

Multi-Phase Texture Segmentation Using Gabor Features Histograms Based on Wasserstein Distance

Motong Qiao^{1,*}, Wei Wang² and Michael Ng³

¹ Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong.

² Department of Mathematics, Tongji University, Shanghai 200092, China.

³ Centre for Mathematical Imaging and Vision and Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong.

Received 6 December 2012; Accepted (in revised version) 11 October 2013

Available online 14 March 2014

Abstract. We present a multi-phase image segmentation method based on the histogram of the Gabor feature space, which consists of a set of Gabor-filter responses with various orientations, scales and frequencies. Our model replaces the error function term in the original fuzzy region competition model with squared 2-Wasserstein distance function, which is a metric to measure the distance of two histograms. The energy functional is minimized by alternative minimization method and the existence of closed-form solutions is guaranteed when the exponent of the fuzzy membership term being 1 or 2. We test our model on both simple synthetic texture images and complex natural images with two or more phases. Experimental results are shown and compared to other recent results.

AMS subject classifications: 68U10

Key words: Multi-phase texture segmentation, Wasserstein distance, Gabor filter, Mumford-Shah model.

1 Introduction

Image Segmentation has been one of the most talked about issues in recent decades, due to its fundamental role in many image processing applications, such as image coding, image synthesis, pattern recognition and so on. From the point of view of human vision, most images, including both synthetic ones and natural ones, can be easily segmented into several phases by human naked eyes. However, this human visual system has complex neurobiological criterion for segmentation that would be challenging to imitate exactly for a computer. In a machine vision system, the segmentation criterion is usually

*Corresponding author. *Email address:* qiao.motong@gmail.com (M. Qiao)

based on the difference in intensity value, regional statistics and other features of disjoint phases.

The general N -phase segmentation problem can be illustrated as follows: Given a gray scale image $I: \Omega \rightarrow \mathbb{R}$, where the image domain Ω is a bounded, smooth and open subset of \mathbb{R}^2 , the aim is to partition Ω into N disjoint connected open subsets $\{\Omega_i\}_{i=1}^N$, i.e., $\Omega_i \cap \Omega_j = \emptyset$, $j \neq i$ and $\bigcup_{i=1}^N \Omega_i \cup \Gamma = \Omega$, by certain suitable measures, where Γ is the union of the part of boundaries of the Ω_i inside Ω . The segmentation is achieved by minimizing the summary energy functional:

$$\min \left\{ E(\Gamma, \Omega) = \sum_{i=1}^N \left(\int_{\Gamma_i} ds + \lambda \int_{\Omega_i} r_i(x) dx \right) \right\}, \quad (1.1)$$

where the first term is to restrict the boundary of the segment as short as possible and the second term is to use the error function $r_i(x)$ to measure the similarity between the features of the underlying x and of different phases, so as to determine which phase it should belong to. Many edge-based and region-based models in literature share this idea in nature, while differs in the representation of the region Ω_i , the error function r_i and some other regularization terms might be added.

Snakes/Active Contours [22] and Balloons [13] are classical edge-based segmentation methods which minimize the energy through deforming an initial contour towards the object boundaries. However this energy model cannot deal with topological changes. Incorporating the techniques of curve evolution and geometric flows, many active-contour-based models were proposed (e.g., see [5, 6, 24]). These models works for those images whose object boundaries are smooth and clearly defined by intensity gradient, but in many cases the boundaries might not be simply defined in such a way, especially for texture images. This difficulty inspires many researchers to integrate regional features with edge information. Chan and Vese [7] incorporated the classical Active Contour model with Mumford-Shah functional [35] and assumed each region can be approximated by a piece-wise constant function. The difference with the classical Active Contour model is that its stopping criterion does not rely on gradient of the image. Thus the blurry edges or gaps caused by missing edges would not be an issue. Later they extended this model to handle vector-valued images [8] and recently many other variances were proposed (see [17, 41–43]). M-S model which studied by Mumford and Shah [35] holds the idea that each sub-region Ω_i can be approximated by a piece-wise smooth function s_i . Hence their model consists of three terms, one is the error function $r_i = (I - s_i)^2$, the other two are regularizer on the approximate function $|\nabla s_i|^2$ and on the total length of the boundaries $|\Gamma|$. It has become one of the most dominating region-based models due to its great compatibility with other models and variability of the interpretation, for instance, Bresson et al. [3] incorporate the boundary information and shape prior to the M-S model, Brox and Cremers [44] introduced a statistical interpretation of the M-S functional, recently Sochen and Bar proposed a generalization of the original M-S model in combination of the Beltrami Framework.

Since their model is continuous, in terms of implementation, the main questions need to be considered are the discretization of the edge $|\Gamma|$ and the representation of the sub-regions Ω_i . The first question can be addressed by some discrete approximation methods (see [2, 10, 15]). As for the latter question, a general approach is to apply the level set method of Osher and Sethian [38]. At the beginning, the level set method can only handle two-phase segmentation problem. Later in 2002, Chan and Vese [45] developed a multiphase level set framework based on M-S model. Recent applications of the level set method on image segmentation can be found in [12, 26, 27, 39].

Another dominating representation is based on the theory of fuzzy logic (see [25]) and region competition (see [50]). Zhu and Yuille first illustrate a unified theory of region competition and analysis in [49]. In [34], Mory and Ardon proposed a convex two-phase segmentation framework for fuzzy region competition, using the convex relaxation approach. Moreover, it has been proved that the relaxed convex and the original non-convex problems share the same global minimizers. Readers can refer to [5, 9] and others. Based on their work, Ng and Li developed two multi-phase segmentation methods using nonparametric probability density function [29] and piece-wise constant Mumford-Shah function [28] separately.

In addition to the two questions mentioned above, the feature space chosen to be based on is another essential aspect. Most of the models, such as classical M-S model and C-V model, are based on the intensity value itself of the original image, while other feature spaces often considered includes the regional statistics or histograms, LBP (Local Binary Patterns) and Gabor-filter responses. Among which the Gabor filter has been recognized as a good tool for texture analysis [21] formulated in 2-D case as:

$$G_{\sigma,f,\gamma} = g_{\sigma}(x,y) \exp [2\pi j f (x \cos \gamma + y \sin \gamma)], \quad (1.2)$$

where

$$g_{\sigma}(x,y) = \frac{1}{2\pi\sigma^2} \exp \left[-\frac{x^2+y^2}{2\sigma^2} \right].$$

It can be viewed as an oriented complex sinusoidal grating modulated by the 2-D Gaussian kernel g_{σ} with the frequency f of the span-limited sinusoidal grating, the orientation γ and the scale σ . It is a complex function with real and imaginary part as $G_{\sigma,f,\gamma} = G_R + jG_I$. To obtain the Gabor response of an image, one can just convolve the original image with the Gabor function and compute the magnitude as $G_{\sigma,f,\gamma} * I = \sqrt{(G_R * I)^2 + (G_I * I)^2}$. In real applications on the texture analysis or segmentation, people usually use multiple Gabor functions with different frequencies, orientations and scales, so called the filter bank, to achieve a Gabor feature space. It brings the problem of Gabor filter design, which aims at how to determine the parameters of the Gabor feature space.

In this paper, we take into consideration the histogram of the Gabor feature space based on the frequency f . For simplicity, we experimentally fix the parameter as $f = 0.01: 0.01: 1$, $\sigma = 0.5$ in most cases and the orientation γ need to be tuned for different image. The reason we choose this characteristic is that we experimentally find that the most dominant factor to the segmentation performance is the orientation of the Gabor filter.

