

Denoising in Wavelet Domain Using Probabilistic Graphical Models

Maham Haider

Military College of Signals
National University of Sciences and Technology, Islamabad,
Pakistan

Imran Touqir

Military College of Signals
National University of Sciences and Technology, Islamabad,
Pakistan

Muhammad Usman Riaz

Military College of Signals,
National University of Sciences and Technology, Islamabad,
Pakistan

Adil Masood Siddiqui

Military College of Signals
National University of Sciences and Technology, Islamabad,
Pakistan

Abstract—Denoising of real world images that are degraded by Gaussian noise is a long established problem in statistical signal processing. The existing models in time-frequency domain typically model the wavelet coefficients as either independent or jointly Gaussian. However, in the compression arena, techniques like denoising and detection, states the need for models to be non-Gaussian in nature. Probabilistic Graphical Models designed in time-frequency domain, serves the purpose for achieving denoising and compression with an improved performance. In this work, Hidden Markov Model (HMM) designed with 2D Discrete Wavelet Transform (DWT) is proposed. A comparative analysis of proposed method with different existing techniques: Wavelet based and curvelet based methods in Bayesian Network domain and Empirical Bayesian Approach using Hidden Markov Tree model for denoising has been presented. Results are compared in terms of PSNR and visual quality.

Keywords—Gaussian Mixture Models (GMM); Hidden Markov Model (HMM); Discrete Wavelet Transform (DWT); Hidden Markov Tree (HMT)

I. INTRODUCTION

Removing noise from an image with less loss of information is referred as Image Denoising. Even the simplest acquisition, processing and transmission of image can subject it to unwanted noise. Denoising is required at every level whenever there is a question about dealing with images [1]. It finds its applications ranging from image enhancement in terms of highlighting some useful aspects of information from an image to achieving better quality of medical images (CT scan, Ultrasound etc).

There have been a number of techniques and algorithms designed in wavelet domain. All of them worked successfully because of primary properties that wavelet transform has that is; locality, multi-resolution and compression. In [2], a framework for signal denoising in time frequency analysis domain using Hidden Markov Model (HMM) has been proposed. A neighbouring coefficients thresholding method is applied in multi-wavelet framework for denoising in [3]. A non-parametric tree based model for joint statistics of wavelet coefficients has been discussed in [4]. To realize neighbouring

dependency between wavelet coefficients across scales, a generalized Multivariate Gaussian distribution has been proposed in [5]. In [6], another tree model using hidden markov tree structure has been developed that refers to local parameterization for image denoising. In [7], wavelet shrinkage function is used to check neighbouring coefficients dependencies for image denoising. In [8], a denoising scheme has been realized that incorporates dependency of wavelet coefficients with three scale dependency. In [9], another image denoising technique based on neighbouring wavelet coefficients has been proposed.

Although there exists a number of denoising algorithms but there is still a need for improvement. Further research and study can result in an improved performance and better image quality. In this work, we have developed HMM based image denoising algorithm which is used in the context of 2D Gaussian Mixture Models (GMM) and 2D DWT. EM algorithm iteratively finds the maximum likelihood of a fundamental distribution from a given data set when the data is said to have some missing values. It is best suited in HMMs where the hidden states are not observed and the data is said to be missing. This modelling framework summarizes the nature of wavelet coefficients in a probabilistic way. This model finds its flexibility to the two main features; one is the Mixture Densities for dealing with the non-Gaussian nature of coefficients of wavelet transform by modelling them with hidden states as Mixture Distributions. The second feature is that of Probabilistic Graphs which models coefficients interdependencies and represent them in the form of a tree structure. Several local and standard images are put to test to check the performance parameters comparable to different other existing techniques for denoising. Results are shown both in terms of Peak Signal to Noise Ratio (PSNR) and the quality of denoised image.

II. STATISTICAL IMAGE MODELLING USING HIDDEN MARKOV TREE MODEL

Wavelet transform is known to be favorable because of its properties such as non-Gaussianity, Clustering and Persistence [5], [7], [8]. Fig.1, shows 2D discrete wavelet transform

(DWT). In this figure the decomposition upto 3 levels is shown. HMM combined with these properties provides an attractive model for capturing the non-Gaussian parameters of the wavelet coefficients that are persistent across scales [2], [10]. The details of wavelet transform and wavelet domain HMM along with applications in statistical image processing can be found in detail in [1], [2], [11], [12], [13] and [14].



