

WAVELET DOMAIN WATERMARK DETECTION
AND EXTRACTION USING THE VECTOR-BASED
HIDDEN MARKOV MODEL

Marzieh Amini

A Thesis
in
The Department
of
Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada

14 September 2016

© Marzieh Amini, 2016

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Marzieh Amini

Entitled: Wavelet domain watermark detection and extraction using the vector-based hidden Markov model

and submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY (Electrical & Computer Engineering)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. D. Dysart-Gale

_____ External Examiner
Dr. W.B. Mikhael

_____ External to Program
Dr. C.Y. Su

_____ Examiner
Dr. H. Rivaz

_____ Examiner
Dr. W.P. Zhu

_____ Supervisor
Dr. M. O. Ahmad

_____ Supervisor
Dr. M.N.S. Swamy

Approved by _____
Dr. A. R. Sebak

September 14, 2016 _____
Dr. A. Asif

ABSTRACT

WAVELET DOMAIN WATERMARK DETECTION AND EXTRACTION USING THE VECTOR-BASED HIDDEN MARKOV MODEL

Marzieh Amini,

Concordia University, 2016.

Multimedia data piracy is a growing problem in view of the ease and simplicity provided by the internet in transmitting and receiving such data. A possible solution to preclude unauthorized duplication or distribution of digital data is watermarking. Watermarking is an identifiable piece of information that provides security against multimedia piracy. This thesis is concerned with the investigation of various image watermarking schemes in the wavelet domain using the statistical properties of the wavelet coefficients. The wavelet subband coefficients of natural images have significantly non-Gaussian and heavy-tailed features that are best described by heavy-tailed distributions. Moreover the wavelet coefficients of images have strong inter-scale and inter-orientation dependencies. In view of this, the vector-based hidden Markov model is found to be best suited to characterize the wavelet coefficients. In this thesis, this model is used to develop new digital image watermarking schemes. Additive and multiplicative watermarking schemes in the wavelet domain are developed in order to provide improved detection and extraction of the watermark. Blind watermark detectors using log-likelihood ratio test, and watermark

decoders using the maximum likelihood criterion to blindly extract the embedded watermark bits from the observation data are designed.

Extensive experiments are conducted throughout this thesis using a number of databases selected from a wide variety of natural images. Simulation results are presented to demonstrate the effectiveness of the proposed image watermarking scheme and their superiority over some of the state-of-the-art techniques. It is shown that in view of the use of the hidden Markov model characterize the distributions of the wavelet coefficients of images, the proposed watermarking algorithms result in higher detection and decoding rates both before and after subjecting the watermarked image to various kinds of attacks.

ACKNOWLEDGEMENTS

I would like to thank those who made this PhD thesis possible: my advisors, family and friends. First and foremost, I would like to thank my advisors, Professor M. Omair Ahmad and Professor M. N. S. Swamy, whose expertise, continuous guidance, and understanding helped me during the course of my PhD studies. They gave me freedom to explore on my own, and at the same time guided me to remain focused. The time spent in our meetings has truly enriched my experience at Concordia. I could not have wished for better and friendlier supervisors.

Most importantly, I would like to thank my best friend and beloved husband, Hamidreza, whose encouragement, patience and unwavering love has been the driving force throughout my PhD. A big hug and a heartfelt “thanks” to you. I also thank my mom and dad, for their faith in me and for supporting me in every way to be as ambitious as I wanted. Their sacrifices have really brought me here.

TABLE OF CONTENTS

TABLE OF CONTENTS	vi
LIST OF FIGURES.....	ix
LIST OF TABLES.....	xvii
LIST OF SYMBOLS.....	xx
LIST OF ABBRIVIATIONS.....	xxii
CHAPTER 1: Introduction	1
1.1 General.....	1
1.2 A Brief Literature Review of Statistical Watermark Detection and Decoding.....	4
1.3 Objectives.....	6
1.4 Organization of the Thesis.....	7
CHAPTER 2: Modelling of Image Wavelet Coefficients and an Introduction to Image Watermarking	9
2.1 Introduction.....	9
2.2 Discrete Wavelet Transform.....	10
2.3 Statistical Properties of the Wavelet Coefficients.....	13
2.3.1 Wavelet Marginal Models.....	14
2.3.2 Wavelet Joint Models.....	17
2.3.3 Inter-scale and Intra-scale Dependencies.....	18
2.4 Hidden Markov Model.....	20
2.5 Results of Modeling the Wavelet Coefficients Using the Vector-based HMM.....	27
2.6 Watermarking.....	31

2.7 Summary	34
CHAPTER 3: Locally-optimum Detector and Optimum Decoder for Additive Watermarking Schemes	35
3.1 Introduction	35
3.2 Locally Optimum Watermark Detector	37
3.2.1 Watermark Embedding	37
3.2.2 Watermark Detector	38
3.2.3 Performance Analysis of Vector-based HMM Detector	42
3.3 Watermark Decoder	49
3.3.1 Error Analysis	52
3.4 Experimental Results	55
3.4.1 Watermark Detection Results	56
3.4.2 Watermark Decoder Results	70
3.6 Summary	81
CHAPTER 4: Optimum Multiplicative Watermark Detector and Decoder	82
4.1 Introduction	82
4.2 Watermarking Detection	84
4.2.1 Watermark Embedding	84
4.2.2 Watermark Detection	85
4.3 Watermarking Decoder	88
4.3.1 Watermark Embedding	88
4.3.2 Watermark Decoder	89
4.3.3 Error Analysis	92
4.4. Simulation Results	95

4.4.1 Detection Results	97
4.4.2 Decoder Results	104
4.5. Summary	132
CHAPTER 5: Conclusion	133
5.1 Concluding Remarks.....	133
5.2 Scope for Future Work.....	134
REFERENCES	136

List of Figures

Figure 2.1:	Two-level DWT subband representation.....	13
Figure 2.2:	Definition of PX , NX and CX . For a wavelet coefficient X , PX is its parent in the coarser band, NX is its neighborhood and CX is its cousins at the same level but in different orientations.....	19
Figure 2.3:	(a) Scalar and (b) Vector-based HMM.....	23
Figure 2.4:	PDFs of the empirical data as well as vector-based HMM, Cauchy, GG and BKF distributions for second level of the wavelet transform of Lena image. a) LH, a) HL and c) HH.....	30
Figure 3.1:	Block diagram of the proposed watermark embedding procedure.....	38
Figure 3.2:	Locally-optimum watermark detection scheme using the vector-based HMM.....	42
Figure 3.3:	Block diagram of the embedding procedure of the proposed watermark decode.....	50
Figure 3.4:	Block diagram of the decoding procedure of the proposed watermark decoder.....	51
Figure 3.5:	(a)-(e) Original test images, (f)-(g) Watermarked images corresponding to the test images in (a)-(e) for WDR = -34 dB.....	57
Figure 3.6:	Theoretical (solid) and experimental (dashed) ROC curves averaged over 96 test images for the vector-based HMM detector for different values of WDR.....	61

Figure 3.7:	ROC curves averaged over 96 test images for various detectors when WDR = -34 dB.....	61
Figure 3.8:	Boxplots of the probability of detection for various detectors averaged over a number of test images when WDR = -34 dB and P_{FA} ranging from 10^{-8} to 10^{-2}	63
Figure 3.9:	ROC curves averaged over a number of test images obtained using various detectors for WDR = -34 dB when images are JPEG-compressed with QF = 30.....	64
Figure 3.10:	ROC curves averaged over a number of images obtained using various detectors and WDR = -34 dB when images are rotated by 2°	64
Figure 3.11:	ROC curves averaged over a number of test images obtained using various detectors for WDR = -34 dB when images undergo median filtering with a window size of 3×3	65
Figure 3.12:	ROC curves averaged over a number of test images for the various detectors for WDR = -34 dB when Gaussian filtering with window mask size of 3×3	65
Figure 3.13:	ROC curves averaged over a number of test images obtained using various detectors and WDR = -34 dB when images undergo Gaussian noise with SNR = 25 dB.....	66
Figure 3.14:	Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image JPEG compressed with QF = 30.....	67

Figure 3.15:	Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image is rotated by 2 degree.....	67
Figure 3.16:	Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image undergoes median filtering with a window size of 3×3.....	68
Figure 3.17:	Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when Gaussian filtering with mask 3×3.....	68
Figure 3.18:	Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image undergoes Gaussian noise with SNR = 25 dB.....	69
Figure 3.19:	Theoretical and experimental BER of the proposed decoder averaged over a number of test images with message length of 128 bits for different WDR values.....	70
Figure 3.20:	BER of the extracted watermark obtained using the proposed vector-based HMM decoder when watermarked images, are JPEG-compressed with different quality factors.....	72
Figure 3.21:	BER of the extracted watermark using the proposed vector-based HMM decoder when watermarked images are contaminated by the AWGN with different noise levels.....	73

Figure 3.22:	BER of the extracted watermark using the proposed vector-based HMM decoder when watermarked images, Barbara, Baboon, Peppers and Lena are rotated by different angles.....	73
Figure 3.23:	BER (%) obtained using the proposed vector-based HMM based watermarking scheme as well as that obtained using the schemes in [36], [68] and [69], under AWGN with different noise standard deviations for the <i>Lena</i> image. (Message length = 128 bits, PSNR = 45 dB).....	77
Figure 4.1:	Proposed multiplicative watermarking scheme; embedding and detection parts.....	87
Figure 4.2:	The proposed multiplicative watermark decoder, embedding scheme.....	89
Figure 4.3:	The proposed multiplicative watermark decoder scheme using the vector-based HMM.....	91
Figure 4.4:	Original images (a-e) and corresponding watermarked (f-j) images obtained using the proposed multiplicative watermarking scheme for WDR = -42 dB.....	96
Figure 4.5:	Theoretical (solid) and experimental (dashed) ROC curves averaged over 96 test images for the multiplicative vector-based HMM detector for different values of WDR.....	97
Figure 4.6:	ROC curves averaged over 96 test images for various detectors when WDR = -50 dB.....	98
Figure 4.7:	ROC curves averaged over a number of images obtained using various detectors when image is compressed by QF = 30 and WDR = -50 dB....	99

Figure 4.8:	ROC curves averaged over a number of images obtained using various detectors and WDR = -50 dB when images are rotated by 2°	100
Figure 4.9:	ROC curves averaged over a number of test images obtained using various detectors for WDR = -50 dB when images undergo median filtering with a window size of 3×3	100
Figure 4.10:	ROC curves averaged over a number of images for the various detectors for WDR = -50dB when Gaussian filtering with mask 3×3	101
Figure 4.11:	ROC curves averaged over a number of test images obtained using various detectors when image is contaminated by the Gaussian noise with $SNR = 25$ and WDR = -50 dB.....	101
Figure 4.12:	Theoretical and experimental BER of the proposed decoder averaged over a number of test images with message length of 128 bits for different WDR values.....	105
Figure 4.13:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when, the <i>Lena</i> image is JPEG-compressed with different QFs.....	108
Figure 4.14:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when, the <i>Baboon</i> image is JPEG compressed with different QFs.....	108
Figure 4.15:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Peppers</i> image is JPEG-compressed with different QFs.....	109

Figure 4.16:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Barbara</i> image JPEG-compressed with different QFs.....	109
Figure 4.17:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Boat</i> image is JPEG-compressed with different QFs.....	110
Figure 4.18:	BER values of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy decoders when the images are JPEG-compressed with different QFs.....	110
Figure 4.19:	BER of the extracted watermark using the proposed VB-HMM, GG and Cauchy decoders when the <i>Lena</i> image is corrupted by the additive Gaussian noise with different σ_n values.....	111
Figure 4.20:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when the <i>Baboon</i> image is corrupted by the additive Gaussian noise with different σ_n values.....	112
Figure 4.21:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Peppers</i> image is corrupted by additive Gaussian noise with different σ_n values.....	112
Figure 4.22:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Barbara</i> image is corrupted by the additive Gaussian noise with different σ_n values.....	113

Figure 4.23: BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the *Boat* image is corrupted by the additive Gaussian noise with different σ_n values.....113

Figure 4.24: BER of the extracted watermark averaged over 96 test images using the proposed VB-HMM, GG and Cauchy decoders, when the images are corrupted by additive Gaussian noise with different σ_n values114

Figure 4.25: BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when the *Lena* image undergoes median filtering with different window sizes.....116

Figure 4.26: BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when the *Baboon* image undergoes median filtering with different window sizes..... 117

Figure 4.27: BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when the *Peppers* image undergoes median filtering with different window sizes.....117

Figure 4.28: BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when the *Barbara* image undergoes median filtering with different window sizes.....118

Figure 4.29: BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when the *Boat* image undergoes median filtering with different window sizes.....118

Figure 4.30:	BER of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy-based decoders when the images undergo median filtering with different window size.....	119
Figure 4.31:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy-based decoders, when the <i>Lena</i> image is rotated with different angles.....	120
Figure 4.32:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Baboon</i> image is rotated with different angles.....	120
Figure 4.33:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Peppers</i> image is rotated with different angles.....	121
Figure 4.34:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Barbara</i> image is rotated with different angles.....	121
Figure 4.35:	BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when the <i>Barbara</i> image is rotated with different angles.....	122
Figure 4.36:	BER of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy-based decoders, when the images are rotated by different angles.....	122

List of Tables

Table 2.1:	Mutual information between the wavelet coefficients of the <i>Lena</i> image.....	20
Table 2.2:	KSD values between the empirical data and vector-based HMM with different number of states for the Barbara image.....	28
Table 2.3:	CPU time for modeling the wavelet coefficients with vector-based HMM with different number of states, for the Barbara image.....	29
Table 2.4:	KSD values between the empirical data and different distributions averaged over a number of images for the second level of the wavelet transform.....	31
Table 3.1:	Comparison of the vector-based HMM parameters of the second level of the wavelet coefficients of \mathbf{x} for the original image and the corresponding coefficients \mathbf{y} for the watermarked image for the test images (a) - (e) of Figure 3.5.....	58
Table 3.2:	Mean square difference between the corresponding elements of the covariance matrices C_q^m for \mathbf{x} and \mathbf{y} for the test images (a) - (e) of Figure 3.5.....	59
Table 3.3:	Averaged values of the area under ROC curves over 96 test images for various detectors and different values of WDR.	62
Table 3.4:	Mean and variance of the BERs (%) values of 96 test images obtained using the proposed watermarking scheme under various attacks for the different message lengths. (PSNR= 50 dB).....	74

Table 3.5:	BER (%) obtained using the proposed VB-HMM based watermarking scheme as well as that obtained using the schemes in [36] and [35], under various attacks for the <i>Lena</i> images. (Message length = 128 bits, PSNR = 45 dB).....	76
Table 3.6:	BER (%) obtained using the proposed watermarking scheme as well as that obtained using the schemes in [33], [35], [65]-[68], under various attacks when message length is 256 bits and PSNR = 42 dB.	78
Table 3.7:	BER values obtained using the proposed watermarking scheme as well as that obtained using the schemes in [51], when watermarked images are under various attacks. (Message length= 64 bits, PSNR = 42 dB).....	80
Table 4.1:	PSNR and area under ROC curves for the region [0 , 1], averaged over 96 test images for various detectors against different attacks with WDR = -40 dB.....	103
Table 4.2:	BER (%) obtained using various decoders for different test images, with message length 64 and 128 bits and WDR = -42 dB.....	106
Table 4.3:	BER (%) of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy decoders when various test images are corrupted by salt and pepper noise with different noise density.....	115
Table 4.4:	BER values obtained using different decoders when the images undergo gamma correction with different values of gamma.....	124
Table 4.5:	BERs (%) obtained using the proposed additive and multiplicative watermarking scheme as well as that obtained using the schemes in [33],	

	[37] and [35] under various attacks for the different test images. (Message length = 256 bits, PSNR = 42 dB).....	125
Table 4.6:	BER (%) using the proposed additive and multiplicative watermarking schemes as well as that obtained using the schemes in [53] and [35], when watermarked images are under various attacks for the different test images. (Message length = 128 bits).....	127
Table 4.7:	BER (%) obtained using the proposed multiplicative watermarking scheme as well as that obtained using the schemes in [36] and [35], when watermarked <i>Lena</i> image is under various attacks. (Message length=128 bits, PSNR = 45 dB).....	128
Table 4.8:	BER values obtained using the proposed multiplicative watermarking scheme as well as that obtained using the schemes in [37], [35], [61] and [51], when watermarked images are under various attacks. (Message length = 64 bits, PSNR = 42 dB).....	129
Table 4.9:	Extracted watermark logo for watermarked <i>Lena</i> image of size 512× 512 with multiplicative decoder in presence of different attacks when message length = 1024 bits.....	131

List of Symbols

H_1	Alternative hypothesis
f_J^L	Approximation coefficients
f_{BKF}	Bessel K-Form probability density function
P_e	Bit error probability
B	Block
f_{Cauchy}	Cauchy probability density function
C^m	Covariance matrix of the m^{th} state
f_l^H	Detail coefficients
f_X	Distribution of the vector-based HMM
$\Gamma(\cdot)$	Gamma function
$N(\cdot)$	Gaussian distribution
f_{GG}	Generalized Gaussian probability density function
Θ	Hidden Markov model
S	Hidden state
I	Host image
$f_{XY}(x, y)$	Joint PDF of x and y
$D(\ \cdot\)$	Kullback-Liebler distance
J	Largest level of wavelet
g_{LO}	Locally optimum nonlinearity

$l(\mathbf{y})$	Log-likelihood ratio
Λ	Maximum likelihood
μ^m	Mean of the m^{th} state
b_k	Message bit
$I(X, Y)$	Mutual information between X and Y
H_0	Null hypothesis
N_b	Number of blocks
L	Number of coefficients in the selected subband
M	Number of states
p^m	Probability mass function of the m^{th} state
P_{Det}	Probability of detection
P_{FA}	Probability of false alarm
$\phi(\cdot)$	Scaling function
β	Shape parameter of generalized Gaussian distribution
N	Size of image
A_{ij}	State transition probability matrices
τ	Threshold level
σ^m	Variance of m^{th} state
\mathbf{w}	Watermark bits
$\psi(\cdot)$	Wavelet function
α	Weighting factor

List of Abbreviations

1-D	One-dimensional
2-D	Two-dimensional
AWGN	Additive white Gaussian noise
BER	Bit error rate
BKF	Bessel K-form
DCT	Discrete cosine transform
DFT	Discrete Fourier transform
DWT	Discrete wavelet transform
EM	Expectation maximization
GG	Generalized Gaussian
GMM	Gaussian scale mixture
HMM	Hidden Markov models
JPEG	Joint Photographic Experts Group
KL	Kullback-Liebler
KSD	Kolmogorov-Smirnov distance
LO	Locally-optimum
MRF	Markov random field
PSNR	Peak signal to noise ratio
QF	Quality factor
RGB	Red, Green, Blue
ROC	Receiver operating characteristic
WDR	Watermark-to-document ratio

CHAPTER 1

Introduction

1.1 General

Images, like other signals, have certain features that characterize them. In statistical modeling, it is intended to capture these characteristics using a small number of parameters. A model attempts to capture the key characteristics of an image based on which image processing problems can be formulated and solved mathematically and systematically. In image denoising applications, for example, if an image is corrupted by a certain type of noise, the original image can be restored based on a predefined model of the image. In a classification application, the type of a given textured region can be identified by the use of a texture model, which can effectively specify different textures. In an image compression application, an efficient prediction scheme can be devised to encode an image by taking advantage of an accurate image model. Thus, image models play an important role in image processing applications. However, modeling in spatial domain is problematic, since images have large dimensions and are hard to be statistically measured [1]. In addition, in spatial domain pixels are highly dependent on one another and thus, modeling an image with only a few parameters is a difficult task. In recent years, statistical image modeling has been focused mostly on transform domains such as in the discrete Fourier transform (DFT), discrete cosine transform (DCT) and discrete wavelet transform (DWT) domains, in which energy density has more local structure. Among all the transforms, the wavelet transform has drawn more attention, due to its

superior performance and multiresolution properties over that of DFT and DCT. Recently, statistical models for the wavelet coefficients of images have been employed in many areas of image processing, such as denoising [2]-[7], encoding [8], [9], compression [10], classification [11], [12], image retrieval [13] and watermarking [14]-[38]. The wavelet subband coefficients have been previously assumed to be independent and modeled simply by marginal statistics such as the Gaussian [39], generalized Gaussian (GG) [40], [20]-[22], [26], Cauchy [40], [41], [26], alpha-stable [41], [41], Gauss-Hermite [15] and Bessel K-form (BKF) [18], [19], [42] distributions. However, marginal probability density functions (PDFs) cannot capture adequately the dependency of the wavelet coefficients in a single subband or between subbands and therefore, such PDFs cannot be made to fit well the empirical PDF of the wavelet coefficients. It is known that the wavelet coefficients of images have strong dependencies across the scales. In view of this, joint statistical models, such as hidden Markov models (HMMs) [1], [5]-[7], [11]-[12], [29]-[32], [43]-[47] and Markov random field (MRF) priors [49], have been proposed in order to capture the inter-scale dependencies of the wavelet coefficients. Hidden Markov model for modeling the wavelet coefficients has been proposed in [43] to solve an image denoising problems. This model in the wavelet domain was later employed in segmentation [11] and texture retrieval [46] problems.

During the last decade, due to the development of the internet, distribution of digital multimedia data to a large number of users has been increasingly growing. However, duplication and manipulation of the media data can be easily made without any noticeable quality loss. Consequently, violation of intellectual property rights especially on the internet has become a greater concern. Data hiding techniques such as digital

multimedia watermarking are proposed to prevent online piracy. Watermarking is a way of embedding a secret message into the original data in order to protect intellectual property and facilitate copyright protection. In other words, digital watermarks covertly embed a message into the data for the purpose of ownership verification or tracing the copyright infringements.

Digital image watermarking has been extensively studied in the literature during the past decades [14]-[38], [49]-[69], [71], [115]-[122]. Digital watermarking techniques can be classified in many ways such as embedding method, embedding domain, perceptibility, robustness and reversibility. According to the domain used for embedding the watermark, image watermarking algorithms can be classified into two categories: spatial [49] and frequency [14]-[41], [50]-[55], [58]-[62]. In spatial domain, image pixels are directly modified to imperceptibly embrace a piece of information. On the other hand, in frequency domain, image pixels are first projected into lower dimensional bases and the resulting coefficients are then modified. It is known that the frequency domain watermarking techniques can provide greater robustness in comparison to their spatial domain counterparts [25].

