A Generative Perspective on MRFs in Low-Level Vision

Uwe Schmidt Qi Gao Stefan Roth

Department of Computer Science, TU Darmstadt, Germany

TECHNISCHE UNIVERSITÄT DARMSTADT

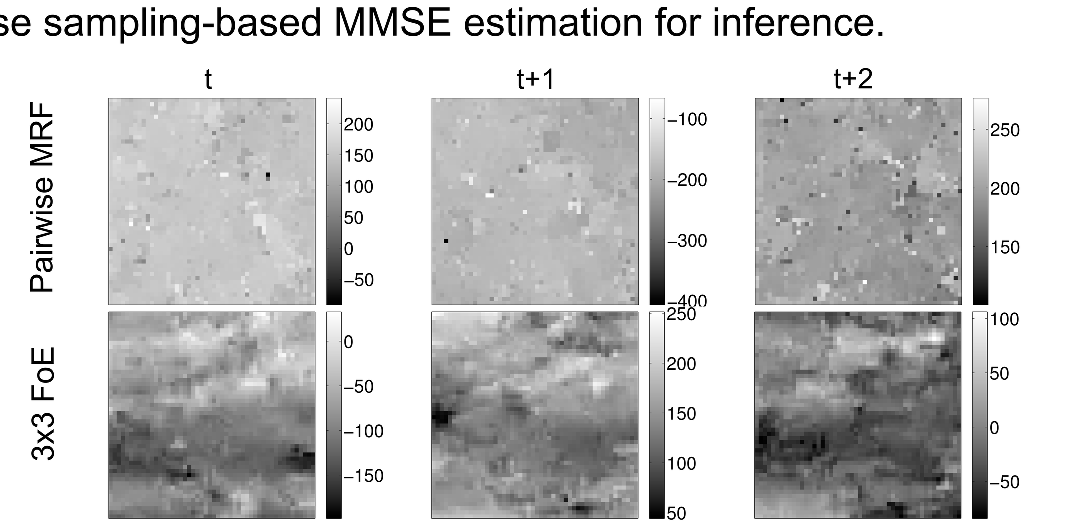
Introduction

Goal:
- Understand generative properties of MRF models of natural images.

Questions:
- How can we evaluate generative MRF models?
- How well do typical image priors capture the statistical properties of natural images?

Previous methods:
- Use application-specific evaluation in the context of MAP estimation.
- Analyzing the statistical properties via single-site Gibbs sampling not very practical [9].

Our approach:
- Exploit rapidly mixing Gibbs sampler for MRF priors [2, 3].
- Analyze the generative properties with efficient sampler.
- Learn better generative MRFs based on flexible GSM potentials.
- Use sampling-based MMSE estimation for inference.



Consecutive samples from MRF image priors using the efficient Gibbs sampler

Flexible MRF framework

Fields-of-Experts image prior [5]:

$$p(\mathbf{x};\Theta) = \frac{1}{Z(\Theta)} e^{-\epsilon\|\mathbf{x}\|^2/2} \prod_{c\in\mathcal{C}} \prod_{i=1}^{N} \phi(\mathbf{J}_i^{\mathrm{T}}\mathbf{x}_{(c)};\boldsymbol{\alpha}_i)$$

- Use flexible Gaussian scale mixtures (GSMs) as experts [8]:

$$\phi(\mathbf{J}_i^{\mathrm{T}}\mathbf{x}_{(c)};\boldsymbol{\alpha}_i) = \sum_{j=1}^{J} \alpha_{ij}\cdot\mathcal{N}(\mathbf{J}_i^{\mathrm{T}}\mathbf{x}_{(c)};0,\sigma_i^2/s_j)$$

- Subsumes many popular pairwise and high-order models.

Efficient auxiliary-variable Gibbs sampler [2]:

- Retain the scales of the GSM as a hidden random vector $\mathbf{z}$.
- Define a joint distribution $p(\mathbf{x},\,\mathbf{z};\Theta)$ such that

$$\sum_{\mathbf{z}} p(\mathbf{x},\mathbf{z};\Theta) = p(\mathbf{x};\Theta)$$

- This allows to define an auxiliary-variable Gibbs sampler that alternates between sampling

$$\mathbf{z}^{(t+1)} \sim p(\mathbf{z}|\mathbf{x}^{(t)};\Theta) \quad \text{and} \quad \mathbf{x}^{(t+1)} \sim p(\mathbf{x}|\mathbf{z}^{(t+1)};\Theta)$$

