\[ \int_{\Omega} f(x, u(x), \nabla u(x)) \, dx \quad \Leftrightarrow \quad \sup_{\phi \in K} \int_{\Omega \times \mathbb{R}} \phi \cdot D1_u \]
• Energy minimization approaches
  • Bayesian Approach
  • Markov Random Fields
  • Total Variation
Bayesian approach

- Images are not deterministic
- Is there a probabilistic approach to deal with image processing problems?
- Let‘s have a look at Bayesian decision theory!
- The joint probability of observing both, \( u \) and \( f \) is given by:

\[
p(u, f) = p(f \mid u)p(u) = p(u \mid f)p(f)
\]

\( u \) ... unknown quantity
\( f \) ... given data
Bayesian approach

- Re-arranging leads to the well-known Bayesian formula.
- The Bayesian approach provides a framework to infer "optimal solutions".
- Converts the prior probability to a posterior probability.

\[
p(u|f) = \frac{p(u)p(f|u)}{p(f)}
\]

- \(p(u|f)\): a-posteriori
- \(p(u)p(f|u)\): a-priori likelihood
- \(p(f)\): normalization
Bayesian approach

- Two main questions:
  1. How to model the a-priori distribution and the likelihood distribution? Convex/non-convex models, learning, …
  2. Bayesian approach delivers a complete distribution. Which sample should be selected as the solution? MAP, expectation, median, …

- Both questions are subject to current research and we are far away from having good answers
- In a few cases, good answers are available
Ad 1. A generic distribution

- Let us consider the following generic distribution, well known in statistical mechanics [Gibbs, 1889]

\[ p(u|f) = \frac{1}{Z(f)} e^{-\left( P(u) + D(u,f) \right)} \]

- \( P(u) \) … prior, regularization, or smoothness term
- \( D(u) \) … data fitting, data fidelity, or likelihood term

- How to model these two terms?
- Let us consider a simple image restoration example
Modeling the likelihood distribution

- Assume a simple, linear image formation model
  \[ f = Au + n \]
  Lin. Operator Noise
- Assume, density of the noise follows a i.i.d. Gaussian
  \[ n(x) = f(x) - (Au)(x) \sim N(0, \sigma^2) \]
- Leads to the following density for the likelihood
  \[ p(f|u) = \prod_{x} e^{-\frac{|Au(x) - f(x)|^2}{2\sigma^2}} \]
Modeling the a-priori distribution

- Major aim of the prior term is to favor physically feasible solutions
- Natural images exhibit a lot of spatial regularity
- A natural choice is
  \[ p(u) = \prod_{x} e^{-\frac{1}{2} |\nabla u(x)|^\alpha}, \quad \alpha > 0 \]
- Image gradient: \( \nabla u(x) = \left( \frac{\partial u(x)}{\partial x}, \frac{\partial u(x)}{\partial y} \right)^T \)
- What is a good \( \alpha \)?
Look at natural images
Statistics of natural images

- Statistics of zero mean linear filters [Huang&Mumford, 1999]
- Natural images exhibit heavy-tailed distributions

True statistics

Models

\[ f(x) = |x|^\alpha \]
Example

- Denosing of a color image using different priors

\[
\alpha = 0.5 \quad \alpha = 1 \quad \alpha = 2
\]
Discussion

• **Best** results for the model that matches the true statistics \( (\alpha = 0.5) \)

• **Good** results for the linear model \( (\alpha = 1) \)

• **Bad** results for the quadratic model \( (\alpha = 2) \)

• Optimization problem is
  - non-convex for \( \alpha < 1 \)
  - Convex for \( \alpha \geq 1 \)
Convex versus non-convex

- **Convex:**
  - local minimum = global minimum
  - Gradient descend leads to global minimum
- **Non-convex:**
  - Many local minima
  - Gradient descend leads to local minimum
The complete distribution

- **Data model**
  \[ p(f|u) = \prod_x e^{-\frac{|Au(x) - f(x)|^2}{2\sigma^2}} \]

- **Prior model**
  \[ p(u) = \prod_x e^{-|\nabla u(x)|^\alpha}, \quad \alpha = 1 \]

- **Posterior Probability**
  \[
  p(u|f) \sim p(f|u)p(u) = \prod_x e^{-\left(|\nabla u(x)| + \frac{1}{2\sigma^2} |Au(x) - f(x)|^2\right)} \\
  = e^{-\left(\sum_x |\nabla u(x)| + \frac{1}{2\sigma^2} \sum_x |Au(x) - f(x)|^2\right)}
  \]
Ad 2. Which sample should be selected?

