16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

ON WIDE-SENSE MARKOV RANDOM FIELDS AND THEIR APPLICATION TO IMPROVED TEXTURE INTERPOLATION

Shira Nemirovsky and Moshe Porat

Electrical Engineering Department, Technion
Haifa, 32000, Israel
phone: + (972) -4- 8294684, fax: + (972) -4-8295757,
nshira@tx.technion.ac.il mp@.ee.technion.ac.il
http://vision.technion.ac.il/nshira http://vision.technion.ac.il/mp

ABSTRACT

Random field models characterize the correlation between neighboring pixels in an image. Specifically, a wide-sense Markov model is obtained by assuming a separable correlation function for a 2D auto-regressive (AR) model. In this work we analyze the effect of sub-sampling on statistical features of an image such as histogram and the autocorrelation function. We show that the Markovian property is preserved for the 2nd-order case (of the wide-sense model) and use this result to prove that, under mild conditions, the histogram of such images is invariant under sub-sampling. Furthermore, we develop relations between the statistics of the image and its sub-sampled version in terms of moments and noise characteristics. Motivated by these results, we propose a new method for texture interpolation, based on orthogonal decomposition. Experiments with natural texture images demonstrate the advantages of the proposed method over presently available interpolation methods.

1. INTRODUCTION

The 2D auto-regressive image model is well known for its various applications in image processing. Specifically, it is useful in the analysis and synthesis of textures [1], as well as in texture classification and segmentation. In its general from, this model may be described by the equation

$$w(m,n) = -\sum_{(k,l)\in\chi_{m,n}} a(k,l)\,w(m-k,n-l) + u(m,n), \quad (1)$$

where $\{w(m,n)\}$ is the modelled image field, $\{a(k,l)\}$ are the AR model parameters, $\{u(m,n)\}$ is zero mean white noise (also termed the *innovation process*) and $\chi_{m,n}$ is some neighborhood of the pixel at location (m,n).

A special case of the model in (1) is obtained by assuming a 2D causal pixel neighborhood [1], i.e.,

$$w(m,n) = \sum_{(i,j)\in D} c_{ij}\,w(m-i,n-j) + u(m,n), \quad (2)$$

where c_{ij} are the model parameters (coefficients) and the neighborhood is given by

$$D = \{(0,1),(1,0),(1,1)\}. \quad (3)$$

Furthermore, assuming a separable correlation function of the form

$$R_{ww}(\alpha,\beta) = C\rho_h^{|\alpha|}\rho_v^{|\beta|}, \quad 0<\rho_h,\rho_v<1,$$

where C is some constant, and applying the Yule-Walker equations, the following straightforward solution for the model is obtained:

$$c_{10}=\rho_h,\ c_{01}=\rho_v,\ c_{11}=-\rho_h\rho_v. \quad (4)$$

The AR model, which is described by Equations (2), (3) and (4) is known as the *wide-sense Markov* model. In this work we focus our interest on this model, and show that when sub-sampling an image that obeys this model, (i) the down sampled image also possesses the wide-sense Markov property; and (ii) the image gray level distribution is preserved. In addition, we formulate expressions that relate the white noise statistics of the image and its sub-sampled version.

The paper is organized as follows. Section 2 presents the invariance of the wide-sense Markov property, as well as the statistical features of the sub-sampled image. Motivated by these results, we propose in Section 3 a method for texture interpolation. Experimental results of the proposed method are shown in Section 4. Section 5 summarizes the main results.

2. WIDE-SENSE MARKOV RANDOM FIELDS AND SUB-SAMPLING OPERATION

2.1 Sub-sampling

Denoting by f a given $2M\times 2M$ square image, we define its low-resolution version f^* as an $M\times M$ image such that

$$f^*(k,l) = f(2k-1,2l-1).$$

The relation between the high and low-resolution versions is illustrated in Figure 1.

2.2 Theoretical Results

In this subsection we show that the wide-sense Markov property is retained under sub-sampling as defined in Section 2.1.

