CSIC Technical Report #54. Deposito Legal: M-10904. April 11, 1999.
Available at: ftp://ftp.optica.csic.es/pub/papers/portillaTR54.ps.gz
or: ftp://ftp.cns.nyu.edu/pub/eero/portillaTR54.ps.gz

Texture Representation and Synthesis Using Correlation of Complex Wavelet Coefficient Magnitudes

Javier Portilla
Instituto de Optica
CSIC
Serrano 121
28006 Madrid, SPAIN
iodpm79@pinar2.csic.es

Eero P. Simoncelli
Center for Neural Science, and
Courant Inst. of Mathematical Sciences
New York University
New York, NY 10003
eero.simoncelli@nyu.edu

Abstract

We present a statistical characterization of texture images in the context of an overcomplete complex wavelet transform. The characterization is based on empirical observations of statistical regularities in such images, and parameterized by (1) the local auto-correlation of the coefficients in each subband; (2) both the local auto-correlation and cross-correlation of coefficient *magnitudes* at other orientations and spatial scales; and (3) the first few moments of the image pixel histogram. We develop an efficient algorithm for synthesizing random images subject to these constraints using alternated projections, and demonstrate its effectiveness on a wide range of synthetic and natural textures. We also show the flexibility of the representation, by applying to a variety of tasks which can be viewed as constrained image synthesis problems.

Vision is arguably our most important sensory system, judging from both the ubiquity of visual forms of communication, and the large proportion of the human brain devoted to visual processing. Nevertheless, it has proven difficult to establish a good mathematical definition (in the form of a statistical model) for visual images. The set of visual images is enormous, and yet only a small fraction of these are likely to be encountered in a natural setting [43, 25, 19, 54]. Thus, a statistical prior model, even one that partially captures these variations in likelihood, can substantially benefit image processing and artificial vision systems. In addition, many authors have proposed that biological visual systems have evolved to take advantage of the statistical properties of the signals to which they are exposed [*e.g.*, 3, 4, 2], thus suggesting a direct link between image statistics and visual processing.

In order to characterize visual images statistically, one must make some sort of restriction on the probability model. The most common assumptions are locality (the characterization is specified on local spatial neighborhoods), stationarity (the statistics depend only on *relative* spatial position within the image), and parameterized forms of density (*e.g.*, Gaussian). The subclass of images that we commonly call "visual texture" seems most consistent with local stationary density models. Traditional definitions of texture can be classified into "structural" and "statistical" [34], the first consisting of a set of repeated deterministic features, and the second corresponding to a sample drawn from a

• JP is supported by a fellowship from the Consejo Superior de Investigaciones Cientificas (CSIC), and the Comision Interministerial de Ciencia y Tecnologia (CICYT, Spain), under grant TIC97-325. EPS is supported by an Alfred P. Sloan Research Fellowship, NSF CAREER grant MIP-9796040, and the Sloan Center for Theoretical Neurobiology at NYU.

probability density. Such a distinction has turned out to be somewhat artificial: For example, Zhu *et al.* [70] have demonstrated that is possible to capture and reproduce structural elements in texture using purely statistical models. Furthermore, many real-world textures seem to incorporate both aspects, in that they can be described as a set of repeating structural elements subject to some randomness in their location, size, color, orientation, etc. This observation leads us to seek a single method of representing textures. In this paper, we develop a fully statistical description, and demonstrate that it is also able to capture and reproduce a wide variety of structural elements.

Julesz pioneered the statistical characterization of textures by hypothesizing that the Nth-order joint empirical densities (for some unspecified N) of neighborhoods of image pixels, could be used to partition textures into classes that are indistinguishable to a human observer [40]. This work thus established the use of both the locality and stationarity assumptions, the goal of determining a minimal set of statistical constraints, and the validation of texture models using human observers. Since then, researchers have explored a wide variety of approaches for texture characterization and synthesis.

One of the most basic distinctions between the various approaches is the choice of representation. Starting with Julesz, many authors have worked directly on the statistical attributes of local spatial neighborhoods of pixels, typically in the form of a Markov random field [*e.g.*, 35, 42, 17]. But most others attempt to simplify the description of the density by first processing with a set of linear filters, such as Gabor filters or wavelet bases. The use of localized multi-scale multi-orientation sets of bandpass filters is inspired by what is known of biological visual processing, and justified by recent studies of the higher-order statistical properties of such representations. We discuss this further in section 1. These filters may be held fixed, or chosen adaptively depending on the image statistics [*e.g.*, 10, 26, 22, 60, 51].

Assuming a preconditioning subband decomposition, most texture characterizations are defined by a set of statistical measurements that are used to constrain the model. The majority are based only on marginal statistics. In particular, much of the work on texture segmentation is based on estimates of variances of a set of subbands [*e.g.*, 61, 6, 7, 64]. The use of variance constraints implicitly assumes an underlying Gaussian marginal density, but there is abundant evidence that bandpass filtered natural images have highly non-Gaussian marginals [*e.g.*, 25, 18, 45]. Specifically, the densities tend to have sharp peaks at zero, and longer tails than a Gaussian. In addition, a number of researchers have shown that Gaussian statistics are not sufficient to constrain the appearance of visual textures [52, 24]. For purposes of synthesis, some authors have utilized filtered non-Gaussian white noise [*e.g.*, 24, 14] which have provided good results when based on a multi-scale representation [*e.g.* 15]. Two recent highly successful algorithms (described below) are based on marginal densities of linear filter outputs [36, 70].

Joint densities are more difficult to characterize, even when limited to pairwise ($N = 2$) interactions. The models discussed above capture joint statistics implicitly through the preconditioning linear transformation. Some authors have used covariance (or correlation) matrices to characterize the joint densities of local neighborhoods of subband coefficients for purposes of classification or segmentation [*e.g.*, 16, 8]. In the pixel domain, many authors have utilized the pairwise joint histograms (also known as "co-occurrence matrices") [*e.g.*, 35, 17, 29, 31, 23]. Popat and Picard [50] have developed a clustering approach for representing densities of local pixel neighborhoods (including those at different scales). They have applied this to compression, classification and restoration, and have demonstrated impressive synthesis results [49].

A final distinguishing aspect of previous texture synthesis work is the algorithm used to sample from the model. If the model is Gaussian, one can simply draw samples of white Gaussian noise, and linearly transform these to achieve the desired covariance relationships. This technique may also be used for models based on non-Gaussian marginals along a linearly independent set of axes. But if the model is based on marginals of an overcomplete basis or on a density that does not factorize along a set of linear axes, a direct sampling technique is not possible. Many authors use so-called "synthesis-by-analysis" techniques [14, 38], in which the synthesized is adjusted to match a desired set of statistical measurements taken from an example texture. Monte Carlo sampling techniques, such as the Metropolis algorithm [17] or Gibbs sampling [21], fall into this category.

Our approach shares common features with those of four recent successful models. Portilla *et al.* [51] constrain the auto-correlation of Gabor subbands, in which the filter bandwidths are adaptively chosen for each texture. Their technique captures the dominant scales and orientations, and is effective for representing both highly random and quasi-periodic textures, but fails to capture local structures. Heeger and Bergen [36] constrain the marginal statistics of coefficients in a fixed overcomplete basis. They developed a heuristic synthesis-by-analysis algorithm, in which they iteratively alternate between matching the subband histograms, and matching the pixel histogram. This method captures and reproduces the richness of random features of some natural textures, but fails to reproduce texture having straight lines, quasi-periodicity or well-defined structural elements. Zhu *et al.* [70] used Gibbs sampling to draw from the maximal-entropy distribution with coefficient marginals matched to those estimated from an example image. In [69], they propose a minimal entropy criterion for adaptively choosing the filters used to compute the coefficients. This technique may be used with both linear and non-linear operators and it provides good results on a variety of textures (including highly structured and quasi-periodic ones), but it is extremely expensive computationally. Debonet and Viola [20] developed an efficient coarse-to-fine algorithm, in which they sample from the empirical conditional density over scale (i.e., they use a type of statistical bootstrap). This method is very successful at capturing repetitive structural elements, but may not be as well suited to many natural textures, which have a strong random component. We discuss these previous results in section 5.

Our goal in this paper is to develop a texture model based on a set of statistical descriptors that are both necessary and sufficient for representing a wide variety of visual textures. As a test of the model, we estimate the descriptors from individual texture images, and use these descriptors to synthesize new samples whose visual appearance may then be compared with the original texture. We work within a fixed overcomplete multi-scale wavelet representation, and choose a set of statistical measurements that includes the central samples of the auto-correlation of the subbands. In addition, we have recently observed that the pairwise joint densities of coefficients at different spatial locations, orientations and scales exhibit striking dependencies that are not captured by simple variance or correlation measures [13, 57]. Thus, unlike previous techniques, we also include the cross-correlations of coefficient *magnitudes*. We develop an efficient synthesis-by-analysis algorithm for synthesizing images subject to these constraints, which utilizes iterative projections onto solution sets. Finally, we show a large set of examples of texture synthesis and constrained texture synthesis, demonstrating the power and flexibility of the model. Portions of this work have been more briefly described in [57, 55].

1 Multi-Scale Oriented Image Decomposition

Our texture representation is based on a linear decomposition whose basis functions are spatially localized, oriented, and roughly one octave in bandwidth. Such decompositions are inspired by what is known of biological visual processing [*e.g.*, 32], both from human psychophysical experiments, and electro-physiological measurements from neurons in primary visual cortex (i.e., area V1) of mammals. The importance of choosing a representation inspired by human vision has been demonstrated in texture synthesis comparisons [51, 46]. These decompositions are also widely used in computer vision, have been quite successful for texture representation and synthesis, and are justified by recent studies of their higher-order statistical properties [*e.g.*, 25, 65, 18, 68, 57]. Especially relevant are recent results on choosing bases to optimize statistical criterion, which suggest that a basis of localized oriented operators at multiple scales is optimal for image representation [47, 5]. A conceptually similar proposal by Zhu *et al.* is that the set of filters should be selected according to a minimum-entropy criterion [71].

In addition, we wanted our basis functions to be translations, rotations and dilations of a common function. We chose to use a "steerable pyramid" [27, 59, 33, 58], since this transform has nice reconstruction properties (specifically, it is a tight frame), in addition to properties of translation-invariance and rotation-invariance. Similar representations have been used by Watson for coding [65] and Unser for texture segmentation [62]. For this paper, we have extended the original steerable pyramid representation to include complex analytic (i.e., "quadrature pair") filters. That is, the real and imaginary parts of each filter form a Hilbert transform pair. This allows us to utilize measures of local phase and energy in some of our texture descriptors. Such measures have proved important throughout computer vision [*e.g.*, 44, 48], and are again motivated by what is known of biological vision [*e.g.*, 32].

Figure 1 shows the Fourier magnitudes of the subband filters used in this transformation. The oriented filters are polar-separable in the Fourier domain, and may be written as:

$$F_{n,k}(r,\theta) = B_n(r) G_k(\theta), \quad n \in [0,N], k \in [0, K-1], \tag{1}$$

where

$$B_n(r) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2\left(\frac{2^n r}{\pi}\right)\right), & \frac{2^n r}{\pi} \in [\frac{1}{2}, 2] \\ 0, & \text{otherwise} \end{cases}$$

$$G_k(\theta) = \begin{cases} \left[\cos(\theta - \frac{\pi k}{K})\right]^{K-1}, & |\theta - \frac{\pi k}{K}| < \frac{\pi}{2} \\ 0, & \text{otherwise,} \end{cases}$$

where r, θ are polar frequency coordinates. Subbands are subsampled by a factor of 2^n along both axes. This does not produce aliasing artifacts, as the support of each filter obeys the Nyquist sampling criterion.

