TOWARDS RANDOM FIELD MODELING OF WAVELET STATISTICS

Z. Azimifar *P. Fieguth* *E. Jernigan*

Department of Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada, N2L-3G1

ABSTRACT

This paper investigates the statistical characterization of signals and images in the wavelet domain. In particular, in contrast to common decorrelated-coefficient models, we find that the correlation between wavelet scales can be surprisingly substantial, even across several scales. In this paper we investigate possible choices of statistical-interaction models. One efficient and fast strategy which describes the wavelet-based statistical correlations is illustrated. Finally, the effectiveness of the proposed tool towards an efficient hierarchical MRF modeling of within-scale neighborhoods and across-scale dependencies will be demonstrated.

1. INTRODUCTION

This paper presents a fast and efficient strategy for modeling the statistical dependencies of 2-D wavelet coefficients. Specifically, we are interested in studying the Markovian nature of wavelet coefficient interactions, both within and across scales. We propose multiscale and Markov random field (MRF) models for the wavelet correlation structures.

Our motivation is model-based statistical image processing, which requires some probabilistic description of the underlying image characteristics. Because of the complexity of spatial behaviour and pixel interactions, the raw statistics of pixels are extremely complicated and inconvenient to specify. It is much more convenient to consider describing the statistics of a transformed image, where the transform is chosen to simplify or decorrelate, as much as possible, the starting statistics, analogous to the preconditioning of complicated linear system problems. The popularity of the wavelet transform (WT) stems from its effectiveness in this task: many operations, such as interpolation, estimation, compression, and denoising are simplified in the wavelet domain, because of its energy compaction and decorrelative properties [1, 2].

A conspicuously common assumption is that the WT is a perfect whitener, such that all of the wavelet coefficients are independent, and ideally Gaussian. There is, however, a growing recognition that neither of these assumptions are accurate, nor even adequate for many image processing needs. There have been several recent efforts to study the wavelet statistics; most of these focus on the individual (marginal) statistics, only very little literature is present on the interrelationship (joint) statistics:

1. Marginal Models:
   (a) Non-Gaussian, i.e., heavy tail distribution [3],
   (b) Mixture of Gaussians [3],
   (c) Generalized Gaussian distribution [2],
   (d) Bessel functions [4].
2. Joint Models:
   Hidden Markov tree models [1].

In virtually all marginal models, currently being used in wavelet shrinkage [2], the coefficients are treated individually and as independent, i.e., only the diagonal elements of wavelet based covariance matrix are considered. This approach, however, is not optimal in a sense that WT is not a perfect whitening process.

The latter approach, however, examines the joint statistics of coefficients. Normally an assumption is present that the correlation between coefficients does not exceed the parent-child dependencies, e.g. given the state of its parent, a child is decoupled from the entire wavelet tree [1].

It is difficult to study both aspects simultaneously: that is, the development of non-Gaussian joint models with non-trivial neighborhood. The study of independent non-Gaussian models has been thorough; the complementary

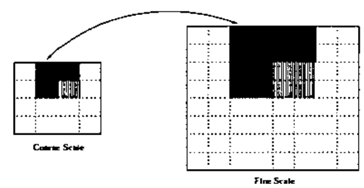


Fig. 1. Illustration of a coefficient in the fine scale (shown by •), whose spatial neighbors come from different parents in the coarser scale.

The support of the Natural Science & Engineering Research Council of Canada and Communications & Information Technology Ontario are acknowledged.

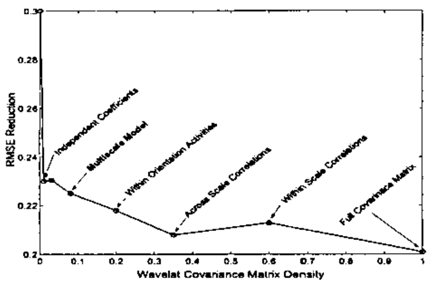


Fig. 2. RMSE noise reduction as a function of covariance density. It is evident how a tiny fraction of coefficients already provides the majority of the improvement.

study, the development of Gaussian joint models, is the focus of this paper. The goal, of course, is the ultimate merging of the two fields. However for the purpose of this paper, we are willing to limit ourselves to simplifying marginal assumptions (Gaussianity) which we know to be incorrect, but which allow us to undertake a correspondingly more sophisticated study of joint models.