Fig. 1. Three level decomposition using 2D Discrete Wavelet Transform

For each wavelet coefficient magnitude $|c_{j,i}|$, we can associate a set of Hidden states ($S_{j,i}$) with it. Given $S_{j,i} = k$, the pdf will be Gaussian with mean and variance as $\mu_{j,k}$ and $\sigma_{j,k}^2$ respectively. The overall pmf will be;

$$f(c_{j,i}) = \sum_{k=1}^K P(S_{j,i} = k) f(c_{j,i} | S_{j,i} = k) \quad (2)$$

where $P(S_{j,i} = k)$ is the probability mass function (pmf) and $f(c_{j,i} | S_{j,i} = k)$ is the conditional pmf given by the following equation;

$$f(c_{j,i} | S_{j,i} = k) = \frac{1}{\sigma_{j,k} \sqrt{2\pi}} \exp\left(-\frac{(c_{j,i} - \mu_{j,k})^2}{2\sigma_{j,k}^2}\right) \quad (3)$$

Owing to persistence across scale, state transition matrices A_t shows the parent \rightarrow child state-to-state links between hidden states given as;

$$A_t = \begin{bmatrix} p_t^{m \rightarrow m} & p_t^{m \rightarrow n} \\ p_t^{n \rightarrow m} & p_t^{n \rightarrow n} \end{bmatrix} \quad (4)$$

where $p_t^{s \rightarrow s'}$ shows that given s' being the parent coefficient state, the child coefficient is in state s [15]. HMT model on the whole is specified by $P(S_0 = k)$, pmf of the node S_0 , a state transition matrix A_t and $\mu_{j,k}, \sigma_{j,k}^2$, means and variances, of the wavelet coefficient $c_{j,i}$ given $S_{j,i}$ being in the state k . All these parameters are combined together in a vector θ . The state transitions and variances are generally different for each wavelet coefficient. This can lead to a more complex HMT model. To reduce the computational complexity, a method referred to as tying within scale is implemented as discussed in [2].

H. Chipman has shown that GMM is capable of well approximating this non-Gaussian density in [16]. GMM has a generative model with a random variable Z and marginal distribution as $\sum_{k=1}^m \alpha_k N(c_{j,i} | \mu_{j,k}, C_{j,k})$ where $C_{j,k}$ represents the co-variance matrix of coefficients. A multidimensional GMM is referred to as HMT. To capture the non-Gaussianity

of the wavelet coefficients which is referred to the Clustering property, it associates each of the wavelet coefficients magnitude $|c_{j,i}|$ with a hidden variable ($S_{j,i}$) that is said to be unobserved. We find few of the coefficients with large magnitudes and they tend to contain the maximum amount of information regarding the image. On the other hand, the coefficients with smaller magnitudes have the lesser amount of relevant information about the image but they tend to exist in large numbers. This leads to the simplest model with only two states; 'high' and 'low' referring to large and small magnitudes of wavelet coefficients respectively. Several results have proved this model to be simple in nature but effective in estimation point of view.

To capture the inter-scale dependencies between the wavelet coefficients, referring to the Persistence property [15], [17], GMM performs Markovian Chain across scales that are tree-structured as the magnitude of the coefficients is said to be only depending on the size of their respective parents alone. This means that the probability of a coefficient to be 'large' or 'small' is determined only by the magnitude of its parent. HMT in time-frequency domain was developed in [2] that connect the state variables vertically. For images it forms a structure in a quad-tree fashion that associates each hidden state variable as parent to its four children. This model satisfies both secondary properties that wavelet transform has such as Clustering and Persistence. HMT is applied to image processing in [11], [13] and [15].

III. MODEL TRAINING VIA EM ALGORITHM

In this section a denoising model has been discussed that uses EM algorithm to determine noise-free coefficients from noisy elements. Translating the problem in image denoising, there is a need to determine noise-free coefficient c from the noisy coefficients y . Estimate $c_{j,i}$ such that $y_{j,i} = c_{j,i} + n_{j,i}$, where $n_{j,i} = \mathcal{N}(0, \sigma_n^2)$. Generalizing it, this model refers to determining the vector denoted by θ . The following relation of expectation describes the sufficient statistics S_t of the model for variable X and the hidden state variable Z ;

$$E_{\theta_0} (S_t (X, Z) | X = x) = E_{\theta} S_t (X, Z) \quad (5)$$

The conditional pmf of hidden states $S_{j,i}$ and its maximization is given by following expressions;

$$P(S_{j,i} = k | c_{j,i}, \theta') = \frac{P(S_0 = k) g(c_{j,i}; 0, \sigma_{j,k}^2)}{\sum_{l=0}^1 P(S_0 = l) g(c_{j,i}; 0, \sigma_{j,l}^2)} \quad (6)$$

$$P(S_0 = k) = \frac{1}{N_j} \sum_{i \in \mathbb{Z}^2} P(S_{j,i} = k | c_{j,i}, \theta') \quad (7)$$