In addition, one must retain (non-oriented) highpass and lowpass residual bands, which are computed using the following filters:

$$H(r) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2(\frac{r}{\pi})\right), & r \in [\frac{\pi}{2}, \pi] \\ 0 & \text{otherwise} \end{cases}$$

$$L(r) = \begin{cases} \cos\left(\frac{\pi}{2}\log_2(\frac{2^{(N+1)} r}{\pi})\right), & r \in [\frac{\pi}{2^{N+1}}, \frac{\pi}{2^N}] \\ 1 & r < \frac{\pi}{2^{N+1}} \\ 0 & r > \frac{\pi}{2^N}. \end{cases}$$

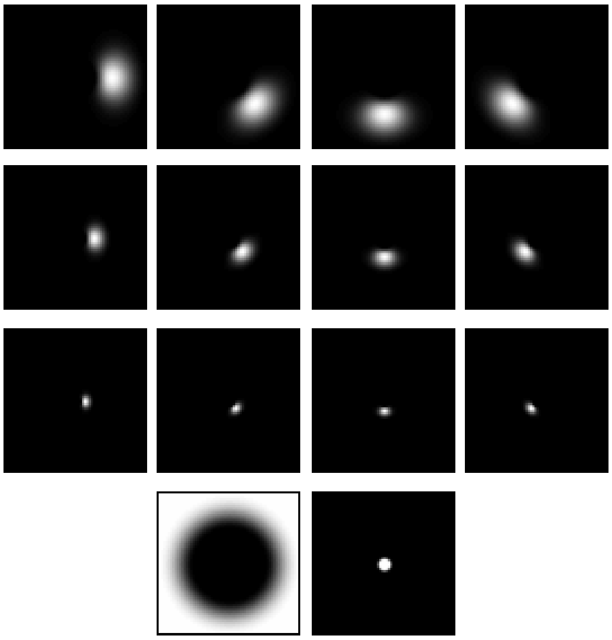

Fig. 1. Fourier spectra of complex-analytic filters of a complex steerable pyramid with four orientations and three scales. Bottom two images correspond to highpass and lowpass residual filters. Vertical and horizontal frequency axes cover the range $[-\pi, \pi]$ radians.

We used $K = 4$ orientation bands, and $N = 4$ pyramid levels for our examples. The transform is implemented directly in the Fourier domain. This set of functions forms a tight frame, and thus the transformation may be inverted by convolving each complex subband with its associated complex-conjugated filter and adding the results. Alternatively, one may reconstruct from either the real or imaginary portions alone. Figure 2 shows an example image decomposition.

2 Statistical Characterization of Pyramid Coefficients

In this section, we describe a set of statistical measurements which are sufficient to parameterize a wide variety of textures. We describe these measurements sequentially, from simplest to most complex. For each type of measurement, we illustrate the information captured by that measurement by synthesizing images that match that statistic. We show four images, chosen as exemplars of different types of visual texture. The first, an artificial "Texton" [41] image, contains randomly located structured elements. The second, a photograph of tree bark, is fairly random, but still contains important features. The third, a photograph of a brick wall, is quasi-periodic. The last, a photograph of herringbone fabric, contains alternating strips of oriented "staircases".

2.1 Pixel Statistics

The statistics of gray-level texture pixels express the relative amount of each intensity in the texture. Many texture models make direct use of the pixel histogram to characterize this distribution (e.g.

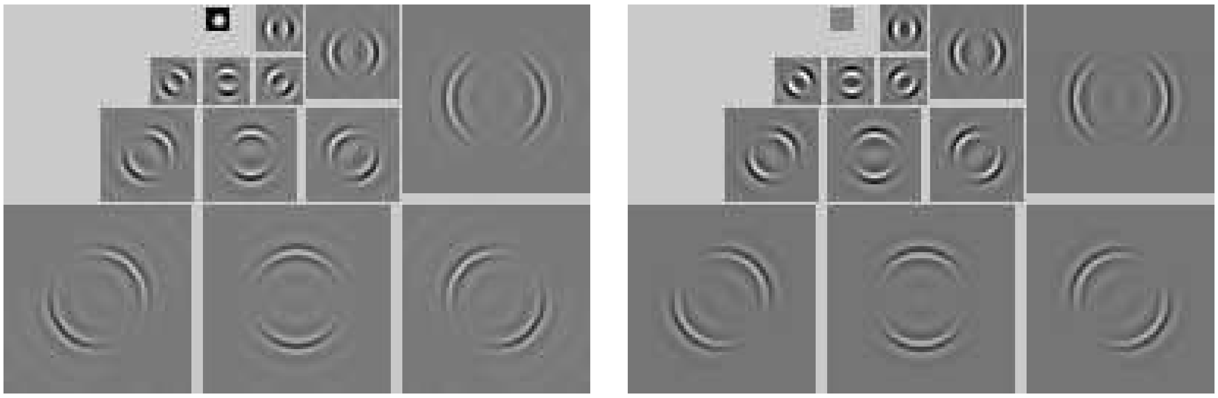

Fig. 2. A 3-scale, 4-orientation complex steerable pyramid representation of a disk image. Shown are the oriented bandpass images at each scale and the final lowpass image. Left: real parts. Right: imaginary parts.

[28, 36, 70]). Figure 3 shows four pairs of images with matching pixel histograms. That is, each synthesized image contains white noise drawn from the appropriate density. Clearly, this measurement characterizes only the most rudimentary aspect of the visual appearance of these textures.

In previous work [51] we used a subsampled version (after low-pass filtering) of the original 256-bin histogram. In our current model we have used, instead, the first four normalized sample moments, i.e., mean, variance, skewness and kurtosis, together with the range (minimum and maximum values) to characterize the first order statistics. This solution yields histograms that are closely matched to the original ones in the synthesis examples.

2.2 Subband Variance Statistics

The subband variances describe the total amount of spectral power in each region of the frequency domain. Natural image spectra are known to have a power-law form [54, 63], but texture images are often less generic. Figure 4 shows four pairs of images with matching subband variances (within a steerable pyramid decomposition). Although they capture more structure than the raw pixel marginals, these are clearly insufficient for representation of textures.

2.3 Subband Covariance Statistics

The coefficients of our wavelet decomposition are typically correlated for two independent reasons. First, the representation is highly overcomplete, and therefore the coefficients are constrained to lie within a linear subspace. More importantly, covariances of subband coefficients can arise from spectral peaks (i.e., periodicity) in a texture. In order to represent such spectral features, we include the local auto-correlation of each subband as a texture descriptor, as was done in [9, 53, 8].

It has been known for some time that correlation measurements are not sufficient to capture the structure of many natural textures [52, 24]. It is simple to generate examples that demonstrate this. Figure 5

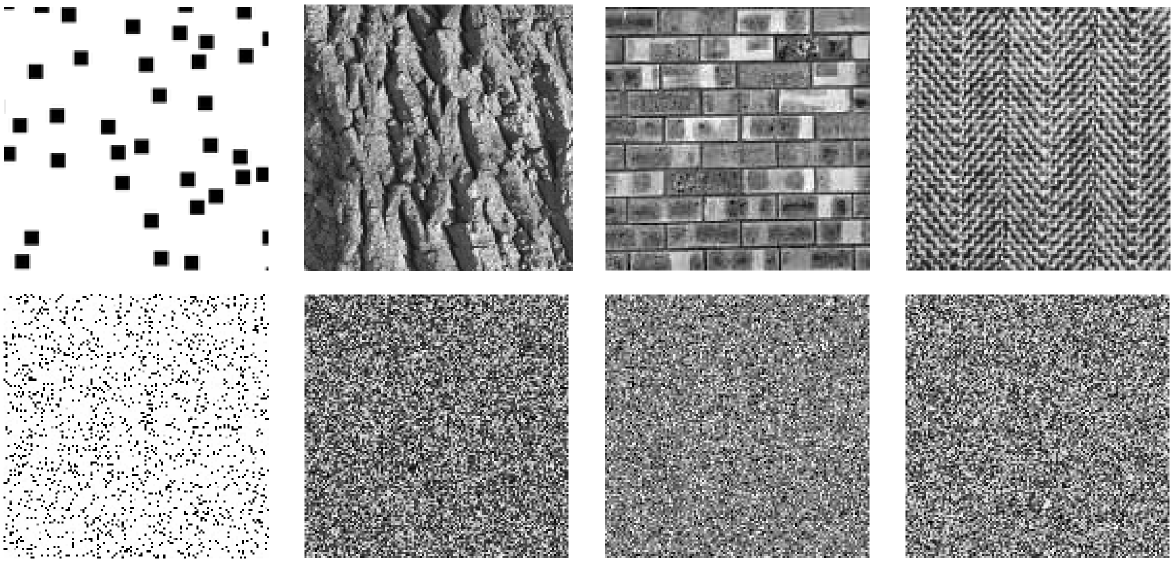

Fig. 3. Inadequacy of pixel marginal statistics. Shown are four pairs of original and synthesized texture images. For each pair, the histograms of the pixels are identical.

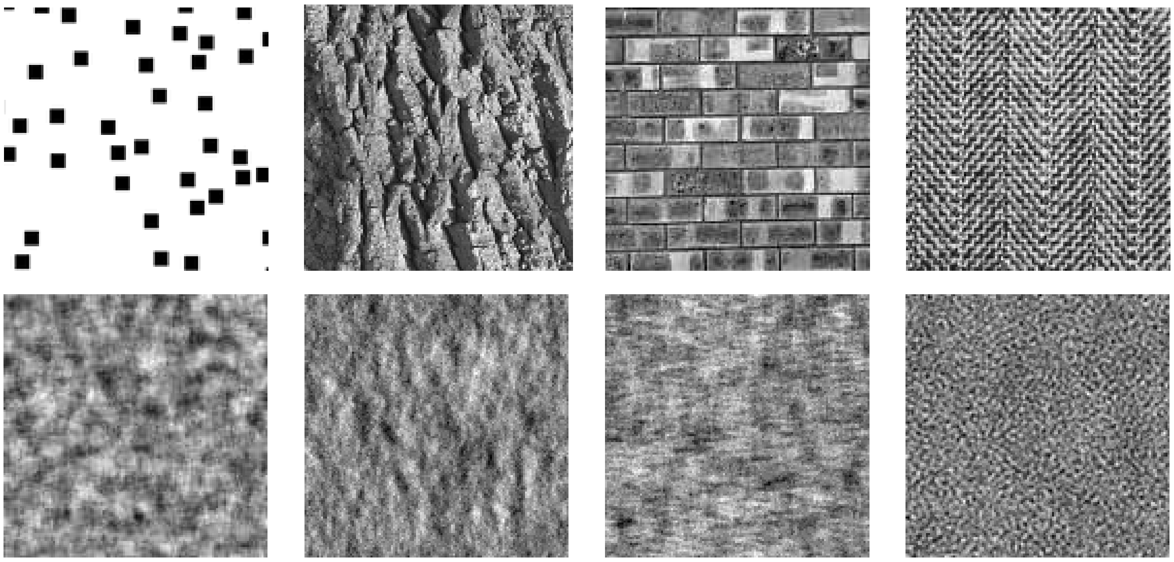

Fig. 4. Inadequacy of subband variance statistics. Shown are four pairs of original and synthesized texture images. For each pair, the steerable pyramid subbands have the same variance.

shows four example texture images, each accompanied by a synthesized image having the same subband auto-correlation. The examples were synthesized by computing the Fourier amplitude of the sample, generating an image of Gaussian noise samples with standard deviations matched to this amplitude, inverse Fourier transforming, and matching the pixel histogram. If one does not include this last step, the results are significantly worse visually. The auto-correlation clearly fails to adequately capture the visual appearance of these textures. The best of the four synthesized examples is the herringbone image. In this case, the Fourier amplitude spectrum contains sharp spectral peaks corresponding to the periodicity of the diagonal staircase stripes.

