IMAGE DENOISING WITH AN ORIENTATION-ADAPTIVE GAUSSIAN SCALE MIXTURE MODEL

David K. Hammond and Eero P. Simoncelli

Center for Neural Science, and Courant Institute of Mathematical Sciences
New York University

We develop a statistical model for images that explicitly captures variations in local orientation and contrast. Patches of wavelet coefficients are described as samples of a fixed Gaussian process that are rotated and scaled according to a set of hidden variables representing the local image contrast and orientation. An optimal Bayesian least squares estimator is developed by conditioning upon and integrating over the hidden orientation and scale variables. The resulting denoising procedure gives results that are visually superior to those obtained with a Gaussian scale mixture model that does not explicitly incorporate local image orientation.

1. INTRODUCTION

Natural photographic images occupy only a small portion of the space of all possible two dimensional signals. Human observers can easily identify images that have been corrupted by noise, indicating that the difference between natural image signals and noise are substantial. Understanding and exploiting this difference underlies essentially all methods for image denoising. In this work we examine the problem of restoring images corrupted by additive Gaussian noise of known covariance. Two of the most salient features of natural images not shared by this noise process are the presence of strongly oriented features and large variations in local contrast.

When images are decomposed in a multiscale wavelet-type representation, variations in local image contrast manifest themselves as clustering of high magnitude coefficients. This observation has been succesfully exploited in image coding (e.g., [1]). A tractable statistical model for clusters of wavelet coefficients that is consistent with this behavior can be constructed using a fixed homogenous distribution that is modulated by a spatially varying hidden variable that controls the local magnitude [2, 3]. The Gaussian scale mixture (GSM) is a model of this type where the homogenous distribution is Gaussian and the hidden variable is a scalar field that controls the variance by simple multiplication.

There has been significant amount of recent research in developing image representations that are well suited for representing image geometry. The curvelet [4] and contourlet [5] representations are fixed bases designed to efficiently approximate images that contain discontinuities along smooth contours, while wedglet [6] and bandlet [7] representations use bases that are directly adapted to the local image geometry.

In this paper we create a model for image patches that is explicitly adapted to both local image amplitude and orientation. Specifically, we extend the GSM concept to include local orientation. Oriented features are an important component of natural images. Their orientation can vary greatly across the image, but the structure of wavelet patches with different orientations are often similar up to a rotation. An example is shown in Fig. 1. In order to capture this behavior, we describe wavelet patches as arising from a GSM process that is then rotated according to an inhomogenous hidden orientation variable. We develop a Bayes least squares estimator based on this model, and apply it to the problem of denoising.

2. IMAGE TRANSFORM

Multiscale linear transforms (loosely known as “wavelets”) have become the representation of choice for a wide variety of image processing tasks. In the current work, we require a representation that allows us to measure local orientation and to rotate coefficient patches by an arbitrary angle. As the sample locations within rotated patches do not in general match the original lattice points, it is necessary to spatially interpolate between the original coefficient sample grid points. Critically sampled seperable orthogonal wavelet representations are unsuitable for this as aliasing effects prevent successful interpolation. Instead, we use the two-band Steerable Pyramid (SP) [8], an overcomplete linear transform that has basis functions comprised of oriented multiscale derivative operators.

The two-band SP of height J decomposes an image into a highpass band, a lowpass band and oriented bandpass bands $B_{j,x}$ and $B_{j,y}$ for $1 \leq j \leq J$. The bandpass filters are derivatives in the x and y directions, and thus pairs of coefficients at a given location in two subbands at a given scale may be considered as vector components of the image gradient at that location and scale. This representation of the gradient provides a direct measure of local orientation. The SP bandpass bands are designed to be free of aliasing which allows interpolation between lattice points.

3. MODEL

We construct a probability model for patches of bandpass pyramid coefficients. Letting v be a vector representating

To appear in: *Proc. IEEE Intl. Conf. Image Processing, Atlanta, Georgia, Oct. 8-11, 2006.*