In previous work [5], we proposed a MS model, which described the wavelets coefficients as a first-order Markov process in scale. The virtue of the model is its ability to capture the most significant statistical information between tree parents and children, however the interrelationship of pixels within a scale is only implicit, and very limited. In this paper we extend our work to the proper modeling of statistical dependencies on spatial neighbors.

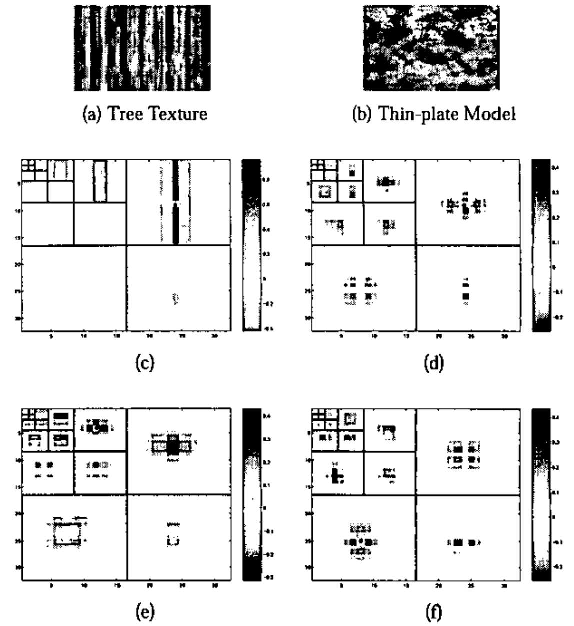
2. WAVELET SPATIAL NEIGHBORHOODS

Figure 1 illustrates the arrangement of a typical coefficient at a fine scale (right panel) and the corresponding parents (left) at the coarser scale. It is immediately obvious that first-order neighbors of a pixel are not necessarily spawned from the same parent.

The purpose of these illustrations is to point toward an important issue: although two coefficients may be spatially close, they can be located on distantly separated branches of the wavelet tree. Consequently a standard wavelet quadtree, modeling only parent-child relationships, will only poorly represent spatial interrelationships, in those cases where they are found to be significant.

3. WAVELET NEIGHBORHOOD MODELING

In order to study the exact correlations between wavelet coefficients we considered a variety of prior models based on Gaussian Markov random field (GMRF) covariance structures. The chosen priors, shown in Figure 3(a),(b) are the



(a) Tree Texture (b) Thin-plate Model

(c) (d)

(e) (f)

Fig. 3. The correlation structure of some wavelet coefficient, marked by •, with all other coefficients. (a)-(b): Two typical GMRF textures, (c): Correlation structure of a horizontal coefficient for the tree texture of (a), (d)-(f) Correlation structures for the thin-plate model of (b), for three different choices of subband and scale.

tree-bark and thin-plate models. They are spatially stationary, an assumption for convenience only and is not fundamental to our analysis.

The selected covariance structure Σ_f is transformed into the wavelet domain by computing the 2-D wavelet transform W, containing all translated and dilated versions of the selected wavelet basis functions:

$$\Sigma_{Wf} = W \Sigma_f W^T \qquad (1)$$

where we have restricted our attention to the set of Daubechies basis functions.

In past work [5] we have already examined a variety of possible wavelet correlation structures to consider: from complete independence (diagonal elements only) to full dependency (entire covariance matrix preserved), with six other intermediate variations. Each variation clearly will differ in its complexity (matrix density) and statistical accuracy, a comparison which is shown in Figure 2 for the image denoising problem.

It is clear that MSE performance is not even necessarily monotonic in matrix density! Furthermore, it is also clear that the vast bulk of the benefit is to be held from relatively few coefficients. Our goal in this paper is a study of probabilistic models which describe the wavelet random field with a small fraction of coefficients, but which accurately absorb each pixel's dependency on the rest of the wavelet tree.

3.1. Hierarchical Representation of Correlations

Our goal is to obtain a clear neighborhood structure, such as for a Markov random field, which is capable of describing the statistical interactions of wavelet coefficients.