In some applications of watermarking, it may only be necessary to determine whether a specific watermark is present or not in the received signal [14]-[30], [39]-[41], [52], [60], [62], whereas in the others, the embedded watermark is considered as a hidden unknown message that needs to be decoded accurately [31]-[39], [50], [51], [56]-[58], [62], [64]-[69]. Depending on the detection methods, existing watermarking schemes can be classified into two categories: informed detection and blind detection depending on whether or not the host signal is available at the detector during the watermark detection

process. In informed detection techniques, in order to prevent the interferences of the original image on the watermark message, some side information is transmitted to the receiver [56], [57], [71]. On the other hand, blind watermark detection techniques have no knowledge about the original image at the receiver [15]-[32], [52]. In order for blind watermark detection or decoding to be realized, advantage can be taken of the statistical properties of the image. Efforts in this direction have been mostly on the statistical modeling of images in the frequency domain [14]-[32], [52].

1.2 A Brief Literature Review of Statistical Watermark Detection and Decoding

In the case of model-based watermark detection algorithms, most of the existing methods in the wavelet domain are based on the assumption that the wavelet coefficients follow the Gaussian distribution, so that the common correlation detector can be used for the purpose of detection. However, correlation-based detectors are not optimal for non-Gaussian data, and in addition, they ignore the dependencies among the wavelet coefficients. In view of this, optimal or locally-optimum (LO) detectors based on the signal statistics have been proposed and shown to provide significantly better detection results than that provided by the correlation-based detectors in various transform domains [70]. In [21] and [22], LO detectors have been designed for watermarking schemes in which the DFT, DCT or DWT coefficients of images have been modeled by the GG distribution. In [39], the GG modeling has been used for the DCT coefficients of images, and a detector has been designed based on the maximum likelihood decision rule. In [20], a LO detector has been developed using GG modeling. In [40], a LO watermark detector

has been designed by modeling the DCT coefficients of images using the Cauchy distribution. In [26], the Cauchy and GG PDFs are applied to model the detail subband coefficients of DWT. In [18], a LO watermark detector has been proposed in which the BKF distribution is used for modeling the DWT coefficients.

There exist several schemes focusing on watermark decoding using the statistical properties of transformed domain coefficients. In [39], additive watermarking has been performed in the DCT domain, and the decoding has been performed by using the GG distribution as a prior model for the DCT coefficients. In [64], an optimum decoder for multiplicative watermark has been proposed in the DFT domain using the Weibull distribution, in which the performance of the decoder has been evaluated by Monte Carlo simulations. In [37], a scaling-based watermarking in the wavelet domain has been proposed by assuming a Gaussian distribution for modeling the wavelet coefficients. In [38], a multiplicative watermarking decoder has been proposed for a fingerprint application in the wavelet domain using the GG distribution. In [36], a quantization-based method has been proposed in the logarithmic domain. In [35], a robust quantization-based image watermarking has been proposed in which the watermark bits are embedded by quantizing the angles of significant gradient vectors in the wavelet domain.

Many of the watermarking schemes mentioned above for detecting or extracting the watermark bits are wavelet-domain based. These works have mostly focused on the marginal statistical properties of the wavelet coefficients of images to develop watermark detectors or decoders. However, the marginal PDFs are not the best choices for modeling the wavelet coefficients of images as such PDFs ignore the inter-scale and inter-orientation dependencies. The use of the marginal models may result in parameter

estimation with relatively lower precision, watermark detection with lower detection rate and a watermark decoder with higher bit error rate. Since the performance of such techniques is highly dependent on the accuracy of the model employed.

1.3 Objectives

The objective of this thesis is to investigate the suitability of the vector-based hidden Markov model in characterizing the distribution of the wavelet coefficients of images and to develop robust watermark detection and extraction schemes based on this model. The main focus of this study is on enhancing the rate of watermark detection and extraction by taking advantage of vector-based HMM in capturing the subband marginal distribution and the inter-scale and cross-orientation dependencies of the wavelet coefficients of images.

In the first part of the thesis, a locally-optimum additive watermark detector and decoder using the vector-based HMM in the wavelet domain are proposed. A formulation for watermark detection is derived using the log-likelihood ratio test. A closed-form expression for the test statistics of the receiver operating characteristic curve of the proposed detector is obtained for a low-complexity detection of the possible presence of a watermark in the original image. An optimum additive watermark decoder is next designed by using the maximum likelihood criterion to extract the hidden watermark message from the watermarked image. The performances of the proposed watermark detector and decoder are comprehensively investigated and the robustness of the proposed watermarking scheme against various known distortions is also studied.

The second part of the thesis deals with the development of vector-based HMM multiplicative watermarking schemes. Closed-form expressions for the test statistics of the proposed watermark detector and decoder are derived. The theoretical bit error rate is also obtained and validated experimentally. The performance of the proposed detector and decoder is investigated through several experiments and robustness of the proposed scheme is examined when the watermarked images are subjected to various kinds of distortions.

1.4 Organization of the Thesis

The thesis is organized as follows:

In Chapter 2, a brief review of the discrete wavelet transform is presented. Statistical properties of the wavelet coefficients of images are studied. Different statistical distributions for modeling the wavelet coefficient are presented. The suitability of the vector-based HMM for modeling the wavelet coefficients of images is comprehensively studied. The basic idea and mechanism generally used in image watermarking problem is briefly introduced.

In Chapter 3, based on the modeling results, a novel blind additive image watermarking scheme in the wavelet domain is proposed. A locally-optimum watermark detector and an optimum watermark decoder using the vector-based HMM in the wavelet domain are proposed. In a Bayesian framework, closed-form expressions for the mean and variance of a test statistics are derived, experimentally validated and used in evaluating the performance of the proposed detector. The watermark decoder is designed based on the

maximum likelihood criterion and theoretical expression for bit-error-rate is derived. The performance of the proposed additive detector and decoder is evaluated for the additive embedding of the watermark using a number of test images.

In Chapter 4, new schemes for blind multiplicative watermark incorporating the vector-based HMM in the wavelet domain are devised. A watermark detector is developed and the theoretical expressions for its test statistics are derived. A watermark decoder based on the maximum likelihood criterion is designed and closed-form expression for the bit error rate is derived and validated experimentally with Monte Carlo simulations. The performance of the proposed multiplicative watermark detector and decoder is evaluated using image datasets. A performance comparison of the proposed additive and multiplicative detectors and decoders is also carried out in this Chapter.

Finally, in Chapter 5, some concluding remarks and scope for further research are presented.

CHAPTER 2

Modelling of Image Wavelet Coefficients and an Introduction to Image Watermarking

2.1 Introduction

Wavelet transform has been developed as a powerful tool for signal analysis and processing [72]-[77]. The wavelet domain provides a natural setting for processing of signals such as their estimation, detection, classification, compression and watermarking. Properties of the wavelet transform including multiresolution, localization and compression, have led to the development of powerful signal processing methods. In image watermarking applications, the wavelet transform has gained considerable popularity due to a number of advantages offered by it. The main features of this transform that makes it specifically suited to watermark applications are as follows [14], [80]-[82]:

- *Localization in Space-frequency Domain*: The wavelet transform is capable of analyzing image features in view of its time-frequency localization property. This may increase the robustness of watermarking technique against the geometric distortions.
- *Multi-resolution representation*: Multiresolution analysis highlights the local and global properties of an image, which are of significant importance in embedding and detecting of specific watermarks.

- *Human visual system modeling*: A watermarking scheme can benefit from the desirable feature of match of the wavelet transform to the characteristics of HVS such as its frequency-sensitivity. For instance, since the human eye is less sensitive to the high frequencies, the watermark can be masked into these frequency bands with a suitable choice of watermark strength that is based on the local sensitivity of the image to the watermark bits [83], [84].
- *Linear complexity*: Wavelet domain watermarking schemes require computational costs lower than that of the Fourier or cosine transform domain-based schemes.

This chapter starts with the basics of the wavelet transform and the statistical properties of the wavelet coefficients. Common probabilistic models describing the wavelet statistics and their associated issues are presented. The vector-based HMM as the most powerful model to characterize the wavelet coefficients of images is discussed. The results of modeling the wavelet coefficients using this model are presented [6], [29], [30]. Finally, the basic mechanism of image watermarking is introduced.

2.2 Discrete Wavelet Transform

The one-dimensional (1-D) and two-dimensional (2-D) wavelets are defined as follows.

a. One-dimensional DWT

Let $I(x)$ represent a 1-D signal with length of N . The DWT of the signal is given by [78]

$$I(x) = N^{-\frac{1}{2}} \left[\sum_{i=1}^N f_J^L(i) \phi_J^L(x,i) + \sum_{l=1}^J \sum_{i=1}^N f_l^H(i) \psi_l^H(x,i) \right] \quad (2.1)$$

where f_J^L denotes the approximation coefficients in the largest level J , f_l^H denotes the detail coefficients in the decomposition level l , $l=1,2,\dots,J$. The wavelet $\psi(t)$ and the scaling functions ϕ are defined as

$$\begin{aligned} \phi_J^L(x,i) &= 2^{\frac{J}{2}} \phi(2^J x - i) \\ \psi_j^H(x,i) &= 2^{\frac{j}{2}} \psi(2^j x - i) \end{aligned} \quad (2.2)$$

The approximation and detail coefficients are then given by [75]

$$\begin{aligned} f_J^L(i) &= N^{-\frac{1}{2}} \sum_{x=1}^N I(x) \phi_J^L(x,i) \\ f_l^H(i) &= N^{-\frac{1}{2}} \sum_{x=1}^N I(x) \psi_l^H(x,i) \end{aligned} \quad (2.3)$$

In order to have a perfect reconstruction, the functions $\phi(\cdot)$ and $\psi(\cdot)$ are chosen in a way that the signal can be reconstructed from the wavelet coefficients without any difference.

b. Two-dimensional DWT

Let $I(x,y)$ represent a 2-D signal with size of $N_1 \times N_2$. The DWT of the 2-D signal is given by [78]

$$I(x, y) = (N_1 N_2)^{-\frac{1}{2}} \left[\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} f_J^{LL}(i, j) \phi_J^{LL}(x, y, i, j) + \sum_{l=1}^J \sum_{d \in \{LH, HL, HH\}} \sum_{j=1}^{N_1} \sum_{i=1}^{N_2} f_l^d(i) \psi_l^d(x, y, i, j) \right] \quad (2.4)$$

where $f_J^d, d \in \{LH, HL, HH\}$ represents the detail coefficients in the level l of orientation d , and $\phi(\cdot)$ and $\psi(\cdot)$ are the 2-D scaling and wavelet functions, respectively.

The 2-D scaling and wavelet functions can be represented using 1-D scaling and wavelet functions as

$$\begin{aligned} \phi_J^{LL}(x, y, i, j) &= 2^{\frac{J}{2}} \phi(2^J x - i) \phi(2^J y - j) \\ \psi_J^{LH}(x, y, i, j) &= 2^{\frac{J}{2}} \psi(2^J x - i) \phi(2^J y - j) \\ \psi_J^{HL}(x, y, i, j) &= 2^{\frac{J}{2}} \phi(2^J x - i) \psi(2^J y - j) \\ \psi_J^{HH}(x, y, i, j) &= 2^{\frac{J}{2}} \psi(2^J x - i) \psi(2^J y - j) \end{aligned} \quad (2.5)$$

Accordingly, the 2-D approximation and detail representation of the wavelet coefficients can be obtained as

$$\begin{aligned} f_J^A(i, j) &= (N_1 N_2)^{-\frac{1}{2}} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} I(x, y) \phi_J^{LL}(x, y, i, j) \\ f_l^d(i, j) &= (N_1 N_2)^{-\frac{1}{2}} \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} I(x, y) \psi_l^d(x, y, i, j), d \in \{LH, HL, HH\} \end{aligned} \quad (2.6)$$

These coefficients can be grouped into different subbands and orientations. The detail coefficients of LH_l , HL_l and HH_l subbands are represented as horizontal f_l^{LH} , vertical

f_i^{HL} and diagonal f_i^{HH} orientations, respectively. Figure 2.1 shows a typical two-level DWT.

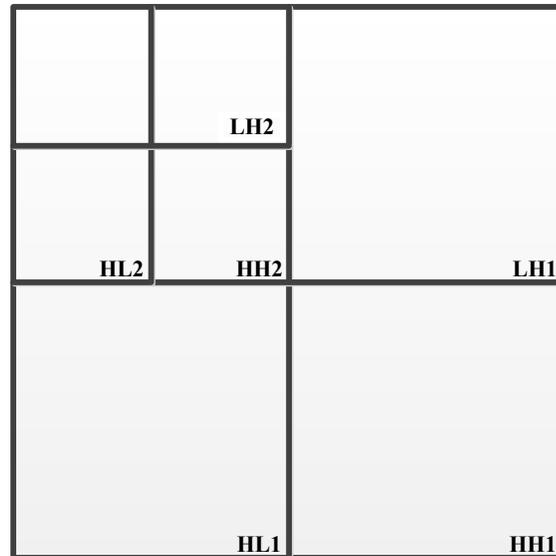


Figure 2.1 Two-level DWT subband representation.

2.3 Statistical Properties of the Wavelet Coefficients

The statistical properties of the wavelet coefficients play an important role in many image processing algorithms, such as watermarking. By exploiting the histograms of the wavelet coefficients, researchers have incorporated a number of distributions to study the suitability of these models. Wavelet transform has many properties that make it attractive in image processing applications. The primary properties of the wavelet coefficients are as follows [43], [44].

Locality: The wavelet coefficients represent the image content which are localized simultaneously in both the spatial location and frequency domains.

Multiresolution: With the use of wavelet transform, an image can be represented by a nested set of scales.

Compression: The wavelet coefficients of natural images are sparse.

The locality and multiresolution properties result in a quad-tree structure of the wavelet coefficients with three subband orientations in each scale. According to the compression property, an image can be approximated by only a few wavelet coefficients with large magnitudes.

Beside these primary features, the wavelet transform has the following important properties [43]-[44], [85]-[90].

Non-Gaussianity: Wavelet coefficients have peaky, heavy-tailed marginal distributions [85], [86].

Persistence across scales: Large/small values of wavelet coefficients tend to spread across scales [87]-[90].

In view of the above properties of the wavelet transform, statistical properties of the wavelet coefficients and their modeling are of great importance in many estimation and detection algorithms in image processing applications.

There exist several works studying the wavelet coefficients statistics, mostly focusing on the marginal statistics, and only a few providing models representing the joint statistics of the wavelet coefficients [43], [91], [92]. In the following, the marginal and joint statistical models of the wavelet coefficients will be studied.

2.3.1 Wavelet Marginal Models

A common assumption in marginal modeling of the wavelet coefficients is that these coefficients are independent and identically distributed (i.i.d). In addition, the peakiness

and heavy-tail properties of the wavelet coefficients are taken into account in the marginal PDFs. A few of these models are discussed below.

a. Generalized Gaussian (GG) Distribution

The GG distribution has been used to model the wavelet coefficients of images. The zero-mean GG probability density function is defined as [13], [93]

$$f_{GG}(x) = \frac{\beta}{2\alpha\Gamma(\frac{1}{\beta})} e^{-\left(\frac{|x|}{\alpha}\right)^\beta} \quad (2.7)$$

where α is a scale parameter and β is shape parameter and $\Gamma(\cdot)$ is the Gamma function defined by $\Gamma(t) = \int_0^\infty e^{-u} t^{t-1} dt, t > 0$. The parameter α controls the width of the peak of PDF, while β is inversely proportional to the decreasing rate of the peak. Note that the GG distribution includes the Gaussian and Laplacian distributions as special cases.

b. Cauchy Distribution

The Cauchy distribution, as a non-Gaussian distribution with one degree of freedom has been used to model the wavelet coefficients of images [26], [40], [41]. The PDF of the zero-mean symmetric Cauchy distribution is given by

$$f_{Cauchy}(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + x^2} \quad (2.8)$$

where $\gamma > 0$ is a dispersion parameter, which corresponds to the spread of the PDF, supplying the same information as the variance.

c. Bessel K-Form Distribution

The Bessel K-Form (BKF) distribution has been recently used as an alternative to the GG and Cauchy distributions to provide a better fit to the empirical distributions of the wavelet coefficients. The density function of a zero-mean BKF distribution is given by [42], [79]

$$f_{BKF}(x) = \frac{\left(\frac{c}{2}\right)^{-\frac{p}{2} - \frac{1}{4}} |x|^{\frac{p}{2} - \frac{1}{2}}}{\sqrt{\pi} \Gamma(p)} K\left(p - \frac{1}{2}\right) \left(\sqrt{\frac{2}{c}} |x|\right) \quad (2.10)$$

where $p > 0$ is a shape parameter, $c > 0$ is a scale parameter, K is the modified Bessel function of the second kind and $\Gamma(\cdot)$ is the gamma function. It should be noted that the distribution becomes more heavy-tailed when p is close to zero.

The above distributions provide models for capturing the non-Gaussian behavior of the wavelet coefficients. However, they do not consider the inter-scale dependencies of the coefficients. In the following the joint models for the wavelet coefficients of images are discussed.

2.3.2 Wavelet Joint Models

The conventional PDFs usually disregard the dependency of the wavelet coefficients in a single subband and between subbands. Consequently, these PDFs cannot fit very well the empirical PDF of the wavelet coefficients of images. In order to have a better modeling of the wavelet coefficients, joint models have been proposed to not only take into account the non-Gaussian behavior of the wavelet coefficients, but also to capture inter-scale, intra-scale and cross orientation dependencies of the wavelet coefficients. A number of wavelet joint models such as the hidden Markov models (HMMs) [43]- [45], Markov random field priors (MRFs) [48], [92], [94] and Gaussian scale mixtures (GSMs) [95], [96] have been proposed. It should be noted that only very little literature has studied statistical models to describe cross orientation correlations. This is mostly due to the fact that unlike marginal models, joint models have much more complicated structures that deal with the characteristics of the wavelet coefficients. In general, joint models are a combination of three types of dependencies: inter-scale [100], [97], intra-scale [45], [98], [99], [101] and combined intra-scale and inter-scale [102], [103], [104] dependencies.

Crouse et al. [43] have studied a probabilistic model that captures coefficient dependencies across scales. In this model, the hidden states describe each coefficient the significance of the values of coefficients. Then, statistical models are fitted the coefficient's hidden state dependencies. To realize the influence of the inter-scale and intra-scale dependencies, one needs to be familiar with the concept of information theory.

2.3.3 Inter-scale and Intra-scale Dependencies

Joint statistics of transform domain coefficients show significant improvement in performance over the marginal statistics since they take into account the correlation between the two random variables results in a more precise expression of them. It is known that there exist non-zero dependencies between the wavelet coefficients. Neighboring coefficients contain substantial information on one another, and exploiting this information in modeling can improve the accuracy and usefulness of the resulting models. The mutual information has been used as a measure of dependencies between the wavelet coefficients [105]. The mutual information $I(X, Y)$, between two variables X and Y having a joint PDF $f_{XY}(x, y)$ is defined as

$$I(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x, y) \log \frac{f_{XY}(x, y)}{f_X(x)f_Y(y)} dx dy \quad (2.11)$$

Mutual information can be interpreted as to how much information one variable contains about the other. Another expression of mutual information is given by

$$\begin{aligned} I(X, Y) &= E_{XY} \left[\log \frac{f_{XY}(x, y)}{f_X(x)f_Y(y)} \right] \\ &= D(f_{XY}(x, y) \| f_X(x)f_Y(y)) \end{aligned} \quad (2.12)$$

where $D(\cdot \| \cdot)$ denotes the Kullback-Liebler (KL) distance or relative entropy. Mutual information can be interpreted as a difference measure between the joint density of X

and Y , and the product of their marginal densities. This difference increases as X and Y become increasingly dependent on each other.

It is challenging to calculate the mutual information between every pair of coefficients due to the large dimensions of images. Therefore, it is assumed that mutual information remains constant within each subband, i.e., stationarity and ergodicity in each subband. The dependencies between coefficients and other generalized neighborhoods are ignored. Thus, only the dependencies between coefficients and their parents, neighbors and cousins are computed. Figure 2.2 shows the dependencies between coefficients and their parents, neighbors and cousins are shown in Fig 2.2.

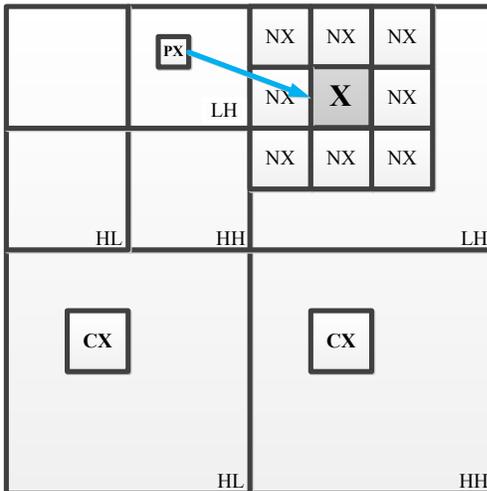


Figure 2.2. Dependencies between wavelet coefficient X and their parents PX in the coarser band, neighbors NX and cousins CX at the same level but in different orientations.

TABLE 2.1: Mutual information between the wavelet coefficients of the *Lena* image [106]

Filter types	$I(X, PX)$	$I(X, NX)$	$I(X, CX)$
Haar	0.20	0.27	0.14
Daubechies 4-taps	0.14	0.23	0.08

Table 2.1 gives an estimate on the mutual information of the wavelet coefficients for the *Lena* image. When $I(X, Y)$ increases, the dependency increases. In this table, the following mutual information is presented

- $I(X, PX)$, where X denotes a wavelet coefficient and PX its parent in the next coarser subband.
- $I(X, NX)$, where NX is a predefined neighborhood of X (excluding X).
- $I(X, CX)$, where CX is a predefined cousin of X in the same level and position but in a different orientation.

It is observed from this table that there exist inter-scale, intra-scale and cross orientation dependencies between the wavelet coefficients of images. Therefore, taking advantage of a model that considers such dependencies can improve the performance in different image processing applications using wavelet-domain statistical models.

2.4 Hidden Markov Model

As mentioned earlier, the wavelet transform has some attractive features such as locality, multiresolution and compression, which make it a desirable choice in statistical signal

processing. Beside these primary features, the wavelet transform also has the properties of non-Gaussianity, i.e., peaky, heavy-tailed marginal distributions and persistence across scales, i.e., large/small values of wavelet coefficients tend to spread across scales.

Taking into account these properties of the wavelet transform, the hidden Markov model in the wavelet domain has been proposed in [43]. It is known that the wavelet transform of a typical signal consists of a small number of large coefficients and a large number of small coefficients. Each coefficient can be considered as being in one of two states, “high” or “low” depending on the level of energy it contains. The result is a two-state mixture model for each wavelet coefficient called a two-state HMM. The two-state HMM models the non-Gaussian marginal PDF as a two-component Gaussian mixture. If a wavelet coefficient is small (large), its hidden state is labeled as small (high). The small state corresponds to the Gaussian component with a relatively small variance and captures the peakiness around the mean value, whereas the high state corresponds to the high variance Gaussian components, capturing the heavy tails. It should be noted that although each wavelet coefficient is conditionally Gaussian, due to the randomness of states, the overall density function is non-Gaussian. The two-state HMM can readily be extended to an M-state HMM [43].

