

Weberized Mumford-Shah Model with Bose-Einstein Photon Noise

JIANHONG SHEN AND YOON-MO JUNG *

Abstract.

Human vision works equally well in a large dynamic range of light intensities, from only a few photons to typical midday sunlight. Contributing to such remarkable flexibility is a famous law in perceptual (both visual and aural) psychology and psychophysics known as *Weber's Law*. The current paper develops a new segmentation model based on the integration of both Weber's Law and the celebrated Mumford–Shah segmentation model (*Comm. Pure Applied Math.*, **42**, pp. 577-685, 1989). Explained in details are issues concerning why the classical Mumford–Shah model lacks light adaptivity, and why its “weberized” version can more faithfully reflect human vision's superior segmentation capability in a variety of illuminance conditions from dawn to dusk. It is also argued that the popular Gaussian noise model is physically inappropriate for the weberization procedure. As a result, the intrinsic thermal noise of photon ensembles is introduced based on Bose and Einstein's distributions in quantum statistics, which turns out to be compatible with weberization both analytically and computationally. The current paper focuses on both the theory and computation of the weberized Mumford–Shah model with Bose–Einstein noise. In particular, Ambrosio-Tortorelli's Γ -convergence approximation theory is adapted (*Boll. Un. Mat. Ital.*, **6-B**, pp. 105-123, 1992), and stable numerical algorithms are developed for the associated pair of nonlinear Euler-Lagrange PDEs.

Key words. Weber's law, light adaptivity, retina, Mumford–Shah segmentation, Bose–Einstein distribution, noise, Bayesian, variational, free boundary, Γ -convergence, computational PDE.

AMS subject classifications. Primary: 94A08; Secondary: 68U10, 65K10.

1. Introduction and Motivations. The current work targets at integrating two facts into the celebrated segmentation model of Mumford and Shah [34]: Weber's Law in vision psychology and retinal physiology, and the Bose–Einstein statistics for photon ensembles.

Light adaptivity is a remarkable feature of human visual perception. Entering a relatively dimmer room from an outdoor stroll in the sun, few people ever feel much difference in image interpretation, although physically the average ambient light intensities have undergone a major change. This function has been quantified into a famous law in psychology and psychophysics due to Weber [47] and later Fechner [20], whose physiological foundations have also been extensively explored in contemporary mathematical biology [19, 31, 35, 38, 41, 44, 45]. Section 2 will give a further detailed account on the important implications of Weber's Law.

Light adaptivity has not yet drawn commensurate attention from the communities of image processing and vision modelling. It is partially due to the fact that many multimedia image processing tasks only involve *localized still* images, typically like those captured by a digital camera in a fixed environment at a fixed time. Similarly, many application-driven problems such as medical imaging only deal with images with very much *fixed* patterns of intensity distributions, often computed by inverse problem solvers in imaging or signal acquisition technologies.

For dynamic vision systems of either mobile robots or surveillance cameras, light adaptivity then becomes indispensable due to the constant changes of incident light in a dynamic environment. In [43],

*Shen and Jung are with the School of Mathematics, University of Minnesota, Minneapolis, MN 55455. ({jhshen, ymjung}@math.umn.edu); Tel: (612) 625-3570; Fax: (612) 626-2017.

Research is supported by (USA) NSF Program in Applied Mathematics under the grant number DMS-0202565.

the first author has made the preliminary attempt to introduce it into the total variation based denoising scheme by Rudin, Osher, and Fatemi [40]. In the current work, we take a more systematic and detailed step in integrating light adaptivity (i.e., Weber’s Law) into the celebrated segmentation model of Mumford and Shah [34].

Segmentation is a fundamental task in both machine and human vision. To segment is to partition an image scene Ω into disjoint regions corresponding to individual objects, built upon which all vision tasks of higher levels can be further developed using tools from pattern analysis and recognition [26] and learning theory [37].

In the deterministic (or *cartoonish* [34]) modelling, David Mumford and Jayant Shah formulated the segmentation problem as: given an image observation sample u_0 which might be degraded by noise, a cartoonish image u and an edge set Γ are to be sought so that on each connected component of $\Omega \setminus \Gamma$, u is smooth or stationary. The solution to segmentation is therefore modelled as an estimation problem, which inevitably falls into the general framework of Bayesian inference [22, 33]:

$$\max_{u, \Gamma} p(u, \Gamma \mid u_0) = \frac{p(u, \Gamma, u_0)}{p(u_0)} = p(\Gamma)p(u \mid \Gamma)p(u_0 \mid u, \Gamma)\frac{1}{p(u_0)}.$$

For the sake of estimation, the probability of the observation $p(u_0)$ is simply a fixed normalization constant. By taking the logarithmic likelihood $E[\cdot] = -\log p(\cdot)$, or formally the “energy” according to Gibbs’ canonical ensembles [23, 33], one arrives at the deterministic segmentation formulation

$$\min_{u, \Gamma} E[u, \Gamma \mid u_0] = E[\Gamma] + E[u \mid \Gamma] + E[u_0 \mid u, \Gamma].$$

Mumford and Shah modelled the three terms by

$$E_{\text{ms}}[u, \Gamma \mid u_0] = \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx dy + \lambda \int_{\Omega} (u - u_0)^2 dx dy, \quad (1.1)$$

where \mathcal{H}^1 stands for 1-D Hausdorff measure, and α, β , and λ are weighting parameters, and will be called *visual potentials* in the present paper¹. This Mumford-Shah free-boundary model has attracted a great deal of attention from the communities of partial differential equations (PDE), calculus of variations, geometric measure theory, and computational mathematics [1, 2, 3, 4, 7, 8, 15, 17, 24]. Some interesting new applications can also be found in, e.g., [10, 11, 13, 14, 18, 42, 46].

The following proposition first shows that the Mumford–Shah model lacks light adaptivity.

PROPOSITION 1.1 (Gray scale shift-invariance of the Mumford–Shah model). *Suppose that*

$$(u_*, \Gamma_*) = \operatorname{argmin}_{u, \Gamma} E_{\text{ms}}[u, \Gamma \mid u_0]$$

is an optimal estimator of the Mumford–Shah model. Then for any constant gray scale shift θ ,

$$(u_* + \theta, \Gamma_*) = \operatorname{argmin}_{u, \Gamma} E_{\text{ms}}[u, \Gamma \mid u_0 + \theta].$$

It immediately implies that an edge can be detected in a dimmer image scene u_0 if and only if when it can be in a brighter scene, say $u_0 + 100$. In particular, for example, a same intensity jump of

¹Physically, the parameters are like pressure, surface tension, or chemical potentials in statistical mechanics [16, 23].

strength $\delta u_0 = 10$ is perceived with no difference no matter if it is embedded in a dimmer environment of $u_0 \approx 20$ or a brighter one of $u_0 \approx 200$. These conclusions are inappropriate according to Weber’s Law on human visual perception and the principle of light adaptivity.

Following the initial attempt of the first author in [43], the current work intends to introduce light adaptivity into the Mumford–Shah segmentation model (1.1) by incorporating Weber’s Law. The new model is called the *weberized Mumford–Shah segmentation model*, where the procedure “weberization” was first introduced in [43].

Furthermore, since weberization necessarily requires to interpret images as light intensity distributions, the shift-invariant Gaussian noise $E[u_0 | u, \Gamma]$ in the original Mumford–Shah model becomes physically unjustifiable (though mathematically still a good approximation). For instance, in the Gaussian noise model

$$u_0(x, y) = u(x, y) + n(x, y) \quad \text{with } n \text{ subject to the normal distribution } N(0, \sigma^2),$$

the image observation u_0 could become negative (esp. in a low-luminance scene when u is small), which is incompatible with Weber’s Law where u_0 and u are both light intensities. By exploring the very physical nature of light, i.e., ensembles of photons subject to the Bose–Einstein statistics, we are able to propose a new noise model called *Bose–Einstein noise* which is compatible to weberization.

The organization of the paper goes as follows. Section 2 reviews the role and essence of Weber’s Law in both vision psychology and retinal physiology. Section 3 shows the lack of light adaptivity in the original Mumford–Shah model by analyzing a concrete family of ideal step edges in Theorem 3.1. The weberization procedure is developed in Section 4, together with the development of the physically compatible noise model derived from the Bose–Einstein statistics of photon ensembles. Section 5 then studies the admissible space, the existence theorem (in 1-D case), and the Euler-Lagrange equations of the weberized Mumford–Shah model with Bose–Einstein noise. The techniques employed therein are very similar to those for the classical model [2, 3, 4, 24, 17, 34]. In Section 6, Ambrosio and Tortorelli’s Γ -convergence approximation techniques [3, 4] are adapted to the new model, and a linearization based iterative algorithm is designed for integrating the system of nonlinear Euler-Lagrange equations. The performance of the new model, as well as its comparison with the classical Mumford–Shah model, are finally demonstrated through two examples.

2. Weber’s Law and Light Adaptivity at Retinas. In this section, we briefly introduce the background and roles of Weber’s Law in both psychology and physiology.

2.1. Weber’s Law in Psychology and Psychophysics. Weber’s Law was first qualitatively described by German psychologist E. H. Weber, and later quantified by the great psychologist G. T. Fechner, the father of modern psychophysics.

Generally, for both visual and aural perception, let u denote the strength of the mean-field background, and δu the difference or contrast between the background signal and a target signal embedded within. Let $T(\delta u | u)$ be the yes-no binary decision of an average human observer for detecting the target signal under a fixed background level u . Then psychological evidences show that $T(\cdot | u)$ behaves like a Heaviside function $H(\delta u - \delta_* u)$ for some threshold $\delta_* u$ depending on u :

$$T_u(\delta u | u) = H(\delta u - \delta_* u) = 1, \quad \delta u \geq \delta_* u; \quad 0, \quad \text{otherwise.}$$

For each fixed background level u , this threshold δ_*u is often called the just-noticeable-difference (JND) in psychology. Weber’s Law says that the JND is linearly proportional to the mean-field u .

Weber’s Law. There exists a constant W , so that for a large range of background mean field u ,

$$\frac{\delta_*u}{u} \equiv W. \tag{2.1}$$

W has been naturally called *Weber’s constant* or *Weber’s ratio* in the literature.

This *empirical* statement of Weber’s Law can get deepened in the light of perceptual modelling. First, the classical statement (2.1) implies that the detection function $T(\delta u | u)$ could also be expressed as

$$T(\delta u | u) = H\left(\frac{\delta u}{u} - W\right), \tag{2.2}$$

again based on the Heaviside 0-1 function. In the continuum or physiological setting (see next subsection), such binary output can be further mollified as follows.

