On the Slicing Moments of BV Functions and Applications to Image Dejittering *

Sung Ha Kang †and Jianhong (Jackie) Shen‡

Abstract

Since the celebrated works of Rudin-Osher-Fatemi (*Physica D*, **60**:259-268, 1992), the functions in the space of bounded variations (BV) has become a powerful mathematical model for approximating generic image signals in contemporary imaging and vision sciences. Motivated by the important application of image and video dejittering, we study the mathematical properties of the slicing moments of BV images. The regularity characterization leads to a novel model for the dejittering problem based upon the Bayesian/Tikhonov principle. Mathematical as well as computational properties are developed.

keyword: Bounded variation, slicing moments, Bayesian, inverse problem, dejittering, variational, regularization, existence. (AMS 2000 Subj: 94A08 and 49N45.)

1 Introduction

Growing popularity of image processing and vision analysis within the mathematics community has been determined by two basic facts:

(a) images and visual signals are first of all *functions* [7], and

(b) understanding the *patterns* [19] of the functions (i.e. visual perception (e.g., for robots) and image analysis (e.g., medical CT or MRI images)) is, mathematical properties, such as geometrical, algebraic, topological, or stochastic invariants.

Thus, processing of images or visual signals is the tantamount to the analysis of a special class of functions called *images*, which serves as the mathematical foundation of image processing.

In recent two decades, the marriage of image processing, vision analysis, and mathematics has nurtured numerous exciting discoveries as well as revived various classical subjects. For example, revisiting many classical subjects includes wavelets, multiresolution analysis, oscillatory patterns, fractals, moving fronts, multiphase problems with free boundaries, and Gibbs' random fields, just to name a few [10, 12, 17, 20, 21, 28]. Mathematics has provided the solid ground for solving many challenging imaging and vision problems in unified and mass-production manners. At the same

*This work has been partially supported by the NSF (USA) under grant numbers DMS-0312223 and DMS-0202565.
†Department of Mathematics, University of Kentucky, Lexington, KY 40515, USA, (skang@ms.uky.edu)
‡School of Mathematics, University of Minnesota, Minneapolis, MN 55455, (jhshen@math.umn.edu)

time, countless emerging applications of imaging and vision technologies in this information era have provided fertile soils for nurturing new problems and theories in mathematics. The recent expository article [9] and research monograph [7] provide further details along this line.

This current work is easily embedded into this general picture of contemporary mathematical image and vision analysis (Miva). Inspired by an important application called image(video) dejittering, we introduce and explore the properties of the slicing moments of multi-dimensional functions with *bounded variations* (BV).

The BV image model was first introduced into image analysis by the celebrated work of Rudin, Osher, and Fatemi [21]. It has become one of the most powerful image models that reach a good balance between geometric fidelity and computational complexity (e.g., [1, 3, 5, 6, 8, 22, 24, 27]). Numerous applications have shown that except for oscillatory textures of small amplitudes [2, 25], the BV image model performs sufficiently well in describing visually important geometric features like edges.

Motivated by the image dejittering problem, in the current paper, we first introduce and study the properties of the slicing moments of BV functions, and propose a novel dejittering model based upon the idea of moment regularization. Our mathematical framework is intentionally kept general (in terms of dimensions and assumptions), and aims at contributing to solving many other problems in related applied sciences.

As shown in Fig. 1, image jittering occurs when the slices of a high dimensional image signal are *randomly* displaced along the slicing space (e.g., a line or a plane). Three major technological areas where jittering frequently arises are: (a) *video jittering* due to the corruption of synchronization signals in analog video tapes; (b) *video interlacing* due to the temporal difference between the fast motions of objects in a scene and the refreshing speed of a digital display device; and (c) *slice jittering* in imaging devices such as CT (computer tomography) and MRI (magnetic resonance imaging) scanning, when patients or devices undergo random spatial displacements during an imaging acquisition process.

Figure 1: (a) an ideal image $u(x, y)$, and (b) its randomly jittered image $u_J(x, y)$.

To restore an ideal image u from its jittered version u_J is the problem called *image dejittering*. For corrupted analog videos, in [15, 16], Kokaram and his colleagues first explored dejittering methods that only rely upon the jittered video images instead of other irrelevant tape information. Such approaches are said to be *intrinsic* in contrast with most conventional video dejittering techniques, which employ extra image-irrelevant information. In [23], the second author developed an intrinsic variational dejittering model based on Bayesian estimation theory. In [14], the two authors further proposed a flexible two-step model called "bake and shake" for intrinsic image dejittering using nonlinear diffusion partial differential equations.

The aforementioned works could be considered as "differential" since they all depend upon the characterizations of local image structures. The current work ,therefore, distinguishes in its "integral" nature since slicing moments are integrated quantities. In general, integral methods are more robust to small perturbations. Furthermore, integrated quantities like moments naturally achieve dimensionality reduction and gain substantial computational efficiency.

The organization of the paper goes as follows. In Section 2, we first introduce the notion of slicing moments for high-dimensional BV images, and prove that they generally inherit the BV regularity. In Section 3, based on the regularity of slicing moments as well as Bayesian estimation theory, we propose a novel variational dejittering model in arbitrary dimensions, and establish its wellposedness by showing the existence and uniqueness of the optimal solution. In Section 4, algorithm and numerical examples are presented to demonstrate the performance of the new dejittering model. A brief conclusion is made in Section 5.

2 Slicing Moments of BV Functions

In this section, we first show that the slicing moments of a typical BV image is also a BV function, which enables us to employ the Bayesian restoration framework for image dejittering [7]. In this paper, we shall study BV functions in $\mathbb{R}^n$ which are compactly supported and nonnegative:

$$BV_c^+ = BV_c^+(\mathbb{R}^n) = \{v \in L^1(\mathbb{R}^n) \mid v \geq 0, \text{compactly supported, and } \int_{\mathbb{R}^n} |Dv| < \infty\}.$$

Nonnegativity is a plausible assumption in imaging and vision since physically image values represent photon counts. Recall that the total variation (TV) Radon measure is defined by, for any open domain $U \subseteq \mathbb{R}^n$,

$$\int_U |Dv| = \sup_{\vec{g} \in C_c^1(U, \mathcal{B}^n)} \int_U v \, \mathrm{div}(\vec{g}) dz, \quad \text{with } \; dz = dz_1 \cdots dz_n, \tag{1}$$

where $\mathcal{B}^n$ denotes the n-dimensional unit ball centered at the origin in $\mathbb{R}^n$. Fixing any $d = 0, 1, \ldots n-1$, we write $z = (x, y) \in \mathbb{R}^n$ with

$$x = (z_1, \cdots, z_{n-d}) \in \mathbb{R}^{n-d} \quad \text{and} \quad y = (z_{n-d+1}, \cdots, z_n) \in \mathbb{R}^d.$$

For any multi-exponent $\alpha = (\alpha_1, \cdots, \alpha_{n-d}) \in \{0, 1, 2, \cdots\}^{n-d}$, define x^α to be

$$x^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \cdots z_{n-d}^{\alpha_{n-d}} \in \mathbb{R}.$$

