Learning High-Order MRF Priors of Color Images

Julian J. McAuley
National ICT Australia, Canberra, ACT 0200, Australia
University of New South Wales, Sydney, NSW 2052, Australia
j.mcauley@student.unsw.edu.au

Tibério S. Caetano
National ICT Australia, Canberra, ACT 0200, Australia
RSISE, Australian National University, Canberra, ACT 0200, Australia
tiberio.caetano@nicta.com.au

Alex J. Smola
National ICT Australia, Canberra, ACT 0200, Australia
RSISE, Australian National University, Canberra, ACT 0200, Australia
alex.smola@nicta.com.au

Matthias O. Franz
Max Planck Institute for Biological Cybernetics, Tuebingen 72012, Germany
matthias.franz@tuebingen.mpg.de

Abstract

In our paper [1], we use large neighborhood Markov random fields to learn rich prior models of color images. Our approach extends the monochromatic *Fields of Experts* model [2] to color images. In the *Fields of Experts* model, the curse of dimensionality due to very large clique sizes is circumvented by parameterizing the potential functions according to a product of experts. We introduce simplifications to the original approach by Roth and Black which allow us to cope with the increased clique size (typically 3x3x3 or 5x5x3 pixels) of color images. Experimental results are presented for image denoising which evidence improvements over state-of-the-art monochromatic image priors.

Overview of Our Paper

In [2], the authors have developed a high-order Markov random field (MRF) model for natural images. In their model, cliques are square image patches of 3x3 and 5x5 pixels (figure 1). The potential functions over these cliques are assumed to be *products of experts* [3], in which each expert takes the form of a Student-T distribution. Specifically, their potential functions are given by

$$\phi_c(\mathbf{x}_c; J, \alpha) = \prod_{f=1}^{F} (1 + \frac{1}{2}\langle J_f, \mathbf{x}_c\rangle^2)^{-\alpha_f}.$$

By invoking the Hammersly-Clifford theorem, the joint probability distribution of their model is just

$$p(\mathbf{x}) = \frac{1}{Z(\Theta)} \prod_{c \in \mathcal{C}} \phi_c(\mathbf{x}_c; J, \alpha).$$

They apply contrastive-divergence to learn the J_f's and α_f's and gradient-based approaches are used to infer the most likely correction of a noisy image. Gradient-ascent makes sense in this setting, since the noisy image is 'close to' the global maximum of the distribution.

Although their model is trained entirely on greyscale images, it is still possible to denoise color images, simply by performing gradient-ascent on each of the three channels (RGB, or YCbCr) independently. However, this does not take advantage of the fact that the channels in a multiband image appear to be highly correlated.

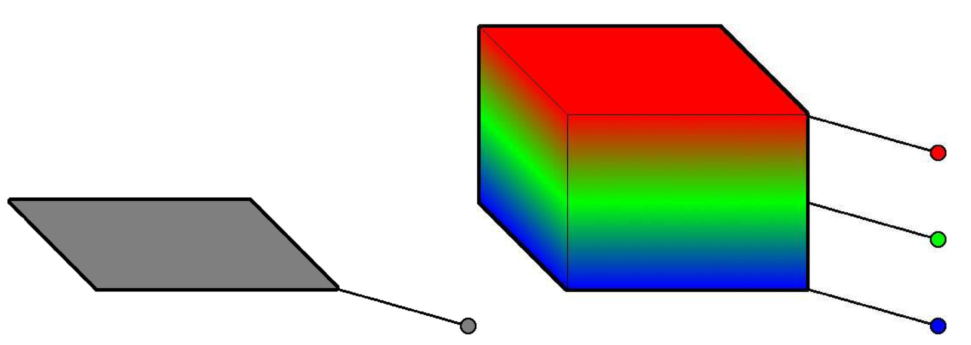

Figure 1. A visualization of the intensity model (left), and our color model (right). Here the grey patch represents a 3x3 or 5x5 maximal clique (fully-connected), whereas the colored cube represents a 3x3x3 or 5x5x3 clique. The attached nodes correspond to the 'noisy pixels' whose correct values we are trying to infer. Since noise is independently distributed, these nodes are only singly-connected.

In our paper [1], we treat the three channels as being dependent objects, and the maximal cliques in our MRF therefore become patches of 3x3x3 and 5x5x3 pixels (figure 1). Since this increases the number of parameters we are required to learn by an order of magnitude, we propose significant simplifications to the learning algorithm in order to render it computationally feasible.

Overview of Our Results

We have found that the improvements achieved by treating the channels of a multiband image as being correlated are statistically significant, despite the suboptimality of our learning procedure. In fact, in many cases, our 3x3x3 color model outperforms the 5x5 greyscale model of [2]. This is an interesting result, since the two models require approximately equivalent computational resources. This result tells us that there appears to be a greater statistical benefit in using color than there is in increasing the neighborhood size.

