IMPLICIT PRIORS FOR MODEL-BASED INVERSION

Eri Haneda, Pengchong Jin, and Charles. A. Bouman
Electrical and Computer Engineering, Purdue University, West Lafayette, IN USA 47907

1) Overview

Problem:

- MRFs are commonly used prior models, however they are restricted to very simple Gibbs distributions.

$$\log p(x) = u(x) = \sum_{\{s,r\}\in C} \rho(x_s - x_r) + const$$

- **How can we make MRFs more expressive?**

Our Approach:

1) Model conditional probability of pixels given neighbors, $p(x_s \mid x_{\partial s})$.
2) Use local approximation to the implicit Gibbs distribution of the MRF.
3) Iteratively minimize the MAP cost.

Result:

An MRF prior which adapts to local image structure.

2) MRFs and Inverse Problems

- MRFs can be expressed as Gibbs distributions

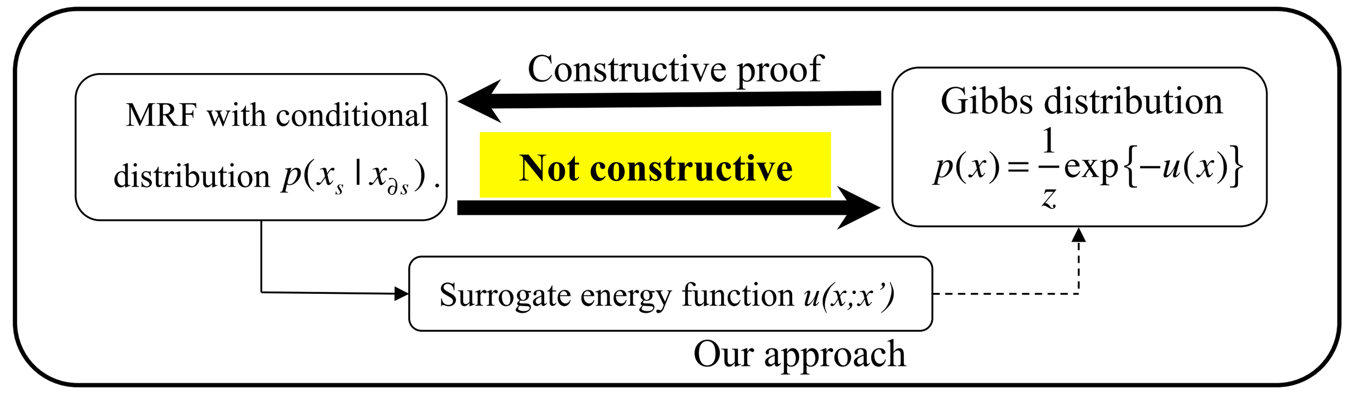
$$p(x) = \frac{1}{z}\exp\{-u(x)\}$$

- Then inverse problems can be solved as MAP estimate with MRF prior

$$\hat{x} = \arg\min_x \left\{ \|y - Ax\|_\Lambda^2 + u(x) \right\}$$

3) The Problems with MRFs

- An MRF is defined by the property that $p(x_s \mid x_{\partial s}) = p(x_s \mid x_{r\neq s})$.
- However, the Hammersley-Clifford theorem provides no way to compute compute Gibbs energy, $u(x)$, from $p(x_s \mid x_{\partial s})$.
- Therefore, current MRF models are restricted to very simplistic Gibbs distributions that are not sufficiently expressive for real images.
- Question: How can we create more complex and expressive MRFs?



Hammersley-Clifford Theorem

4) Our Solution: Implicit Gibbs distributions

Our Approach:

1) Model conditional probability of pixels given neighbors.
   - Estimate the conditional distribution using off-line training procedure
   - We use a Gaussian mixture model, but many choices are possible
2) Locally estimate the energy of the Gibbs distribution.
   - Compute a local approximation to the energy function about the point, x'

$$u(x) \cong u(x;x')$$

   - Ensure that $u(x;x')$ is a surrogate function for $u(x)$.
3) Iteratively minimize the MAP cost function with the surrogate approximation.

