Bayesian Video Dejittering by BV Image Model

Jackie (Jianhong) Shen

School of Mathematics

University of Minnesota, Minneapolis

jhshen@math.umn.edu

www.math.umn.edu/~jhshen

Group webpage:

www.math.ucla.edu/~imagers

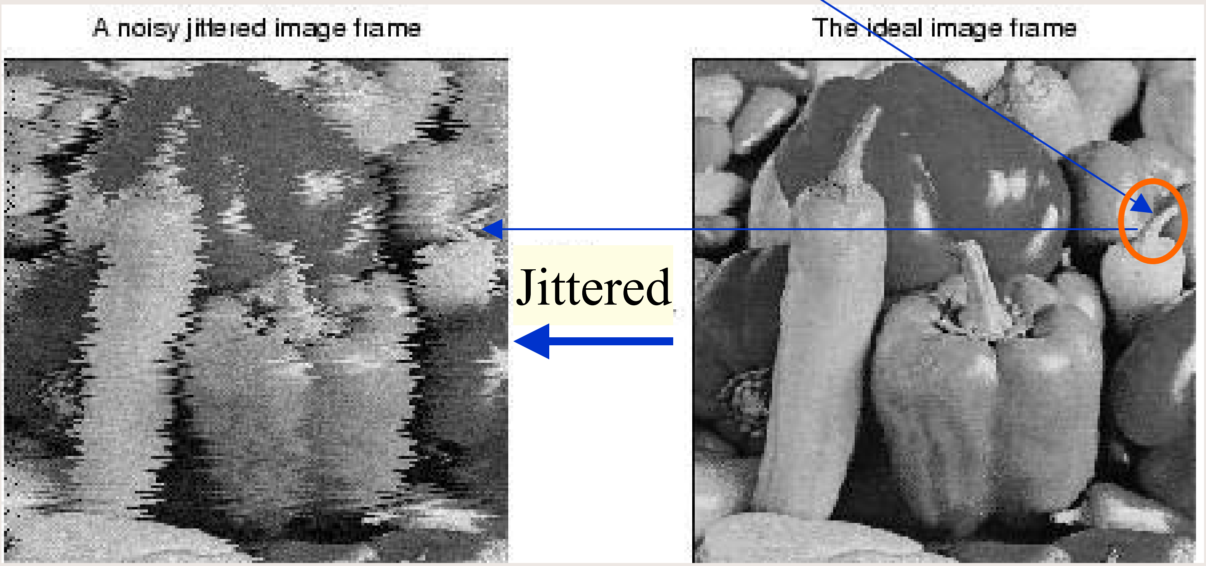
Prepared exclusively for the LANL workshop, December 2002

Overview

- Why Video Jittering and De-jittering
- Ingredients of Our New Model:
  - Statistical Models of Jittering and Noise
  - Dejittering as a Bayesian Inference Problem
  - From Bayesian to Variational: Data Model and Prior Models
- Model Analysis: Existence, Uniqueness, Convergence…
- Model Computation
  - The Alternating Minimization (AM) Algorithm
  - Nonlinear PDE Method for Image Estimation
  - Newton-Raphson Method for Jitter Estimation
- Conclusion