In M-state HMM, for each wavelet coefficient x_{ij} , i and j representing the node and scale, respectively, there is a hidden state S_{ij} with the probability mass function $P(S_{ij} = m) = P_{ij}^m$, $m = 1, 2, \dots, M$. Conditioning on $S_{ij} = m$, x_{ij} follows a Gaussian density with mean μ_{ij}^m and variance $(\sigma_{ij}^m)^2$. The marginal distribution of the wavelets coefficients in the i^{th} node and the j^{th} scale can be written as

$$f_X(x_{ij}) = \sum_{m=1}^M \frac{p_{ij}^m}{\sqrt{2\pi}\sigma_{ij}^m} \exp\left\{-\frac{(x_{ij} - \mu_{ij}^m)^2}{2(\sigma_{ij}^m)^2}\right\} \quad (2.13)$$

where $\sum_{m=1}^M p_{ij}^m = 1$. There exists an inter-scale dependency between each of the wavelet coefficients at a coarse level, parent and the corresponding four coefficients at the next level, children, as shown in Figure 2.3 (a). The persistence across scales can be captured through state transition probability matrices, A_{ij} given by

$$A_{ij} = \begin{bmatrix} p_{ij}^{1 \rightarrow 1} & p_{ij}^{1 \rightarrow 2} & \dots & p_{ij}^{1 \rightarrow M} \\ p_{ij}^{2 \rightarrow 1} & \ddots & & p_{ij}^{2 \rightarrow M} \\ \vdots & & \ddots & \vdots \\ p_{ij}^{M \rightarrow 1} & \dots & \dots & p_{ij}^{M \rightarrow M} \end{bmatrix}_{M \times M} \quad j = 2, 3, \dots, J \quad (2.14)$$

where $p_{ij}^{m \rightarrow m'}$ is the probability of a child coefficient being in state m given its parent coefficient in state m' , where $m' = 1, 2, \dots, M$. By denoting the parent of the node i by $\rho(i)$ in the wavelet coefficient tree, we have

$$P(S_{ij} = m) = \sum_{m'} P(S_{\rho(i)} = m') P(S_{ij} = m | S_{\rho(i)} = m') \quad (2.15)$$

To reduce the number of the model parameters, the tied version of HMM is used, i.e., all the nodes at the same scale j have the same statistics. Hence, we may write

$A_{ij} = A_j$, $p_{ij}^m = p_j^m$, $\mu_{ij}^m = \mu_j^m$ and $\sigma_{ij}^m = \sigma_j^m$, $\forall i$. Thus,

$$p_j^m = \sum_{m'} p_{j-1}^{m'} p_j^{m' \rightarrow m}, \forall j = 2, 3, \dots, J \quad (2.16)$$

If $p_j = [p_j^1, p_j^2, \dots, p_j^M]$, then $p_j = p_{j-1} A_j$. Thus

$$p_j = p_1 A_2 A_3 \dots A_j, \forall j = 2, 3, \dots, J \quad (2.17)$$

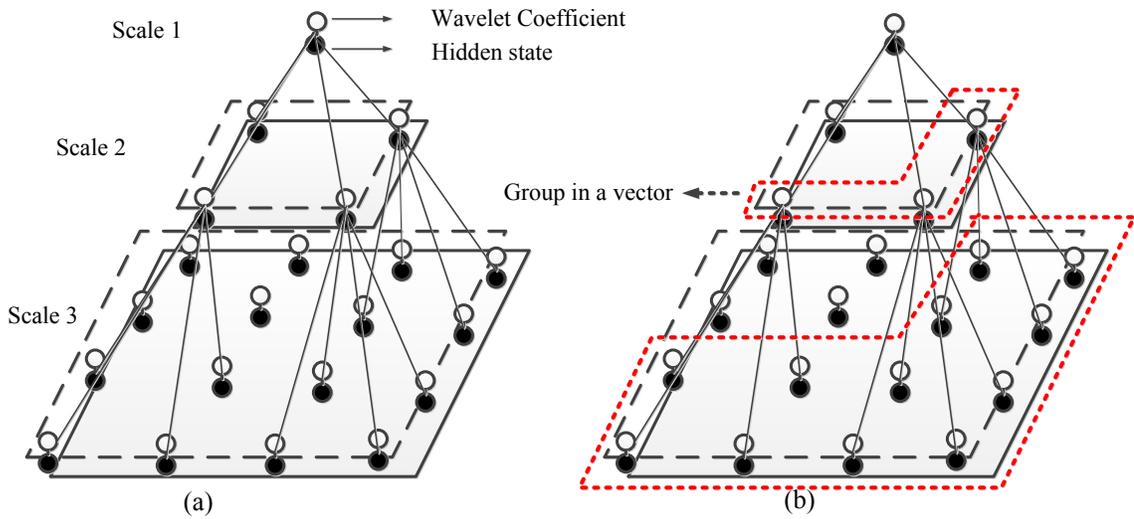


Figure 2.3. (a) Scalar and (b) Vector-based HMM.

The HMM is completely defined by a set of model parameters for each orientation $d = LH, HL$ or HH , as

$$\Theta^d = \left\{ p_1, A_2, \dots, A_J; \mu_j^m, \sigma_j^m, \forall j = 1, 2, \dots, J; m = 1, 2, \dots, M \right\}^d \quad (2.18)$$

To enhance the capability of the wavelet domain HMM model to capture the cross-orientation dependency of the wavelet coefficients, grouping coefficients at the same location and scale into vectors, and then modeling them by a single multidimensional HMM has been proposed in [46]. This results in a single vector HMM Θ for the entire input image. If x_{ij}^d denotes the wavelet coefficients at orientation $d: HL, LH$ or HH , the grouping process yields vectors of coefficients as $\mathbf{x}_{ij} = [x_{ij}^{LH}, x_{ij}^{HL}, x_{ij}^{HH}]^T$. The cross-correlation of these three wavelet coefficients for the tied version can be described by their covariance matrix C_j^m . The diagonal elements of the 3×3 covariance matrix C_j^m are the variances of the three orientations of the wavelet coefficients, whereas the non-diagonal elements are the cross-correlations between pairs of these coefficients. Figure 2.3 (b) depicts the result of the grouping. The marginal distribution function of the wavelet coefficient vectors at scale j in a vector-based HMM can be expressed as [46]

$$f_X(\mathbf{x}_{ij}) = \sum_{m=1}^M \frac{p_j^m}{\sqrt{(2\pi)^3 |\det(C_j^m)|}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_{ij} - \mu_j^m)^T (C_j^m)^{-1} (\mathbf{x}_{ij} - \mu_j^m) \right\} \quad (2.19)$$

In the vector-based HMM, the wavelet coefficients at the same scale and location, but in

different orientations, are tied so as to have the same hidden state. For modeling an image in a vector-based HMM, the set of parameters are given by

$$\Theta = \left\{ p_1, A_2, \dots, A_J; \mu_j^m, C_j^m, \forall j = 1, 2, \dots, J; m = 1, 2, \dots, M \right\} \quad (2.20)$$

There exists an efficient expectation maximization (EM) algorithm for fitting HMM to the observed signal data using the maximum likelihood criterion [107]. The EM algorithm is used to train the current model Θ in (2.20), to find the model Θ' by maximizing the expectation $E_s[\ln P(\mathbf{x}, \mathbf{S} | \Theta')]$ over the hidden states, where the wavelet coefficients and their corresponding hidden states for the image are denoted by \mathbf{x} and \mathbf{S} . In the E-step, the EM algorithm computes the marginal state probability mass functions $P(S_{ij} | \mathbf{x}, \Theta)$ and the joint probability mass functions between the parent nodes and their children nodes $P(S_{ij}, S_{\rho(i)} | \mathbf{x}, \Theta)$ given the current model Θ and the observation \mathbf{x} . In the M-step, these probabilities are used to update Θ' . The Baum–Welch algorithm can be used to calculate $P(S_{ij} | \mathbf{x}, \Theta)$ and $P(S_{ij}, S_{\rho(i)} | \mathbf{x}, \Theta)$ [107], [108]. To update the model parameters, the following formulas are used [47].

$$P(S_{ij} = m | S_{\rho(i)} = n) = \frac{P(S_{ij} = m, S_{\rho(i)} = n | \mathbf{x}_j, \Theta)}{n_j P(S_{\rho(i)} = m)}$$

$$\mu_j^m = \frac{\sum_i x_{ij} P(S_{ij} = m | \mathbf{x}_j, \Theta)}{\sum_i P(S_{ij} = m | \mathbf{x}_j, \Theta)} \quad (2.21)$$

$$C_j^m = \frac{\sum_i \Psi_{ij}^m (\Psi_{ij}^m)^T P(S_{ij} = m | \mathbf{x}_j, \Theta)}{\sum_i P(S_{ij} = m | \mathbf{x}_j, \Theta)}$$

where $\Psi_{ij}^m = [x_{ij}^{LH} - \mu_j^m, x_{ij}^{HL} - \mu_j^m, x_{ij}^{HH} - \mu_j^m]^T$, n_j is the number of nodes at scale j and the vector \mathbf{x}_j denotes the wavelet coefficients of the image at scale j . This procedure is repeated until the error between Θ and Θ' is less than a specified value.

The EM procedure can be summarized as follows.

-
1. Initialize: Select an initial model Θ^l , $l = 0$.
 2. E-Step: Compute $P(S_{ij} | \mathbf{x}, \Theta^l)$ and $P(S_{ij}, S_{\rho(i)} | \mathbf{x}, \Theta^l)$.
 3. M-Step: Estimate $P(S_{ij} = m | S_{\rho(i)} = n)$, μ_j^m and C_j^m using (2.21).
 4. Set $l = l + 1$.
 5. Repeat Steps 2 to 4 until convergence to a predefined error.
-

2.5 Results of Modeling the Wavelet Coefficients Using the Vector-based HMM

As discussed earlier, many probability density functions have been used to provide a model for the wavelet coefficients of images. Among them, the joint models such as the HMM have been shown to provide better fits to the empirical distribution of the wavelet coefficients. We now investigate the performance of the vector-based HMM in modeling the wavelet coefficients of images and compare it to that yielded by the other existing distributions. To this end, the Kolmogorov-Smirnov distance (KSD) values between the empirical data and the vector-based HMM are computed. The KSD metric is a measure of closeness of the fit and is given by [109]

$$KSD = \max_x \left| \int [f_X(x) - \hat{f}_X(x)] dx \right| \quad (2.22)$$

where $f_X(x)$ and $\hat{f}_X(x)$ represent the theoretical and empirical PDFs of the random variable x , respectively. In order to see if the accuracy of the fit to empirical data in using the vector-based HMM can be improved by increasing the number of states M , the KSD values between the empirical data and vector-based HMM are computed for various values of M for a number of test images. Table 2.2 gives the KSD values and Table 2.3 the CPU time for modeling the second-level wavelet coefficients for one of the test images, *Barbara*, for various values of the number of states M . The simulations are carried out in Matlab on an Intel core i7 2.93GHz personal computer with 8 GB RAM. It is seen from these tables that there is little effect on the KSD values when the number of states are increased; on the other hand, the complexity as introduced by CPU time

increases significantly with M . Similar results have been observed for other images. In view of this, we choose the two-state HMM for modeling the wavelet coefficients.

Table 2.2: KSD values between the empirical data and vector-based HMM with different number of states for the *Barbara* image.

Number of States (M)	KSD values		
	LH	HL	HH
2	0.0742	0.0737	0.0796
3	0.0729	0.0733	0.0792
4	0.0731	0.0730	0.0802
5	0.0731	0.0728	0.0792
6	0.0730	0.0732	0.0790
7	0.0731	0.0732	0.0791
8	0.0731	0.0733	0.0790

Table 2.3: CPU time for modeling the wavelet coefficients with vector-based HMM with different number of states, for the *Barbara* image.

Number of States (M)	Time (Sec)
2	0.4492
3	5.8155
4	23.4395
5	34.1187
6	51.3816
7	106.2668
8	178.6773

The performance of the vector-based HMM in modeling the wavelet coefficients of images is studied. To this end, histograms of the actual data as well as the PDFs of various distributions such as the GG, Cauchy and BKF are plotted. Figure 2.4 shows the modeling results of the second level of wavelet transform using the vector-based HMM for different subbands of the *Lena* image. It is seen from this figure that the vector-based HMM provides a better fit to the empirical data than the GG, Cauchy and BKF distributions do. Similar results have also been observed for other test images. In addition, to quantify the performance of the PDFs, the KSD values are obtained. Table 2.4 gives the averaged KSD values between the empirical PDF and the vector-based HMM, Cauchy, BKF and GG PDFs over a number of images taken from [41]. It is observed from this table that the vector-based HMM provides a better fit to the empirical data than other distributions do, as indicated by the lower KSD values.

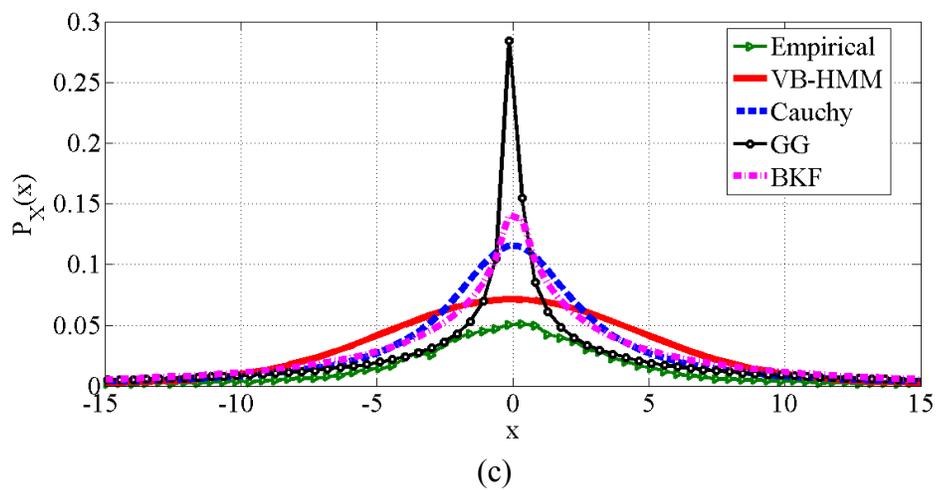
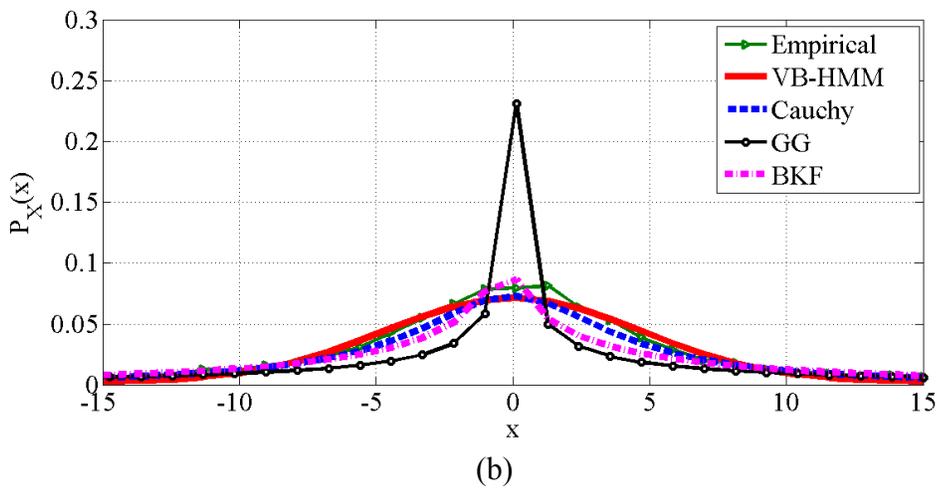
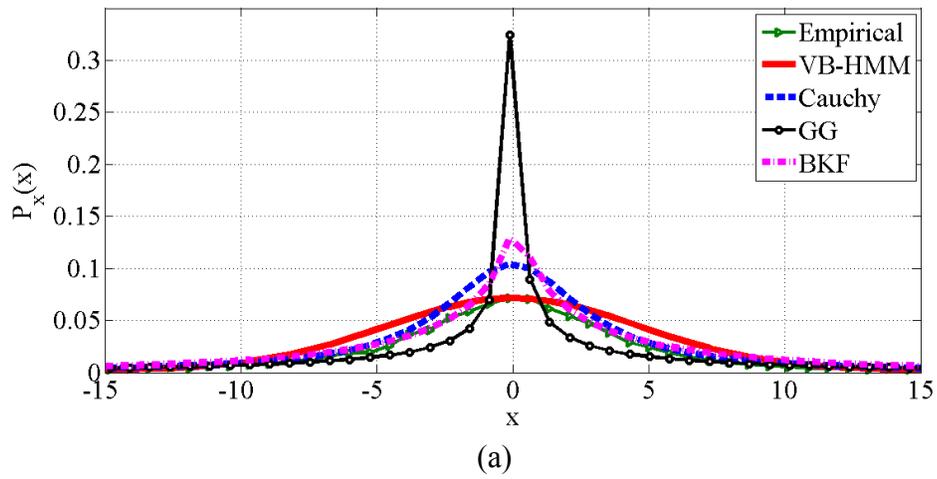


Figure 2.4. PDFs of the empirical data as well as vector-based HMM, Cauchy, GG and BKF distributions for second level of the wavelet transform of *Lena* image. a) *LH*, a) *HL* and c) *HH*.

Table 2.4: KSD values between the empirical data and different distributions averaged over a number of images for the second level of the wavelet transform.

Distribution	LH	HL	HH
GG	0.1797	0.2109	0.1941
Cauchy	0.0951	0.0943	0.0972
BKF	0.0824	0.0836	0.0980
VB-HMM	0.0702	0.0698	0.0719

2.6 Watermarking

As mentioned in Chapter 1, due to the development of the Internet, it has been possible to easily distribute digital multimedia data to a large number of users, and it can be duplicated very fast and without any loss of quality. Therefore, the possibility of unauthorized duplication and distribution of copyrighted material such as photographs, music, and movies, without an appropriate compensation to the copyright holders, is becoming increasingly problematic. As a remedy, development of data hiding techniques such as watermarking is essential. Watermarking is a way of embedding a secret message into the original data in order to increase its security and facilitate copyright protection [14]-[38], [49]-[69], [115]-[122].

In order to design a watermarking technique, two following important properties should be considered.

- a) Invisibility: The hidden message should be perceptually invisible in the original image. In other words, the quality of the watermarked image should be almost the same as the original image.
- b) Robustness: The hidden message should be detectable in the watermarked image, even after degradation due to any intentional or unintentional attacks or other processing. A scheme is usually called “robust” if the hidden message can be precisely decoded from the distorted watermarked image.

2.7.1. Classification of Watermarking Techniques

Image watermarking algorithms may be classified into two main categories according to the domain used for embedding the watermark, spatial [49] or frequency [14]-[41], [50]-[55], [58]-[62]. In the spatial domain, pixels of the image are modified in watermark embedding process. Although spatial domain watermarking schemes are considered simpler to be implemented, it may not have the satisfactory level of robustness to common image processing operations. On the other hand, in the frequency domain watermarking schemes, the transform domain coefficients of the image are modified for the embedding purpose. Frequency domain methods usually offer more robustness against different kinds of distortions. Some of the transforms that have been commonly used are those based on DFT [62],[111], DCT [39]-[41], DWT [14]-[38], ridgelet [50], [51], [117] and contourlet transform [52],[53], [109], [118], [119].

For embedding the watermark bits, there exist many approaches such as additive [15]-[31], [32], [39]-[41],

$$y_i = x_i + \alpha w_i \quad (2.23)$$

multiplicative [33],[38], [52], [53], [55], [58]

$$y_i = (x_i + \alpha)w_i \quad (2.24)$$

scaling [37]

$$y_i = \begin{cases} x_i \alpha & \text{for embedding 1} \\ \frac{x_i}{\alpha} & \text{for embedding 1} \end{cases} \quad (2.25)$$

and quantization-based [33]-[36], [65]-[68].

$$y_i = \text{round}\left(\frac{x_i}{Q_i}\right)Q_i \quad (2.26)$$

Watermarking schemes can also be classified according to the detection method employed into two major categories: informed detection, where the host signal is available at the detector during the watermark detection process, and blind detection, where the host signal is not available [15]-[32], [52].

In some applications of watermarking, it may be necessary only to determine whether a specific watermark is present or not in the received signal [14]-[30], [39]-[41], [52], [60], [62], whereas in the others, the embedded watermark is considered as a hidden unknown message that needs to be decoded accurately [31]- [39], [50], [51], [56]-[58], [62], [64]-[69]. In order to implement a blind watermark detector or decoder, statistical properties of images are commonly used. In this direction, efforts have been mostly made on the statistical modeling of the transform domain coefficients [14]-[32], [38]-[41].

2.7 Summary

In this chapter, an introduction of the wavelet transform has been first presented. The statistical properties of the wavelet coefficients of images have then been studied. Several marginal and joint statistical models for the wavelet coefficients have been presented. Since the joint models can not only take into account the non-Gaussian behavior of the wavelet coefficients, but also capture the inter-scale and cross orientation dependencies of the wavelet coefficients, the vector-based HMM has next been studied in detail as an example of the joint model for the wavelet coefficients. The performance of this model has been evaluated in detail by conducting several experiments, and comparing the results with that of the other existing distributions for the wavelet coefficients, namely, the GG, Cauchy and BKF distributions. Simulation results have shown that the vector-based HMM can model the wavelet coefficients more accurately than other distributions do in terms of the Kolmogorov-Smirnov distance and the visual comparison between the model and empirical distribution. Finally, an introduction on the image watermarking problem has been given and the various ways of classifying watermarking techniques have been described. In Chapter 3, a more detailed study on image watermarking in the wavelet domain is carried and novel methods for detection and extraction of additively embedded watermarks are developed.

