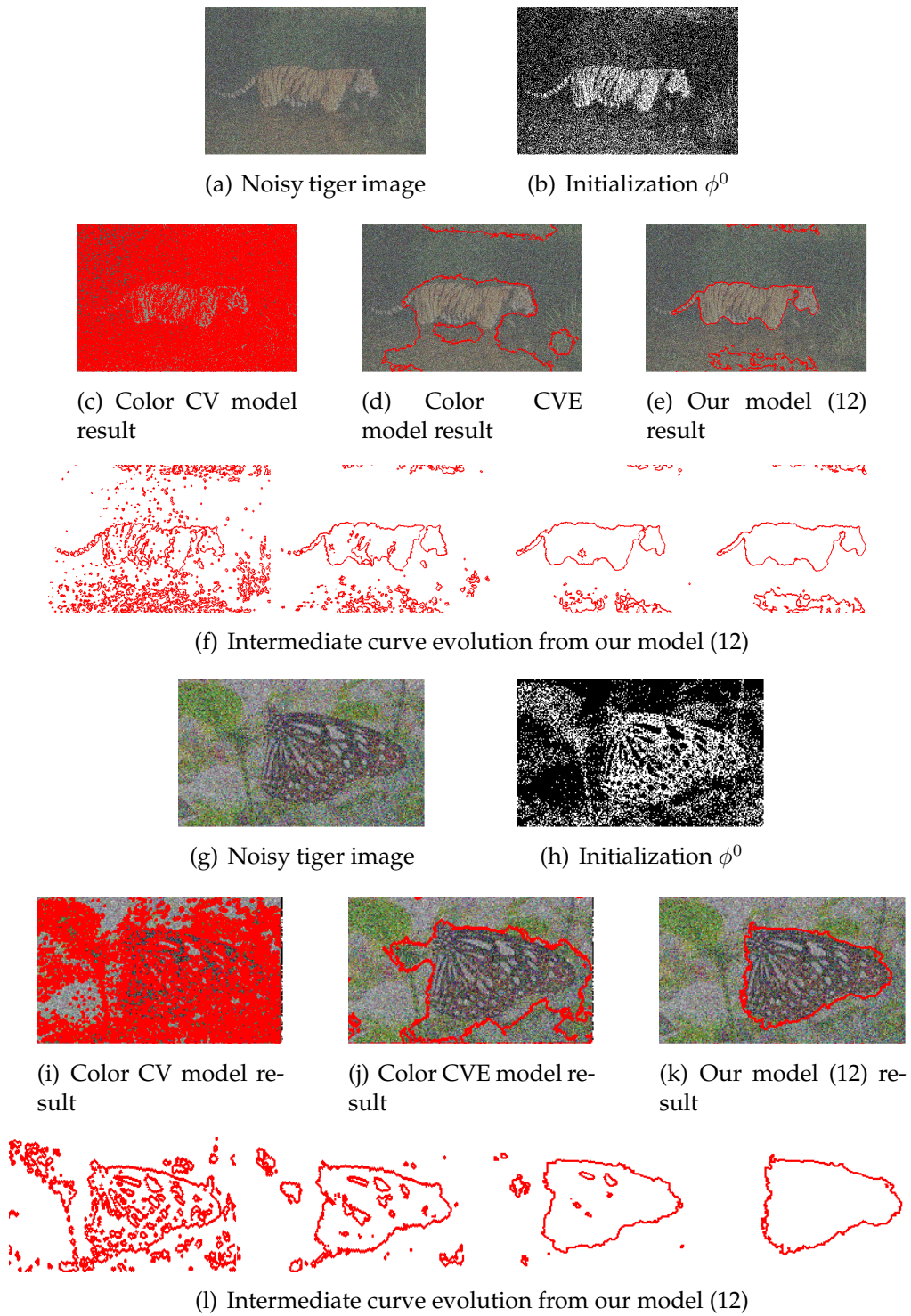
Real tiger image of size  $481 \times 321$  and butterfly image of size  $230 \times 137$  are used as the test images. Visual comparisons with the results and the evolution process of our model are given subsequently. In Figure 4, some results of the color CV model [31], color CVE model and our proposed model (12) are shown respectively. The color CVE model is stated as

$$\begin{aligned}
 E(\phi, c) &= \alpha_1 \int_{\Omega} \sum_{l=1}^m (f_l - c_{1l})^2 \phi dx + \alpha_2 \int_{\Omega} \sum_{l=1}^m (f_l - c_{2l})^2 (1 - \phi) dx \\
 &+ \int_{\Omega} (\alpha + \beta \kappa^2) |\nabla \phi| dx.
 \end{aligned} \tag{54}$$