We then proceed to statistical relations between the high and low resolutions. Due to the limited space, only the main steps of the proofs are presented - the detailed derivations are available in [2].

f_{11}	f_{13}	f_{35}
f_{31}	f_{33}	f_{35}
f_{51}	f_{53}	f_{55}

Figure 1(a) –low-resolution image.

f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
f_{21}	f_{22}	f_{23}	f_{24}	f_{25}	f_{26}
f_{31}	f_{32}	f_{33}	f_{34}	f_{35}	f_{36}
f_{41}	f_{42}	f_{43}	f_{44}	f_{45}	f_{46}
f_{51}	f_{52}	f_{53}	f_{54}	f_{55}	f_{56}
f_{61}	f_{62}	f_{63}	f_{64}	f_{65}	f_{66}

Figure 2(b) – high-resolution image.

From now on we denote $f(m,n) \equiv f_{mn}$ for simplicity. We start with our first theorem.

Theorem 1: Let f be a wide-sense Markov image, i.e.

$$f(m,n) = \sum_{(i,j)\in D} c_{ij} f(m-i,n-j) + u(m,n),$$

and let f^* be its low-resolution version. Then, f^* is also wide-sense Markov, and its coefficients are given by

$$c_{01}^* = c_{01}^2,\ c_{10}^* = c_{10}^2,\ c_{11}^* = -c_{10}^* c_{01}^*. \qquad (5)$$

Proof: See Figure 1 for the notation of f, f^*. Using the wide-sense Markov model (2), we can express f_{55} in terms of f_{35}, f_{33}, f_{53} to obtain the pixels' relation in the low-resolution image f^*. Doing so and re-arranging yields the following result:

$$f_{55} = c_{01}^2 f_{53} + c_{10}^2 f_{35} - c_{01}^2 c_{10}^2 f_{33} + N_{total}, \qquad (6)$$

where N_{total} is white noise given by

$$N_{total} = c_{01}c_{10}u_{44} + c_{01}u_{54} + c_{10}u_{45} + u_{55}. \quad (7)$$

It is important to note that due to the wide-sense relation $c_{11} = -c_{10}c_{01}$, the terms in N_{total} that depend on $f_{24}, f_{42}, f_{23}, f_{32}$ vanish, and so N_{total} is orthogonal to f_{33}, f_{35}, f_{53} in the low-resolution image. This property is required in a valid autoregressive process, in which the innovation process should maintain orthogonality in relation to 'past' samples of the process.

Without loss of generality, this calculation holds for all the pixels in the low-resolution image f^*, and thus, by comparing (6) to (2) we obtain the above result. □

Based on Theorem 1, the next 3 corollaries follow.

Corollary 1: Denoting by $\sigma_u^2, \sigma_{u^*}^2$ the variances of the generating white noise of f, f^* respectively, the following relation holds:

$$\sigma_{u^*}^2 = \left(c_{01}^2 c_{10}^2 + c_{01}^2 + c_{10}^2 + 1\right)\sigma_u^2.$$

Corollary 2: Denoting by $R_{ff}^*(\alpha,\beta), R_{ff}(\alpha,\beta)$, the autocorrelation functions (matrices) of the low and high-resolution versions, respectively,

$$R_{ff}^*(\alpha,\beta) = R_{ff}(2\alpha,2\beta), \qquad (8)$$

i.e., the autocorrelation function of the low resolution image is given by sub-sampling the autocorrelation function of the high-resolution image.

Corollary 3: The second-order moments of f, f^* are equal, i.e.,

$$Ef^2 = E\left(f^*\right)^2.$$

For detailed proofs of the above mentioned corollaries, the reader is referred to [2].

If we further assume that the white noise $\{u_{mn}\}$ is normally distributed, we can prove the following result.

Theorem 2: The moments of f^* up to any order are equal to those of f; i.e., the gray level probability distributions of the high-resolution image f and its lower resolution version f^* are equal.