CHAPTER 3

Locally-optimum Detector and Optimum Decoder for Additive Watermarking Schemes

3.1 Introduction

Additive watermarking, i.e., adding the watermark bits directly to the spatial or frequency domain, has been known to be one of the best watermarking techniques due to its simplicity and ability in providing a high degree of imperceptibility. Many detection and extraction techniques have been proposed, especially in the transform domains, for an additive embedding of the watermark. In the case of watermark detection techniques using the statistical modeling of images, most of the existing methods in the transform domain [14], [33], [116] are based on the assumption that the transform domain coefficients have Gaussian distribution so that a correlation detector can be used for the purpose of detection. However, it is well known that the correlation-based detectors are not optimal for non-Gaussian data, and in addition, they ignore the dependencies among the transform domain coefficients. Thus, the use of locally optimum (LO) detectors designed based on the signal statistics have been proposed and shown to provide considerably better detection results than that provided by the correlation-based detectors in various transform domains [70]. A LO detector have been designed in [22] for an additive watermark embedding, where the image coefficients have been modeled by the GG distribution. In [39], GG model has been employed for the DCT coefficients of images, and the corresponding detector has been designed based on a maximum

likelihood decision rule. A LO watermark detector has been designed in [40] by modeling the DCT coefficients using the Cauchy distribution. In [26], the same PDF has been applied to the detail subband coefficients of DWT. In [18], a LO watermark detector has been proposed in which the BKF distribution is used for modeling the DWT coefficients. Image watermark decoding techniques have been established to extract the watermark bits from the watermarked image that is possibly distorted rather than verifying the existence of the watermark. There exist several works [27], [28], [33], [39], [50], [128] focusing on the watermark extraction specifically by using the statistical properties of the transform domain coefficients. In [39], additive watermarking has been used in the DCT domain, and a decoder has been designed by using the GG distribution as a prior model for the DCT coefficients.

Although there exist a number of detectors and decoders for additive watermarking approach, there is still a scope for further research and study on new watermarking schemes to improve the performance of watermark detector and decoder against various distortions.

It is known that the performance of a statistical model-based watermark detection or extraction is highly influenced by the accuracy of the model employed. There exist a number of distributions that have been used for watermark detection or extraction. In order to improve the watermark detection and extraction performance for an additively embedded watermark, in this chapter, locally optimum robust blind watermark detector and decoder are designed using the vector-based hidden Markov model in the wavelet domain [29], [30], [32]. As discussed in Chapter 2, the vector-based HMM provides a better fit to the distributions of the wavelet coefficients of images in comparison to that

provided by other existing distributions. This is mostly due to its ability to capture the inter-scale and cross orientation dependencies between the wavelet coefficients. The scheme of watermark detection is formulated using the log-likelihood ratio test. A closed-form expression for the test statistics of the receiver operating characteristic curve of the proposed detector is obtained. The optimum watermark decoder is designed by using the maximum likelihood criterion. A closed form expression for the bit-error-rate of the proposed additive decoder is derived and validated experimentally. The performances of the proposed detector and decoder are investigated experimentally and compared with those of the other existing detectors and decoders. The robustness of the proposed watermarking scheme against various attacks is also studied.

3.2 Locally Optimum Watermark Detector

A watermarking technique consists of two steps: the embedding part in which the watermark signal is inserted into the original image and the detection part, wherein the image is received and the presence of the watermark detected.

3.2.1 Watermark Embedding

In this thesis, we consider the host image I to be a grayscale image of size $N \times N$. The embedding part procedure consist of generating a watermark using a pseudo-random sequence w taking values $\{+1, -1\}$ with equal probabilities and a watermark weighting factor α . In order to embed the watermark bits, the variance of each approximation subband in the second level is calculated and the subband with the maximum variance is selected for inserting the watermark bits. The coefficients in the selected subband is modified using the embedding equation, given by

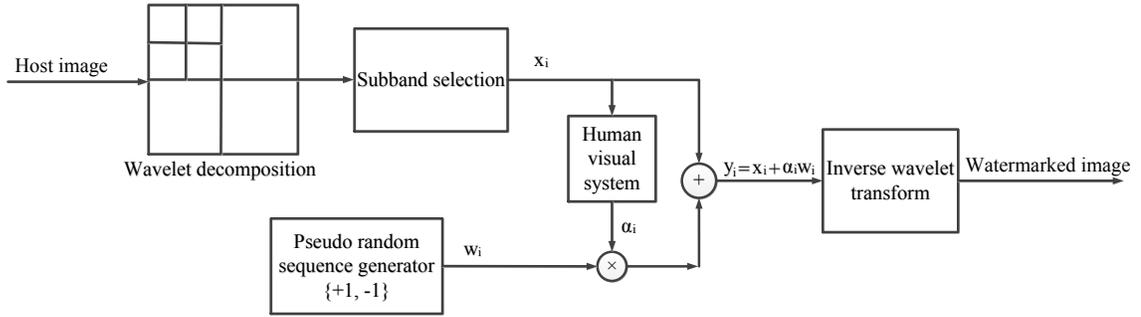


Figure 3.1. Block diagram of the proposed watermark embedding procedure.

$$y_i = x_i + \alpha_i w_i \quad (3.1)$$

By applying the inverse wavelet transform, the watermarked image is then obtained.

Figure 3.1 shows the block diagram of the proposed watermark embedding procedure.

3.2.2 Watermark Detector

In this section, we propose a new locally optimum watermark detector. Since the proposed watermarking detection method is performed without any knowledge of the original image, the watermarking technique is called blind. Signal detection, revealing the presence of a signal in a noisy observation, is a problem that can be regarded as binary hypothesis testing [122]. The binary hypothesis test can be formulated using the likelihood ratio test. The conventional watermark detectors were designed based on the assumption that the wavelet coefficients are Gaussian. This assumption leads to linear detectors. However, as discussed in Chapter 2, transform domain signals can be characterized more accurately by non-Gaussian distributions such as the vector-based

HMM that takes into account inter-scale and cross-orientation dependencies of the wavelet coefficients. Since a linear detector is optimal only for Gaussian data, the detector in the case of the vector-based HMM will be nonlinear. For the detection of non-Gaussian data, an optimal detector in the Neyman-Pearson sense can be designed for weak signals. This detector is known as a locally optimum (LO) detector, since it achieves asymptotically optimum performance for low signal levels [70]. We now give a method of designing a LO detector using the vector-based HMM and study its performance.

The problem of watermark detection can be formulated as a binary hypothesis test concerning the existence of a watermark. The two hypotheses for the test are formulated as [123]

$$\begin{cases} H_1 : y_i = x_i + \alpha w_i \\ H_0 : y_i = x_i \end{cases} \quad (3.2)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_L]$ and $\mathbf{y} = [y_1, y_2, \dots, y_L]$ are the wavelet coefficients of the selected subband of the original and watermarked images, respectively, $\mathbf{w} = [w_1, w_2, \dots, w_L]$ is the watermark, α is the watermarking weighting factor and L is the number of coefficients in the selected subband. The symbols x_i , y_i and w_i represent the values of the random variables X , Y and W , respectively. The detector is designed based on the maximum likelihood (ML) decision rule [39], [122] given by

$$\Lambda(\mathbf{y}) = \frac{f_Y(\mathbf{y} | H_1)}{f_Y(\mathbf{y} | H_0)} = \prod_{i=1}^L \left(\frac{f_Y(y_i | H_1)}{f_Y(y_i | H_0)} \right) \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \quad (3.3)$$

where η is the threshold. It should be noted that the magnitude of the watermark bits is low and therefore, embedding the watermark does not change the statistical characteristics of the image. The PDFs $f_Y(y_i | H_1)$ and $f_Y(y_i | H_0)$ follow the vector-based HMM distribution. Using (3.3), the log-likelihood ratio test can be written as

$$\begin{aligned} l(\mathbf{y}) &= \sum_{i=1}^L \ln \left(\frac{f_Y(y_i | H_1)}{f_Y(y_i | H_0)} \right) \\ &= \sum_{i=1}^L \ln \left(\frac{f_X(y_i - \alpha_i w_i)}{f_X(y_i)} \right) = \sum_{i=1}^L l(y_i) \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \quad (3.4) \end{aligned}$$

where $l(\mathbf{y}) \stackrel{\Delta}{=} \ln[\Lambda(\mathbf{y})]$ is the log-likelihood ratio, and $\tau = \ln(\eta)$. The log-likelihood term $l(y_i)$ in (3.4) can be approximated by using the Taylor series as

$$\begin{aligned} l(y_i)|_{w_i} &= l(y_i)|_{w_i=0} + \left. \frac{\partial l(y_i)}{\partial w_i} \right|_{w_i=0} \cdot w_i + O(w_i^2) \\ &\cong - \frac{\frac{\partial f_X(y_i)}{\partial y_i}}{f_X(y_i)} \cdot \alpha_i w_i \\ &= g_{LO}(y_i) \cdot \alpha_i w_i \quad (3.5) \end{aligned}$$

where $g_{LO}(y_i)$ is the ‘‘locally optimum nonlinearity’’ [39], [29], [30] defined by

$$g_{LO}(y_i) = -\frac{\frac{\partial f_X(y_i)}{\partial y_i}}{f_X(y_i)} \quad (3.6)$$

We now make use of the distribution of the vector-based HMM as given in (2.19) to obtain the PDF $f_X(y_i)$. Since the watermarking is performed at the second level of the wavelet transform, j in (2.19) assumes the value $(\log_2 N) - 2$, which for simplicity is denoted by q . Then, the locally optimum nonlinearity can be expressed after some algebraic manipulations, as

$$g_{LO}(y_i) = \frac{\sum_{m=1}^M \frac{1}{\sqrt{|\det(C_q^m)|}} p_q^m (y_i - \mu_q^m)^T (C_q^m)^{-1} \exp\left\{-\frac{1}{2}(y_i - \mu_q^m)^T (C_q^m)^{-1} (y_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{1}{\sqrt{|\det(C_q^m)|}} p_q^m \exp\left\{-\frac{1}{2}(y_i - \mu_q^m)^T (C_q^m)^{-1} (y_i - \mu_q^m)\right\}} \quad (3.7)$$

Now using (3.7) and (3.4) in (3.5), the log-likelihood ratio used to achieve LO detection is obtained as

$$l_{LO}(\mathbf{y}) = \sum_{i=1}^L \left(\frac{\sum_{m=1}^M \frac{1}{\sqrt{|\det(C_q^m)|}} p_q^m (y_i - \mu_q^m)^T (C_q^m)^{-1} \exp\left\{-\frac{1}{2}(y_i - \mu_q^m)^T (C_q^m)^{-1} (y_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{1}{\sqrt{|\det(C_q^m)|}} p_q^m \exp\left\{-\frac{1}{2}(y_i - \mu_q^m)^T (C_q^m)^{-1} (y_i - \mu_q^m)\right\}} \cdot \alpha_i w_i \right) \begin{matrix} H_1 \\ > \tau \\ H_0 \end{matrix} \quad (3.8)$$

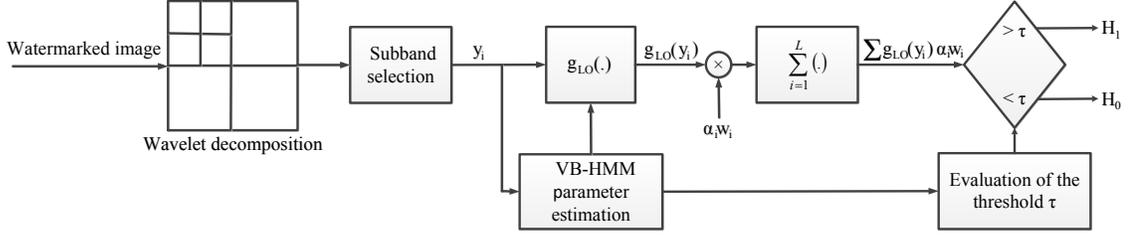


Figure 3.2. Locally-optimum watermark detection scheme using the vector-based HMM.

The detection is performed by comparing $l_{LO}(\mathbf{y})$ with a threshold τ , determined by the Neyman-Pearson criterion by maximizing the *probability of detection* P_{Det} for a predefined *probability of false alarm* P_{FA} [123]. It is noted that P_{Det} is the probability that the detector decides the proposition H_1 when the image is watermarked and that P_{FA} is the probability that it decides H_1 to be true when the image is, in fact, not watermarked. Figure 3.2 shows the block diagram of the proposed locally-optimum vector-based HMM detection scheme.

3.2.3 Performance Analysis of Vector-based HMM Detector

The analytical expression $l_{LO}(\mathbf{y})$ given in (3.12) allows theoretical measurement and experimental verification of the performance of the proposed vector-based HMM detector. The LO detector ratio $l_{LO}(\mathbf{y})$ in (3.8) is the sum of a large number of statistically independent random variables and hence, according to the central limit theorem, it can be assumed to follow a Gaussian distribution [39], [40], [15]. Therefore, we can consider the PDFs of $l_{LO}(\mathbf{y})$ conditioned on each of these hypotheses H_0 and

H_1 to follow Gaussian distributions, $N(m_0, \sigma_0)$ and $N(m_1, \sigma_1)$, respectively. The probabilities of false alarm and detection are computed using the mean and variance under each of these hypotheses as [40]

$$\begin{aligned} P_{FA} &= Q\left(\frac{\tau - m_0}{\sigma_0}\right) \\ P_{Det} &= Q\left(\frac{\tau - m_1}{\sigma_1}\right) \end{aligned} \quad (3.9)$$

where $Q(x)$ is given by

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \quad (3.10)$$

For a given probability of false alarm P_{FA} , the threshold is given by [40]

$$\tau = m_0 + \sigma_0 Q^{-1}(P_{FA}) \quad (3.11)$$

where $Q^{-1}(P_{FA}) = x$ if $Q(x) = P_{FA}$.

Using (3.7) in (3.5) gives the relationship between P_{FA} and P_{Det} as

$$P_{Det} = Q\left(\frac{\sigma_0}{\sigma_1} Q_C(P_{FA}) - \frac{m_1 - m_0}{\sigma_1}\right) \quad (3.12)$$

The performance of the LO vector-based HMM detector is evaluated in terms of its receiver operating characteristic (ROC) based on (3.12).

We now evaluate the theoretical performance of the proposed detector in terms of the ROC curves. For this purpose, expression for the mean and variance of the test statistic for two hypotheses H_0 and H_1 are obtained as follow. The mean under hypothesis H_0 is given by

$$m_0 = E[l_{LO}(\mathbf{y}) | H_0] = E[l_{LO}(\mathbf{x})]$$

$$= E \left[\sum_{i=1}^L \alpha_i w_i \cdot \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}} \right] \quad (3.13)$$

which can be simplified to

$$\begin{aligned}
m_0 = & \sum_{i=1}^L \left(\frac{1}{2} \right) (1) \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}}{\sqrt{|\det(C_q^m)|}}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}} + \\
& \left(\frac{1}{2} \right) (-1) \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}}{\sqrt{|\det(C_q^m)|}}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}} \right) \quad (3.14) \\
= & 0
\end{aligned}$$

In a similar manner, the mean of the likelihood ratio under hypothesis $H_1 : \mathbf{y} = \mathbf{x} + \alpha \mathbf{w}$ is given by

$$\begin{aligned}
m_1 &= E[l_{LO}(\mathbf{y}) | H_1] = E[l_{LO}(\mathbf{x} + \alpha \mathbf{w})] \\
&= E \left[\sum_{i=1}^L w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} \exp \left\{ -\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m) \right\}}{\sqrt{|\det(C_q^m)|}}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m) \right\}} \right] \\
&= \sum_{i=1}^L (a_i + b_i) \quad (3.15)
\end{aligned}$$

where

$$a_i = \left(\frac{\alpha_i}{2}\right) \frac{\sum_{m=1}^M \frac{p_q^m (x_i + \alpha_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i - \mu_q^m)\right\}}$$

(3.16)

and

$$b_i = \left(\frac{-\alpha_i}{2}\right) \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \alpha_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i - \alpha_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \alpha_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i - \alpha_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \alpha_i - \mu_q^m)\right\}}$$

(3.17)

The variance under H_0 is

$$\begin{aligned}
\sigma_0^2 &= \text{var}[l(\mathbf{y}) | H_0] = \text{var}[l(\mathbf{x})] \\
&= E \left[\left(\frac{\sum_{i=1}^L w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}} \right)^2 \right] \\
&= \sum_{i=1}^L E \left[\left(w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}} \right)^2 \right] \\
&\quad + \sum_l \sum_{l \neq i} E \left[\frac{w_l \alpha_l \frac{\sum_{m=1}^M \frac{p_q^m (x_l - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_l - \mu_q^m)^T (C_q^m)^{-1} (x_l - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_l - \mu_q^m)^T (C_q^m)^{-1} (x_l - \mu_q^m)\right\}} \cdot \frac{w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\}} \right] \\
&= \sum_{i=1}^L \alpha_i^2 \left(\frac{\left(\sum_{m=1}^M \frac{p_q^m (x_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\} \right)^2}{\left(\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2}(x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m)\right\} \right)^2} \right) \tag{3.18}
\end{aligned}$$

The variance under H_1 is given by

$$\begin{aligned}
\sigma_1^2 &= E\left[\left((LLO(y) | H_1) - m_1\right)^2\right] \\
&= \sum_{i=1}^L E \left[\left[\frac{w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m)\right\}} - a_i - b_i}{w_l \alpha_l \frac{\sum_{m=1}^M \frac{p_q^m (x_l + w_l \alpha_l - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_l + w_l \alpha_l - \mu_q^m)^T (C_q^m)^{-1} (x_l + w_l \alpha_l - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_l + w_l \alpha_l - \mu_q^m)^T (C_q^m)^{-1} (x_l + w_l \alpha_l - \mu_q^m)\right\}} - a_l - b_l}{w_i \alpha_i \frac{\sum_{m=1}^M \frac{p_q^m (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1}}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m)\right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp\left\{-\frac{1}{2} (x_i + \alpha_i w_i - \mu_q^m)^T (C_q^m)^{-1} (x_i + \alpha_i w_i - \mu_q^m)\right\}} - a_i - b_i} \right]^2 \right] \\
&= \sum_{i=1}^L (a_i - b_i)^2
\end{aligned} \tag{3.19}$$

In order to obtain the experimental ROC curves, Monte Carlo simulations are carried out in which 1000 pseudo-random watermark sequences are generated and embedded in the test image at every run for a given WDR. Then, experimental values of the mean and variance of the test statistic conditioned on each hypothesis are computed and the resulting ROC curve, obtained.

3.3 Watermark Decoder

In this section, a new additive watermark decoder using the vector-based HMM is proposed. In order to design a watermark decoder, some modification needs to be done in the embedding procedure discussed in Section 3.2.1. The original image is decomposed by a two-level wavelet transform, and for the purpose of embedding the watermark bits, the variance of each subband in the second level is calculated and the subband with the maximum variance is selected for inserting the watermark. Let $\mathbf{x} = \{x_1, \dots, x_L\}$, the set of wavelet coefficients of the selected subband of the original image, be divided into N_b nonoverlapping blocks B_1, B_2, \dots, B_{N_b} and let $\mathbf{m} = \{m_1, \dots, m_L\}$ be a pseudo-random sequence, where m_i takes the value “-1” or “1” with equal probability. The set of watermarked coefficients $\mathbf{y} = \{y_1, \dots, y_L\}$ is then given by

$$y_i = x_i + \alpha m_i b_k, \quad i = 1, \dots, L, \quad k = \left\lceil \frac{i N_b}{L} \right\rceil \quad (3.20)$$

The weighting factor α is used to provide a trade-off between the robustness of the watermarking scheme and the imperceptibility of the embedded watermark $\mathbf{b} = \{b_1, \dots, b_{N_b}\}$, b_k assuming the values +1 or -1 with equal probability. The watermarked image is then obtained by applying the inverse wavelet transform. Figure 3.3 shows the diagram of the new embedding procedure for of the proposed watermark decoder.

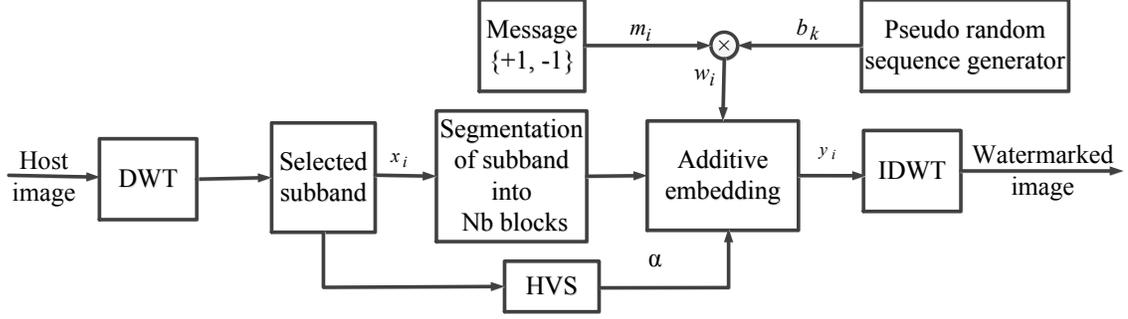


Figure 3.3. Block diagram of the embedding procedure of the proposed watermark decoder.

The function of a decoder in a watermarking scheme is to extract the hidden binary sequence from a set of observed wavelet coefficients. In order to extract the hidden bits in the wavelet subband coefficients, an optimum decoder based on the maximum likelihood criterion is developed using the vector-based HMM. To this end, for the k^{th} block with

$\left\lfloor \frac{L}{N_b} \right\rfloor$ coefficients, the maximum likelihood decision rule can be formulated as

$$\prod_{i \in B_k} f_Y(y_i | b_k = 1) \underset{H_0}{\overset{H_1}{>}} \prod_{i \in B_k} f_Y(y_i | b_k = -1) \quad (3.21)$$

where

$$\begin{cases} H_1: & y_i = x_i + \alpha m_i, & b_k = 1 \\ H_0: & y_i = x_i - \alpha m_i, & b_k = -1 \end{cases} \quad (3.22)$$

y_i being the i^{th} coefficient in the k^{th} block. Therefore, the optimum decoder $l_k(y)$ can be obtained by applying the natural logarithm on both sides of (3. 21) resulting in

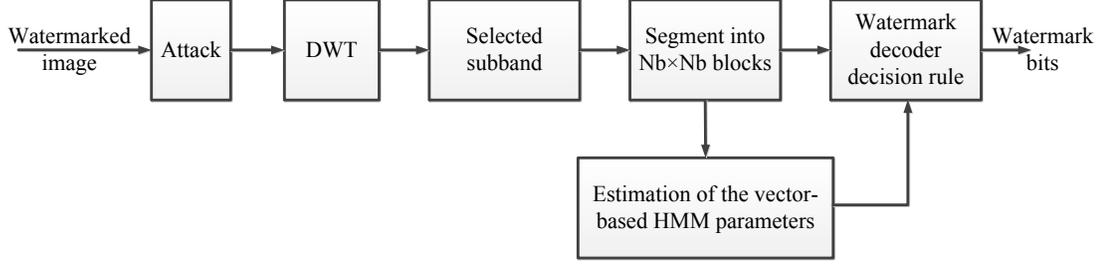


Figure 3.4. Block diagram of the decoding procedure of the proposed watermark decoder.

$$l_k(y) = \sum_{i \in B_k} \ln \frac{f_Y(y_i | b_k = 1)}{f_Y(y_i | b_k = -1)} \begin{matrix} > & H_1 \\ < & H_0 \end{matrix} < 0 \quad (3.23)$$

In order to calculate $l_k(y)$, we assume that the statistical models for $f_Y(y_i | b_k = \pm 1)$ are given by $f_X(y_i \pm \alpha)$, where $f_X(x)$ indicates the PDF of the wavelet coefficients of the selected subband of the host image. To obtain the PDF, $f_X(y_i \pm \alpha m_i)$, we make use of the M-state vector-based HMM marginal distribution given by (2.19). Thus, $l_k(y)$ can be obtained as

$$l_k(y) = \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (y_i - \alpha m_i - \mu_q^m)^T (C_q^m)^{-1} (y_i - \alpha m_i - \mu_q^m) \right\}}{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (y_i + \alpha m_i - \mu_q^m)^T (C_q^m)^{-1} (y_i + \alpha m_i - \mu_q^m) \right\}} \quad (3.24)$$

where $q = (\log_2 N) - 2$.