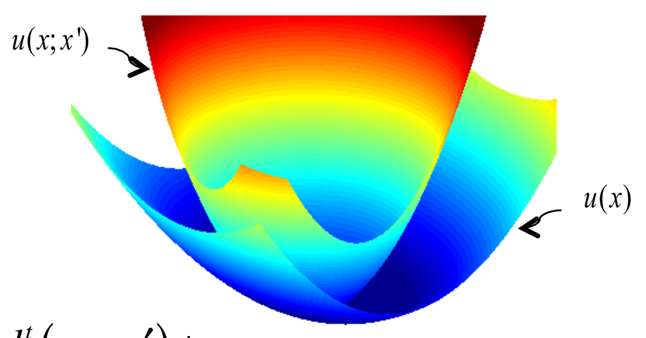
Observation:

- We never explicitly computed the energy u(x) .
- The true energy and prior remains implicit !

5) Computing the Surrogate Energy Function

- Surrogate energy function must satisfy the upper-bound conditions.

$$u(x') = u(x';x')$$
$$u(x) \le u(x;x')$$



- We formulate the surrogate energy function as a quadratic form such as:

$$u(x;x') = \frac{1}{2}(x-x')B(x-x') + d^t(x-x') + c$$

where $d_s = -\frac{\partial}{\partial x_s}\log p(x_s \mid x_{\partial s})\Big|_{x=x'}$.

- For B, our approach is to first find B which satisfies the three strong necessary conditions, then adjust the matrix by $B \leftarrow B + \alpha\ diag\{B\}$ to ensure an upper bound.

- Condition 1

  The symmetric matrix B must be positive definite (i.e. $B > 0$)
- Condition 2

  Surrogate energy function must have greater 2nd derivative than true energy function. (i.e. $B - H \ge 0$)

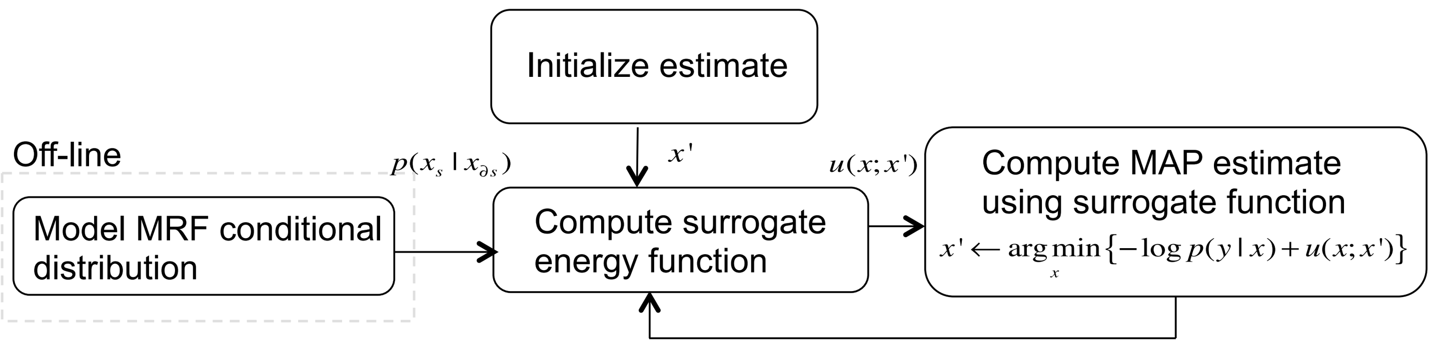
$$H_{s,r} = -\frac{\partial^2}{\partial x_s \partial x_r}\log p(x_s \mid x_{\partial s})\Big|_{x=x'}$$

- Condition 3

  Surrogate energy function must upper bound true energy function along each axis.

6) Iterative MAP Optimization with Implicit Prior

Iterative MAP optimization flowchart:



- This iterative optimization guarantees minimization of MAP cost with implicit energy function.