Proof: We first show that the gray level of f is normally distributed. To see that, we observe that each pixel in f may be expressed as a linear combination of samples of the normally distributed white noise u. For example,

$$f_{11} = u_{11}$$
$$f_{22} = c_{10}f_{12} + c_{01}f_{21} + c_{11}f_{11} + u_{22} =$$
$$= c_{10}\left[c_{01}f_{11} + u_{12}\right] + c_{01}\left[c_{10}f_{11} + u_{21}\right] + c_{11}f_{11} + u_{22} =$$
$$= 2c_{10}c_{01}f_{11} - c_{10}c_{01}f_{11} + c_{10}u_{12} + c_{01}u_{21} + u_{22} =$$
$$= c_{10}c_{01}u_{11} + c_{10}u_{12} + c_{01}u_{21} + u_{22}.$$

Now, since u is normally distributed *white* noise, the last result implies that each pixel in f is a linear combination of statistically independent normally distributed random variables, and thus each pixel in f is also a normally distributed random variable. Furthermore, the low-resolution image f^* is also normally distributed, since as we saw in previous sections, f^* is also wide-sense Markov and the same arguments regarding the normal distribution are still valid.

Recalling Corollary 3, according to which the variances of the low and high resolutions are equal, and using the assumption that the expectation value of the noise is zero to conclude that the low and high resolutions also have zero mean, we obtain that the first two moments of f, f^* are

equal. But since f and f^* are normally distributed and have equal first and second moments, their gray level distributions are also equal, which concludes the proof. □

We now proceed to propose a method for texture interpolation. We first describe the texture model used, and then describe our proposed method.

3. A PROPOSED TEXTURE INTERPOLATION METHOD

3.1 Texture Model

We use a texture model proposed by Francos et al. [3-5]. In this model, the texture field is assumed to be a realization of a 2D homogeneous random field. Based on a 2D Wold-like decomposition of homogeneous random field, the texture field is decomposed into a sum of two mutually orthogonal, spatially homogeneous components: (i) the global structural component, which is **deterministic** in the prediction theory sense; and (ii) the purely stochastic component, or a purely **non-deterministic** component. Denoting the homogeneous 2D texture field by $\{y(m,n)\}$, the above decomposition may be expressed as

$$y(m,n) = w(m,n) + v(m,n), \qquad (9)$$

where $w(m,n)$ is the purely non-deterministic component and $v(m,n)$ is a deterministic field. It is important to note that the deterministic component $v(m,n)$ may be further decomposed into

$$v(m,n) = h(m,n) + g(m,n),$$

where $h(m,n)$ is called the *harmonic field*, and $g(m,n)$ is termed the *generalized evanescent field*. Generally speaking, the harmonic field generates the periodic features of the texture field, while the evanescent component generates global directional features in the texture field. The spectral density function of the harmonic field is a sum of 2D Dirac Delta functions, while the spectral density function of the generalized evanescent field is a sum of 1D delta functions that are supported on lines of rational slope in the frequency domain [3-5]. For simplicity, in this paper we assume that $g(m,n) = 0$, that is, we treat the deterministic component $v(m,n)$ as composed of only the harmonic field, i.e., $v(m,n) = h(m,n)$. According to this assumption, we confine our discussion to textures that possess periodic features, and not directional ones. Experiments with Brodatz and VisTex textures [6],[8] reveal that this assumption is not restrictive in practice, and in fact it enables us to simply treat a wide variety of textures.

To complete the description of this model, we note that the non-deterministic component is modelled as a 2D AR model, i.e., it may be expressed as in (1).

3.2 A New Method for Texture Interpolation

The problem of texture interpolation may be stated as follows: given a down-sampled (low resolution) version of the texture, we wish to obtain a high-resolution texture image. This problem is of practical interest, since many natural images are a collage of texture patches, or relatively smooth areas and texture patches.

The framework of our method is as follows. Given a down-sampled (low-resolution) texture, we extract its purely non-deterministic component, evaluate its optimal (low-resolution) auto-regressive model parameters, and use these parameters to generate the high-resolution purely non-deterministic component.