Definition 1 (Slicing Moments) *Given an image $u \in BV_c^+$ and an exponent α, the slicing moment of u of codimension d is defined by*

$$m_d(y|u, \alpha) = \int_{\mathbb{R}^{n-d}} x^\alpha u(x, y) dx. \tag{2}$$

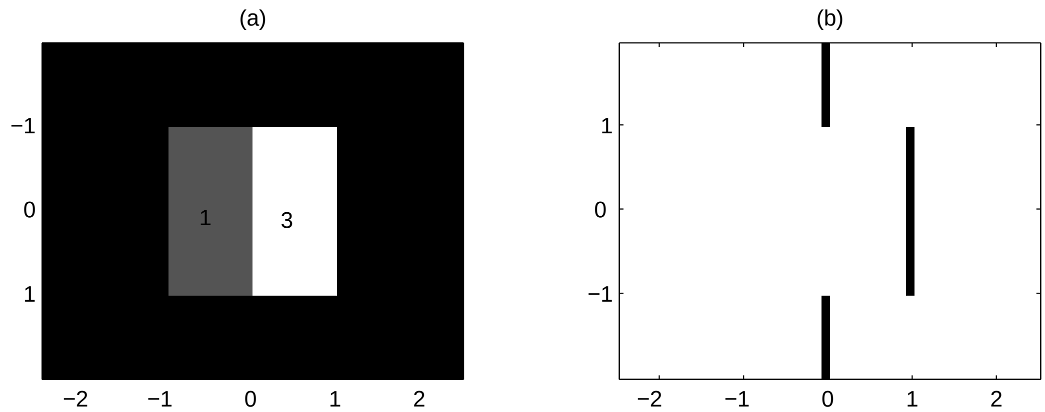

Figure 2: (a) an image sample u; (b) the (horizontal) slicing moment $m_1(y \mid u, 1)$.

Notice that m_d is a function in $\mathbb{R}^d$ for any given u and α. The integral is indeed well defined since $u \in BV_c^+$ is assumed to be compactly supported. Fig. 2 shows an example of slicing moments of a simple image with dimension $n = 2$ and codimension $d = 1$. The image on the left panel is a synthetic BV image, and the graph plotted on the right panel is its (horizontal) slicing moment with $\alpha = 1$ and $d = 1$. It is clear that the slicing moment is piecewise constant and still a BV function of y. If image (a) is jittered, the moment function in (b) would become noisy, and effective noise estimation can reveal the important information about the unknown jitters. This is the key observation leading to our novel dejittering model later. We now first show that the slicing moment function is also a BV function when the given image u is. This theorem is crucial for our new model, since it allows to make use of regularization techniques for *degraded* BV functions [21].

Theorem 1 *For any given image $u \in BV_c^+(\mathbb{R}^n)$, codimension $d \in \{0, 1, \cdots, n-1\}$, and multi-exponent $\alpha \in \{0, 1, \cdots\}^{n-d}$,*

$$m_d(y|u, \alpha) \in BV_c(\mathbb{R}^d).$$

Proof. We show that m_d is compactly supported, belongs to $L^1(\mathbb{R}^d)$, and $\int |Dm_d| < \infty$.

[1] Since u is compactly supported, there exists some $\gamma > 0$ such that

$$\text{supp } u \subseteq \{z \in \mathbb{R}^n : |z|_\infty = \max_{1 \le i \le n} |z_i| \le \gamma\}. \tag{3}$$

In particular, for any $z = (x, y)$ with $x \in \mathbb{R}^{n-d}$ and $|y|_\infty > \gamma$, one has $u(x, y) = 0$ and

$$m_d(y|u, \alpha) = \int_{\mathbb{R}^{n-d}} x^\alpha u(x, y) dx = 0.$$

Therefore, $m_d(y|u, \alpha)$ is also compactly supported and

$$\text{supp } m_d(y|u, \alpha) \subseteq \{y \in \mathbb{R}^d : \ |y|_\infty \le \gamma\}.$$

[2] Next, we show that $m_d \in L^1(\mathbb{R}^d)$. With $z = (x, y)$, one has

$$\begin{aligned}
\int_{\mathbb{R}^d} |m_d(y|u,\alpha)| dy &= \int_{\mathbb{R}^d} |\int_{\mathbb{R}^{n-d}} x^\alpha u(x,y) dx| dy \\
&\le \int_{\mathbb{R}^n} |x^\alpha|\ u(z) dz \qquad \text{(by Fubini's Theorem)} \\
&= \int_{\{z : |x|_\infty \le \gamma\}} |x^\alpha|\ u(z) dz \qquad \text{(by (3))} \\
&\le \gamma^{|\alpha|} \int_{\mathbb{R}^n} u(z) dz < \infty, \qquad \text{(since } u \in L^1(\mathbb{R}^n))
\end{aligned}$$

where $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_{n-d}$. Therefore, $m_d(y|u,\alpha) \in L^1(\mathbb{R}^d)$.

[3] By the definition of the TV Radon measure (1),

$$\int_{\mathbb{R}^d} |D\ m_d(y|u,\alpha)| = \sup_{\vec{\varphi} \in C_c^1(\mathbb{R}^d, \mathcal{B}^d)} \int_{\mathbb{R}^d} m_d(y|u,\alpha)\ \mathrm{div}_y(\vec{\varphi})\ dy, \tag{4}$$

where $y = (y_1, y_2, \ldots, y_d)$, $\vec{\varphi} = \vec{\varphi}(y) = (\varphi_1, \cdots, \varphi_d)$, and

$$\mathrm{div}_y(\vec{\varphi}) = \partial_{y_1}\varphi_1 + \cdots + \partial_{y_d}\varphi_d.$$

For any fixed γ in (3), choose $\rho_\gamma(x) \in C_c^1(\mathbb{R}^{n-d})$ such that $\rho_\gamma(x) \in [0, 1]$ and

$$\rho_\gamma(x) = \begin{cases} 1, & \text{for } |x|_\infty \le \gamma \\ 0, & \text{for } |x|_\infty > \gamma + 1 \end{cases}. \tag{5}$$

Then, $\forall z \in \mathbb{R}^n$ (with $x \in \mathbb{R}^{n-d}$ and $y \in \mathbb{R}^d$), one has

$$u(z) = u(x, y) \equiv u(x, y)\rho_\gamma(x). \tag{6}$$

For any given α and $\vec{\varphi}(y) \in C_c^1(\mathbb{R}^d, \mathcal{B}^d)$, define a new flow on the entire space $\mathbb{R}^n$ by

$$\vec{g}(z) = \vec{g}(x, y) = \big(0^{n-d},\ x^\alpha \vec{\varphi}(y)\rho_\gamma(x)\big), \tag{7}$$

where 0^{n-d} denotes the origin of $\mathbb{R}^{n-d}$. Then,

$$\mathrm{div}(\vec{g}(z)) = \mathrm{div}_y\big(x^\alpha \rho_\gamma(x)\vec{\varphi}(y)\big) = x^\alpha \rho_\gamma(x)\ \mathrm{div}_y(\vec{\varphi}(y)). \tag{8}$$