2.4 Subband Marginal Densities

In the previous section, we considered 2nd-order (covariance) statistical constraints. These constraints, coupled with a maximal entropy assumption, imply Gaussian densities. But natural images are known to give rise to non-Gaussian subband marginal densities [*e.g.*, 25, 18, 45], that are sharply peaked at zero with heavier tails than a Gaussian of the same variance. Intuitively, the explanation for this is that images have spatial structure consisting of smooth areas interspersed with occasional edges or other abrupt transitions. The smooth regions lead to near-zero coefficients, and the edges give sparse large-amplitude coefficients.

The marginal densities of subbands of natural textures are also non-Gaussian, but are not as readily described as those of natural images. Figure 2.4 shows empirical marginal densities (histograms) for example wavelet subbands of our four texture images, compared with Gaussian densities of the same variance. Such histogram measurements have been used as the constraining statistics for texture synthesis algorithms by Heeger and Bergen [36], and by Zhu *et al.* [70, 71].

Constraining a set of marginals of a probability density can impose joint statistical relationships between the variables. In particular, both [36, 70] impose marginals on an *overcomplete* set of linear (and possibly nonlinear) measurements. Although it may be shown that large numbers of marginals are sufficient to uniquely constrain a high-dimensional probability density [70][1], we believe that the marginals of a fixed finite linear basis are insufficient to properly capture the structure of many textures. In particular, long-range structures (such as straight or curved contours), nearly periodic patterns, and second-order textures are not well represented. Such structures introduce striking statistical dependencies between coefficients that are not explicitly represented in techniques based on marginals. We discuss this in greater detail in the next section.

As an example, figure 7 shows pairs of images whose subbands have the same empirical marginal densities (histograms). The marginals were imposed using the technique of [36]. The synthesized "jittered squares" texture is preferable to that shown in the previous section. But the representation of the herringbone cloth is worse. The brick wall has lost some periodicity, but gained some local structure.

In our own work, we have found that imposing the second and fourth moments is usually sufficient to produce marginal densities well-matched to those of the original texture. An indirect way of imposing these moments is to adjust the mean and variance of the energy (squared magnitude) of the complex channels. The mean of the energy controls the second moment of the coefficients, while the variance

[1] The theorem is a variant of the Fourier projection-slice theorem used for tomographic reconstruction.

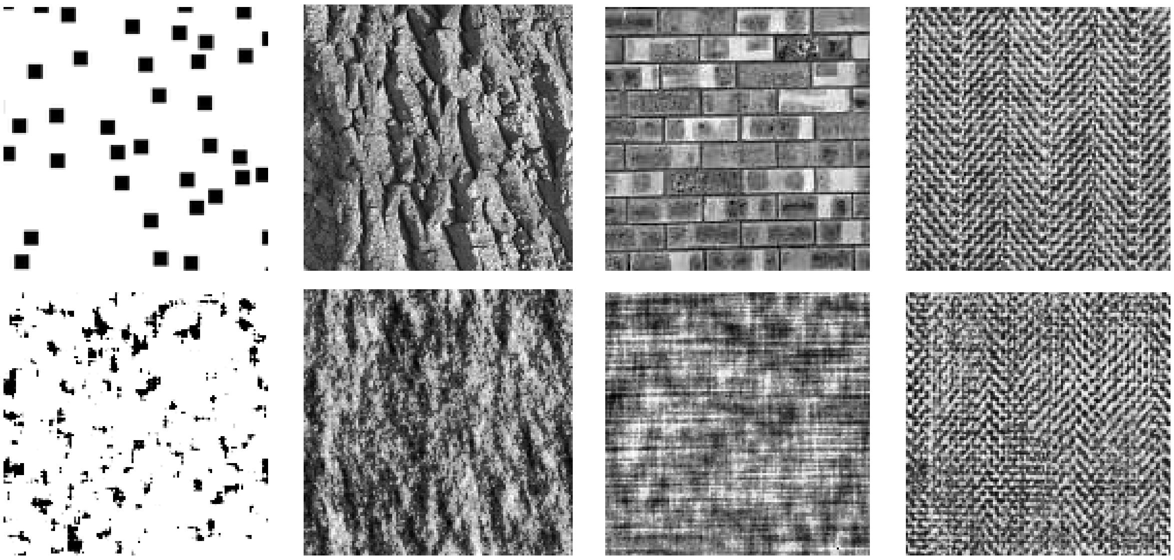

Fig. 5. Inadequacy of covariance statistics. Shown are four pairs of original and synthesized texture images. For each pair, the steerable pyramid subbands of the two images have identical mean and covariance statistics.

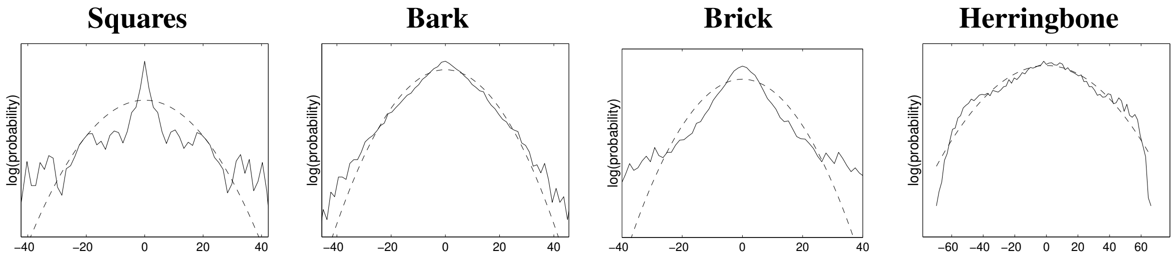

Fig. 6. Marginal histograms for one steerable pyramid subband of our four different texture images. Solid lines: log empirical densities (histograms). Dashed lines: Gaussian density with matched variance.

controls the fourth moment.

2.5 Subband Joint Magnitude Statistics

Each of the statistical measures described thus far have been used by previous authors for texture synthesis. In this section, we describe a novel statistical measurement that captures important structural information about textures. In recent work, we have studied the joint statistics of wavelet coefficient amplitudes [13, 12, 57], and have found that these are quite regular. In particular, we have examined the conditional histograms of pairs of coefficient magnitudes at adjacent spatial locations, orientations, and scales. We find that the variance of the conditional density often scales with the square of the conditioning coefficient, even when the raw coefficients are uncorrelated. There is an intuitive explanation for this: the "features" of real images give rise to large coefficients in local spatial neighborhoods, as well as at adjacent scales and orientations.

In the context of this paper, we examine the joint statistics of the complex coefficient magnitudes, and attempt to characterize and impose these directly. Figure 8 shows the steerable pyramid coefficient magnitudes of the four texture images of figures 5 and 7. One can see that the magnitudes (and the relationship between them) capture important structural information about each of these textures. In particular, large magnitudes appear at the same locations in the "squares" texture, and large magnitudes of the two diagonal orientations are anti-correlated in the herringbone texture. To make this more precise, we examine conditional histograms of pairs of coefficients, as shown in figure 9. The left histogram shows statistics (gathered over spatial position, from the squares image) from the magnitudes of two coefficients at adjacent spatial scales. Note the strong linear correlation. The right histogram shows statistics taken from the herringbone texture for a pair of coefficients at the two diagonal orientations. Here the two magnitudes are anti-correlated. Such correlations are often present despite the fact that the raw coefficients may be uncorrelated. This occurs because variations in phase across the image lead to cancellation. We need to select a descriptor of this dependence that is easy to measure and impose in a synthesis algorithm. As a simple summary of this dependence, we compute the covariance of the complex magnitude of pairs of coefficients.

In general, we observe that edges, the most basic features of natural images, account for much of the dependency between neighboring magnitudes. In particular, we typically find:

1. Positive covariance of neighbors at the same location and orientation, at different scales. This dependence extends through several octaves.
2. Positive covariance with neighbors at the same orientation and scale, at location in the direction corresponding to the orientation of the channel, due to the continuous arrangement in edges and bars of the significant responses.
3. For neighbors at the same location and scale, the covariance varies from positive to negative as a function of the preferred orientation. Adjacent orientation bands tend to respond to the same edges, giving rise to positive correlation. But the majority of edge structures in images represent a single orientation, and thus will produce little or no response in orthogonally or near-orthogonally oriented filters.

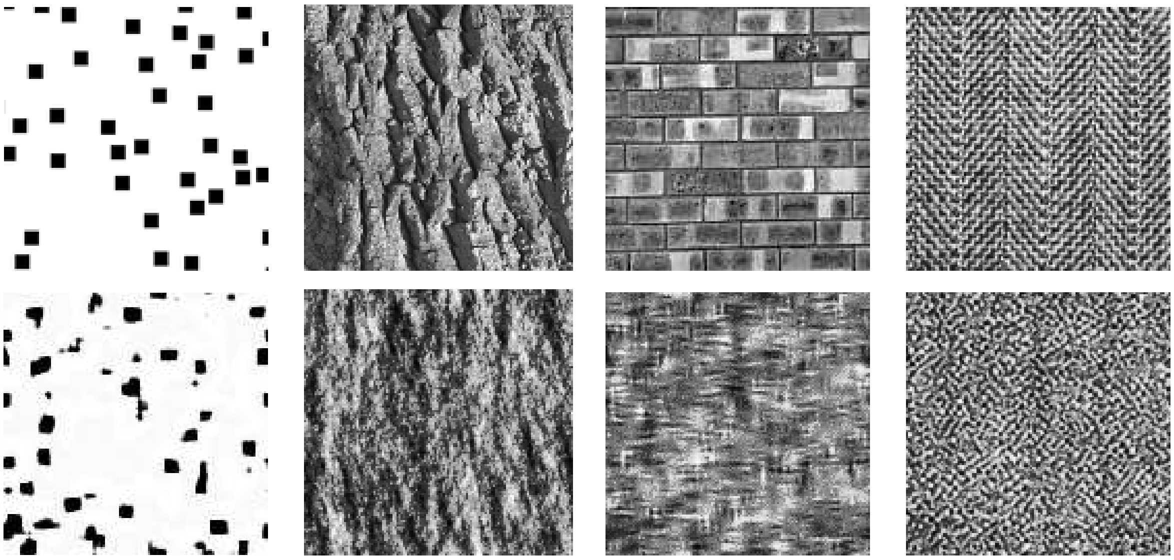

Fig. 7. Inadequacy of subband marginal statistics. Shown are four pairs of original and synthesized texture images. For each pair, the pyramid subbands of the two images have identical histograms.