The noise-free coefficients $c_{j,i}$ can be obtained from the following expression;

$$c_{j,i} = E [c_{j,i} | y, \theta] = \sum_{i \in \mathbb{Z}^2} P(S_0 = k | y, \theta) \frac{\sigma_{i,k}^2}{\sigma_n^2 + \sigma_{i,k}^2} y_i \quad (8)$$

Fig.2 shows the denoising scheme used in this research work. The proposed scheme for image denoising can be summarized as follows;

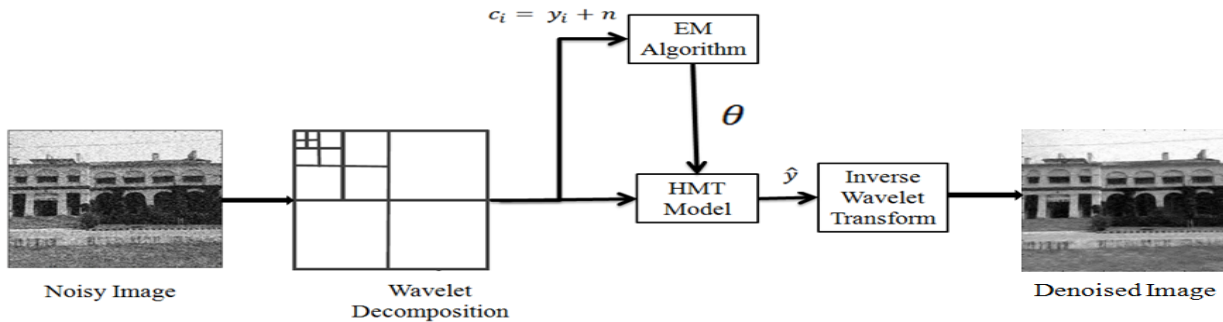


Fig. 2. Proposed Denoising Scheme

- 1) Add AWGN noise in the real-world image ($N \times N$ dimensions) after converting it into gray scale
- 2) Apply Daubechies-8 DWT on the image to divide it into three levels of DWT
- 3) Estimate the initial GMM parameters of each sub-band from the noisy coefficients.
- 4) Train the HMT model using EM algorithm with reference to tying within scales method
- 5) Apply inverse wavelet transform to estimate the noise-free coefficients to obtain the denoised image

IV. EXPERIMENTAL RESULTS

We have tested standard as well as local gray scale images in order to compare this scheme with several other denoising methods that are used. Each image is of 256x256 dimensions. We have assumed that each image is corrupted with Gaussian noise with three different known variances $\sigma = 10, 20$ and 30 . The image corrupted by noise is decomposed into three levels of wavelet transform by using Daubechies-8 Wavelet. Each sub-band is then applied by the proposed denoising scheme.

TABLE I. PSNR VALUES OF DENOISED IMAGES USING DIFFERENT METHODS

Image	σ_n	Wavelet methods using Bayesian Approach [18]	Curvelet Transform using Bayesian Approach [18]	Empirical Bayesian Approach Using HMT [15]	Proposed
Lenna	10	27.58	31.93	30.51	32.50
	20	26.02	28.49	28.56	28.74
	30	22.03	26.70	26.70	26.63
Camera Man	10	24.63	30.09	30.42	30.89
	20	24.86	26.67	25.32	27.94
	30	22.72	24.80	25.42	25.51
Peppers	10	29.00	29.93	32.61	32.33
	20	22.39	26.46	27.11	28.85
	30	18.78	24.67	25.33	26.60
Barbara	10	28.79	29.57	-	31.89
	20	24.14	26.41	-	27.88
	30	22.46	24.69	-	25.79

Standard images include Lena, Camera Man, Peppers and Barbara. These images are tested using the proposed method comparable to other techniques including Wavelet Transform using Bayesian Network Approach [18], Curvelet Transform using Bayesian Network Approach [18] and Empirical

Bayesian Approach using HMT [15]. Also, some local images are also used that are tested based on the PSNR of the noisy image compared with the denoised one.