Figure 2. From left right: The original image, the degraded image with $\sigma = 128$ in the green channel only (PSNR = 13.87), the image restored using our single-channel model (PSNR = 24.15), the image restored using our color model (PSNR = 28.81).

Furthermore, our model shows significant improvements over existing techniques in cases where noise is not equally distributed across all channels. In figure 2 we show an image in which a large amount of noise ($\sigma = 128$) has been applied to the green channel only. A model trained only on green intensities is unable to denoise the image, whereas by exploiting the correlations between the channels, our model achieves very satisfactory results.

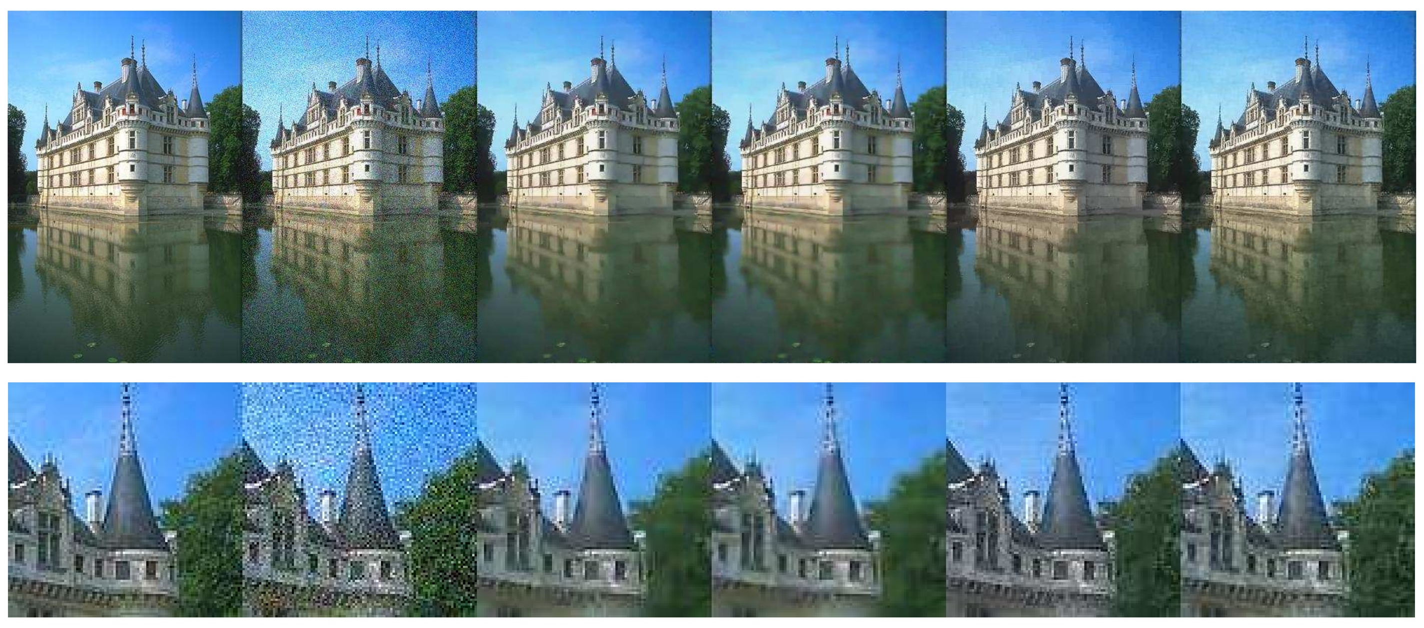

Figure 3. Top – from left to right: Original image, corrupt image with $\sigma = 25$ in all channels (PSNR = 20.46), denoised image using 3x3 model from [2] (PSNR = 29.91), using 5x5 model from [2] (PSNR = 29.82), using our 3x3 model (PSNR = 29.98), using our 5x5 model (PSNR = 30.41). Bottom – close-ups of all images.

Finally, figure 3 shows an image with equal noise applied to each channel ($\sigma = 25$). We compare our model to the model of [2], and see that even our 3x3x3 model produces competitive results, while our 5x5x3 model achieves results which significantly surpass the state-of-the-art.

References

[1] Julian J. McAuley, Tibério S. Caetano, Alex J. Smola, and Matthias O. Franz. Learning high-order mrf priors of color images. In *23rd International Conference on Machine Learning*, June 2006.

[2] Stefan Roth and Michael J. Black. Fields of experts: A framework for learning image priors. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 860–867, June 2005.

[3] G. E. Hinton. Products of experts. In *Ninth International Conference on Artificial Neural Networks, ICANN*, volume 1, pages 1–6, 1999.