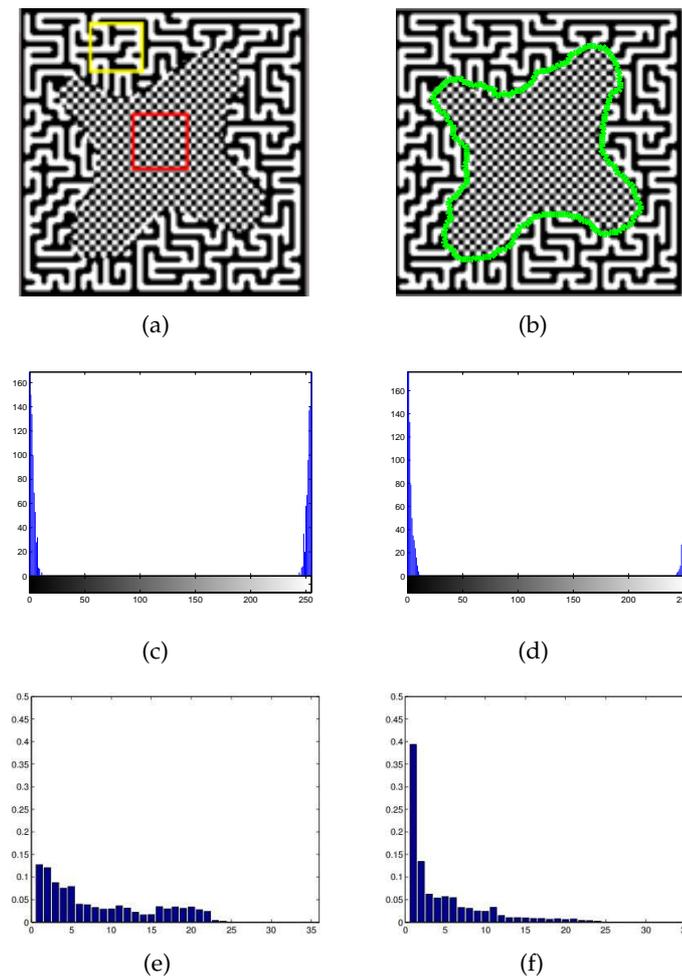


Figure 1: (a): Original Image of size 256×256 (without the red and yellow frame); (b): Segmentation contour using our method; (c): Local intensity histogram of 50×50 region in red box; (d): Local intensity histogram of 50×50 region in yellow box; (e): Histogram based on frequency ranging from 0.01 to 1 in step length 0.01 calculated from our algorithm of the same region in red box; (f): Histogram based on frequency ranging from 0.01 to 1 in step length 0.01 calculated from our algorithm of the same region in yellow box.

Once γ is chosen appropriately, reasonable performance can be obtained while tuning the other two parameters can achieve subtle improvement, but not be necessary in most cases. The advantage of this Gabor feature space towards intensity value is shown in Fig. 1, which consist of two different texture patterns with exactly the same local intensity histogram. Making use of the Gabor filter and the Wasserstein distance, we proposed a generalized multi-phase fuzzy region competition model which includes the model of Chan et al. [36] as a special case.

The other parts of this paper are arranged as follows. In Section 2, some related segmentation models are reviewed, which bring us the inspiration and motivation. Then

we introduce our proposed model in Section 3, followed by numerical analysis and algorithm. Finally, we test our method on some synthetic and natural images with two or more phases in Section 4 and present some discussions in Section 5.

2 Related works

In 2007, Mory and Ardon [33] proposed a two-phase segmentation method based on Fuzzy Region Competition model [34], using $\lambda \int_{\mathcal{A}} (p_i(y) - K(I(x) - y))^2 dy$ as the error function r_i to measure the similarity between a probability density function $p_i(y)$ and a symmetric kernel $K(I(x) - y)$. Their model is as following:

$$\min \left\{ E(u, p_1, p_2) = \int_{\Omega} |\nabla u| + \lambda \int_{x \in \Omega} u(x) \int_{y \in \mathcal{A}} (p_1(y) - K(I(x) - y))^2 dy dx \right. \\ \left. + \lambda \int_{x \in \Omega} (1 - u(x)) \int_{y \in \mathcal{A}} (p_2(y) - K(I(x) - y))^2 dy dx \right\}$$

subject to

$$\int_{\mathcal{A}} K(y) dy = 1 \quad \text{and} \quad K(y) > 0, \quad \forall y \in \mathcal{A}.$$

Their model is non-parametric and utilizes a point-wise distance between two probability density functions p_1 and p_2 . Those models using this kind of point-wise distance, such as χ^2 statistics, Kullback-Leibler distance (see [18, 19, 48]) and the Bhattacharyya distance (see [14, 32]), have some certain disadvantage that, for instance the distance between two delta functions with disjoint supports is the same no matter how close or how far the two supports are from each other.

In the same year, Chan et al. [9] proposed a level set based model utilizing different error function which measures the similarity between two cumulative distribution functions instead, so called the Wasserstein distance (or Monge-Kantorovich distance, details refer to [23] and [40]). Although this distance metric overcomes the drawback mentioned above, this model would face the problem of non-convexity of the level set functions, which means the segmentation results are sensitive to initialization. Later in 2009, Ni et al. [36] modified this model to a convex version with fuzzy membership $u \in [0, 1]$ as following:

$$\min \left\{ E(u, P_1, P_2) = \int_{\partial\Omega} ds + \lambda \int_{\Omega} W_1(P_1, P_x) u(x) dx + \lambda \int_{\Omega} W_1(P_2, P_x) (1 - u(x)) dx \right. \\ = \int_{\partial\Omega} ds + \lambda \int_{\Omega} \int_0^L |F_1(y) - F_x(l)| u(x) dl dx \\ \left. + \lambda \int_{\Omega} \int_0^L |F_2(l) - F_x(l)| (1 - u(x)) dl dx \right\}, \quad (2.1)$$

where P_1 and P_2 represent the histograms of intensity of object and background region separately, F_1 and F_2 are their corresponding cumulative distribution functions and $W_1(\cdot)$

is the linear Wasserstein distance. The Wasserstein distance in literature can be defined in general form:

$$p\text{-Wasserstein Distance} = W_p(P_a, P_b) := \left(\int_0^L |F_a(l) - F_b(l)|^p dl \right)^{1/p}, \quad (2.2)$$

where P_a, P_b stand for two normalized histograms with F_a, F_b being their corresponding cumulative distributions and $l \in [0, L]$ is the bar index of the histogram. However, the proof of the minimizer existence still remains to be resolved when using W_1 distance and the solution of F_1, F_2 requires calculation of weighted median for each bar in the histogram in each iteration. It would be much more costly using the method as they mentioned when the image is larger.

Other variants of this region-based active contour model in combination with regional statistics are as follows. In 2010, Xie [47] proposed another level set based model combining the Gabor filtered feature distribution instead of intensity. Wang et al. [46] extend the model to vector-valued version, which aims at integrating the responses from a filter bank with various orientation, scale and frequency and they introduced a way to select the optimal bins in the histogram. In 2011, Ma and Yu [31] incorporated the global convex Chan-Vese model with quadratic-chi histogram distance as the error function term, as well as using feature from Gabor-filter bank.

The afore mentioned works gave us the inspiration to develop a multi-phase model that overcomes those shortcomings, such as non-convexity, the expensive calculation and the bias caused by the point-wise distance, while makes an appropriate choice among different features like intensity, Local Binary Pattern (LBP) map and Gabor feature space. In the next section, we propose our model which extends Tony et al.'s model to a multi-phase case using the squared W_2 distance based on the Gabor feature space and the proof of existence is given.