- From the defined model, we have the conditionals

$$p(z_{ic}|\mathbf{x};\Theta) \propto \alpha_{iz_{ic}}\cdot\mathcal{N}(\mathbf{J}_i^{\mathrm{T}}\mathbf{x}_{(c)};0,\sigma_i^2/s_{z_{ic}})$$

$$p(\mathbf{x}|\mathbf{z};\Theta) \propto \mathcal{N}\left(\mathbf{x};\mathbf{0},\left(\epsilon\mathbf{I}+\sum_{i=1}^{N}\mathbf{W}_i\mathbf{Z}_i\mathbf{W}_i^{\mathrm{T}}\right)^{-1}\right) = \mathcal{N}\left(\mathbf{x};\mathbf{0},\left(\mathbf{W}\mathbf{Z}\mathbf{W}^{\mathrm{T}}\right)^{-1}\right)$$

- Obtain a sample $\mathbf{x}$ by solving a least-squares problem [3, 7]

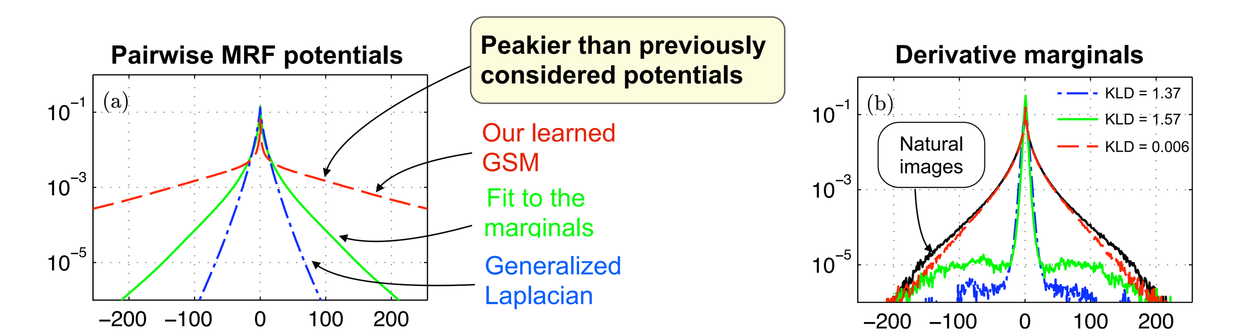
$$\mathbf{W}\mathbf{Z}\mathbf{W}^{\mathrm{T}}\mathbf{x} = \mathbf{W}\sqrt{\mathbf{Z}}\mathbf{y}, \qquad \mathbf{y}\sim\mathcal{N}(\mathbf{0},\mathbf{I})$$

- Convergence monitoring shows that the sampler mixes very rapidly.

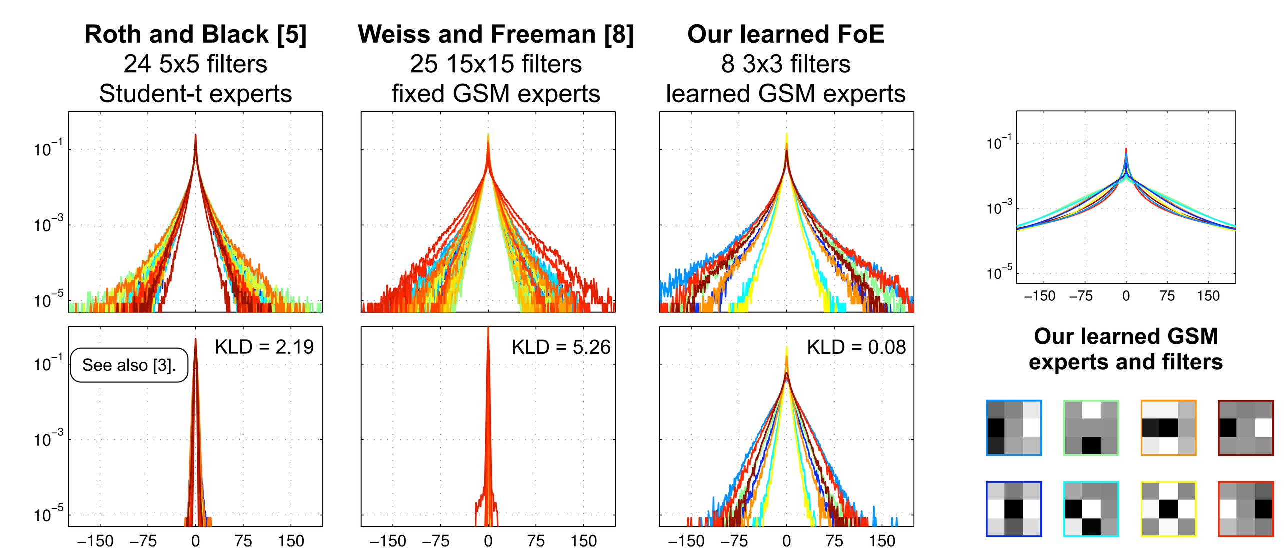
Generative properties of MRF image priors