The k^{th} message bit can then be decoded as

$$\hat{b}_k = \begin{cases} 1 & l_k(y) > 0 \\ -1 & l_k(y) < 0 \end{cases} \quad (3.25)$$

The extraction part of the proposed watermarking scheme is shown in Figure 3.4.

3.3.1 Error Analysis

The bit error probability, also called bit error rate (BER) is used to analyze the performance of the proposed watermark decoder. The bit error probability is first computed in the absence of any attack. For the optimum decoder, the bit error probability is given by

$$P_e = \frac{1}{N_B} \sum_{k=1}^{N_B} \frac{1}{2} [P(l_k(y) > 0 | H_0) + P(l_k(y) < 0 | H_1)] \quad (3.26)$$

To find the probability under the condition of H_0 , $y_i = x_i - \alpha m_i$; hence $l_k(y | H_0)$ is equal to

$$\begin{aligned}
& l_k(y | H_0) \\
&= \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - 2\alpha m_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - 2\alpha m_i - \mu_q^m) \right\}}{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - \mu_q^m) \right\}} \quad (3.27)
\end{aligned}$$

Under the condition H_1 , $y_i = x_i + \alpha m_i$; hence $l_k(y | H_1)$ is equal to

$$\begin{aligned}
& l_k(y | H_1) \\
&= \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - \mu_q^m) \right\}}{\sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i + 2\alpha m_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i + 2\alpha m_i - \mu_q^m) \right\}} \quad (3.28)
\end{aligned}$$

It is noted that the sequence m_i is an independent identical random process that can have two values “- 1” and “1” with equal probability. Since $l_k(y | H_0)$ is the sum of a large number of independent random variables, according to the central limit theorem, it can be approximated by the Gaussian distribution with finite mean and variance under each hypothesis, i.e., (μ_0, σ_0) and (μ_1, σ_1) . The mean under the H_0 hypothesis, μ_0 is given by

$$\begin{aligned}
& \mu_0 = E[l_k(y | H_0)] \\
&= \sum_{i \in B_k} \left(\begin{aligned} & \frac{1}{2} \ln \sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - 2\alpha - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - 2\alpha - \mu_q^m) \right\} \\ & + \frac{1}{2} \ln \sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i + 2\alpha - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i + 2\alpha - \mu_q^m) \right\} \\ & - \ln \sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - \mu_q^m) \right\} \end{aligned} \right) \quad (3.29)
\end{aligned}$$

which can be simplified to

$$\mu_0 = \sum_{i \in B_k} (\ln \frac{\sqrt{a_i b_i}}{c_i}) \quad (3.30)$$

where

$$\begin{cases} a_i = \sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - 2\alpha - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - 2\alpha - \mu_q^m) \right\} \\ b_i = \sum_{m=1}^M p_q^m \exp \left\{ -\frac{1}{2} (\mathbf{x}_i + 2\alpha - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i + 2\alpha - \mu_q^m) \right\} \\ c_i = \sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - \mu_q^m)^T (C_q^m)^{-1} (\mathbf{x}_i - \mu_q^m) \right\} \end{cases} \quad (3.31)$$

The variance σ_0^2 is given by

$$\begin{aligned} \sigma_0^2 &= E \left[(l_k(y | H_0) - \mu_0)^2 \right] \\ &= \sum_{i \in B_k} \left(\ln \frac{a_i}{b_i} \right)^2 \end{aligned} \quad (3.32)$$

In the same way the mean under the H_1 hypothesis, μ_1 is given by

$$\begin{aligned} \mu_1 &= E[Z_k(y | H_0)] \\ &= \sum_{i \in B_k} (\ln \frac{c_i}{\sqrt{a_i b_i}}) \\ &= -\mu_0 \end{aligned} \quad (3.33)$$

The variance σ_1^2 is given by

$$\begin{aligned} \sigma_1^2 &= E \left[(l_k(y | H_1) - \mu_1)^2 \right] \\ &= \sum_{i \in B_k} \left(\ln \frac{a_i}{b_i} \right)^2 \\ &= \sigma_0^2 \end{aligned} \quad (3.34)$$

The error probability P_e^k for decoding a watermark bit is obtained as

$$\begin{aligned}
P_e^k &= \frac{1}{2} \{P(l_k(y) > 0 | H_0) + P(l_k(y) < 0 | H_1)\} \\
&= \frac{1}{2} \left[1 + Q\left(\frac{-\mu_0}{\sigma_0}\right) - Q\left(\frac{-\mu_1}{\sigma_1}\right) \right] \\
&= \frac{1}{2} \left[1 + Q\left(\frac{-\mu_0}{\sigma_0}\right) - Q\left(\frac{\mu_0}{\sigma_0}\right) \right] \\
&= 1 - Q\left(\frac{\mu_0}{\sigma_0}\right)
\end{aligned} \tag{3.35}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$. Thus, if the binary message bits “- 1” or “1” are embedded in the host image with the same probability, then the total BER is given by

$$P_e = \frac{1}{N_B} \sum_{k=1}^{N_B} P_e^k. \tag{3.36}$$

The performance of the proposed decoder is evaluated in terms of BER based on (3.36).

3.4 Experimental Results

Extensive experiments are conducted on a large set of test images taken from [127] to evaluate the performance of the proposed watermark detector and decoder based on vector-based HMM. The power of the watermark embedding can be determined by the watermark-to-document ratio (WDR), defined as [124]

$$WDR = 10 \log_{10} \left(\frac{\sum_i (\alpha_i w_i)^2}{\sum_i x_i^2} \right) \quad (3.37)$$

the numerator being the energy of the weighted pseudo-watermark bits and the denominator the energy of the host wavelet coefficients.

Figure 3.5 illustrates a few original and their corresponding watermarked images with $WDR = -34$ dB. It can be seen from this figure that there is no noticeable difference between the original and watermarked images, and hence, the proposed watermark embedding scheme provides a good performance in terms of the invisibility of the embedded watermark. The objective measure of the peak signal-to-noise-ratio (PSNR) between the original and watermarked images is also computed and given in Figure 3.5, in order to evaluate the imperceptibility of the watermark. The high PSNR values support the superior performance of the watermark embedding scheme.

3.4.1 Watermark Detection Results

Since we do not have access to the vector-based HMM parameters of the original image \mathbf{x} in order to calculate $l_{LO}(y)$ given by (3.8), we investigate the influence of using the parameters of the wavelet coefficients of the watermarked image, \mathbf{y} , instead of the parameters of the original image \mathbf{x} . To this end, we obtain the state probabilities p_q^m , the means μ_q^m and the covariance matrices C_q^m for a two-state vector-based HMM for the original image \mathbf{x} and the watermarked image \mathbf{y} . Table 3.1 gives the parameters p_q^m and

μ_q^m for the images (a)-(e) shown in Figure 3.5. Table 3.2 gives the mean square difference between the corresponding elements of the covariance matrices C_q^m for \mathbf{x} and \mathbf{y} . It is seen from these tables that the values of the estimated parameters of the watermarked wavelet coefficients, \mathbf{y} , are very close to that of the original wavelet coefficients, \mathbf{x} . Therefore, in the watermark detection scheme, we use the parameters obtained for the watermarked wavelet coefficients \mathbf{y} instead of that of \mathbf{x} for evaluating $l_{LO}(\mathbf{y})$.

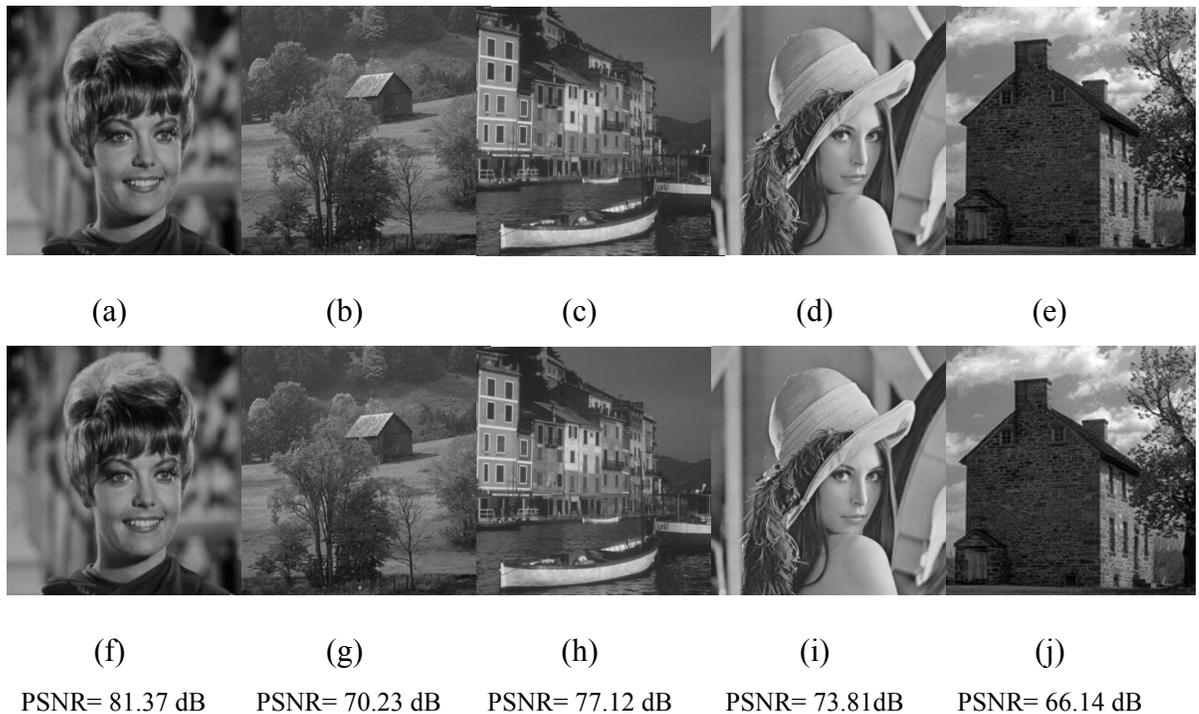


Figure 3.5. (a) - (e) Original test images, (f) - (g) Watermarked images corresponding to the test images in (a) - (e) for WDR = -34 dB.

TABLE 3.1: Comparison of the vector-based HMM parameters of the second level of the wavelet coefficients of x for the original image and the corresponding coefficients y for the watermarked image for the test images (a)- (e) of Fig. 3.5.

Modeling		State Probability p_q^m		Mean μ_q^m	
Image		m=1	m=2	m=1	m=2
(a)	x	0.9240	0.0760	-0.0602	2.9680
	y	0.9220	0.0780	-0.0562	2.8439
(b)	x	0.3906	0.6094	0.8570	-0.1115
	y	0.3839	0.6161	0.8655	-0.1063
(c)	x	0.6905	0.3095	-0.1268	0.6776
	y	0.6906	0.3094	-0.1267	0.6775
(d)	x	0.8053	0.1947	-0.0402	0.9009
	y	0.8055	0.1945	-0.0404	0.9029
(e)	x	0.4205	0.5795	0.1415	0.2714
	y	0.4204	0.5796	0.1415	0.2713

TABLE 3.2: Mean square difference between the corresponding elements of the covariance matrices C_q^m for \mathbf{x} and \mathbf{y} for the test images (a) - (e) of Figure 3.5.

Image	Mean Square Difference	
	m=1	m=2
(a)	0.0071	0.0231
(b)	0.0000	0.0001
(c)	0.0128	0.0170
(d)	0.0049	0.0039
(e)	0.0001	0.0059

- **Detection performance without attack**

In order to evaluate the theoretical performance of the proposed detector in terms of the ROC curves, we make use of the expressions for the mean and variance of the test statistic for the two hypotheses H_0 and H_1 obtained in Section 3.2.3. It is observed from (3.16), (3.18) and (3.19), that m_1 , σ_0 and σ_1 are dependent on the power of the watermark, WDR, given by (3.26). The theoretical ROC curves can be obtained using (3.12) for different values of WDR for a given image. In order to obtain the experimental ROC curves, Monte Carlo simulations are carried out. Experimental values of the means and variances of the test statistic conditioned on each hypothesis are computed and the resulting ROC curves, obtained. Figure 3.6 depicts the averaged theoretical as well as the experimental ROC curves over a number of test images [127] for P_{FA} varying from

10^{-8} to 10^{-2} for various values of WDR. It is seen from this figure that the experimental ROC curves are very close to the theoretical ones, thus establishing the validity of the expression derived in Section 3.2.3.

In view of this result, henceforth we use the theoretical values of m_0 , m_1 , σ_0 and σ_1 in order to compare the performance of the proposed detector with that of the locally optimum detectors using Cauchy [40], Gaussian [40], BKF [18], and GG [39] distributions, in terms of the ROC curves.

The ROC curves of the various detectors are obtained for a given watermarked image with P_{FA} varying from 10^{-8} to 10^{-2} . Figure 3.7 shows the ROC curves averaged over a number of test images for various detectors when WDR= - 34 dB. It is seen from this figure that the proposed LO vector-based HMM detector has the best performance in that it provides the highest probability of detection for a given probability of false alarm. Similar results have been obtained for various values of WDR.

Table 3.3 gives the area under ROC curve averaged over a number of test images for various detectors for different WDR values, for two regions for P_{FA} , $[0, 10^{-4}]$ and $[0, 1]$. It is seen from this table that the proposed method yields the best performance in that it provides the largest value for the both regions, irrespective of the watermark strength. Figure 3.8 shows the boxplots of the probability of detection, averaged over a number of test images, for different detectors when P_{FA} varies from 10^{-8} to 10^{-2} . It is also confirmed from this figure that the vector-based HMM detector outperforms the other detectors by providing higher detection rates.

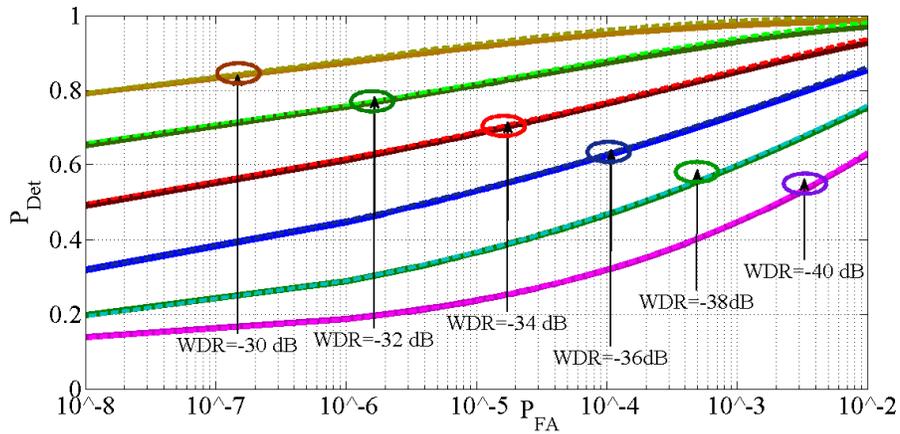


Figure 3.6. Theoretical (solid) and experimental (dashed) ROC curves averaged over 96 test images for the vector-based HMM detector for different values of WDR.

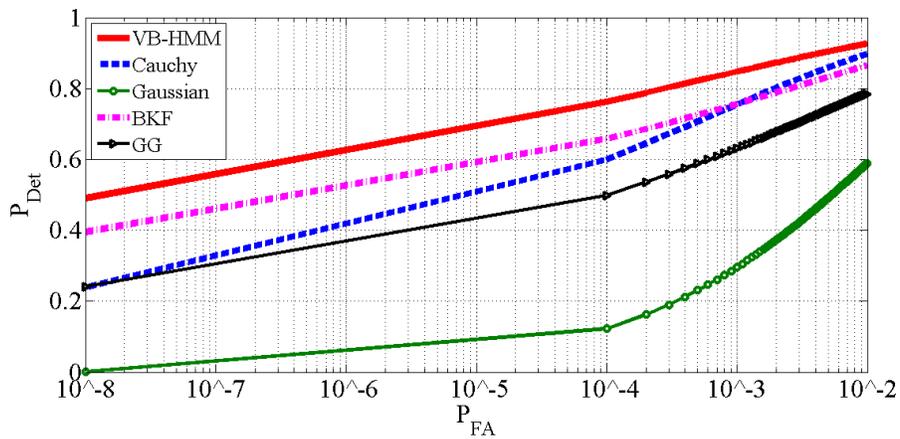


Figure 3.7. ROC curves averaged over 96 test images for various detectors when WDR = -34 dB.

TABLE 3.3: Averaged values of the area under ROC curves over 96 test images for various detectors and different values of WDR. (Best results are shown in bold)

WDR	Region	VB-HMM	Cauchy	Gaussian	BKF	GG
-34	[0, 10 ⁻⁴]	0.72627×10⁻⁴	0.5992×10 ⁻⁴	0.1220×10 ⁻⁴	0.6574×10 ⁻⁴	0.4988×10 ⁻⁴
	[0, 1]	0.9932	0.9855	0.9645	0.9802	0.9571
-36	[0, 10 ⁻⁴]	0.6248×10⁻⁴	0.4998×10 ⁻⁴	0.0455×10 ⁻⁴	0.5249×10 ⁻⁴	0.3912×10 ⁻⁴
	[0, 1]	0.9800	0.9768	0.9243	0.9702	0.9383
-38	[0, 10 ⁻⁴]	0.4662×10⁻⁴	0.3488×10 ⁻⁴	0.0175×10 ⁻⁴	0.4283×10 ⁻⁴	0.2802×10 ⁻⁴
	[0, 1]	0.9551	0.9627	0.8728	0.9481	0.9208
-40	[0, 10 ⁻⁴]	0.3580×10⁻⁴	0.2459×10 ⁻⁴	0.0073×10 ⁻⁴	0.3481×10 ⁻⁴	0.2346×10 ⁻⁴
	[0, 1]	0.9405	0.9191	0.8174	0.9162	0.8895

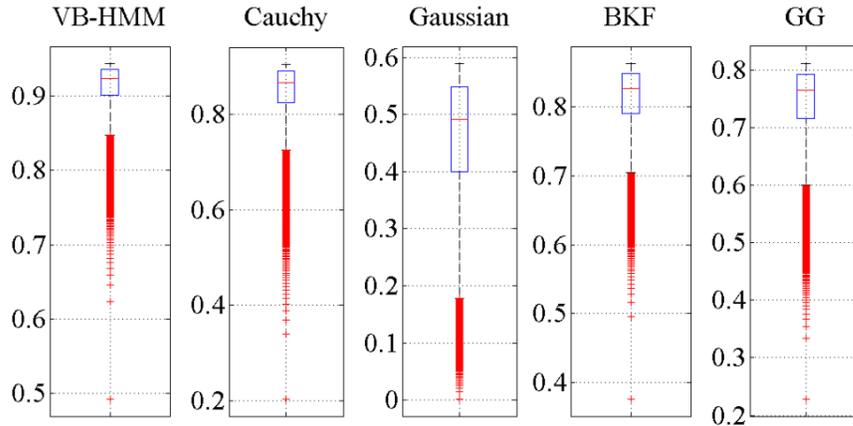


Figure 3.8. Boxplots of the probability of detection for various detectors averaged over a number of test images when $WDR = -34$ dB and P_{FA} ranging from 10^{-8} to 10^{-2} .

- **Detection performance in presence of attacks**

We now study the robustness of the proposed detector against various attacks: JPEG compression, geometric distortion (rotation), median filtering, Gaussian filtering and Gaussian noise. Figures 3.9 to 3.13 depict ROC curves averaged over 96 test images obtained using the proposed watermark detector as well as those obtained using the Cauchy, Gaussian, BKF and GG-based detectors when the watermarked images with $WDR = -34$ dB are JPEG compressed with a quality factor (QF) of 30, rotated counter clockwise by 2° , median filtered with a window of size 3×3 , Gaussian filtered with a window of size 3×3 , and corrupted by the Gaussian noise ($SNR = 25$ dB), respectively. It can be seen from these figures that the proposed detector is more robust than the other detectors against any of the attacks considered.

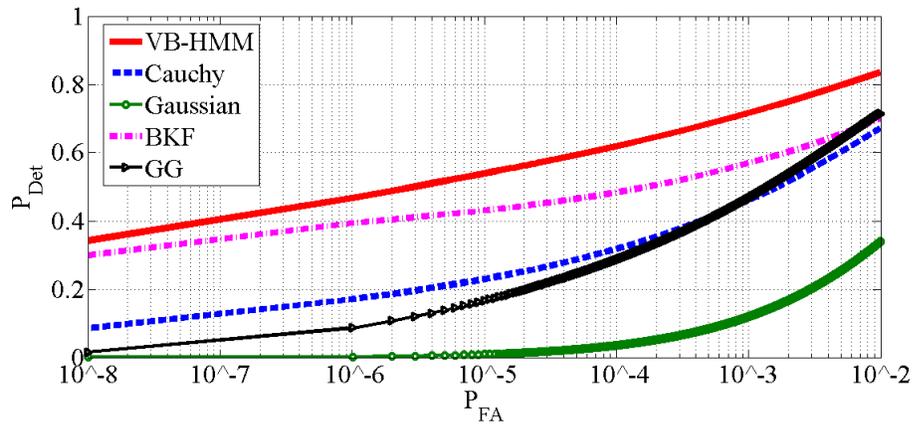


Figure 3.9. ROC curves averaged over a number of test images obtained using various detectors for WDR = -34 dB when image is JPEG-compressed with QF = 30.