Answering this question is challenging because of the issues raised in Figure 1: the tree-relationship between a pixel and its spatial neighbors is pixel dependent. So notions of stationarity, obvious in the spatial domain, become subtle (or completely invalid) in the wavelet domain. In short, do we need to specify a different neighborhood structure for *every* wavelet coefficient (since each coefficient occupies a unique position on the tree), or perhaps one structure for all of the "lower-left" children of parents and another for "upper-right" etc., or is there some degree of uniformity that applies?

We have chosen to begin by studying the problem visually, and without any particular spatial assumptions. As shown in Figure 3(c)-(f), we have devised a tool which utilizes the traditional 2-D WT structure to display the correlation between any specified coefficient and all other coefficients on the entire wavelet tree. Figure 3(c) shows the correlation of a typical horizontal coefficient (indicated by •) of the tree texture, exhibiting a strong vertical correlation both within and across scales. Similarly Figure 3(d)-(f) display correlation structures for the thin-plate model and for three different choices of subband and scale.

In these illustrations the coefficient interactions show a clear preference to locality, as must be expected. This locality increases toward finer scales, which supports the persistency property of wavelet coefficients [1]. The local neighborhood definition for any given pixel does not confine to the pixel's subband: it extends to dependencies across directions and resolutions. Besides the long range across scale correlations, every typical coefficient exhibits strong correlation with its immediate neighbors both within subband and scale.

It is well-known that the WT is a sparse representation of original data by a few *essential* coefficient values [1]. To take this fact into account as a sufficient condition, these empirical evaluations have been extended to dependency structure of those significant coefficients.

3.2. Significant Coefficients Correlation

Although intriguing, it is not remotely obvious that the correlation coefficients plotted in Figure 3 necessarily quantitatively correspond to importance in considering coefficient interactions. That is, can we more objectively quantify what it means for some correlation to be important or significant?

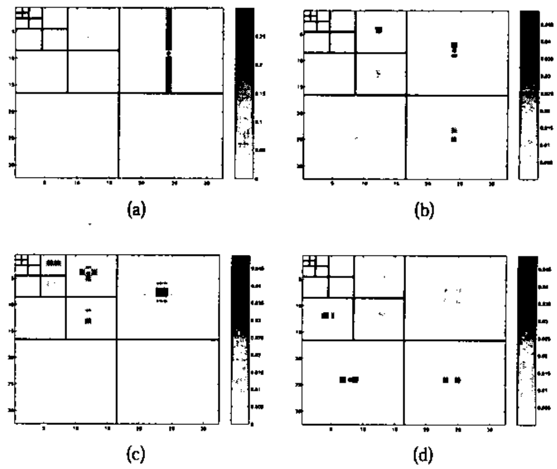


Fig. 4. The same panels as in Figure 3(c)-(f), but now plotting the *significance* of each interrelationship, rather than its correlation. Significance is measured in terms of MSE improvement $\hat{\lambda}$ from (5).

For small test problems the wavelet-based covariance Σ_{Wf} can be determined exactly. Suppose two coefficients c_1, c_2 are observed in the presence of noise:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix}, \quad \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} \sim N\left(0, \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\right) \qquad (2)$$

Under the standard independence assumption, if only the coefficient variances are kept from the full covariance, then their estimation error is given by

$$\tilde{P}_1 = \left(\begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}^{-1} + \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1} \right)^{-1} \qquad (3)$$

On the other hand, if we model the two coefficients with their correct correlation, the estimation error proceeds as

$$\tilde{P}_2 = \left(\begin{bmatrix} \sigma_x^2 & \lambda_{x,y} \\ \lambda_{y,x} & \sigma_y^2 \end{bmatrix}^{-1} + \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}^{-1} \right)^{-1} \qquad (4)$$

In other words, the importance of this particular correlation can be quantified as the degree to which it affects the accuracy of the estimation, which although related to the correlation coefficient, is not proportional to it. We define the significance to be the difference of the total MSE under the two approaches:

$$\hat{\lambda}_{x,y} = tr(\tilde{P}_1) - tr(\tilde{P}_2) \qquad (5)$$

Figure 4(a)-(d) shows the significance of correlations for the corresponding coefficients displayed in Figure 3(c)-(f). It is evident from these diagrams that within scale dependency range reduces to shorter locality, but across scale activities still present up to several scales. The computation of significant covariances, thus, confirms that the well-structured coefficients dependencies to be hierarchical.