3 Proposed model

Firstly, we can rewrite the multi-phase fuzzy region competition model in [28] based on the p -Wasserstein distance (2.2) into a general form as following:

$$\min \left\{ E(u_i, P_i) = \sum_{i=1}^N \int_{\Omega} |\nabla u_i| dx + \frac{\lambda}{q} \sum_{i=1}^N \int_{\Omega} W_p(P_x, P_i)^p u_i^q dx \right\}, \quad (3.1)$$

which incorporates the several previously mentioned model (see [9, 31, 36, 47]) as special cases: when $N = 2$, $p = q = 1$ and $u_i = \{0, 1\}$ being the Heavyside Function of a level set variable, the above model is the one in [9]; with the same setting, if the histogram is based on the Gabor feature space instead of intensity, this is the model proposed in [47]; when $u_i = \{0, 1\}$ changes to a fuzzy membership function $u_i \in [0, 1]$, this is the model in [36]; further on, if change the Wasserstein distance to quadratic-chi histogram distance, it becomes the model in [31].

In our multi-phase segmentation model, we set $p = 1$ or 2 , $q = 1$ or 2 , which means that the W_1 distance or squared W_2 distance are both considered and the fuzziness of the membership function u_i is controlled by q and the phase number N can be greater than 2. Thus our model is given as:

$$\begin{aligned} \min_{u_i, P_i} \left\{ E(u_i, P_i) = \sum_{i=1}^N \int_{\Omega} |\nabla u_i| dx + \frac{\lambda}{q} \sum_{i=1}^N \int_{\Omega} W_p(P_x, P_i)^p u_i^q dx \right. \\ \left. = \sum_{i=1}^N \int_{\Omega} |\nabla u_i| dx + \frac{\lambda}{q} \sum_{i=1}^N \int_{\Omega} \int_0^1 |F_{x,r}(l) - F_i(l)|^p u_i^q dl dx \right\} \end{aligned} \quad (3.2)$$

subject to:

$$(i) \sum_{i=1}^N u_i = 1, \quad (ii) 0 \leq u_i \leq 1, \text{ for } i = 1:N,$$

where $u_i(x)$ is the fuzzy membership function representing the probability of each pixel $x \in \Omega$ belonging to the corresponding phase Ω_i , $P_{x,r}$, P_i are the histograms of Gabor-filter responses corresponding to the local region $\mathcal{N}_r(x)$ centered at x and the resulting segment Ω_i , $F_{x,r}(l)$, $F_i(l)$ are their normalized cumulative histograms with l denoting each bar.

The numerical analysis of the solution of this minimization problem is given in the next section.

3.1 Numerical analysis

In this section, the existence of the minimizer of the following energy functional in the case of $p = 2$ is proved, however, the existence proof when $p = 1$ still remains to be resolved. Here we are trying to minimize the following energy:

$$\min_{u_i, F_i} \left\{ E(u_i, F_i) = \sum_{i=1}^N \int_{\Omega} |\nabla u_i| dx + \frac{\lambda}{q} \int_{\Omega} \int_0^1 (F_{x,r}(l) - F_i(l))^2 u_i^q(x) dl dx \right\} \quad (3.3)$$

subject to:

$$\begin{aligned} (a) \sum_{i=1}^N u_i &= 1, \\ (b) 0 \leq u_i &\leq 1, \text{ for } i = 1, 2, \dots, N, \\ (c) 0 \leq F_i &\leq 1, \text{ for } i = 1, 2, \dots, N, \end{aligned}$$

where λ, q, r are positive parameters. The energy functional is well defined in the following admissible set: $\Lambda = \{(u_i, F_i) | u_i \in BV(\Omega), \text{ satisfies (a) and (b); } F_i \in L^2(0,1), \text{ satisfies (c)}\}$. Then for the problem,

$$\min_{(u_i, F_i) \in \Lambda} \{E(u_1, u_2, \dots, u_N, F_1, F_2, \dots, F_N)\}, \quad (3.4)$$

we have the existence theorem.

Theorem 3.1. Assume $F_{x,r} \in L^\infty(0,1)$ for all $x \in \Omega$, then for fixed parameters N, λ, p, r , there exists at least one solution of the problem (3.4) in the admissible set Λ .

Proof. First, if we let u_i and F_i be constants, the energy will be finite, which implies that problem (3.4) is the correct setting.

Suppose $\{(u_i^n, F_i^n)\}$ is a minimizing sequence of E ; then there exists a constant $M > 0$ such that

$$E(u_1^n, u_2^n, \dots, u_N^n, F_1^n, F_2^n, \dots, F_N^n) \leq M.$$

The above inequality reads as

$$\sum_{i=1}^N \int_{\Omega} |\nabla u_i^n| dx + \frac{\lambda}{q} \int_{\Omega} \int_0^1 (F_{x,r}(l) - F_i^n(l))^2 (u_i^n)^q dl dx \leq M.$$

The boundedness of $\int_{\Omega} |\nabla u_i^n| dx$ and the constraint (b) guarantee that the sequence $\{u_i^n\}$ is uniformly bounded in $BV(\Omega)$ for each $i=1,2,\dots,N$. Noting the compactness property of $BV(\Omega)$ with respect to BV_w^* topology, up to a subsequence also denoted by $\{u_i^n\}$, there exists a function $u_i^* \in BV(\Omega)$ such that,

$$u_i^n \xrightarrow{L^1(\Omega)} u_i^* \quad \text{and} \quad u_i^n \rightarrow u_i^* \quad \text{a.e. } x \in \Omega.$$

Then by using the lower semicontinuity of total variation,

$$\liminf_{n \rightarrow \infty} \int_{\Omega} |\nabla u_i^n| dx \geq \int_{\Omega} |\nabla u_i^*| dx. \tag{3.5}$$

Meanwhile since u_i^n satisfies constraints (a) and (b), by using the convergence result, u_i^* also satisfies (a) and (b). Noting that $F_i^n \in L^2(0,1)$, together with the constraint (c) we can derive that the sequence $\{F_i^n\}$ is uniformly bounded in $L^2(0,1)$ for each $i=1,2,\dots,N$. Therefore, there exists $F_i^* \in L^2(0,1)$ such that, up to a subsequence,

$$F_i^n \rightharpoonup F_i^* \in L^2(0,1) \quad \text{and} \quad F_i^n \rightarrow F_i^* \quad \text{a.e. } l \in (0,1).$$

Therefore,

$$\lim_{n \rightarrow \infty} \int_0^1 F_{x,r}(l) F_i^n(l) dl = \int_0^1 F_{x,r}(l) F_i^*(l) dy,$$

then we can easily deduce that for almost every $x \in \Omega$,

$$\begin{aligned} & \int_0^1 (F_{x,r}(l))^2 dl (u_i^n)^p - 2 \int_0^1 F_{x,r}(l) F_i^n(l) dl (u_i^n)^q \\ \rightarrow & \int_0^1 (F_{x,r}(l))^2 dl (u_i^*)^q - 2 \int_0^1 F_{x,r}(l) F_i^*(l) dl (u_i^*)^q. \end{aligned}$$

Fatou’s lemma gives,

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \int_{\Omega} \left(\int_0^1 (F_{x,r}(l))^2 dl (u_i^n)^q - 2 \int_0^1 F_{x,r}(l) F_i^n(l) dl (u_i^n)^q \right) dx \\ & \geq \int_{\Omega} \left(\int_0^1 (F_{x,r}(l))^2 dl (u_i^*)^q - 2 \int_0^1 F_{x,r}(l) F_i^*(l) dl (u_i^*)^q \right) dx \end{aligned}$$

and

$$\liminf_{n \rightarrow \infty} \int_{\Omega} (u_i^n)^q dx \geq \int_{\Omega} (u_i^*)^q dx.$$

