ISSN: 2321-7782 (Online)
Volume 3, Issue 5, May 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

Dynamic Statistical Image Denoising

Nisam P M
Dept. Electronics and Communication
College Of Engineering Cherthala
India

Abstract: From the perspective of the Bayesian approach, the denoising problem is essentially a prior probability modeling and estimation task. The proposed approach exploits a hidden Bayesian network, constructed from wavelet coefficients, to model the prior probability of the original image. Then, use the belief propagation (BP) algorithm, which estimates a coefficient based on all the coefficients of an image, as the maximum-a-posterior (MAP) estimator to derive the denoised wavelet coefficients. Also try whether the network is a spanning tree, the standard BP algorithm can perform MAP estimation efficiently. In the experiment results demonstrate that, in terms of the peak-signal-to-noise-ratio and perceptual quality, the proposed approach outperforms state-of-the-art algorithms on several images, particularly in the textured regions, with various amounts of white Gaussian noise.

Keywords: Bayesian network, image denoising, wavelet transform

I. INTRODUCTION

The class of natural images that we encounter in daily life is only a small subset of the set of all possible images. This subset is called an image manifold. Digital image processing applications are becoming increasingly important and they all start with a mathematical representation of the image. In Bayesian restoration methods, the image manifold is encoded in the form of prior knowledge that express the probabilities that specified combinations of pixel intensities can be experiential in an image. Because image spaces are high-dimensional, one often isolates the manifolds by decomposing images into their components and by fitting probabilistic models on it. The construction of a Bayesian network involves prior knowledge of the probability relationships between the variables of interest. Learning approaches are widely used to construct Bayesian networks that best represent the joint probabilities of training data. In practice, an optimization process based on a heuristic search technique is used to find the best structure over the space of all possible networks. However, the approach is computationally intractable because it must explore several combinations of dependent variables to derive an optimal Bayesian network.

The difficulty is resolved in this paper by representing the data in wavelet domains and restricting the space of possible networks by using certain techniques, such as the maximal weighted spanning tree. Three wavelet properties - sparsity, cluster, and motion - can be oppressed to reduce the computational complexity of learning a Bayesian network. During the last decades, multi resolution image representations, like wavelets, have received much attention for this purpose, due to their sparseness which manifests in highly non-Gaussian statistics for wavelet coefficients. Marginal histograms of wavelet coefficients are typically leptokurtotic and have heavy tails. In literature, many wavelet-based image denoising methods have arisen exploiting this property, and are often based on simple and elegant shrinkage rules. In addition, joint histograms of wavelet coefficients have been studied in. Taking advantage of correlations between wavelet coefficients either across space, scale or orientation, additional improvement in denoising performance is obtained. The Gaussian Scale Mixture (GSM) model, in which clusters of coefficients are modeled as the artifact of a Gaussian random vector and a positive scaling variable, has been shown to produce outcome that are appreciably better than marginal models. Image restoration aims to construct an estimate sharing the significant features still present in the degraded image, but with the artifacts censored.