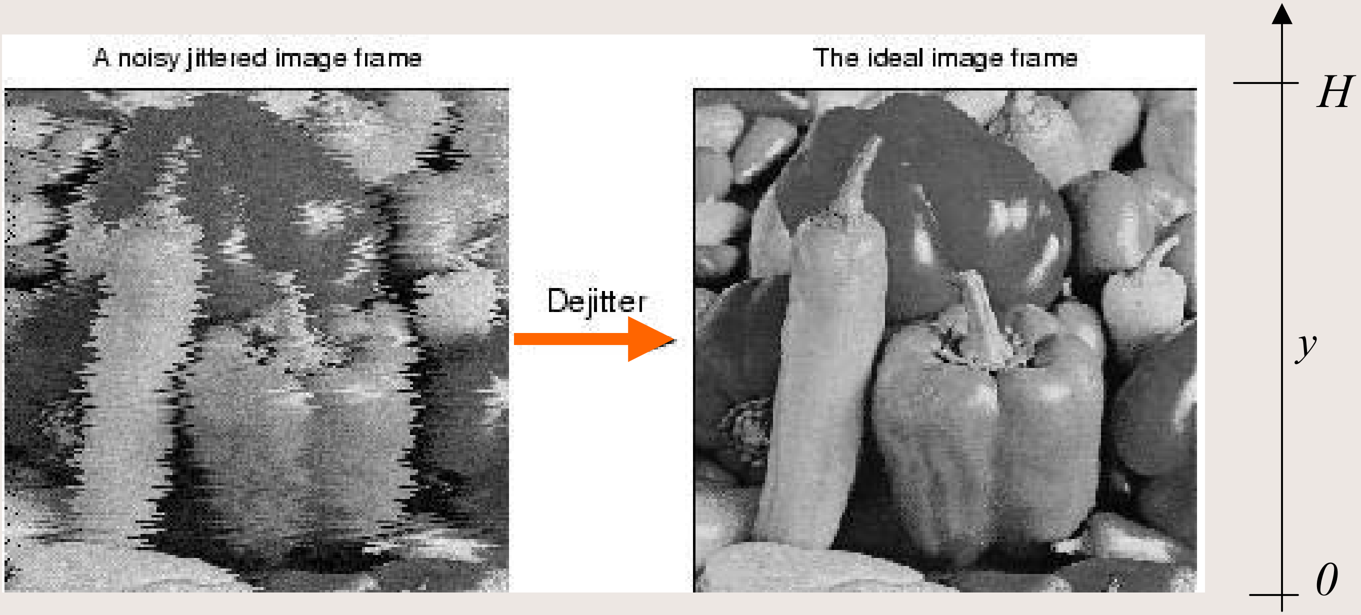
Video Jittering: Why and How

- Why: Line synchronization signals (in video tapes or wireless communication) are destructed or lost.
- How look like: (a) Annoying (b) Small details hard to see



Mission of Video Dejittering

- Given a single observation, recover the original one
- Simultaneously deal with possible intensity noise



Statistical Description of Jittering

- Random jittering:

  Associated to each horizontal line labeled by y, there is a random jitter (or displacement) s(y), so that the y-line is jittered:

  $$u(x, y) \rightarrow u_s(x,y)=u(x+s(y), y).$$

  We model s by a Gaussian white noise with variance σ_s^2 .

- Intensity noise:

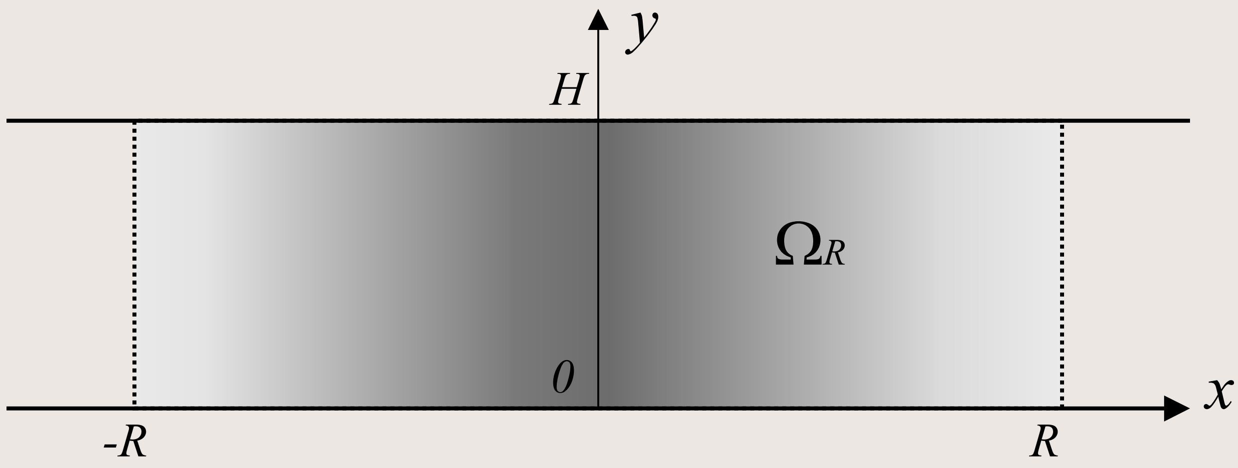
  Due to the degradation of video storage media, image signal is often further polluted by some intensity noise:

  $$u_s(x, y) \rightarrow u_o(x,y)= u_s(x,y)+n(x,y).$$

  We assume that the noise is additive and Gaussian with σ_n^2 .

- Dejittering as an inverse problem: $u_o \rightarrow u$.