As a consequence of the lower semicontinuity for the L^2 norm,

$$\liminf_{n \rightarrow \infty} \int_0^1 (F_i^n(l))^2 dl \geq \int_0^1 (F_i^*(l))^2 dl.$$

Finally

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \int_{\Omega} \int_0^1 (F_{x,r}(l) - F_i^n(l))^2 (u_i^n)^q dl dx \\ & = \liminf_{n \rightarrow \infty} \int_{\Omega} \int_0^1 (F_{x,r}(l))^2 (u_i^n)^q dl dx - 2 \int_{\Omega} \int_0^1 F_{x,r}(l) F_i^n(l) (u_i^n)^q dl dx \\ & \quad + \int_{\Omega} (u_i^n)^q dx \int_0^1 (F_i^n(l))^2 dl \\ & \geq \int_{\Omega} \left(\int_0^1 (F_{x,r}(l))^2 dl (u_i^*)^q - 2 \int_0^1 F_{x,r}(l) F_i^*(l) dl (u_i^*)^q \right) dx + \int_{\Omega} (u_i^*)^q dx \int_0^1 (F_i^*(l))^2 dl \\ & = \int_{\Omega} \int_0^1 (F_{x,r}(l) - F_i^*(l))^2 (u_i^*)^q dl dx. \end{aligned}$$

Combining the above inequality and (3.5), we have

$$\begin{aligned} & \min E(u_1, u_2, \dots, u_N, F_1, F_2, \dots, F_N) \\ & = \liminf_{n \rightarrow \infty} E(u_1^n, u_2^n, \dots, u_N^n, F_1^n, F_2^n, \dots, F_N^n) \\ & \geq E(u_1^*, u_2^*, \dots, u_N^*, F_1^*, F_2^*, \dots, F_N^*). \end{aligned} \tag{3.6}$$

Meanwhile, we have $(u_i^*, F_i^*) \in \Lambda$, this completes the proof. □

3.2 Numerical algorithm

The alternative minimization method is used in the algorithm, i.e., iteratively minimizing the energy functional (3.2) in respect to u_i or P_i with the other fixed until the stopping criterion is reached. In the case of $q = 1$, the model is indeed the relaxed Potts model, which could be solved by many efficient approach. We adopt the continuous max-flow approach proposed by Yuan et al. [47], which rewrite the original minimization problem with respect to the membership function u_i as a augmented Lagrangian function of the max-flow problem with u_i being the multiplier. In the case of $q = 2$, we add an auxiliary variable and follow the method in [27] to find the approximated solution. The details of each case is as following.

3.2.1 Solving cumulative histograms F_i

Fix u_i , the minimization problem (3.2) with respect to P_i can be simplified as:

$$\min_{P_i} \left\{ E_1 = \sum_{l=1}^L \int_{\Omega} W_p(P_{x,r}, P_i)^p u_i^q dx \right\}. \tag{3.7}$$

Case $p = 1$:

In this case, we are minimizing:

$$\min_{F_i} \left\{ E'_1 = \sum_{i=1}^N \int_{\Omega} \int_0^1 |F_{x,r}(l) - F_i(l)| u_i^q dl dx \right\}.$$

For each bar $y = 1, 2, \dots, L$, the minimizer is given by:

$$F_i(l) = \text{weighted (by } u_i^q(x) \text{) median of } F_{x,r}(l). \tag{3.8}$$

As for the details of how to compute the weighted median of a cumulative histogram, please refer to [36].

Case $p = 2$:

For each bar $y = 1, 2, \dots, L$ in the normalized histogram P_i , the minimization problem E_1 is similar to that with the Chan-Vese error function:

$$\min_{F_i} \left\{ E''_1 = \sum_{i=1}^N \int_{\Omega} \int_0^1 |F_{x,r}(l) - F_i(l)|^2 u_i^q dl dx \right\}.$$

Take the derivative of E''_1 with respect to $F_i(l)$ and setting it to zeros, we can achieve the closed form solution as:

$$F_i(l) = \frac{\int_{\Omega} F_{x,r}(l) u_i^q dx}{\int_{\Omega} u_i^q dx}. \tag{3.9}$$

3.2.2 Solving membership function u_i

Fixing F_i , the minimization problem (3.2) with respect to u_i becomes:

$$\min \left\{ E_2(u_i) = \frac{\lambda}{q} \sum_{i=1}^N \int_{\Omega} W_p(P_x, P_i)^p u_i^q dl dx + \sum_{i=1}^N \int_{\Omega} |\nabla u_i^n| dx \right\} \tag{3.10}$$

subject to

- (a) $\sum_{i=1}^N u_i(x) - 1 = 0,$
- (b) $0 \leq u_i(x) \leq 1, \text{ for } i = 1, \dots, N.$

Case $q = 1$:

In this case, we are minimizing indeed the relaxed Potts model subject to (a) and (b):

$$\min_{u_i} \left\{ E'_2(u_i) = \lambda \sum_{i=1}^{N-1} \int_{\Omega} W_p(P_x, P_i)^p u_i dx + \sum_{i=1}^N \int_{\Omega} |\nabla u_i^n| dx \right\}. \tag{3.11}$$

It has been proved in [5, 9] that the above minimization problem is equivalent to the continuous max-flow model and share the same global minimizer. Thus we can rewrite (3.11) as its primal-dual formulation:

$$\begin{aligned} \max_{p_s, p, q} \min_u \left\{ E'_{2,dual}(p_s, p, q; u) = \int_{\Omega} p_s dx + \sum_{i=1}^N \int_{\Omega} u_i (\operatorname{div} q_i - p_s + p_i) dx \right\} \\ \text{s.t. } p_i(x) \leq W_p(P_x, P_i)^p, \quad |q_i(x)| \leq \frac{1}{\lambda}, \quad i = 1, \dots, N, \end{aligned} \tag{3.12}$$

where p_s denotes the unconstrained source flow, $p_i(x)$ and $q_i(x)$ denotes the constrained sink flow and spatial flow respectively. Then the augmented Lagrangian formulation can be defined as:

$$\begin{aligned} \min_u \max_{p_s, p, q} \left\{ E'_{2,dual}(p_s, p, q; u) = \int_{\Omega} p_s dx + \sum_{i=1}^N \langle u_i, \operatorname{div} q_i - p_s + p_i \rangle \right. \\ \left. - \frac{c}{2} \sum_{i=1}^N \|\operatorname{div} q_i - p_s + p_i\|^2 \right\}, \end{aligned} \tag{3.13}$$

where $c > 0$. Note that the min and max operators can be interchanged because the conditions of the min-max theorem are all satisfied. Multiplier-based max-flow algorithm [47] can be adopted to find the final $u_i(x)$.

Case $q = 2$:

In this case, we can add an auxiliary variable v_i to approximate u_i , thus we are minimizing:

$$\min_{u_i} \left\{ E''_2(u_i) = \frac{\lambda}{2} \sum_{i=1}^N \int_{\Omega} W_p(P_x, P_i)^p u_i^2 dx + \frac{1}{2\theta} \sum_{i=1}^N \int_{\Omega} (v_i - u_i)^2 dx \right\} \tag{3.14}$$

subject to

- (a) $\sum_{i=1}^N u_i(x) - 1 = 0,$
- (b) $0 \leq u_i(x) \leq 1, \quad \text{for } i = 1, \dots, N.$

In this case, the minimization problem with respect to v_i is:

$$\int_{\Omega} |\nabla v_i| dx + \frac{1}{2\theta} \int_{\Omega} (v_i - u_i)^2 dx. \tag{3.15}$$

Applying the Chambolle's fast dual projection algorithm [13], this problem can be solved by:

$$v_i = u_i - \theta \operatorname{div} p_i^*, \quad i = 1, \dots, N, \quad (3.16)$$

where the vector p_i^* can be solved by the fixed point method: Initialize $p_i^0 = 0$ and iterate:

$$p_i^{n+1} = \frac{p_i^n + \tau \nabla (\operatorname{div} p_i^n - u_i / \theta)}{1 + \tau |\nabla (\operatorname{div} p_i^n - u_i / \theta)|}$$

with $\tau \leq 1/8$ to ensure convergence.