- Consider three simple posterior distributions $p(u \mid f)$

1. Unimodal, symmetric: Max = Mean = Median
2. Bimodal, symmetric: Max ≠ Mean = Median
3. Unimodal, non-symmetric: Max ≠ Mean ≠ Median
Computing the expectation

• Computing the **expectation** (weighted mean) amounts to evaluate the high-dimensional integral

\[
\bar{u} = \frac{1}{Z(f)} \int u e^{-\left(\mathcal{P}(u) + \mathcal{D}(u, f)\right)} \, du
\]

• Consider an image of $512 \times 512$ pixels and 256 gray values. This leads to $4.54 \cdot 10^{631305}$ possible images!

• Subtle algorithms such as MCMC are needed to estimate the integral.
Computing the MAP

- The so-called \textbf{maximum a-posteriori} (MAP) leads to

\[ u^* = \max_u p(u | f) = \min_u \mathcal{P}(u) + \mathcal{D}(u, f) \]

- Leads to a well-defined minimization problem
- Very often, \textbf{fast algorithms} are available for minimization
- Sometimes, computing the MAP leads to model distortions in the solution
MAP vs. Expectation

• In case of the ROF model, it was recently shown that the expectation leads to improved results.
Putting all together

• Posterior distribution of the model

\[ p(u|f) \sim e^{-\left( \sum_x |\nabla u(x)| + \frac{1}{2\sigma^2} \sum_x |Au(x) - f(x)|^2 \right)} \]

Maximum-a-posteriori (MAP) estimation:

– Discrete setting: pixels are nodes of a graph: „MRF Approach“

\[ u^* = \arg\min_u \left\{ \sum_x |\nabla u(x)| + \frac{1}{2\sigma^2} \sum_x |Au(x) - f(x)|^2 \right\} \]

– Continuous setting: images are functions: „Variational Approach“

\[ u^* = \arg\min_u \left\{ \int_\Omega |\nabla u|dx + \frac{1}{2\sigma^2} \int_\Omega |Au - f|^2dx \right\} \]
Discrete versus continuous

The world is continuous, but the mind is discrete

[D. Mumford]

• Advantages of discrete MRF approaches
  – Often work on arbitrary graph structures
  – Powerful combinatorial optimization algorithms
  – Sometimes easier to solve (finite-dim. solution space)

• Advantages of continuous Variational approaches
  – Analysis of the properties of a minimizer (existence, uniqueness, regularity, …)
  – Great flexibility in choosing discretizations
  – Powerful algorithms from continuous (convex) optimization
MRF approach

• A Markov Random Field (MRF) is a graphical model in which a set of random variables have a Markov property described by an undirected graph

• Markov property ⇔ memoryless random variables

• Images
  – Random variables are image pixels
  – They are connected with adjacent pixels via a neighborhood system, which is described by the Markov property
  – Images are organized on a regular grid, which can be seen as an undirected graph
Graphical models

- Images are organized as graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
- Node set $\mathcal{V} = \{v_1, v_2, \ldots, v_n\}$
- Edge set $\mathcal{E} = \{(v_1, v_2), \ldots, (v_i, v_j), \ldots\}$

Neighborhood structure
Image labeling

• Many image processing problems can be seen as labeling problems

• Formal task:
  – Assign a label \( l(v_i) \in \mathcal{L} = \{l_1, l_2, \ldots, l_m\} \) to each node \( v_i \in \mathcal{V} \)
  – Such that the labeling \( l = \{l(v_1), \ldots, l(v_n) \in \mathcal{L}^{\mid\mathcal{V}\mid}\} \)
  – Minimizes a certain energy

\[
E(l) = \sum_{(v_i, v_j) \in \mathcal{E}} V(l(v_i), l(v_j)) + \sum_{v_i \in \mathcal{V}} D(l(v_i))
\]

• Image labels can be
  – Object labels
  – Image intensities
  – Depth values
  – Indicator for an edge, …
How many labelings?