Image domain Ω suitable for the dejittering problem



- We work with an infinite stripe domain Ω: IR X $(0, H)$, to practically allow any horizontal Gaussian jittering.
- In numerical computation and other steps, we may also pay attention to finite box regions $\Omega_R = (-R, R) \times (0, H)$.

Existing Methods

- Time Base Corrector (non-intrinsic): recover the line synchronization signatures by denoising the *non-picture* parts of video signals. It demands information *unrelated* to the image or video contents.

- Intrinsic Dejittering : recover the original ideal image frame *directly* based on the observed jittered image data. It relies on the intrinsic image *structures and features*. Works by **Kokaram-Rayner (1992), Kokaram et al. (1997):**
  - Use autoregressive (AR) image models
  - Line registration techniques
  - Local or semi-local filtering
  - Compensated by various image processing tools

Contributions of Our Approach

- Propose the first variational model in the literature, which is built upon the rationale of Bayesian inference.
- The model is clean (a single formula for the objective), self-contained (no pre- or post- processing is needed), and explicitly combines dejittering with denoising.
- Propose to apply the BV (*bounded variation*) image model of Rudin-Osher-Fatemi to restore the regularity of jittered object boundaries.
- Reveal some important properties (existence, uniqueness, and convergence…) of the model by the direct method in Calculus of Variations and the BV function space.
- Design an effective iterative algorithm for the nonlinear and non-convex objective, which is implemented by computational PDEs and nonlinear optimization tools.

Our Approach: Bayesian Inference

- Dejittering as Bayesian inference:

$$p(u, s \mid u_o) = p(u_o \mid u, s)\, p(u, s) \,/\, p(u_o).$$

  Meaning:

  - $p(u, s \mid u_o)$: <u>posterior probability</u> (after observation is given).
  - $p(u_o \mid u, s)$: <u>data model</u>, telling how the observation is generated.
  - $p(u, s)$: <u>prior model</u>. Jittering is often independent of image information. Therefore, $p(u, s) = p(u)\, p(s)$.
  - $p(u_o)$: is probability normalization after observation is made, makes no essential contribution to the inference.

- Bayesian inference is MAP : max $p(u, s \mid u_o)$

Bayesian Goes Variational: Gibbs' Formula

- Taking logarithm likelihood function $E[X] = -\beta \ln p(X)$:

$$E[u, s \mid u_o] = E[u_o \mid u, s] + E[u, s] + \textit{const.}$$

And MAP becomes "energy" minimization:

$$\max\ E[u, s \mid u_o].$$

If $\beta = \kappa T =$ Boltzmann constant * Absolute temperature, then we are formally working with the ensemble energy instead of probability according to *Gibbs formula*: $p = 1/Z \exp(-E/\beta)$.

- The data model is simple due to Gaussian noise model:

$$\boxed{E[u_0 \mid u, s] = \lim_{R \to \infty} \frac{1}{2\sigma_n^2 \mid \Omega_R \mid} \int_{\Omega_R} (u_0 - u_s)^2 dxdy.}$$

- The prior is *separable*: $E[u, s] = E[u]\, E[s]$.

(cont'd) prior models for jittering and image

$$E\,[u, s] = E\,[s] + E[u]$$

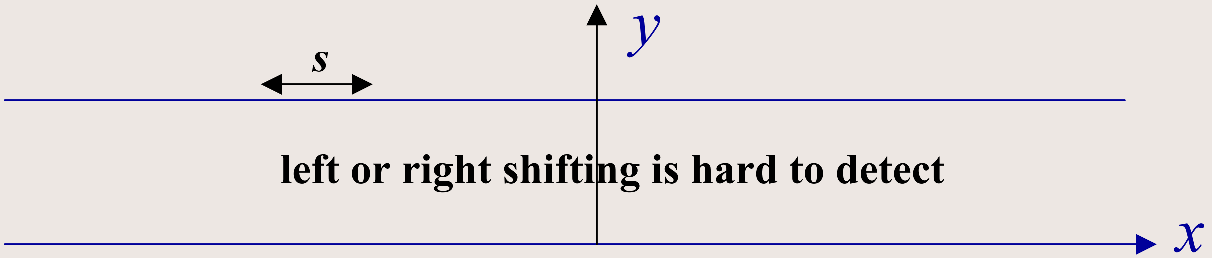
- The prior model for jittering s:

