IEEE Transactions on Image Processing, Aug. 1996

Texture Coding Using a Wold Decomposition Model*

R. Sriram[†], Joseph M. Francos[‡], and William A. Pearlman [§]

Abstract

A novel approach for coding textured images is presented. The texture field is assumed to be a realization of a regular homogeneous random field, which can have a mixed spectral distribution. On the basis of a 2-D Wold-like decomposition, the field is represented as a sum of purely indeterministic, harmonic, and a countable number of evanescent fields. We present an algorithm for estimating and coding the texture model parameters, and show that the suggested algorithm yields high quality reconstructions at low bit rates. The model and the resulting coding algorithm are seen to be applicable to a wide variety of texture types found in natural images.

EDICS Categories: IP 1.1, IP 1.8

1 Introduction

Many natural images can be described as a finite ensemble of patches of uniform textures. In a previous paper, [1], we have presented a parametric texture model, and an estimation algorithm for estimating the parameters of individual textures. The previously derived texture model and estimation/synthesis algorithms are employed in the present paper as the basic building block of a contour-texture image coding method. In this method the image is segmented into its individual texture patches. For each patch the corresponding parametric model and the contour of its boundaries are estimated and coded. This type of method can be used for coding textured images, or for coding textured regions in larger images.

The texture model which is the basis for the image coding scheme suggested in the present paper

*This work was performed at Rensselaer Polytechnic Institute and was presented at the Twelfth International Conference on Pattern Recognition, Jerusalem, Israel, October 1994.

†Current address is Philips Laboratories, Briarcliff Manor, NY 10510; E-mail: ram@philabs.philips.com.

‡Current address is Electrical and Computer Engineering Dept., Ben-Gurion University, Beer-Sheva 84105, Israel; E-mail: francos@newton.bgu.ac.il.

§Current address is Electrical, Computer and Systems Engineering Dept., Rensselaer Polytechnic Institute, Troy, NY 12180: E-mail: pearlman@ecse.rpi.edu.

is based on a 2-D Wold-like decomposition for homogeneous random fields [2]. The texture field is decomposed into a sum of two mutually orthogonal components: a *deterministic* component which results in the structural attributes of the observed realization, and a *purely indeterministic* component which is the structureless, "random looking" component of the texture field. The deterministic field is further orthogonally decomposed into a harmonic component and a countable number of evanescent components. The harmonic field results in the periodic attributes of the texture, whereas the evanescent components result in directional ones. In other words, the parametric representation of the texture has meaningful interpretation in terms of the visual properties of the observed texture. Since the proposed texture model is based on the Wold decomposition results, it enables a rigorous mathematical treatment of the texture modeling and parameter estimation problems. Moreover, due to the generality of the decomposition, the model is not tailored to any specific type of texture.

In the present paper we elaborate on the problem of coding the estimated parameters of the individual texture fields. We show that the suggested coding scheme yields high compression ratios, while producing high quality synthesis results on the decoder side. The idea of contour-texture coding is not new, e.g., [3]. However, in the coding scheme proposed in [3], the texture is represented using 2-D polynomials, whereas in the present paper we propose a coding scheme which is based on a much more advanced texture model.

This paper is organized as follows: In Section 2 we present the texture model which is based on the results of the 2-D Wold decomposition, and an overview on the estimation algorithm for the parameters of the harmonic, evanescent and purely indeterministic components of the texture. In Section 3, we describe the coding methods chosen for these parameters and in Section 4, the results are compared with those of the JPEG algorithm. Finally, in Section 5, we state our conclusions.

2 Estimation of Texture Parameters

The presented texture model is based on the results of the 2-D Wold-type decomposition of discrete, homogeneous random fields, [1], [2]. It can be shown that any 2-D regular and homogeneous random field $\{y(n,m)\}$, can be uniquely represented by the orthogonal decomposition

$$y(n,m) = w(n,m) + v(n,m) \quad . \tag{1}$$

The field $\{w(n,m)\}$ is purely indeterministic and has a unique white innovations driven moving average (MA) representation. The field $\{v(n,m)\}$ is a deterministic random field. It can also be shown that it is possible to define a family of NSHP total-order definitions such that the boundary line of the NSHP is of rational slope. Let α, β be two coprime integers, such that $\alpha \neq 0$. The angle θ of the slope is given by $\tan\ \theta = \beta/\alpha$. (See, for example, Figure 1.) Each of these supports is called *rational non-symmetrical half-plane* (RNSHP). We denote by O the set of all possible RNSHP definitions on the 2-D lattice, (i.e., the set of all NSHP definitions in which the boundary line of the NSHP is of rational slope). The introduction of the family of RNSHP total-ordering definitions results in a corresponding decomposition of the deterministic random field into a countable number of mutually orthogonal fields:

$$v(n,m) = p(n,m) + \sum_{(\alpha,\beta)\in O} e_{(\alpha,\beta)}(n,m)\ . \tag{2}$$