- Assume, we have a graphical model with \( n \) nodes and all nodes can take a label of one and the same label set \( \mathcal{L} \).
- The space of all labelings is given by \( \mathcal{L} \times \ldots \times \mathcal{L} = \mathcal{L}^n \).
- Then there are \( |\mathcal{L}|^n \) different labelings.
- Example:
  - Image with 512x512 pixels
  - 256 different gray values
  - There are \( 256^{262144} \) different labelings
  - Too much to try them all
- How can we find the labeling, that minimizes the energy?
- We need smart algorithms (and a few assumptions)
Binary MRFs

- If the label set is binary, i.e. $\mathcal{L} = \{0, 1\}$ and the pairwise terms $V(l(v_i), l(v_j))$ are submodular:

\[
V(0, 0) + V(1, 1) \leq V(1, 0) + V(0, 1)
\]

- the energy $E(l)$ can be minimized globally in polynomial time using graph cut algorithms (e.g. Boykov-Kolmogorov)

- Example image segmentation:
Ising model

- Well-known model from quantum physics
- Binary label set $\mathcal{L} = \{0, 1\}$
- Figure/ground image segmentation

$$E_{\text{Ising}} = \sum_{(v_i, v_j) \in \mathcal{E}} |l(v_i) - l(v_j)| + \sum_{v_i \in \mathcal{V}} f_i l(v_i)$$

- Pairwise term counts label jumps in x- and y-direction and hence approximates the boundary length
- Data term assign figure/ground costs to each pixel, e.g.

$$f_i = (I_i - \mu_{\text{figure}})^2 - (I_i - \mu_{\text{ground}})^2$$

- Minimizer can be computed in polynomial time using graph cuts
Example

User selected regions

Data term \( f \)

Graph cut

Extracted object
Multilabel MRF‘s

- Multi-label MRFs cannot be solved globally in general
- **Exception:** If the label set is ordered and the pairwise terms are convex functions, then a globally optimal solution can be computed by a graph cut on a special graph in higher dimensions (Ishikawa et al.)
- This is for example the case for disparity estimation, where the label set has a natural ordering
Computing approximate solutions

• If the labels do not provide a natural ordering or the pairwise functions are non-convex, there are still algorithms that can find a good local minimizer

• Move-making algorithms (Boykov et al.)
  – Alpha-beta-swap
  – Alpha-expansion

• Solve a sequence of binary submodular problems

(a) Standard-move
(b) Alpha-beta-swap
(c) Alpha-expansion
Potts Model

- Generalization of the Ising model
- Arbitrary label set $\mathcal{L} = \{0, 1, ..., m\}$
- Application to image labeling (street, sky, facade, ...)

$$E_{\text{Potts}} = \sum_{(v_i, v_j) \in \mathcal{E}} \delta(l(v_i), l(v_j)) + \sum_{v_i \in \mathcal{V}} f_i(l(v_i))$$

$$\delta(a, b) = \begin{cases} 
0 & \text{if } a = b \\
1 & \text{else}
\end{cases}$$

- The Potts model is NP-complete
- Can be solved approximately for example by alpha-expansion
Examples

• Uses an additional ordering constraint, which can be incorporated into the Potts model

Strekalovskiy, Cremers, 2011
Nonlinear total variation based noise removal algorithms*

Leonid I. Rudin¹, Stanley Osher and Emad Fatemi²
Cognitech Inc., 2800, 28th Street, Suite 101, Santa Monica, CA 90405, USA

A constrained optimization type of numerical algorithm for removing noise from images is presented. The total variation of the image is minimized subject to constraints involving the statistics of the noise. The constraints are imposed using Lagrange multipliers. The solution is obtained using the gradient-projection method. This amounts to solving a time dependent partial differential equation on a manifold determined by the constraints. As \( t \to \infty \) the solution converges to a steady state which is the denoised image. The numerical algorithm is simple and relatively fast. The results appear to be state-of-the-art for very noisy images. The method is noninvasive, yielding sharp edges in the image. The technique could be interpreted as a first step of moving each level set of the image normal to itself with velocity equal to the curvature of the level set divided by the magnitude of the gradient of the image, and a second step which projects the image back onto the constraint set.

- Defined as the Variational Problem [Rudin, Osher, Fatemi, 1992], [Chambolle, Lions, 1997]

\[
\min_u \int_\Omega |Du| + \frac{\lambda}{2} \|Au - f\|_2^2
\]

Total Variation  \quad L_2 \text{ data term}
Model of Rudin Osher and Fatemi

- “The prototype“ of continuous convex optimization in computer vision
- Probably one of the best image denoising methods in terms of performance versus simplicity
- Does preserve sharp discontinuities in the solution
- Has been extended in various ways
What is the TV norm?