The harmonic component of the down-sampled texture is also extracted, and by filtering and zero padding in the frequency domain is perfectly interpolated to create its up-sampled version. The generated deterministic (i.e., harmonic) and non-deterministic components are then combined to yield the high-resolution output texture.

As noted above, we use the evaluated optimal AR model parameters of the low resolution as an approximation to those of the high resolution. This approximation is motivated by the results of Section 2, that show that for the case of the wide-sense Markov field, generated by white Gaussian noise, the gray level distribution was the same in the high and low resolutions. Noting that the wide-sense Markov field is a special case of the auto-regressive model, and as such may be treated as a special case of an non-deterministic texture component, it is thus motivating to assume that the probability distributions of the low and high versions of the non-deterministic components of general auto-regressive causal model (not necessarily wide-sense Markov) will also be approximately equal and will have similar visual features, and as such – also similar autocorrelations [5]. Thus, it is motivating to approximate the AR model parameters of the high-resolution version by the parameters of the low resolution, and in this manner obtain an interpolation from low to high resolution for the non-deterministic component.

The method we propose for texture interpolation is summarized as follows:

Input: y^* - a $N \times M$ low-resolution texture image.

Output: y - a $2N \times 2M$ high-resolution texture image.

Step 1: Extraction and interpolation of the harmonic component:

1.1 Calculation of the DFT and periodogram (i.e., squared absolute value of the DFT) of y^*.

1.2 Finding the peaks (2D delta functions) of the periodogram from Step 1.1, and creating a frequency domain filter (mask) whose value is 1 at the locations of those peaks and 0 elsewhere.

1.3 Filtering the DFT of y^* with the frequency domain mask from step 1.2, to obtain the DFT of the harmonic component of y^*.

1.4 Zero padding the filtered DFT of y^* from Step 1.3 to obtain the DFT that corresponds to the harmonic component of the high resolution y.

1.5 Applying the inverse DFT on the output of 1.4, to obtain the estimated harmonic component of the high resolution, denoted by $\hat{h}$.

Step 2: Extraction and interpolation of the purely non-deterministic component:

2.1 Filtering the DFT of y^* with a mask that is the negative version of the mask in Step 1.2, to obtain the DFT of the purely non-deterministic component of the low resolution.

2.2 Applying inverse DFT to the output of Step 2.1, to obtain an estimation of the (space domain) purely non-deterministic component in low resolution, denoted by $\hat{w}^*$.

2.3 Choosing the number of AR model parameters (AR order), or in other words, the size of the past support that is used in the model.

2.4 Estimating the optimal AR model parameters according to the minimum innovation process variance criterion, using a least squares solution. These AR model parameters are estimated from the low-resolution version.

2.5 Re-arranging the pixels of $\hat{w}^*$ in a $2N\times 2M$ matrix, such that sub-sampling of this matrix yields again the low resolution $\hat{w}^*$ (as shown in Figure 1).

2.6 Calculating the gray levels at the locations of missing pixels (i.e., that are not populated by the original low-resolution pixels' values) according to the AR model (1), and using the model parameters that were obtained from the low resolution (in Step 2.4). The innovation process u that generates the model is white Gaussian noise with some predetermined variance σ_u^2. Pixels whose support neighborhood exceeds the boundaries of the image are initialized simply by white noise with the same distribution like the innovation process.

2.7 The output of Step 2.6 is the estimated purely non-deterministic component in the high resolution, denoted by $\hat{w}$.

Step 3: Combining the estimated harmonic and purely non-deterministic components from Steps 1 and 2, to obtain the estimated texture image in high resolution, denoted by $\hat{y}$, i.e., $\hat{y}=\hat{h}+\hat{w}$.

4. EXPERIMENTAL RESULTS

In this section experimental results of the proposed method are shown. The textures were taken from the Brodatz album [6] and from the VisTex [8] database. Each original $2N\times 2M$ image texture was sub-sampled such that an $N\times M$ texture was obtained. Then, various interpolation methods were applied to the sampled texture and compared to the original (high-resolution) texture in terms of visual impression and PSNR.