We can find an approximate numerical solution for the membership function u_i by considering the relaxation of the original problem without the constraint (b), by adding point-wise Lagrange multipliers $\delta(x)$, thus giving the approximate solution for the relaxation problem as (details refer to the method III in [28]):

$$u_i = \frac{v_i}{1 + \lambda \theta r_i} - \frac{\sum_{j=1}^N \frac{v_j}{1 + \lambda \theta r_j} - 1}{\sum_{j=1}^N \frac{1 + \lambda \theta r_i}{1 + \lambda \theta r_j}}.$$

And then apply the inequality constraints by projecting u_i on $[0, 1]$,

$$\hat{u}_i := \min\{\max\{u_i, 0\}, 1\}. \quad (3.17)$$

Although \hat{u}_i here gives an approximate numerical solution for the membership function u_i due to the difficulty to find an exact solution, we experimentally found that this approximated \hat{u}_i could obtain an acceptable segmentation results already. Besides, we also found that the model with $p = 2$ and $q = 2$ gives the best performance for most cases. In Section 4, the comparison of the results between these four models and other previous mentioned algorithms are shown.

3.2.3 Summary algorithm

Here we can summarize the minimization algorithm as the following procedures:

Algorithm 3.1.

1. Calculate the Gabor-filtered responses of the original image $I(x, y)$ with a set of k Gabor filters $G_{\sigma, f, \gamma}$ with different frequencies f or orientations γ or scales σ . Denote the smallest and the greatest magnitude of the responses as M_{\min} and M_{\max} separately. Divide the interval between M_{\min} and M_{\max} into L sub-intervals. Each sub-interval value would build a bar in the histogram.
2. For each pixel (x, y) compute the counts of magnitude of responses belong to each sub-interval, in a local region $\mathcal{N}_r(x, y)$ centered at (x, y) with pre-determined radius r , to build a histogram $P_{(x, y), r}$ with L bars. Normalize it and compute the cumulative histogram $F_{(x, y), r}$. Hence a histogram map can be obtained finally for the original image.
3. Initialize the membership function u_i by random matrices where each entry follows a uniform distribution in $[0, 1]$ and then normalize them such that membership constraints (a) and (b) are both satisfied.

4. Iteratively update each bar value in the histogram $F_i^k(l)$, $l = 1, 2, \dots, L$ by formula (3.8) when $p = 1$ and by (3.9) when $p = 2$; update v_i^k by formula (3.16); update u_i^k by solving (3.13) when $q = 1$, or by formula (3.17) when $q = 2$.
5. Repeat updating iteration till $\sum_{l=1}^L \|F^k - F^{k-1}\| \leq \epsilon$, where $\|\cdot\|$ denotes the Euclidean distance and ϵ is a small positive number defined by the user.

4 Numerical results

4.1 Gabor filter selection

Theoretically the full Gabor feature space should be calculated from a filter bank with different frequencies, orientations and scales. It can be imagined that the resulting histograms have three dimensions, like a tensor of histograms, with each dimension represents the histogram of Gabor filter responses with one varying parameter (say frequencies) and the other two parameters (say orientations and scales) fixed. It would be computationally costly to do so. To avoid such a time-costly prestage, many related Gabor filter design techniques could be applied. However, whether to choose the best Gabor filter bank is not our greatest concern, since the essence of our method is to build a larger histogram containing more information about how each texture phase responds with the varying feature (orientation or frequency or scale). In our experiments, we fix the orientation of the Gabor filter at some certain angle and set the frequency f ranging from 0.01 to 1 in step length of 0.01, thus building the final histogram with 100 bars. As for the choice of the orientation γ , one can follow the simple scheme in [30] to get a hint, which seeks for the biggest Fisher ratio separability measure of features between different texture phases. As for the segmentation iterations, the initial fuzzy membership was randomly uniform distributed within $[0, 1]$, the stopping criterion $\epsilon = 1 \times 10^{-6}$. The local histograms of the Gabor feature space need to be calculated for only once before the segmentation iterations.

4.2 Numerical examples

Firstly we test our model (3.2) with $p = 2$, $q = 2$ on two-phase, three-phase and four-phase synthetic texture images, as shown in Figs. 2-4, where the segmentation results are represented by contours. The parameters need to be determined are the frequency f , the orientation γ , the radius R of the Gaussian kernel in the Gabor filter and two regularization parameters λ and θ . The local histograms are usually computed in a neighborhood block region $\mathcal{N}_r(x)$ centered at each x with the size of $(2r+1) \times (2r+1)$. For simplicity, we fix $f = 0.01:0.01:1$, $R = 6$, $r = 5$, $\lambda = 0.05$, $\theta = 0.1$ in most cases, thus only the orientation γ need to be tuned for each test image. Experimentally we found that this parameter setting could achieve an reasonable segmentation result, except for some certain cases where different sizes of texture patterns are included in the same phase, as in the Fig. 8. In order to illustrate the efficiency of our model, the average computational time for segmenting the

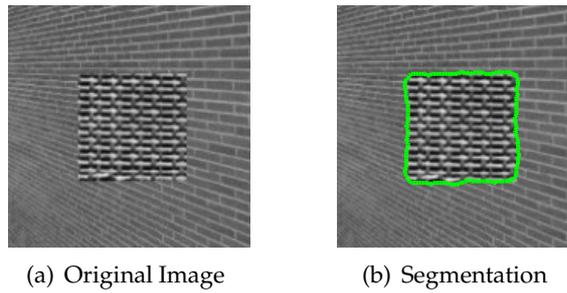


Figure 2: Two-phase texture segmentation (size: 160×160).

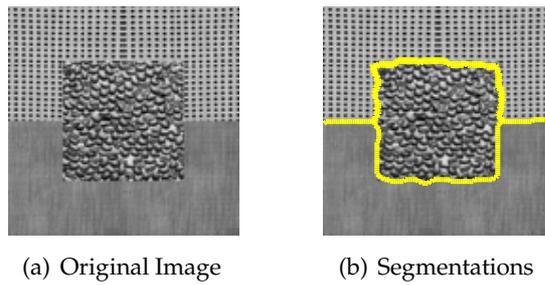


Figure 3: Three-phase texture segmentation (size: 256×256).

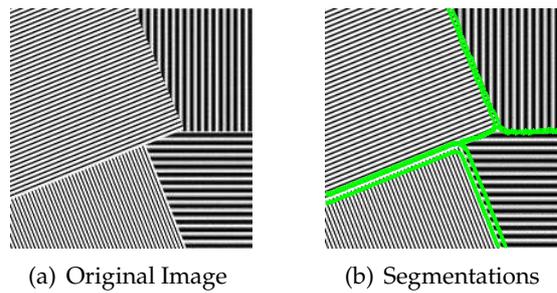


Figure 4: Four-phase texture segmentation (size: 256×256).

same two-phase texture image with different scales is shown in Table 1. The experiments are implemented on computer with Intel Core i7-2600 CPU 3.40GHz (8 Cores) and 16GB RAM.

Table 1: Scales and average computational time.