Modified Weber’s Law. Let $T(\delta u | u)$ denote the continuous confidence function of perceptual target detection. Then there exists a monotonically increasing function $f(\delta w)$ on $\delta w \in (0, \infty)$, so that $f(0) = 0, f(\infty) = 1$, and $T(\delta u | u) = f(\delta u/u)$. Moreover, according to the classical statement of Weber’s Law in (2.1), f should bear a sharp transition near Weber’s constant $\delta w = W$.

Both statements clearly suggest that to human visual or aural perception, the relative contrast $\delta w = \delta u/u$ is a variable more suitable to work with, and can be considered as the net input to the perceptive decision algorithms that human subjects subconsciously employ.

A very enlightening inference from Weber’s Law was made by Shapley and Enroth-Cugell [41], and is also highlighted in Keener and Sneyd’s book [28].

Shapley and Enroth-Cugell’s Inference from Weber’s Law. Under Weber’s Law, what human vision perceives are the surface material properties of the objects in a visual field, instead of their reflected light.

From Darwin’s *Natural Selection* point of view, one can immediately feel the message encoded into Shapley and Enroth-Cugell’s inference. Vision systems that remain stable and insensitive to illuminance changes are apparently superior to those of the opposite. It is simply because in 3-D environments lighting conditions constantly vary from dawn to dusk, while objects mostly remain more stable. Such variation is more noticeable in forest environments where shadows of trees or tall bushes constantly change the apparent intensity values of moving animals. Being able to perceive objects instead of their reflectance therefore helps the survival of both preys and predators.

We now give a brief derivation of Shapley and Enroth-Cugell's inference and refer the reader to [28, 41] for further details. Consider a simple scene only consisting of a background and a target object. Further assume that the background surface as well as the target surface are planar, and the incident light strikes perpendicularly to the two. Let R_B and R_T denote the reflectance indices of the background and the target separately, and I the incident light intensity. Then the reflected light intensities (i.e., observations) u_B and u_T are given by

$$u_B = I \times R_B \quad \text{and} \quad u_T = I \times R_T.$$

According to Weber's Law, the visually perceived difference is given by

$$\frac{|u_B - u_T|}{u_B} = \frac{|IR_B - IR_T|}{IR_B} = \frac{|R_B - R_T|}{R_B},$$

which is indeed independent of the incident light!

2.2. Weber's Law as Modelled in Mathematical Biology. The contemporary theory in mathematical biology believes that Weber's law is realized at the retinal portion of the visual system, and is implemented by the biochemical mechanisms of the photoreceptors, i.e., via various ion channels or pumps across the membranes of rods and cones [19, 31, 35, 38, 41, 44, 45].

Let I_0 denote the steady background luminance (or light intensity), and $V_0(I_0)$ the corresponding steady membrane potential (of a tagged cone or rod). When I_0 is suddenly increased by an amount of δI to a new level I , the membrane voltage undergoes a transient change $V(t)$:

$$V(0) = V_0(I_0) \quad \text{and} \quad V(t \rightarrow \infty) = V_0(I),$$

and in between there is a peak or valley turning point. Let $V(I | I_0)$ denote the turning point potential during such a transition, with a possible normalization on the sign and base level. Then $V_0(I_0) = V(I_0 | I_0)$.

Based on the response measurement of a red cone of the turtle, Normann and Perlman in 1979 [35] (also reproduced in Keener and Sneyd [28]) obtained a remarkable experimental result that demonstrates a strong pattern well fitted by the Naka-Rushton equation [28]:

$$\frac{V(I|I_0)}{V_{\max}} = \frac{I}{I + A(I_0)}, \tag{2.3}$$

where V_{\max} denotes the saturated and constant potential level, and $A = A(I_0)$ the half-saturation level, so that when $I = A$, $V(I | I_0) = V_{\max}/2$. For convenience, we shall normalize the potential so that $V_{\max} = 1$.

Under the condition of *ideal* or *exact* adaptation: $V_0(I_0) \equiv \text{const.}$, we have

$$\text{const.} = V_0(I_0) = V(I_0 | I_0) = \frac{I_0}{I_0 + A(I_0)},$$

implying that $A(I_0) = aI_0$ for some constant a . In reality, the ideal adaptation condition is only an approximation [35].

Let $\delta I = I - I_0$ denote the contrast change. Then by (2.3), the peak potential response is

$$V(I|I_0) = \frac{I_0 + \delta I}{(a+1)I_0 + \delta I} = \frac{1 + \delta I/I_0}{a+1 + \delta I/I_0} = f(\delta I/I_0), \quad f(\delta w) = \frac{1 + \delta w}{a+1 + \delta w}.$$

This is consistent with the modified Weber's Law in Section 2, and provides a solid *physiological* foundation for an otherwise purely *psychological* law.

Define the net peak response by $\delta V = \delta V(\delta I | I_0) = V(I | I_0) - V_0(I_0)$. Then it is easy to establish

$$\delta V(\delta I | I_0) = \frac{a}{(a+1)^2} \frac{\delta I}{I_0} + O((\delta I/I_0)^2).$$

Thus the so called peak sensitivity $S = S(I_0)$ is

$$S(I_0) = \lim_{\delta I \rightarrow 0} \frac{\delta V}{\delta I} \propto \frac{1}{I_0},$$

which is another equivalent statement of Weber's Law. Response sensitivity decreases as the background luminance grows stronger, resulting in stable potential responses (i.e., without quick saturation of membrane potentials by intense bright light), as well as constant performance (i.e., without losing perceptive accuracy in dimmer environments) across a broad range of background luminance levels.

Mathematical modelling efforts in simulating the Weber's Law can be categorized into two classes: algorithmic simulations based on system designs [44, 45], and biochemical simulations based on the dynamics of ions channels and chemical reactions.

3. Mumford–Shah Segmentation Lacks Light Adaptivity. In the Mumford–Shah model presented in the Introduction:

$$\min_{u, \Gamma} E_{\text{ms}}[u, \Gamma | u_0] = \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx dy + \lambda \int_{\Omega} (u - u_0)^2 dx dy,$$

the penalty on intensity irregularity (i.e., the second term) is enforced solely based on the measurement of ∇u . The underlying assumption is that the gradients are perceived independent of the background mean fields. But this is inappropriate according to Weber's Law, which claims that light changes are perceived by human vision with an adaptive modulation by the ambient mean fields. Therefore, the effectiveness of the Mumford–Shah segmentation model is at risk in dimmer environments.

Besides Proposition 1.1 in the Introduction, we now analyze an intriguing example to further illustrate how the Mumford–Shah segmentation model lacks light adaptivity, as contrast to Weber's Law for human visual perception.

Suppose the image domain $\Omega = \mathbb{R} \times (0, 1)$, and the images to be segmented are a one-parameter family $(u_\delta | \delta > 0)$ of *ideal* vertical edges along the y -axis:

$$u_\delta(x, y) = 2\delta + \delta \times (2H(x) - 1) = \delta H(-x) + 3\delta H(x), \quad (3.1)$$

where $H(t)$ denotes the canonical 0-1 Heaviside transition function. The intensity jumps along $\Gamma_0 = \{0\} \times (0, 1)$ are

$$[u_\delta]_{\Gamma_0}(y) = u_\delta(0^+, y) - u_\delta(0^-, y) \equiv 2\delta.$$

On the other hand, the mean field near the edge is defined by a local averaging

$$\langle u_\delta \rangle = \frac{1}{2A} \int_{(-A, A) \times (0, 1)} u_\delta(x, y) dx dy = 2\delta,$$

which is independent of the windowing size A in this ideal setting.

To human visual perception, by Weber's Law, the strength of the edge Γ_0 for any given δ is measured by Weber's ratio:

$$\frac{[u_\delta]}{\langle u_\delta \rangle} = \frac{2\delta}{2\delta} \equiv 1, \quad \text{independent of } \delta !$$

Therefore, if it is assumed that Weber's threshold constant W in (2.2) is less than 1, human perception should always be able to report or detect the existence of the edge Γ_0 .

But this is not the case for the machine vision of a robot that implements Mumford–Shah's segmentation model:

$$\min E_{\text{ms}}[u, \Gamma \mid u_\delta] = \alpha \text{length}(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx dy + \lambda \int_{\Omega} (u - u_\delta)^2 dx dy. \quad (3.2)$$

THEOREM 3.1 (Mumford–Shah Model Lacks Light Adaptivity). *For any fixed set of positive visual potentials (α, β, λ) , there exists a critical parameter $\delta_c = \delta_c(\alpha, \beta, \lambda)$, such that for any weaker luminance $\delta < \delta_c$,*

$$E_{\text{ms}}[u, \Gamma_0 \mid u_\delta] > \min_v E_{\text{ms}}[v, \phi \mid u_\delta], \quad (3.3)$$

where ϕ denotes an empty edge set, and u any Sobolev output on $\Omega \setminus \Gamma$. That is, if the luminance gets weaker below certain level, the hypothesis of the existence of an edge is not favored in Mumford and Shah's segmentation model.

Proof. First notice that for any desirable segmentation (u, Γ_0) ,

$$E_{\text{ms}}[u, \Gamma_0 \mid u_\delta] \geq \alpha \text{length}(\Gamma_0) = \alpha.$$

Second, we investigate any compatible segmentation $v = v(x, y)$ that fails to report the presence of the ideal edge Γ_0 , for which the Mumford–Shah cost energy is

$$E_{\text{ms}}[v, \phi \mid u_\delta] = \beta \int_{\Omega} |\nabla v|^2 dx dy + \lambda \int_{\Omega} (v - u_\delta)^2 dx dy < \infty.$$

By the Schwarz inequality $A + B \geq 2\sqrt{AB}$,

$$\begin{aligned} E_{\text{ms}}[v, \phi \mid u_\delta] &\geq 2\sqrt{\beta\lambda} \int_{\Omega} |\nabla v| \times |v - u_\delta| dx dy \\ &= \sqrt{\beta\lambda} \left(\int_{\Omega^+} |\nabla(3\delta - v)|^2 dx dy + \int_{\Omega^-} |\nabla(v - \delta)|^2 dx dy \right), \end{aligned} \quad (3.4)$$

where $\Omega^+ = (0, \infty) \times (0, 1)$ and $\Omega^- = (-\infty, 0) \times (0, 1)$. Notice that the equality holds in the first line if almost surely, $A(x, y) \equiv B(x, y)$. By the vertical translation invariance of both the domain and image u_δ , it can be further assumed that $v = v(x)$ is essentially 1-dimensional and $\nabla = d/dx$. As a result,

$$E_{\text{ms}}[v, \phi \mid u_\delta] \geq \sqrt{\beta\lambda} \left(\int_0^\infty |\nabla(3\delta - v(x))|^2 dx + \int_{-\infty}^0 |\nabla(v(x) - \delta)|^2 dx \right). \quad (3.5)$$

On the other hand, the compatibility condition also implies that

$$\int_0^\infty (3\delta - v(x))^2 dx < \infty \quad \text{and} \quad \int_{-\infty}^0 (v(x) - \delta)^2 dx < \infty.$$

In particular, there is a sequence of (x_n) so that as $x_n \rightarrow +\infty$, $\lim_{n \rightarrow +\infty} v(x_n) = 3\delta$. Therefore, according to the definition of total variation,

$$\begin{aligned} \int_0^\infty |\nabla(3\delta - v(x))^2| dx &= \lim_{n \rightarrow \infty} \int_0^{x_n} |\nabla(3\delta - v(x))^2| dx \\ &\geq \lim_{n \rightarrow \infty} |(3\delta - c)^2 - (3\delta - v(x_n))^2| \\ &= (3\delta - c)^2, \quad n \rightarrow \infty, \end{aligned}$$

where $c = v(0)$. Notice that equality holds in the second line if $v(x)$ is monotonic.