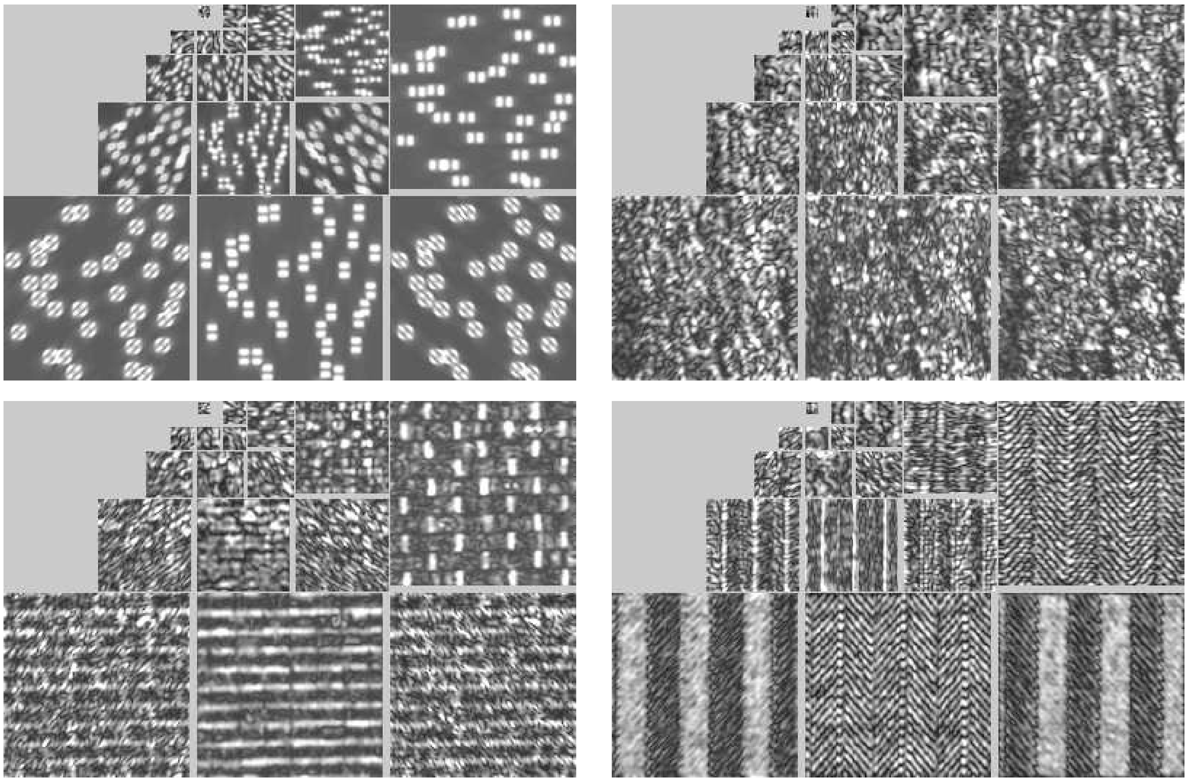

Fig. 8: Normalized magnitude responses in the subbands for the images of figures 5 and 7.

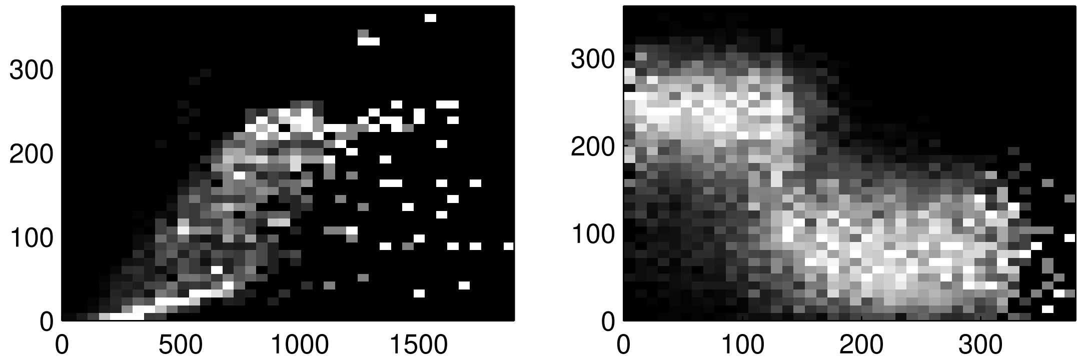

Fig. 9. Empirical conditional densities of neighboring wavelet coefficient magnitudes, displayed as grayscale images. Image brightness corresponds to frequency of occurrence of each pair of values, and each column is independently normalized to fill the range $[0, 1]$. Left: two coefficients at adjacent scales of the "jittered squares" image. Right: two coefficients at opposite diagonal orientations of the "herringbone" image.

2.6 Summary of Statistical Constraints

As a summary of the statistical model proposed, we enumerate the set of statistical descriptors, review the features they capture, and compute their associated number of parameters.

- Image pixel statistics: Mean, variance, skewness, kurtosis, minimum and maximum values. (6 parameters).
- Raw coefficient auto-correlation: Central samples of the auto-correlation of the real part of each subband ($N \cdot K \cdot \frac{M^2+1}{2}$ parameters), and mean, variance, minimum and maximum of lowpass and highpass residual bands (8 parameters). These characterize the regularity (linear predictability) of the texture.
- Coefficient magnitude statistics: Central samples of the auto-correlation of each subband (magnitude) ($N \cdot K \cdot \frac{M^2+1}{2}$ parameters), cross-correlation of each subband with other orientations at the same scale ($N \cdot \frac{K(K-1)}{2}$ parameters), and cross-correlation of each subband with other orientations at a coarser scale ($K^2(N-1)$ parameters). These represent significant structures in images, such as edges, bars, corners, repeated patterns and "second order" structure.

For our texture examples, we have made choices of N=4, K=4 and M=7, resulting in a total of 846 parameters. This set represents the union of the parameters we found to be necessary for each of our example textures. Comparable results may be achieved for many of the individual textures using substantially fewer parameters.

3 Texture Synthesis via Constraint Enforcement

The set of statistical constraints described above does not uniquely determine the image, and thus a synthesis algorithm must necessarily choose from amongst the set of images matching the constraints. If the output of the algorithm is non-deterministic, then this choice will be drawn according to a probability density on the set of images matching the constraints. A desirable solution is to sample from the density with maximal entropy subject to the constraints [39]. As mentioned previously, this formalism was used successfully for texture synthesis by Zhu *et al.* [70]. The maximal entropy distribution has a form:

$$\mathcal{P}\left(\vec{I}\right) \propto \prod_n \exp[-\lambda_n \phi_n(I)],$$

where the set of statistical measurements are of the form $\mathcal{E}\left(\phi_n(I)\right)$. But solving for the λ_n typically requires computationally expensive Monte Carlo techniques such as Gibbs sampling [30, 70]. In addition, this solution presupposes that one has chosen precisely the "correct" set of statistical constraints $\phi_n(I)$, since these completely determine the form of the density. For example, just because one has chosen to measure covariance statistics does not mean that the density of the observations must be Gaussian! Finally, an idealized goal for our choice of statistical features is that all images producing the same statistical measurements should be visually indistinguishable. If we were to achieve this, then the choice of nondeterministic algorithm by which one selects an image from this visual equivalence class becomes much less important.

Once we decide to abandon the maximal-entropy formalism, we have considerably more freedom to develop a non-deterministic algorithm that produces a synthetic image whose sample statistics match those of an original texture image. We based our algorithm on the concept of alternating projections onto constraint surfaces. We start with an image of Gaussian white noise, and force this image to satisfy the statistical constraints by sequentially and repeatedly projecting onto the set of images satisfying each constraint. Specifically, we alternate between projecting the coefficients onto the manifold of coefficients having the desired statistics, and, after projecting the coefficients onto the image subspace (i.e., collapsing the pyramid), projecting the resulting image onto the set of images having the desired sample moments and range. Figure 10 shows a block diagram of this synthesis-by-analysis algorithm. The algorithm bears a close resemblance to the projection onto convex sets (POCS) approaches that have been used in image restoration [66, 37], and to the texture synthesis method of Heeger and Bergen [36].

In general, it is desirable to implement the projection operations so as to select the image in the constraint set that is closest (in a Euclidean sense) to the initial image. Producing a minimal change in the image improves the stability and convergence of the iterative process. But it is often not easy to solve for this projection, and in some cases we choose to simply project in the direction of the gradient of the constraint, until reaching the surface. This is described in more detail in the next section.

Convergence of an infinite sequence of alternated projections is only guaranteed for the case when there are only two sets, both convex, and the projections are orthogonal to the constraint surfaces [67]. Unfortunately, the constraint sets corresponding to our statistical measurements are typically not convex, there are more than two, and our projection operations are not always orthogonal. We note, however, that projection onto non-convex sets has been used successfully for tasks such as image reconstruction from the power spectrum (phase recovery) [66].

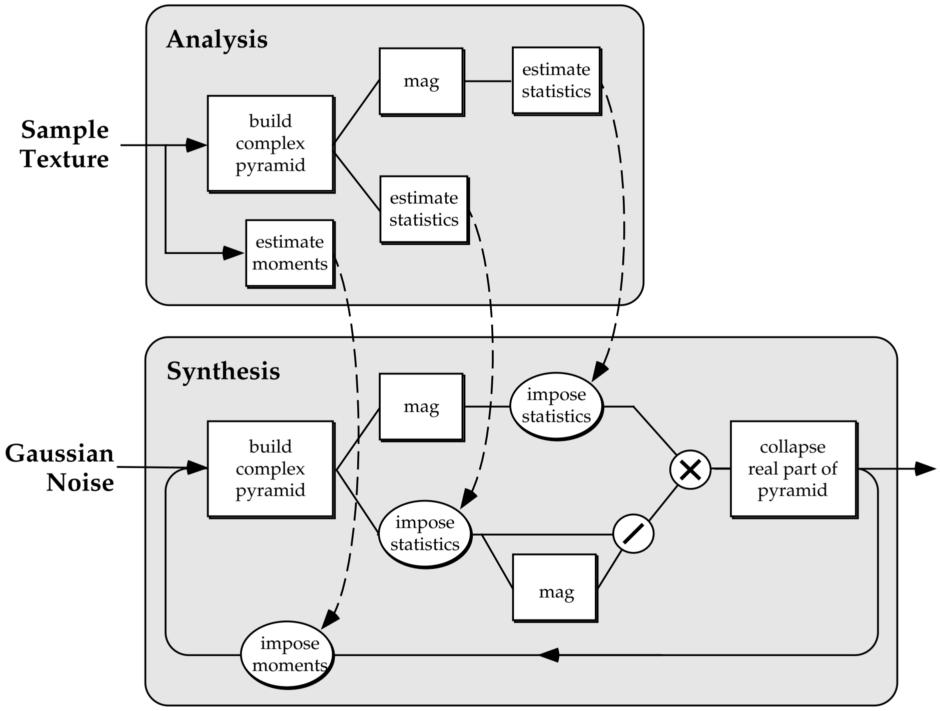

Fig. 10: Block diagram of texture synthesis algorithm. See text.

For the adjustment of magnitude statistics, we first adjust the auto-correlation, and then the joint correlation with other orientation bands and the next coarser scale. Note that the statistics of the raw coefficients are used to adjust the phase of the coefficients, and this is done independently of the magnitude adjustment. The resulting synthetic coefficients will thus have the correct magnitude statistics, but their raw statistics may be incorrect. While this is clearly non-optimal, it leads to a simple implementation that converges rapidly. In the following sections, we describe in more detail the method by which we impose each of the statistical constraints.