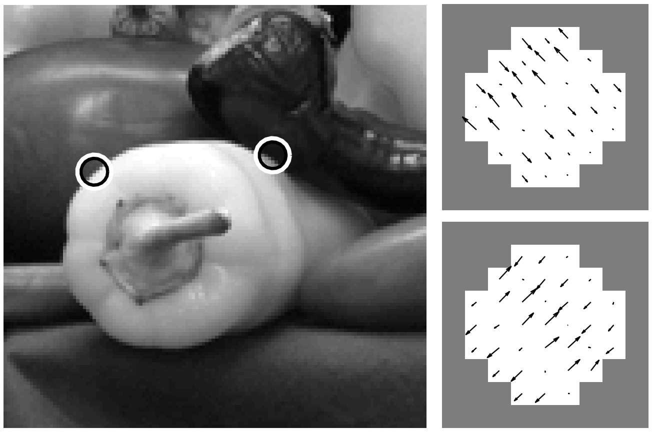

Fig. 1. **Left**: Image with two strongly oriented patches indicated. **Right**: coefficients for each patch at one scale of a two-band steerable pyramid, displayed as vector fields. Patches are similar up to rotation.

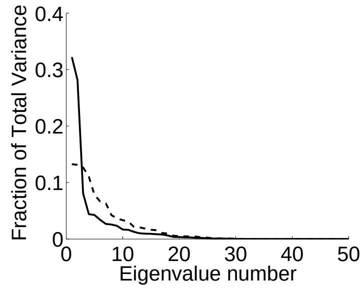

Fig. 2. Normalized eigenvalues of covariance matrix estimated from coefficient patches drawn from single scale of the pyramid representation of an example image ("peppers"). Dashed curve corresponds to raw patches, and solid curve to patches rotated according to dominant orientation.

such a patch, we set

$$v = \sqrt{z} R(\theta) u$$

where z and θ are hidden scalar variables, $R(\theta)$ is a linear spatial rotation operator and u is a zero mean multivariate Gaussian random variable with covariance C_0. We assume the hidden variables z and θ are independent of each other and of u. Thus, the full density of v corresponds to an infinite mixture of Gaussians, with the mixture determined by the density of the hidden variables. We assume a uniform distribution on $[0, 2\pi)$ for $p_\theta(\theta)$, and following [9, 10] place a non-informative Jeffrey's pseudoprior $p_z(z) \propto \frac{1}{z}$ on the multiplier z, truncated to $z \in [z_{min}, z_{max}]$ with same bounds as in [10].

Note that when conditioned on fixed values of the hidden variables z and θ, the patch v is distributed as a zero mean multivariate Gaussian with covariance adapted according to the local amplitude and orientation and written as

$$zC(\theta) = zR(\theta) C_0 R(\theta)^T$$

In this paper we use 5x5 patches augmented with one pair of "parent" coefficients from the immediately coarser subband, to include some cross-scale interaction.

Attempting to undo the action of the multiplier z by dividing each patch by an estimate of the hidden variable at its location yields patches with statistical properties much closer to Gaussian [2]. Thus, the variations of the local contrast can be captured by the hidden multiplier, and when this is removed by division the statistics of the remaining process are more homogenous.

One may similarly attempt to undo the effect of the hidden orientation by rotating the content of image wavelet patches according to an estimate of the local orientation. As a measure of the effectiveness of the model in capturing local image statistics, we can examine the eigenvalues of the covariance matrix. Figure 2 shows the eigenvalues of the covariance matrix formed from the sample outer products of vectorized 5x5 patches of 2-band SP coefficients drawn directly from an example image, compared with that computed from patches that are rotated to align the dominant local orientation with the horizontal axis. The curve for the rotated patches clearly indicates that energy is more concentrated in the first few eigenvalues, suggesting that behavior of the rotated patches is more homogeneous.

4. ESTIMATING MODEL PARAMETERS

The model is specified by knowledge of the covariance matrix $C(\theta)$, or equivalently, of C_0 and $R(\theta)$. In this work we estimate $C(\theta)$ directly from the noisy data, by estimating the orientation of each patch, spatially rotating (and interpolating) the patches contents, and then computing the covariance of these rotated patches. Each of these steps is described in the following subsections.

4.1. Neighborhood Orientation

An $m \times m$ patch of two-band SP coefficients v may be considered as a collection of m^2 gradient vectors h_i for $i = 1...m^2$. We define the neighborhood orientation ϕ for the patch as the angle of the unit vector $k(\phi) = (\cos(\phi), \sin(\phi))^T$ that maximizes the sum of squares of inner products

$$\Sigma_{i=1}^{m^2} (k(\phi)^T h_i)^2.$$

This is equivalent to the orientation of the eigenvector corresponding to the largest eigenvalue of the 2×2 Orientation Response Matrix M = $\Sigma h_i h_i^T$. Setting $h_i = (x_i, y_i)^T$, we have

$$\phi = \frac{1}{2} \angle \left(2\Sigma x_i y_i, \Sigma (x_i^2 - y_i^2) \right),$$

where $\angle$ indicates the angle of the vector whose components are specified by the two arguments.