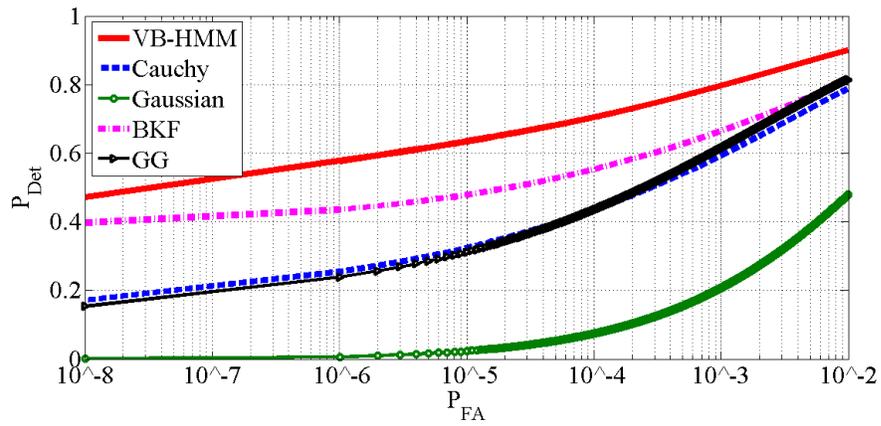


Figure 3.10. ROC curves averaged over a number of images obtained using various detectors and WDR = -34 dB when images are rotated by 2° .

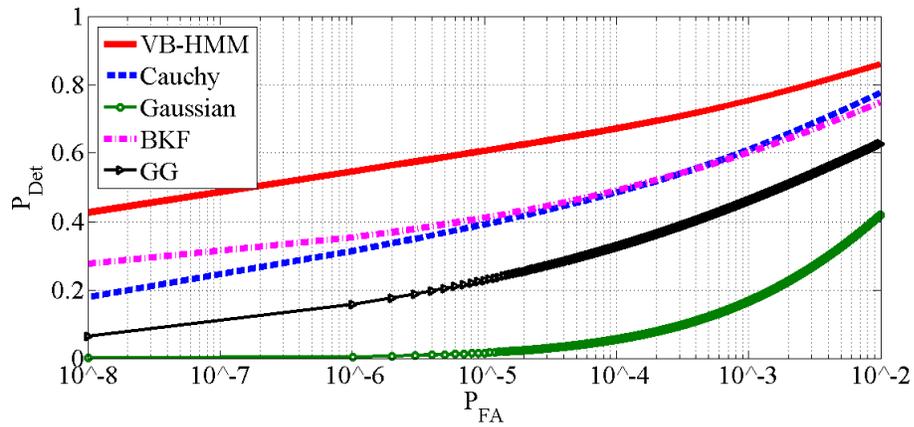


Figure 3.11. ROC curves averaged over a number of test images obtained using various detectors for WDR = -34 dB when images undergo median filtering with a window size of 3×3 .

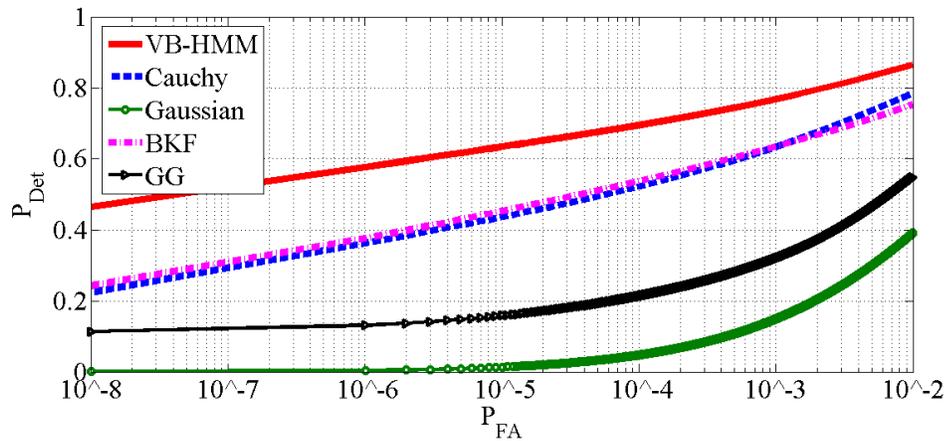


Figure 3.12. ROC curves averaged over a number of test images for the various detectors for WDR = -34 dB when Gaussian filtering with mask 3×3 .

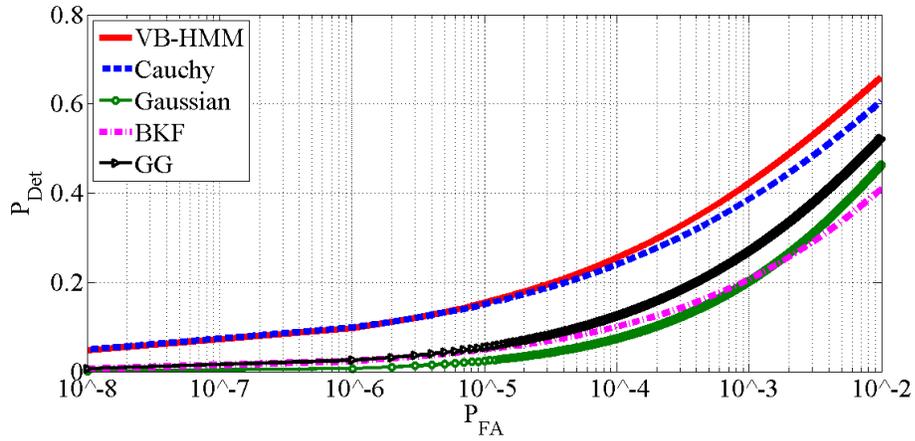


Figure 3.13. ROC curves averaged over a number of test images obtained using various detectors and WDR = -34 dB when images undergo Gaussian noise with SNR = 25 dB.

In order to further investigate the performance of the proposed watermark detector using the vector-based HMM and compare it to that of the other detectors, we show in Figures 3.14 to 3.18 the boxplots of P_{Det} obtained using the proposed detector as well as that obtained using the detectors based on the Cauchy, Gaussian, BKF and GG distributions, when the watermarked images are JPEG-compressed with QF = 30, rotated counter clockwise by 2° , median filtered with a window of size 3×3 , Gaussian filtered with a window of size 3×3 , and corrupted with Gaussian noise with SNR = 25 dB, respectively. It can be seen from these figures that the proposed detector is more robust than the other detectors against any of these attacks.

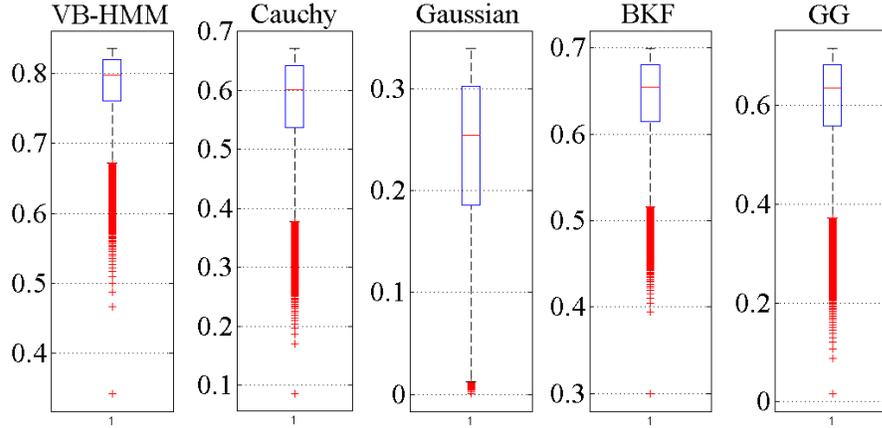


Figure 3.14. Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image JPEG compressed with QF = 30.

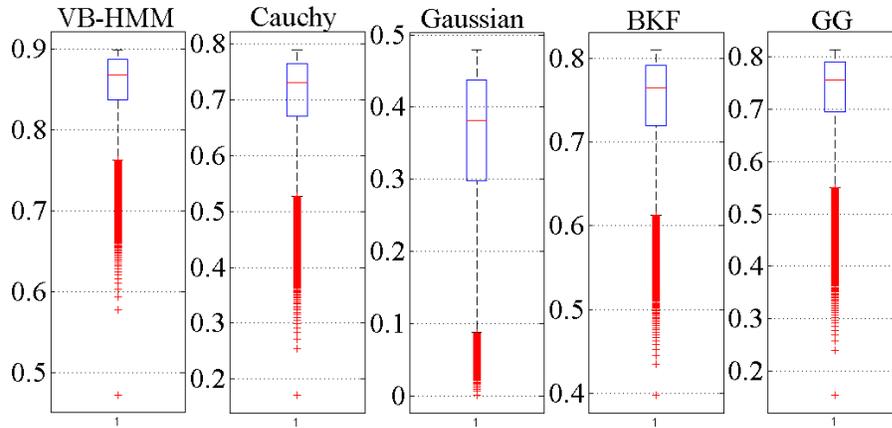


Figure 3.15. Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image is rotated by 2° .

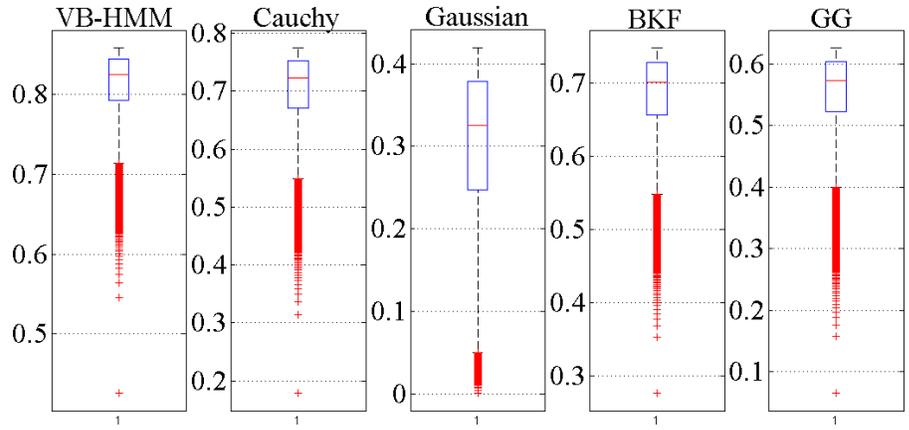


Figure 3.16. Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when image undergoes median filtering with a window size of 3×3 .

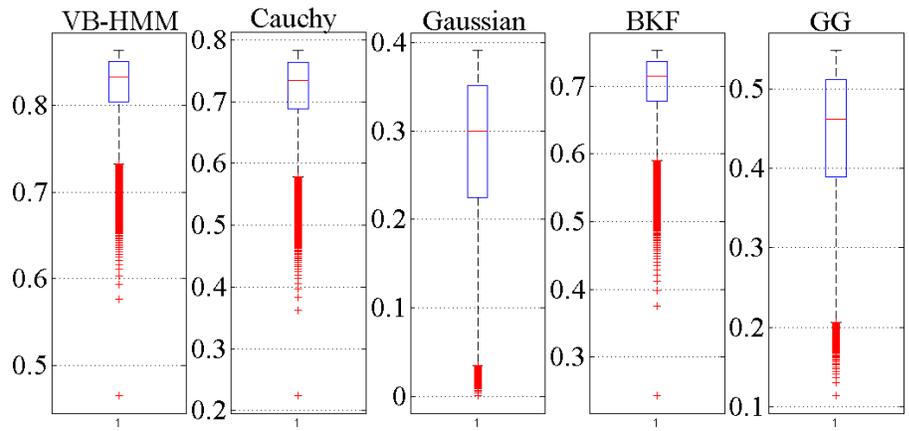


Figure 3.17. Boxplots of the probability of detection for various detectors averaged over a number of test images for WDR = -34 dB when Gaussian filtering with mask 3×3 .

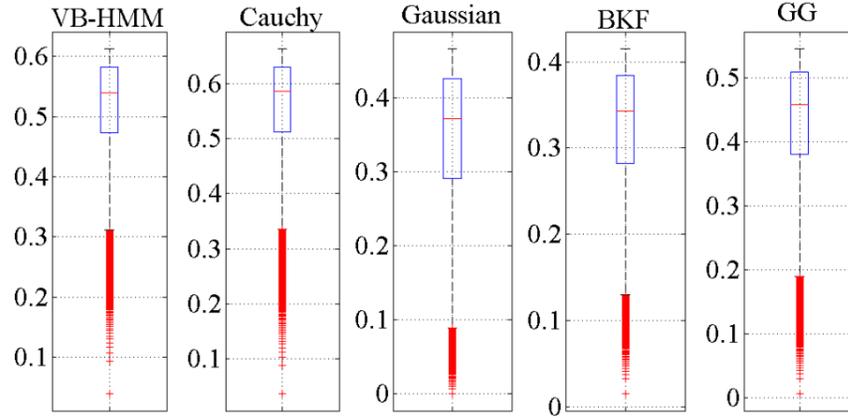


Figure 3.18. Boxplots of the probability of detection for various detectors averaged over a number of test images for $WDR = -34$ dB when image undergoes Gaussian noise with $SNR = 25$ dB.

We now obtain the CPU times of the various detectors by averaging over a number of test images to evaluate the computational complexity of these detectors. The CPU times, averaged over 96 images, required by Gaussian, Cauchy, GG, VB-HMM and BKF-based detectors are 0.51, 0.96, 0.98, 1.01 and 1.34 seconds, respectively. It is seen from Figures 3.7 to 3.13 that in spite of the low computational time, the performance of the Gaussian-based detector is unacceptable. Amongst the remaining detectors, the proposed VB-HMM detector provides the highest detection rate with comparable or lower computational time.

3.4.2 Watermark Decoder Results

Performance of the proposed decoder is studied without attack and also in presence of attacks. In order to validate the theoretical values of BER obtained from (3.36), comparisons are made with experimental BER obtained from Monte Carlo simulations. For this purpose, for each of the test images, 1000 pseudo-random message sequences are generated, each sequence embedded in the test image for a given WDR, and decoded using (3.36). The number of errors is computed for each run, and the experimental BER averaged over the 1000 runs. Figure 3.19 shows the theoretical and experimental BER values of the proposed decoder averaged over the 96 test images for various values of WDR. It is seen from this figure that the BER values obtained theoretically are very close to the experimental ones, thus validating the expression for BER given by (3.36).

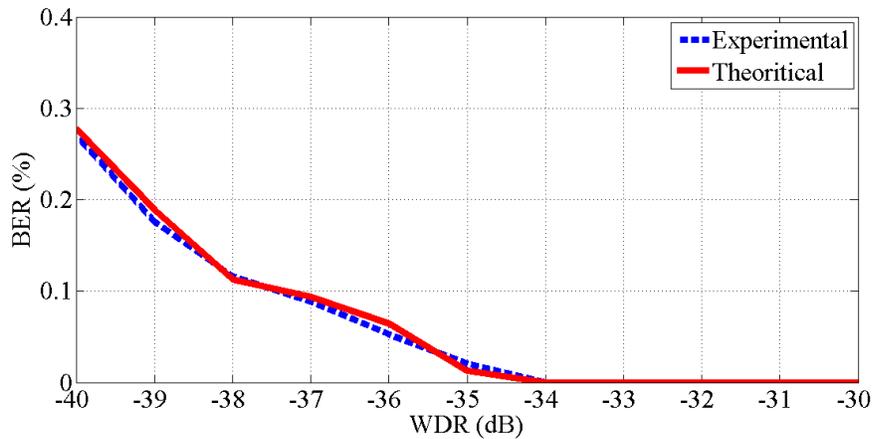


Figure 3.19. Theoretical and experimental BER of the proposed decoder averaged over a number of test images with message length of 128 bits for different WDR values.

We also study the robustness of the proposed additive watermark decoder using the vector-based HMM against common signal processing attacks such as JPEG compression, additive Gaussian noise and rotation. Figure 3.20 shows the BER values obtained using the proposed decoder for the test images, *Barbara*, *Baboon*, *Peppers* and *Lena*, when the watermarked images are JPEG-compressed with QF changing from 10 to 80, when PSNR = 50dB and message length is 128 bits. It is seen from this figure that the proposed watermarking scheme using the vector-based HMM decoder is highly robust against compression attack. Similar results have also been obtained for the other test images.

The results of BER when the test images, *Barbara*, *Baboon*, *Peppers* and *Lena*, with PSNR = 50 dB and message length of 128 bits are contaminated by the additive white Gaussian noise (AWGN) with noise standard deviations σ_n varying from 0 to 30 are shown in Figure 3.21. It is seen from this figure that the proposed decoder using the vector-based HMM is highly robust against additive Gaussian noise.

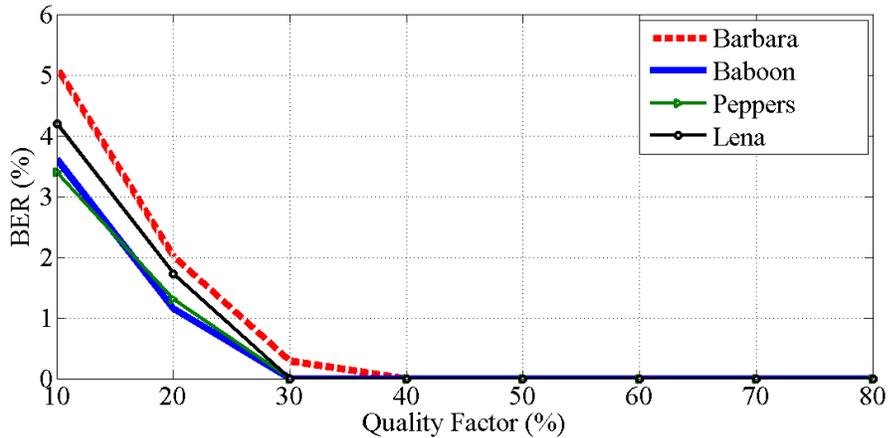


Figure 3.20. BER of the extracted watermark obtained using the proposed vector-based HMM decoder when watermarked images, are JPEG-compressed with different quality factors.

We now investigate the robustness of the proposed watermarking scheme against rotation attack. Figure 3.22 shows the results of BER when the test images, *Barbara*, *Baboon*, *Peppers* and *Lena*, are rotated by different angles when PSNR = 50dB and message length is 128 bits. It is seen from this figure that the proposed watermarking scheme using the vector-based HMM is highly robust against rotation attack as indicated by low values of BER. It is to be noted that we compensate the desynchronization caused by the rotation attack using the method proposed in [37].

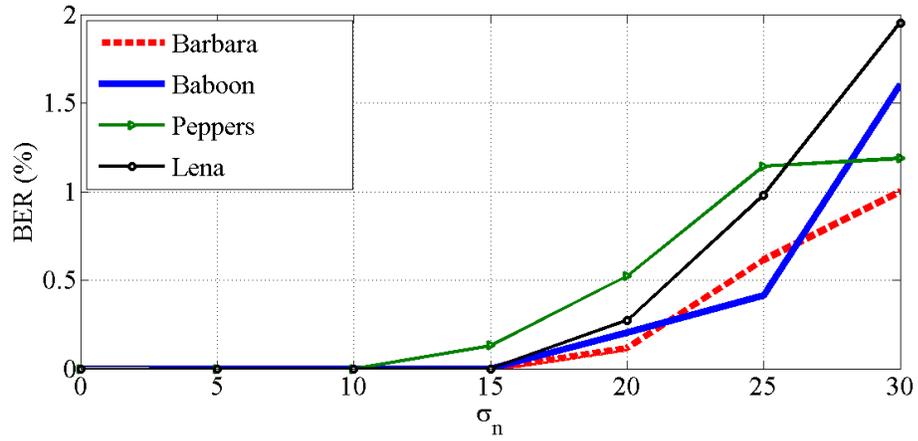


Figure 3.21. BER of the extracted watermark using the proposed vector-based HMM decoder when watermarked images are contaminated by the AWGN with different noise levels.

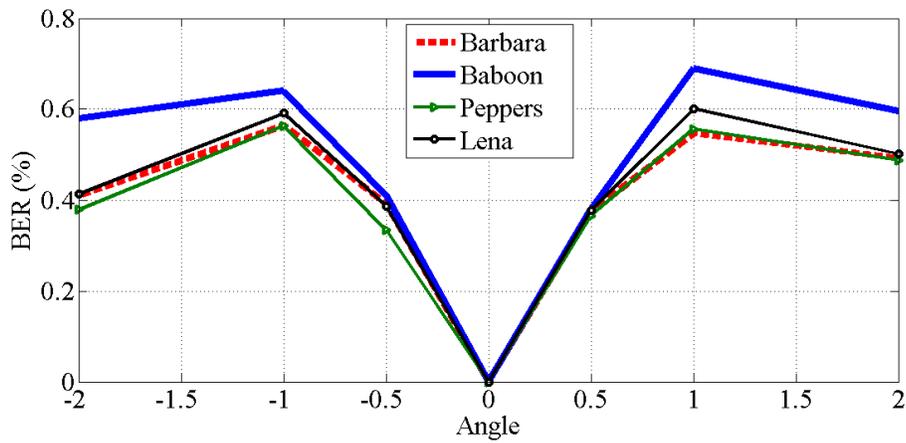


Figure 3.22. BER of the extracted watermark using the proposed vector-based HMM decoder when watermarked images, *Barbara*, *Baboon*, *Peppers* and *Lena* are rotated by different angles.

In order to investigate the performance of the proposed method for much larger set of images, we compute the mean and variance of the BER values for 96 test images with and without different kinds of attacks. These are computed for three different message lengths and given in Table 3.4. As seen from this table both the mean and variance of the BER values are small, which indicate the good performance of the proposed schemes as well as the repeatability of the result for different images.

TABLE 3.4: Mean and variance of the BERs (%) values of 96 test images obtained using the proposed watermarking scheme under various attacks for the different message lengths. (PSNR= 50 dB)

Type of attacks	Message length					
	64 bits		128 bits		256 bits	
	Mean	Variance	Mean	Variance	Mean	Variance
No attack	0.0	0.0	0.0	0.0	0.0	0.0
JPEG QF=10	2.38	0.3053	4.07	0.3890	5.76	3.9640
AWGN $\sigma_n = 20$	0.0	0.0	2.23	0.3045	3.99	1.5181
Rotation $\theta = 0.5^\circ$	0.38	0.0063	0.43	0.0078	0.75	0.24
Median filtering 3×3	0.0	0.0	0.05	0.0009	0.79	0.16

We now compare the performance of the proposed decoder using vector-based HMM with that of the works in [33], [35], [36], [65]-[69]. In order to make a fair comparison, for a given message length, we set the PSNR values of the watermarked images in our proposed method to be the same as the values reported in these works.