3.1 Imposing Pixel Moments

In the image domain, we impose the first through fourth moments of the pixel histogram. We would like to make a change to the image that is minimal but that gives the correct values for the sample moments. The first and second moments (mean and variance) are adjusted through a linear mapping of the pixel intensities. The normalized third and fourth moments (skew and kurtosis) are adjusted by adding a multiple of the gradient of the moment function to the image. In particular, the normalized moment is written

$$m_k(I) = \frac{\mathcal{E}\left((I_n - \bar{I})^k\right)}{\mathcal{E}\left((I_n - \bar{I})^2\right)^{k/2}},$$

where $\mathcal{E}(\cdot)$ indicates the expected value, which we estimate via the sample mean. Let M_k be the desired value for the moment. Then we solve for the λ that satisfies

$$m_k(I + \lambda \vec{\bigtriangledown} m_k(I)) = M_k.$$

In the limit as the projection distance shrinks to zero, this adjustment becomes an orthogonal projection onto the surface of images having the desired moment. The skew and kurtosis are adjusted sequentially within an iteration of the synthesis loop. In cases where no solution exists, we make no adjustment. This occurs sometimes in the first few iterations of the algorithm. But as the algorithm begins to converge, we have found that we can always find a solution.

3.2 Auto-Correlation of the Subband Responses

We have applied the same method to impose local $M \times M$ auto-correlation within each subband both on the raw coefficients (real part), and their magnitudes. We first solve for an $M \times M$ convolution kernel, $h_{n,m}$, that satisfies:

$$D_{n,m} = \sum_{\alpha,\beta} h_{\alpha,\beta} C_{n-\alpha,m-\beta},$$

where $C_{n,m}$ is the initial auto-correlation image, and $D_{n,m}$ is the desired auto-correlation. We then perform the convolution in the Fourier domain, using the (positive) square root:

$$F(I') = F(I) \cdot \sqrt{|H|},$$

where H is the Fourier transform of $h_{n,m}$, and $F(I)$ the Fourier transform of the original subband. This method is simple and direct, behaving well in terms of stability and convergence. It has an obvious drawback that when H is negative, the adjustment of the correlation using the filter $\sqrt{|H|}$ is not exact.

3.3 Cross-Correlation of the Subband Responses in Magnitude

We use a linear transform to impose correlations on the coefficient magnitudes. In general, consider the problem of adjusting a random vector, $\vec{x}$, to have a correlation matrix C_x, using a linear transform. That is, we seek M such that:

$$\mathcal{E}\left(M\vec{x}\vec{x}^T M^T\right) = C_x. \tag{2}$$

Let $B_x = \mathcal{E}\left(\vec{x}\vec{x}^T\right)$, and let $\{E_C, D_C\}$ and $\{E_B, D_B\}$ be appropriate eigenvector/eigenvector matrices such that:

$$B_x = E_B D_B D_B^T E_B^T \qquad C_x = E_C D_C D_C^T E_C^T \tag{3}$$

Then the complete set of solutions are specified as:

$$M = E_C D_C O D_B^{-1} E_B^T, \tag{4}$$

where O is any orthonormal matrix. Thus, the problem is reduced to choosing an orthonormal matrix O. We would like to choose O to minimize the MSE change in the vector $\vec{x}$. A reasonable (but not necessarily optimal) choice is:

$$O = E_C^T E_B. \tag{5}$$

The adjustment of the magnitude statistics is done in a coarse-to-fine fashion. At the coarsest scale (lowest frequency bands), we adjust the magnitudes as described above to achieve the correct cross-correlation between the set of oriented bands at the scale. For each subsequent finer scale, the magnitudes are adjusted so as to achieve the correct cross-correlation between bands at that scale, and also also with the bands at the (previously adjusted) coarser scale.

Consider the general problem of adjusting a random vector, $\vec{x}$, to have a self-correlation matrix C_x, as well as a cross-correlation matrix with variable $\vec{y}$ of C_{xy}. We compute a modified vector, $\vec{x}_{\mathrm{mod}} = M\vec{x} + K\vec{y}$, using matrices $\{K, M\}$ such that:

$$\begin{aligned} C_x &= \mathcal{E}\left(\vec{x}_{\mathrm{mod}}\vec{x}_{\mathrm{mod}}^T\right) \\ &= \mathcal{E}\left((M\vec{x} + K\vec{y})(M\vec{x} + K\vec{y})^T\right) \end{aligned}$$

and

$$\begin{aligned} C_{xy} &= \mathcal{E}\left(\vec{x}_{\mathrm{mod}}\vec{y}^T\right) \\ &= \mathcal{E}\left((M\vec{x} + K\vec{y})\vec{y}^T\right) \end{aligned}$$

Let $B_{xy} = \mathcal{E}\left(\vec{x}\vec{y}^T\right)$ and $B_y = \mathcal{E}\left(\vec{y}\vec{y}^T\right)$. Then solving the second equation for K gives:

$$K = [C_{xy} - MB_{xy}]B_y^{-1}.$$

Substituting this into the first equation and manipulating algebraically gives:

$$M[B_x - B_{xy}B_y^{-1}B_{xy}^T]M^T + C_{xy}B_y^{-1}C_{xy}^T = C_x.$$

Now let $\{E_C, D_C\}$ and $\{E_B, D_B\}$ be appropriate eigenvector/eigenvector matrices such that:

$$\begin{aligned} B_x - B_{xy}B_y^{-1}B_{xy}^T &= E_B D_B D_B^T E_B^T, \qquad \text{and} \\ C_x - C_{xy}B_y^{-1}C_{xy}^T &= E_C D_C D_C^T E_C^T \end{aligned}$$

Then, as before, the complete set of solutions for M may be specified as:

$$M = E_C D_C O D_B^{-1} E_B^T, \tag{6}$$

for O any orthogonal matrix. As above, we choose

$$O = E_C^T E_B. \tag{7}$$

3.4 Convergence, stability and computational cost

The algorithm we have described is fairly simple, but we cannot guarantee convergence. The projection operations are not exactly orthogonal (for example, the gradient projections are only orthogonal to the constraint surface in the small-displacement limit), and the constraint surfaces are not all convex. Nevertheless, we find that convergence is achieved after about 50 iterations, for the several hundred textures we have synthesized. In addition, once convergence has been achieved (to within some tolerance), the synthetic texture oscillates only slightly in its parameters, and no serious problems of stability have been recorded.

Figure 12 shows the evolution of the adjustment of the parameters as a function of the number of iterations, for a representative example (the "herringbone" texture). The dependence of the error is approximately linear with n in a log-log plot, implying that the error in the decreases roughly as $k/(n^c)$, with k and c constants dependent on the choice of parameter and texture. Figure 11 shows the corresponding synthetic texture for 0, 1, 3, 7, 15, and 31 iterations, illustrating the rapid visual convergence of the algorithm. Our current implementation (in MatLab) requires roughly 5 minutes to synthesize a 128×128 texture on a 200 Mhz Pentium workstation.

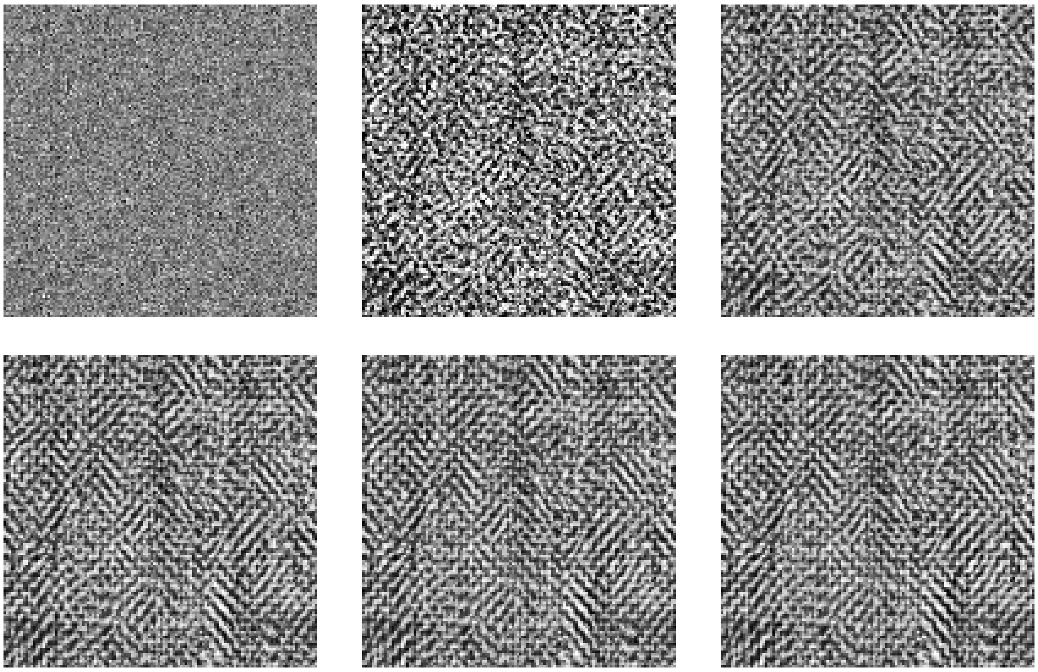

Fig. 11: Herringbone synthesis, for 0, 1, 3, 7, 15 and 31 iterations

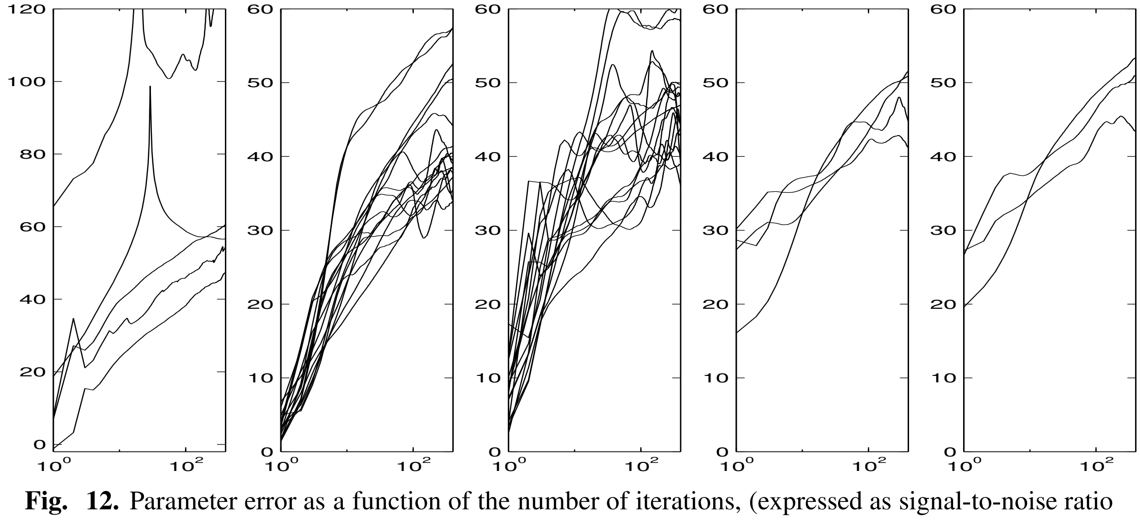

Fig. 12. Parameter error as a function of the number of iterations, (expressed as signal-to-noise ratio in decibels). (a) First-order statistics descriptors (from top to bottom: mean, variance, kurtosis, range and skewness). (b) Auto covariance of the real part of each subbands (16); (c) Auto-covariance of the magnitude responses; (d) Cross-covariance of the magnitude responses at different orientations, for each scale (4); (e) Cross-covariance of the parent scales with the child scale (3).