$$\boxed{E[s] = \frac{1}{2\sigma_s^2 H}\int_0^H s^2(y)\,dy\ .}$$

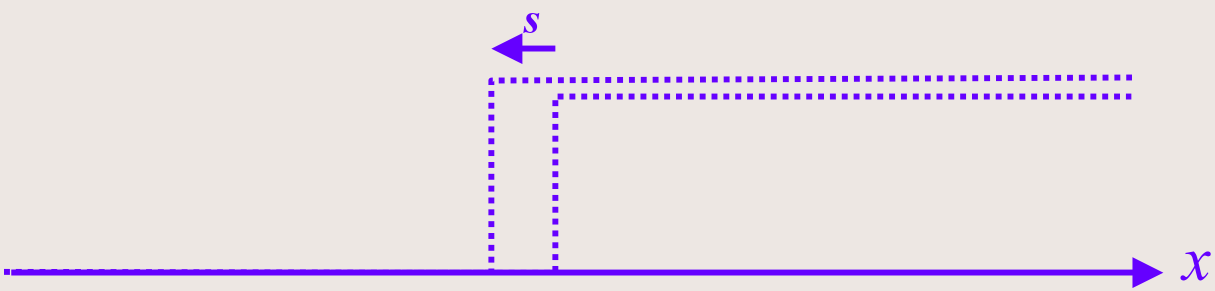
- The prior model for image u: $E[u]$
  - Looking for good image priors is the fundamental problem in IP.
  - Roughly classified into two categories:
    - Stochastic models: lattice models/Gibbs Random fields **(Geman-Geman, 1984)**, and statistical learning based on filtering and the maximum entropy principle **(Zhu-Wu-Mumford, 1997)**.
    - Function spaces : deterministic; regularity based: square integrable functions (classical spectral/Fourier method); Sobolev functions (linear filtering theory); functions with bounded variations (BV, **ROF,1992**); Mumford-Shah piecewise smooth model (**1989**)…

The Role of Edges in Video Dejittering

Clue or Clueless (1-D image for example)?



Scenario I. Image is very smooth and featureless.



Scenario II. Image contains clearly identifiable edges.

Edge-Aware Deterministic Image Models

- Edges give crucial visual info for objects recognition and representation (such as wavelets), important in computer vision from the very beginning **(David Marr).**

- Mumford-Shah piecewise smooth image model **(1989)**

  $$E[u, G] = E[u \mid \Gamma] + E[\Gamma] = \text{Sobolev} + \text{Hausdorff}$$

  Edge set is singled out; the rest is measured by Sobolev.

  Very powerful for segmentation. Could be costly for other applications (where edge is still a crucial hidden feature but not particularly interesting or explicitly required, such as inpainting & dejittering).

- Rudin-Osher-Fatemi's BV image model **(1992):** $E[u]=\mathrm{TV}[u]$.

Why BV Image Model

- BV(Ω)={u|integrable & with finite total variation TV[u]}:

$$\mathrm{TV}\,[u] = \int_\Omega |Du| = \sup_{\text{smooth}\ \mathbf{f}:|\mathbf{f}|\le 1} \int_\Omega u\nabla\cdot\mathbf{f}\,dx\,.$$

The Sobolev space W(1,1) is its subspace, for which

$$\mathrm{TV}\,[u] = \int_\Omega |\nabla u|\,dx = \int_\Omega \sqrt{u_{x_1}^{\ 2} + u_{x_2}^{\ 2}}\,dx\,.$$

Generally, TV is a Radon measure, and legalizes edges.

- But edges are not explicitly singled out, which greatly reduces cost, especially in computation.
- Geometry comes clear: **The Co-area Formula** **(De Giorgi, 1961)**

$$E[u] = \int_{-\infty}^{\infty} \mathrm{Per}(u<\lambda,\Omega)\,d\lambda \overset{\text{smooth } u}{\Rightarrow} \int_{-\infty}^{\infty} \mathrm{length}(u=\lambda)\,d\lambda.$$

Variational Dejittering Model on an Infinite Stripe

- The combination of the above three models:
  - Gaussian jittering $E[s]$ (define $\mu = 1/[\,\mathrm{var}(s)\, H\,]$).
  - Gaussian noise $E[u_o \mid u,\ s]$ (define $\lambda_R = 1/[\,\mathrm{var}(n)\, |\Omega_R|\,]$).
  - BV image model $E[u]$

  leads to the variational dejittering model on Ω .