The textures ‘wood grain’ and 'Fabric-0004', and a close-up on their interpolations using various interpolation methods (nearest neighbor, bilinear, bicubic and Spline) and the proposed method are shown in Figures 2, 3. For these textures, a causal 4×4 pixel neighborhood, i.e., the 15 nearest neighbors composing the top-left corner of a given pixel, was found to yield good results.

Additional experiments with other textures yielded similar results [2].

Inspection of Figures 2 and 3 reveals that the proposed texture interpolation method outperforms other interpolation results visually, in terms of preservation of fine details of the textures. For most of them it also obtains higher PSNR values (except for Spline). The average improvement in PSNR was 0.92dB, 0.71dB and 1.37dB for the textures 'Water', 'Wood Grain' and 'Herringbone Weave', respectively. For specific PSNR values see [2].

It is important to note that due to the stochastic nature of the non-deterministic component of the texture, the PSNR is clearly not the best measure for texture reconstruction quality. It is thus desirable to obtain a measure for texture reconstruction. Such a measure is currently developed by the authors of this work, and is based on the decomposition of the textures to their deterministic (harmonic) and non-deterministic components. The comparison of the deterministic components may be obtained by well-known measures, while the comparison of the non-deterministic components may be done according to their second-order statistics. Preliminary results show that such a combined measure is more suitable than PSNR for the evaluation of texture reconstruction quality.

5. SUMMARY

In this work we have shown that, under sub-sampling, the wide-sense Markovian property is retained. In addition, we have proved that under mild conditions, the gray level distribution is invariant. This result motivated us to propose a new method for texture interpolation, based on an orthogonal decomposition texture model [7]. The proposed method performance was tested on Brodatz and VisTex textures. The experimental results demonstrate the advantages of the proposed method over existing interpolation methods in terms of PSNR values (except for Spline) and conservation of fine texture details. In other words, the proposed method achieves less blurry reconstructions.

Future research will focus on extension of the proposed algorithm to color textures, as well as on further analysis of the relations between the high and low resolutions of a general 2D AR model.

REFERENCES

[1] R. Chellapa and R. L. Kashyap, *Texture Synthesis Using 2-D Non-causal Autoregressive Models*, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-33, No. 1, pp. 194-203, 1985.

[2] S. Nemirovsky and M. Porat, *On the Statistics of Wide-Sense Markov Random Fields and its Application to Texture Interpolation*, CCIT Report No. 682, January 2008.
[3] J. M. Francos and A. Z. Meiri, *A 2-D Autoregressive, Finite Support, Causal Model for Texture Analysis and Synthesis*, 1989 International Conference on Acoustics, Speech and Signal Processing.
[4] J. M. Francos, A. Z. Meiri and B. Porat**,** *On a Wold-Like Decomposition of 2-D Discrete Random Fields*, 1990 International Conference on Acoustics, Speech and Signal Processing.
[5] J. M. Francos, A. Z. Meiri and B. Porat, *A Unified Texture Model Based on a 2-D Wold-Like Decomposition,* IEEE Transactions on Signal Processing, Vol. 41, No. 8, August 1993.
[6] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover Publications, New York, 1966.
[7] M. Porat and Y.Y. Zeevi, "*Localized Texture Processing in Vision: Analysis and Synthesis in the Gaborian Space*"**,** IEEE Trans. Biomedical Engineering, Vol. 36, pp. 115-129, 1989.
[8] Website of VisTex database: http://vismod.media.mit.edu/vismod/imagery/VisionTexture/vistex.html