Here we extend CVE model from gray scale to color scale according to the coupling approach used in [31]. Figure 4 (a), (g) show the real noisy images. (b), (h) give the initialization of  $\phi^0$ . The segmented images by using the color CV model are shown in (c) and (i). Figure 4 (d) and (j) show the results of color CVE model. Our proposed model's performance is reflected in (e) and (k). (f) and (l) show the intermediate evolution process of the contour  $\phi_\xi$  obtained from the proposed model. We directly choose the results from color CV model as the input of  $\phi^0$ . One can also initialize  $\phi^0$  randomly while the final results vary a little. One feature in both of these two images is that a few discontinuous stripes on the tiger's tail or sparse spots on the butterfly's wings. With the proceeding of iteration, it can be observed that CV model fails to capture the

correct boundaries of objects, while both CVE and our model are able to complete an intact shape regardless of the existing gaps within the objects. However, CVE model is inevitably influenced by the stochastic noises especially when these noises smear the elongated structures like the tiger's tail or they increase the homogeneity degree. The parameters used to obtain Figure 4 (c)-(e) and (i)-(k) are

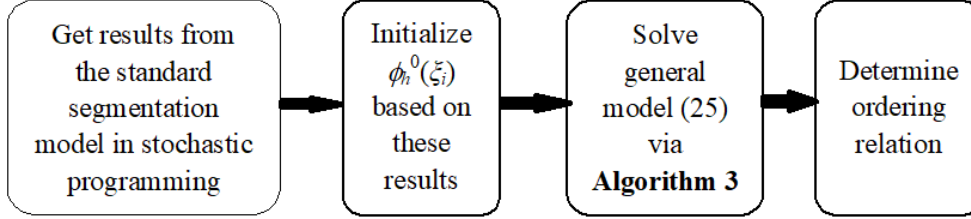
Color CV model [31]		Color CVE model (54)	
Figure 4 (c):	$\mu = 3, \gamma = 3$ $\alpha_1 = 8, \alpha_2 = 5$	Figure 4 (d):	$\alpha = 3, \beta = 8, \mu = 20$ $\alpha_1 = 5, \alpha_2 = 5$
Figure 4 (i):	$\mu = 3, \gamma = 3$ $\alpha_1 = 8, \alpha_2 = 5$	Figure 4 (j):	$\alpha = 3, \beta = 20, \mu = 20$ $\alpha_1 = 7, \alpha_2 = 7$
Our proposed model (12) via PHA with ADMM-C			
Figure 4 (e):	$\alpha = 3, \beta = 4, \tau = 5, \mu = 80$ $\alpha_1 = 4, \alpha_2 = 4, p(\xi) = (0.6, 0.1, 0.2, 0.1)$		
Figure 4 (k):	$\alpha = 3, \beta = 16, \tau = 3, \mu = 50$ $\alpha_1 = 4, \alpha_2 = 4, p(\xi) = (0.6, 0.1, 0.2, 0.1)$		



**Figure 4:** Two phase segmentation for real noisy images with incomplete shapes. (a) and (g): noisy images; (b) and (h): initial  $\phi^0$ ; (c) and (i): results obtained by CV model; (d) and (j): results obtained by CVE model; (e) and (k): final results from our proposed model (12); (f) and (l): intermediate curve evolution by our proposed model.

