

Bayesian Wavelet Domain Segmentation

Patrice Brault* and Ali Mohammad-Djafari†

**Institut d'Electronique Fondamentale,
Unité mixte de recherche 8622 (CNRS-UPS),
Université Paris-Sud, Orsay, 91405, France.
patrice.brault@lss.supelec.fr*

*†Laboratoire des Signaux et Systèmes,
Unité mixte de recherche 8506 (CNRS-Supélec-UPS),
Supélec, Plateau de Moulon, 91192 Gif-sur-Yvette, France.
djafari@lss.supelec.fr*

Abstract. We have recently demonstrated that fully unsupervised segmentations of still images and 2D+T sequences is possible by Bayesian methods, on the basis of a Hidden Markovian Model (HMM) and a Potts-Markov Random Field (PMRF), in the pixel domain. The use of a high number of iterations to reach convergence in a segmentation where the number of segments, or “classes” labels, is important makes the algorithm rather slow for the processing of a large quantity of data like in image sequences. We more recently have worked out a new version of this algorithm by first operating our segmentation in the wavelet transform domain rather than in the direct domain. Doing so, we take advantage of the local decay property, or “peaky” distribution of the wavelet coefficients, in an orthogonal decomposition. This decomposition is a fast pyramidal, $O(N^2)$, decomposition, so the Bayesian segmentation is performed only once on the first coarse image then on all subbands up to the highest resolution level. Moreover, we have improved our Potts-Markov model in order to take into account the three main orientations of the wavelets band-pass, or so-called, detail, subbands. The main advantage of such an algorithm, in comparison with the direct domain Bayesian segmentation, is that the high frequency coefficients, i.e. the coefficients of all sub-bands except the coarsest, are segmented in only 2 classes : 1 for the weak energy coefficients and 2 for the few, and most representative, high energy coefficients, thus enabling to speed up the convergence process of the segmentation.

INTRODUCTION

The supervised, partially-unsupervised and fully unsupervised segmentation of still images has now been widely studied and described in the literature [8]. Several developments of the Bayesian segmentation [2, 3, 6] based on multiscale, wavelet domain and Hidden Markov Tree (HMT), take also advantage of this interesting property of the wavelets to exhibit quite sparse significant coefficients, thus reducing the computation complexity. In our group, one method is based on a fully unsupervised segmentation in the direct domain [7, 4]. This method uses a HMM for the z variable which labels the different regions of an image. This HMM is based on a PMRF modeling for the pixels, which enables to render the strength of dependency of neighboring pixels to insure a good homogeneity of the region being segmented. In this direct domain approach the PMRF uses a first order neighborhood. The same model has been used more recently to segment $2D + T$ videos sequences, i.e. regularly sampled sequences of images [1]. We have thus demonstrated a significant improvement in the segmentation of a sequence of

images. Turning this result to motion detection and quantification and also to sequence compression, is being under way and might turn to be an efficient method amongst other recent developments for motion acquisition, quantification and compression [1]. Nevertheless, though very efficient for long time dependence segmentation processes, we are also very much interested in reducing the segmentation time. This has motivated a new approach based on the use of a wavelet transform domain to perform the segmentation. Due to their specific property of fast “local-decay”, the wavelet coefficients, obtained by decomposition on an orthogonal basis, behave like a mixture of zero-centered Gaussians [3, 6, 5]: a first, high variance, distribution representative of a few strong coefficients of major importance, and a second, “peaky” low variance distribution representative of a large number of low magnitude and low importance coefficients. Thus the segmentation of these band-pass “subband” coefficients can be made using only 2 classes labels. Furthermore, by zeroing coefficients pertaining to the class of low importance coefficients, we denoise the data and increase the performances of the segmentation. Based now on a decomposition in the wavelet domain we found also pertinent to improve the PMRF model used in the direct domain. This is why we adopted a first + second order, 8-connexity model, for the PMRF. This was motivated by the fact that wavelet subbands are split in vertical, diagonal and horizontal details. We thus tuned the Markov Field neighboring to the wavelet subbands orientations through a specific α dependency parameter for the Potts model. So we naturally fixed three parameters: α_V , α_D and α_H as dependency parameters for our Potts model.