It can be noticed from the table I that our proposed scheme gives a better performance in terms of PSNR values. We have compared our method in terms of PSNR at different noise variances. Our technique shows comparable results to that of other techniques. It can be noticed that Empirical Bayesian approach performs well when used in HMT domain. Comparable results are found in proposed work when the HMT is initialized by EM algorithm. Also, the standard image 'Lenna' is tested for visual quality. Using our proposed method, it has been shown in fig. 3 that image quality is comparable to other techniques with less blurring.

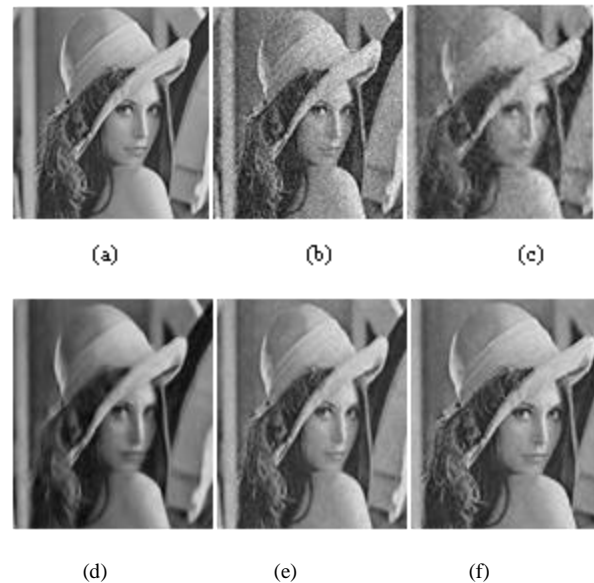


Fig. 3. Images of denoising experiment corresponding to first row of table I. (a) Original 256x256 'Lena' image (b) Noisy Image with $\sigma = 20$, PSNR=22.41 db. Image Denoised by (c) Wavelet Transform using Bayesian Approach [18], PSNR=26.02db (d) Curvelet Transform using Bayesian Approach [18], PSNR=28.49db (e) Empirical Bayesian Approach using HMT [15], PSNR=28.56db (f) Proposed Method, PSNR=28.74db

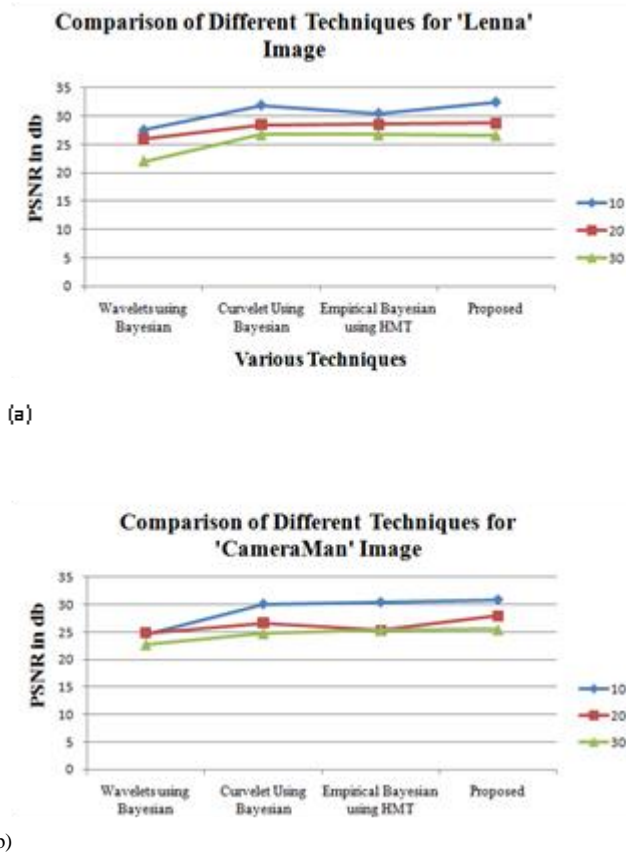


Fig. 4. Comparison of various denoising methods (a) 256x256 'Lenna' image (b) 256x256 'CameraMan' image