### 4.3 Experiments for segmentation with depth cases on Synthetic Images

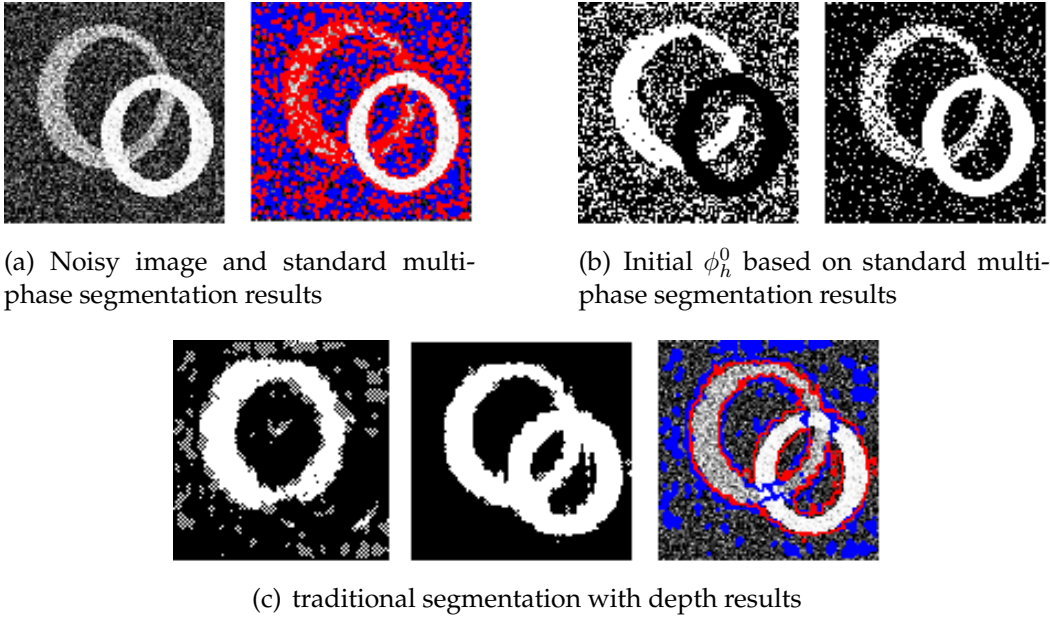
The proposed models (28) and (29) using PHA with ADMM-C algorithm into synthetic images will be compared with classic segmentation with depth model (2) without stochastic programming. The detailed implementation framework is shown as follows and all experiments for segmentation with depth applications follow the same procedure.



In order to speed up the evolution of contours and improve the accuracy of results, we initialize the binary level-set functions  $\phi_h^0(\xi_1) = \dots = \phi_h^0(\xi_N)$  using the results from the the standard multiphase segmentation model in stochastic programming, which is given by

$$\begin{aligned}
 & \arg \min_{\theta_\xi, \phi_\xi \in \{0,1\}} \left\{ E_{\text{SP-general}}^{\text{Multi}}(\theta_\xi, \phi_\xi) \right. \\
 &= \sum_{h=1}^n \int_{\Omega} |\nabla \phi_h(\xi)| dx + \sum_{h=1}^n \int_{\Omega} Q_h^*(x, \theta_h(\xi)) \chi_h(\xi) dx + \int_{\Omega} Q_{n+1}^*(x, \theta_{n+1}(\xi)) \chi_{n+1}(\xi) dx \\
 & \left. + \sum_{h=1}^n \int_{\Omega} (v_h^k(\xi) \cdot \phi_h(\xi) + \frac{\tau}{2} (\phi_h(\xi) - \phi_{h(\xi)}^k)^2) dx \right\} \quad (55)
 \end{aligned}$$