MODELING FOR DIRECT DOMAIN SEGMENTATION

The description of still image segmentation and fusion by a Bayesian approach and a Hidden Markov Model (HMM) has been developed and well described by Féron et al in [4]. We summarize in this section the main steps for the computation of the posterior laws necessary in the final segmentation algorithm. We write the basic relations, explaining that the inverse problem is to find a non-noisy form f of the observed image g and also to extract a segmentation in statistically homogeneous regions of f . We thus can write down the basic expression :

$$g(\mathbf{r}) = f(\mathbf{r}) + \epsilon(\mathbf{r}), \quad \mathbf{r} \in \mathcal{R} \quad (1)$$

where \mathcal{R} is the image set of sites.

If we note by $\mathbf{g} = \{g(\mathbf{r}), \mathbf{r} \in \mathcal{R}\}$, $\mathbf{f} = \{f(\mathbf{r}), \mathbf{r} \in \mathcal{R}\}$ and $\boldsymbol{\epsilon} = \{\epsilon(\mathbf{r}), \mathbf{r} \in \mathcal{R}\}$, and if we assume $\{\epsilon(\mathbf{r}), \forall \mathbf{r} \in \mathcal{R}\}$ are centered, iid and Gaussian, we can write :

$$p(g(\mathbf{r})|f(\mathbf{r})) = \mathcal{N}(0, \sigma_\epsilon^2) \quad (2)$$

and

$$p(\mathbf{g}|\mathbf{f}) = \mathcal{N}(0, \sigma_\epsilon^2 \mathbf{I}) \quad (3)$$

Because the goal is to have a reconstructed image f segmented in a limited number of homogeneous regions, a hidden variable z which can take the discrete values $k \in \{1, \dots, K\}$ is introduced which represents the f image classified in K classes thus

segmenting the image in regions $\mathcal{R}_k = \{\mathbf{r} : Z(\mathbf{r}) = k\}$ that can be gathered, or not, to retrieve a segmentation in homogeneous regions:

$$p(f(\mathbf{r})|z(\mathbf{r}) = k) = \mathcal{N}(m_k, \sigma_k^2) \quad (4)$$

In order to find homogeneous regions, we introduce a spatial dependency via a Potts-Markov random field which fixes the dependency between the homogeneous zones of an image by means of an attractive/repulsive “ α ” parameter:

$$p(z(\mathbf{r})_{\mathbf{r} \in \mathcal{R}}) = \frac{1}{T(\alpha)} \exp \left\{ \alpha \sum_{\mathbf{r} \in \mathcal{R}} \sum_{\mathbf{s} \in V(\mathbf{r})} \delta(z(\mathbf{r}) - z(\mathbf{s})) \right\} \quad (5)$$

where $V(\mathbf{r})$ represents the neighborhood of \mathbf{r} . In the sequel we consider $V(\mathbf{r})$ as the first order neighborhood (4-connectivity) of the pixel \mathbf{r} .

The prior laws, defined formerly, need some parameters which are σ_ϵ , σ_k and m_k . So we also assign prior laws to these parameters by means of “hyper-hyperparameters” α_0 , β_0 , m_0 and σ_0 . We refer to [7] for the choice of these hyper-hyperparameters. So:

$$\theta = (\sigma_\epsilon^2, m_k, \sigma_k^2) \quad (6)$$

is the set of hyper-parameters, with laws:

$$\begin{cases} p(\sigma_\epsilon^2) & \sim \mathcal{IG}(\alpha_0, \beta_0) \\ p(m_k) & \sim \mathcal{N}(m_0, \sigma_0^2), k = \{1 \dots K\} \\ p(\sigma_k^2) & \sim \mathcal{IG}(\alpha_0, \beta_0), k = \{1 \dots K\} \end{cases} \quad (7)$$

$\forall k \in \mathbb{N}$, and \mathcal{IG} is the Inverse-Gamma.