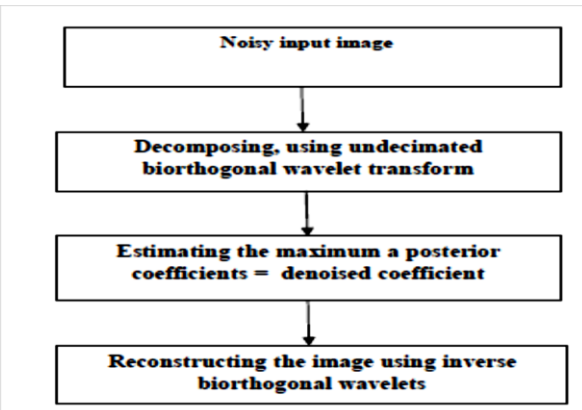
II. FORMULATION OF PROBLEM

For the construction, we use image patches to take into account complex spatial interactions in images. In contrast to exemplar-based approaches for image modeling. An unsupervised method that uses no collection of image patches and no computational intensive training algorithms. Our adaptive smoothing works in the joint spatial-range domain as the nonlocal means filter but have a more powerful adaptation to the local structure of the data since the size of windows and control parameters are estimated from local image statistics. We create the presentation of the proposed denoising algorithm by first introducing how sparsity and redundancy are brought to exploit. We do that via the beginning of the Sparse land reproduction Once this is set, we will talk about how local management on image patches turns into a global prior in a Bayesian rebuilding framework. The second part of the paper attempts to further validate recent claims that lossy compression can be used for denoising. The Bayes Shrink threshold can aid in the parameter selection of a coder designed with the intention of denoising, and thus achieving concurrent denoising and looseness. Specifically, the zero-zone in the quantization step of compression is analogous to the threshold value in the thresholding function. The left behind coder design parameters are selected based on a criterion derived from Rissanens minimum description length (MDL) theory. Experiments show that this compression method does indeed remove noise extensively, especially for great noise power. Although it introduces quantization noise and should be used only if bit rate were an additional concern to denoising. In meticulous, the transform-domain denoising methods normally assume that the true signal can be well approximated by a linear combination of few basis elements. That is, the signal is sparsely represent in the transform domain. Thus, by preserving the few high-magnitude transform coefficients that convey typically the accurate-signal energy and discarding the rest which are mainly due to noise, the correct signal can be successfully estimated. The sparsity of the representation depends on both the transform and the true-signals properties. The multi resolution transforms can achieve first-class sparsity for spatially localized fine points, for instance edges and singularities. When this prior-learning plan is combined with sparsity and redundancy, it is the glossary to be used that we target as the learned set of parameters.

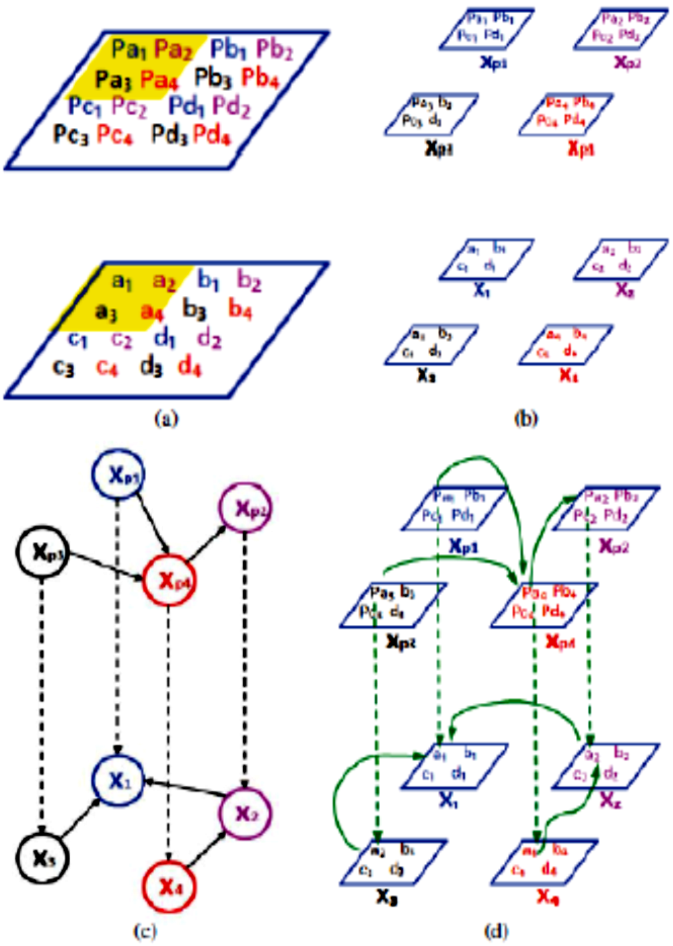
III. SOLUTION OF PROBLEM

(a) the proposed wavelet bayes adaptive image denoising

In this approach, the denoising problem is basically a prior probability modeling and estimation task. The proposed approach that exploits a hidden Bayesian system, constructed from wavelet coefficients, to model the previous probability of the original image. Then, we use the belief propagation (BP) method, which estimates a coefficient based on all the coefficients of an image, as the maximum-a-posterior (MAP) estimator to develop the denoised wavelet coefficients. Also explain that if the network is a spanning tree, the standard BP algorithm can execute MAP estimation competently. Our experiment results demonstrate that, in conditions of the peak-signal-to-noise-ratio and perceptual quality, the projected approach outperforms state-of-the-art algorithms on various images, particularly in the textured regions, with various amounts of white Gaussian noise.