4.2. Rotation of Patches

We describe a patch v located at position (m,n) of the j'th gradient band by

$$v_p(r,s) = B_{j,p}(m+r, n+s)$$

for $p = \{x,y\}$. Resampling the x and y components along a coordinate system rotated by θ gives

$$S_p(r,s) = B_{j,p}(m + r\sin\theta + s\cos\theta, n + r\cos\theta - s\sin\theta)$$

These vector components must then be transformed together as a vectors, we thus define $R(\theta)v$ by

$$\begin{aligned}(R(\theta)v)_x &= \cos\theta S_x + \sin\theta S_y \\ (R(\theta)v)_y &= -\sin\theta S_x + \cos\theta S_y\end{aligned}$$

The resampling requires values for the bands $B_{j,x}$ and $B_{j,y}$ at locations between the regular sample lattice points. We first upsample each band by a factor of 2^{U_f} in each direction by padding the fourier transform with zeros and taking the inverse fourier transform. We then perform bilinear interpolation from the four nearest upsampled lattice points. Results presented in this paper use $U_f = 4$.

4.3. Estimation of $C(\theta)$

Our model describes an image patch v corrupted with additive Gaussian noise by

$$w = \sqrt{z}R(\phi)u + n$$

where n is a sample from the zero mean Gaussian noise process with known covariance C_n.

The rotator variables ϕ are unknown, and are estimated by computing the neighborhood orientation ϕ^* of the noisy patch. Under the assumption that the noise process is rotationally invariant and that z, u and n are independent, we have

$$\begin{aligned}&E[R(\theta-\phi^*)v(R(\theta-\phi^*)v)^T] \\ &\qquad = E[z]E[R(\theta)u(R(\theta)u)^T] + C_n \\ &\qquad = C(\theta) + C_n\end{aligned}$$

where we assume without loss of generality that $E[z]$=1. Note that ϕ^* is different for every patch.

We estimate $C(\theta)$ by forming the average outer product of rotated patches and subtracting C_n. To ensure positive definiteness we diagonalize the estimate and replace all negative eigenvalues by the smallest positive eigenvalue.

5. DENOISING ALGORITHM

We perform denoising in the SP coefficient domain by decomposing the noisy image, denoising the subbands and then inverting the SP transform. We assume the noise is additive, Gaussian and stationary with known covariance. Each subband contains noise filtered by the SP basis functions; the covariance for the subband noise can be computed from the pyramid decomposition of the power spectra of the noise process in the pixel domain (as in [10]).

Given a patch of noisy coefficient data w, we write $w = v + n$ where v is the original image patch we wish to estimate, and n is a zero mean Gaussian with covariance C_n.

The Bayes least squares estimator is then

$$\hat{v} = \int v p(v|w) dv$$

Analogous to the method of [10], we compute this first by introducing and integrating over the hidden variables θ and z

$$\begin{aligned}\hat{v} &= \int v \iint p(v,z,\theta|w) dz d\theta \;\; dv \\ &= \iint p(z,\theta|w)\left(\int v p(v|w,z,\theta) dv\right) \;\; dz d\theta\end{aligned}$$

As v is Gaussian with covariance $zC(\theta)$ when conditioned on z and θ, the inner integral over v is a standard linear (Wiener) estimate:

$$\hat{v}(w;z,\theta) = \int v p(v|w,z,\theta) dv = zC(\theta)(C_n + zC(\theta))^{-1} w$$

$\hat{v}$ is thus a weighted integral of these Wiener estimates. By Bayes rule we have

$$p(z,\theta|w) = \frac{p(w|z,\theta)p(z,\theta)}{p(w)}$$

When conditioned on z and θ, w is a zero mean Gaussian with covariance $zC(\theta) + C_n$. Transforming the integral over z to the log domain by setting $y = \log(z)$ yields

$$\hat{v} = \frac{1}{N}\iint \frac{\exp(-\frac{1}{2}w^T(e^y C(\theta) + C_n)^{-1} w)}{|e^y C(\theta) + C_n|^{1/2}} \hat{v}(w;z,\theta) d\theta dy$$

with appropriate normalization constant N. This procedure gives an estimate of the entire patch; in practice we estimate each coefficient using a patch centered on it, taking the center coefficient. These integrals are computed as simple double sums by discretizing $\log(z)$ and θ to a finite number of points. Results shown in this paper use 13 points for $\log z$ and 16 for θ. The scalar highpass residual is denoised using the GSM procedure described in [10], while the lowpass band is left unchanged.