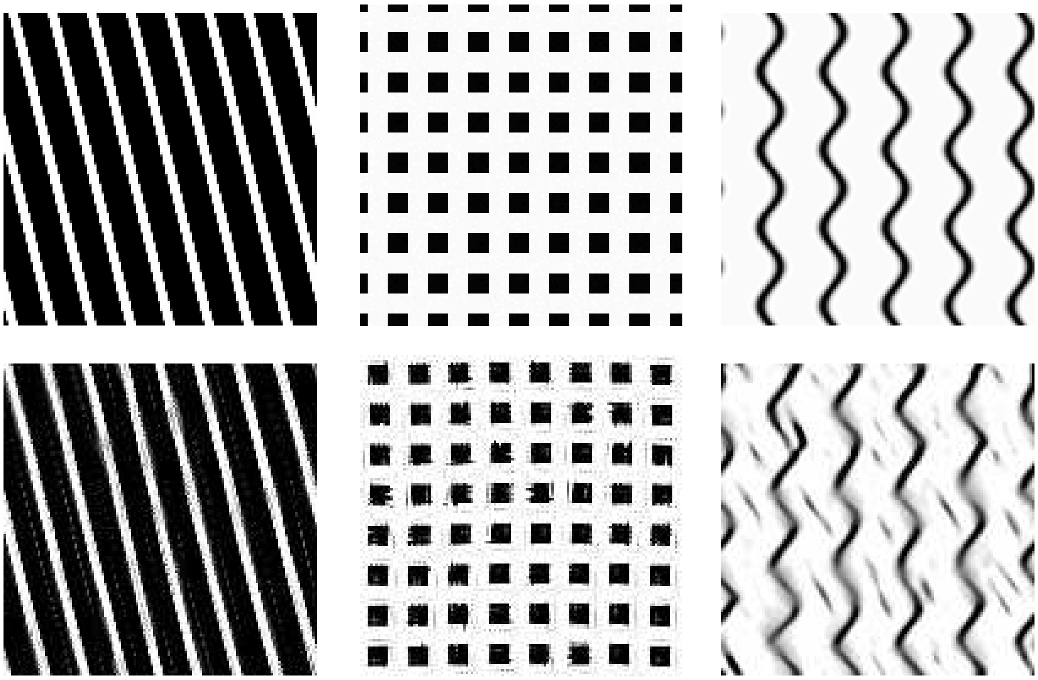

Fig. 13. Synthesis results on artificial periodic textures. For each pair of textures, the upper image is the original texture, and the lower image is the synthesized texture.

4 Results

4.1 Synthesis Examples

Texture synthesis provides a method to validate a texture model. In this section we present a set of synthesis results obtained using 4 scales and 4 orientations, a neighborhood of 7x7 pixels and 50 iterations of the algorithm on images of 128x128 pixels.

Figure 13 shows a set of synthesis results on artificial periodic texture images. The algorithm handles such input quite well, producing output with the same periodicities and structures. Note that the absolute phase of the synthesized images is random, due to the translation-invariant nature of the algorithm, and the fact that we are treating boundaries periodically in all computations.

Figure 14 shows a set of synthesis results on classical "texton" images, as pioneered by Julesz [40]. The textons are placed at independent random positions within the image. The algorithm does a reasonable job of recreating these images, although there are significantly more artifacts than in the periodic examples.

Figures 15 and 16 show synthesis results for natural textures. The first set contains highly random textures, such as the animal fur or wood grain. The synthesized images are virtually indistinguishable from the originals. The second set contains more regular structured textures, such as a brick wall and various types of woven fabric. These are also reproduced quite well. Comparing these results with

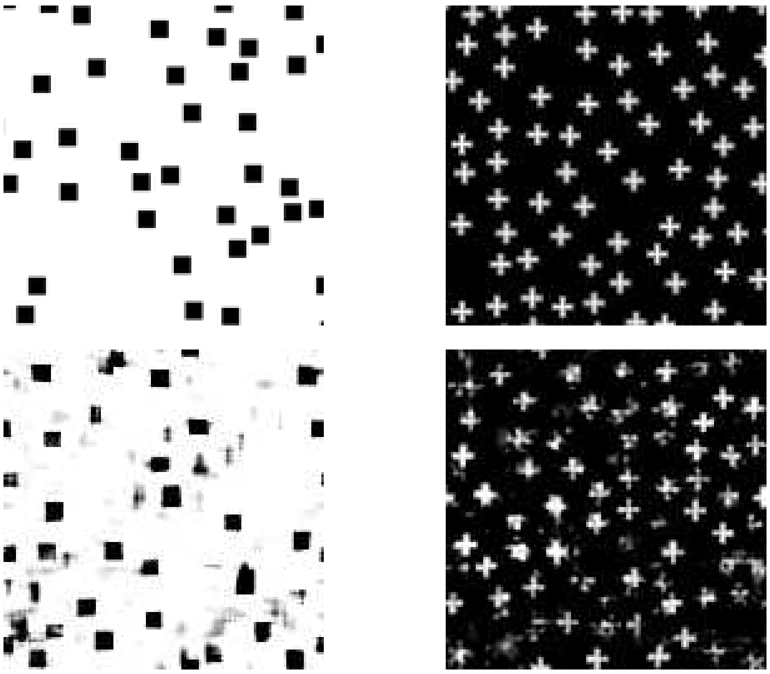

Fig. 14: Synthesis results on artificial "texton" images [40].

those obtained using the statistical models analyzed (see figures 5 and figures 7), the higher capacity of the proposed model is apparent.

Finally, it is instructive to apply the algorithm to images that are structured and highly inhomogeneous. Figure 17 shows a set of synthesis results on an artificial image of a "bull's eye", a real image of a crowd of people, and a face. The first two produce quite compelling results that capture the local structure of the original images, albeit in a globally disorganized fashion. The third is also interesting, in that it seems to contain face-like structures, again in a globally disorganized arrangement.

Not all the tested textures have provided positive results. Figure 18 shows an image of wet stones. The image is highly structured, with reflection of light on the stone surfaces creating a strong impression of three-dimensionality. The algorithm produces a synthesized image that bears some resemblance to the original, but lacks this three-dimensionality. Also in figure 18 is an image containing stripes with a "sawtooth" (linear ramp) cross section. The result contains stripes of the correct periodicity and orientation, but the cross sectional shape is irregular. The algorithm is unable to represent the local phase structure of this pattern, due to our reliance on the real symmetric portion of the complex transform. It is interesting to note, however, that the synthesized image is quite visually similar to the original. Another clear example of a failure to properly represent local phase is an image of piecewise constant polygonal patches. We can see that the model, although detecting and reproducing discontinuities, is not able to differentiate between odd-symmetric edges and even-symmetric lines.

Finally, as a demonstration of the flexibility of the technique, we consider the problem of creating textures that lie visually in between two other textures. Figure 19 shows a sequence of four textures synthesized using sample statistics that are weighted averages of those corresponding to two original

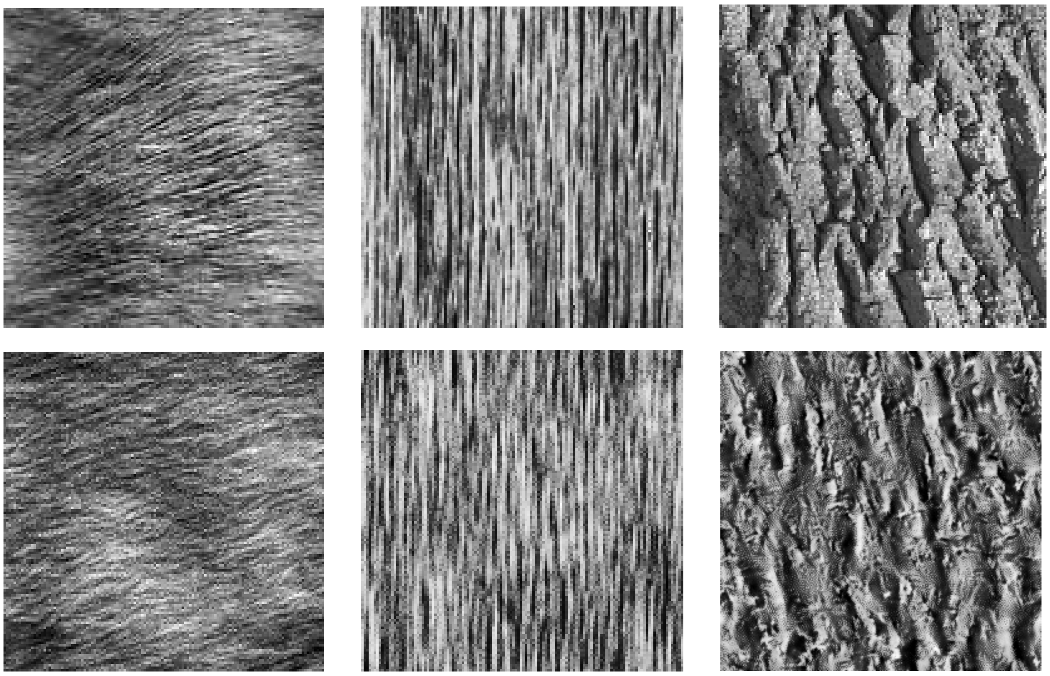

Fig. 15. Synthesis results on natural “random” textures. See caption of figure 13. These are from the Brodatz texture album [11]

.

textures.

4.2 Constrained Synthesis Examples

In addition to the synthesis application demonstrated in the previous section, it is often of interest to synthesize images subject to some set of constraints. In particular, we consider the problems of extending an image beyond its spatial boundaries (spatial extrapolation), and extending an image to higher resolution (scale extrapolation). In the former case, the additional constraints are that the original image pixels are known. In the latter, the additional constraints are that the coefficients of the pyramid representation at all resolutions coarser than some the starting resolution are known.

Spatial extrapolation. In this application, we want to synthesize an image in which a subset of the pixels are known (i.e., taken from an existing image). The set of all images with this subset of pixels matching the known values is convex, and the projection onto this convex set is easily inserted into the iterative loop of the synthesis algorithm. Specifically, we simply re-set the constrained pixels on each iteration of the synthesis loop. This technique is applicable to the restoration of pictures which have been destroyed in some region (”filling holes”) [e.g., 37]. Figure 20 shows two examples that have been spatially extrapolated using this method. In these examples, the set of statistical parameters have been obtained from the original images. Notice that the border between real and synthetic data is barely noticeable.