Basic flow chart



Constructing a multilayer hidden network. (a) Two sub bands, with the coarser sub band on top. (b) Procedure creates two wavelet patches, each of which is associated with a sub graph. Sub band coefficients are assigned to the nodes, as specified in Fig. 1(c). (c) Nodes in the two-layer network are linked by intra-scale (solid) arcs and inter-scale (dashed) arcs. (d) To derive the prior probability

Wavelet transform: The Fourier transform is a useful tool to analyze the frequency components of the signal. However, if we take the Fourier transform over the whole time axis, we cannot tell at what instant a particular frequency rises. Short-time Fourier transform (STFT) uses a sliding window spectrogram, which gives the information of both time and frequency. But still another problem exists: The length of window limits the resolution in frequency. Wavelet transform seems to be a solution to the problem above. Wavelet transforms are based on small wavelets with limited duration. The translated version wavelets locate where we concern. Whereas the scaled-version wavelets allow us to analyze the signal in different scale.

Most of the signals in practice, are TIME-DOMAIN signals in their raw format. That is, whatever that signal is measuring, is a function of time. In other words, when we plot the signal one of the axes is time (independent variable), and the other (dependent variable) is usually the amplitude. When we plot time-domain signals, we obtain a time-amplitude representation of the signal. This representation is not always the best representation of the signal for most signal processing related applications. In many cases, the most distinguished information is hidden in the frequency content of the signal. The frequency SPECTRUM of a signal is basically the frequency components (spectral components) of that signal. The frequency spectrum of a signal shows what frequencies exist in the signal.

(b) Multi resolution analysis

The basis functions for all the nested subspaces of this structure maybe generated from the translates and dilates of a single function called the scaling function. Different scaled version can see different frequency resolutions. Combined with the two properties, we can construct a basis from the scaling function and wavelet function with two parameters: scaling and translating, eqtn.

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$$

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$$

Where j is the parameter about dilation, or the visibility in frequency and k is the parameter about the position. In practice, we may want to see the whole data with "desired" resolution.

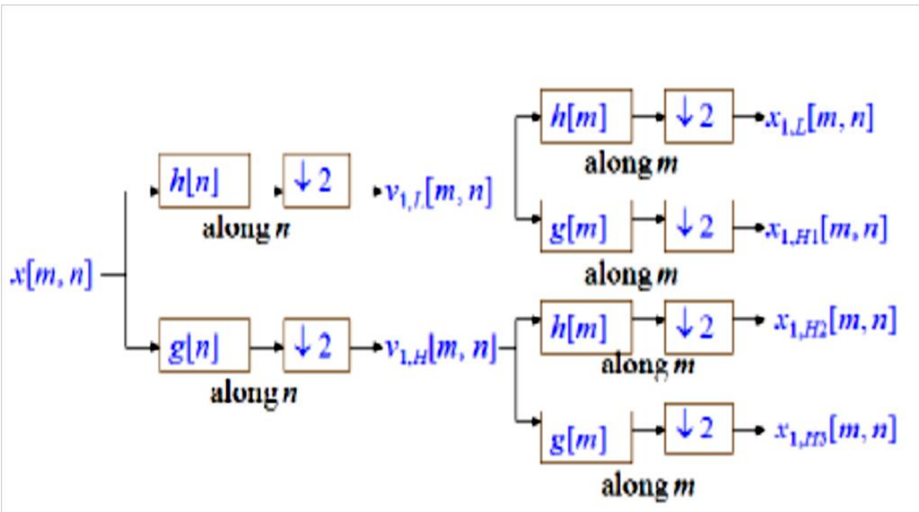
(c) 2D Wavelet transform

The transformed coefficient becomes two variable functions so as the 2D wavelet transform. In the scaling and wavelet function are two variable functions. The scaled and translated basis functions are defined in equation.

$$\phi_{j,m,n}(X,Y) = 2^j\phi(2^j x - m, 2^j y - n)$$

$$\psi^i_{j,m,n}(X,Y) = 2^j\psi^i(2^j x - m, 2^j y - n)$$

There are three different wavelet functions, H(x; y), V (x; y) and D(x; y). Conceptually, the scaling function is the low frequency component of the previous scaling function in 2 dimensions. Therefore, there is one 2D scaling function. However, the wavelet function is related to the order to apply the filters. If the wavelet function is separable, i.e. f(x; y) = f1(x)f2(y). These functions can be easily rewritten in equation.



Filter bank structure.

(d) Un-decimated wavelet transforms (UDWT)

The best two known and most utilized of these are the conventional (decimated) discrete wavelet transform or DWT and the un- decimated discrete wavelet transform or UDWT. A quick note about terminology is in order here. The DWT is actually more complicated than the UDWT and an argument could be made that it should be made that it should be called the decimated or down sampled discrete wavelet transform. In UDWT the number of sample coefficient in the transformed demine is equal to the number of sample coefficients in the input.