The field $\{p(n,m)\}$ is a *half-plane deterministic* random field, and $\{e_{(\alpha,\beta)}(n,m)\}$ is the evanescent field corresponding to the RNSHP total-ordering definition $(\alpha,\beta)\in O$. The spectral measures of the decomposition components in equations (1) and (2) are mutually singular. The spectral measure of the purely indeterministic component $\{w(n,m)\}$ is absolutely continuous with respect to the Lebesgue measure. The spectral measure of the deterministic field is singular with respect to the Lebesgue measure, and therefore it is concentrated on a set of Lebesgue measure zero in the 2-D spectral domain. Hence, the spectral measure of the half-plane-deterministic field, as well as the spectral measures of the evanescent fields are concentrated on sets of Lebesgue measure zero in the spectral domain.

For practical applications we can exclude singular-continuous spectral measures from the framework of our treatment. Hence, a model for the evanescent field which corresponds to the RNSHP defined by $(\alpha,\beta)\in O$ is given by

$$e_{(\alpha,\beta)}(n,m) = \\ \sum_{i=1}^{I^{(\alpha,\beta)}} s_i^{(\alpha,\beta)}(n\alpha - m\beta)\cos 2\pi\frac{\nu_i^{(\alpha,\beta)}}{\alpha^2+\beta^2}(n\beta+m\alpha) + t_i^{(\alpha,\beta)}(n\alpha-m\beta)\sin 2\pi\frac{\nu_i^{(\alpha,\beta)}}{\alpha^2+\beta^2}(n\beta+m\alpha), \tag{3}$$

where $\{s_i^{(\alpha,\beta)}(n\alpha - m\beta)\}$ and $\{t_i^{(\alpha,\beta)}(n\alpha - m\beta)\}$ are mutually orthogonal 1-D purely indeterministic processes of identical autocorrelation function, and $I^{(\alpha,\beta)}$ is infinite in general. Hence, the "spectral density function" of each evanescent field has the form of a countable sum of 1-D delta functions which are supported on lines of rational slope in the 2-D spectral domain.

One of the half-plane-deterministic field components, which is often found in natural textures, is the harmonic random field

$$h(n,m) \;\;=\;\; \sum_{p=1}^{P}\Big(C_p\cos 2\pi(n\omega_p + m\nu_p) + D_p\sin 2\pi(n\omega_p + m\nu_p)\Big), \tag{4}$$

where the C_p 's and D_p 's are mutually orthogonal random variables, and (ω_p,ν_p) are the spatial frequencies of the pth harmonic. In general, P is infinite. This component generates the 2-D delta functions of the "spectral density". (The 2-D delta functions are singular functions supported on discrete points in the frequency plane.) The parametric modeling of deterministic random fields whose spectral measures are concentrated on curves other than lines of rational slope, or discrete points in the frequency plane, is still an open question to the best of our knowledge. Since such components seem to be of very little practical importance for the texture modeling problem, we assume that the half-plane deterministic field consists only of the harmonic random field.

As stated earlier, the most general model for the purely indeterministic component $w(n, m)$ is the MA model. However, if its spectral density function is strictly positive on the unit bicircle and analytic in some neighborhood of it, a 2-D autoregressive (AR) representation for the purely indeterministic field exists as well. In the following, we assume that the above requirements are satisfied. Hence the purely indeterministic component's *autoregressive model* is given by

$$w(n, m) = -\sum_{(0,0)\prec(k,\ell)} b(k, \ell)w(n - k, m - \ell) + u(n, m), \tag{5}$$

where $\{u(n, m)\}$ is the 2-D white innovations field, whose variance is σ^2. In the practical estimation problem, the model support is assumed finite.