Similarly, one can show that

$$\int_{-\infty}^0 |\nabla(v(x) - \delta)^2| dx \geq (c - \delta)^2,$$

and the equality holds if v is monotonic. Hence in combination, (3.5) leads to

$$E_{\text{ms}}[v, \phi \mid u_\delta] \geq \sqrt{\beta\lambda} ((c - \delta)^2 + (3\delta - c)^2).$$

Varying $c = v(0)$, one easily sees that the lowest value on the right hand side is obtained when $c = 2\delta = \langle u_\delta \rangle$. Therefore,

$$E_{\text{ms}}[v, \phi \mid u_\delta] \geq 2\sqrt{\beta\lambda} \delta^2. \quad (3.6)$$

On the other hand, the lower bound in (3.6) is indeed attainable. To see it, define a Sobolev function $v_*(x) \in H_{\text{loc}}^1(-\infty, \infty)$ by the following procedure. On $(0, \infty)$, $v_*(x)$ solves the linear initial value problem

$$v_*(0) = 2\delta \quad \text{and} \quad v'_*(x) = \sqrt{\lambda/\beta} (3\delta - v_*(x)), \quad x > 0 \quad (3.7)$$

and on $(-\infty, 0)$, $v_*(x)$ solves the backward initial value problem

$$v_*(0) = 2\delta \quad \text{and} \quad v'_*(x) = \sqrt{\lambda/\beta} (v_*(x) - \delta), \quad x < 0. \quad (3.8)$$

Then all the equality conditions are indeed met in all the preceding inequalities starting with (3.4). On the other hand, this optimal v_* can be easily solved explicitly:

$$v_*(x) = (3\delta - \delta e^{-\sqrt{\lambda/\beta} x}) H(x) + (\delta + \delta e^{\sqrt{\lambda/\beta} x}) H(-x).$$

Therefore we have proved that

$$\min_v E_{\text{ms}}[v, \phi \mid u_\delta] = E_{\text{ms}}[v_*, \phi \mid u_\delta] = 2\sqrt{\beta\lambda} \delta^2.$$

Finally, define the critical luminance level $\delta_c = \delta_c(\alpha, \beta, \lambda) = \sqrt{\alpha/(2\sqrt{\beta\lambda})}$. Then the inequality (3.3) holds for all $\delta < \delta_c$, which completes the proof. \square

4. Weberized Mumford–Shah Model with Bose–Einstein Photon Noise (WMS-BE).

In this section, we introduce the main model by incorporating Weber’s Law into the Mumford–Shah model. In addition, we introduce a new noise model based on the quantum statistics of photons - the Bose–Einstein statistics, that is compatible to the idea of weberizing the Mumford–Shah model, in particular, the interpretation of pixel values as *positive* light intensities.

4.1. Weberized Mumford–Shah Regularity. To introduce light adaptivity into the classical Mumford–Shah segmentation model, we propose to incorporate Weber’s Law through the following procedure.

The gradient at any pixel (x, y) is $|\nabla u(x, y)|h \simeq \sqrt{(\delta_{x,h}u(x, y))^2 + (\delta_{y,h}u(x, y))^2}$, where $h > 0$ denotes the grid size, and

$$\delta_{x,h}u(x, y) = u(x + h, y) - u(x, y) \quad \text{and} \quad \delta_{y,h}u(x, y) = u(x, y + h) - u(x, y)$$

are the contrasts along x and y . In many digital applications, h could be treated as a fixed constant.

To apply Weber’s Law in a localized setting, one attempts to have the local contrasts $\delta_{x,h}u$ or $\delta_{y,h}u$ modulated by the associated local mean fields. As well practiced in nonlinear image filtering and the theory of scale spaces (see Catté, Lions, Morel, and Coll [6] for example), the local mean field can be defined through Gaussian smoothing with some window size $\sigma \ll 1$. That is, define

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

and define the local mean field $\langle u \rangle$ at (x, y) to be

$$\langle u \rangle(x, y) = u_\sigma(x, y) = G_\sigma * u(x, y).$$

Here for convenience it has been assumed that the imaging domain $\Omega = \mathbb{R}^2$. For a finite domain, one may run the linear heat equation to the image and terminate the process at a stopping time T . Then according to Weber’s Law, the perceptual responses to the luminance contrasts $\delta_{x,h}u(x, y)$ and $\delta_{y,h}u(x, y)$ are $|\delta_{x,h}u(x, y)|/u_\sigma(x, y)$ and $|\delta_{y,h}u(x, y)|/u_\sigma(x, y)$. Thus the perceived gradient is

$$\sqrt{\frac{(\delta_{x,h}u)^2}{(u_\sigma)^2} + \frac{(\delta_{y,h}u)^2}{(u_\sigma)^2}} \simeq \frac{|\nabla u|}{u_\sigma} h, \quad h \ll 1.$$

As a result, this suggests that as a penalty for irregularities, the bulky energy $E[u | \Gamma]$ in the original Mumford–Shah model should be replaced by its perceived version

$$E_w[u | \Gamma] = \beta \int_{\Omega \setminus \Gamma} \frac{|\nabla u|^2}{u_\sigma^2} dx dy.$$

On the other hand, as long as $u \in L^\infty(\Omega \setminus \Gamma)$ (a property that will be justified later), $E_w[u | \Gamma] < \infty$ requires that u belongs to the Sobolev space $H^1(\Omega \setminus \Gamma)$, making unnecessary the local smoothing procedure $u \rightarrow u_\sigma = G_\sigma * u$ for robustly getting the localized mean field $\langle u \rangle$. Thus one could simply take $\langle u \rangle(x, y) = u(x, y)$ at each pixel, and refresh the weberized regularity to

$$E_w[u | \Gamma] = \beta \int_{\Omega \setminus \Gamma} \frac{|\nabla u|^2}{u^2} dx dy. \tag{4.1}$$

4.2. Compatible Noise Model: Bose–Einstein Statistics. Parallel to the weberization of the bulky regularity energy $E[u | \Gamma] \rightarrow E_w[u | \Gamma]$, it is also necessary to update the noise model or the fidelity cost function $E[u_0 | u]$.

In Weber’s Law, the signal u represents light intensity, a nonnegative quantity well defined in both wave and particle mechanics. In particular, for example, $u = 0$ corresponds to the very specific and important physical meaning that there is no single photon or light at all (the *black hole* condition, as called by the first author in [43]). Therefore, the phase space of u is not translation invariant and the associated noise model must demonstrate such characteristics.

Due to the *central limit theorem*, Gaussian white noise has been the most popular noise model in the literature and has led to the *least square fitting* energy:

$$E_2[u_0 | u, \Gamma] = E_2[u_0 | u] = \lambda \int_{\Omega} (u - u_0)^2 dx dy.$$

Many works [9, 39, 40] have also studied the multiplicative noise model

$$u_0(x, y) = u(x, y) \times (1 + n(x, y)),$$

with n again modelled by Gaussian. As a result, the fitting energy becomes [39, 40]

$$E_{\times}[u_0 | u] = \lambda \int_{\Omega} \left(\frac{u_0}{u} - 1 \right)^2 dx dy.$$

Mathematically and computationally convenient, the physical foundations of both these noise models are unclear, however. In particular, there have been no physical mechanisms of light (or photons) that can reproduce these noise models when the u -values are interpreted as light intensities.

Our approach to noise modeling below is based on the very nature of photons, or the Bose–Einstein statistics, and therefore is *physically* consistent with the weberization procedure. Moreover, later analysis and computation show that it indeed leads to a well behaved new segmentation model.

By the quantum theory of (free) electromagnetic field, the Hamiltonian can be decomposed into a sum of singleton harmonic oscillators, each of which is identified by a specific base frequency ω_k , $k = 1, 2, \dots$, or equivalently, quantal energy $\varepsilon_k = \hbar\omega_k$, where \hbar is the Plank constant. The energy spectra of an ε_k -harmonic oscillator, modulated by its ground state $\varepsilon_k/2$, consist of

$$0, \varepsilon_k, 2\varepsilon_k, \dots, n \varepsilon_k, \dots.$$

In terms of the basic photon ε_k , the quantal energy $n\varepsilon_k$ can be interpreted as a group of n particles.

A complete description of a state of the light is therefore specified by the so called *occupation numbers*: $\mathbf{u} = (u_1, u_2, \dots) \in \mathbb{Z}_+^{\infty}$, where \mathbb{Z}_+ denotes all nonnegative integers, and each u_k counts the number of ε_k -photons. Since photons are bosons, each u_k can indeed be any nonnegative integers, as contrast to the fermions (e.g., electrons), for which u_k can only be either 1 or 0, representing the on-off states of being occupied or not [16, 23].