(e) Biorthogonal wavelet transform

It is well known that bases that span a space do not have to be orthogonal. In order to gain greater flexibility in the construction of wavelet bases, the orthogonality condition is relaxed allowing semi-orthogonal, biorthogonal or non-orthogonal wavelet bases Biorthogonal Wavelets are families of compactly supported symmetric wavelets. The symmetry of the filter coefficients is often desirable since it results in linear phase of the transfer function. In the biorthogonal case, rather than having one scaling and wavelet function, there are two scaling functions that may generate different multi resolution analysis, and accordingly two different wavelet functions.

(f) Bayes' rule

Let B1 and B2 be disjoint events whose union is Let also A be another event. We can writen in following equations

$$P[B_1/A] = \frac{P(B_1 \cap A)}{P(A)} = \frac{P(A/B_1)P(B_1)}{P(A)}$$

$$P(A) = P(B_1 \cap A) + P(B_2 \cap A) = P(A/B_1)P(B_1) + P(A/B_2)P(B_2)$$

$$P(B_1/A) = \frac{P(A/B_1)P(B_1)}{P(A/B_1)P(B_1) + P(A/B_2)P(B_2}$$

This formula extends to a finite number of events Bn that partition. The result is known as Bayes rule. Think of the Bn as possible of some effect A. You know the prior probabilities P (Bn) of the causes and also the probability that each cause provokes the effect A. The formula tells you how to calculate the probability that a given cause provoked the observed effect.

(g) Wavelet transform and probability for denoising

Complex phenomena usually involve a large number of hidden variables and data sources. Graphical models provide a unifying framework for modeling the probability distributions of such phenomena by decomposing joint probability distributions into a set of local constraints and dependencies. After formulating a problem as a graphical model, a wide range of statistical learning and inference algorithms can be applied directly to derive a solution. Bayesian networks are probably the most popular type of (directed) graphical model. In this paper, our objective is to construct a Bayesian network from a single image for denoising purposes. To do this, we need to overcome two difficulties: 1) constructing a Bayesian network is computationally inefficient and 2) the data over-fitting problem, which exaggerates minor fluctuations in the input data.

The construction of a Bayesian network involves prior knowledge of the probability relationships between the variables of interest. Learning approaches are widely used to construct Bayesian networks that best represent the joint probabilities of training data. In practice, an optimization process based on a heuristic search technique is used to find the best structure over the space of all possible networks. However, the approach is computationally intractable because it must explore several combinations of dependent variables to derive an optimal Bayesian network. The difficulty is resolved in this paper by representing the data in wavelet domains and restricting the space of possible networks by using certain techniques, such as the maximal weighted spanning tree (MWST). Three wavelet properties sparsity, clustering, and persistence - can be exploited to reduce the computational complexity of learning a Bayesian network. First, the wavelet transform of a natural image tends to be sparse with large coefficients at the edges. The sparsity reduces the number of variables required to construct a graph. Second, the adjacent wavelet coefficients tend to have similar values as a cluster. Third, wavelet coefficients at the same location and orientation tend to be positively correlated in adjacent scales.

The Bayesian approach is also widely used to resolve the image denoising problem. The Bayesian formula indicates that the denoising problem is essentially a prior probability modeling and estimation task. If y = x + n, where n is white Gaussian noise with known variance, then the Bayesian formula is P(x—y) = P(y—x) P(x), where P(y—x) = Pn(y-x) is the noise probability. The maximum-a-posterior (MAP) solution of P(x—y) is determined by the priori probability P(x). The structure of image priors is usually modeled by Markov Random Fields (MRFs), where the probability of a pixel depends solely on the joint probability of the pixels nin its neighborhood, the probability distribution of an MRF is the Gibbs distribution whose energy function is the sum of the potential functions defined on the cliques (i.e., maximal complete subgraphs) in image neighborhoods.