Hence, the observed field $y(n, m)$ is uniquely represented by the model $y(n, m) = w(n, m) + h(n, m) + \sum_{(\alpha,\beta)\in O} e_{(\alpha,\beta)}(n, m)$. In [6] we present a maximum-likelihood (ML) estimation procedure for *jointly* estimating the parameters of the harmonic, evanescent, and purely indeterministic components, of the texture field. However, due to its high computational complexity, we have based the coding algorithm suggested in the present paper, on the suboptimal algorithm, [1], which is computationally much simpler. Contrary to the ML algorithm in [6] which enables us to obtain a complete parameter estimation for the parameters of the evanescent components, in the suboptimal algorithm, the evanescent components are approximated by a linear combination of harmonic components whose frequencies are along a "line" in the sampled frequency domain. This estimation/synthesis algorithm enables the synthesis of purely random, as well as structured textures from the estimated parameters. Although the algorithm is suboptimal, the synthesis results obtained using the estimated parameters were far superior to those obtained by the frequently used AR and Markov random field models. The algorithm in [1] is a sequential, periodogram based estimation algorithm. In the first stage the parameters of the harmonic and evanescent components are estimated and their contribution to the observed realization is removed. Ideally, the obtained residual is the purely indeterministic component of the texture. In a second stage, a 2-D AR model is fitted to the residual by solving the 2-D normal equations system using a 2-D Levinson type algorithm for the reflection coefficients representation of the AR model. The estimation algorithm is summarized in Table 5. For additional details on the estimation algorithm we refer the reader to [1].

3 Coding of the Component Fields

3.1 Introduction

The problem of encoding a texture image is now one of encoding the parameters of the model of the image. The model separates the image into two independent fields, deterministic and indeterministic. As these fields have different characteristics, different coding procedures were adopted for each one. For the indeterministic field, the significant reflection coefficients of the model support region are quantized and the indices of the quantizer levels are conveyed to the decoder through a binary code of rate close to the entropy of these indices. This encoding of the quantizer levels is called an entropy code and the levels are said to be entropy-coded. The locations of the significant reflection coefficients must also be transmitted to inform the decoder where in the support region these coefficients should be placed. For the deterministic field, DFT coefficients of significant magnitude may occur anywhere in the Fourier plane. The real and imaginary parts of these significant coefficients are independently quantized and entropy-coded. The Fourier plane locations of these significant coefficients are transmitted along with the entropy-coded quantized components to reconstruct the field at the decoder. The reconstructed indeterministic and deterministic fields are then added to produce the texture image reconstruction. In the following, we describe the above coding procedures in detail. For the limited purpose of validating this model-based approach, the coding procedures for the deterministic and indeterministic fields were chosen to be simple and non-rigorous. Nevertheless, these procedures brought good quality reconstructions at low rates.

3.2 Coding of the Indeterministic Field

First, consider the indeterministic field determined through the 2-D AR model. Through experiments on a few textures, it was determined that a (6, 6) NSHP support yielded a sufficiently accurate reconstruction for most textures. The eighty-four (84) reflection coefficients for this model's support are extracted directly by a 2-D Levinson type algorithm and are the parameters of choice for encoding. Such a strategy has proved successful in speech coding, due to the fact that the reflection coefficients are bounded by unity in magnitude and that the reconstructions are less sensitive to their perturbations than to similar perturbations of AR coefficients [4].

Two reflection coefficients, the (0, 1) and (1, 0) ones, were deemed sufficiently important to reproduce with eight bit accuracy. All the remaining coefficients were quantized with the same quantizer. This quantizer was empirically designed through calculation of a histogram of the magnitude distribution and setting eight non-uniform decision intervals experimentally with the quantization levels as the midpoint of these intervals. This experimental design was selected, due to the difficulty of relating the quantization error in reflection coefficients to the visual error in the reconstructed texture. Finer quantization with more than eight levels for the magnitude produced little visible advantage.

A reflection coefficient was selected for quantization if its magnitude exceeded 0.05, otherwise, it was set to zero. Since in most circumstances, few of the eighty-four coefficients were quantized, it was decided to transmit their location indices by run-length coding. The run-length symbols describing the lengths of runs of zeros were then entropy coded using an adaptive arithmetic code. The indices of the non-zero quantizer levels were also entropy coded, using a Huffman code with fixed probabilities calculated from the histogram of the reflection coefficients and the quantizer characteristic.