Avg. Comp. Time	Models	
	$p=2, q=2$	$p=2, q=1$
Scales		
64×64	0.13s	0.05s
128×128	0.82s	0.14s
192×192	2.11s	0.31s
256×256	4.33s	1.24s

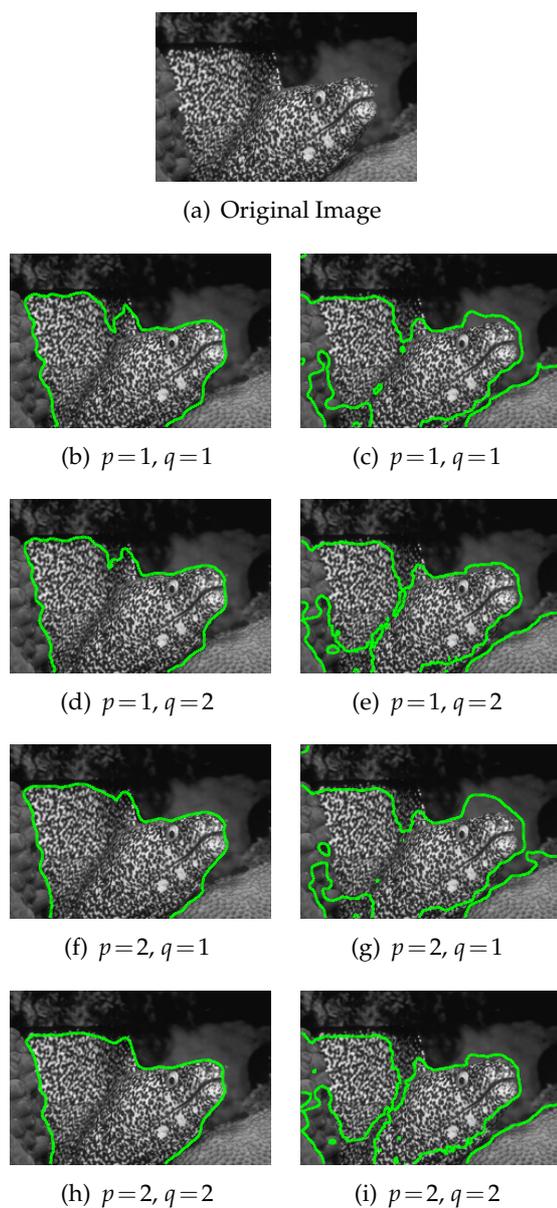


Figure 5: Performance Comparison based on Gabor feature space (left column) and Intensity (right column). (The parameters are tuned differently for each case trying to achieve the best visual performance).

We didn't show the computational time for the other two models with $p=1$ because sometimes these models using the linear 1-Wasserstein distance might not lead to a satisfying segmentation results. To illustrate the difference of the four models with $p=1$ or $p=2$ and $q=1$ or $q=2$, we show the comparison of the segmentation performance on a natural image corresponding to the four models based on intensity and Gabor feature

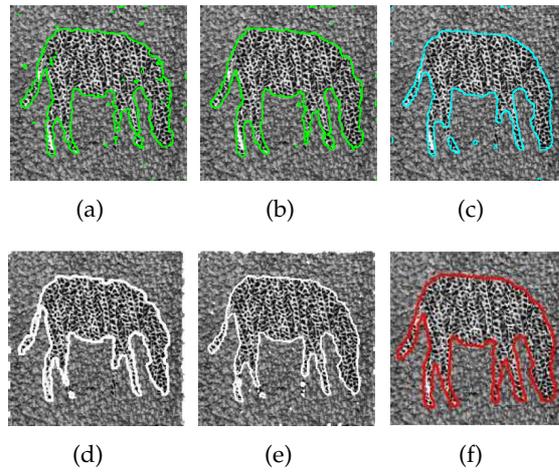


Figure 6: (a), (b): Our model with $p=2, q=1$ and $p=2, q=2$; (c): Result by F. Li and M. Ng [29]; (d): Result in C. Sagiv et al. 2006 [42]; (e): Result in N. Houhou et al. 2008 [18]; (f): Result in N. Houhou et al. 2009 [19].

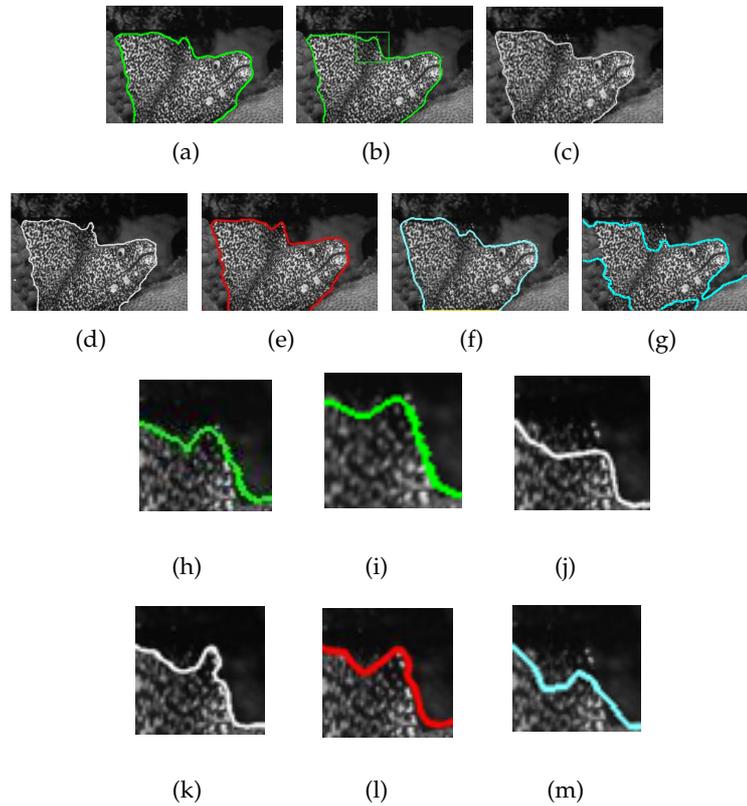


Figure 7: (a), (b): Our model with $p=2, q=1$ and $p=2, q=2$; (c): Result in C. Sagiv et al. 2006 [42]; (d): Result in N. Houhou et al. 2008 [18]; (e): Result in N. Houhou et al. 2009 [19]; (f): Result in X. Xie 2010 [47]; (g): Result by F. Li and M. Ng [29]; (h)-(m): Comparison of details selected from the same position of the fish corresponding to (a)-(f).