(h) Data-adaptive hidden network approach

In this section, we describe the proposed framework for constructing a data-adaptive graph structure and formulating the prior probability of the original image in the wavelet domain. We also explain how the graph structure is used with the BP algorithm to derive the MAP solution. We assume that the wavelet coefficients of the original image represent a realization of a hidden graph. To construct the graph, we first create a matrix of random variables (a wavelet patch) for each subband (indexed by orientation and scale), and assign the subband coefficients as the observed data of the random variables. Then, a subgraph (network) is constructed from the wavelet patch as follows. First, we associate each random variable with a node; hence, the coefficients assigned to the random variable can be regarded as nodes observed data. Second, the arcs (directed edges) in the subgraph are derived according to a data dependence measurement between the observed data in any pair of nodes In this way, we can construct a one-layer network from the wavelet subband and associate the subbands coefficients with the nodes in the network. Finally, the one-layer networks of adjacent subbands can be linked by inter-scale arcs, from coarser scales to finer scales, to form a multi-layer network structure. Two problems may arise during the above construction procedure: 1) the coefficient and wavelet patch association problem, which involves associating subband coefficients with a wavelet patch; and 2) the graph selection problem, i.e., determining the type of graph to construct. To solve the first problem, we propose the following heuristic procedure.

Assume that the wavelet patch is a matrix of m*m random variables. Let the size of a subband be N*N and let m divide N. We partition the subband into (Nm)2 rectangular blocks, each of which contains m*m coefficients. Then, the coefficient at ocation (i, j) in each block is assigned as a realization of the random variable at location (i, j) in the wavelet patch. Thus, each random variable has (Nm)2 sampled observations.

(i) Constructing wavelet bayesian networks

A Bayesian network, denoted as B = (V, E, P), comprises a set of random variables and their conditional dependencies represented by a directed acyclic graph in which the nodes represent the elements in V. Each edge element in E takes the form of a directed arc x y, where x and y V. The likelihood p(y — x) P of an edge x y E is the conditional probability of observing y given that x exists. We call the Bayesian networks constructed in wavelet domains wavelet Bayesian networks (WBNs). Our primary objective is to construct a WBN from the undecimated discrete wavelet transform (DWT) of a single image. Initially, wavelet decomposition of an image F yields three images of wavelet coefficients with horizontal, vertical, and diagonal orientations respectively, and one approximate image of F.

Then, at the next scale, the approximated image is further decomposed to obtain three images of the wavelet coefficients and one coarser approximate image of F. Let ,F(u, v), ,F(u, v), and ,F(u, v) denote, respectively, the horizontal, vertical, and diagonal images of the wavelet coefficients at scale 2 j ; and let F(u, v) represent the approximated image at the same scale. If the undecimated DWT is decomposed J times, we will have wavelet coefficients, F and F with j = 1J and the coarsest approximate image. To construct a WBN, we first group subbands with the same orientation together to obtain a horizontalgroup, a vertical-group, and a diagonal-group of wavelet coefficients. Then, we construct a Bayesian network B for each group. Let Bh = (Vh, Eh, Ph), Bv = (V v , Ev , Pv), and Bh = (Vd , Ed , Pd) denote the Bayesian networks constructed from the horizontal-group, vertical-group, and diagonal-group of wavelet coefficients respectively.

(j) Vertex set

Let the size of the input image F be N*N. If J wavelet decompositions are applied to F, there will be J sub-bands of size N*N in each orientation. Let a wavelet patch (matrix of random variable) be of size m*m. For each sub-band, a graph of m^2 variable nodes are formed. We then associate each random variable in the wavelet patch to a variable node in the graph. Without loss of generality, we assume that m divides N. Each sub-band can then be partitioned into $(Nm)^2$ blocks, each of size m*m. Then, $(Nm)^2$ wavelet coefficients sampled from the sub-band are assigned to each variable node. Let $x_{h,j}$ (i, k), with j = 1, . . . , J and i, k = 0, . . . ,m 1, denote the (i, k) variable node in the j^{th} sub-band. In our construction, the $(Nm)^2$ wavelet coefficients assigned to node x_j^u (i, k) are sampled from w_j^u F(i+mp, k+mq) with p, q =0, . . . , Nm 1 . If we represent each node as a vertex in the Bayesian network B_u, then the vertex set of B_u will be and the $(Nm)^2$ wavelet coefficients can be regarded as sampled from some (unknown) distribution of a random variable. Figs. 1(a), (b), and (c) show the procedure used to construct the vertex set for a sub-band of 4*4 coefficients.