Furthermore, by the definitions in (5) and (7),

$$\mathrm{supp}\ \vec{g} \subseteq \{x :\ |x|_\infty \le \gamma\} \times \mathrm{supp}\vec{\varphi}(y),$$

implying that $\vec{g} \in C_c^1(\mathbb{R}^n, \mathbb{R}^n)$. With $z = (x, y)$ and (5),

$$|\vec{g}(z)|_2 = |x^\alpha \rho_\gamma(x)| \cdot |\vec{\varphi}(y)|_2 \le \gamma^{|\alpha|} \|\vec{\varphi}\|_\infty, \tag{9}$$

where $\|\vec{\varphi}\|_\infty = \sup_y |\vec{\varphi}(y)|_2$. Therefore, $\gamma^{-|\alpha|}\vec{g} \in C_c^1(\mathbb{R}^n, \mathcal{B}^n)$.

For any test flow $\vec{\varphi}(y) \in C_c^1(\mathbb{R}^d, \mathcal{B}^d)$, by Fubini's Theorem,

$$\begin{aligned}
\int_{\mathbb{R}^d} m_d(y|u,\alpha)\ \mathrm{div}_y\vec{\varphi}(y)\ dy &= \int_{\mathbb{R}^d} \big(\int_{\mathbb{R}^{n-d}} x^\alpha u(x,y)dx\big)\ \mathrm{div}_y\vec{\varphi}(y)\ dy \\
&= \int_{\mathbb{R}^n} u(x,y)\ x^\alpha \mathrm{div}_y\vec{\varphi}(y)\ dz \\
&= \int_{\mathbb{R}^n} u(x,y)\ \rho_\gamma(x)\ x^\alpha \mathrm{div}_y\vec{\varphi}(y)\ dz \qquad \text{(by (6))} \\
&= \int_{\mathbb{R}^n} u(z)\ \mathrm{div}\vec{g}(z)dz \qquad \text{(by (8))} \\
&\le \gamma^{|\alpha|} \int_{\mathbb{R}^n} |Du|. \qquad \text{(by (9))}
\end{aligned}$$

Since $\vec{\varphi}$ is arbitrary and $u \in BV_c^+(\mathbb{R}^n)$, we conclude that

$$\int_{\mathbb{R}^d} |D\ m_d(y|u,\alpha)| \le \gamma^{|\alpha|} \int_{\mathbb{R}^n} |Du| < \infty. \tag{10}$$

The proof is complete. □

In particular when $\alpha = 0^{n-d}$, we have the following corollary for marginal projections, which is need for later developments. (The term "marginal" has been motivated by the term "marginal distribution" in multivariate probability theory.)

Corollary 1 (Marginal Projections) *Define $M_d(y|u) = m_d(y|u, 0^{n-d})$ to be the marginal projection of codimension d. Then, $M_d(y|u) \in BV_c^+(\mathbb{R}^d)$, and*

$$\int_{\mathbb{R}^d} |D\ M_d(y|u)| \le \int_{\mathbb{R}^n} |Du|. \tag{11}$$

Proof. Notice that $M_d \ge 0$ due to $u \ge 0$. Then, (11) follows from (10) for $\alpha = 0^{n-d}$. □

In Theorem 1, the slicing moment functions have been shown to belong to the BV space. We now remark via the example in Fig. 2 that the BV regularity cannot be upgraded to the Sobolev regularity $W^{1,1}$. The image on the left panel of Fig. 2 is defined by, with $z = (x, y)$,

$$u(z) = \begin{cases} 0, & |z|_\infty > 1 \\ 1, & |z|_\infty \le 1, x \le 0 \\ 3, & |z|_\infty \le 1, x > 0 \end{cases}.$$

For $\alpha = 1$, define the (horizontal) linear slicing moment $m(y|u) = m_1(y|u, 1) = \int_{\mathbb{R}} xu(x, y)dx$. Then, for $\forall y$ with $|y| > 1$, one has $m(y|u) \equiv 0$, and for $\forall y \in (-1, 1)$,

$$m(y|u) = \int_{-1}^0 xdx + \int_0^1 3xdx = \int_0^1 2xdx \equiv 1.$$

Therefore, as illustrated on the right panel of Fig. 2, $m(y|u) = 1_{|y|\le 1}(y)$, and the signed total variation Radon measure is only expressible via Dirac's delta function:

$$Dm(y|u) = \delta(y + 1) - \delta(y - 1),$$

which does not belong to $L^1(\mathbb{R})$. Thus, $m(y|u) \in BV(\mathbb{R}) \setminus W^{1,1}(\mathbb{R})$, and the regularity result in Theorem 1 is optimal.

For the application to *intrinsic* image dejittering, such regularity information will be the key to our novel model which is to be discussed next.

3 Moments Regularization for Image Dejittering

In this section, we apply the above regularity results to the problem of image dejittering.

3.1 Formulation of the jittering problem

In the language of inverse problems, dejittering is to invert the *forward problem* of jittering. Thus, we first propose a generic forward model for the jittering process.

Definition 2 (Jitter s) *A q-dimensional jitter (field) on $\mathbb{R}^d$ is a random map:*

$$s : \mathbb{R}^d \to \mathbb{R}^q, \;\; y \to s(y),$$

such that, for any finite set of points $E \subseteq \mathbb{R}^d$,

$$\{s(y) \mid y \in E\},$$

are independent and identically distributed (i.i.d) random variables.

As an example, for any fixed $y \in \mathbb{R}^d$, jitter $s(y)$ could be subject to the Gaussian normal distribution $\mathcal{N}(0^q, \Sigma)$ with a covariance matrix Σ. In term of the probability density function (p.d.f.), one has

$$p(s = \hat{s}) = \frac{1}{\sqrt{(2\pi)^q |\Sigma|}} e^{-\frac{1}{2}\hat{s}^T \Sigma^{-1} \hat{s}}. \tag{12}$$

Definition 3 (Jittered Image u_J) *Let $u \in BV_c^+(\mathbb{R}^n)$ and $d \in \{0, 1, \ldots, n-1\}$. For any given $(n-d)$-dimensional jitter $s(y)$ on $\mathbb{R}^d$, the jittered image u_J is defined to be :*

$$u_J(z) = u_J(x, y) = u(x - s(y), y), \;\; z \in \mathbb{R}^n, \; x \in \mathbb{R}^{n-d}, \; \text{and}\, y \in \mathbb{R}^d. \tag{13}$$

Definition 4 (Dejittering) *The dejittering problem is the inverse problem of restoring the original image $u(z)$ from its jittered observation $u_J(z)$ (see Fig. 1).*

3.2 Linear slicing moments and Bayesian inference

Definition 5 (Linear Slicing Moments) *Let the codimension d linear moments $\vec{m}_d(y|u)$ for $u \in BV_c^+(\mathbb{R}^n)$ be the vectorial function*

$$\vec{m}_d(y|u) = \big(m_d(y|u, e_1), \cdots, m_d(y|u, e_{n-d})\big), \tag{14}$$

where $e_i = (0, \cdots, 0, 1_{i^{th}}, 0, \cdots, 0)$, $\;i = 1, \ldots, n-d$. Equivalently, it is given by

$$\vec{m}_d(y|u) = \int_{\mathbb{R}^{n-d}} x u(x, y) dx, \;\; x = (z_1, \cdots, z_{n-d}).$$

By Theorem 1, one immediately has the following regularity.