Posterior distributions and Gibbs sampling algorithm

The Bayesian approach consists now in estimating the whole set of variables (f, z, θ) following the joint *a posteriori* distribution $p(f, z, \theta|g)$ using the following Gibbs sampling algorithm :

$$\begin{cases} f & \sim p(f|g, z^{(n-1)}, \theta^{(n-1)}) \\ z & \sim p(z|g, \theta^{(n-1)}) \\ \theta & \sim p(\theta|f^{(n)}, z^{(n)}, g) \end{cases} \quad (8)$$

We refer here to [4] for details on the expressions and more detailed computational implementation of the conditional *posterior* laws.

This algorithm is iterated a “sufficient” number of times (*itermax*) in order to reach the convergence of the segmentation. We do not use any real “convergence criterion” and by convergence we mean that the segmented regions do not change significantly in the next iterations. After convergence, we take the max(histogram) for each pixel value and for all iterations. Also, our experience on the image analyzed has showed us

that $itermax$ depends essentially on the complexity of the image and the number of luminance levels, as well as on the number of classes taken for the segmentation. In general for a number of classes $K = 4$ we take $itermax$ between 20 and 50.

WAVELET TRANSFORM PROPERTIES

A multiresolution wavelet transform is described by scaling (approximation) and wavelet (detail) coefficients. For a special choice of the scaling and wavelet functions, the total decomposition (all the coefficients) of a one dimensional signal $T(x)$ can be written:

$$T(x) = \sum_K u_{(J_0,K)} \cdot \varphi_{(J_0,K)}(x) + \sum_{J=-\infty}^{J_0} \sum_K w_{(J,K)} \cdot \psi_{(J,K)}(x) \quad (9)$$

where φ and ψ are respectively the scaling and wavelet functions. On another hand, it has been established [3, 6] that the wavelet coefficients of “real-world” signals exhibit a local decay property which means that the coefficients of highest energy, and of utmost representativity for the signal, are very sparse and that the coefficients of low energy, and of low importance for the signal, are in large quantity. In the probability domain, we can then express the marginal density of the wavelet coefficients by one spread and heavy-tailed, i.e. with a large variance, Gaussian density and by another very peaky, and low variance, Gaussian density. Such a density is well expressed by a mixture of Gaussians and the coefficients can be considered approximately decorrelated due to the DWT decomposition used (Fig.2-b). It has also been established that for real-world signals, the wavelet coefficients are not statistically independent in general and especially that neighbor wavelet coefficients tend to be dependent within and across scales, creating “clusters” of large and small coefficients. This second property has led to Wavelet-domain Hidden Markov Models in which the hidden states have a Markov dependency structure. More specifically the dependency across scales has been modeled by a HMT, a Hidden-Markov Tree model, with mixture parameters $\mu_{i,m}, \sigma_{i,m}^2$ and transition probabilities $P_{S_i|S_{\rho(i)}}(m|n)$.

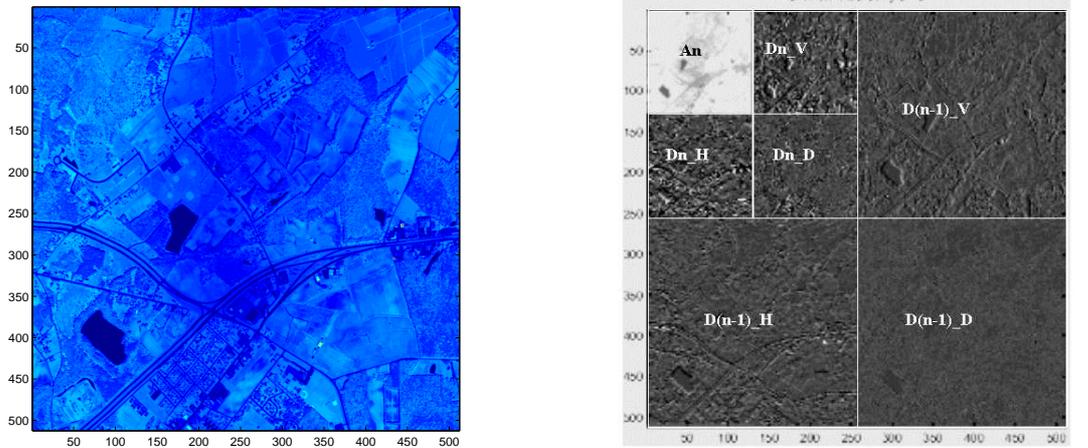


Figure 1: a) Original image of the Channel 100 of an hyperspectral satellite image composed

of 224 frequency bands. b) Mallat Pyramidal representation of the Fast Orthogonal Wavelet 2-levels decomposition applied to the the channel 100 of our hyperspectral image.