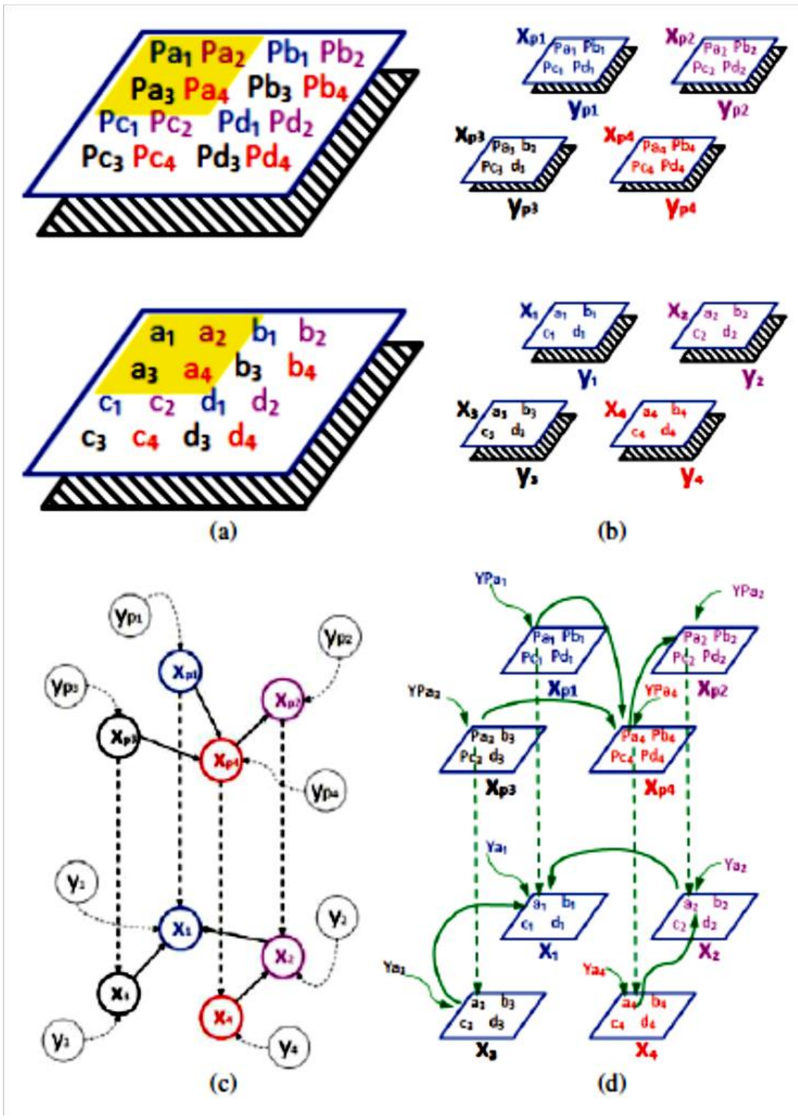
(k) Edge set Eu

The arcs (directed edges) in Bu can be divided into two disjoint sets, Eu (o) and Eu (i), where Eu (o) comprises the (inter-scale) edges incident to vertices at different scales, and Eu (i) comprises the (intra-scale) edges incident to vertices at the same scale. The persistence property of the wavelet transforms indicates that large/small values of wavelet coefficients tend to occur at the same spatial locations in sub-bands at adjacent scales. The property can be used to construct arcs in Eu (o) by linking a vertex at the coarser scale j + 1 to the vertex of the same index at the edges of finer scale j; that is, Eu (o) = x_j^u +1(i, k) x_j^u (i, k)— i, k = 0, . . . ,m 1 and j = 1., J 1. (8) The edges in E_u (i) represent the connections between vertices the same scale and orientation. Constructing the edges corresponds to deriving the Bayesian network on the nodes (i, k) > x_j^u (i,k)that best represent the joint probability of the nodes at the same scale j and orientation u. However, as discussed in Section II, BP inference is computationally intractable if the Bayesian network is a general graph. Thus, we limit the solution space to spanning trees so that we can derive an efficient solution by using the maximal weighted spanning tree (MWST) algorithm. A maximal weighted spanning tree is a spanning tree whose weight is greater than or equal to the total weight of every other spanning tree. The optimum weighted spanning tree can be derived by minimizing the relative entropy (Kullback-Leibler distance) D(p/q) between the probability functions p and q. In the following, we show how the spanning tree that minimizes D(p/q) is equivalent to the tree that maximizes the weighted summation of conditional mutual information on all the edges of the tree. Let x be the vector of variables x1.xn; let p(i) and b(i) denote the indices of the parent nodes and the sibling nodes of xi respectively; and let q be the induced probability of the spanning tree. Note that nodes at the coarsest scale do not have parent nodes. To find the optimal spanning tree, we minimize thebrelative entropy between q(x) and the joint probability p(x) given in equation.

$$D(P/Q) = \sum_x P(x) log \frac{p(x)}{q(x)} = \sum_x p(x) log p(x) - \sum_x p(x) log q(x)$$

(l) Probability model Pu

There are two types of arcs in a Bayesian network Bu: 1) the inter-scale parent-child arc, which connects a node with its coarser-scale parent; and 2) the intra-scale sibling arc, which connects two nodes of the same scale. To obtain the probability inference, we need to model the probability function on each arc. exploited the persistence property of wavelet transforms and proposed a joint statistical model of a ”child” coefficient conditioned on the coarse-scale ”parent” coefficients at the same spatial locations in all orientations Let (x pk) comprise the parent coefficients of the child coefficient x. Then, the probability function of x conditioned on (x pk).