At thermal equilibrium with a fixed temperature T , an ensemble of photons obey the Bose–Einstein statistics. That is, for any state \mathbf{u} with total energy

$$E[\mathbf{u}] = \sum_{k=1}^{\infty} u_k \varepsilon_k = \hbar \sum_{k=1}^{\infty} u_k \omega_k,$$

the Bose–Einstein probability $p(\mathbf{u})$ is analogous to Gibbs’ canonical ensemble:

$$p(\mathbf{u}) = \frac{1}{Z_\beta} \exp(-\beta E[\mathbf{u}]),$$

where $\beta = 1/(kT)$ is the inverse temperature normalized by the Boltzmann constant k . Therefore the partition function is given by

$$\begin{aligned} Z_\beta &= \sum_{\mathbf{u} \in \mathbb{Z}_+^\infty} \exp(-\beta E[\mathbf{u}]) = \sum_{\mathbf{u} \in \mathbb{Z}_+^\infty} \prod_{k=1}^{\infty} \exp(-\beta \varepsilon_k u_k) \\ &= \prod_{k=1}^{\infty} \sum_{u_k=0}^{\infty} \exp(-\beta \varepsilon_k u_k) = \prod_{k=1}^{\infty} \frac{1}{1 - e^{-\beta \varepsilon_k}}. \end{aligned} \quad (4.2)$$

Hence, for any ε_k -species of photons, the average occupation number is

$$\langle u_k \rangle = \frac{\partial(-\ln Z_\beta)}{\partial(\beta \varepsilon_k)} = \frac{1}{e^{\beta \varepsilon_k} - 1}, \quad k = 1, 2, \dots. \quad (4.3)$$

As a simplified model, let us now consider monochromatic light or an ensemble of photons of a same species with quantal frequency ω_0 and energy $\varepsilon_0 = \hbar\omega_0$. Let u_0 denote the observed number of photons each time, and $u = \langle u_0 \rangle$ its mean.

At thermal equilibrium with temperature T , the Bose–Einstein distribution becomes

$$p(u_0) = \frac{1}{Z_\beta} \exp(-\beta \varepsilon_0 u_0) = \frac{1}{Z_\beta} \exp\left(-\frac{\varepsilon_0 u_0}{kT}\right), \quad u_0 = 0, 1, \dots.$$

And the partition function and the average number are

$$Z_\beta = \frac{1}{1 - e^{-\beta \varepsilon_0}} \quad \text{and} \quad u = \langle u_0 \rangle = \frac{1}{e^{\beta \varepsilon_0} - 1}. \quad (4.4)$$

Therefore,

$$e^{-\beta \varepsilon_0} = \frac{u}{1 + u} \quad \text{and} \quad Z_\beta = \left(1 - \frac{u}{1 + u}\right)^{-1} = 1 + u.$$

When the light is not too weak so that the mean number of photons $u \gg 1$, one has $\beta \varepsilon_0 \ll 1$ and

$$u \simeq \frac{1}{\beta \varepsilon_0} \quad \text{and} \quad p(u_0) = p(u_0 | u) \simeq \frac{1}{u} \exp\left(-\frac{u_0}{u}\right).$$

In particular, in the continuum limit, one can assume that $u_0 \in (0, \infty)$, and the above is precisely the exponential distribution with average observation u .

Therefore, the logarithmic likelihood function is given by

$$e(u_0 | u) = -\ln p(u_0 | u) = \ln u + \frac{u_0}{u}.$$

For a given imaging plane or lattice Ω , the sum of this likelihood function over the entire domain (assuming free photons without spatial interactions) yields the new fidelity energy

$$E_{\text{be}}[u_0 | u] = \lambda \int_{\Omega} e(u_0(x, y) | u(x, y)) dx dy = \lambda \int_{\Omega} \left(\ln u(x, y) + \frac{u_0(x, y)}{u(x, y)} \right) dx dy, \quad (4.5)$$

which is different from the popular Gaussian or least square models either additive or multiplicative. We shall call it the Bose–Einstein noise model.

Notice that the form of $e(u_0 | u)$ does require the purity of the photon species at a target sensor. However across the entire imaging plane or domain Ω , the independent summation in E_{be} does not demand a spatial uniformity of the species. That is, near two distinct sensors A and B , the photon species can be ω_A and ω_B with $\omega_A \neq \omega_B$, which makes E_{be} more plausible in approximating real situations. Notice that in the wave theory of light, all the above can be similarly explained via the stationary phase method in asymptotic analysis (see, e.g., [5]).

4.3. The Model: Weberized Mumford–Shah with Bose–Einstein Noise. In combination of the preceding two subsections, we thereby propose the following weberized Mumford–Shah model with Bose–Einstein noise:

$$\begin{aligned} E_{\text{wmsbe}}[u, \Gamma | u_0] &= E[\Gamma] + E_{\text{w}}[u | \Gamma] + E_{\text{be}}[u_0 | u] \\ &= \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} \frac{|\nabla u|^2}{u^2} dx dy + \lambda \int_{\Omega} \left(\ln u + \frac{u_0}{u} \right) dx dy. \end{aligned} \quad (4.6)$$

Unlike the classical Mumford–Shah model, here the images u and u_0 are both understood as light intensity distributions as in Weber’s Law, or equivalently, both count the number of photons. In particular, we must have $u(x, y), u_0(x, y) > 0$, on Ω .

Or more conveniently, define the logarithmic intensities

$$w(x, y) = \ln u(x, y) \quad \text{and} \quad w_0(x, y) = \ln u_0(x, y).$$

Then the weberized Mumford–Shah energy naturally applies to w and w_0 :

$$\begin{aligned} E_{\text{wmsbe}}[w, \Gamma | w_0] &= E[\Gamma] + E_{\text{w}}[w | \Gamma] + E_{\text{be}}[w_0 | w] \\ &= \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla w|^2 dx dy + \lambda \int_{\Omega} (w + e^{w_0 - w}) dx dy. \end{aligned} \quad (4.7)$$

Notice that w and w_0 can now both take negative values. In fact the last model (4.7) looks more similar to the original Mumford–Shah model except for the noise model and the logarithmic transform.

We have thus successfully completed the modelling stage, and the rest of the paper will focus on the analysis and computation of the model (4.6) or (4.7). One can benefit profoundly from the rich existing literature on the classical Mumford–Shah model [2, 3, 4, 17, 24].

5. Study of the WMS-BE Model.

5.1. Admissible Spaces. From the last term on Bose–Einstein noise in the new model (4.7), it is natural to assume that $w(x, y) \in L^1(\Omega)$. On the other hand, as in the classical Mumford–Shah model, the first two terms on the logarithmic prior suggest that w belongs to $\text{SBV}(\Omega)$, the special functions of bounded variations: $Dw = \nabla w + [w] \times \mathcal{H}^1|_{\Gamma}$, where Dw denotes the vectorial Radon measure associated with the distributional derivative of w , ∇w the Lebesgue continuous part (or the Sobolev component), and $[w]$ the jump along the oriented jump set $\Gamma = \Gamma_w$, which is well defined almost surely according to the 1-D Hausdorff measure.

The logarithmic transform of the noise model also requires $e^{w_0(x,y)-w(x,y)} \in L^1(\Omega)$. By the principle of L^p duality, it is therefore natural (pragmatic, convenient, but not necessarily unique) to assume that

$$e^{w_0} \in L^\infty(\Omega) \quad \text{and} \quad e^w \in L^1(\Omega).$$

Consequently, $w_0 \in L^\infty(\Omega)$. The next theorem shows that one can simply demand $w \in L^\infty(\Omega)$ as a result.

THEOREM 5.1. *Let $w_*(x, y)$ be a minimizer to the weberized Mumford–Shah model $E_{\text{wmsbe}}[w, \Gamma \mid w_0]$ (4.7) with Bose–Einstein noise. Assume that the logarithmic light intensity observation $w_0 \in L^\infty(\Omega)$. Then $w_* \in L^\infty(\Omega)$ and $\|w_*\|_\infty \leq \|w_0\|_\infty$.*

Proof. For any real parameter a , define a nonlinear function by

$$\psi(t \mid a) = t + e^{a-t}, \quad t \in \mathbb{R}. \quad \text{Then,}$$

$$\psi'(t \mid a) = 1 - e^{a-t} \quad \text{and} \quad \psi''(t \mid a) = e^{a-t}.$$

In particular, $\psi(\cdot \mid a)$ is strictly convex for any fixed a . Since $\psi(\pm\infty) = +\infty$, and $\psi'(a \mid a) = 0$, we conclude that $t_* = a$ is the unique minimizer of $\psi(\cdot \mid a)$ for any a . As a result, for any t_1, t_2 with

$$(t_1 - a)(t_2 - a) \geq 0 \quad \text{and} \quad |t_1 - a| \leq |t_2 - a|,$$

one has $\psi(t_1 \mid a) \leq \psi(t_2 \mid a)$ and the inequality holds if and only if $t_1 = t_2$.

Now assume $\|w_0\|_\infty = A$ and define the truncated function from the minimizer w_* by

$$[w_*]_A(x, y) = w_*(x, y) \times \left(1 \wedge \frac{A}{|w_*(x, y)|} \right), \quad (x, y) \in \Omega.$$

Following the preceding preparation, at any pixel (x, y) , identifying

$$a = w_0(x, y), \quad t_1 = [w_*]_A(x, y), \quad \text{and} \quad t_2 = w_*(x, y),$$

one arrives at

$$\psi([w_*]_A(x, y) \mid w_0(x, y)) \leq \psi(w_*(x, y) \mid w_0(x, y)),$$

as well as

$$\int_{\Omega} \psi([w_*]_A(x, y) \mid w_0(x, y)) dx dy \leq \int_{\Omega} \psi(w_*(x, y) \mid w_0(x, y)) dx dy. \quad (5.1)$$

The equality in the last equation holds if and only if *a.e.*,

$$[w_*]_A(x, y) = w_*(x, y), \quad \text{or equivalently,} \quad \|w_*\|_\infty \leq A \quad (= \|w_0\|_\infty).$$

On the other hand, according to [4, 24], it is sufficient to work with $w_* \in C^1(\Omega \setminus \Gamma)$. In particular, one can further assume that A is a *regular* value of w_* , meaning that the level sets $w_* \equiv \pm A$ are 1-D manifolds. (Otherwise simply replace A by some $A + \varepsilon$ with $|\varepsilon| \rightarrow 0$. This is guaranteed by the

denseness and openness of regular values in differential topology [32].) Then, as a Radon measure on $\Omega \setminus \Gamma$ (whose notation will be omitted in the next two lines for simplicity), one has

$$\begin{aligned} D[w_*]_A &= D[w_*]_A|_{\{|w_*|<A\}} + D[w_*]_A|_{\{|w_*|>A\}} + D[w_*]_A|_{\{|w_*|=A\}} \\ &= \nabla w_*|_{\{|w_*|<A\}}. \end{aligned}$$

The second term vanishes since $[w_*]_A$ is constant on the open set $|w_*| > A$, while the last term vanishes because w_* , and therefore $[w_*]_A$, are continuous along the 1-D manifolds $|w_*| = A$ (see, e.g., [25]). Remind that here all terms are restricted on $\Omega \setminus \Gamma$. Therefore,

$$\int_{\Omega \setminus \Gamma} |\nabla [w_*]_A|^2 dx dy \leq \int_{\Omega \setminus \Gamma} |\nabla w_*|^2 dx dy. \quad (5.2)$$

In combination of (5.1) and (5.2), one arrives at

$$E[[w_*]_A, \Gamma \mid w_0] \leq E[w_*, \Gamma \mid w_0].$$

Since w_* is a minimizer, the equality must hold. Then the theorem directly follows from the equality condition of (5.1). \square

Notice that in the proof one could circumvent the C^1 characterization of [24] by turning to the continuity of the truncation operator $[\cdot]_A$ (both weak and *a.s.*), mollification approximations of $H^1(\Omega \setminus \Gamma)$, and the weak lower semi-continuity properties of Sobolev norms [29]. The proof also clearly demonstrates the good behavior of the Bose–Einstein noise for mathematical analysis.