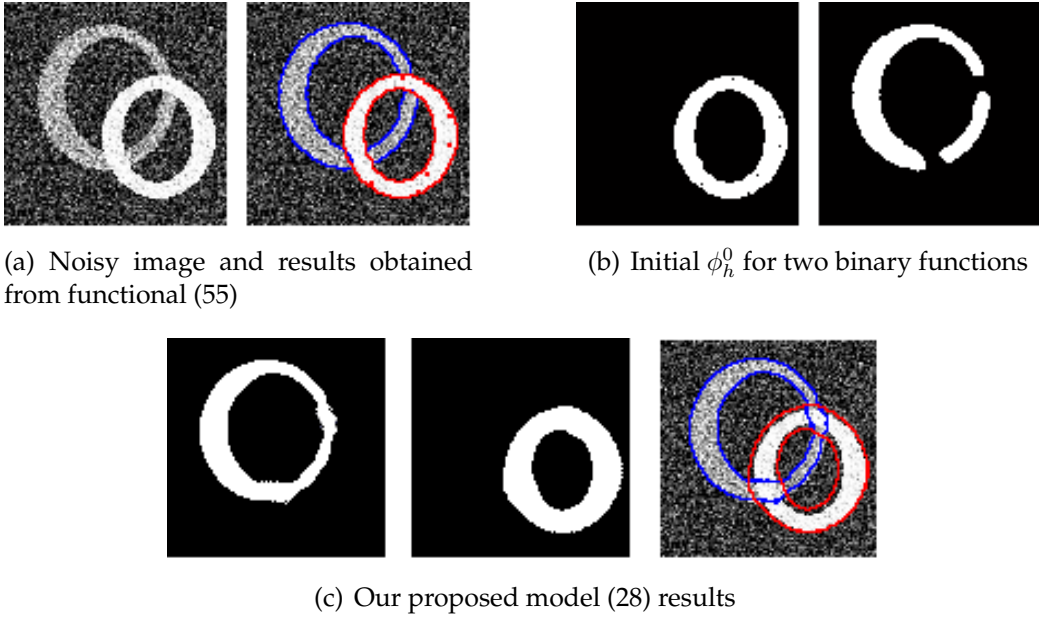
Different from the traditional ones [12, 14] that initialized contours by the standard multiphase segmentation model without stochastic programming, we take the situation of unknown noises into consideration. In experiments, we find the initialization of  $\phi_h^0(\xi_i)$  will be inevitably influenced when big stochastic noises contained in the original image. Then it may lead to a failure for entire framework to obtain expected results. An example is shown below in Figure 5 to explain this situation clearly. Figure 5 (a) gives the synthetic image (size  $100 \times 100$ ) with two circles corrupted by noises randomly for testing and standard multiphase segmentation results. (b) shows the initialization of the binary level set functions by using the results of standard segmentation method. And the final results obtained by traditional segmentation with depth model [12, 14] are presented in (c). It can be observed that traditional implementation framework will not work under the impact of big noises. What's more, the CVE model was designed only for one foreground shape recovery problem without considering segmentation with depth information, which can not be used for comparison.



**Figure 5:** Traditional segmentation with depth for an image with two circles. (a): noisy images and results obtained by the standard multiphase segmentation model; (b): the initialization for two binary functions  $\phi_h^0$ ; (c): results obtained by traditional segmentation with depth model [13,15].

Figure 6 presents the novel implementation framework for segmentation with depth in stochastic programming. The same testing image is used in (a). With the application of (55), we obtain separate objects shown in (a) and initialize  $\phi_h^0$  in (b). Obvious progress in performance of our proposed model (28) compared with the traditional one lies in (c), the final shapes are reconstructed successfully even though there exist big noises. The parameters used for our proposed model (28) are

Our proposed model (28) via PHA with ADMM-C
Figure 6 (c): $\alpha = 3, \beta = 10, \tau = 5, \mu = 30, p(\xi) = (0.5, 0.1, 0.2, 0.2)$



**Figure 6:** Our proposed model (28) for an image with two circles. (a): noisy images and results obtained by the standard multiphase segmentation in stochastic programming model; (b): the initialization for two binary functions  $\phi_h^0$ ; (c): final results obtained by the proposed model

In order to determine the ordering relations of the white circle and gray circle, we minimize the energy functional (28) based on the assumptions that the white circle is occluded by the gray circle or the gray circle is occluded by the white circle. The results are listed in Table 3, from which we can deduce that the white circle, the gray circle and the background are ordered from the nearest to farthest with respect to the observer.