(a) Top layer and the bottom (dashed) layer of each group are composed of hidden subband coefficients and coefficients observed at the corresponding subband, respectively. (b) Coefficients in observed (dashed) layers are organized in the same way as the coefficients in the hidden layers. (c) Observation nodes {ypi , yi } are created and linked to hidden nodes {x pi , xi }. (d) For denoising proposes, data in the network is organized into four groups and BP inference is applied to each group. For example, {ypai , pai , yai , ai |i = 1, 2, 3, 4} is a group. The other three data groups can be derived in a similar manner.

IV. WAVELET BAYESIAN NETWORKS FOR DENOISING

Here we consider using the wavelet Bayesian network to model the prior probability of the original image for the image denoising problem, which involves removing white and homogenous Gaussian noise with zero mean and known variance from an image. To infer the probability for denoising, we associate each variable node *x* in Bayesian network *B* with an observation node *y* and create the arc *y* → *x*. The probability function of *x* conditioned on the observed value of *y* is modeled as

$$f_n(x|y) \propto \frac{1}{\sigma_n} \exp(-\frac{(x-y)^2}{2\sigma_n^2 \rho})$$

Where σ_n^2 is the variance of the zero mean Gaussian white noise and ρ depends on the scale and wavelets. Recall that the variable nodes in the Bayesian network B are represented by $\{x_j^u$ (i, k)|u ∈ {h, v, d}; j = 1, . . . , J ; i, k = 0, . . . ,m − 1}. Let y_j^u (i, k) denote the observation node corresponding to x_j^u (i, k); and let Y , En, and Pn denote the collections of {yju (i, k)}, the arcs { y_j^u (i, k) →x_j^u (i, k)}, and the probability functions { fn(x_j^u (i, k)| y_j^u (i, k))} respectively. The WBN Bn for image denoising is constructed and represented as Bn = (V ∪ Y, E ∪ En, P ∪ Pn). Let a noisy image Z = F + N, where F is the original image and N is zero-mean white Gaussian noise. Each wavelet coefficient of Z is assigned to one observation node in Bn. That is, the coefficient Wuj Z(i + mp, k + mq) is assigned to observation node y_j^u (i, k), where p, q = 0, · · · Nm − 1; thus, each observation node has $(Nm)^2$ observation values and realizations. If we take one wavelet coefficient from each observation node, we can obtain $(Nm)^2$ realizations of Bn, denoted as (p, q) with p, q = 0, · · · Nm − 1. Note that, similar to the example shown in Fig. 3(d), the wavelet coefficients assigned to the $(p, q)^{th}$ realization are taken locally from the (i, k)-blocks of all subbands at scale 2 j and orientation u.

We use the message passing algorithm to obtain the estimated wavelet coefficients of each realization. First, we convert WBN Bn to a factor graph Fn, and then use the maxproduct algorithm to derive the estimated wavelet coefficients. The conversion of Bn to Fn and the max-product message passing schemes are standard techniques. For completeness, we provide them in Appendix A. The last step of the max product algorithm calculates the marginal probability of each variable node V in Fn. Let N(x) represent the neighboring factor nodes of variable node x in Fn. In addition, let x p and xc denote, respectively, the parent variable node and child variable node of x in Bn; and let {x j } denote the sibling variable nodes of x in Bn. The value of ˆ x can be estimated based on whether x has a child node.

V. Conclusion

The Bayesian formula indicates that the denoising problem is essentially a prior probability modeling and estimation task. In this paper, we present a constructive data-adaptive procedure that derives a hidden graph structure from the wavelet coefficients. The graph is then used to model the prior probability of the original image for denoising purposes. Moreover, we show that if the network is a spanning tree, the standard BP algorithm can estimate MAP efficiently. We compare our denoised results with those derived by other approaches, including BM3D, and demonstrate that our method yields a better PSNR and better perceptual quality on the textured areas of an image. Extending our method to content sensitive wavelet patches is an issue that merits future study. We will also investigate ways to speed up our algorithm's execution time.