5.2. Existence Theorem. The remarkable works [2, 3, 4, 24, 17] on the existence of solutions to the classical Mumford–Shah segmentation can be naturally adapted to the weberized Mumford–Shah model with Bose–Einstein noise. For completeness, in what follows, we prove the existence of E_{wmsbe} in the 1-D case. After some necessary adaptation being made, the techniques are nothing very different from the existing literature. For much more involved high dimensional cases, we refer to the aforementioned works in the classical literature.

We assume that the image domain is the unit interval $I = (0, 1)$, and denote by Γ_u or equivalently Γ_w the jump set of an admissible image u on I .

THEOREM 5.2 (Existence of Optimal Segmentation by WMS-BE). *Suppose $w_0 = \ln u_0 \in L^\infty(I)$. Define $w = \ln u$, and*

$$\mathcal{D} = \{u(x) : \|w\|_\infty \leq \|w_0\|_\infty, \#\Gamma_w < \infty, w \in H^1(I \setminus \Gamma_w)\}.$$

Then the minimizer in \mathcal{D} exists to the weberized Mumford–Shah model with Bose–Einstein noise:

$$E_{\text{wmsbe}}[u, \Gamma \mid u_0] = \alpha \#\Gamma + \beta \int_{I \setminus \Gamma} \frac{|\nabla u|^2}{u^2} dx + \lambda \int_I \left(\ln u + \frac{u_0}{u} \right) dx,$$

or equivalently, its logarithmic version $E_{\text{wmsbe}}[w, \Gamma \mid w_0]$.

Notice that the restriction $\|w\|_\infty \leq \|w_0\|_\infty$ has been well justified by Theorem 5.1.

Proof.

- (a) First notice that $E_{\text{wmsbe}}[w, \phi \mid w_0] < \infty$ for $w \equiv 0 \in \mathcal{D}$. Thus one can assume that there is a (logarithmic) minimizing sequence $\{w_n\}_{n=1}^\infty \subset \mathcal{D}$ whose E_{wmsbe} energies are bounded by some finite value, say, M .
- (b) Let Γ_n denote the associated jump set of w_n . Then $\{\#\Gamma_n\}$ is uniformly bounded. With a possible subsequence refinement, one could assume that $\#\Gamma_n \equiv N$ for $n = 1, 2, \dots$, and that

$$\Gamma_n = \{x_1^{(n)} < x_2^{(n)} < \dots < x_N^{(n)}\} \subset I.$$

By the pre-compactness of bounded sequences in \mathbb{R} , with possibly another round of subsequence refinement, one could assume that there exists a limiting set (possibly a multiset)

$$\Gamma^* = \{x_1^* \leq x_2^* \leq \dots \leq x_N^*\} \subset \bar{I} = [0, 1],$$

and for each $j = 1 : N$, $x_j^{(n)} \rightarrow x_j^*$ as $n \rightarrow \infty$.

- (c) For any $\varepsilon > 0$, define the (relatively) closed ε -neighborhood of Γ^* in $I = (0, 1)$ by

$$\Gamma_\varepsilon^* = \{y \in I : \text{dist}(y, \Gamma^*) \leq \varepsilon\},$$

and its open complement $I_\varepsilon = I \setminus \Gamma_\varepsilon^*$. Then as $\varepsilon \rightarrow 0$, I_ε monotonically expands to I . For any fixed ε , there exists an n_ε , so that for any $n > n_\varepsilon$, w_n 's are bounded in $H^1(I_\varepsilon)$ since

$$\|w_n\|_{H^1(I_\varepsilon)}^2 = \int_{I_\varepsilon} w_n^2(x) dx + \int_{I_\varepsilon} (w_n'(x))^2 dx \leq \|w_0\|_\infty^2 + \frac{M}{\beta},$$

as long as Γ_n starts to be included within Γ_ε^* .

- (d) Choose any sequence (ε_k) with $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$. For ε_1 , by the compactness of $H^1(I_{\varepsilon_1})$ in $L^2(I_{\varepsilon_1})$ [30], one could refine a subsequence of $(n)_{n=1}^\infty$, denoted by $(n \mid 1)_{n=1}^\infty$, so that $(w_{(n \mid 1)})_{n=1}^\infty$ is a Cauchy sequence in $L^2(I_{\varepsilon_1})$.

Repeating this process, for each step $k \geq 2$, one could find a subsequence $(n \mid k)$ of $(n \mid k - 1)$ so that $(w_{(n \mid k)})_{n=1}^\infty$ is a Cauchy sequence in $L^2(I_{\varepsilon_k})$. Finally, define $w_k^* = w_{(k \mid k)}$, $k = 1 : \infty$.

- (e) Then it is clear there that there exists a unique $w^* \in L_{\text{loc}}^2(I \setminus \Gamma^*)$, so that $w_k^* \rightarrow w^*$ in any $L^2(I_\varepsilon)$. In particular,

$$\|w^*\|_\infty \leq \|w_0\|_\infty.$$

Furthermore, by the L^2 lower semi-continuity of Sobolev norms, on any I_ε ,

$$\int_{I \setminus \Gamma^*} |(w^*)'|^2 dx = \sup_\varepsilon \int_{I_\varepsilon} |(w^*)'|^2 dx \leq \sup_\varepsilon \liminf_{k \rightarrow \infty} \int_{I_\varepsilon} |(w_k^*)'|^2 dx \leq \liminf_{k \rightarrow \infty} \int_{I \setminus \Gamma_{(k \mid k)}^*} |(w_k^*)'|^2 dx.$$

Since (w_k^*) is a minimizing sequence, we conclude that $w^* \in H^1(I \setminus \Gamma^*)$, $\Gamma_{w^*} \subset \Gamma^*$, and $w^* \in \mathcal{D}$.

- (f) With a possible subsequence refinement, one could assume

$$w_k^*(x) \rightarrow w^*(x), \quad a.e. \quad x \in I.$$

Following the notation in the proof of Theorem 5.1, we then must have

$$\psi(w_k^*(x) \mid w_0(x)) \rightarrow \psi(w^*(x) \mid w_0(x)) \quad a.e. \quad x \in I,$$

as a result of the continuity of $\psi(t \mid a)$. Since w_0 and all w_k^* 's are bounded by $A = \|w_0\|_\infty$, so is ψ 's by

$$|\psi(w_k^*(x) \mid w_0(x))| \leq A + e^{2A}, \quad \forall k, x.$$

Then, by *Lebesgue's Dominated Convergence Theorem*,

$$\int_I \psi(w^* \mid w_0) dx = \lim_{k \rightarrow \infty} \int_I \psi(w_k^* \mid w_0) dx.$$

(g) In combination of the last two itemized results, we arrives at

$$E_{\text{wmsbe}}[w^*, \Gamma_{w^*} \mid w_0] \leq \liminf_{k \rightarrow \infty} E_{\text{wmsbe}}[w_k^*, \Gamma_{w_k^*} \mid w_0].$$

Since (w_k^*) is a minimizing sequence, $w^* \in \mathcal{D}$ has to be a minimizer, which concludes the proof.

□

5.3. Euler-Lagrange Equations. To derive the Euler-Lagrange equations, one defines two *conditional* cost functions associated with E_{wmsbe} , similar to the classical literature.

For a given jump set Γ , define the conditional bulk energy of w to be

$$E[w \mid \Gamma, w_0] = \int_{\Omega \setminus \Gamma} B(w, \nabla w \mid w_0) dx dy, \quad (5.3)$$

where the bulk function

$$B = B(w, \nabla w \mid w_0) = \beta |\nabla w|^2 + \lambda(w + e^{w_0 - w}).$$

Let $w_\Gamma = \operatorname{argmin} E[w \mid \Gamma, w_0]$ be the minimizer for the given jump set Γ , which is unique due to the strict convexity. Then $w_\Gamma \in H^1(\Omega \setminus \Gamma)$ must solve the following Euler-Lagrange equation in the distributional sense

$$0 = \frac{\partial B}{\partial w} - \nabla \cdot \left(\frac{\partial B}{\partial \nabla w} \right) = -2\beta \Delta w + \lambda(1 - e^{w_0 - w}), \quad (5.4)$$

on each connected components Ω_i of $\Omega \setminus \Gamma$, with Neumann adiabatic boundary condition $\partial w / \partial \vec{n} = 0$. Notice that $\partial \Omega_i \subset \partial \Omega \cup \Gamma$.