7) Conditional probability model

Our choice for conditional probability model:

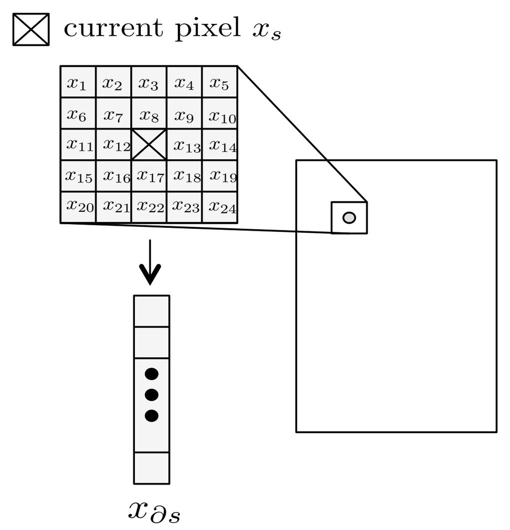
- Gaussian mixture form
- Each pixel is assumed to be fallen into "classes" based on edge orientations
- For a given class k, a pixel is formulated as a weighted sum of its neighbors

$$x_s \mid x_{\partial s}, k \sim \mathcal{N}(A_k x_{\partial s} + \beta_k, \sigma_k)$$

$$p(x_s \mid x_{\partial s}) = \sum_{k=1}^{M} p(x_s \mid x_{\partial s}, k)\, p(k \mid z)$$

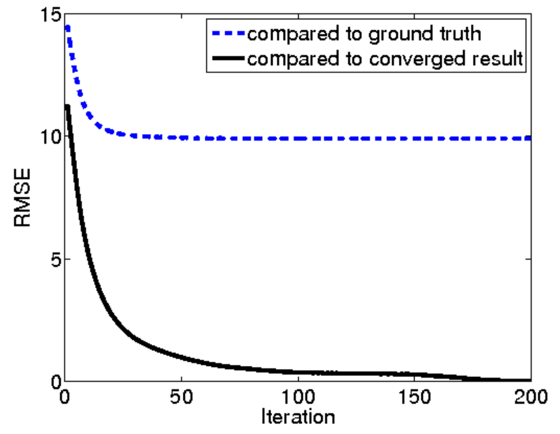
where A_k is a coefficient row vector, β_k is a scalar, and z is a edge feature.

- The model parameters are trained off-line using a linear least squares regression.

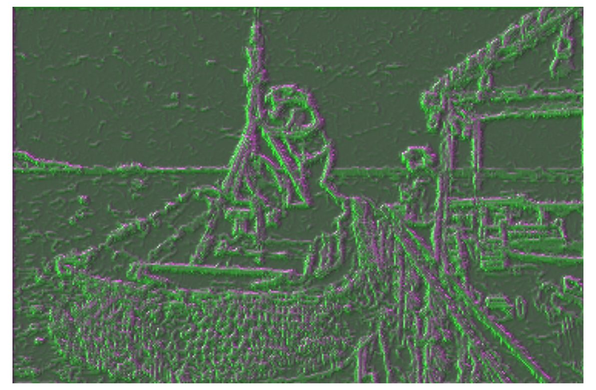


8) Experimental results

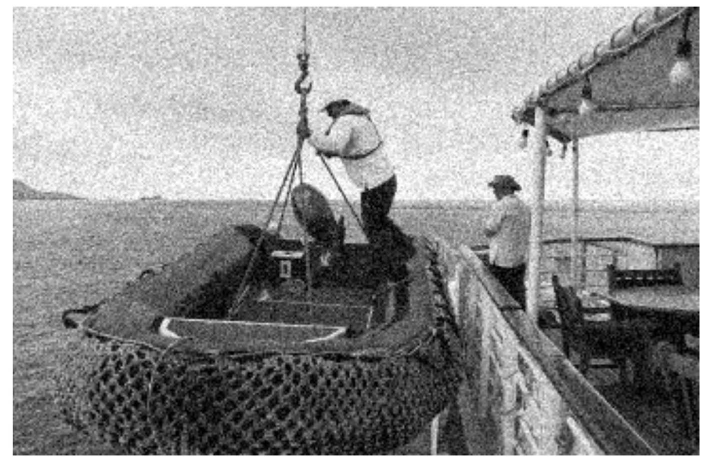
- We performed a simple denoising experiment of removing additive white Gaussian noise (σw=20).
- The 25 grayscale training images were used.
- We ran comparisons with different parameters of GGMRF and qGGMRF which are the current state-of-the-art priors for inverse problems.