Evaluation methodology:
- Draw a large number of samples from the model.
- Compute marginal distributions of certain image features (e.g., derivatives, model features) from the samples.
- Compute the marginal KL-divergence between statistics of natural images and samples from the model.

Compare the generative properties of different MRFs:
- Enable sampling other models by fitting GSMs to their potentials.
- For pairwise MRFs, directly compare the derivative statistics:

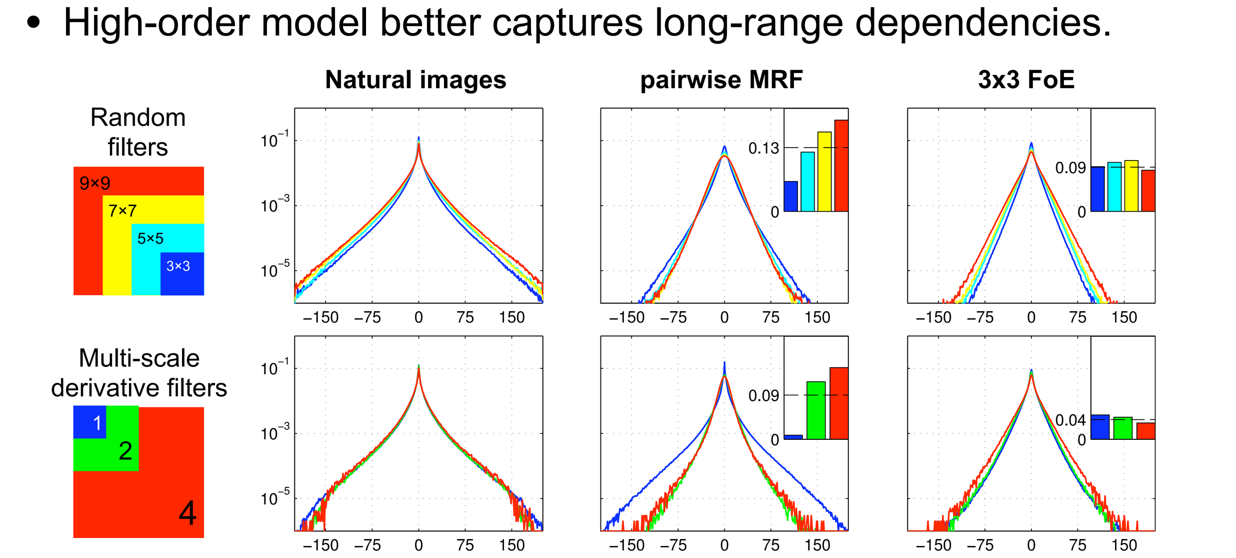


- For high-order MRFs, evaluate each model w.r.t. its learned filters:



Filter statistics of natural images (top) and of FoE models (bottom)

Some additional generative properties of our models:
- Random filter statistics and multi-scale derivative statistics.
- High-order model better captures long-range dependencies.



Improved, learned MRFs:
- Pairwise MRF: fixed horizontal and vertical derivative filters, learn shape of single, more flexible GSM potential.
- FoE: 3x3 cliques, 8 GSM experts; learn expert shapes and filters.
- Trained with Contrastive Divergence.
- Avoid boundary artifacts of samples by conditional sampling of the interior pixels, given the boundary pixels.