Next, for the current estimation $w = w_\Gamma$, to compute the variational effect caused by the perturbation of Γ , define the conditional energy

$$E[\Gamma \mid w, w_0] = E[\Gamma \mid B] = \alpha \mathcal{H}^1(\Gamma) + \int_{\Omega \setminus \Gamma} B(w, \nabla w \mid w_0) dx dy.$$

Assuming that Γ is piecewise C^1 , near any regular point $\mathbf{x}(s) \in \Gamma$ parameterized by the arc length s , a perturbation $\Gamma \rightarrow \Gamma_\varepsilon$ can be locally expressed by

$$\mathbf{x}(s) \rightarrow \mathbf{x}_\varepsilon(s) = \mathbf{x}(s) + \varepsilon(s)N(s), \quad |\varepsilon(s)| \ll 1,$$

where $N(s)$ is unit normal at $\mathbf{x}(s)$. Adapted to the perturbation of $\Gamma \rightarrow \Gamma_\varepsilon$ is the local necessary extension of the bulk function $B(x, y) = B(w, \nabla w \mid w_0) \rightarrow B^\varepsilon(x, y)$. Locally according to the

orientation of N , the region is segmented by Γ into the tail half and the head half, labelled by “−” and “+” separately. Accordingly, B is segmented into B_{\pm} components on each half. Following the perturbation of $\Gamma \rightarrow \Gamma_{\varepsilon}$, locally, B_{σ} with $\sigma(s) = -\text{sign}(\varepsilon(s))$ can be extended onto the new stripe territory by

$$B_{\sigma}^{\varepsilon}(\mathbf{x}(s) + tN(s)) = B_{\sigma}(\mathbf{x}(s)), \quad t \in [0, \varepsilon(s)).$$

The conditional energy $E[\Gamma \mid B]$ is naturally perturbed to $E[\Gamma_{\varepsilon} \mid B^{\varepsilon}]$, and the standard calculation then gives rise to

$$\delta E = E[\Gamma_{\varepsilon} \mid B^{\varepsilon}] - E[\Gamma \mid B] = (-\alpha\kappa N - [B]N) \cdot (\varepsilon N) + O(\varepsilon^2), \quad (5.5)$$

where $[B] = [B](s) = B_{+}(s) - B_{-}(s)$ denotes the jump across Γ , and κ the signed curvature, both coupled with N so that although N can have binary uncertainty, both κN and $[B]N$ are unique (and κN points to the center of curvature). Therefore, if (w, Γ) is a minimizer, the jump curve must satisfy

$$\alpha\kappa + [B(w, \nabla w \mid w_0)] = 0.$$

For the sake of numerical computation, as done in Chan and Vese [15], the variational formula (5.5) offers a strategy for local interface motion that is similar to the active contour technique [27]:

$$\frac{\partial \mathbf{x}(s, t)}{\partial t} = \alpha\kappa N + [B]N. \quad (5.6)$$

To the jump set or interface, this is a combination of two motion mechanisms: the mean curvature motion and the motion driven by an external bulk energy B .

The free boundary nature of the Mumford–Shah model makes the computation highly nontrivial. There are many remarkable works on effective implementations of the model, including the finite element or finite difference methods [7, 8], the level-set method of Osher and Sethian [15, 36, 46], and the Γ -convergence elliptic approximations first proposed by Ambrosio and Tortorelli [3, 4]. In what follows, for the weberized Mumford–Shah model with Bose–Einstein noise, we adopt the approach of Γ -convergence approximations and explore the corresponding computational issues.

6. Γ -Convergence Approximation and Numerical Computation.

6.1. Γ -Convergence Approximation of Ambrosio and Tortorelli. For Γ -convergence approximation, the edge set Γ is encoded by the so called *auxiliary* function $z = z(x, y)$, or the *canyon* function, which is close to 1 almost everywhere except near the edge set, where it sharply drops to zero or close. Thus, Γ looks like a canyon if z is understood as the height field of a territory. The rate as well as the amount of dropping along the edge set are controlled by an approximation parameter $\varepsilon \rightarrow 0$. Hence z depends on ε : $z = z_{\varepsilon}(x, y)$.

First, the length cost $E[\Gamma] = \alpha\mathcal{H}^1(\Gamma)$ is replaced by its phase-field approximation [3, 4]:

$$E_{\varepsilon}[z] = \alpha \int_{\Omega} \left(\varepsilon |\nabla z|^2 + \frac{(z-1)^2}{4\varepsilon} \right) dx dy. \quad (6.1)$$

Next, one approximates the weberized regularity energy $E_w[u \mid \Gamma]$ using the canyon function by

$$E_w[u \mid z] = \beta \int_{\Omega} (z^2 + o_{\varepsilon}) \frac{|\nabla u|^2}{u^2} dx dy, \quad (6.2)$$

where o_ε stands for any positive scalar parameter that vanishes faster than ε . The role of o_ε is to enforce the uniform ellipticity of the integral even near the edge set Γ where the well function z could be zero or close. Notice that in terms of the logarithmic luminance function $w = \ln u$, one has

$$E_w[w | z] = \beta \int_{\Omega} (z^2 + o_\varepsilon) |\nabla w|^2 dx dy, \quad (6.3)$$

Finally recall that the Bose–Einstein noise leads to the generative data model

$$E_{\text{be}}[w_0 | w] = \lambda \int_{\Omega} (w + e^{w_0 - w}) dx dy = \lambda \int_{\Omega} \psi(w | w_0) dx dy, \quad (6.4)$$

where $\psi(t | a) = t + \exp(a - t)$ as in the preceding sections.

In combination of (6.1) and (6.3), we propose the Γ -convergence approximation to the weberized Mumford–Shah model with Bose–Einstein noise

$$\begin{aligned} E_{\text{wmsbe}}[w, z | w_0, \varepsilon] &= E_\varepsilon[z] + E_w[w | z] + E_{\text{be}}[w_0 | w] \\ &= \alpha \int_{\Omega} \left(\varepsilon |\nabla z|^2 + \frac{(z-1)^2}{4\varepsilon} \right) dx dy \\ &\quad + \beta \int_{\Omega} (z^2 + o_\varepsilon) |\nabla w|^2 dx dy + \lambda \int_{\Omega} (w + e^{w_0 - w}) dx dy. \end{aligned} \quad (6.5)$$

Following Ambrosio and Tortorelli’s work on the classical Mumford–Shah model, it is possible to establish the existence and convergence theorem, whose proof will be omitted.

THEOREM 6.1. *Assume that $w_0 \in L^\infty(\Omega)$ as in the preceding theorems. Then*

- (a) *For any fixed $\varepsilon > 0$, the minimizer to $E_{\text{wmsbe}}[w, z | w_0, \varepsilon]$ exists.*
- (b) *As $\varepsilon \rightarrow 0$, the sequence of cost energies $E_{\text{wmsbe}}[w, z | w_0, \varepsilon]$ converge to $E_{\text{wmsbe}}[w, \Gamma | w_0]$ in the Γ -convergence sense.*

6.2. Euler-Lagrange Equations. To derive the Euler-Lagrange equations for (6.5), define the two conditional cost energies as in Section 5. First, for a fixed current estimation of the well function z , one defines the conditional energy

$$E_{\text{wmsbe}}[w | w_0, \varepsilon, z] = \beta \int_{\Omega} (z^2 + o_\varepsilon) |\nabla w|^2 dx dy + \lambda \int_{\Omega} (w + e^{w_0 - w}) dx dy. \quad (6.6)$$

Variation on w : $w \rightarrow w + \delta w$ in $H^1(\Omega)$ leads to the Euler-Lagrange equation

$$0 = -2\beta \nabla \cdot (z^2 + o_\varepsilon) \nabla w + \lambda (1 - e^{w_0 - w}), \quad (6.7)$$

with Neumann adiabatic conditions $\partial w / \partial \vec{n} = 0$ along $\partial \Omega$. Notice that unlike the Gaussian noise model, the Bose–Einstein noise model makes the Euler-Lagrange equation nonlinear.

For the current best estimation of the logarithmic image w , define the conditional cost energy

$$E_{\text{wmsbe}}[z | w_0, \varepsilon, w] = \alpha \int_{\Omega} \left(\varepsilon |\nabla z|^2 + \frac{(z-1)^2}{4\varepsilon} \right) dx dy + \beta \int_{\Omega} (z^2 + o_\varepsilon) |\nabla w|^2 dx dy. \quad (6.8)$$

Variation on z : $z \rightarrow z + \delta z$ in $H^1(\Omega)$ leads to the Euler-Lagrange equation:

$$0 = \alpha \left(-2\varepsilon \Delta z + \frac{z-1}{2\varepsilon} \right) + 2\beta z |\nabla w|^2, \quad (6.9)$$

similarly with Neumann boundary condition. Any minimizer $(w_\varepsilon, z_\varepsilon)$ to $E_{\text{wmsbe}}[w, z | w_0, \varepsilon]$ in $H^1(\Omega)$ therefore must satisfy (6.7) and (6.9) in the weak sense.

6.3. An Iterative Algorithm. In this section, we propose an iterative algorithm for solving the system of nonlinear Euler-Lagrange equations just established above.

For a given w , define the linear elliptic operator

$$L_w = -4\varepsilon^2 \Delta + \left(1 + \frac{4\beta\varepsilon}{\alpha} |\nabla w|^2\right).$$

Then the z -equation (6.9) simply becomes

$$L_w z(x, y) = 1 \quad (x, y) \in \Omega \quad (6.10)$$

with Neumann boundary condition.

For the w -equation (6.7), define a positive analytic function

$$\phi(t) = \frac{\lambda}{2\beta} \frac{1 - e^{-t}}{t} = \frac{\lambda}{2\beta} \int_0^1 e^{-tp} dp > 0.$$

Then (6.7) becomes

$$0 = -\nabla \cdot (z^2 + o_\varepsilon) \nabla w + (w - w_0) \phi(w - w_0) \quad (6.11)$$

Furthermore, for *given* z and w , define the positive linear elliptic operator

$$M_{z,,w} = -\nabla \cdot (z^2 + o_\varepsilon) \nabla + \phi(w - w_0).$$

Then (6.11) can be written as

$$M_{z,,w} w = \phi(w - w_0) w_0, \quad (6.12)$$

which of course becomes nonlinear for w since the operator $M_{z,,w}$ itself involves w .

The combination of the structures in (6.10) and (6.12) suggests a natural iterative algorithm. Starting with an initial guess $(w^{(0)}, z^{(0)})$, at step n , $(w^{(n+1)}, z^{(n+1)})$ solves the follow system of linear equations on (w, z) :

$$\begin{aligned} M_{z^{(n)}, w^{(n)}} w &= \phi(w^{(n)} - w_0) w_0, \\ L_{w^{(n)}} z &= 1, \end{aligned} \quad (6.13)$$

with Neumann boundary conditions for both w and z . This system of decoupled and linear elliptic equations can be efficiently integrated using any linear elliptic solvers, among which we adopt finite difference based iterative methods.

6.4. Numerical Examples and Comparison with the Mumford–Shah Model.

6.4.1. Γ convergence approximation to the Mumford–Shah model. To make comparison, we also need to establish the Γ -convergence approximation to the ordinary Mumford–Shah model, with either Gaussian noise or Bose–Einstein noise.