3.3 Coding of Deterministic Field

In the extraction of the deterministic field, the cosinusoidal and sinusoidal components in (3) and (4) are real and imaginary parts of DFT coefficients. If the texture generator were a homogeneous random field having the strong mixing property, (i.e., the autocorrelation function decays sufficiently fast with increasing lag values), the extracted DFT components would be asymptotically Gaussian and statistically independent [5]. In general, the texture random field is non-ergodic and therefore does not possess the strong mixing property. Nevertheless, lacking no better model, we assume Gaussian and quantize, using the same Gaussian quantizer, each component of every (non-redundant) DFT coefficient whose magnitude exceeds a certain threshold. The minimum mean-squared-error uniform quantizer of 32 quantization levels proved to be a satisfactory choice. Theoretically, the entropy of $H = 4.4$ bits/component of this quantizer can be achieved by entropy coding if the Gaussian model is accurate. Otherwise, the Gaussian entropy serves as an upper bound for a given mean square error. Because the outputs of these quantizers were entropy coded with an adaptive arithmetic code, the actual average number of bits per component usually fell below the Gaussian entropy.

The suprathreshold DFT coefficients are relatively few in number, and can occur at any spatial frequency with equal probability. Therefore, the location indicators of the suprathreshold coefficients were encoded with a run-length code, where the runs of zeros signaling subthreshold coefficients were entropy coded with an adaptive arithmetic code.

4 Coding Results

Six original textures of size 64×64 and 256 gray levels were chosen for coding. For any of these textures, the reconstruction quality could not be anticipated by setting beforehand a coding rate. Even if one could set such a rate, the allocation of that rate between the deterministic and indeterministic fields can not be determined, because the usual mathematical criteria such as squared error are of questionable value in evaluating reconstruction quality. Therefore, neither a rate target nor a rate allocation was attempted. Instead, magnitude thresholds were set for the reflection coefficients and the DFT coefficients and whatever total rate resulted from the subsequent quantization of the suprathreshold coefficients and their run-length encoded locations was accepted.

Table 5 presents the number of bits which were used for representation of the indeterministic and deterministic fields for each of the six textures. These numbers are actual bits counted, not entropies. Note that there seems to be no consistent division of the bits between the two fields. The coding rate measured by the average number of bits per pixel varied from 0.08 to 0.23 bits per pixel. In order to assess the quality and efficiency of the reconstructions, we decided to encode the same textures using the JPEG algorithm. Two alternative JPEG coding schemes were tested. The first one uses a uniform quantizer, while in the second the quantization matrix is computed for each texture according to its statistics (and hence must be transmitted as well). Since the uniform quantization yields higher quality reconstructions, at rates that are not much higher than the adaptive scheme, only results using the uniform quantization JPEG coding are presented. The JPEG coding rate is such that it gives a roughly similar visual quality as our reconstructions. The JPEG rates turned out to be much higher than the coding rates with our parametric texture representation as shown in the two right columns of Table 5. The other obvious strategy is to set the JPEG coding rate to be the same rate obtained for the parametric model. However, at such low rates, the JPEG procedure broke down entirely, as the recommended tables supplied with the software could not be adapted to work for such low rates.

We present in Figures 2 and 3 the original, the model based reconstruction, and the JPEG reconstruction for the six textures at the coding rates given in Table 5. Despite the higher rates, the JPEG reconstructions are inferior to the model-based ones, as they display blocking effects and tend to smooth the characteristic roughness of the textures. It is evident that our coding method obtains a high quality rendition of the original and is superior to the JPEG method applied at much higher rates.

5 Conclusions

We have presented a novel approach for coding textured images, whereby the images are decomposed into a 2-D Wold-like representation and the parameters of this representation are encoded. We have proved, through enactment of a coding scheme for the parameters, that both high quality reconstructions and low bit rates are simultaneously achievable for a wide variety of natural textures. We conclude that this new approach is promising and viable for encoding of textured images.

References

[1] 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.*, vol. 41, pp. 2665-2678, Aug. 1993.

[2] J. M. Francos, A. Z. Meiri and B. Porat, "A Wold-Like Decomposition of 2-D Discrete Homogeneous Random Fields," *Annals Appl. Prob.*, vol. 5, pp. 248–260, Feb., 1995.

[3] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second Generation Image Coding Techniques", *Proc. IEEE*, vol. 73, pp. 549-574, April 1985.

[4] B. Friedlander, "Lattice Filters for Adaptive Processing," *Proc. IEEE*, Vol. 70, pp. 829-867, Aug. 1982.

[5] W.A. Pearlman and R. M. Gray, "Source Coding of the Discrete Fourier Transform," *IEEE Trans. Inform. Theory*, Vol. IT-24, pp. 683-692, Nov. 1978.