• The TV norm is the $\ell_1$ norm of the $\ell_2$ vector norm of the image gradient

$$TV(u) = \int_\Omega |Du| \approx \int_\Omega |\nabla u| \, dx = \int_\Omega \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2} \, dx$$

• Extension to color images (vectorial TV)

$$TV(u) = \int_\Omega \sqrt{\sum_{k=1}^{n} |\nabla u_k|^2} \, dx, \quad u = (u_1, u_2, \ldots, u_n)$$

• Many other variants possible
Why does it preserve discontinuities?

\[ \int_{\Omega} |Du| = 1.0 \]
\[ \int_{\Omega} |\nabla u|^2 \, dx = 0.01 \]

Total Variation has no bias against discontinuities
Geometric interpretation

• Total variation has a nice geometric interpretation

• Can be written in terms of its level sets

\[ \int_{\Omega} |D u| = \int_{-\infty}^{\infty} \text{Per}(\{x : u(x) > t\}) dt \]

• The total variation of a binary shape is the boundary length of the shape
ROF image denoising

\[
\min_u \int_{\Omega} |Du| + \frac{\lambda}{2} \| u - g \|_2^2
\]
ROF color image denoising

$$\min_u \int_\Omega |Du| + \frac{\lambda}{2} \|u - f\|_2^2$$
Discretization

– We consider an $M \times N$ Cartesian grid

\[
\{(ih, jh) : 1 \leq i \leq M, 1 \leq j \leq N\}
\]

– We define a scalar product in $X$

\[
\langle u, v \rangle_X = \sum_{i,j} u_{i,j} v_{i,j}, \quad u, v \in X
\]

– The gradient operator is defined as

\[
(\nabla u)_{i,j} = \begin{pmatrix}
(\nabla u)^{1}_{i,j} \\
(\nabla u)^{2}_{i,j}
\end{pmatrix}
\]

\[
(\nabla u)^{1}_{i,j} = \begin{cases}
\frac{u_{i,j+1} - u_{i,j}}{h} & \text{if } i < M \\
0 & \text{if } i = M
\end{cases},
(\nabla u)^{2}_{i,j} = \begin{cases}
\frac{u_{i+1,j} - u_{i,j}}{h} & \text{if } j < N \\
0 & \text{if } j = N
\end{cases}
\]

– We also define a scalar product in $Y$

\[
\langle p, q \rangle_Y = \sum_{i,j} p_{i,j}^1 q_{i,j}^1 + p_{i,j}^2 q_{i,j}^2, \quad p = (p^1, p^2), \quad q = (q^1, q^2) \in Y
\]

– Note that

\[
\langle \nabla u, p \rangle_Y = - \langle u, \text{div } p \rangle_X
\]
The discrete energy

- The discrete ROF model is given by

\[ h^2 \min_{u \in X} \| \nabla u \|_1 + \frac{\lambda}{2} \| u - g \|_2^2 \]

- The discrete total variation is defined as

\[ \| \nabla u \|_1 = \sum_{i,j} |(\nabla u)_{i,j}| , \quad |(\nabla u)_{i,j}| = \sqrt{((\nabla u)^1_{i,j})^2 + ((\nabla u)^2_{i,j})^2} \]

- The squared data term is defined as

\[ \| u - f \|_2^2 = \sum_{i,j} (u_{i,j} - f_{i,j})^2 \]
Solving the ROF model

- The standard approach is to perform a gradient descent on the ROF energy
- Formally, the gradient is given by

\[
\nabla^T \begin{pmatrix}
\text{vec}(\frac{(\nabla u_{i,j})^1}{||(\nabla u_{i,j})||_2}) \\
\text{vec}(\frac{(\nabla u_{i,j})^2}{||(\nabla u_{i,j})||_2}) \\
\end{pmatrix} + \lambda(u - f)
\]

- Note that problems occur, if \(|(\nabla u_{i,j})||_2 = 0\)
- Can be avoided by replacing \(|(\nabla u_{i,j})||_2\) by \(\max(\varepsilon, |(\nabla u_{i,j})||_2)\)
- The corresponding model performs a quadratic regularization for small values of \(|(\nabla u_{i,j})||_2\)
- The resulting approximated ROF model can be solved via a gradient descent algorithm
Gradient descent on the primal