In the case of Bose–Einstein noise, the Mumford–Shah model becomes

$$E_{\text{msbe}}[u, \Gamma \mid w_0] = \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx dy + \lambda \int_{\Omega} \left(\ln u + \frac{u_0}{u} \right) dx dy,$$

whose Γ -convergence approximation could similarly be established:

$$\begin{aligned} E_{\text{msbe}}[u, z \mid u_0, \varepsilon] &= \alpha \int_{\Omega} \left(\varepsilon |\nabla z|^2 + \frac{(z-1)^2}{4\varepsilon} \right) dx dy \\ &+ \beta \int_{\Omega} (z^2 + o_\varepsilon) |\nabla u|^2 dx dy + \lambda \int_{\Omega} \left(\ln u + \frac{u_0}{u} \right) dx dy, \end{aligned} \quad (6.14)$$

Similar to the preceding section, one obtains the Euler-Lagrange equations:

$$\begin{aligned} 0 &= -2\beta \nabla \cdot (z^2 + o_\varepsilon) \nabla u + \frac{\lambda}{u^2} (u - u_0), \\ 0 &= \alpha \left(-2\varepsilon \Delta z + \frac{z-1}{2\varepsilon} \right) + 2\beta |\nabla u|^2 z, \end{aligned} \quad (6.15)$$

with Neumann adiabatic boundary conditions. Notice that without weberization, the equations are naturally about (u, z) instead of $(w = \ln u, z)$, as in contrast with Eqn. (6.7) and (6.9).

This system of nonlinear elliptic equations can be similarly integrated by iterations based on stepwise linearization. That is, at each step n , $(u^{(n+1)}, z^{(n+1)})$ solves the follow system on (u, z) :

$$\begin{aligned} 0 &= -2\beta \nabla \cdot ([z^{(n)}]^2 + o_\varepsilon) \nabla u + \frac{\lambda}{[u^{(n)}]^2} (u - u_0), \\ 0 &= \alpha \left(-2\varepsilon \Delta z + \frac{z-1}{2\varepsilon} \right) + 2\beta |\nabla u^{(n)}|^2 z. \end{aligned} \quad (6.16)$$

As before, o_ε denotes any scalar sequence that vanishes faster than ε as $\varepsilon \rightarrow 0^+$.

These are the equations to be solved for comparison with their weberized versions in (6.13).

In the case of Gaussian noises, the system of equations (6.15) and (6.16) still hold after the removal of u^2 or $[u^{(n)}]^2$ in the denominators associated with the fitting terms, which is of course a coupled linear system.

6.4.2. Strength of edges based on Γ -convergence approximation. Γ -convergence approximation also provides a natural way to define the *strength* of an optimal edge well function z_ε , which can be a useful measure in evaluating and comparing the performances of different models. To the best knowledge of the authors, we are the first to endow the edge well or wall function with such meaning.

We define the edge *strength* at bandwidth ε by

$$\hat{\eta}_\varepsilon = \int_{\Omega} |\nabla s_\varepsilon(x, y)| dx dy = \text{TV}(s_\varepsilon),$$

i.e., the total variation of the “wall” function $s_\varepsilon = (1 - z_\varepsilon)^2$. This has been primarily motivated by the analysis in the beginning of Section 6.1, in particular, the Schwarz inequality (??), as well as the construction technique in the proof by Ambrosio and Tortorelli [4].

For the optimal wall function s_ε , denote its height by $d_\varepsilon \in [0, 1]$. Then the edge strength is approximately given by

$$\hat{\eta}_\varepsilon = \text{TV}(s_\varepsilon) \simeq 2d_\varepsilon \text{length}(\Gamma), \quad (6.17)$$

which justifies $\hat{\eta}_\varepsilon$ as a measure of edge strength - larger d_ε values correspond to stronger edges. Here the factor 2 comes from the two sides of the “wall.” Approximation (6.17) at least holds locally along any stationary (i.e., with almost constant height) segment of the walls, and can be further justified by the celebrated co-area formula of De Giorgi [25], and Fleming and Rishel [21] on the total variation Radon measure (also see [12, 25]):

$$\mathrm{TV}(s_\varepsilon) = \int_{\Omega} |Ds_\varepsilon| = \int_0^1 \mathrm{length}(s_\varepsilon \equiv \lambda) d\lambda, \quad (\text{the co-area formula})$$

where $(s_\varepsilon \equiv \lambda)$ denotes the λ -level set of s_ε . For simplicity, s_ε has been assumed to be smooth, otherwise, the length measure should be replaced by the *perimeter* of the cumulative level sets $\{s_\varepsilon > \lambda\}$. Therefore, if the wall is steep in the transition from the ground value s_- to the peak value s_+ , as should be in the case of small ε , each level set $s_\varepsilon \equiv \lambda$ consists of two slightly translated copies of the edge curve Γ on either side of the wall. As a result,

$$\mathrm{TV}(s_\varepsilon) \simeq \int_{s_-}^{s_+} 2\mathrm{length}(\Gamma) d\lambda = 2(s_+ - s_-) \mathrm{length}(\Gamma),$$

which is precisely (6.17).

Practically, or at least locally along a stationary edge segment, it is also convenient to simply consider the *relative* strength

$$\eta_\varepsilon := \max_{(x,y) \in \Omega} z_\varepsilon(x,y) - \min_{(x,y) \in \Omega} z_\varepsilon(x,y),$$

which is proportional to $d_\varepsilon = s_+ - s_-$ just defined above. It could be considered as the absolute strength $\hat{\eta}_\varepsilon$ modulated by the length of the underlying segment according to (6.17), and can be useful in comparing the relative edge strengths output from different models (e.g., Figure 6.1 and 6.2). Visually, one simply inspects the darkness of the well functions along the targeted edges, since their maximal values are always close to 1 for small ε (as guaranteed by the term $(z - 1)^2/(4\varepsilon)$ in the Γ -convergence approximations). Thus, darker edges are stronger.

6.4.3. Two examples and comparisons. We illustrate the performance of our new model through two image examples: one synthetic and the other an image of a real 3-dimensional scene designed by the first author.

The synthetic image example in Figure 6.1 is to verify and highlight the light adaptivity of our new model, while the real image in Figure 6.2 is to compare the performances of our new model with the classical Mumford–Shah model (6.15) and (6.16) in terms of light adaptivity.

Both figures have been generated by Matlab codes, which implement the two algorithms (6.13) and (6.16). The Matlab codes for the several files involved, as well as the real image designed by the first author (in Figure 6.2), are available under request.

Both noises have been simulated by using Matlab’s random number generators, and Figure 6.1 involves simulated Gaussian noise while Figure 6.2 simulated Bose–Einstein noise. Notice that the least square fitting energy is employed in both models in the case of Gaussian noise.

Since it is better to read the interpretation of the computational results while staring at the figures, our comments, views, and explanations are detailed directly in the captions of the two figures.

From these examples, we conclude that

- (a) The *weberized* Mumford–Shah model indeed faithfully embodies Weber’s Law, and demonstrates clear light adaptivity parallel to the way that human vision functions.
- (b) Due to the gray-scale shift invariance, the classical Mumford–Shah model does not faithfully simulate light adaptivity of human visual perception as predicted by both Weber’s Law and retinal physiology [28, 20, 47, 45].

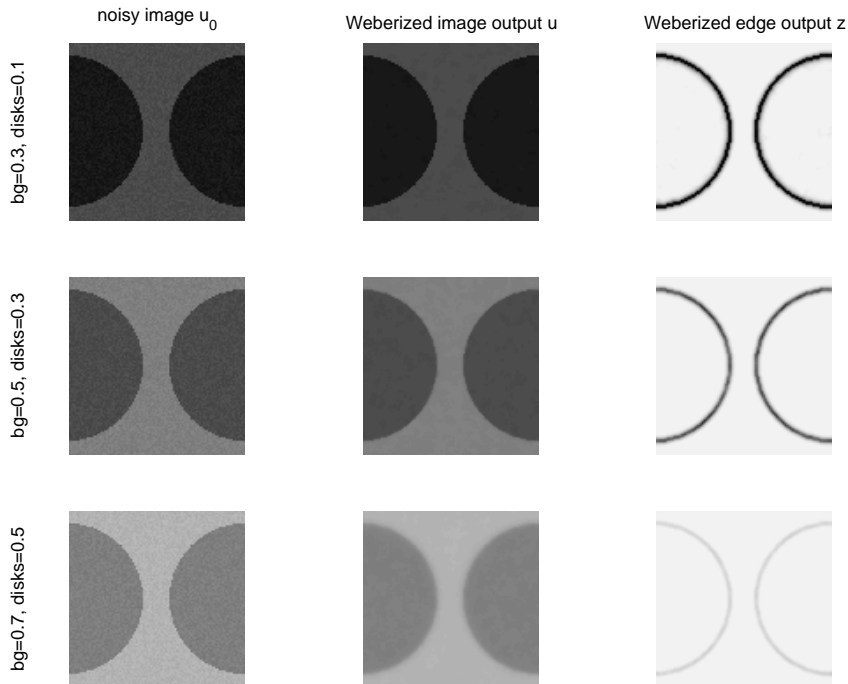


FIG. 6.1. *Performance via a synthetic image: light adaptivity of the weberized Mumford–Shah model (via Γ -convergence approximation). Each noisy image is obtained from the immediately above by a simple intensity increment of 0.2 everywhere in the scale of $[0,1]$. By the weberized Mumford–Shah model, as consistent with Weber’s Law, the edge functions z are distinct and have relative strengths $\eta_\varepsilon = \max z_\varepsilon - \min z_\varepsilon = 0.98, 0.75, 0.20$ separately. The classical Mumford–Shah model is contrast shift-invariant (see Proposition 1.1), and therefore makes no distinction among the three, which is less faithful in vision modelling according to the real perceptual responses characterized by Weber’s Law. All three examples are computed using a same set of visual potentials (α, β, λ) as well as the Γ -convergence bandwidth ε .*

7. Conclusion. Towards a more faithful model of human visual segmentation, the current paper introduces the light adaptivity feature into the celebrated Mumford–Shah segmentation model [34], as inspired by Weber’s Law in both vision psychology and retinal physiology [19, 20, 31, 35, 38, 41, 44, 45, 47]. To be consistent with the intensity (or photon counting) interpretation of images in the weberization procedure, the traditional gray scale shift-invariant Gaussian noise model is accordingly replaced by Bose–Einstein photon noise.