[6] J. M. Francos, A. Narasimhan and J. W. Woods, "Maximum Likelihood Parameter Estimation of Textures Using a Wold Decomposition Based Model," *IEEE Trans. on Image Processing*, Vol. 4, pp. 1655–1666, Dec. 1995.

List of Tables

List of Figures

1. Test for the existence of deterministic components (i.e., test for the existence of 1-D and 2-D delta functions in the periodogram).
2. Construct a filter which is "1" at the frequencies of the detected 2-D delta functions, and "0" elsewhere.
3. Filter the sample DFT through the filter of step 2, and compute the inverse transform to obtain the estimated harmonic component.
4. Remove the contribution of the estimated harmonic component from the data.
5. Construct a filter which is "1" at the frequencies of the detected 1-D delta functions, and "0" elsewhere.
6. Filter the sample DFT through the filter of step 5, and compute the inverse transform to obtain the estimated evanescent components.
7. Remove the contribution of the estimated evanescent components from the data.

   At this stage the residual image contains no deterministic components.

8. Apply a 2-D Levinson-type algorithm to estimate the 2-D AR model parameters of the residual (the purely indeterministic component).

Table 1: The Estimation Algorithm

Texture Number	*Indeterministic Components*	*Deterministic Components*	*Total (Including 8-bits for Mean)*	*Number of Bits/Pixel*	
				Model Based Coding	*JPEG*
1	74	480	554	0.14	1.16
2	74	592	666	0.16	1.28
3	90	520	610	0.15	1.56
4	60	864	924	0.23	1.24
5	202	136	338	0.08	0.76
6	267	584	851	0.21	1.36

Table 2: Coding results. Total bits and rates required for the experiment textures. The right column gives the required JPEG rate.

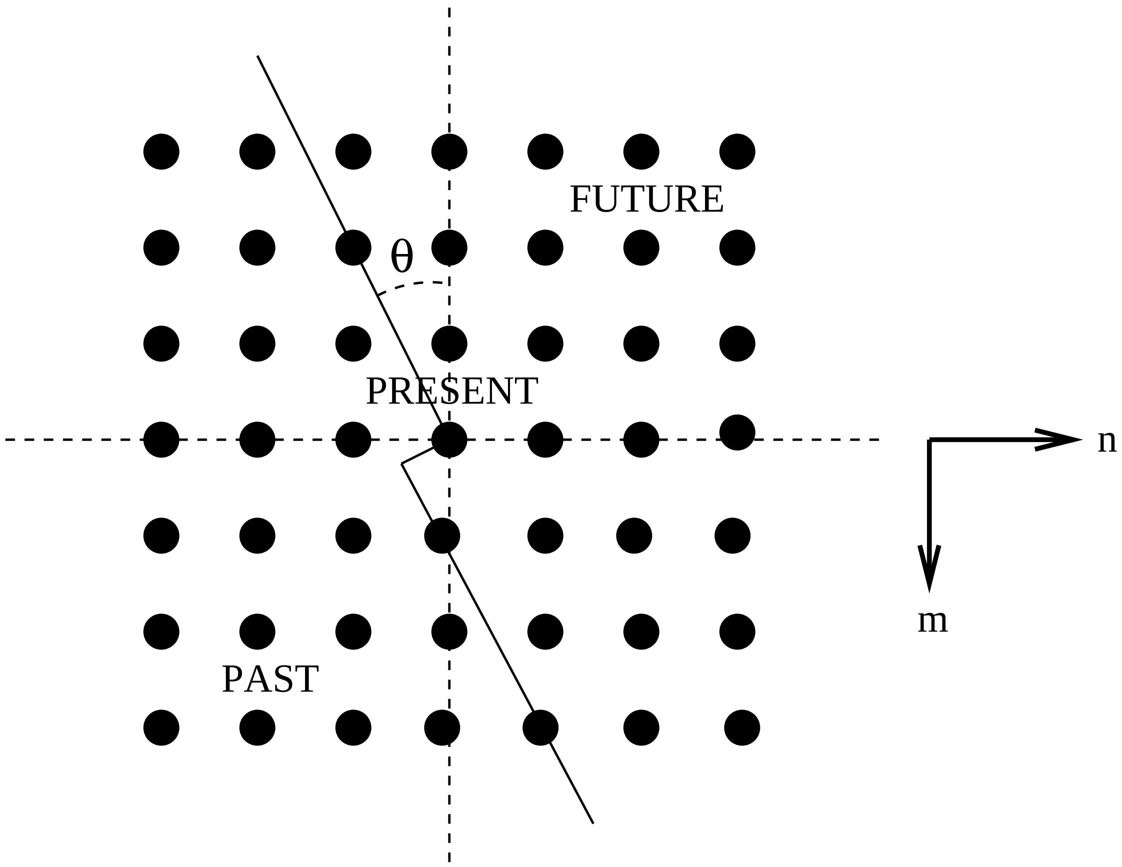

Figure 1: RNSHP total-order definition.