Image restoration

Common method – MAP:

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y};\Theta) = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})\cdot p(\mathbf{x};\Theta)^{\lambda}$$

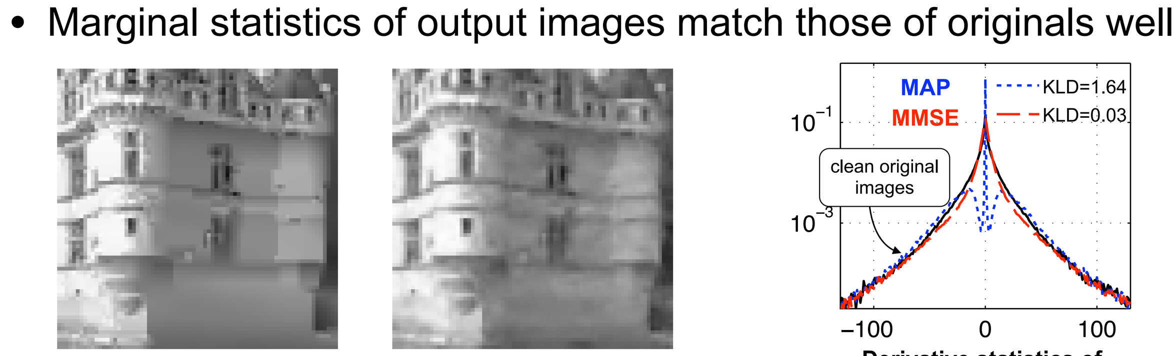
- Regularization weight λ often used to improve performance.
- Point estimate - makes no use of uncertainty.
- Often see staircasing artifacts with heavy-tailed potentials.
- Low correlation between generative quality of the model and its performance in terms of image restoration.

Our method – MMSE:
- *Bayesian minimum mean squared error* estimate:

$$\hat{\mathbf{x}} = \arg\min_{\tilde{\mathbf{x}}} \int ||\tilde{\mathbf{x}}-\mathbf{x}||^2 p(\mathbf{x}|\mathbf{y};\Theta)\,\mathrm{d}\mathbf{x} = E[\mathbf{x}|\mathbf{y}]$$

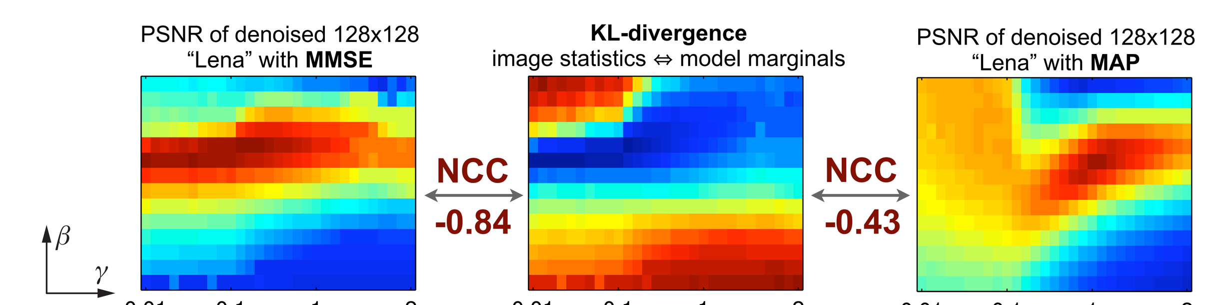
- Compute the average of a large number of posterior samples.
- The auxiliary-variable Gibbs sampler extends to the posterior for Gaussian likelihood models [3].
- Use multiple Markov chains to assess sampler convergence.
- Readily applies to image denoising, inpainting, interpolation, ...

Benefits from MMSE estimation:
- Improved quantitive results: PSNR, SSIM.
- No staircasing artifacts (piecewise-constant regions).
- Marginal statistics of output images match those of originals well.



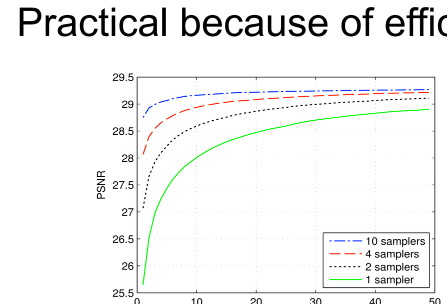
Staircasing artifacts MAP denoising | **MMSE denoising** | **Derivative statistics of 10 denoised test images**

- Better generative models lead to better denoising performance.

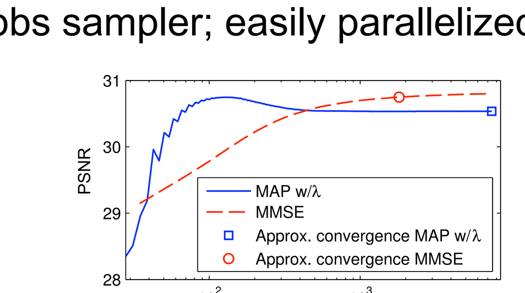


Correlation test using pairwise MRF with generalized Laplacian potential and different parameters.