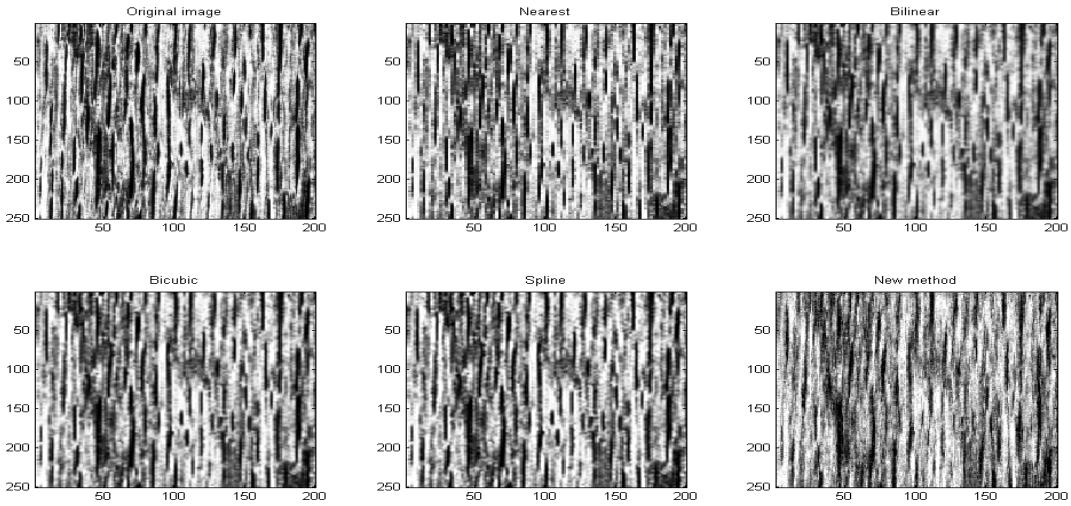

Figure 2 - A close-up on a patch from the original and interpolated textures ('wood grain'). Top left to bottom right: Original texture, Nearest Neighbor, Bilinear, Bicubic, Spline and New Method.

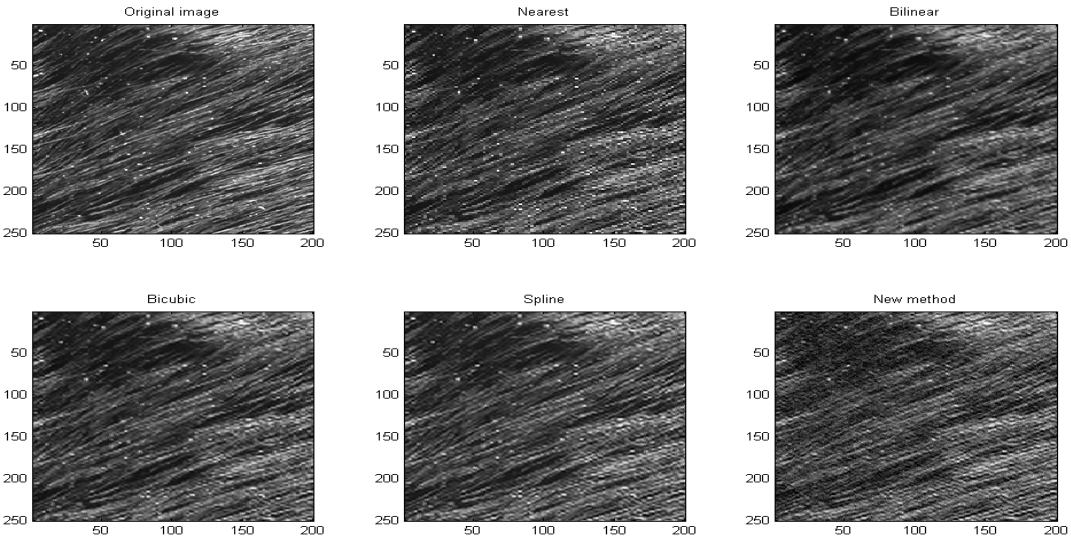

Figure 3 - A close-up on a patch from the original and interpolated textures ('Fabric-0004'). Top left to bottom right: Original texture, Nearest Neighbor, Bilinear, Bicubic, Spline and New Method.