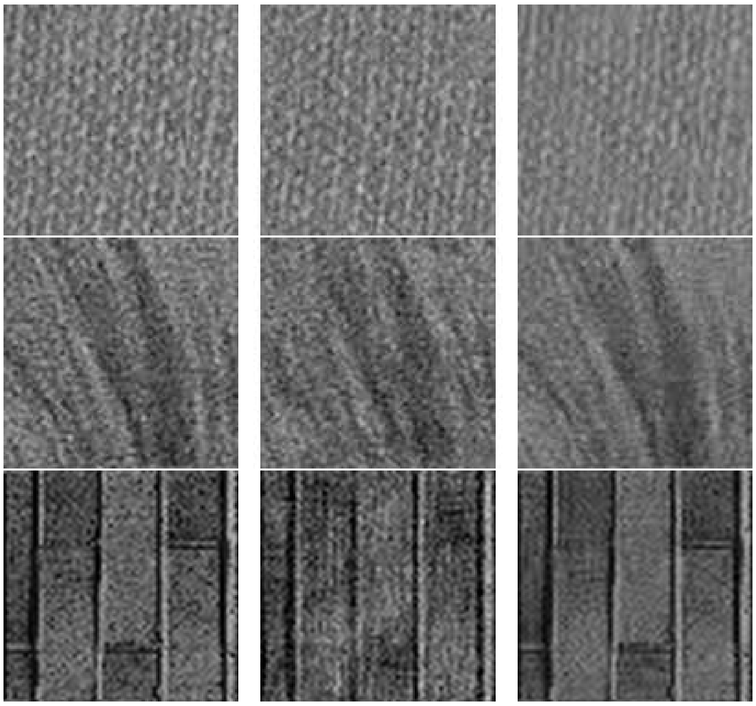

Figure 2: Coding results for three texture fields. From left to right: original, reconstruction from model based coding, and reconstruction from JPEG coding of the textures. From top to bottom: Texture 1 through Texture 3.

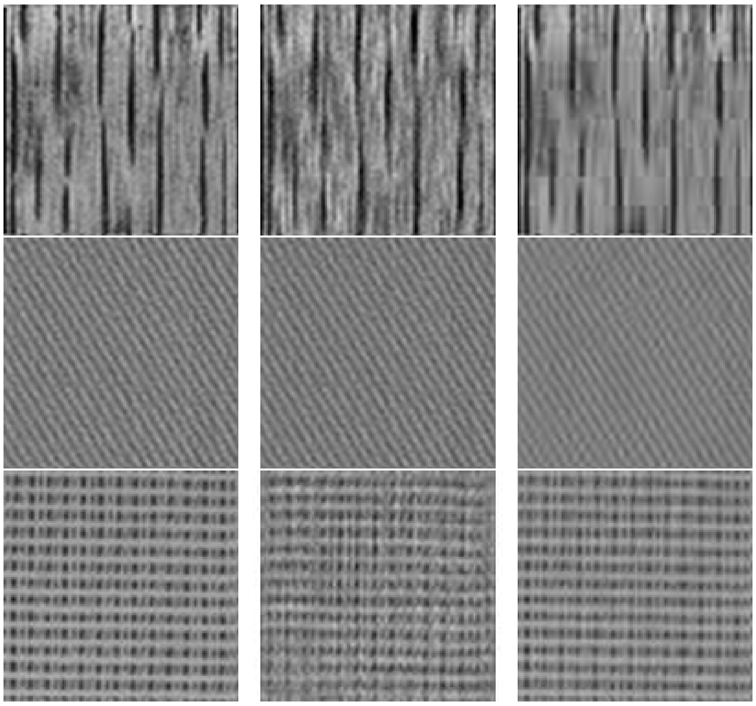

Figure 3: Coding results for three texture fields. From left to right: original, reconstruction from model based coding, and reconstruction from JPEG coding of the textures. From top to bottom: Texture 4 through Texture 6.