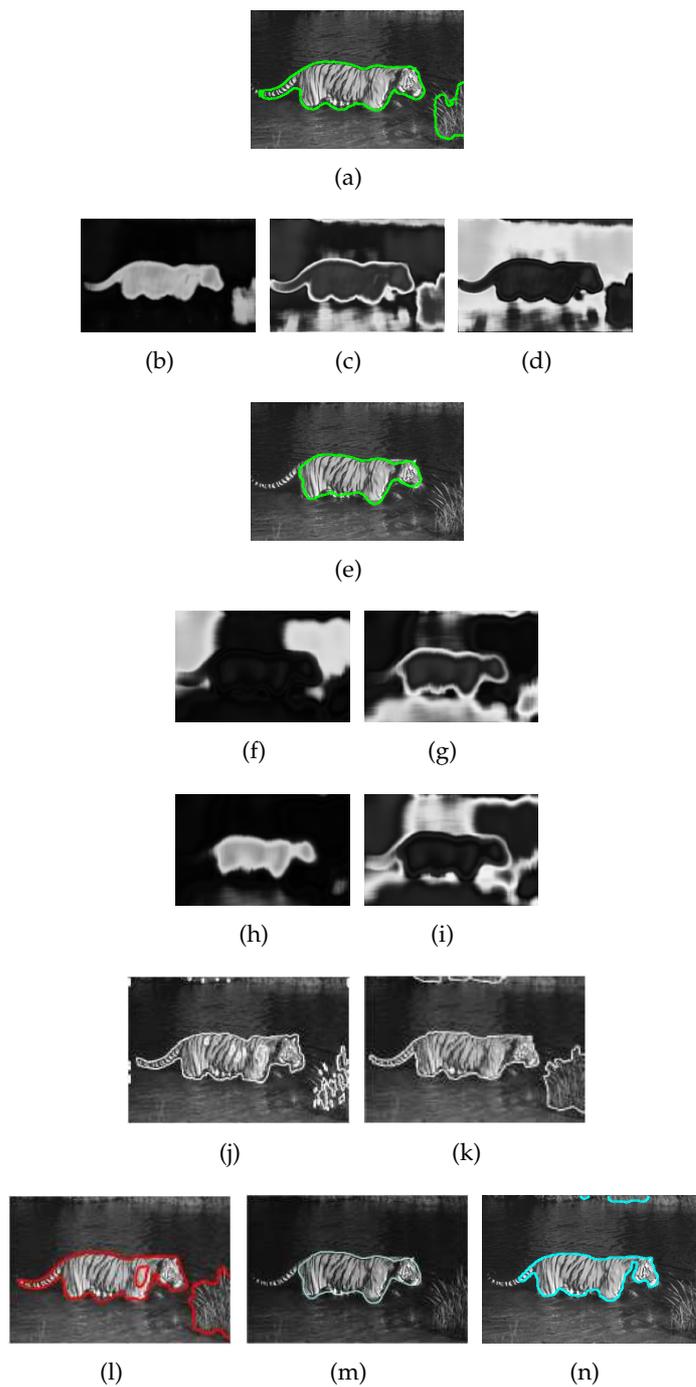


Figure 8: (a)-(d): Result by our model assuming $N=3$ and the resulting fuzzy membership; (e)-(i): Result by our model assuming $N=4$ and the resulting fuzzy membership; (j): Result in C. Sagiv et al. 2006 [42]; (k): Result in N. Houhou et al. 2008 [18]; (l): Result in N. Houhou et al. 2009 [19]; (m): Result in Kangyu Ni et al. 2009 [36]; (n): Result by F. Li and M. Ng [29].

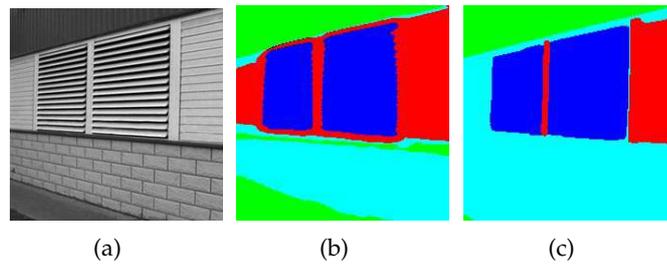


Figure 9: (a): Original Image; (b): 4-phase segmentation results by our method; (c): 4-phase segmentation results by F. Li and M. Ng [29].

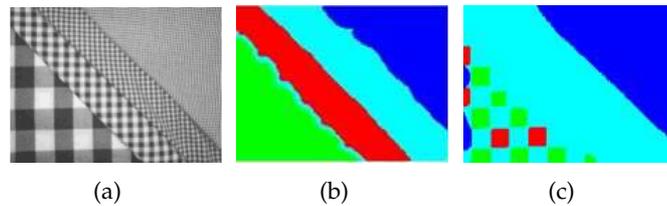


Figure 10: (a): Original Image; (b): 4-phase segmentation results by our method; (c): 4-phase segmentation results by F. Li and M. Ng [29].

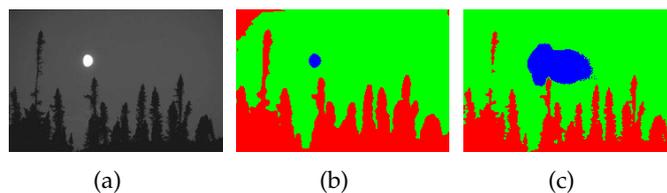


Figure 11: (a): Original Image; (b): 3-phase segmentation results by our method; Parameters setting: (c): 3phase segmentation results by F. Li and M. Ng [29].

space separately in Fig. 5, where (a) is the original image, the left column (b), (d), (f), (h) are the segmentation results based on the Gabor feature space and the right column (c), (e), (g), (i) based on the intensity. Overall speaking, the segmentation performances based on Gabor feature space surpass that on intensity due to that the Gabor transform itself has the ability to distinguish the texture pattern of the fish body from that of the coral at the bottom right corner. From the left column, we can find that the model (f), (h) using the squared 2–Wasserstein distance would lead to better results than that (b), (d) using the linear 1–Wasserstein distance. In terms of the choice of $q=1$ or $q=2$, the performance (f) and (h) are close to each other while setting $q=2$ would give more flexibility to modify the segmentation a little by choosing different threshold for the resulting fuzzy membership.

Nextly, we test our models on some natural images out from the Berkeley segmentation database *BSDS 500* [1] as shown in Figs. 5-8 and some multi-phase natural texture image in Figs. 9-10. The competitive results illustrate the advantages of using the squared 2–Wasserstein distance and the Gabor feature space.

5 Conclusions and discussion

We developed a multi-phase segmentation model using the squared 2–Wasserstein distance based on the Gabor feature space, which combines the advantages of the Gabor filter and the fuzzy region competition model both of which are good tools for texture image segmentation. The computational time are shorten due to that the closed form solution of $F_i(l)$ is to take the mean value instead of the weighted median in the case of using linear 1–Wasserstein distance. Besides, the existence proof is given for our model.

The segmentation performance are competitive to others within an acceptable computational cost. However, we only apply our model on the Gabor feature space with varying frequencies and the other two parameters are manually tuned and fixed. Of course the full Gabor feature space is also useable and beneficial for enhancement of the performance, but might increased the computational cost. So it still remains in our future work. Another question worthy of consideration is that the use of non-convex regularizer on the membership function, which is hopeful to preserve the geometric shapes of the homogeneous regions and prevent the over-smoothing problem.

References

- [1] The Berkeley Segmentation Dataset and Benchmark, <http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/>.
- [2] A. Braides and G. D. Maso, Non-local approximation of the Mumford-Shah functional, *Calculus of Variations and Partial Differential Equations*, 5(4) (1997), 293–322.
- [3] X. Bresson, P. Vandergheynst and J. P. Thiran, A variational model for object segmetation using boundary information and shape prior driven by the Mumford-Shah functional, *Int. J. Comput. Vision*, 68(2) (2006), 145–162.
- [4] X. Bresson, S. Esedoglu, P. Vandergheynst, J. P. Tharan and S. Osher, Fast global minimization of the active contour/snake model, *J. Math. Imaging Vision*, 28 (2007), 151–167.
- [5] E. Bae, J. Yuan and X-C. Tai, Global minimization for continuous multiphase partitioning problems using a dual approach, *Int. J. Comput. Vision*, 92(1) (2011), 112–129.
- [6] V. Caselles, Geometric models for active contours, *International Conference on Image Processing*, 3 (1995), 9–12.
- [7] V. Caselles, R. Kimmel and G. Sapiro, Geodesic active contours, *Int. J. Cumput. Vision*, 22(1) (1997), 61–79.
- [8] T. Chan and L. Vese, Active contours without edges, *IEEE Tran. Image Process.*, 10(2) (2001).
- [9] T. Chan, S. Esedoglu and M. Nikolova, Algorithms for finding global minimizers of image segmentation and denoising models, *SIAM J. Appl. Math.*, 66(5) (2006), 1632–1648.
- [10] T. Chan, B. Sandberg and L. Vese, Active contours without edges for vector-valued images, *J. Visual Commun. Image Representation*, 11(2) (2000), 130–141.
- [11] T. Chan, S. Esedoglu and K. Ni, Histogram based segmentation using wasserstein distances, scale space and variational methods in computer vision, *Lecture Notes in Computer Science*, 4485 (2007), 697–708.
- [12] A. Chambolle, Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations, *SIAM J. Appl. Math.*, 55(3) (1995), 827–863.