**TABLE 3** Minimal energies of different ordering relations

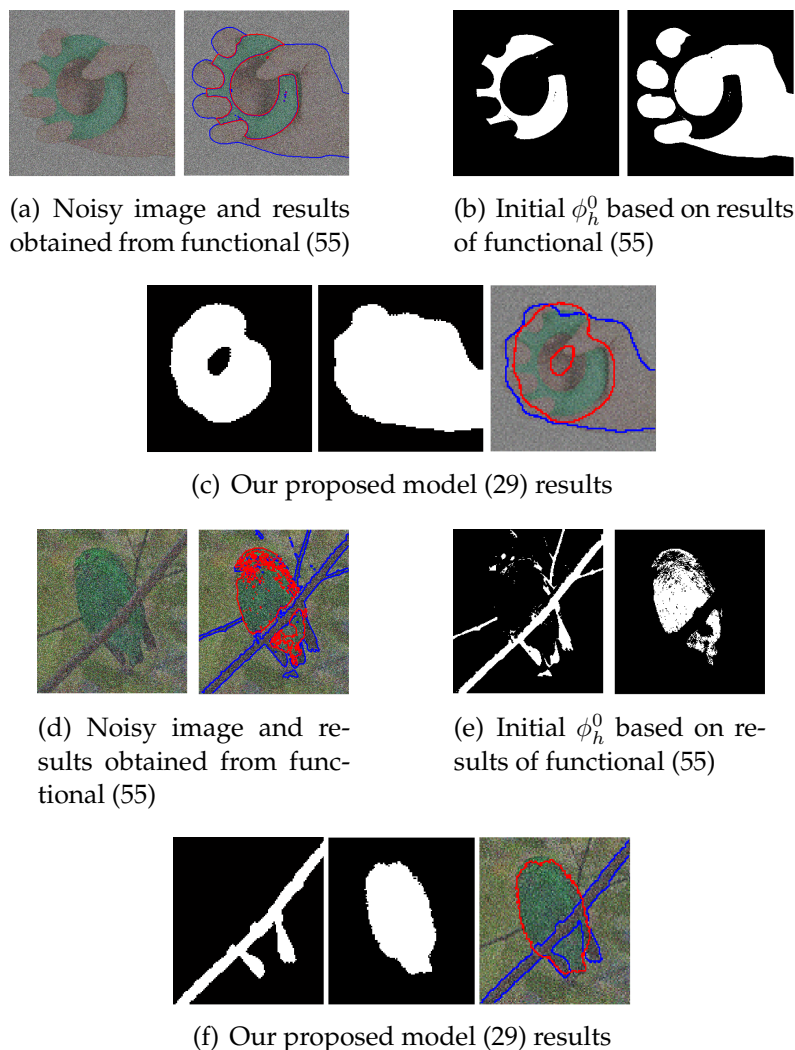
Possible Order	Minimum of energy functional
1. white circle $\Rightarrow$ gray circle	18.56
2. gray circle $\Rightarrow$ white circle	20.83

#### 4.4 Experiments for segmentation with depth cases on real Images

In the last experiment, one real image with a circle and a hand (size  $360 \times 360$ ) and the other with a bird and a trunk (size  $220 \times 241$ ) are shown in Figure 6. Figure 6 (a) and (d) show the original noisy image and the result from standard segmentation model in stochastic programming plotted on the original noisy image. The two initial values for  $\phi_h^0$  are given in (b) and (e). The final results from our model (29) are provided in (c) and (f). Our model can clearly perform well in real images. The parameters for our proposed model (29) are given as follows

Our proposed model (29) via PHA with ADMM-C	
Figure 6 (c):	$\alpha = 3, \beta = 25, \tau = 5, \mu = 20, p(\xi) = (0.4, 0.1, 0.3, 0.2)$
Figure 6 (f):	$\alpha = 3, \beta = 25, \tau = 3, \mu = 10, p(\xi) = (0.5, 0.1, 0.3, 0.1)$





**Figure 7:** Our proposed model (29) for real image. (a) and (d): noisy images and results obtained by the standard multiphase segmentation in stochastic programming model; (b) and (e): the initialization for two binary functions  $\phi_h^0$ ; (c) and (f): final results obtained by the proposed model.