- Set $u_0 = f$ and choose $0 < \tau < 2/(8/\varepsilon + \lambda)$
- For $n = 0, 1, 2, \ldots$

$$g_n = \nabla^T \left( \begin{array}{c} \text{vec} \left( \frac{(\nabla u_n)_{i,j}^1}{\max(\varepsilon, |(\nabla u_n)_{i,j}|_2)} \right) \\ \text{vec} \left( \frac{(\nabla u_n)_{i,j}^2}{\max(\varepsilon, |(\nabla u_n)_{i,j}|_2)} \right) \end{array} \right) + \lambda(u_n - f)$$

$$u_{n+1} = u_n - \tau g_n$$

- End
Dual approach

– On order to avoid the non-differentiability, one can instead solve a dual problem

– The basic observation is the following

\[ \| \nabla u \|_1 = \max_{\| p \|_\infty \leq 1} \langle \nabla u, p \rangle \iff |(\nabla u)_{i,j}|_2 = \max_{|p_{i,j}|_2 \leq 1} (\nabla u)_{i,j} \cdot p_{i,j} \]

– Hence the ROF model can be written as

\[ \min_u \max_{\| p \|_\infty \leq 1} \langle \nabla u, p \rangle + \frac{\lambda}{2} \| u - g \|_2^2 \]

– Computing the minimum for \( u \) and substituting back one obtains

\[ \max_{\| p \|_\infty \leq 1} -\frac{1}{2\lambda} \| \nabla^T p \|^2 + \langle f, \nabla^T p \rangle \iff \min_{\| p \|_\infty \leq 1} \frac{1}{2} \| \nabla^T p - \lambda f \|^2 \]
Dual approach

- The dual energy is quadratic and hence continuously differentiable.
- The non-differentiability of the total variation has moved to the pointwise constraints on the dual variable $p$.
- The dual approach allows to compute the so-called primal-dual gap which is the difference between the primal and the dual energy.

\[
gap(u, p) = \| \nabla u \|_1 + \frac{\lambda}{2} \| u - g \|_2^2 + \frac{1}{2\lambda} \| \nabla^T p \|_2^2 - \langle f, \nabla^T p \rangle
\]

- The gap is zero, if $(u, p)$ is an optimal solution.
- The dual ROF model can be solved via a projected gradient descend algorithm.
- Also note that

\[
p_{i,j} = \frac{(\nabla u)_{i,j}}{|(\nabla u)_{i,j}|} \quad \text{if } (\nabla u)_{i,j} \neq 0
\]
Projected gradient descent on the dual

- Set $p_0 = 0$ and choose $0 < \tau < 2/8$
- For $n = 0, 1, 2, ...$
  
  $g_n = \nabla (\nabla^T p_n - \lambda f)$
  
  $p_{n+\frac{1}{2}} = p_n - \tau g_n$

  $p_{n+1} = \text{proj}_{\|p\|_\infty \leq 1}(p_{n+\frac{1}{2}}) \iff (p_{i,j})_{n+1} = \frac{(p_{i,j})_{n+\frac{1}{2}}}{\max(1, |(p_{i,j})_{n+\frac{1}{2}}|_2)}$

- End

- Projection:
Shortcomings of the ROF model

• **Staircasing effect**: The total variation tends to favor piecewise constant solutions
  - Huber norm
  - Total generalized variation (TGV) [Bredies, Kunisch, Pock, 2010]

\[
\min_{u,v} \lambda_1 \int_{\Omega} |\nabla u - v| \, dx + \lambda_2 \int_{\Omega} |\nabla_s v| \, dx + \frac{1}{2} \| u - f \|_2^2
\]

  - Tries to additionally fit a function \( v \) into affine regions
  - Still a convex problem
The TGV norm is an extension of the TV norm to higher order derivatives.

- TGV of first order is equivalent to TV.
- TGV of $k$th order can be used to reconstruct piecewise polynomial functions of $k$th order.