Table 3.5 gives BER values obtained using the proposed decoder for an embedded message of 128 bits as well as that of the decoders in [35] and [36], when the *Lena* image is contaminated by the AWGN with various values of the noise standard deviation and is JPEG-compressed with different QF values. It is seen from this table that the proposed vector-based HMM decoder outperforms the methods in [35] and [36], by providing the lower BER values under attacks.

The performance of the proposed decoder using vector-based HMM is compared with that of the works in [36], [68], [69], when PSNR = 45dB and message length is 128 bits. Figure 3.23 gives BER values obtained using various methods under AWGN when the noise level varies from 5 to 30. It is seen from this figure that the proposed decoder has the best performance in the presence of AWGN in comparison to the other decoders.

TABLE 3.5: BER (%) obtained using the proposed VB-HMM-based watermarking scheme as well as that obtained using the schemes in [36] and [35], under various attacks for the *Lena* images. (Message length = 128 bits, PSNR = 45dB)

	VB-HMM	[36]	[35]
σ_n	AWGN		
5	0	0	0
20	2.25	10.16	2.34
35	11.31	13.44	20.31
QF	JPEG		
4	6.39	37.5	32.03
10	3.43	3.91	6.25
16	0	0	0
20	0	0	0

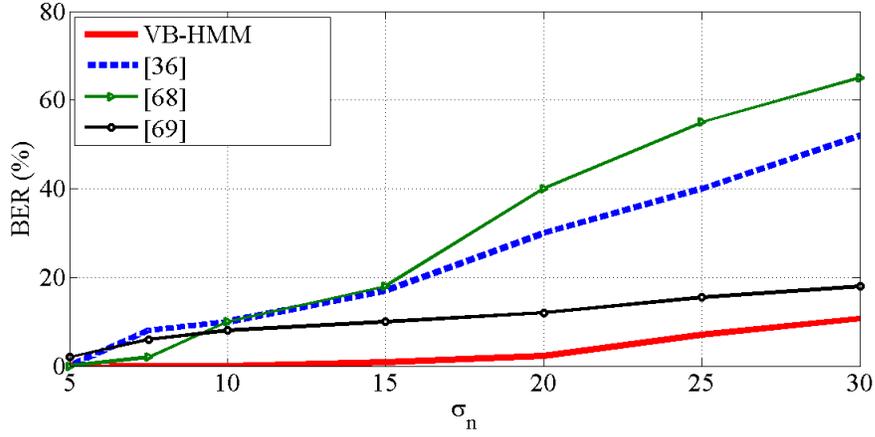


Figure 3.23. BER (%) obtained using the proposed vector-based HMM based watermarking scheme as well as that obtained using the schemes in [36], [68] and [69], under AWGN with different noise standard deviations for the *Lena* image. (Message length = 128 bits, PSNR = 45dB).

The performance of the proposed decoder is also compared to that of the works in [33], [35], [65]-[68]. Table 3.6 gives BER values obtained using the proposed decoder and that of the decoders in [33], [35], [65]-[68], for an embedded message of 256 bits against different attacks, namely, JPEG compression with QF = 11, AWGN with $\sigma_n = 10$, and median filtering with a window of size 3×3 for some of the test images, namely, *Barbara*, *Baboon*, *Peppers* and *Lena*. It is seen from this table that the proposed watermark decoder using the vector-based HMM is generally more robust than the other methods against various attacks as seen by the lower values of BER when watermarked images are under various attacks.

TABLE 3.6: BER (%) obtained using the proposed watermarking scheme as well as that obtained using the schemes in [33], [35], [65]-[68], under various attacks when Message length is 256 bits and PSNR = 42dB. (Best results are shown in bold and second best in parentheses)

	JPEG (QF=10)	AWGN $\sigma_n = 10$	Median filter 3×3
	<i>Barbara</i>		
VB-HMM	4.34	(1.15)	0.89
[35]	9.64	1.40	(1.10)
[65]	20.11	11.47	19.52
[66]	24.1	4.48	15.82
[67]	(4.69)	0.39	1.17
[33]	16.45	1.45	24.95
[68]	20.11	11.47	19.52
	<i>Baboon</i>		
VB-HMM	(3.81)	0.00	0.87
[35]	9.86	(1.28)	5.03
[65]	15.23	6.13	17.38
[66]	15.08	1.09	15.14
[67]	1.95	0.00	(2.73)
[33]	16.95	1.3	31.65
[68]	15.23	6.13	19.52

	<i>Peppers</i>		
VB-HMM	4.05	(1.19)	0.00
[35]	10.68	1.32	(1.17)
[65]	11.65	3.62	8.3
[66]	18.4	1.87	4.41
[67]	(10.16)	0.00	0.00
[33]	26.10	1.25	29.35
[68]	11.65	3.62	8.30
	<i>Lena</i>		
VB-HMM	7.93	(1.24)	0.00
[35]	(8.64)	1.85	0.00
[65]	15.81	5.01	(6.24)
[66]	24.96	3.34	6.25
[67]	9.77	0.00	0.00
[33]	29.80	1.45	30.80
[68]	15.81	5.01	(6.24)

Table 3.7 gives BER values for the proposed decoder and those in [51] for an embedded message of 64 bits against median filtering with window sizes 3×3 , 5×5 , 7×7 and 9×9 . It is seen from this table that the proposed watermark decoder is more robust against median filtering than the one in [51].

TABLE 3.7: BER values obtained using the proposed watermarking scheme as well as that obtained using the schemes in [51], when watermarked images are under various attacks. (Message length= 64 bits, PSNR = 42 dB)

	Method	3×3	5×5	7×7	9×9
Peppers	[51]	0.0	5.31	17.18	28.75
	Proposed	0.0	0.5	1.6	2.8
Baboon	[51]	1.56	20.93	30.62	35
	Proposed	0.0	0.18	2.34	3.17

3.6 Summary

In this chapter, algorithms for the blind detection and extraction of an additively embedded watermark using the vector-based HMM model for the image wavelet coefficients have been developed. The proposed watermark detector has been designed to be locally optimal and closed-form expressions for the mean and variance of a test statistic in a Bayesian log-likelihood framework have been derived using the vector-based HMM as a prior for the wavelet coefficients. To validate these theoretical expressions, experiments have been conducted using a large set of test images. The optimum watermark decoder has been designed using the maximum likelihood criterion. Closed form expression for the bit-error-rate of the proposed decoder is derived and validated experimentally. The performance of the proposed watermark detector has been evaluated in terms of the ROC curves and area under ROC curve values, using a large number of test images. The proposed detector has been shown to provide a rate of detection for a given probability of false alarm that is higher than that of the detector using the Cauchy, Gaussian, BKF or GG distribution for the wavelet coefficients. It has been shown that the proposed watermark decoder is superior to other decoders in terms of providing a lower bit error rate. It has also been shown that the proposed watermark detector and decoder based on the vector-based HMM are highly robust against various kinds of attacks. The next chapter deals with the problems of detection and extraction of a watermark embedded in the wavelet-transformed image multiplicatively.

CHAPTER 4

Optimum Multiplicative Watermark Detector and Decoder

4.1 Introduction

An important requirement for a robust image watermarking is to be able to achieve a desirable tradeoff between the quality of the watermarked image and the accuracy of the watermark detection. An alternative to additive embedding of the watermark is a multiplicative embedding [33], [38], [52], [53], [55], [58]. It is known that due to the data-dependent nature of the multiplicative embedding, the resulting watermarked images are not as transparent as their additive counterparts. However, schemes employing the multiplicative embedding approach have been shown to provide improved watermark detection performance against various attacks [20], [62]. Accordingly, several watermark detectors and decoders have been proposed [20], [21], [33], [38], [41], [52], [53], [55], [58], [62], [64]. Among these multiplicative watermarking schemes, there exist some works focusing on the watermark detection and extraction using the statistical properties of the image coefficients. In [21] LO detectors have been designed for watermarking schemes in which the DFT, DCT or DWT coefficients of images modeled by the GG distribution. In [20], a multiplicative LO detector has been developed using GG modeling the DWT coefficients. In [41], a multiplicative watermark detector has been designed by modeling the DCT coefficients of images using the Cauchy distribution. In [64], an optimum watermark decoder has been proposed in the Fourier domain using the Weibull distribution for a multiplicative embedding of the watermark. In [37], a scaling-based

watermark decoder in the wavelet domain has been proposed by assuming a Gaussian distribution for the modeling the wavelet coefficients. In [53], a multiplicative watermark decoder has been proposed using the GG distribution as a prior for the transform domain coefficients. In [38], a multiplicative watermark decoder has been proposed for fingerprint application in the wavelet domain using the GG distribution.

Although there exist several statistical watermark detectors and decoders in the wavelet domain, none of them has taken into account the correlation of the wavelet coefficients that exist across scales and orientations. Considering these dependencies, one can improve the accuracy of the parameter estimation and develop watermark detector and decoder schemes that can achieve higher rate of detection and extraction. To this end and to improve the robustness of a multiplicative watermarking scheme against various distortions, in this chapter, we propose an optimum blind multiplicative watermark detector as well as decoder in the wavelet domain using the vector-based HMM distribution [31]. A formulation for watermark detection is derived using the log-likelihood ratio test. A closed-form expression for the test statistics of the receiver operating characteristic curve of the proposed detector is obtained. The decoder is designed using the maximum likelihood criterion that uses the vector-based HMM as the statistical prior for the image wavelet coefficients. Closed form expression for the bit-error-rate of the proposed decoder is derived and validated experimentally. The performances of the proposed detector and decoder are investigated by conducting several experiments and compared with those of the other existing detectors and decoders. The robustness of the proposed scheme is examined when the watermarked

images are subjected to various kinds of distortions and compared to that of the other decoders.

4.2 Optimum Watermark Detection

The proposed multiplicative watermark detection has two steps; embedding and detection. In the embedding part, the watermark signal is inserted into the host image through the multiplicative approach whereas in the detection part, the existence of the watermark signal is detected from the watermarked image.

4.2.1 Watermark Embedding

A host image I is grayscale with size of $N \times N$. The watermark signal W is generated using a pseudo random sequence taking values $\{-1, 1\}$ with equal probabilities. The watermark signal is assumed to be independent from the host image coefficients. The host image is decomposed into subbands using a two-level wavelet transform. In order to embed the watermark bits, variance of each subband in the second level is calculated and the subband with maximum variance X is selected for inserting the watermark. The proposed multiplicative embedding equation is given by

$$Y = (1 + \alpha W)X \quad (4.1)$$

where α is a weighting factor that provides a trade-off between the robustness of the watermarking scheme and the imperceptibility of the embedded watermark. The watermarked image is then obtained by applying the inverse wavelet transform to the marked coefficients.

4.2.2 Watermark Detection

The two hypotheses for the test statistics are formulated as follows

$$\begin{cases} H_1 : Y = (1 + \alpha W) X \\ H_0 : Y = X \end{cases} \quad (4.2)$$

where $X = (x_1, x_2, \dots, x_L)$ and $Y = (y_1, y_2, \dots, y_L)$ are the wavelet coefficients of the selected subband for the original and watermarked images, respectively, $W = (w_1, w_2, \dots, w_L)$ is the watermark sequence and L is the number of coefficients in the selected subband. The detector is designed based on the log-likelihood ratio test given by

$$l(Y) = \ln \left(\frac{f_Y(Y | H_1)}{f_Y(Y | H_0)} \right) \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \tau \quad (4.3)$$

where τ is the threshold. Since

$$f_Y(y_i | H_1) = \frac{1}{1 + \alpha w_i} f_X \left(\frac{y_i}{1 + \alpha w_i} \right) \quad (4.4)$$

The log-likelihood ratio can be rewritten as

$$l(Y) = \sum_{i=1}^L \ln \left(\frac{1}{1 + \alpha w_i} f_X \left(\frac{y_i}{1 + \alpha w_i} \right) \right) \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \quad (4.5)$$

After inserting the M-state vector-based HMM distribution, given in (2.19), into (4.5), the log-likelihood ratio becomes

$$l(Y) = \sum_{i=1}^L \ln \frac{\frac{1}{1 + \alpha w_i} \sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_j^m)|}} \exp \left\{ \frac{1}{2} \left(\frac{y_i}{1 + \alpha w_i} - \mu_j^m \right)^T (C_j^m)^{-1} \left(\frac{y_i}{1 + \alpha w_i} - \mu_j^m \right) \right\}}{\sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_j^m)|}} \exp \left\{ -\frac{1}{2} (y_i - \mu_j^m)^T (C_j^m)^{-1} (y_i - \mu_j^m) \right\}} \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \quad (4.6)$$

The threshold τ is obtained using Neyman–Pearson criterion [115]. The performance of the detector can be measured by relating the probabilities of false alarm P_{FA} and detection P_{Det} , resulting in the ROC curves. To generate the ROC curves, the mean and variance of $l(Y)$ under each hypothesis need to be first estimated. The mean and variance under H_0 are obtained as

$$m_0 = E[l(Y | H_0)] = \sum_{i=1}^L \ln \frac{\sqrt{a_i b_i}}{c_i} \quad (4.7)$$

where

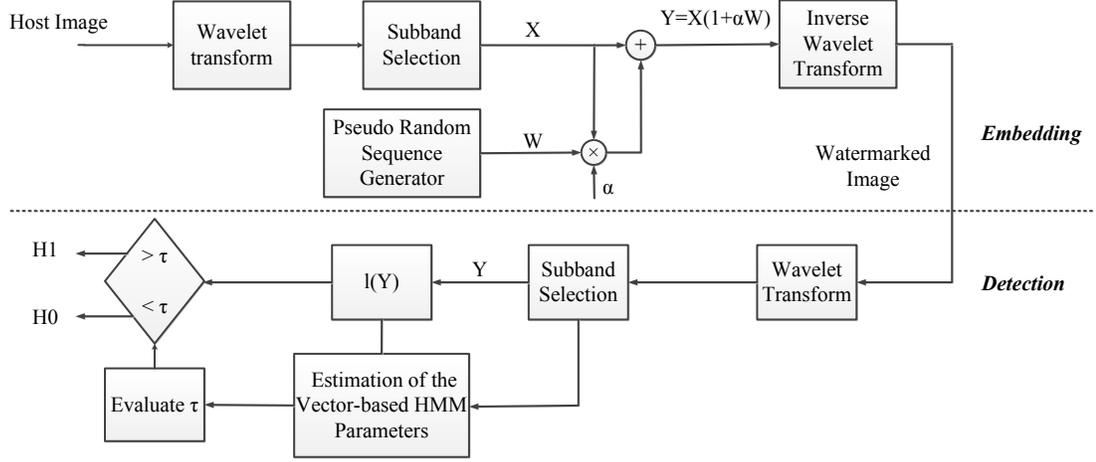


Figure 4.1. Proposed multiplicative watermarking scheme; embedding and detection parts.

$$\begin{cases}
 a_i = \frac{1}{(1+\alpha)} \sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_j^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i}{1+\alpha} - \mu_j^m \right)^T (C_j^m)^{-1} \left(\frac{y_i}{1+\alpha} - \mu_j^m \right) \right\} \\
 b_i = \frac{1}{(1-\alpha)} \sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_j^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{y_i}{1-\alpha} - \mu_j^m \right)^T (C_j^m)^{-1} \left(\frac{y_i}{1-\alpha} - \mu_j^m \right) \right\} \\
 c_i = \sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_j^m)|}} \exp \left\{ -\frac{1}{2} (y_i - \mu_j^m)^T (C_j^m)^{-1} (y_i - \mu_j^m) \right\}
 \end{cases} \quad (4.8)$$

and

$$\sigma_0^2 = E \left[(l(Y|H_0) - m_0)^2 \right] = \frac{1}{4} \left(\sum_{i=1}^L \ln \frac{a_i}{b_i} \right)^2 \quad (4.9)$$

Similarly, the mean and variance under H_1 can be obtained, where $m_1 = -m_0$ and $\sigma_1 = \sigma_0$. Embedding and detection scheme are summarize in Fig. 4.1.

4.3 Optimum Watermark Decoder

A watermark decoder is designed to extract the hidden watermark bits from the watermarked image. To this end, advantage can be taken of the statistical properties of the wavelet coefficients of the image. As discussed in Chapter 2, the wavelet coefficients can be modeled more accurately by using the vector-based HMM than the other distribution by taking into account the inter-scale and cross-orientation dependencies between the coefficients. Thus, a blind watermark decoder is designed using the vector-based HMM.

4.3.1 Watermark Embedding

The host image is first decomposed by a two-level wavelet transform. In order to embed the watermark bits, the variance of each subband in the second level is calculated and the subband with the maximum variance is selected for inserting the watermark. Let $\mathbf{x} = \{x_1, \dots, x_L\}$ be the set of the magnitudes of the wavelet coefficients of the selected subband. The set \mathbf{x} is divided into N_b non-overlapping equal-sized blocks B_1, B_2, \dots, B_{N_b} , and let $\mathbf{m} = \{m_1, \dots, m_L\}$ be a pseudo-random sequence, where m_i takes the value “-1” or “1” with equal probability. The watermark bits w are generated using $w_i = m_i b_k$, $i = 1, \dots, L$, $k = \left\lceil \frac{i N_b}{L} \right\rceil$, where $\mathbf{b} = \{b_1, \dots, b_{N_b}\}$ are message bits that can have values “- 1” and “1”. It should be noted that the same bit b_k is used for all the coefficients in the block B_k to obtain the watermark bits.

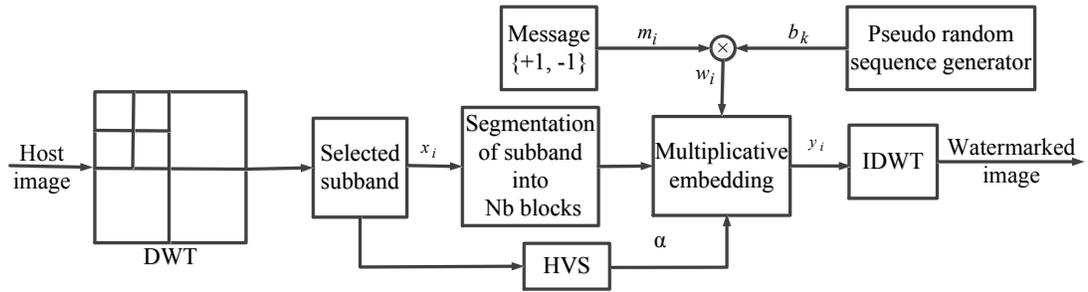


Figure 4.2. The proposed multiplicative watermark decoder; embedding scheme.

The set of watermarked coefficients $\mathbf{y} = \{y_1, \dots, y_L\}$ is obtained as

$$y_i = (1 + \alpha w_i) x_i, \quad i = 1, \dots, L \quad (4.10)$$

where α is a positive weighting factor that provides a trade-off between the robustness of the watermarking scheme and the imperceptibility of the embedded watermark. The weighting factor α is obtained by taking into account the human visual system (HVS) properties. The watermarked image is then obtained by applying the inverse wavelet transform to the marked coefficients. The block diagram for the proposed embedding scheme is shown in Figure 4.2.

4.3.2 Watermark Decoder

The proposed blind watermark decoder is based on the statistical properties of the wavelet coefficients of the image. The wavelet coefficients are modeled using the vector-based HMM, which is superior to other models in characterizing the statistical properties of

the wavelet coefficients by taking into account their dependencies across scales and orientations. Since the performance of a decoder is highly dependent on the accuracy of the model, the proposed decoder is expected to provide a performance better than that of the other decoders. The scheme for the proposed watermark decoder is shown in Figure 4.3.

In extraction part, the watermarked image is decomposed by a two-level wavelet transform and the coefficients y of the selected subband (the one with the maximum variance) are divided into N_b non-overlapping equal-sized blocks. A binary bit message b_k of “- 1” or “1” is embedded in the k^{th} block as

$$\begin{cases} H_1 : y_i = (1 + \alpha m_i)x_i, & b_k = 1 \text{ is embedded} \\ H_0 : y_i = (1 - \alpha m_i)x_i, & b_k = -1 \text{ is embedded} \end{cases} \quad (4.11)$$

where $i \in B_k$, the k^{th} block and x_i 's and y_i 's are the corresponding host and watermarked coefficients. It should be noted that the bits of the binary message sequence are assumed to be equally probable and the wavelet coefficients in each block are assumed to be independent. In order to extract the hidden message bit in the block B_k of the wavelet coefficients of the selected subband of the watermarked image, an optimum decoder based on the maximum likelihood decision is developed and formulated as

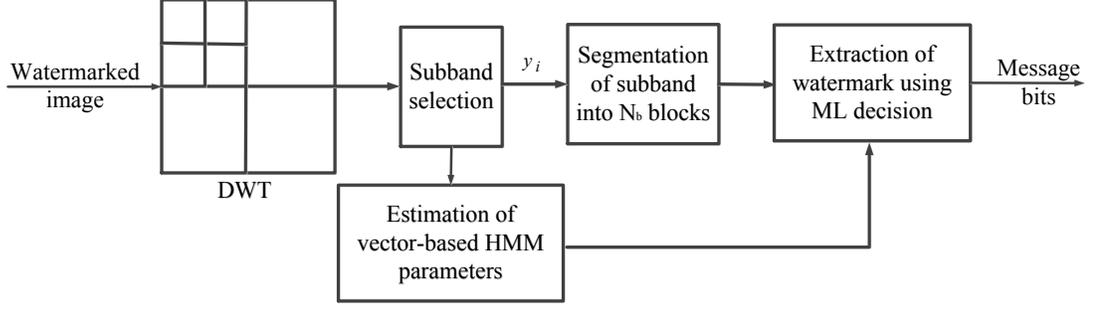


Figure 4.3. The proposed multiplicative watermark decoder scheme using the vector-based HMM.

$$\prod_{i \in B_k} f_Y(y_i | b_k = 1) \underset{H_0}{\overset{H_1}{>}} \prod_{i \in B_k} f_Y(y_i | b_k = -1) \quad (4.12)$$

Applying the natural logarithm on both sides of this equation, the optimum decoder $l_k(y)$ can be obtained as

$$l_k(y) = \sum_{i \in B_k} \ln \frac{f_Y(y_i | b_k = 1)}{f_Y(y_i | b_k = -1)} \underset{H_0}{\overset{H_1}{>}} 0 \quad (4.13)$$

In order to calculate $l_k(y)$, we note that the statistical models for $f_Y(y_i | b_k = \pm 1)$ are

$f_Y(y_i | b_k = \pm 1) = \left(\frac{1}{1 \pm \alpha m_i} \right) f_X\left(\frac{y_i}{1 \pm \alpha m_i} \right)$, where $f_X(x)$ indicates the PDF of the wavelet

coefficients of the selected subband of the host image. To obtain the PDF, $f_X\left(\frac{y_i}{1 \pm \alpha m_i} \right)$,

we make use of the M-state vector-based HMM marginal distribution given in (2.19).