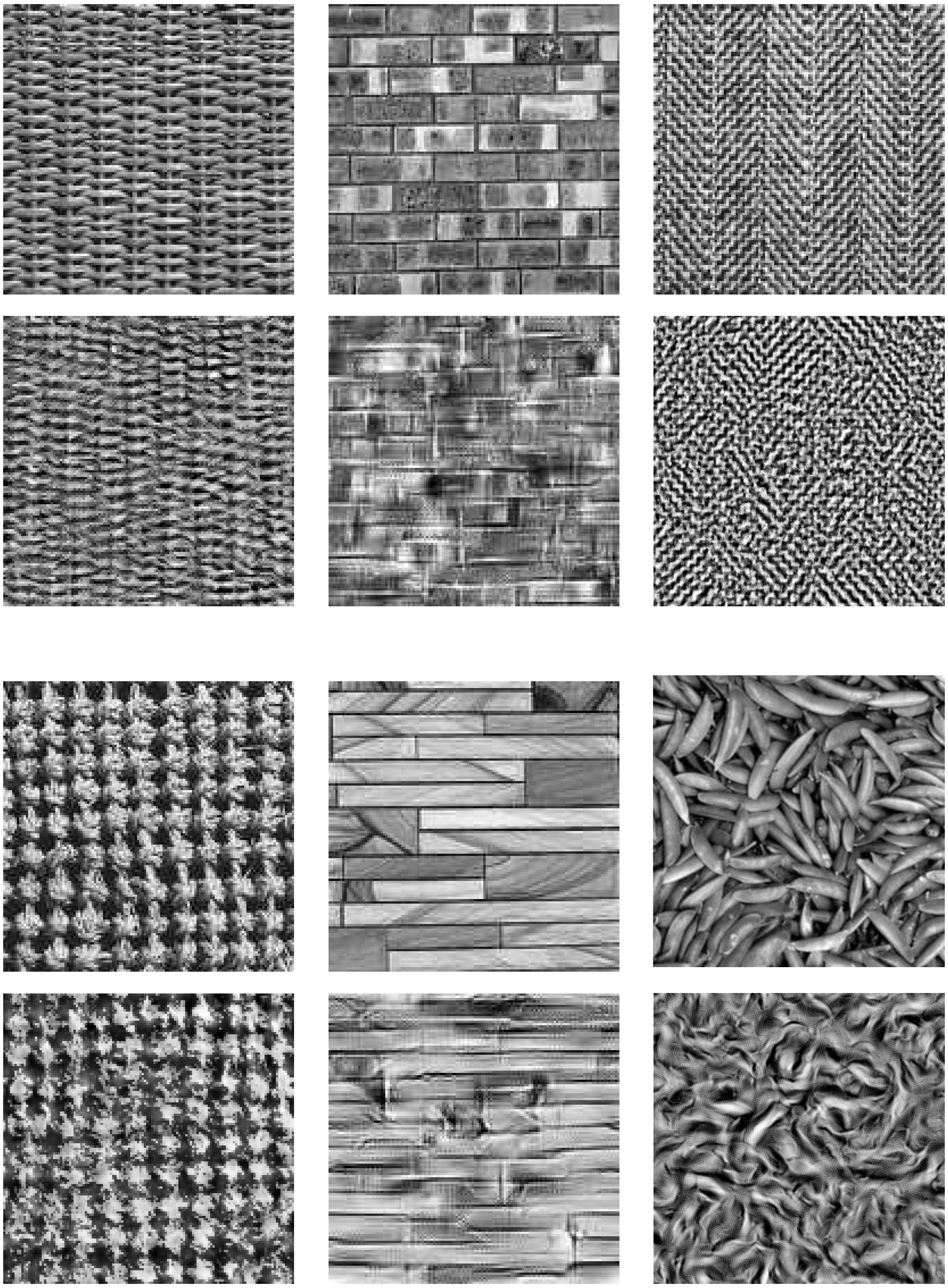

Fig. 16: Synthesis results on natural “structured” textures. See caption of figure 13.

.

Fig. 17: Synthesis results on structured images not usually considered to be "texture".

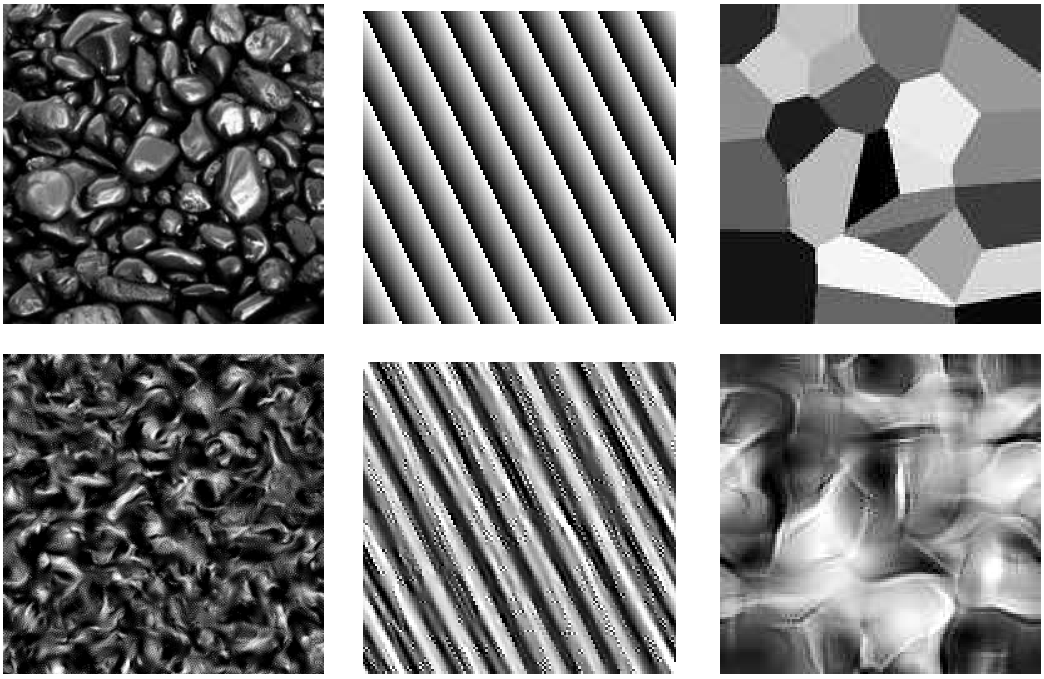

Fig. 18: Synthesis failures.

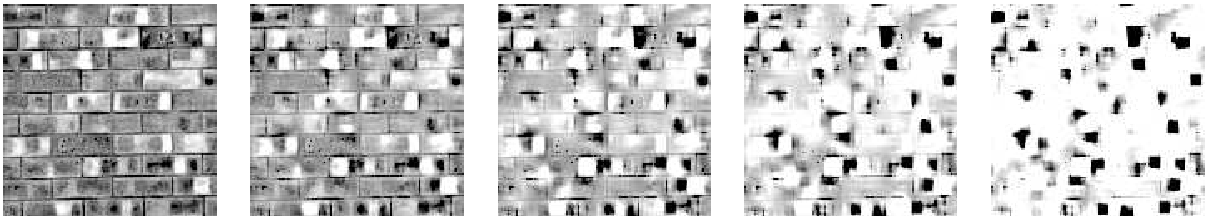

Fig. 19. Sequence of five synthesized textures that span the range "between" the "brick" and "jittered squares" images.

Scale extrapolation. Considering that an important part of the energy of natural images correspond to the edges, and that these are very redundant over a range of scales , it seems feasible to extrapolate some of the finer scale information from the coarser scale coefficients. This is a variant of the so-called "super-resolution" problem. We modified our algorithm to accomplish this by replacing the coarse-scale subbands with those of a low-resolution image on each iteration. This replacement is done just after constructing the pyramid. Figure 21 shows an example. Much of the appearance of the original has been recovered.

5 Discussion

We have described a model for visual texture, based on a set of statistical measurements made in an overcomplete multi-scale image representation. The most significant contribution of our model is the introduction of non-Gaussian joint statistical relationships between coefficients. We demonstrated the quality of the model by synthesizing a wide variety of visual textures using an iterative synthesis-by-analysis statistical sampling algorithm. We believe these synthesis results to be competitive, both in terms of quality of the results and in computational cost, with the current best techniques. We also demonstrated the flexibility of this algorithm by using it in several applications that may be described as "constrained synthesis" problems.

The most fundamental issue in our model is the choice of statistical measurements. While we have tried to motivate these by empirical studies on real textures, better choices undoubtedly exist. In particular, the failures of figure 18 are primarily due to the lack of a representation of local phase information (e.g., even-symmetric lines vs. odd-symmetric intensity edges). We view this an important area for future development. Another important area for exploration is the use of adaptive bases. Our technique currently uses a fixed basis whose properties are appropriate for typical images. Adapting the basis to the statistical properties of each individual texture (as is described, for example, in [71]) could give both visually improved results and a reduction in the size of the basis.

A secondary issue is the somewhat ad hoc nature of our sampling algorithm. In our experience, the algorithm has never failed to converge to an image that nearly satisfies the statistical constraints (as shown in figure 12). Nevertheless, we do not have a proof of convergence. A deeper problem with the algorithm is that our probability model is *implicit*. Specifically, we do not not sample from the density with maximal entropy, but rather impose the statistical constraints through our iterative procedure. This induces a probability density on the space of output images, but we do not have a form for this

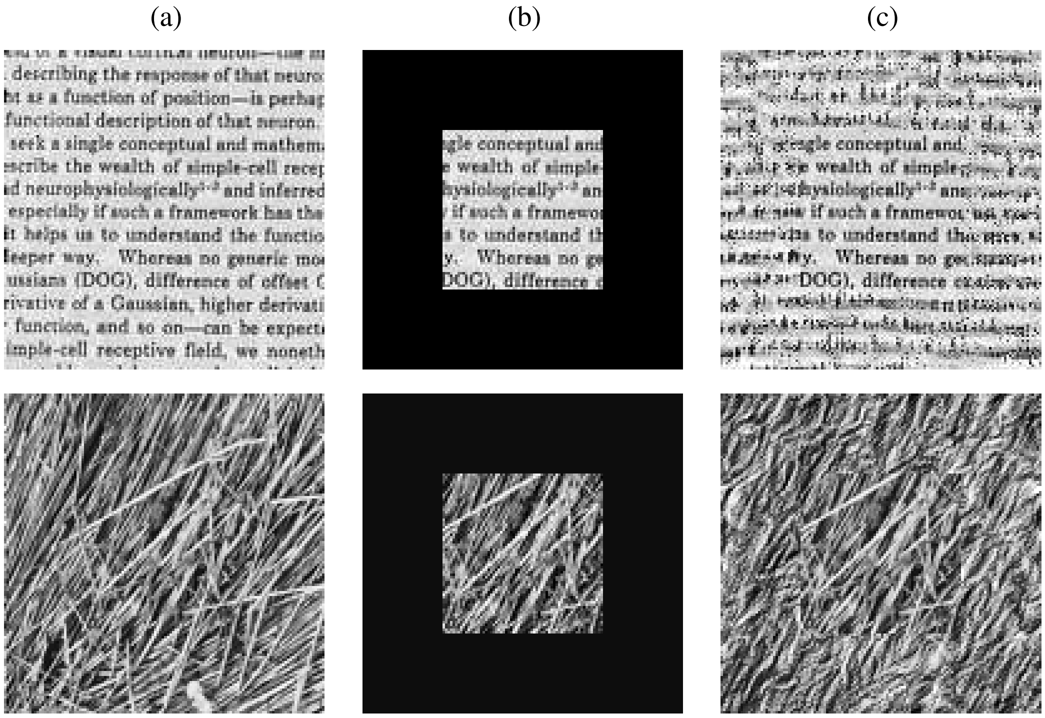

Fig. 20. (a) Original images; (b) Conserved samples; (c) Extended images, synthesized using statistical measurements from (a) while retaining the samples of (b).

density and we do not know its entropy. Nevertheless, as described earlier, we believe that maximal entropy, although desirable, is not an indispensable property of a texture synthesis algorithm. Our technique achieves a good compromise between computational cost and what in practice seems to be a high entropy of the associated implicit density.

Our choice of constraints bears a resemblance to observations made in a number of other areas. In particular, we motivated the second-order statistics as a means of capturing periodicity and linear features, as represented by uneven spectral distribution within subbands. We used the second-order magnitude statistics as a means of capturing so-called local structures. We can also interpret these as constraints on local coefficient magnitude and phase, since the second-order coefficient statistics determine the local arrangement of coefficient phases (see Figure 10). Anderson [1] has synthesized textures using statistical constraints based on the amplitude and sign of subband coefficients.