Both theory and computation have been developed for the weberized Mumford–Shah model with Bose–Einstein noise. In particular, Ambrosio and Tortorelli’s Γ -convergence approximation procedure has been adapted to the new model, and the resulting system of nonlinear Euler-Lagrange equations

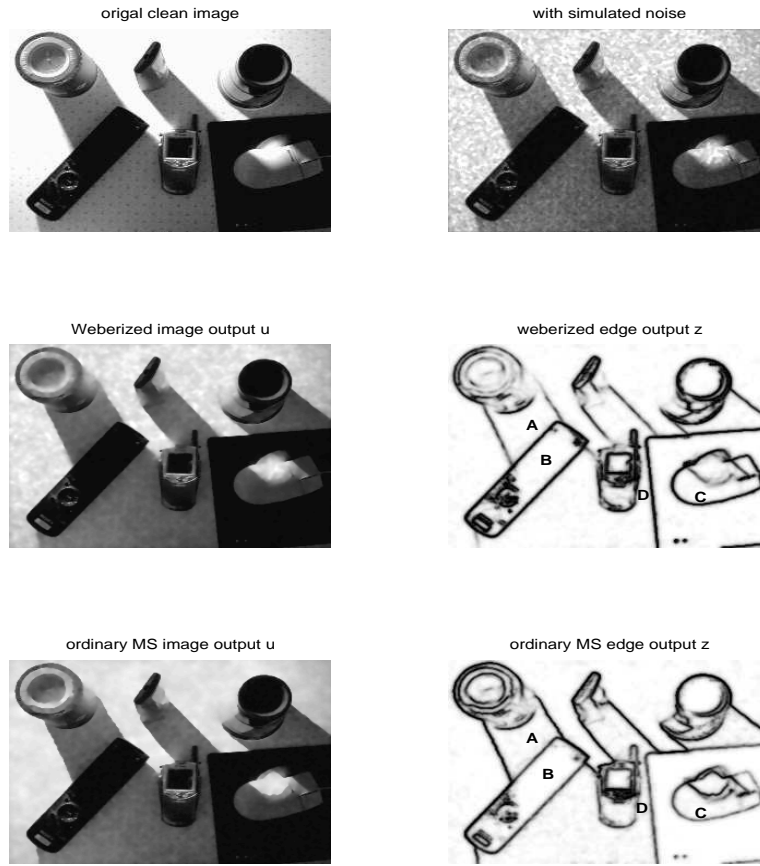


FIG. 6.2. Performance comparison via a real image between the ordinary Mumford–Shah model and its weberization (via Γ -convergence approximation). The 3-D scene has been arranged to create an image with a large dynamic range of intensity values. In particular, the shadows have been intentionally introduced to create dimmer regions. Notice that for objects in the shadows (i.e., the control panel **B**, cell phone **D**, and the mouse and mouse pad **C**), the detected edges are still strong in the weberized Mumford–Shah model while much weaker in the ordinary Mumford–Shah model. That is, along the shady edges, the relative strengths η_ε are much stronger from the weberized model. On the other hand, the boundaries of the shadows themselves (e.g., shadow **A** of the pen holder) against a brighter background are stronger in the ordinary Mumford–Shah model than its weberization. From the constancy of the edge strengths of the control panel (**B**) or the mouse (**C**), it is clear that the weberized model is insensitive to external illuminance conditions, and able to capture object edges equally well whether under direct illumination or in the shadows, which is consistent with Shapley and Enroth-Cugell’s inference in Section 2 and Weber’s Law.

are numerically integrated by a stable iterative algorithm based on a linearization technique.

Computational results have convincingly confirmed the light adaptivity feature of the new model. Comparison with the classical Mumford–Shah model further highlights the noticeable improvement achieved by the new model.

Acknowledgments. The first author would like to thank Professors Gilbert Strang, Tony Chan, Stan Osher, and David Mumford for their invaluable teaching and inspirations, as well as the Applied Mathematics Program of NSF (USA) under DMS-0202565 for the support to the current project.

REFERENCES

- [1] G. Alberti and C. Mantegazza. A note on the theory of SBV functions. *Bollettino U. M. I.*, 7(11-B):375–382, 1997.
- [2] L. Ambrosio. A compactness theorem for a new class of functions of bounded variation. *Bollettino U. M. I.*, 7(3-B):857–881, 1989.
- [3] L. Ambrosio and V. M. Tortorelli. Approximation of functionals depending on jumps by elliptic functionals via Γ -convergence. *Comm. Pure Appl. Math.*, 43:999–1036, 1990.
- [4] L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital.*, 6-B:105–123, 1992.
- [5] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. McGrawHill, Inc., 1978.
- [6] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29:182–193, 1992.
- [7] A. Chambolle. Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations. *SIAM J. Appl. Math.*, 55(3):827–863, 1995.
- [8] A. Chambolle. Finite-differences discretizations of the Mumford-Shah functional. *M2AN Math. Model. Numer. Anal.*, 33(2):261–288, 1999.
- [9] A. Chambolle and P. L. Lions. Image recovery via Total Variational minimization and related problems. *Numer. Math.*, 76:167–188, 1997.
- [10] T. F. Chan, S.-H. Kang, and J. Shen. Euler’s elastica and curvature based inpainting. *SIAM J. Appl. Math.*, 63(2):564–592, 2002.
- [11] T. F. Chan and J. Shen. Mathematical models for local nontexture inpaintings. *SIAM J. Appl. Math.*, 62(3):1019–1043, 2001.
- [12] T. F. Chan and J. Shen. On the role of the BV image model in image restoration. *Amer. Math. Soc. Contemporary Mathematics*, volume on *Recent Advances in Scientific Computing and Partial Differential Equations*, Editors: S. Y. Cheng, C.-W. Shu, and T. Tang, 330:25–41, 2003.
- [13] T. F. Chan and J. Shen. *Image Analysis and Processing: variational, PDE, wavelet, and stochastic methods*. SIAM Publisher, Philadelphia, 2005.
- [14] T. F. Chan, J. Shen, and L. Vese. Variational PDE models in image processing. *Notices Amer. Math. Soc.*, 50:14–26, 2003.
- [15] T. F. Chan and L. A. Vese. A level set algorithm for minimizing the Mumford-Shah functional in image processing. *IEEE/Computer Society Proceedings of the 1st IEEE Workshop on “Variational and Level Set Methods in Computer Vision”*, pages 161–168, 2001.
- [16] D. Chandler. *Introduction to Modern Statistical Mechanics*. Oxford University Press, New York and Oxford, 1987.
- [17] G. Dal Maso, J.-M. Morel, and S. Solimini. A variational method in image segmentation: Existence and approximation results. *Acta. Math.*, 168:89–151, 1992.
- [18] S. Esedoglu and J. Shen. Digital inpainting based on the Mumford-Shah-Euler image model. *European J. Appl. Math.*, 13:353–370, 2002.
- [19] G. Fain and H. R. Matthews. Calcium and the mechanism of light adaptation in vertebrate photoreceptors. *Trends Neurosci.*, 13:378–384, 1990.
- [20] G. T. Fechner. Über ein wichtiges psychophysisches Grundgesetz und dessen Beziehung zur Schätzung der Sterngrößen. *Abh. k. Ges. Wissensch. Math.-Phys.*, K1, 4, 1858.
- [21] W. H. Fleming and R. Rishel. An integral formula for total gradient variation. *Arch. Math.*, 11:218–222, 1960.
- [22] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6:721–741, 1984.
- [23] W. Gibbs. *Elementary Principles of Statistical Mechanics*. Yale University Press, 1902.
- [24] E. De Giorgi, M. Carriero, and A. Leaci. Existence theorem for a minimization problem with free discontinuity set. *Arch. Rational Mech. Anal.*, 108:195–218, 1989.
- [25] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Birkhäuser, Boston, 1984.
- [26] U. Grenander. *Lectures in Pattern Theory. I. II. and III*. Springer, 1976-1981.
- [27] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int’l J. Comp. Vision*, 1(4):321–331,

- 1987.
- [28] J. Keener and J. Sneyd. *Mathematical Physiology*. Interdisciplinary Applied Mathematics. Springer, New York, 1998.
 - [29] E. H. Lieb and M. Loss. *Analysis*. Amer. Math. Soc., second edition, 2001.
 - [30] J. T. Marti. *Introduction to Sobolev Spaces and Finite Element Solution Elliptic Boundary Value Problems*. Academic Press, 1986.
 - [31] P. A. McNaughton. Light response of vertebrate photoreceptors. *Physiological Reviews*, 70:847–883, 1990.
 - [32] J. W. Milnor. *Topology from the Differentiable Viewpoint*. Princeton Univ. Press, revised edition, 1997.
 - [33] D. Mumford. *Geometry Driven Diffusion in Computer Vision*, chapter “The Bayesian rationale for energy functionals”, pages 141–153. Kluwer Academic, 1994.
 - [34] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Applied. Math.*, 42:577–685, 1989.
 - [35] R. A. Norman and I. Perlman. The effects of background illumination on the photoresponses of red and green cones. *J. Physiology*, 286:491–507, 1979.
 - [36] S. Osher and J. A. Sethian. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *J. Comput. Phys.*, 79(12), 1988.
 - [37] T. Poggio and S. Smale. The mathematics of learning: Dealing with data. *Notices Amer. Math. Soc.*, 50(5):537–544, 2003.
 - [38] E. N. Pugh and T. D. Lamb. Cyclic GMP and calcium: messengers of excitation and adaptation in vertebrate photoreceptors. *Vision Research*, 30:1923–1948, 1990.
 - [39] L. Rudin and S. Osher. Total variation based image restoration with free local constraints. *Proc. 1st IEEE ICIP*, 1:31–35, 1994.
 - [40] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.
 - [41] R. M. Shapley and C. Enroth-Cugell. Visual adaptation and retinal gain controls. In: *Progress in Retinal Research*, Ed: N. Osborne and G. Chader; Pergamon Press, London, 1984.
 - [42] J. Shen. Inpainting and the fundamental problem of image processing. *SIAM News*, 36, 2003.
 - [43] J. Shen. On the foundations of vision modeling I. Weber’s law and Weberized TV restoration. *Physica D: Nonlinear Phenomena*, 175:241–251, 2003.
 - [44] D. Tranchina, J. Gordon, and R. Shapley. Retinal light adaptation - evidence for a feedback mechanism. *Nature*, 310:314–316, 1984.
 - [45] D. Tranchina and C. S. Peskin. Light adaptation in the turtle retina: embedding a parametric family of linear models in a single nonlinear model. *Visual Neuroscience*, 1:339–348, 1988.
 - [46] A. Tsai, Jr. A. Yezzi, and A. S. Willsky. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation and magnification. *IEEE Trans. Image Process.*, 10(8):1169–1186, 2001.
 - [47] E. H. Weber. De pulsu, resorptione, audita et tactu. *Annotationes anatomicae et physiologicae*, Leipzig: Koehler, 1834.