**TV versus TGV**

![Clean image](image1)
![Noisy image](image2)
![TV](image3)
![TGV-2](image4)
TV-$\ell_1$ model

- Slight change of the original ROF model
- Replaces the quadratic data term by an $\ell_1$ data term

$$\min_{u} \int_{\Omega} |Du| + \lambda \| u - f \|_1$$

- Contrast invariance
- Energy acts independently on each level set
- Applications
  - Multiscale image decomposition
  - Denoising of shapes
Denosing using the TV-$\ell_1$ model

- Better performance for impulse noise
Minimizing the TV-$\ell_1$ model

- The discrete TV-$\ell_1$ model is given by

$$\min_u \| \nabla u \|_1 + \lambda \| u - f \|_1$$

- Again approximate the TV norm such that it becomes differentiable

- Then, a variant of the projected gradient descend algorithm can be used

- The algorithm takes forward steps in the total variation part and takes backward steps in the $\ell_1$ data term

- The resulting algorithm is called forward-backward algorithm
Forward-backward algorithm

- Set $u_0 = f$ and choose $0 < \tau < 2/(8/\varepsilon)$
- For $n = 0, 1, 2, \ldots$

$$g_n = \nabla^T \left( \begin{pmatrix} \frac{(\nabla u_n)^1_{i,j}}{\max(\varepsilon, |(\nabla u_n)^1_{i,j}|_2)} \\ \frac{(\nabla u_n)^2_{i,j}}{\max(\varepsilon, |(\nabla u_n)^2_{i,j}|_2)} \end{pmatrix} \right)$$

$$u_{n+\frac{1}{2}} = u_n - \tau g_n$$

$$u_{n+1} = f + \text{shrink}(u_{n+\frac{1}{2}} - f, \tau \lambda)$$

- End

- Shrinkage

$$\text{shrink}(v, \theta) = \arg \min_w ||w - v||^2/2 + \theta ||w||_1 = \max(0, |w| - \theta)\text{sign}(w)$$
Multiscale Image decomposition

• „Denoising“ of the image using different values of $\lambda$
Geodesic Active Contour Model

• Reformulation of the GAC model using total variation

\[
\min_{u(x) \in [0,1]} \int_{\Omega} g |Du|
\]

• Problem: optimal solution is the empty set!
• Solution: User defined seed points for object and background
Continuous Ising model

- Continuous formulation of the discrete Ising model using total variation

\[
\min_{u(x) \in [0,1]} \int_{\Omega} g|Du| + \lambda \int_{\Omega} u(x)(I(x) - c_1)^2 + (1 - u(x))(I(x) - c_2)^2 \, dx
\]

- Total variation of a binary function equals the boundary length

- Yields a convex optimization problem for given mean values

- Joint optimization of contour and mean values is still a non-convex optimization problem
Example
Relation between the ROF model and the Ising model

It can be shown that that minimizer $v^*$ of the Ising model

$$v^* = \arg \min_{v(x) \in [0,1]} \int_{\Omega} g|Dv| + \int_{\Omega} v(x)f(x)dx$$

Can be obtained by $v^* = 1_{\{x:u^*(x)>0\}}$, where $u^*$is the minimizer of the g-weighted ROF model

$$\min_{u(x)} \int_{\Omega} g|Du| + \|u - f\|_2^2$$

Hence, the segmentation model can be minimized by solving the ROF model
TV-Deconvolution

- Incorporate blur kernel into the denoising model

\[
\min_u \int_\Omega |Du| + \frac{\lambda}{2} \|Au - f\|_2^2
\]

- Examples for the blur kernel \( Au = h \ast u \)
  - Gaussian blur
  - Motion blur
Original

Motion Blur + 1% Noise

Wiener Filter

Quadratic

Total Variation
TV-Superresolution

- The linear operator $A$ can also be a blur kernel followed by a downsampling

\[
\min_u \int_\Omega |Du| + \frac{\lambda}{2} \|SBu - f\|_2^2
\]

- Downsampling Operator $S$
- Blur-Operator $B$
TV-Superresolution

Original images  Bicubic interpolation  TV-Superresolution
TV – Interpolation (Inpainting)

\[
\min_u \int_{\Omega} |Du| + \int_{\Omega} \lambda(x)(u(x) - f(x))^2 \, dx
\]

\[
\lambda(x) = \begin{cases} 
0 & \text{... } x \in \Omega_I \\
\infty & \text{... else}
\end{cases}
\]
Blind Deconvolution

• Aim is to recover a sharp image from a single blurred image
• Basic assumption is the model

\[ f = Au + n \]

• Both, noise and blur kernel are unknown
• Total variation approach with quadratic kernel regularization

\[ \min_{u, A} \int_\Omega |Du| + \frac{\lambda}{2} \| Au - f \|_2^2 + \frac{\mu}{2} \| A \|_2^2 \]

• Still one of the most challenging inverse problems in image processing