EIGHT-CONNEXITY AND POTTS-MARKOV RANDOM FIELD

In order to find statistically homogeneous regions for the segments, our Markovian segmentation method in the direct domain uses a Potts-Markov random field [4] to define the spatial dependency of the labels in the HMM. In this direct model, the dependency of the pixels is searched in a first order neighboring. Nevertheless and due to the fact that the segmentation is done now in the wavelet domain, we have to take into account that this spatial dependency is quite different than in the direct domain due to the fact that the three “detail” subbands are oriented in the vertical, diagonal ($D_1 = \pi/4$ and $D_2 = 3\pi/4$) and horizontal directions. This prior knowledge is taken into account by introducing these orientations in a new “eight-connexity” (or neighboring of orders 1 and 2) Potts-Markov Random Field. The new model of the PMRF can be written:

$$p(z) = \frac{1}{T(\alpha_V, \alpha_{D_1}, \alpha_{D_2}, \alpha_H)} \times \exp\{\alpha_V \sum_{r(i,j) \in S} \delta(z(i,j) - z(i,j-1)) + \alpha_{D_1} \sum_{r \in S} \delta(z(i,j) - z(i-1,j)) + \alpha_{D_2} \sum_{r \in S} \delta(z(i,j) - z(i,j-1)) + \alpha_H \sum_{r \in S} \delta(z(i,j) - z(i-1,j))\} \quad (10)$$

The parameters α_V , α_{D_1} , α_{D_2} and α_H control respectively the degree of spatial dependency of the z variable in the directions V , D_1 , D_2 and H . For the wavelet “diagonal” subband we gather the D_1 and D_2 dependencies in one diagonal $D_1 D_2$ dependency. For the first, coarse, approximation subband, we also gather the V and H dependency in a VH dependency which is equivalent to the 4-connexity dependency (first order neighboring) used in the direct-domain implementation [4]. Also, as described in [4], the implementation of the MCMC-Gibbs algorithm, can be made in parallel by considering two independent sets of pixels referenced by the indexes of the black and white cases of a chessboard. This implementation in parallel is also used for the horizontal and vertical neighborings (first order) but the “chessboard scheme” is replaced by an alternance of white and black rows for the implementation in the diagonal neighboring (second order). The application of this model to the wavelet subbands is done by annealing the PMRF terms that do not apply to the subband concerned. Moreover the α_V parameter used in this “vertical” term can be adjusted to any value > 0 , which means that we assign a specified, non null, dependency strength between the labels z of the region concerned. As we have seen in [4], this means that the higher the α parameter, the higher the *a priori* of a little number of great homogeneous regions.

BAYESIAN SEGMENTATION IN THE WAVELET DOMAIN

The initial operation to perform on the image to be segmented, is first to do its decomposition on an orthogonal wavelet basis. The decomposition used here is the classical “Mallat” pyramidal transform of complexity $O(N^2)$. This transform enables to get wavelet coefficients which are split in two main classes : the weak and the strong coefficients. Noticing this point on the wavelet coefficients has an important incidence for segmentation purposes : it enables to segment the “detail” images with only two classes ($K = 2$). On complex images, i.e. the first “approximation” subband, the segmentation will be led with a higher number of classes, e.g. $K = 8$, and a high number of iterations.