Convergence of RMSE



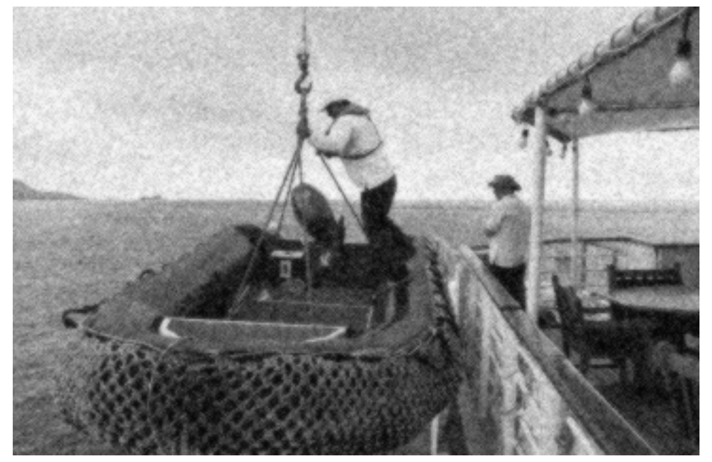
The matrix Bs,r entry



Original noisy image

Implicit prior

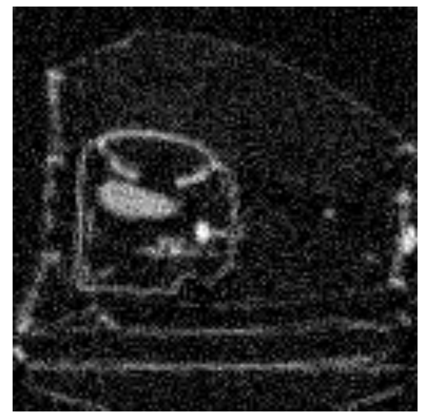
GGMRF (p=1.2)



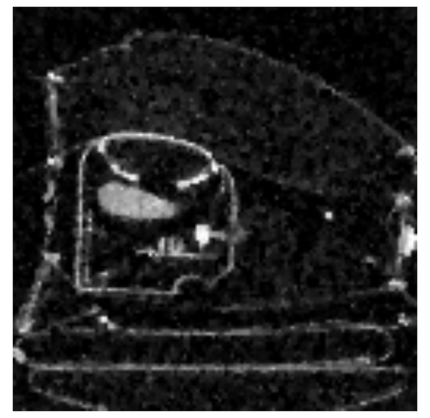
qGGMRF (p=2, q=1, c=1.5)

CT reconstruction simulation

Ground Truth

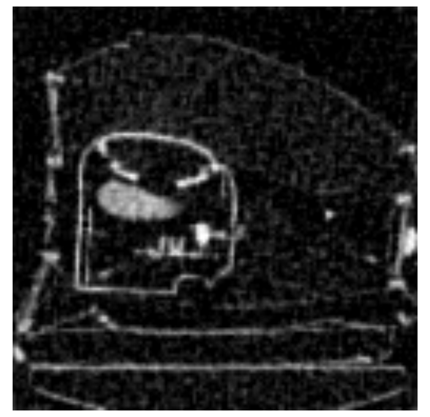


FBP

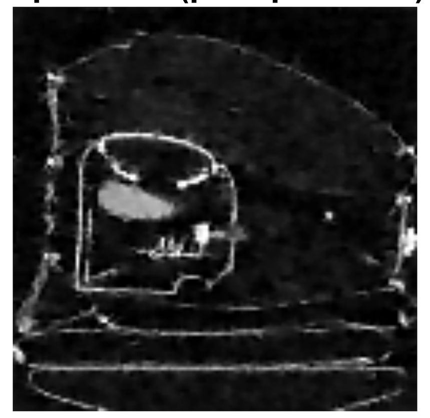


qGGMRF (p=2 q=1.2 c=5)

- 2D parallel beam CT
- 128x128 resolution, 1mm width
- 180 views, 1 degree per view
- 186 detectors, 1mm each
- White noise added to sinogram



kSVD prior



Implicit prior

Conclusion

- We introduced a new MRF modeling which is only implicitly specified through the conditional probabilities.
- We provided a simple example of image denoising, but the method is generally applicable to any continuously valued MRF prior model.

References

[1] C. B. Atkins, C. A. Bouman and J. P. Allebach, "Optimal image scaling using pixel classification," IEEE Int'l Conf. on Image Proc. (ICIP) vol. 3, pp. 864–867, 2001.

This research was supported by ALERT DHS center Northeastern University and the U. S. Army Research Office and the Army Research Laboratory under contract number W911NF-09-1-0540.