6. RESULTS

Figure 3 shows a denoising example, using 5×5 patches including parent and SP height 3. The lower right image is denoised using the model described above. The lower left image is denoised using the GSM model of [10] with two SP bands, which is similar to the current method without adaptation to orientation. The best results in [10] were obtained using an 8-band SP, and are comparable or slightly better than

Fig. 3. **Upper left**: Cropped original image. **Upper right**: image with added white Gaussian noise with $\sigma = 40$ (PSNR = 16.08). **Lower left**: denoised with two-band GSM model (PSNR = 25.60) (see [10]). **Lower right**: denoised with orientation-adapted two-band GSM. (PSNR=26.22)

the orientation-adapted GSM results presented here. Note that the edges in the image denoised with the orientation-adapted method are sharper, and suffer less from blockiness on oblique edges. Table 1 compares PSNR results for several test images.

7. DISCUSSION

We've introduced a statistical model for local patches of wavelet coefficients using an infinite mixture of Gaussians, in which the mixture depends on two hidden variables representing local amplitude and local orientation. We've demonstrated use of this new model in denoising, comparing its performance with previous models that utilize only a single hidden amplitude variable [10]. We find that the addition of the orientation variable leads to improvements in PSNR, as well as visual appearance. More generally, the orientation variable allows efficient representation of locally oriented structures which should render the model useful for other applications such as compression, restoration, inpainting, or synthesis.

The current denoising method is well suited for capturing oriented image content; however in regions of the image that do not have a strong dominant orientation such as constant or nonoriented texture areas the method produces spurious oriented artifacts. Preliminary work indicates that these artifacts may be supressed by adapting the model to the local "orientedness" of the signal. We are also pursuing a number of other potential improvements, including alternative methods

	Barbara	Lena	Peppers	Boats
Two band GSM	29.19	31.51	31.56	30.06
Current method	29.83	31.74	31.69	30.05

Table 1. Denoising results in PSNR for 512×512 test images with gaussian white noise with $\sigma = 20$. Noisy images all have PSNR=22.10

of estimating $C(\theta)$ and a more systematic exploration of the effects of neighborhood size.

Finally, as in previous work [10], we do not currently make use of the full global model implied by our local description (we ignore the overlap of the patches, computing independent estimates of each coefficient based solely on its surrounding patch). We do not include any description of the spatial structure of the hidden orientation variables. Incorporating such effects is difficult, but is likely to lead to substantial additional improvements in performance.

8. REFERENCES

[1] J M Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3445–62, December 1993.

[2] M J Wainwright and E P Simoncelli, "Scale mixtures of Gaussians and the statistics of natural images," in *Adv. Neural Information Processing Systems (NIPS*99)*, S. A. Solla, T. K. Leen, and K.-R. Müller, Eds., Cambridge, MA, May 2000, vol. 12, pp. 855–861, MIT Press.

[3] M K Mihçak, I Kozintsev, K Ramchandran, and P Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 300–303, December 1999.

[4] E. Candés and D Donoho, "Curvelets : A surprisingly eddective nonadaptive represenentation of objects with edges," *in Curves and Surfaces fitting, L L Schumaker, A Cohen and C Rabut, eds, Vandebuilt University Press*, 1999.

[5] M N Do and M Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.

[6] J K Romberg, M B Wakin, and R G Baraniuk, "Multiscale geometric image processing," in *Proceedings of the SPIE*, 2003, vol. 5150, pp. 1265–1272.

[7] Erwan Le Pennec and Stéphane Mallat, "Sparse geometric image representations with bandlets," *IEEE Trans. on Image Processing*, vol. 14, no. 4, pp. 423–438, 2005.

[8] E P Simoncelli and W T Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *2nd Int'l Conf on Image Proc*, Washington, DC, October 1995, vol. III, pp. 444–447.

[9] M Figueiredo and R Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffrey's noninformative prior," *IEEE Trans. Image Processing*, vol. 10, no. 9, pp. 1322–1331, September 2001.

[10] J Portilla, V Strela, M Wainwright, and E P Simoncelli, "Image denoising using a scale mixture of Gaussians in the wavelet domain," *IEEE Trans. Image Processing*, vol. 12, no. 11, pp. 1338–1351, November 2003.