We envision a number of extensions to our synthesis algorithm. The results could be made much more visually compelling by introducing color (e.g., using the approach of [36]. Since the representation of a texture image is quite compact (846 parameters, for our examples), the synthesis technique might be used in conjunction with a compression system in order to "fabricate" detail rather than encode it exactly.

(a) (b) (c)

Fig. 21. (a) Original images; (b) Images reconstructed without including the finest scale subbands of the pyramid; (c) Images extended in scale, using the course-scale subbands of (b) and the statistics of (a).

References

[1] C H Anderson and W D Langer. Statistical models of image texture. Technical report, Washington U. Medical School, 1997. Available at ftp://shifter.wustl.edu/pub/.

[2] J J Atick. Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems*, 3:213–251, 1992.

[3] F Attneave. Some informational aspects of visual perception. *Psych. Rev.*, 61:183–193, 1954.

[4] H B Barlow. Possible principles underlying the transformation of sensory messages. In W A Rosenblith, editor, *Sensory Communication*, page 217. MIT Press, Cambridge, MA, 1961.

[5] A J Bell and T J Sejnowski. The 'independent components' of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338, 1997.

[6] J R Bergen and E H Adelson. Visual texture segmentation based on energy measures. *J. Opt. Soc. Am. A*, 3, 1986.

[7] J R Bergen and E H Adelson. Early vision and texture perception. *Nature*, 333:363–364, 1988.

[8] C A Bouman and M Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. Image Proc.*, 3(2), 1994.

[9] A C Bovik, M Clark, and W S Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Pat. Anal. Mach. Intell.*, 12(1), 1990.

[10] A C Bovik, M Clark, and W S Geisler. Localized measurements of emergent image frequencies by Gabor wavelets. *IEEE Pat. Anal. Mach. Intell.*, 38:691–712, 1992.

[11] P Brodatz. *Textures: A Photographic Album for Artists and Designers*. Dover, New York, 1966.

[12] R W Buccigrossi and E P Simoncelli. Image compression via joint statistical characterization in the wavelet domain. Technical Report 414, GRASP Laboratory, University of Pennsylvania, May 1997.

[13] R W Buccigrossi and E P Simoncelli. Progressive wavelet image coding based on a conditional probability model. In *ICASSP*, volume IV, pages 2957–2960, Munich, Germany, April 1997. IEEE Sig Proc Society.

[14] J A Cadzow, D M Wilkes, R A Peters II, and X Li. Image texture synthesis–by–analysis using moving–average models. 29(4):1110–1122, 1993.

[15] D Cano. Texture synthesis using hierarchical linear transforms. *Signal Processing*, 15:131–148, 1988.

[16] P C Chen and T Pavlidis. Segmentation by texture using correlation. *IEEE Pat. Anal. Mach. Intell.*, 5(1):64–69, 1983.

[17] G Cross and A Jain. Markov random field texture models. *IEEE Trans PAMI*, 5:25–39, 1983.

[18] J Daugman. ?? title ?? *IEEE Trans. Acoust. Speech Signal Proc.*, 36(7):1169–1179, 1988.

[19] John G. Daugman. Entropy reduction and decorrelation in visual coding by oriented neural receptive fields. *IEEE Trans. Biomedical Engineering*, 36(1):107–114, 1989.

[20] J De Bonet and P Viola. A non-parametric multi-scale statistical model for natural images. In *Adv in Neural Info Processing*, volume 9. MIT Press, December 1997.

[21] H Derin and H Elliott. Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE Pat. Anal. Mach. Intell.*, 9(1):39–55, 1987.

[22] D Dunn, W E Higgins, and J Wakeley. Texture segmentation using 2-D Gabor elementary functions. *IEEE Pat. Anal. Mach. Intell.*, 16(2), 1994.

[23] I M Elfadel and R W Picard. Gibbs random fields, co-occurrences, and texture modeling. *IEEE Pat. Anal. Mach. Intell.*, 16(1):24–37, Jan 1994.

[24] O D Faugeras and W K Pratt. Decorrelation methods of texture feature extraction. *IEEE Pat. Anal. Mach. Intell.*, 2(4), 1980.

[25] D J Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394, 1987.

[26] J M Francos, A Z Meiri, and B Porat. A unified texture model based on a 2-D Wold-like decomposition. *IEEE Trans. Signal Proc.*, 41(8):2665–2678, 1993.

[27] W T Freeman and E H Adelson. The design and use of steerable filters. *IEEE Pat. Anal. Mach. Intell.*, 13(9):891–906, 1991.

[28] A Gagalowicz. A new method for texture fields synthesis: Some applications to the study of human vision. *IEEE Pat. Anal. Mach. Intell.*, 3(5), 1981.

[29] A Gagalowicz and S D Ma. Sequential synthesis of natural textures. *Comp. Vis. Graphics Image Proc.*, 30:289–315, 1985.

[30] S Geman and D Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Pat. Anal. Mach. Intell.*, 6:721–741, 1984.

[31] C C Gotlieb and H E Kreyszig. Texture descriptors based on co-occurance matrices. *Comp. Vis. Graphics Image Proc.*, 51:70–86, 1990.

[32] N Graham. *Visual pattern analyzers*. Oxford University Press, New York, 1989.

[33] H. Greenspan, S. Belongie, R. Goodman, P. Perona, S. Rakshit, and C. H. Anderson. Overcomplete steerable pyramid filters and rotation invariance. In *Proceedings, CVPR*, pages 222–228, 1994.

[34] R M Haralick. Statistical and structural approach to texture. *Proc. IEEE*, 67:786–804, May 1979.

[35] M Hassner and J Sklansky. The use of Markov random fields as models of texture. *Comp. Graphics Image Proc.*, 12:357–370, 1980.

[36] D Heeger and J Bergen. Pyramid-based texture analysis/synthesis. In *Proc. ACM SIGGRAPH*, August 1995.

[37] A N Hirani and T Totsuka. Combining frequency and spatial domain information for fast interactive image noise removal. In *ACM SIGGRAPH*, pages 269–276, 1996.

[38] H Iversen and T Lonnestad. An evaluation of stochastic models for analysis and synthesis of gray scale texture. *Pattern Recognition Letters*, 15:575–585, 1994.

[39] E T Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.

[40] B Julesz. Visual pattern discrimination. *IRE Trans Info Theory*, IT-8, 1962.

[41] B Julesz. Textons, the elements of texture perception and their interactions. *Nature*, 290:91–97, 1981.

[42] R L Kashyap, R Chellappa, and A Khotanzad. Texture classification using features derived from random field models. *Patt Rec Letters*, 1:43–50, Oct 1982.

[43] D Kersten. Predictability and redundancy of natural images. *J. Opt. Soc. Am. A*, 4(12):2395–2400, 1987.

[44] H Knutsson and G H Granlund. Texture analysis using two-dimensional quadrature filters. In *Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, pages 206–213. IEEE Computer Society, 1983.

[45] S G Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Pat. Anal. Mach. Intell.*, 11:674–693, July 1989.

[46] R Navarro and J Portilla. Robust method for texture synthesis–by–analysis based on a multiscale Gabor scheme. In *Proc. of the SPIE, vol. 2657*, pages 86–96, San Jose, CA, Jan 1996.

[47] B A Olshausen and D J Field. Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7:333–339, 1996.

[48] P Perona and J Malik. Detecting and localizing edges composed of steps, peaks and roofs. In *Proc. 3rd Intl. Conf. Computer Vision*, Osaka, Japan, 1990.

[49] A C (K) Popat. *Conjoint probabilistic Subband Modeling*. PhD thesis, Massachusetts Institute of Technology, Program in Media Arts and Sciences, Cambridge, MA, September 1997.

[50] K Popat and R W Picard. Cluster-based probability model and its application to image and texture processing. *IEEE Trans Im Proc*, 6(2):268–284, 1997.

[51] J Portilla, R Navarro, O Nestares, and A Tabernero. Texture synthesis-by-analysis based on a multiscale early-vision model. *Optical Engineering*, 35(8), 1996.

[52] W K Pratt, O D Faugeras, and A Gagolowicz. Visual discrimination of stochastic texture fields. *IEEE Trans. on Systems Man and Cybernetics*, Nov 1978.

[53] T R Reed and H Wechsler. Segmentation of textured images and gestalt organization using spatial/spatial-frequency representations. *IEEE Pat. Anal. Mach. Intell.*, 12(1), 1990.

[54] D L Ruderman and W Bialek. Statistics of natural images: Scaling in the woods. *Phys. Rev. Letters*, 73(6), 1994.

[55] E Simoncelli and J Portilla. Texture characterization via joint statistics of wavelet coefficient magnitudes. In *Fifth IEEE Int'l Conf on Image Proc*, volume I, Chicago, October 4-7 1998. IEEE Computer Society.

[56] E P Simoncelli. Normalized component analysis and the statistics of natural scenes. In *Natural Scene Statistics Meeting*, Hancock, MA, September 11-14 1997.

[57] E P Simoncelli. Statistical models for images: Compression, restoration and synthesis. In *31st Asilomar Conf on Signals, Systems and Computers*, pages 673–678, Pacific Grove, CA, November 1997. IEEE Computer Society.

[58] E P Simoncelli and W T Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Second Int'l Conf on Image Proc*, volume III, pages 444–447, Washington, DC, October 1995. IEEE Sig Proc Society.

[59] E P Simoncelli, W T Freeman, E H Adelson, and D J Heeger. Shiftable multi-scale transforms. *IEEE Trans Information Theory*, 38(2):587–607, March 1992. Special Issue on Wavelets.

[60] A Teuner, O Pichler, and B J Hosticka. Unsupervised texture segmentation of images using tuned matched Gabor filters. *IEEE Trans. Image Proc.*, 4(6), 1995.

[61] M R Turner. Texture discrimination by Gabor functions. *Biol. Cybern.*, 55:71–82, 1986.

[62] M Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. Image Proc.*, 4(11), 1995.

[63] A van der Schaaf and J H van Hateren. Modelling the power spectra of natural images: Statistics and information. *Vision Research*, 28(17):2759–2770, 1996.

[64] H Voorhees and T Poggio. Computing texture boundaries from images. *Nature*, 333:364–367, 1988.

[65] A. B. Watson. Efficiency of a model human image code. *J. Opt. Soc. Am. A*, 12:2401–2417, 1987.

[66] D C Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits and Systems*, 25:694–702, 1978.

[67] D C Youla and H Webb. Image restoration by the method of convex projections. 1:81–101, 1982.

[68] C Zetzsche, B Wegmann, and E Barth. Nonlinear aspects of primary vision: Entropy reduction beyond decorrelation. In *Int'l Symposium, Society for Information Display*, volume XXIV, pages 933–936, 1993.

[69] S Zhu and D Mumford. Prior learning and Gibbs reaction-diffusion. *IEEE Pat. Anal. Mach. Intell.*, 19(11), 1997.

[70] S Zhu, Y Wu, and D Mumford. Filters, random fields and maximum entropy (FRAME) – towards the unified theory for texture modeling. In *IEEE Conf Computer Vision and Pattern Recognition*, June 1996.

[71] S C Zhu, Y N Wu, and D Mumford. Minimax entropy principle and its application to texture modeling. In *Neural Computation*, volume 9, pages 1627–1660, 1997.