Corollary 2 *The linear slicing moment $\vec{m}_d(y|u)$ belongs to $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$.*

Notice that in terms of linear structures, one has $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d}) = BV_c(\mathbb{R}^d, \mathbb{R})^{n-d} = BV_c(\mathbb{R}^d)^{n-d}$. As for the TV Radon measure in $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$, we follow the general definition of total variations for product measures [11]. Recall that for any given p measures $\mu_1, \cdots, \mu_p$ on a measurable space (X, Σ) (where Σ is a σ-algebra on X), the total variation $|\vec{\mu}|$ of the vectorial measure $\vec{\mu} = (\mu_1, \cdots, \mu_p)$ is defined by ;

$$\text{For any } E \in \Sigma, \quad |\vec{\mu}|(E) = \sup_{\|\vec{\varphi}\|_\infty \le 1} \sum_{i=1}^{p} \int_E \varphi_i d\mu_i = \sup_{\|\vec{\varphi}\|_\infty \le 1} \int_E \vec{\varphi} \cdot d\vec{\mu},$$

where $\vec{\varphi}$ is a Σ-measurable vectorial function, and

$$\|\vec{\varphi}\|_\infty = \sup_{x \in X} |\vec{\varphi}|_2(x) = \sup_{x \in X} \sqrt{\varphi_1^2(x) + \cdots + \varphi_p^2(x)}.$$

One symbolically writes $|\vec{\mu}| = \sqrt{\mu_1^2 + \cdots + \mu_p^2}$. If there exists a (positive) measure v on (X, Σ), such that all the Radon-Nikodym derivatives exist:

$$\rho_i = \frac{d\mu_i}{dv}, \quad i = 1, \ldots, p,$$

then, $|\vec{\mu}|$ must be differentiable with respect to v, and

$$\frac{d|\vec{\mu}|}{dv} = |\vec{\rho}|_2 = \sqrt{\rho_1^2 + \cdots + \rho_p^2}$$

or equivalently $|\vec{\mu}|(E) = \int_E |\vec{\rho}|_2 dv$ for any $E \in \Sigma$.

By this general framework, the natural total variation measure in $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$ for $\vec{m}_d(y|u)$ is: for any Borel set $U \subseteq \mathbb{R}^d$,

$$\int_U |D\, \vec{m}_d(y|u)| = \int_U \left[\sum_{i=1}^{n-d} [D m_d(y|u, e_i)]^2 \right]^{1/2}, \tag{15}$$

where e_i's are as in (14). In particular, if $\vec{m}_d \in W^{1,1}(\mathbb{R}^d, \mathbb{R}^{n-d})$, one has

$$\int_U |D\, \vec{m}_d(y|u)| = \int_U \left[\sum_{i=1}^{n-d} [\nabla m_d(y|u, e_i)]^2 \right]^{1/2} dy. \tag{16}$$

By Corollary 2, $\vec{m}_d$ is a BV vectorial function under the definition in (15). In the following proposition, we consider the link between the linear slicing moment and the jitter $s(y)$.

Proposition 1 *Let $u_J(z)$ denote the jittered image generated from $u(z)$ by jitter $s(y)$ as in (13). Then, the linear slicing moment of u_J and u are connected by:*

$$\vec{m}_d(y|u_J) = \vec{m}_d(y|u) + s(y) M_d(y|u), \tag{17}$$

where $M_d(y|u)$ is the codimension d marginal projection of u as defined in Corollary 1.

Proof. It suffices to carry out the following computation,

$$
\begin{aligned}
\vec{m}_d(y|u_J) &= \int_{\mathbb{R}^{n-d}} x u_J(x,y)dx = \int_{\mathbb{R}^{n-d}} x u(x-s(y),y)dx \\
&= \int_{\mathbb{R}^{n-d}} (t+s(y))u(t,y)dt \\
&= \int_{\mathbb{R}^{n-d}} tu(t,y)dt + s(y)\int_{\mathbb{R}^{n-d}} u(x,y)dx \\
&= \vec{m}_d(y|u) + s(y)M_d(y|u). \qquad \square
\end{aligned}
$$

Therefore, if the true image u were known, one could easily identify the jitter $s(y)$ by Proposition 1. In reality, only u_J and $\vec{m}_d(y \mid u_J)$ are directly available while u and $\vec{m}_d(y \mid u)$ are unknown. The following proposition shows that $M_d(y \mid u)$ is in fact directly readable from the jittered image u_J.

Proposition 2 *The marginal projection is jittering invariant, i.e.,*

$$M_d(y|u_J) = M_d(y|u).$$

The proof is straightforward since the Lebesgue measure dx is translation-invariant. Eqn. (17) now becomes

$$\vec{m}_d(y|u_J) = \vec{m}_d(y|u) + s(y)M_d(y|u_J). \tag{18}$$

To summarize, in terms of estimating the unknown linear slicing moment $\vec{m}_d(y \mid u)$, (which is equivalent to the estimation of the jitter $s(y)$,) we have established the following two key ingredients in the framework of Bayesian inference [12, 18].

1. The prior model: Eqn. (15) specifies the regularity of the linear slicing moment $\vec{m}_d(y|u)$ for any given $u \in BV_c^+(\mathbb{R}^n)$.
2. The (generative) data model: Eqn. (18) specifies how the observable or computable data $\vec{m}_d(y|u_J)$ are generated from the unknown $\vec{m}_d(y|u)$.

In combination, they lead to our novel dejittering model built upon the Bayesian rationale [18], or equivalently in terms of the framework of inverse problems, the Tikhonov method [26].