For the ordering relations determination for Figure 7 (c) and (f), we minimize the energy functional based on the assumptions including all the potential ordering respectively. From the results listed in Table 4 and Table 5, we can choose the correct orderings which are mapping to the minimal functional values.

**TABLE 4** Minimal energies of different ordering relations for Figure 7(a)

Possible Order	Minimum of energy functional
1. green circle $\Rightarrow$ hand	40.57
2. hand $\Rightarrow$ green circle	45.26

**TABLE 5** Minimal energies of different ordering relations for Figure 7(b)

Possible Order	Minimum of energy functional
1. bird $\Rightarrow$ trunk	28.56
2. trunk $\Rightarrow$ bird	20.83

There is one important thing need to be noted. When dealing with the convex optimization problem, we have to use a threshold method to realize the binarization of  $\phi_{h(\xi)}^{k+1}$ . It is an important way to help find the accurate results. The histograms of non-threshold and threshold results from Figure 7 (c) are given in Figure 8. It gives us a good way to choose the threshold of  $\phi_{h(\xi)}^{k+1}$ . In practice, we find the threshold  $\eta = 0.5$  could be applicable.

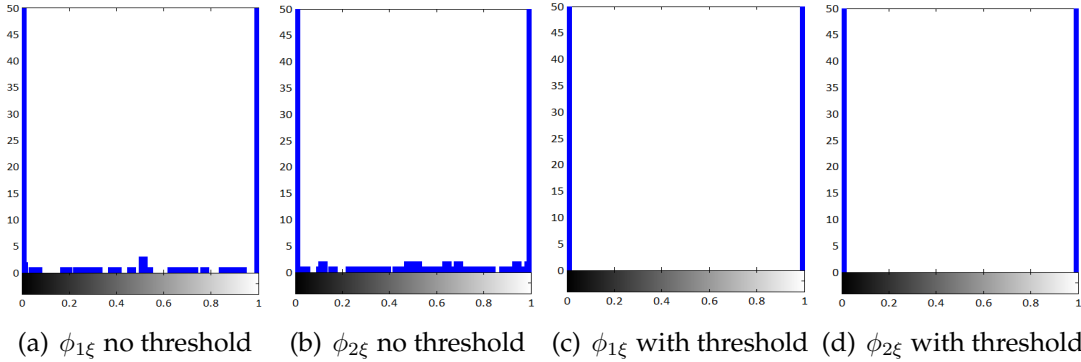


Figure 8: Histograms of the final binary level set functions (non-threshold and threshold). (a) and (b): histogram of non-threshold solutions; (c) and (d): histogram of threshold solutions; (a)-(d) are from Figure 7 (c).

At last, the efficiency of our proposed PHA with ADMM-C algorithm is emphasized by presenting the number of iterations and computation time in Table 6. The iterations and time are shown according to all of our proposed models applied in Experiment 1 to 4. The computational time is measured in seconds.

**TABLE 6** Number of iterations and computational time

Image	Size	Iterations	Time
Figure 2 (d)	$256 \times 256$	80	1.53
Figure 2 (h)	$256 \times 256$	73	1.25
Figure 4 (e)	$481 \times 321$	55	7.6
Figure 4 (k)	$230 \times 137$	50	0.79
Figure 6 (c)	$100 \times 100$	45	1.23
Figure 7 (c)	$360 \times 360$	60	8.9
Figure 7 (f)	$220 \times 241$	55	2.69



## 5 Conclusions

We propose a novel variational approach for image segmentation with stochastic noises and develop a progressive hedging algorithm to solve them. Our approach possesses three outstanding advantages: 1) improving segmentation ability for noisy images without the prerequisite that one given model and one specific noise distribution are counterpoints; 2) realizing completion of meaningful missing boundaries and reconstruction of occluded structures of objects in a highly noisy background; 3) We incorporate the ADMM method and a curvature weighted approach into the calculation procedure to guarantee the segmentation quality on both convergence and efficiency. Extensive experiments were conducted on images with multiple segmentation purposes which is more challenging due to the limited image quality. Experiment results demonstrate the significant performance improvements of our work. Furthermore, for cases with big stochastic noises and damages, our proposed model achieves better performance than the traditional model, which is of great significance for image understanding with problems such as occlusion, large damages or noises, etc.