$$E[u, s \mid u_0] = \lim_{R\to\infty} \frac{\lambda_R}{2} \int_{\Omega_R} (u_0 - u_s)^2\, dxdy \; + \frac{\mu}{2}\int_0^H s^2(y)\,dy + \alpha \int_{\Omega} |\, Du \,| .$$

- Assumptions & admissible conditions:
  - s is in $L_2(0, H)$.
  - u is in BV(Ω). [By Sobolev embedding, u is in $L_2(\Omega)$.]
  - There exists some $w(x,y)$ in BV so that $n = u_o - w_s$ satisfies

$$\lim_{R\to\infty} \frac{1}{|\,\Omega_R\,|} \int_{\Omega_R} n(x,y)^2\, dxdy \; = \sigma_n^2 \quad \text{exists.}$$

Why Does the Model Turn Out to Be Uninteresting

- **Theorem A.** *Let u_o be a given noisy jittered image in $L^2_{\text{loc}}(\Omega)$, and there exists some $w(x,y)$ mentioned in the previous slide. Then,*

$$(u=0,\ s=0) = \operatorname{argmin} E[\,u, s \mid u_o\,].$$

- The problem is caused by the <u>averaging near infinity</u>.
  - Functions in BV die out suitably at infinity, and their behavior on any compact domain is blanked out by the averaging.
  - Or simply speaking, the infinity over reacts.

- How can the model be appropriately fixed…

Improved Model and Admissible Conditions

- In practice, finite images are extended by zero-padding.
- To cool down the over reaction of the infinity, we modify the model to

$$E[u, s \mid u_0] = \frac{\lambda}{2}\int_\Omega (u_0 - u_s)^2\,dxdy + \frac{\mu}{2}\int_0^H s^2(y)dy + \alpha\int_\Omega |Du|.$$

- Assumptions and Parameters:
  - μ, α as before; $\lambda = \beta$/var(n) now introduces the 2nd parameter;
  - s is in $L_2(0, H)$ and u in BV(Ω);
  - the observation u_o is in $L_2(\Omega)$;
  - by Sobolev embedding & Fubini Theorem, u_s is in $L_2(\Omega)$.

Separation to Two "Conditional" Energies

$$E[u,s\,|\,u_0] = \frac{\lambda}{2}\int_\Omega (u_0 - u_s)^2\,dxdy + \frac{\mu}{2}\int_0^H s^2(y)\,dy + \alpha\int_\Omega |\,Du\,|.$$

- Nonlinear, non-quadratic, non-convex (due to jittering)
- Existence and uniqueness become hard to analyze
- Thus, we define two "conditional" energies:

u-estimation:

$$E[u\,|\,u_0, s] = \frac{\lambda}{2}\int_\Omega (u_0 - u_s)^2\,dxdy + \alpha\int_\Omega |\,Du\,|.$$

s-estimation:

$$E[s\,|\,u_0, u] = \frac{\lambda}{2}\int_\Omega (u_0 - u_s)^2\,dxdy + \frac{\mu}{2}\int_0^H s^2(y)\,dy.$$

- Also important for our algorithm.

Existence & Uniqueness of $E[u \mid u_0, s]$

- **Theorem B.** *For any given jitter s, the minimizer in* BV *for $E[u \mid u_0, s]$ is unique. Furthermore, if there exists a minimizing sequence bounded in L_1, then the minimizer exists.*

  - ❖ Unlike in conventional restoration problems (such as denoising) on finite domains (**e.g., Chambolle and Lions**), the L_1 boundedness is not a byproduct of a minimizing sequence.
  - ❖ An example:

$$u^n = \frac{1}{(1+x^2)^{\frac{n+1}{2n}}}, \qquad n = 1,2,\ldots$$

    - with both bounded L_2 norms and TV measures;
    - yet L_1 norms diverge to infinity.

Separability of $E[s \mid u_0, u]$

- For each y, define $f_0(x)=u_0(x, y)$, $f(x)=u(x, y)$; and

$$e(t,y)=\frac{\mu}{2}t^2+\frac{\lambda}{2}\int_{IR}[f_0(x)-f(x+t)]^2\,dx.$$

- Then $E[\, s|\, u_o\, ,\, u]$ could be re-written as

$$E[s\,|\,u_0,u]=\int_0^H e(s(y),y)\,dy.$$

- Therefore, functional minimization becomes the minimization of a collection of 1-D functions $e(t)=e(t, y)$, generally non-quadratic and convex. Thus uniqueness is not guaranteed. We can say more.