- Practical because of efficient Gibbs sampler; easily parallelized.



Efficiency of MMSE denoising when assuming parallel computing with learned pairwise MRF.



Efficiency of sampling-based MMSE denoising, comparing to CG-based MAP w/λ.

Image denoising:
- Assume additive white Gaussian noise with known variance.
- Sampling conditional distribution of scales $\mathbf{z}$ is straightforward.
- Conditional distribution of image $\mathbf{x}$ is a Gaussian

$$p(\mathbf{x}|\mathbf{y},\mathbf{z};\Theta) = \mathcal{N}(\mathbf{x};\tilde{\Sigma}\mathbf{y}/\sigma^2,\tilde{\Sigma}) \quad \text{where} \quad \tilde{\Sigma} = \left(\mathbf{I}/\sigma^2+\Sigma^{-1}\right)^{-1}$$

- For assessing convergence, run 4 parallel Markov chains.

Denoising results

Original image | **Pairwise MRF (Laplacian) MAP w/λ, 25.93dB** | **Pairwise MRF (g. Laplacian) MAP w/λ, 25.36dB** | **Learned pairwise MRF MMSE, 26.12dB**

Noisy image (σ=25) 20.34dB | **5x5 FoE from [5] MAP w/λ, 25.36dB** | **5x5 discrimin. FoE from [6] MAP, 26.19dB** | **Learned 3x3 FoE MMSE, 26.27dB**

Table 1. Denoising results (average PSNR in dB) for 10 test images

Model	MAP		MAP w/λ		MMSE	
	σ=10	σ=20	σ=10	σ=20	σ=10	σ=20
pairwise (marginal fit)	28.35	23.96	30.98	26.92	29.70	24.72
pairwise (generalized Laplacian)	27.35	22.97	31.54	27.59	28.64	23.92
pairwise (Laplacian)	29.36	24.27	**31.91**	**28.11**	30.34	25.47
pairwise (ours)	**30.27**	**26.48**	30.41	26.55	**32.09**	**28.32**
5x5 FoE (Roth & Black [5])	27.92	23.81	**32.63**	**28.92**	29.38	24.95
15x15 FoE (Weiss & Freeman [8])	22.51	20.45	32.27	28.47	23.22	21.47
3x3 FoE (ours)	**30.33**	**25.15**	32.19	27.98	**32.85**	**28.91**

Table 2. Denoising results for 68 test images [4, 5] (σ=25)

Model	Learning	Inference	avg. PSNR (dB)
5x5 FoE (Roth & Black [5])	CD (generative)	MAP w/λ	27.44
5x5 FoE (Samuel & Tappen [6])	discriminative	MAP	27.86
pairwise (ours)	CD (generative)	MMSE	27.54
3x3 FoE (ours)	CD (generative)	MMSE	27.95
Non-local means (Buades et al. [1])	--	(MMSE)	27.50
BLS-GSM (Portilla et al. [4])	--	MMSE	**28.02**

Summary

- Evaluated MRFs through their generative properties, based on a flexible framework with an efficient sampler.
- Common image priors are surprisingly poor generative models.
- Learned better generative MRFs (pairwise & high-order) - peakier potentials than considered before.
- Sampling-based MMSE estimation allows generic generative models to compete with recent discriminative models; the MMSE estimate even has a number of additional benefits.

References

[1] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. CVPR 2005.
[2] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization. IEEE TIP, 4(7), 1995.
[3] E. Levi. Using natural image priors – Maximizing or sampling? Master's thesis, The Hebrew University of Jerusalem, 2009.
[4] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE TIP, 12(11), 2003.
[5] S. Roth and M. J. Black. Fields of experts. IJCV, 82(2), 2009.
[6] K. G. G. Samuel and M. F. Tappen. Learning optimized MAP estimates in continuously-valued MRF models. CVPR 2009.
[7] Y. Weiss. Personal communication. 2005.
[8] Y. Weiss and W. T. Freeman. What makes a good model of natural images? CVPR 2007.
[9] S. C. Zhu and D. Mumford. Prior learning and Gibbs reaction-diffusion. PAMI, 19(11), 1997.