In future, our work will focus on embedding other powerful techniques such as deep network with generative capacity into a variational framework to cope with more complicated situations. For instance, segmenting salient objects from images with complex background or lower resolution, even though with clutter and partial occlusions.

## References

- [1] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5): 577-685, 1989.
- [2] T. F. Chan, L. A. Vese, Active contours without edges. *Image processing, IEEE transactions on*, 10(2): 266-277, 2001.
- [3] F. Li et al., A multiphase image segmentation method based on fuzzy region competition. *SIAM Journal on Imaging Sciences*, 3(3): 277-299, 2010.
- [4] H. Yu, W. W. Wang, and X. C. Feng, A new fast multiphase image segmentation algorithm based on nonconvex regularizer. *Pattern Recognition*, 45(1): 363-372, 2012.
- [5] B. Wang, X. Yuan, X. Gao et al., A Hybrid Level Set With Semantic Shape Constraint for Object Segmentation. *IEEE Transactions on Cybernetics*, 2018.
- [6] L. A. Vese, T. F. Chan, A multiphase level set framework for image segmentation using the Mumford and Shah model. *International journal of computer vision*, 50(3): 271-293, 2002.

- [7] N. Paragios, R. Deriche, Geodesic active regions: A new framework to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, 13(1-2): 249-268, 2002.
- [8] P. Martin, P. Réfrégier, F. Goudail, F. Guérault, Influence of the noise model on level set active contour segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 26: 799-803, 2004.
- [9] A. Sawatzky et al., A variational framework for region-based segmentation incorporating physical noise models. *Journal of Mathematical Imaging and Vision* 47(3): 179-209, 2013.
- [10] W. Zhu, X. C. Tai, and T. F. Chan, Image segmentation using euler's elastica as the regularization. *Journal of scientific computing*, 57(2):414-438, 2013.
- [11] X. C. Tai, and J. M. Duan. A simple fast algorithm for minimization of the elastica energy combining binary and level set representations. *International journal of numerical analysis and modeling*, 14(6): 809-821, 2017.
- [12] L. Tan, Z. Pan, W. Liu, J. Duan, W. Wei, G. Wang: Image Segmentation with Depth Information via Simplified Variational Level Set Formulation. *Journal of Mathematical Imaging and Vision*, 60(1): 1-17, 2018.
- [13] M. Nitzberg, D. Mumford, and T. Shiota, *Filtering, Segmentation, and Depth*, Lecture Notes in Computer Sciences, 662, Springer-Verlag, Berlin, 1993.
- [14] W. Zhu, T. F. Chan and S. Esedoglu, Segmentation with Depth: A Level Set Approach, *SIAM Journal on Scientific Computing*, 28(5):1957-1973, 2006.
- [15] S. H. Kang, W. Zhu, J. H. Shen, Illusory shapes via corner fusion. *SIAM Journal on Imaging Sciences*, 7(4):1907-1936, 2014.
- [16] L. Tan, W. Liu, L. Li et al. A fast computational approach for illusory contour reconstruction, *Multimedia Tools and Applications*, 1-24, 2018.
- [17] W. Zhu, X. C. Tai, and T. F. Chan, Augmented Lagrangian method for a mean curvature based image denoising model. *Inverse problems and imaging*, 7(4): 1409-1432, 2013.
- [18] L. Tan, W. Liu, Z. Pan, Color image restoration and Inpainting via Multi-Channel Total curvature. *Applied Mathematical Modelling*, 61:280-299, 2018.
- [19] L. Tan, A. Liang, L. Li, et al. Automatic prostate segmentation based on fusion between deep network and variational methods. *Journal of X-Ray Science and Technology (Preprint)*: 1-17, 2019.
- [20] M. Yashtini, S. H. Kang. A Fast Relaxed Normal Two Split Method and an Effective Weighted TV Approach for Euler's Elastica Image Inpainting. *SIAM Journal on Imaging Sciences*, 9(4): 1552-1581, 2016.