References

1. B. Goossens, A. Pizurica, and W. Philips, Image denoising using mixtures of projected Gaussian scale mixtures, IEEE Trans. Image Process., vol. 18, no. 8, pp. 16891702, Aug. 2009.
2. A. Srivastava, A. B. Lee, E.P. Simoncelli, and S-C. Zhu, OnAdvances in Statistical Modeling of Natural Images, Journal of Mathematical Imaging and Vision, vol. 18, pp. 17 33, 2003.
3. J. Ho and W.-L. Hwang, Wavelet Bayesian Network Image Denoising, IEEE Trans. Image Process, vol. 22, no. 4, Apr. 2013.
4. G. F. Cooper and E. Herskovits, A Bayesian method for the induction of probabilistic networks from data, Mach. Learn., vol. 9, no. 4, pp.309347, Oct. 1992.
5. D. Heckerman, D. Geiger, and D. M. Chickering, Learning Bayesian networks: The combination of knowlege and statistical data, Mach. Learn., vol. 20, no. 3, pp. 197243, 1995.
6. D. Heckerman, A tutorial on learning Bayesian networks, Microsoft Research, Mountain View, CA, Tech. Rep. MSR-TR-95-06, 1995.
7. D. M. Chickering, D. Heckerman, and C. Meek, A Bayesian approach to learning Bayesian networks with local structure, in Proc. 13th Conf. Uncertainty Artif. Intell., 1997, pp. 8089.
8. D. J. Field, Relations between the statistics of natural images and the response properties of cortical cells, J. Opt. Soc. Am. A, vol. 4, no. 12, pp. 23792394, 1987.
9. S. Mallat, Multifrequency channel decomposition of images and wavelet models, IEEE Trans. Acoust., Speech, Signal Proc., vol. 37, no. 12, pp. 20912110, Dec. 1989.
10. J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli, Image denoising using Gaussian Scale Mixtures in the Wavelet Domain, IEEE Trans. Image Processing, vol. 12, no. 11, pp. 13381351, Nov. 2003.
11. J. Mairal, F. Bach, J. Ponce, and G. Sapiro, Non-local sparse models for image restoration, in Proc. IEEE Int. Conf Comput. Vis., Kyoto, Japan, 2009, pp. 22722279.
12. G. Chang, B. Yu, and M. Baraniuk, DSpatially adaptive wavelet thresholding with context modeling for image denoising, IEEE Trans. Image Process., vol. 9, no. 9, pp. 1522C 1531, Sep.2000.
13. M. Elad and M. Aharon, DImage denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process., vol. 15, no. pp. 3736C3745, Dec. 2006.

14. C. Kervrann and J. Boulanger, DOptimal spatial adaptation for patchbased image denoising, IEEE Trans. Image Process., vol. 15, no. 10, pp. 2866C2878, Oct. 2006

15. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, DA joint inter- and intrascale statistical model for Bayesian wavelet based image denoising, IEEE Trans. Image Process., vol. 11, no. 5, pp. 545C557, May 2002

16. M. Elad and M. Aharon, DImage denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process., vol. 15, no. 12, pp. 3736C3745, Dec. 2006.

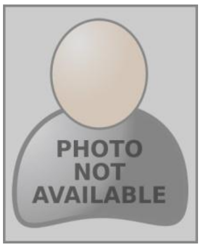
17. G. Chang, B. Yu, and M. Baraniuk, DSpatially adaptive wavelet thresholding with context modeling for image denoising, IEEE Trans. Image Process., vol. 9, no. 9, pp. 1522C 1531, Sep. 2000.

18. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, DImage denoising by sparse 3-D transformdomain collaborative filtering, IEEE Trans. Image Process., vol. 16, no. 8, pp. 2080C2095, Aug 2007

19. B. Goossens, A. Pizurica, andW. Philips, DImage denoising using mixtures of projected Gaussian scale mixtures, IEEE Trans. Image Process., vol. 18, no. 8, pp. 1689C 1702, Aug. 2009.

20. J. Ho and W.-L. Hwang, DWavelet Bayesian Network Image Denoising, IEEE Trans. Image Process, vol. 22, no. 4, Apr. 2013

AUTHOR(S) PROFILE



Nisam P M, received the B.Tech degree in Electronics and Communication from KMEA engineering College. Currently pursuing the M.Tech degree in Signal Processing from College Of Engineering Cherthala affiliated to CUSAT, Kerala.