TABLE II. MEAN OPINION SCORE CLASSIFICATION

5	Excellent Quality
4	Good Quality
3	Fair Quality
2	Poor Quality
1	Bad Quality

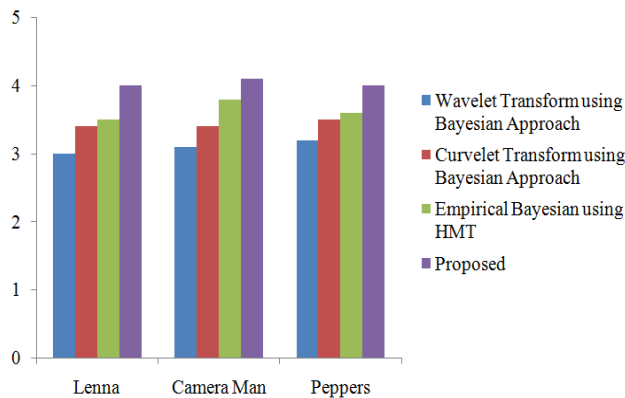


Fig. 5. Graphical Representation of Mean Opinion Score for Test Images

The different denoising techniques have been compared graphically as well. Fig.4, shows the graphical representation of different techniques at 3 different noise variances.

From many years MOS (Mean Opinion Score) method has been used for subjective analysis of an image, video or voice quality. This technique refers to the averaged value of the

opinions taken from the users [19]. This method scores the quality of the image from 1 (worse) to 5 (excellent). Table II shows different classes of MOS ranging from 1 to 5.

We have performed MOS analysis on three different images that are tested using different techniques in order to have an idea about visual quality of the images. Fig. 5, shows the results taken from a group of people and then averaged over the total number of observations. The graph shows that visual quality of images using our proposed technique has the comparable score to other techniques.

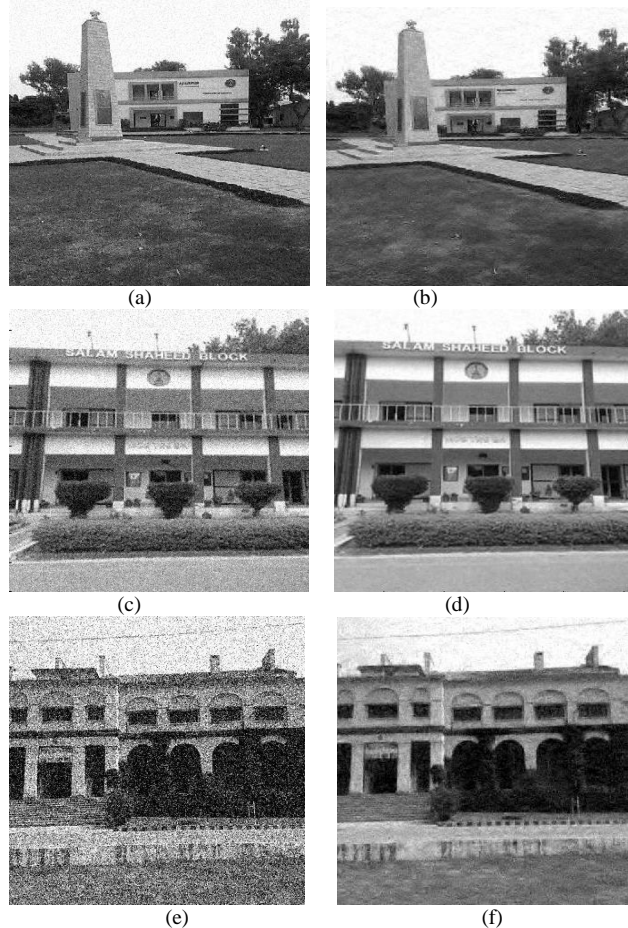


Fig. 6. Images of denoising experiment for some local images. (a) Noisy image with $\sigma = 10$, PSNR= 28.12db (b) Denoised Image, PSNR= 30.44 db. (c) Noisy Image with $\sigma = 20$, PSNR= 22.14db (d) Denoised Image, PSNR= 29.53db (e) Noisy image with $\sigma = 30$, PSNR= 18.62db (f) Denoised Image, PSNR= 25.00db

The denoising algorithm proposed is also applied on several images taken from the institute as well. These images are tested at different noise variances and noisy image is compared with the denoised image in terms of PSNR. Fig.6, shows the results taken after applying the algorithm.

IV. CONCLUSION

We have proposed HMM based image denoising method in Time-Frequency analysis domain. It has been shown that the proposed algorithm outperforms the existing few techniques based on PSNR values. The proposed framework concisely models the non-Gaussian statistics of the individual wavelet

coefficients. The EM algorithm used for training the HMT model captures the statistical dependencies between the coefficients that are tested using some standard and local images comparable to other techniques. This technique gives comparable results as compared to HMT that uses Bayesian approach to model the coefficients. Our method shows improvement of the denoised images as compared to other techniques based on both the PSNR values and in terms of visual quality of the images.