- [21] Y. Yan, F. Nie, W. Li et al., Image classification by cross-media active learning with privileged information. *IEEE Transactions on Multimedia* 18(12):2494-2502, 2016.
- [22] R. T. Rockafellar, R. J. B. Wets, Stochastic variational inequalities: single-stage to multistage. *Mathematical Programming*, 165(1): 331-360, 2017.
- [23] R. T. Rockafellar, J. Sun, Solving monotone stochastic variational inequalities and complementarity problems by progressive hedging, *Mathematical Programming*, 174:453–471, 2019.
- [24] R.T. Rockafellar, Monotone operators and the proximal point algorithms. *SIAM J. Control Optim.*, 14: 877-898, 1976.
- [25] R.T. Rockafellar, R.J.-B. Wets, Scenarios and policy aggregation in optimization under uncertainty. *Math. Oper. Res.*, 16: 119-147, 1991.
- [26] R. Glowinski, T. W. Pan, X. C. Tai, Some facts about operator-splitting and alternating direction methods. *Splitting Methods in Communication, Imaging, Science, and Engineering*, Springer, 19-94, 2016.
- [27] E. Bae, J. Shi, and X. C. Tai. Graph cuts for curvature based image denoising. *IEEE Transactions on Image Processing*, 20(5):1199-1210, 2011.
- [28] L. J. Deng, R. Glowinski, X. C. Tai, A New Operator Splitting Method for Euler’s Elastica Model. *arXiv preprint arXiv:1811.07091*, 2018.
- [29] P. Kellman et al., Classification images reveal that deep learning networks fail to perceive illusory contours. *Journal of vision*, 17(10):569-569, 2017.
- [30] T. Poscoliero, M. Girelli, Electrophysiological Modulation in an Effort to Complete Illusory Figures: Configuration, Illusory Contour and Closure Effects. *Brain topography*, 31(2):202-217, 2018.
- [31] T. F. Chan, B. Y. Sandberg, L. A. Vese. Active Contours without Edges for Vector-Valued Images. *Journal of Visual Communication & Image Representation*, 11(2):130-141, 2000.
- [32] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: nonlinear phenomena*, 60 (1-4): 259-268, 1992.
- [33] T. F. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM journal on applied mathematics*, 66(5):1632-1648, 2006.
- [34] J. Yuan, E. Bae, and X.C. Tai. A study on continuous max-flow and min-cut approaches. In *CVPR, USA, San Francisco*, 2010.
- [35] E. Bae et al., A fast continuous max-flow approach to non-convex multi-labeling problems. *Efficient algorithms for global optimization methods in computer vision*. Springer, Berlin, Heidelberg, 134-154, 2014.

- [36] J. M. Morel, A. B. Petro, C. Sbert, Screened Poisson equation for image contrast enhancement, *Image Processing On Line*, 4: 16-29, 2014.
- [37] M. Hong, Z. Q. Luo, On the linear convergence of the alternating direction method of multipliers. *Mathematical Programming*, 162(1-2): 165-199, 2017.
- [38] Q. Liu, X. Shen, Y. Gu, Linearized admm for non-convex non-smooth optimization with convergence analysis. *arXiv preprint arXiv:1705.02502*, 2017.
- [39] M. Myllykoski, R. Glowinski, T. Karkkainen, T. Rossi, A new augmented Lagrangian approach for L1-mean curvature image denoising. *SIAM Journal of Imaging Sciences*, 8(1):95-125, 2015.
- [40] M. Yashtini, Alternating Direction Method of Multiplier for Euler's Elastica-Based Denoising, *Scale Space and Variational Methods in Computer Vision*. Springer International Publishing, 690-701, 2015.
- [41] E. Bae, X. C. Tai, W. Zhu, Augmented Lagrangian method for an Euler's elastica based segmentation model that promotes convex contours. *Inverse Problems and Imaging*, 11(1): 1-23, 2017.
- [42] J. Sun, H.L. Xu, M. Zhang, A new interpretation of the progressive hedging algorithm for multistage stochastic minimization problems. *Journal of Industrial and Management Optimization* (online publication doi: 10.3934/jimo.2019022), 2019.