Existence and A Priori Bound for E[s | u0, u]

- **Theorem C.** **(following notation of slide –1)** Suppose that $f_0(x)$ and $f(x)$ are both in L_2 (IR). Then
  - $e(t)$ is continuous and the minimizer exists.
  - Suppose that $t = s$ is one minimizer. Then

$$|s| \leq \sqrt{\lambda / \mu}\left(\|f_0\| + \|f\|\right).$$

- Continuity follows from Lebesgue integration theory: $f(x+t) \rightarrow f(x)$ in L_p, for any finite $p >= 1$.

Algorithm and Analysis: The AM-Algorithm

- The Alternating Minimization (AM) algorithm:
  - ❖ starting with the zero jittering estimation: $s_0=0$;
  - ❖ iteratively updating u^{n+1} and s_{n+1} by
    - *u-estimation* $\boxed{u^{n+1} = \arg\min E[u \mid u_0, s_n]}$
    - *s-estimation* $\boxed{s_{n+1} = \arg\min E[s \mid u_0, u^{n+1}]}$

- **Theorem D. The AM algorithm is consistently down-hill.**

$$\boxed{E[u^{n+1}, s_{n+1} \mid u_0] \le E[u^n, s_n \mid u_0]}$$

Weak Convergence Theorem for the u-Estimation

- **Theorem E.** Suppose the jittered image u_0 is L_2 & $u_0(-, y)$ is continuous in x for $a.e.\ y$ in $(0, H)$. Let u^n and s_n be the sequence generated from AM. Suppose that (s_n) $a.e.$ converges to some $s(y)$ in L_2, and the L_1 norms of (u^n) does not converge to ∞. Then there must exist a subsequence (m) of (n), and some u in BV, so that $u^m \to u$ in L_1, and

$$\boxed{E[u, s \mid u_0] \le \lim_n E[u^n, s_n \mid u_0].}$$

u-Estimation Based on Nonlinear PDE

$$\boxed{E[u \mid u_0, s] = \frac{\lambda}{2}\int_\Omega (u_0 - u_s)^2\, dxdy \; + \alpha \int_\Omega | Du | .}$$

- Let T_s denote the jittering operator, then

$$T_s^* = T_{-s} = T_s^{-1}.$$

- The formal Euler-Lagrange equation is (***distributional sense***)

$$0 = -\alpha\nabla \cdot \left[\frac{\nabla u}{|\nabla u|}\right] + \lambda T_s^*(T_s u - u_0) = -\alpha\nabla \cdot \left[\frac{\nabla u}{|\nabla u|}\right] + \lambda(u - u_{0.-s}),$$

  Define $v_0 = u_{0,-s}$. Surprisingly it is identical to the E-L equation of Rudin-Osher-Fatemi applied to v_0.

- Thus techniques such as <u>viscosity approximation</u> and <u>lagged diffusivity iteration</u> can be applied (**Vogel, 1997**).

Differentiating the s-Estimation

$$\boxed{e(t)=\frac{\mu}{2}t^2+\frac{\lambda}{2}\int_{IR}[f_0(x)-f(x+t)]^2\,dx.}$$

- Let <+, +> denote the inner produce in $L_2(IR)$.
- Assume that $f(x)$ is in Sobolev space W(1,2). Then

$$\boxed{e'(t)=\mu t-\lambda\langle f_0(x)-f(x+t), f'(x+t)\rangle.}$$

- Further assume the vanishing conditions at infinity. Take the second order differentiation and apply integration by parts:

$$\boxed{e''(t)=\mu+\lambda\langle f_0(x), -f''(x+t)\rangle.}$$

Newton-Raphson and Its Robustness

$$s = \arg\min e(t) = \frac{\mu}{2}t^2 + \frac{\lambda}{2}\int_{IR}[f_0(x) - f(x+t)]^2\,dx.$$

- Apply the Newton-Raphson method for s:

$$s_{n+1} = s_n - \frac{e'(s_n)}{e''(s_n)} = \frac{s_n\langle f_0, -f''(x+s_n)\rangle + \langle f_0 - f(x+s_n), f'(x+s_n)\rangle}{\mu/\lambda + \langle f_0, -f''(x+s_n)\rangle}$$

- In the noise-free case (i.e. infinite λ):

$$s_{n+1} = s_n + \frac{\langle f_0 - f(x+s_n), f'(x+s_n)\rangle}{\langle f_0, -f''(x+s_n)\rangle}$$

- **Proposition** (robustness of N-R algorithm).