ACKNOWLEDGMENT

This research work has been facilitated by the Image Processing cell at Military College of Signals, National University of Sciences and Technology, Islamabad.

REFERENCES

- [1] R.C. Gonzalez and R.E Woods: 'Digital Image Processing', Prentice Hall, Upper Saddle River, N.J., Second edition, 2002
- [2] M. S. Crouse, R. D. Nowak and R. G. Baraniuk: 'Wavelet-based statistical signal processing using hidden Markov models', IEEE Transaction on Signal Process., vol.46, no.4, pp. 886 -902, 1998.
- [3] G. Y. Chen and T. D. Bui: 'Multi wavelet Denoising using Neighboring Coefficients', IEEE Signal Process. Lett., vol. 10, pp. 211-214, 2003.
- [4] Jyri J.Kivinen, Erric B. Sudderth, Micheal I. Jordan, "Image Denoising with nonparametric Hidden Markov Trees", 2007
- [5] D. Cho and T. D. Bui: 'Multivariate statistical modeling for image denoising using wavelet transforms', Signal Processing: Image Communication, vol. 20, pp. 77-89, 2005.
- [6] Minghui Yang, Zhiyun Xiao, Silong Peng, "A wavelet Domain Hidden Markov Tree Model with Localized Parameters for image denoising", 2006
- [7] D. Cho, T. D. Bui and G. Y. Chen: 'Image denoising based on wavelet shrinkage using neighbour and level dependency', International Journal of Wavelets, Multiresolution and Information Processing, vol. 7, no. 3, pp. 299-311, 2009.
- [8] G. Y. Chen, W. P. Zhu and W. F. Xie: 'Wavelet-based image denoising using three scales of dependency', IET Image Processing, vol. 6, no. 6, pp. 756-760, 2012.
- [9] G. Y. Chen, T. D. Bui and A. Krzyzak: 'Image Denoising using neighboring wavelet coefficients', Integrated Computer-Aided Engineering, vol. 12, no. 1, pp. 99-107, 2005.
- [10] J. Ho and W. L. Hwang, "Wavelet Bayesian network image denoising," *IEEE Transaction on Image Processing*, vol. 22, no. 4, pp. 1277-1290, 2013
- [11] Guoliang Fan and Xiang-Gen Xia: 'Wavelet-based Texture Analysis and Synthesis Using Hidden Markov Models', IEEE Transactions on Circuits and Systems-I, Fundamental Theory and Applications, vol. 50, no. 1, January 2003
- [12] Gilbert Strang and Truong Nguyen: 'Wavelets and Filter Banks', Wellesley-Cambridge Press, 1997
- [13] R. D. Nowak, 'Multiscale hidden Markov models for Bayesian image analysis', in Bayesian Inference in Wavelet Based Models, P. Müller and B. Vidakovic, Eds. New York: Springer Verlag, 1999, pp. 243-266.
- [14] H.Choi and R. Baraniuk: 'Multiscale image segmentation using wavelet-domain hidden Markov models', IEEE Trans. Image Processing, vol. 10, pp. 1309-1321, Sept. 2001.
- [15] J. K. Romberg, H. Choi, and R. G. Baraniuk: 'Bayesian tree-structured image modeling using wavelet-domain hidden Markov models', IEEE Trans. Image Processing, vol. 10, pp. 1056-1068, July 2001.
- [16] H. Chipman, E. Kolaczyk, and R. McCulloch: 'Adaptive Bayesian wavelet shrinkage', J. Amer. Stat. Assoc., vol. 440, no. 92, pp. 1413-1421, Dec. 1997
- [17] M. Amini, M.O. Ahmad and M.N.S. Swamy, "Image denoising in wavelet domain using the Vector-based hidden Markov model," IEEE12th Inter. New Circuits and Systems Conf. (NEWCAS), pp. 29-32, 2014.
- [18] Pallavi Sharma, R.C. Jain, Rashmi Nagvani, "An efficient Curvelet Bayesian Network approach for Image Denoising," IEEE International Conference on Advances in Engineering & Technology, 2014
- [19] Anna Geomi George, A. Kethsy Prabavathy, "A Survey on Different Approaches used in Image Quality Assesment", International Journal of Emerging technology and Advanced Engineering, vol.3, Issue 2, February 2013.