3.3 Dejittering via moment regularization

For any fixed codimension d, we shall simplify the notations by defining $M(y) = M_d(y|u_J) = M_d(y|u)$, $\vec{m}_J(y) = \vec{m}_d(y|u_J)$, and $\vec{m}(y) = \vec{m}_d(y|u)$. For image and video dejittering, as in Eqn. (12), the jitter $s(y)$ is assumed to be of Gaussian type $\mathcal{N}(0^{n-d}, \Sigma)$ with a covariance matrix Σ. Also the data model in (18) reveals

$$s(y) = \frac{1}{M(y)}\big(\vec{m}_J(y) - \vec{m}(y)\big).$$

In combination with the BV regularity, the Bayesian/Tikhonov framework [9, 18], it leads to the following variational model for restoring the ideal linear moment $\vec{m}_d(y)$ from its jittered version $\vec{m}_J(y)$:

$$\min_{\vec{m}(y)\in BV_c(\mathbb{R}^d,\mathbb{R}^{n-d})} \int_{\mathbb{R}^d} |D\vec{m}(y)| + \frac{\lambda}{2}\int_{\mathbb{R}^d} \frac{1}{M^2(y)}(\vec{m}_J(y)-\vec{m}(y))\Sigma^{-1}(\vec{m}_J(y)-\vec{m}(y))^T dy. \tag{19}$$

The weight λ balances the regularity term and the fitting term, and the model is a regularized weighted least-square problem. In the fitting term, $M(y) = M_d(y|u_J)$ and $\vec{m}_J(y) = \vec{m}_d(y|u_J)$ are directly computable from a given jittered image $u_J(z)$, while $\vec{m}(y) = \vec{m}_d(y|u)$ is unknown. Furthermore, they satisfy the following compatibility condition.

Proposition 3 (Compatibility Condition) *For any $u \in BV_c^+(\mathbb{R}^d)$, the condition $M(y) = M_d(y|u_J) = M_d(y|u) = 0$ implies that $\vec{m}_J(y) = 0$ and $\vec{m}(y) = 0$, for any $y \in \mathbb{R}^d$.*

Proof. $\forall y \in \mathbb{R}^d$, $M(y) = 0 \Leftrightarrow u(x,y) = 0$ for a.e. $x \in \mathbb{R}^{n-d}$, which implies that

$$\vec{m}_J(y) = \vec{m}_d(y|u_J) = \int_{\mathbb{R}^{n-d}} xu(x-s(y),y)dx = 0,$$

$$\vec{m}(y) = \vec{m}_d(y|u) = \int_{\mathbb{R}^{n-d}} xu(x,y)dx = 0. \qquad \square$$

Inspired by this proposition, we now study *independently* the properties of the dejittering energy

$$E[\vec{m}|\vec{m}_J, M] = \int_{\mathbb{R}^d} |D\vec{m}| + \frac{\lambda}{2}\int_{\mathbb{R}^d} \frac{1}{M^2}(\vec{m}_J-\vec{m})\Sigma^{-1}(\vec{m}_J-\vec{m})^T dy, \tag{20}$$

for any given $\vec{m}_J$ and M, that are subject to:

(A1) $M(y) \geq 0$, compactly supported, and $M \in L^\infty(\mathbb{R}^d)$;

(A2) $M(y) = 0 \Rightarrow \vec{m}_J(y) = 0^{n-d}$, where $\vec{m}_J : \mathbb{R}^d \to \mathbb{R}^{n-d}$ is Lebesgue measurable; and

(A3) $\vec{m}_J \in L^2(\mathbb{R}^d \to \mathbb{R}^{n-d}, d\mu)$, where $d\mu = \frac{1}{M^2}dy$ denotes the weighted measure on $\mathbb{R}^d$.

Proposition 4 *Let $\vec{m} \equiv 0^{n-d}$ be the zero vectorial function. Then, $E[\vec{m} = 0^{n-d}|\vec{m}_J, M] < \infty$.*

Proof. This is guaranteed by (A3), and the fact that

$$\vec{m}_J\Sigma^{-1}\vec{m}_J^T \leq \frac{1}{\lambda_{\min}(\Sigma)}|m_J|^2, \tag{21}$$

where $\lambda_{\min}(\Sigma) > 0$ denotes the smallest eigenvalue of Σ. $\square$

Proposition 5 *Suppose $\vec{m} \in BV(\mathbb{R}^d, \mathbb{R}^{n-d})$ and $E[\vec{m}|\vec{m}_J, M] < \infty$, then*

$$M(y) = 0 \text{ implies } \vec{m}(y) = 0^{n-d}, \quad \text{a.e. } y \in \mathbb{R}^d. \tag{22}$$

In particular, $\vec{m}(y)$ must be compactly supported and $\vec{m}(y) \in BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$.

Proof. By the assumption,

$$\int_{\mathbb{R}^d} \frac{1}{M^2}(\vec{m}_J - \vec{m})\Sigma^{-1}(\vec{m}_J - \vec{m})^T dy < \infty.$$

Thus, $M = 0$ implies $(\vec{m}_J - \vec{m})\Sigma^{-1}(\vec{m}_J - \vec{m})^T = 0$, for a.e. $y \in \mathbb{R}^d$. Since Σ is positive definite, this further implies $\vec{m}_J = \vec{m}$, a.e. in $\mathbb{R}^d$. Then, (22) follows directly from the assumption (A2) (or Proposition 3), and the compactness of M passes onto $\vec{m}$ as a result. □

With these propositions, we now prove the existence and uniqueness of the minimizers to the dejittering energy (20).

Theorem 2 *Under the assumptions (A1), (A2) and (A3), the minimizer to energy $E[\vec{m}|\vec{m}_J, M]$ in (20) exists and is unique in $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$.*

Proof. First, we prove the existence of the minimizer. By Proposition 4,

$$\inf_{\vec{m} \in BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})} E[\vec{m}|\vec{m}_J, M] \le E[0^{n-d}|\vec{m}_J, M] < \infty.$$

Let $\{\vec{m}^i(y)\}$ be a minimizing sequence in $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$. Then, by Proposition 5, $\{\vec{m}^i(y)\}_{i=1}^{\infty}$ must be uniformly compactly supported, i.e., there exists a *bounded* open set U and a compact set $K \subseteq U$, such that

$$\text{supp } \vec{m}^i \subseteq K \subseteq U, \text{ for } i = 1, \cdots, \infty. \tag{23}$$

In addition, by the assumption (A2) and Proposition 5, one can assume

$$\text{supp } \vec{m}_J, \text{ supp } M \subseteq K \subseteq U. \tag{24}$$

Then,

$$E[\vec{m}^i|\vec{m}_J, M] \equiv E[\vec{m}^i|\vec{m}_J, M, U], \tag{25}$$

where the latter refers to the energy restricted over U:

$$E[\vec{m}^i|\vec{m}_J, M, U] = \int_U |D\vec{m}^i| + \frac{\lambda}{2}\int_U \frac{1}{M^2}(\vec{m}_J - \vec{m}^i)\Sigma^{-1}(\vec{m}_J - \vec{m}^i)^T dy.$$