  If $f_0(x) = g(x+s) + n(x)$ for some $g(x)$ in W(1,2), then as (f, t) gets closer to (g, s), $e''(t) \approx \mu + \lambda\|g'\|_{L_2}^2 \geq \mu > 0.$

Implementation Issues (I): Neumann Jittering Model

- In numerical simulation and real applications, images are given on a <u>finite</u> rectangular domain $\Omega_R=(-R, R) \times (0, H)$.

- We therefore need a boundary jittering model along x.

- In our simulation, we adopted <u>Neumann jittering model</u>:

$$u_s(x, y) = \begin{cases} u(x+s(y), y), & \text{if } |x+s(y)| \le R. \\ u(\pm R, y), & \text{if } \pm(x+s(y)) > R. \end{cases}$$

Implementation Issues (II): Parameter Tuning

- For the finite domain $\Omega_R=(-R, R)$ x $(0, H)$, the data model on the infinite stripe Ω is replaced by:

$$\boxed{E[u_0 \mid u, s] = \frac{\lambda_R}{2} \int_{\Omega_R} (u_0 - u_s)^2 \, dxdy \, .}$$

with $\lambda_R = 1/[\ \mathrm{var}(n) \mid \Omega_R \mid\ ]$.

- Since μ for the jittering energy is also explicitly known, we have <u>only one</u> tunable parameter in implementation, namely, α , the weight for the TV measure.

- Rich literature in Inverse Problems for selecting a good α (see for example, the recent monograph by Vogel, **2002, SIAM**).

Implementation Issues (III): Jitter Simulation via Binomial Distribution

- In digital simulation/real television display, jitters can occur only at whole-pixel length: 1, 2, 3, …
- We could simulate them by the continuous Gaussian, followed by a rounding process (e.g. round (3.14)=3).
- In our simulation, we directly apply the binomial noise model:

$$\text{Prob}\ (s = k) = \binom{2n}{n+k} p^{n+k}(1-p)^{n-k}, \quad k = -n : n.$$

$$E(s) = n(1-2p), \qquad \sigma_s^2 = 2np(1-p).$$

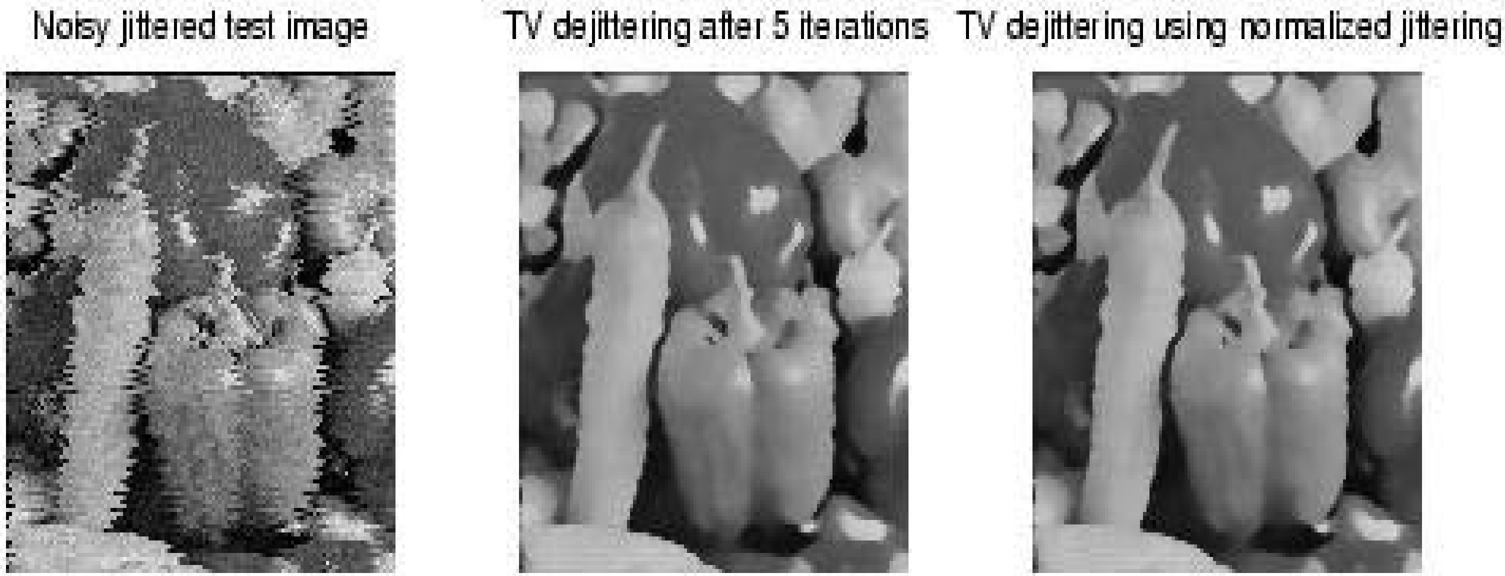
- p=1/2 for zero mean. By the *Central Limit Theorem*, Gaussian is still a good approximation. Thus the squared norm is still valid in the model.