Coarse-to-Fine Scale Segmentation Algorithm

Our segmentation scheme in the wavelet domain is based on a coarse to fine scale scheme (Fig.4). The Bayesian segmentation is performed on the wavelet subband coefficients and at all scales with a low value $K = 2$, except for the approximation coefficients. The segmentation starts from the g observation projected into the wavelet domain. The segmentation algorithm can be described with the following steps:

- 1) Wavelet decomposition to the order n on an orthogonal basis (Haar wavelet).
- 2) We segment the approximation coefficients A_n at scale n with the number of classes desired in the final segmentation, e.g. $K = 8$. For the Gibbs sampling, and with a large value of K , the iteration number is given a high value in order to assure the convergence.
- 3) In the segmented image of the approximations $Z(A_n)$, we detect (by derivation) the regions exhibiting vertical, diagonal ($\pi/4$ and $3\pi/4$) and horizontal discontinuities.
- 4) At this same scale, the 3 detail subbands W_{Vn} , W_{Dn} and W_{Hn} are segmented and we respectively use, as an initialization of these segmentations, the 3 subsets of discontinuities dif_v , dif_d and dif_h computed at the former step. This step is realized with a weak number of classes ($K = 2$) and of iterations of the Gibbs sampling.
- 5) The 3 segmented detail subbands are upsampled by 2 in order to repeat the process at the next upper level.
- 6) We segment the 3 detail subbands $W_{V(n-1)}$, $W_{D(n-1)}$ and $W_{H(n-1)}$ on the basis of the initialization obtained in the former step. The same process is then repeated up to the image resolution level.
- 7) This is the final step where we reconstruct the segmented image starting from the coarsest level n of the decomposition. The reconstruction uses:
 - for the approximation initial subband (level n): the average of the original scaling coefficients within each region of the segmented approximation subband: $(\bar{A}_n(Z_k))$ with $k \in 1, \dots, K$.
 - for all the detail subbands: the original wavelet coefficients for the segments that pertain to the class $k = 2$. The coefficients pertaining to the class $k = 1$ are cancelled.

Advantage of this algorithm

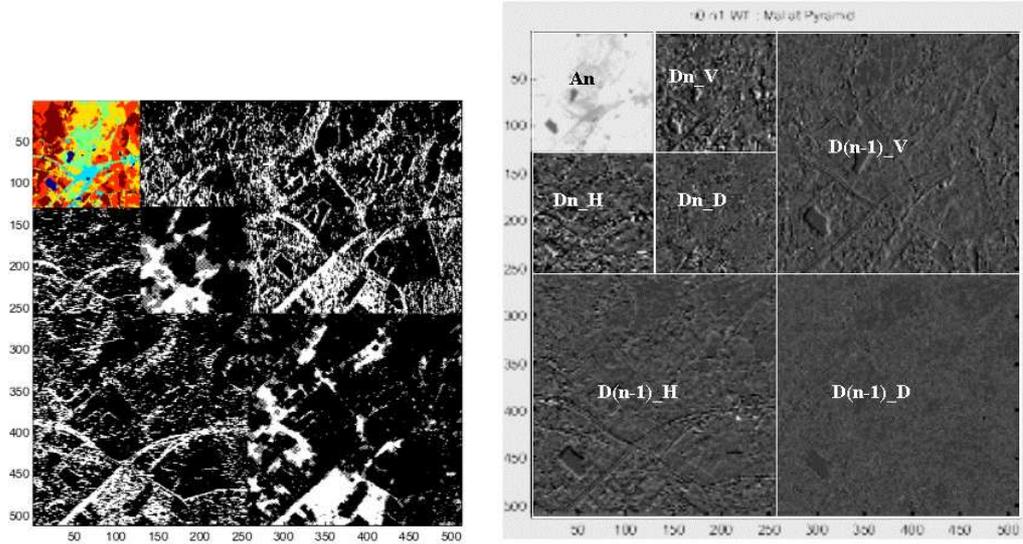


Figure 2: a) Mallat's pyramidal representation of the segmented image Z . The approximation part is segmented in 8 classes (color part). The detail subbands are represented in black for the class of the weak coefficients, that are zeroed, and in white for the class of the strong coefficients that will be used for the final reconstruction. b) Linear and semilog(y) histograms of the $Dn0d$, the detail diagonal coefficients.

This combined wavelet-HMM Bayesian segmentation has a first application in fast segmentation as well as fast compression schemes for still images. A direct application is also to improve our Bayesian segmentation of video sequences as well as their compression and the motion analysis we have initiated in [1].