By the assumption (A1)

$$\frac{1}{M^2}(\vec{m}_J - \vec{m}^i)\Sigma^{-1}(\vec{m}_J - \vec{m}^i)^T \ge \frac{1}{\lambda_{\max}(\Sigma)}\frac{1}{\|M\|_\infty^2}|\vec{m}_J - \vec{m}^i|_2^2,$$

where $\lambda_{\max}(\Sigma)$ denotes the largest eigenvalue of the covariance matrix. Since $L^2(U, \mathbb{R}^{n-d}) \subseteq L^1(U, \mathbb{R}^{n-d})$ for any bounded domain U, the sequence $\{\vec{m}^i(y)|_U\}_{n=1}^{\infty}$ is a bounded sequence in $BV(U, \mathbb{R}^{n-d})$. Therefore, by the L^1-weak compactness, there exists a subsequence $\{\vec{m}^k(y)|_U\} = \{\vec{m}^{i_k}(y)|_U\}$ that converges to some $\vec{m}^\infty$ in $L^1(U, \mathbb{R}^{n-d})$. One can further require that

$$\vec{m}^k(y) \longrightarrow \vec{m}^\infty(y), \text{ a.e. } y \in U. \tag{26}$$

Then, by the lower-semicontinuity property of the TV Radon measure under L^1 convergence,

$$\int_U |D\vec{m}^\infty| \le \liminf_{k\to\infty} \int_U |D\vec{m}^k|. \tag{27}$$

On the other hand, by (26) and Fatou's Lemma:

$$\int_U \frac{1}{M^2}(\vec{m}_J - \vec{m}^\infty)\Sigma^{-1}(\vec{m}_J - \vec{m}^\infty)^T dy \le \liminf_{k\to\infty} \int_U \frac{1}{M^2}(\vec{m}_J - \vec{m}^k)\Sigma^{-1}(\vec{m}_J - \vec{m}^k)^T dy. \qquad (28)$$

In combinations of (27), (28), and (25), we have

$$E[\vec{m}^\infty|\vec{m}_J, M, U] \le \lim_{k\to\infty} E[\vec{m}^k|\vec{m}_J, M, U] = \lim_{k\to\infty} E[\vec{m}^k|\vec{m}_J, M].$$

By (23), one must have supp $\vec{m}^\infty \subseteq K \subseteq U$, and

$$E[\vec{m}^\infty|\vec{m}_J, M, U] = E[\vec{m}^\infty|\vec{m}_J, M].$$

Therefore, we have established

$$E[\vec{m}^\infty|\vec{m}_J, M] \le \lim_{k\to\infty} E[\vec{m}^k|\vec{m}_J, M] = \inf_{\vec{m}} E[\vec{m}|\vec{m}_J, M].$$

Thus $\vec{m}^\infty \in BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$ has to be a minimizer.

Regarding the uniqueness, from the assumption (A1) on $M(y) \in L^\infty(\mathbb{R}^d)$, one has $M < \infty$ and $\frac{1}{M^2} > 0$ a.e. on $\mathbb{R}^d$. Then, it is trivial to see that $E[\vec{m}|\vec{m}_J, M]$ must be strictly convex in $BV_c(\mathbb{R}^d, \mathbb{R}^{n-d})$, and the minimizer has to be unique. □

This theorem secures the feasibility of proper numerical computations of the proposed dejittering model. From the given image u_J, first compute the jittered linear moment $\vec{m}_J$, then apply the dejittering functional (20) to regularize this moment function. The regularized moment function m^* is then employed to estimate the unknown jitter $s(y)$. For the model and algorithm to work effectively, one needs two pieces of input data: the jittered image $u_J \in \mathbb{R}^n$ and the statistics of the $(n-d)$-jitter $s(y) \in \mathbb{R}^d$ (i.e., the covariance matrix Σ as modeled by (12), which is often obtained by suitable statistical estimators).

Algorithm:

1. Compute the marginal projection $M(y)$ and the linear slicing moment $\vec{m}_J(y)$ of image u_J.
2. Find the minimizer of (20), $\vec{m}^*(y) = \text{argmin } E[\vec{m}|\vec{m}_J, M]$.
3. Compute the jitter by
$$s^*(y) = \begin{cases} \frac{\vec{m}_J - \vec{m}^*}{M}, & M(y) \neq 0 \\ 0^{n-d}, & M(y) = 0 \end{cases}.$$
4. Dejitter the image by s^*:
$$u^*(z) = u^*(x, y) = u_J(x + s^*(y), y).$$

In the next section, we discuss how to apply the above general framework to the practical application of 2-D image dejittering, for which $n = 2$, and $d = 1$.

4 Application to Image Dejittering and Examples

Let $\Omega_{R,H} = (-R, R) \times (0, H)$ denote a typical 2-D display domain, and an image defined on $\Omega_{R,H}$ be denoted by $v(x, y) \geq 0$ with $x \in (-R, R)$ and $y \in (0, H)$. A typical jitter can be modeled by a random map,

$$s = (0, H) \to \mathbb{R}, \quad y \to s(y).$$

As in Eqn. (12), assume that $s(y)$'s are i.i.d.'s of Gaussian type $\mathcal{N}(0, \sigma^2)$ with p.d.f.,

$$p(s(y) = a) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{a^2}{2\sigma^2}}, \qquad \text{for any fixed } y.$$

Then, a jittered image v_J is defined as

$$v_J(x, y) = v(x - s(y), y) \geq 0. \tag{29}$$

In practice, both v_J and v are indeed only displayed or available on a finite domain $\Omega_{R,H}$. It is then necessary to specify the boundary filling mechanism when $|s(y)| \neq 0$. Depending on the situation, the filled-in data at the boundaries could be (i) random, (ii) generated by Neumann flat extension, or (iii) generated by other mechanisms such as symmetric extension. To avoid such complications, as well as to illustrate the application of the general theory developed above, we assume that the image domain is an ideal horizontal stripe $\Omega_H = (-\infty, \infty) \times (0, H)$ (as in [23]) and that there exits some $R > 0$, such that

$$\text{supp}_{\Omega_H} \, v \subseteq \Omega_{R,H}, \quad \text{and} \;\; v \in BV^+(\Omega_H). \tag{30}$$

Then, v_J in (29) is always well-defined regardless of $s(y)$. Finally, by zero-padding, both v and v_J on Ω_H are extended to the entire plane $\mathbb{R}^2$, and denoted by u and u_J respectively. The jitter s is also naturally extended from $(0, H)$ to $\mathbb{R}^1$ by i.i.d. sampling. Then,

$$u_J(z) = u_J(x, y) = u(x - s(y), y), \;\; \forall z = (x, y) \in \mathbb{R}^2,$$

and (30) implies that $u \in BV_c^+(\mathbb{R}^2)$.

Notice that $\int_{\mathbb{R}^2} |u(z)|dz = \int_{\Omega_H} |v(z)|dz$, and

$$\int_{\mathbb{R}^2} |D\, u| = \int_{\Omega_H} |D\, v| + \int_{\partial\Omega_H} |f_v| dH^1 < \infty,$$

where $\partial\Omega_H = (\mathbb{R}^1 \times \{0\}) \cup (\mathbb{R}^1 \times \{H\})$ denotes the lower and upper boundaries, dH^1 the 1-dimensional Hausdorff measure, and $f_v = \text{Tr}(v)$ the trace of v along $\partial\Omega_H$ [13].