(a) (b)

Fig. 5. Motivated by the significance maps, such as those shown in Figure 4, the above panels show one possible derived MRF neighbourhood structure for (a) a vertical coefficient and (b) a diagonal coefficient. In both cases, the neighborhood of a single pixel (•) is shown.

3.3. Proposed Model

According to our achievements of statistical dependencies between the wavelet coefficients we propose to model the wavelet coefficients not as independent, but as governed by MRF stochastic processes. Since correlations are present both within and across scales, a random field model for the wavelet coefficients with itself needs to be hierarchical. To develop the MRF model, we start with a simple structure as shown in Figure 5. Figure 5(a) displays the proposed model for a typical vertical component (marked as •), which is formed based on the following hierarchy:

1. The first-order neighbors in the vertical subband,
2. The corresponding second-order neighbors from the horizontal subband,
3. The corresponding first right and first left neighbors from the diagonal subband,
4. Similar neighborhood structure across scales with reduced local activity towards coarser resolutions.

By symmetry, similar MRF model is proposed for horizontally aligned coefficients. Figure 5(b) also displays the locality considered for a diagonal coefficient. Diagonal coefficients are expected to be less correlated, but across subband dependencies are clearly observed.

To obtain the actual Markov random field model coefficients, parameter estimation needs to be done. Figure 6 displays the model parameters calculated for a simple first-order within- and across-scale neighborhood site for a thin-plate MRF 5-level wavelet transformed. The estimated parameters are scale dependent. They increase the MRF model strength as coefficients dependencies increase from coarse to fine resolutions. The within-scale correlations of horizontal and vertical subbands are symmetrically identical and are stronger than those of diagonal subband.

4. CONCLUSIONS

A thorough 2-D wavelet statistical study has been presented in this paper. An examination of the coefficient correlations, within or across scales, revealed the fact the there exists a clear MRF model governing these local dependencies. The proposed MRF model exhibits a sparse neighborhood structure which absorbs correlation of the given coefficient with the rest of the wavelet tree.

Following the modeling stage, there are two primary ongoing research directions: (1) evaluating the model accuracy by comparing it with the existing thresholding methods, in MMSE sense, (2) devising an estimation or denoising algorithm, which takes into account this MRF model and results in optimum error and low computational cost.

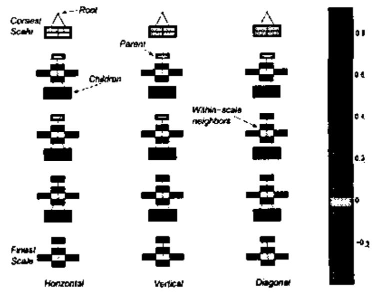


Fig. 6. The MRF model parameters calculated for a first-order within- and across-scale neighborhood site. The parameters are scale dependent and the hierarchical correlation increases from coarser to finer scales. The zero-correlation at the root scale is presumably because of lack of data.

5. REFERENCES

[1] Romberg K., Choi H., and Baraniuk R., "Bayesian tree-structured image modeling using wavelet-domain hidden markov models," *IEEE trans. on IP*, 10(7), 2001.

[2] Portilla J. and Simoncelli E., "Image denoising via adjustment of wavelet coefficient magnitude correlation," *Proceedings of the 7th ICIP*, 2000.

[3] Mumford D. Huang J., "Statistics of natural images and models," *Proceedings of the CVPR'99*, 1999.

[4] Srivastava A., Liu X., and Grenander U., "Analytical models for reduced spectral representations of images," *Proceedings of the 8th ICIP*, 2001.

[5] Azimifar Z., Fieguth P., and Jernigan E., "Hierarchical multiscale modeling of wavelet coefficients correlations," *Proceedings of the 9th SSPR*, 2002.

[6] Azimifar Z., Fieguth P., and Jernigan E., "Wavelet shrinkage with correlated wavelet coefficients," *Proceedings of the 8th ICIP*, 2001.

[7] Crouse M., Nowak R., and Baraniuk R., "Wavelet-based statistical signal processing using hidden markov models," *IEEE trans. on SP*, 46(4), 1998.