Numerical Examples (I): BV Based Bayesian Dejittering

- Left: Noisy jittered test video frame
- Middle: First output from AM, assuming no jittering
- Right: Final output from AM for the model

Numerical Example (II): With Variance Normalization

- Left: Noisy jittered test video frame
- Middel: Output based on s_5 at step 5 of AM algorithm.
- Right: Output based on $\hat{s}_5 = s_5 \times \sigma_s \div \mathrm{std}(s_5)$.



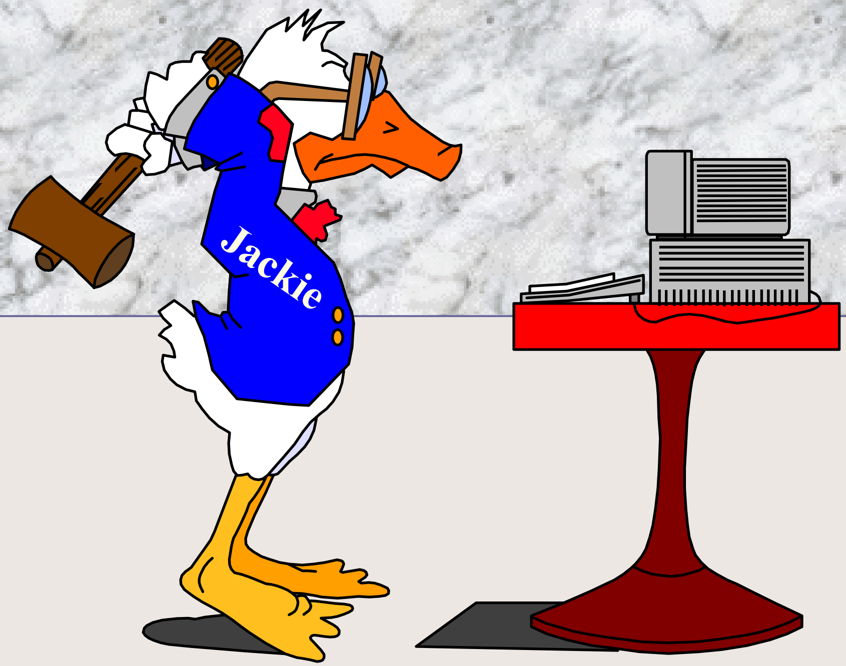
Variance normalization does help recover better small scale details.

Summary, Key Point, and Look Ahead

- (**Modeling**) We have developed the first variational model for intrinsic video dejittering, which is self-contained and naturally combines dejittering with denoising.
- (**Analysis**) We have analyzed some important properties (such as existence, uniqueness, & convergence) of the model.
- (**Computation**) We have proposed the AM (alternating minimization) algorithm for the non-quadratic & non-convex objective, which is then computationally implemented by techniques from nonlinear PDEs and nonlinear optimization.
- (**Key point**) It again shows the power of a good image model.
- (**Look Ahead**) If there exist inter-frame correlations, we expect that dynamic tools such as the Kalman filter can play important roles in modeling and computation.

That is all, folks…

Thank you for your patience!



Acknowledgments

- School of Mathematics & IMA, UMN; IPAM at UCLA
- Tony Chan & Stan Osher (UCLA) for PDE/BV/Variational
- Gil Strang (MIT) for his vision & support on research
- David Mumford & Stu Geman (Brown) for Bayesian
- F. Santosa, P. Olver, R. Gulliver, & W. Miller (UMN)
- Program of Applied Mathematics, NSF.