Thus, we are able to apply the general framework in the previous sections for the dejittering of u_J (and consequently for v_J). Define accordingly,

$$m(y) = \int_{\mathbb{R}} x u(x, y) dx, \;\; m_J(y) = \int_{\mathbb{R}} x\; u_J(x, y) dx, \quad \text{and}$$

$$M(y) = \int_{\mathbb{R}} u_J(x, y) dx \big(= \int_{\mathbb{R}} u(x, y) dx \big).$$

The dejittering model (20) becomes to minimize

$$E[m|m_J, M] = \int_{\mathbb{R}} |D\, m| + \frac{\mu}{2} \int_{\mathbb{R}} \frac{1}{M^2}(m_J - m)^2 dy, \tag{31}$$

where $\mu = \frac{\lambda}{\sigma^2}$. Eqn. (31) is a regularized weighted (by M^{-2}) least-square problem. If M were a constant, this equation would become precisely the 1-D version of the celebrated TV restoration model of Rudin-Osher-Fatemi [21].

For most digital devices, one has $u \in [0, 1]$ or $[1, 255]$ (8-bit). Then, the compactness of u ensures $M \in L^\infty(\mathbb{R})$. As long as $m_J \in L^2(\mathbb{R}, \frac{1}{M^2}dy)$, all the three conditions (A1), (A2) and (A3) of Theorem 2 are naturally satisfied. The optimal estimator $m^* = \text{argmin } E[m|m_J, M]$ therefore must exist uniquely.

In terms of numerical computations, there have been quite a few effective methods in the literature for models like (31), e.g., [3, 4, 21, 26]. One frequently adopted approach is based upon the formal Eular-Lagrange equation of (31),

$$D\left[\frac{D\, m(y)}{|D\, m(y)|}\right] + \frac{\mu}{M^2(y)}(m_J(y) - m(y)) = 0, \tag{32}$$

or equivalently,

$$M^2(y) D\left[\frac{D\, m(y)}{|D\, m(y)|}\right] + \mu(m_J(y) - m(y)) = 0. \tag{33}$$

It is evident from the last equation that $M(y) = 0$ implies $m(y) = m_J(y)$, which further leads to $m(y) = 0$ because of the assumption (A2) in Theorem 2. As common in the literature [7, 21, 26], a regularization parameter $\epsilon > 0$ can be introduced to replace $|Dm(y)|$ in the denominator by $|Dm|_\epsilon = \sqrt{\epsilon^2 + |Dm|^2}$ in (33). The nonlinear equation (33) can be solved iteratively by the lagged diffusivity fix-point method as in Acar and Vogel [1]. We refer to the remarkable monograph of Vogel for more details on the effective computations of models like (31)-(33), including discussions on the selection of the weighting parameter μ.

Numerical Examples

Finally, we demonstrate the computational performance of the new dejittering model through some typical examples. Notice that our model naturally applies to color images as well [2, 5].

The first example in Fig. 3 shows a synthetic piecewise constant image u, its jittered version u_J, and the dejittered image u^* via our new model based upon moment regularization. Since most images in the real world are often noisy, in Fig. 4 we have tested the robustness of our new model in the presence of intensity noises. The dejittered image in (c) clearly confirms such robustness, thanks to the averaging (or lowpass filtering) nature of moment integrals. In Fig. 5 and Fig. 6, via a standard test image in image processing, we have explicitly demonstrated the moment sequence from our dejittering model: the ideal moment $m(y)$, the jittered moment $m_J(y)$, and the optimally estimated moment $m^*(y)$. Finally, Fig. 7 shows the performance of the model on another standard test image of "Barbara" in image processing.

5 Conclusion

Motivated by the image dejittering problem in contemporary imaging science, the current paper introduces the notion of slicing moments of BV functions (or images), and studies their mathemat-

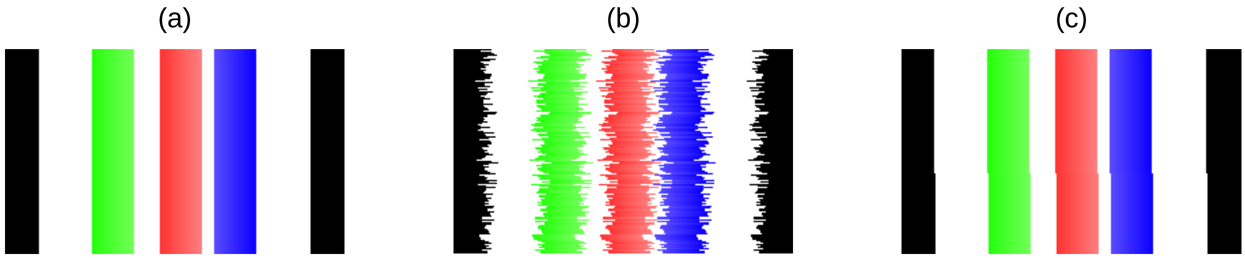

Figure 3: (a) ideal image u, (b) jittered image u_J, (c) dejittered image u^* via moment regularization.

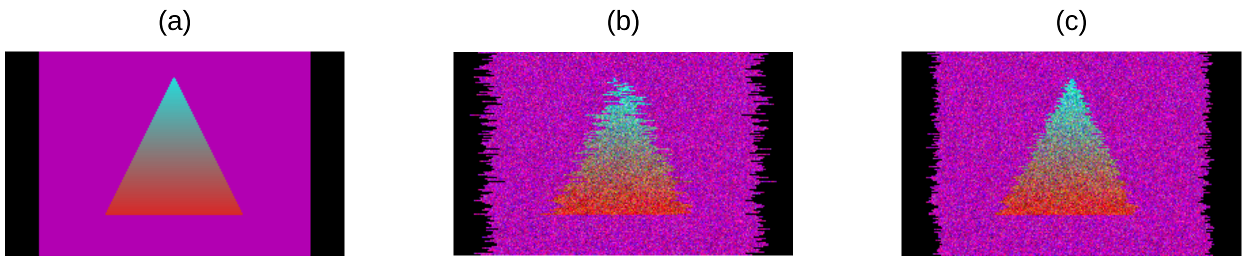

Figure 4: in the same order as Fig. 3, but with intensity Gaussian white noise; the dejittered estimation in (c) shows the robustness of our model to the perturbation of intensity noises.

ical properties and regularization techniques. Under the Bayesian rationale for general restoration problems, the regularities of the slicing moments lead to a variational dejittering model that involves weighted least-square optimization and the total variation Radon measure. The existence and uniqueness of the optimal solutions, as well as the associated computational approaches are all explored under the most general settings and assumptions. In practice, our novel dejittering model introduces dimensionality reduction and gains remarkable computational efficiency.

Our future work will focus on improving the model to achieve maximal degrees of accuracy, performance, and computational efficiency.

Acknowledgments

We thank Stan Osher and Tony Chan for their constant inspirations and encouragement on this emerging field of mathematical image and vision analysis.

References

[1] R. Acar and C. R. Vogel. Analysis of total variation penalty methods for ill-posed problems. *Inverse Prob.*, 10:1217–1229, 1994.