RESULTS AND COMPARISON

We have compared the results of the wavelet-domain segmentation with a direct domain segmentation. The segmentation in the wavelet-domain has been tested only with two levels of decomposition. Qualitatively, it provides more homogeneous regions where many different details appear. Conversely, we can see some artifacts due probably to the segmentation of the detail subbands. Quantitatively, the segmentation is faster: the computation time is reduced by two.

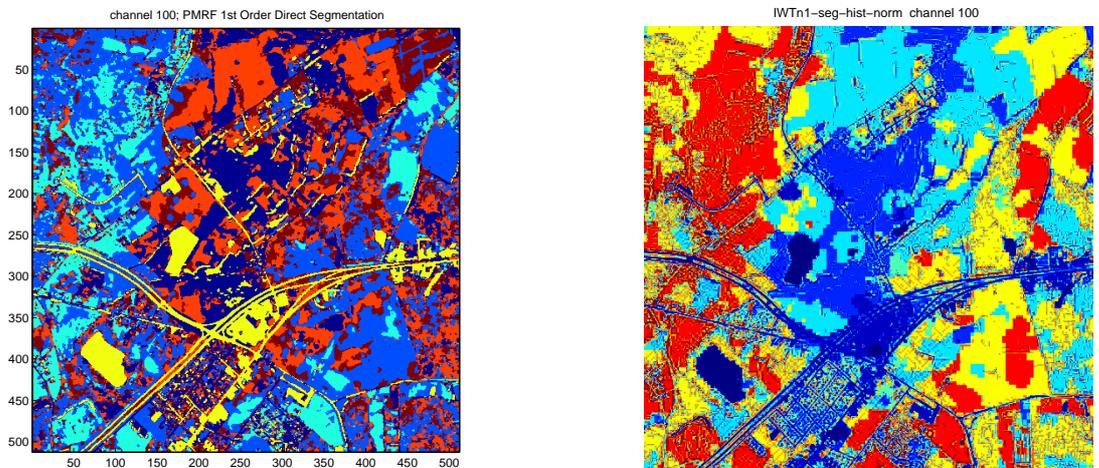


Figure 3: a) Bayesian segmentation in the direct domain based on a first order PMRF. b) Segmentation result obtained by the proposed method using the W.T.

CONCLUSION

We have presented a new algorithm for the segmentation of images based on a Bayesian segmentation operated in the wavelet domain. The first originality of this work resides in the fact that we do the image segmentation in the wavelet domain, but rather than using the HMT property of wavelet coefficients, we concentrate on the initialization of a PMRF approach in a coarse to fine scale scheme. This Wavelet Bayesian segmentation is mainly based on the hypotheses of Gaussian distributions for the image, the noise and the regions segmented. Due to the fact that the segmentations are done with only two classes for all the Band-Pass subbands, we obtain a significant reduction of the complexity. The second originality of this work is that a second order, 8-connexity, Potts-Markov Random Field has been developed to fit to the three main orientation of the detail subbands of the wavelet decomposition.

This Bayesian-Wavelet segmentation scheme has led to a reduction of the computation time by two. It has, until now, been tested with a wavelet decomposition on 2 levels and we think future implementations with more decomposition levels will exhibit faster computation times and better segmentation results.

REFERENCES

1. Brault, P., and Mohammad-Djafari, A., "Bayesian Segmentation and Motion Estimation in Video Sequences using a Markov-Potts Model", *WSEAS04 Int'l Conf. on Applied Math.*, Miami, april 2004. Invited Paper. (Accepted in *WSEAS Transactions on Mathematics*, 2004).
2. Choi, H., and Baraniuk, R.G., "Multiscale image segmentation using wavelet-domain hidden Markov models", *IEEE Transactions on Image Processing*, Vol. 10, Issue 9, pp:1309-1321, Sept. 2001.
3. Crouse, M.S., Nowak, R.D., and Baraniuk, R.G., "Wavelet-based statistical signal processing using hidden markov models", *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 886-902, 1998.
4. Féron, O., and Mohammad-Djafari, A., "Image fusion and unsupervised joint segmentation using a HMM and MCMC algorithms", submitted to *Journ. of Electronic Imaging*, April 2004.
5. Ichir, M., and Mohammad-Djafari, A., "Bayesian wavelet based statistical signal/image separation", in *AIP Conference Proceedings of MaxEnt03*, Jackson Hole, USA, Aug. 2003.