- [13] A. Chambolle, An algorithm for total variation minimization and application, *J. Math. Image Vision*, 20 (2004), 89–97.
- [14] G. Chung and L. Vese, Image segmentation using a multilayer level-set approach, *Comput. Visual. Science*, 12(6) (2009), 267–285.
- [15] L. D. Cohen, On active contour models and balloons, *CVGIP: Image Understanding*, 53(2) (1991), 211–218.
- [16] T. Georgiou, O. Michailovich, Y. Rathi, J. Malcolm and A. Tannenbaum, Distribution metrics and image segmentation, *Linear Algebra and Its Applications*, 405 (2007), 663–672.
- [17] M. Gobbino, Finite difference approximation of the Mumford-Shah functional, *Commun. Pure Appl. Math.*, 51(2) (1998), 197–228.
- [18] R. Goldenberg, R. Kimmel, E. Rivlin and M. Rudzsky, Fast geodesic active contours, *IEEE Tran. Image Process.*, 10(10) (2001), 1467–1475.
- [19] N. Houhou and J. P. Thiran, Fast texture segmentation model based on the shape operator and active contour, *IEEE Conference on Computer Vision and Pattern Recognition*, 1–8, 2008.
- [20] N. Houhou, J. P. Thiran and X. Bresson, Fast texture segmentation based on semi-local region descriptor and active contour, *Numer. Math. Theory Meth. Appl.*, 2(4) (2009), 445–468.
- [21] Y. Huang, M. K. Ng and Y. W. Wen, A fast total variation minimization method for image restoration, *Multiscale Model. Simul.*, 7 (2008), 774–795.
- [22] A. K. Jain and F. Farrokhnia, Unsupervised texture segmentation using gabor filters, *Pattern Recognition*, 24(12) (1991), 1167–1186.
- [23] M. Kass, A. Witkin and D. Terzopoulos, Snakes: active contour models, *Int. J. Comput. Vision*, 1 (1988), 321–331.
- [24] L. Kantorovich, On the translocation of masses, *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, (37) (1942), 199–201.
- [25] S. Kichenassamy, Gradient flows and geometric active contour models, *5th International Conference on Computer Vision*, 810–815, 1995.
- [26] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, ISBN 978-0-13-101171-7, 1995.
- [27] J. Lie, M. Lysaker and X. C. Tai, A binary level set model and some applications to Mumford-Shah image segmentation, *IEEE Trans. Image Processing*, 15(5) (2006), 1171–1181.
- [28] J. Lie, M. Lysaker and X. C. Tai, Piecewise constant level set methods and image segmentation, *Scale Space and PDE Methods in Computer Vision*, 3459 (2005), 573–584.
- [29] F. Li, M. K. Ng, T. Y. Zeng and C. Shen, A multiphase image segmentation method based on fuzzy region competition, *SIAM J. Image. Sciences*, 3(3) (2010), 277–299.
- [30] F. Li and M. K. Ng, Kernel density estimation based multiphase fuzzy region competition method for texture image segmentation, *Commun. Comput. Phys.*, 8 (2010), 623–641.
- [31] W. Li, K. Mao, H. Zhang and T. Chai, Selection of gabor filters for improved texture feature extraction, *Proceedings of IEEE 17th International Conference on Image Processing*, 361–364, 2010.
- [32] L. Y. Ma and J. Yu, Texture segmentation based on local feature histograms, *18th IEEE International Conference on Image Processing, ICIP 2011*, 3349–3352, 2011.
- [33] O. Michailovich, Y. Rathi and A. Tannenbaum, Image segmentation using active contours driven by the Bhattacharyya gradient flow, *IEEE Tran. Image Process.*, 16(11) (2007).
- [34] B. Mory and R. Ardon, Variation segmentation using fuzzy region competition and local non-parametric probability density functions, *11th IEEE International Conference on Computer Vision, ICCV 2007*, 1–8, 2007.

- [35] B. Mory and R. Ardon, Fuzzy region competition: a convex two-phase segmentation framework, scale space and variational methods in computer vision, *Lecture Notes in Computer Science*, 4485 (2007), 214–226, 2007.
- [36] D. Mumford and J. Shah, Optimal approximation by piecewise smooth functions and associated variational problems, *Commun. Pure Appl. Math.*, 42 (1989), 577–685.
- [37] K. Ni, X. Bresson, T. Chan and S. Esedoglu, Local histogram based segmentation using the Wasserstein distance, *Int. J. Comput. Vision*, 84(1) (2009), 97–111.
- [38] S. Osher and J. A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation, *J. Comput. Phys.*, 79 (1998), 12–49.
- [39] N. Paragios and R. Deriche, Geodesic active regions and level set methods for supervised texture segmentation, *Int. J. Comput. Vision*, 46(3) (2002), 223–247.
- [40] S. Rachev and L. Ruschendorf, *Mass Transportation Problems, Vol. I: Theory, Vol. II: Applications Probability and Its Applications*, Springer-Verlag, New York, 1998.
- [41] B. Sandberg, T. Chan and L. Vese, A level-set and gabor based active contour algorithm for segmenting textured images, *UCLA Comput. Appl. Math.*, Rep. 02-39, 2002.
- [42] C. Sagiv, N. A. Sochen and Y. Y. Zeevi, Integrated active contours for texture segmentation, *IEEE Tran. Image Process.*, 15(6) (2006).
- [43] C. Sagiv, N. A. Sochen and Y. Y. Zeevi, Geodesic active contours applied to texture feature space, *Proc. Scale-Space and Morphology in Computer Vision*, M. Kerckhove, Ed. Berlin, Germany: Springer-Verlag, 2106 (2001), 344–352.
- [44] B. Thomas and C. Daniel, On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional, *Scale Space and Variational Methods in Computer Vision*, 4485 (2007), 203–213.
- [45] L. Vese and T. Chan, A multiphase level set framework for image segmentation using the Mumford and Shah Model, *Int. J. Comput. Vision*, 50(3) (2002), 271–293.
- [46] Y. Wang, Y. Xiong, L. Lv, H. Zhang, Z. Cao and D. Zhang, Vector-valued Chan-Vese model driven by local histogram for texture segmentation, *17th IEEE International Conference on Image Processing, ICIP 2010*, 645–648, 2010.
- [47] X. Xie, Level set based segmentation using local feature distribution, *2010 International Conference on Pattern Recognition*, 2010.
- [48] H. Yuan and X. P. Zhang, Texture image retrieval based on a gaussian mixture model and similarity measure using a kullback divergence, *IEEE International Conference on Multimedia and Expo*, 2004.
- [49] J. Yuan, E. Bae, X.C. Tai and Y. Boykov, A continuous max-flow approach to Potts model, *Computer Vision–ECCV 2010*, 379–392, 2010.
- [50] S. C. Zhu and A. L. Yuille, A unified theory for image segmentation: region competition and its analysis, *Harvard Robotics Laboratory Technical Report 95-7*, 1995.
- [51] S. C. Zhu and A. L. Yuille, Region competition: unifying snakes, region growing and Bayes/MDL for multi-band image segmentation, *IEEE Trnas. Pattern Anal. Machine Intell.*, 18 (1996), 884–900.