[2] J.-F. Aujol and S.-H. Kang. Color image decomposition and restoration. *Journal of Visual Communication and Image Representation (to appear)*, 2005.

[3] A. Chambolle and P. L. Lions. Image recovery via Total Variational minimization and related problems. *Numer. Math.*, 76:167–188, 1997.

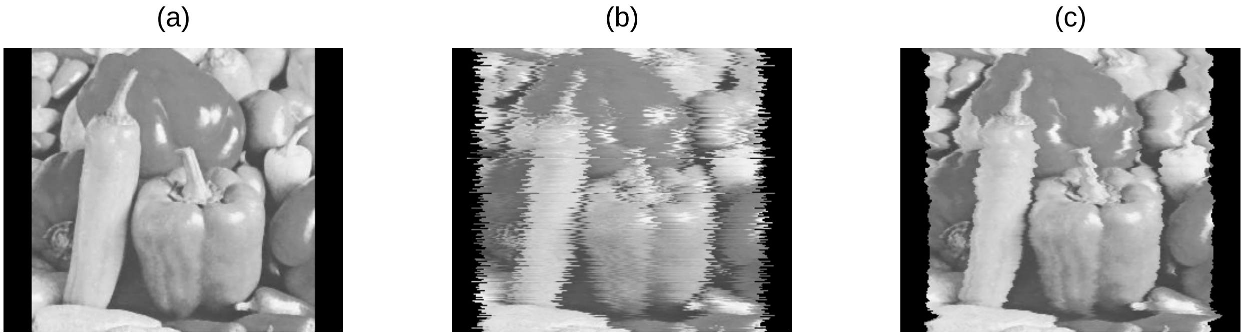

Figure 5: dejittering a standard test image of peppers via moment regularization.

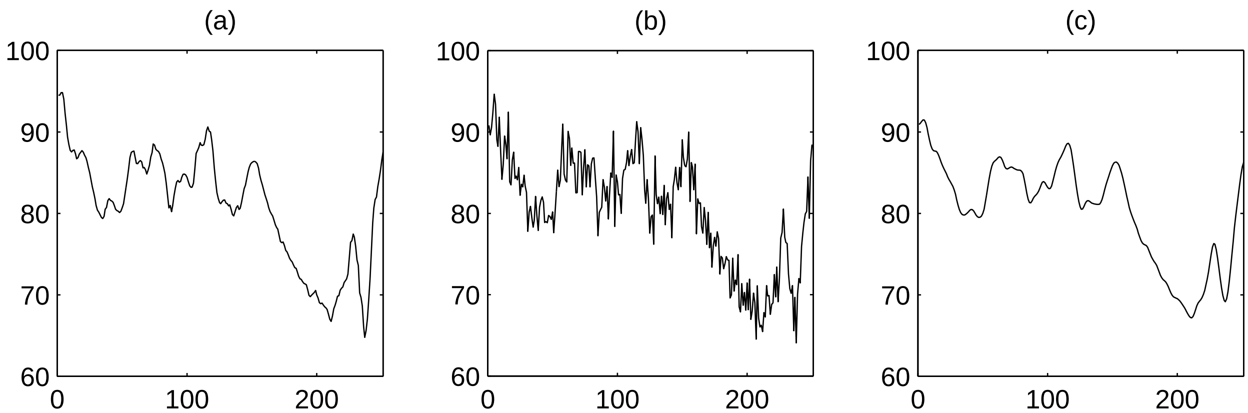

Figure 6: the associated moment profiles corresponding to the images in Fig. 5.

[4] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20:1964–1977, 1999.

[5] T. F. Chan, S.-H. Kang, and J. Shen. Total variation denoising and enhancement of color images based on the CB and HSV color models. *J. Visual Comm. Image Rep.*, 12(4):422–435, 2001.

[6] T. F. Chan, S. Osher, and J. Shen. The digital TV filter and non-linear denoising. *IEEE Trans. Image Process.*, 10(2):231–241, 2001.

[7] T. F. Chan and J. Shen. *Image Processing and Analysis: variational, PDE, wavelets, and stochastic methods*. SIAM Publisher, Philadelphia, 2005.

[8] T. F. Chan and J. Shen. Variational image inpainting. *Comm. Pure Applied Math.,* in press, 2005.

[9] T. F. Chan, J. Shen, and L. Vese. Variational PDE models in image processing. *Notices Amer. Math. Soc.*, 50:14–26, 2003.

[10] I. Daubechies. *Ten lectures on wavelets*. SIAM, Philadelphia, 1992.

Figure 7: The performance of the new model on the standard test image of "Barbara." Dejittering images with rich textures has been a challenging task for PDE (or diffusion) based methods [14, 23].

[11] G. B. Folland. *Real Analysis - Modern Techniques and Their Applications.* John Wiley & Sons, Inc., second edition, 1999.

[12] 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.

[13] E. Giusti. *Minimal Surfaces and Functions of Bounded Variation.* Birkhäuser, Boston, 1984.

[14] S.-H. Kang and J. Shen. Video dejittering by bake and shake. *UCLA CAM 04-65*, 2004.

[15] A. Kokaram and P. Rayner. An algorithm for line registration of TV images based on a 2-D AR model. *Signal Processing VI, Theories and Applications*, pages 1283–1286, 1992.

[16] A. Kokaram, P. M. B. Roosmalen, P. Rayner, and J. Biemond. Line registration of jittered video. *IEEE Int'l Conference on Acoustics, Speech, and Signal Processing*, pages 2553–2556, 1997.

[17] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, volume 22 of *University Lecture Series*. AMS, Providence, 2001.

[18] D. Mumford. *Geometry Driven Diffusion in Computer Vision*, chapter "The Bayesian rationale for energy functionals", pages 141–153. Kluwer Academic, 1994.

[19] D. Mumford. Pattern theory: The mathematics of perception. *Int'l Congress Mathematicians (ICM)*, III, Beijing, 2002.

[20] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Applied. Math.*, 42:577–685, 1989.

[21] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[22] J. Shen. On the foundations of vision modeling I. Weber's law and Weberized TV restoration. *Physica D: Nonlinear Phenomena*, 175:241–251, 2003.

[23] J. Shen. Bayesian video dejittering by BV image model. *SIAM J. Appl. Math.*, 64(5):1691–1708, 2004.

[24] L. A. Vese. A study in the BV space of a denoising-deblurring variational problem. *Appl. Math. Optim.*, 44(2):131–161, 2001.

[25] L. A. Vese and S. J. Osher. Modeling textures with Total Variation minimization and oscillating patterns in image processing. UCLA CAM Tech. Report 02-19, May 2002.

[26] C. Vogel. *Computational Methods for Inverse Problems*. SIAM, Philadelphia, 2002.

[27] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, Germany, 1998.

[28] S. C. Zhu, Y. N. Wu, and D. Mumford. Minimax entropy principle and its applications to texture modeling. *Neural Computation*, 9:1627–1660, 1997.