6. Romberg, J.K., Choi, H., and Baraniuk, R.G., "Bayesian Tree-Structured Image Modeling using Wavelet-Domain Hidden Markov Models", *IEEE Transactions on Image Processing*, Vol.10, Nř 7, July 2001.
7. Snoussi, H., and Mohammad-Djafari, A., "Fast Joint Separation and Segmentation of Mixed Images", *Journal of Electronic Imaging*, Vol. 13(2), april 2004.
8. Song, X., and Fan, G., "A study of supervised, semi-supervised and unsupervised multiscale Bayesian image segmentation", *MWSCAS-2002, 45th Midwest Symposium on Circuits and Systems*, Volume 2, Pages:II-371 - II-374, 24-7 Aug. 2002.

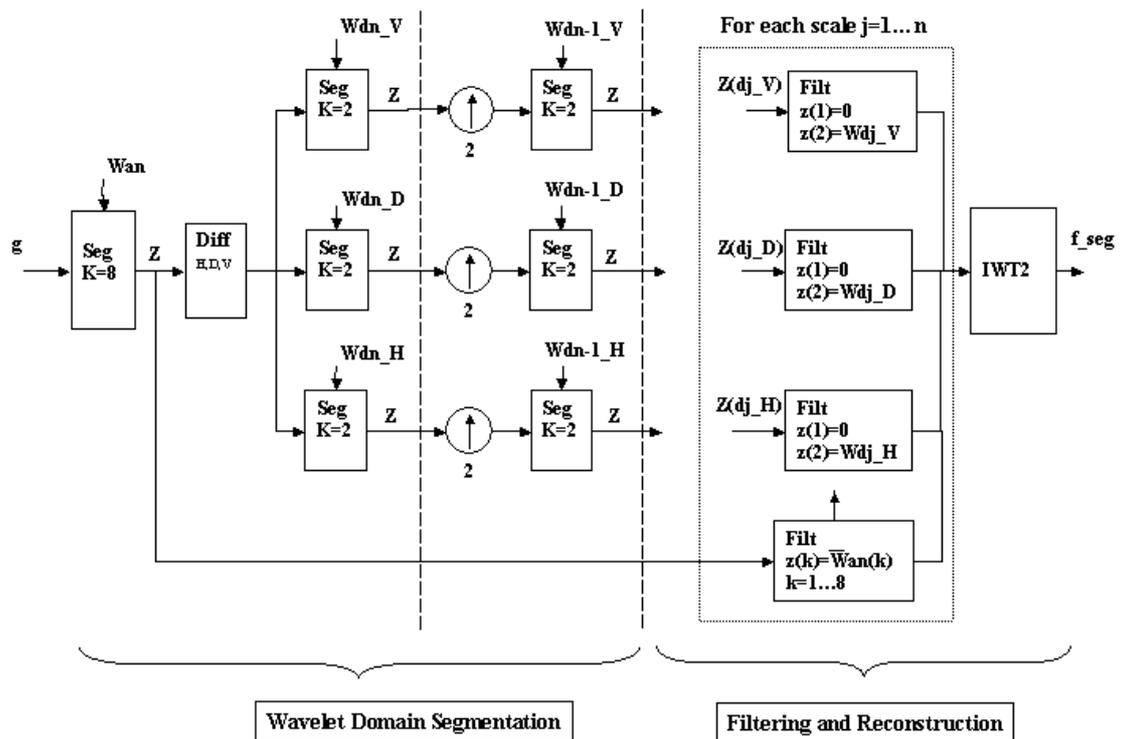


Figure 4: Wavelet domain Bayesian segmentation scheme. The observed data g is decomposed in the wavelet domain (2 scales shown here), segmented in this domain and the final segmentation is obtained by reconstruction in the direct domain..