Thus, $l_k(y)$ can be obtained after some algebraic manipulations as

$$\begin{aligned}
l_k(y) = & \sum_{i \in B_k} \ln \frac{1 - \alpha m_i}{1 + \alpha m_i} + \\
& \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{\mathbf{y}_i}{1 + \alpha m_i} - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{\mathbf{y}_i}{1 + \alpha m_i} - \mu_q^m \right) \right\}}{\sum_{m=1}^M \frac{p_j^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{\mathbf{y}_i}{1 - \alpha m_i} - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{\mathbf{y}_i}{1 - \alpha m_i} - \mu_q^m \right) \right\}}
\end{aligned} \tag{4.14}$$

where $q = (\log_2 L) - 2$. The k^{th} message bit present in the coefficients can be decoded as

$$\hat{b}_k = \begin{cases} 1 & Z_k(y) \geq T_k \\ -1 & Z_k(y) < T_k \end{cases} \tag{4.15}$$

where

$$\begin{aligned}
Z_k = & \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{\mathbf{y}_i}{1 + \alpha m_i} - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{\mathbf{y}_i}{1 + \alpha m_i} - \mu_q^m \right) \right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{\mathbf{y}_i}{1 - \alpha m_i} - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{\mathbf{y}_i}{1 - \alpha m_i} - \mu_q^m \right) \right\}} \\
T_k = & \sum_{i \in B_k} \ln \frac{1 - \alpha m_i}{1 + \alpha m_i}
\end{aligned} \tag{4.16}$$

4.3.3 Error Analysis

BER is used to analyze the performance of the proposed watermark decoder. The bit error probability is first computed in the absence of any attack. For the optimum decoder,

the bit error probability is given by

$$P_e = \frac{1}{N_B} \sum_{k=1}^{N_B} \frac{1}{2} [P(Z_k(y) > T_k) | H_0) + P(Z_k(y) < T_k) | H_1)] \quad (4.17)$$

To find the probability under the condition of H_0 , $y_i = (1 - \alpha m_i)x_i$; hence $Z_k(y)$ is equal to

$$Z_k(y | H_0) = \sum_{i \in B_k} \ln \frac{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{1 - \alpha m_i}{1 + \alpha m_i} x_i - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{1 - \alpha m_i}{1 + \alpha m_i} x_i - \mu_q^m \right) \right\}}{\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\}} \quad (4.18)$$

Under the condition H_1 , $y_i = (1 + \alpha m_i)x_i$; therefore, $Z_k(y | H_0) = -Z_k(y | H_1)$. It is noted that the sequence m_i is an independent identical random process that can have two values “- 1” and “1” with equal probability. Since $Z_k(y | H_0)$ is the sum of a large number of independent random variables, according to the central limit theorem, it can be approximated by the Gaussian distribution with finite mean and variance under each hypothesis, i.e., (μ_0, σ_0) and (μ_1, σ_1) . The mean under the H_0 hypothesis, μ_0 is given by

$$\begin{aligned}
\mu_0 &= E[Z_k(y | H_0)] \\
&= \sum_{i \in B_k} \left(\begin{aligned} &\frac{1}{2} \ln \left(\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{1-\alpha}{1+\alpha} x_i - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{1-\alpha}{1+\alpha} x_i - \mu_q^m \right) \right\} \right) \\ &+ \frac{1}{2} \ln \left(\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{1+\alpha}{1-\alpha} x_i - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{1+\alpha}{1-\alpha} x_i - \mu_q^m \right) \right\} \right) \\ &- \ln \left(\sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\} \right) \end{aligned} \right) \quad (4.19)
\end{aligned}$$

which can be simplified to

$$\mu_0 = \sum_{i \in B_k} \left(\ln \frac{\sqrt{a_i b_i}}{c_i} \right) \quad (4.20)$$

where

$$\begin{cases} a_i = \sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{1-\alpha}{1+\alpha} x_i - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{1-\alpha}{1+\alpha} x_i - \mu_q^m \right) \right\} \\ b_i = \sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} \left(\frac{1+\alpha}{1-\alpha} x_i - \mu_q^m \right)^T (C_q^m)^{-1} \left(\frac{1+\alpha}{1-\alpha} x_i - \mu_q^m \right) \right\} \\ c_i = \sum_{m=1}^M \frac{p_q^m}{\sqrt{|\det(C_q^m)|}} \exp \left\{ -\frac{1}{2} (x_i - \mu_q^m)^T (C_q^m)^{-1} (x_i - \mu_q^m) \right\} \end{cases} \quad (4.21)$$

The variance σ_0^2 is given by

$$\begin{aligned}
\sigma_0^2 &= E \left[(Z_k(y | H_0) - \mu_0)^2 \right] \\
&= \frac{1}{4} \sum_{i \in B_k} \left(\ln \frac{a_i}{b_i} \right)^2 \quad (4.22)
\end{aligned}$$

Since $Z_k(y | H_0) = -Z_k(y | H_1)$, we have $\mu_1 = -\mu_0$ and $\sigma_1 = \sigma_0$. The error probability

P_e^k for decoding a watermark bit is obtained as

$$\begin{aligned}
P_e^k &= \frac{1}{2} \{P(Z_k(y) > T_k | H_0) + P(Z_k(y) < T_k | H_1)\} \\
&= \frac{1}{2} \left[1 + Q\left(\frac{T_k - \mu_0}{\sigma_0}\right) - Q\left(\frac{T_k - \mu_1}{\sigma_1}\right) \right] \\
&= \frac{1}{2} \left[1 + Q\left(\frac{T_k - \mu_0}{\sigma_0}\right) - Q\left(\frac{T_k + \mu_0}{\sigma_0}\right) \right]
\end{aligned} \tag{4.23}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$. Thus, if the binary message bits “- 1” or “1” are embedded in the host image with the same probability, then the total BER is given by

$$P_e = \frac{1}{N_B} \sum_{k=1}^{N_B} P_e^k. \tag{4.24}$$

The performance of the proposed decoder is evaluated in terms of BER based on (4.24).

4.4. Simulation Results

To evaluate the performance of the proposed multiplicative watermarking scheme, extensive experiments are conducted on a large set of test images taken from [127]. The host and watermarked images corresponding to five of the test images are shown in Figure 4.4. The watermarks are embedded in the host images using messages of length 128 bits with a WDR of -42 dB. It is seen from this figure that there is no noticeable difference between the original and the watermarked images, and hence, the proposed watermark embedding scheme thus ensuring the imperceptibility of the embedded watermark. The objective measure of the peak signal to noise ratio (PSNR) between the original and watermarked images used to evaluate this imperceptibility and the values are

also given in Figure 4.4. The high PSNR values confirm the superior performance of the embedding scheme.

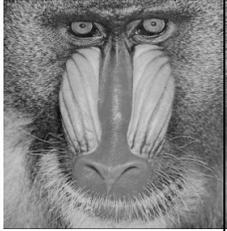
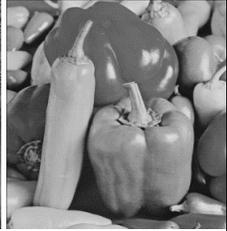
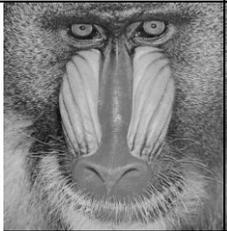
				
(a)	(b)	(c)	(d)	(e)
				
(f)	(g)	(h)	(i)	(j)
PSNR= 48.19dB	PSNR= 42.55dB	PSNR= 49.06 dB	PSNR= 42.60dB	PSNR= 43.84dB

Figure 4.4. Original images (a-e) and corresponding watermarked (f-j) images obtained using the proposed multiplicative watermarking scheme for WDR= - 42 dB.

4.4.1 Detection Results

In order to evaluate the theoretical performance of the proposed detector in terms of the ROC curves, we make use of the expressions for the mean and variance of the test statistic for the two hypotheses H_0 and H_1 obtained in Section 4.2.2. Figure 4.5 depicts the averaged theoretical as well as the experimental ROC curves over a number of test images [47] for P_{FA} varying from 10^{-8} to 10^{-2} for various values of WDR. It is seen from this figure that the ROC curves obtained theoretically are very close to the experimental ones; validating the expressions in (4.7) and (4.8).

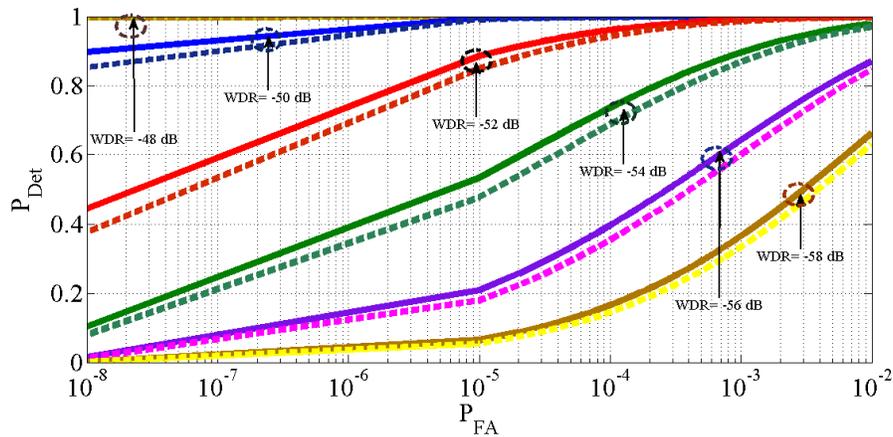


Figure 4.5. Theoretical (solid) and experimental (dashed) ROC curves averaged over 96 test images for the multiplicative vector-based HMM detector for different values of WDR.

In view of this result, henceforth we use the theoretical values of m_0 , m_1 , σ_0 and σ_1 in order to compare the performance of the proposed detector with that of the multiplicative detectors using Cauchy, and GG distributions, in terms of the ROC curves. The ROC

curves of the various detectors are obtained for a given watermarked image with P_{FA} varying from 10^{-8} to 10^{-2} . Figure 4.7 shows the ROC curves averaged over a number of test images for various detectors when $WDR = -50\text{dB}$. It is seen from this figure that the proposed LO vector-based HMM detector has the best performance in that it provides the highest probability of detection for a given probability of false alarm.

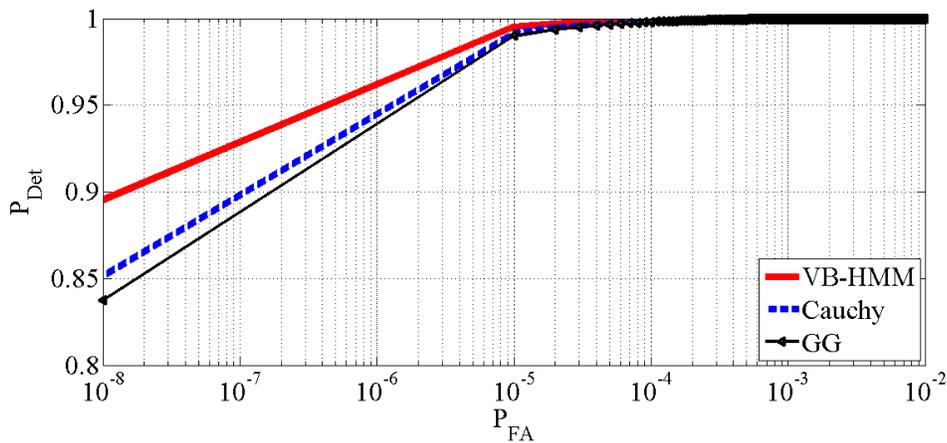


Figure 4.6. ROC curves averaged over 96 test images for various detectors when $WDR = -50\text{dB}$.

In order to study the robustness of the proposed detector against various attacks, we obtain the ROC curved obtained from the proposed multiplicative watermark detector when images are contaminated by JPEG compression, rotation, median filtering, Gaussian filtering and additive Gaussian noise. Figures 4.7 to 4.11 depict ROC curves averaged over 96 test images obtained using the proposed watermark detector as well as those obtained using the Cauchy, and GG-based detectors when the watermarked images

with WDR = -50 dB are JPEG compressed with QF = 30, rotated counter clockwise by 2° , median filtered with a window of size 3×3 , Gaussian filtered with a window of size 3×3 , and corrupted by the Gaussian noise (SNR = 25dB), respectively. It can be seen from these figures that the proposed detector is more robust than the other detectors against any of the attacks considered.

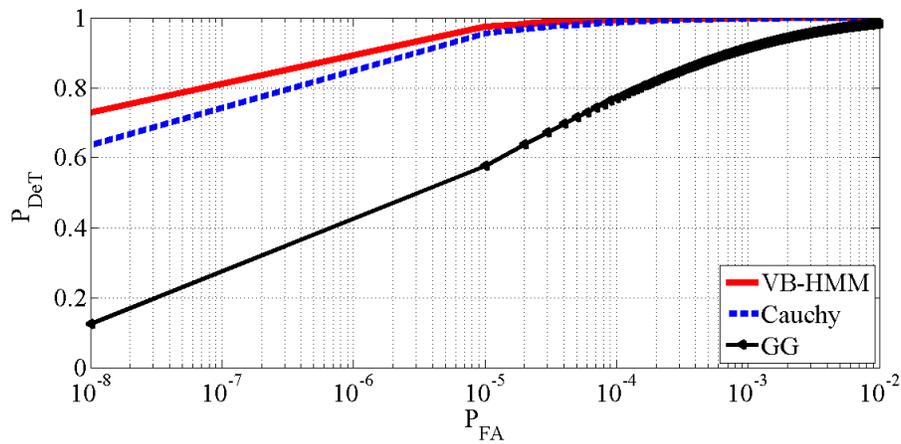


Figure 4.7. ROC curves averaged over a number of test images obtained using various detectors for WDR = -50 dB when image is JPEG-compressed with QF = 30.

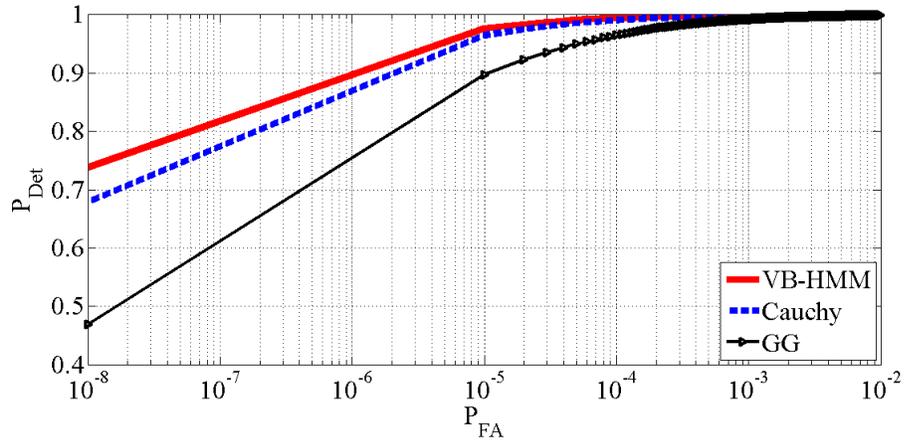


Figure 4.8. ROC curves averaged over a number of images obtained using various detectors and WDR = -50 dB when images are rotated by 2° .

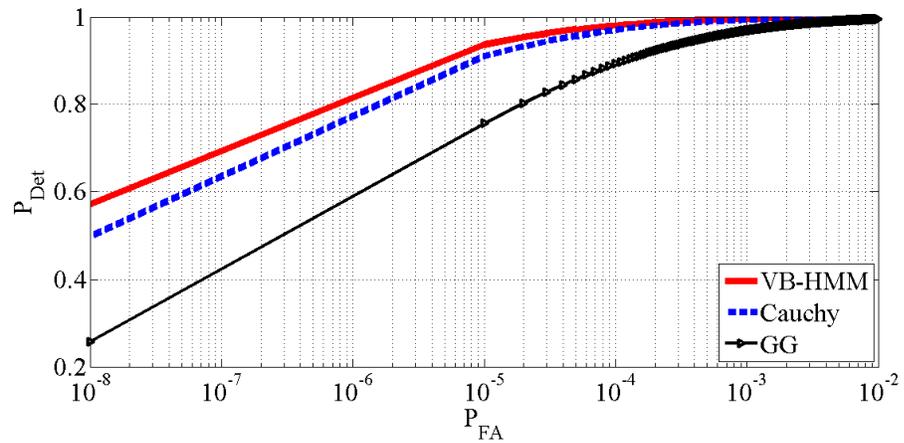


Figure 4.9. ROC curves averaged over a number of test images obtained using various detectors for WDR = -50 dB when images undergo median filtering with a window size of 3×3 .

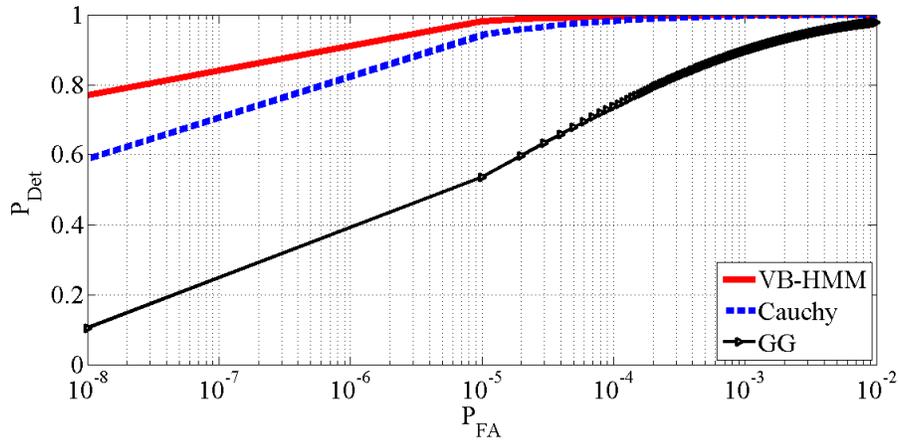


Figure 4.10. ROC curves averaged over a number of test images for the various detectors for WDR = -50 dB when Gaussian filtering with mask 3×3.

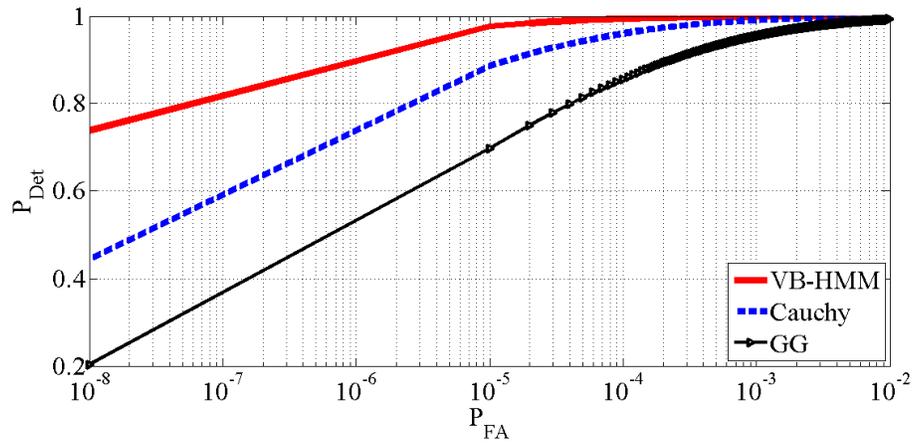


Figure 4.11. ROC curves averaged over a number of test images obtained using various detectors when image is contaminated by the Gaussian noise with $SNR = 25$ and WDR = -50 dB.

In order to compare the performance of the proposed multiplicative detector with the proposed additive LO detector, we compare their results in terms of PSNR and the area under ROC for an equal WDR. Table 4.1 gives the area under ROC curve values average over a number of test images for the proposed additive and multiplicative detectors for the region for a P_{FA} in $[0, 1]$, when the watermarked image is subjected to the following attacks: JPEG compression with QF = 80, 60, 40 and 20, counter clockwise rotations with $\theta = 0.5^\circ, 1^\circ, 1.5^\circ, 2^\circ$, median and Gaussian filtering with window sizes of 3×3 , 5×5 and 7×7 , and Gaussian noise with SNR = 10, 15, 20 and 25 dB. It is seen from this table that that the proposed multiplicative vector-based HMM detector provides the highest area under ROC values, indicating its superior robustness against the various attacks. It is also seen from this table that PSNR value provided by additive watermark detector is higher than the multiplicative detector confirming higher imperceptibility of additive watermarking technique.

Table 4.1: PSNR and area under ROC curves for the region [0, 1], averaged over 96 test images for various detectors against different attacks with WDR = -40 dB.

	Additive VB-HMM	Multiplicative VB-HMM
PSNR (dB)	72.25	44.09
No attack		
Area under ROC	0.9405	1
QF (%)	JPEG	
80	0.9325	1
60	0.9132	1
40	0.9097	0.9956
20	0.9009	0.9820
Angle (degree)	Rotation	
0.5	0.9392	0.9994
1	0.9336	0.9925
1.5	0.9288	0.9889
2	0.9115	0.9721
window size	Median filter	
3×3	0.9370	0.9996
5×5	0.9286	0.9925
7×7	0.9165	0.9898
window size	Gaussian filter	
3×3	0.9386	0.9988
5×5	0.9203	0.9927
7×7	0.9009	0.9801
SNR	Gaussian Noise	
5	0.5537	0.9039
10	0.5909	0.9102
15	0.6559	0.9238
20	0.7506	0.9736

4.4.2 Extraction Results

Performance of the proposed decoder is studied without attack and also in presence of attacks. In Section 4.3.3, an expression for BER was obtained as given by (4.24); in order to calculate BER using (4.24), it is necessary to have the parameter values of the vector-based HMM for the wavelet coefficients of the original image. Since the watermark is embedded with a small value of α , the parameter of the vector-based HMM can be assumed to be the same for the original and watermarked images. Hence, these parameters are estimated from the watermarked coefficient \mathbf{y} .

In order to validate the theoretical values of BER obtained from (4.24), comparisons are made with experimental BER obtained from Monte Carlo simulations. For this purpose, for each of the test images, 1000 pseudo-random message sequences are generated, each sequence embedded in the test image for a given WDR, and decoded using (4.15). The number of errors is computed for each run, and the experimental BER averaged over the 1000 runs. Figure 4.12 shows the theoretical and experimental BER values of the proposed decoder averaged over the 96 test images for various values of WDR. It is seen from this figure that the BER values obtained theoretically are very close to the experimental ones, thus validating the expression for BER given by (4.24).

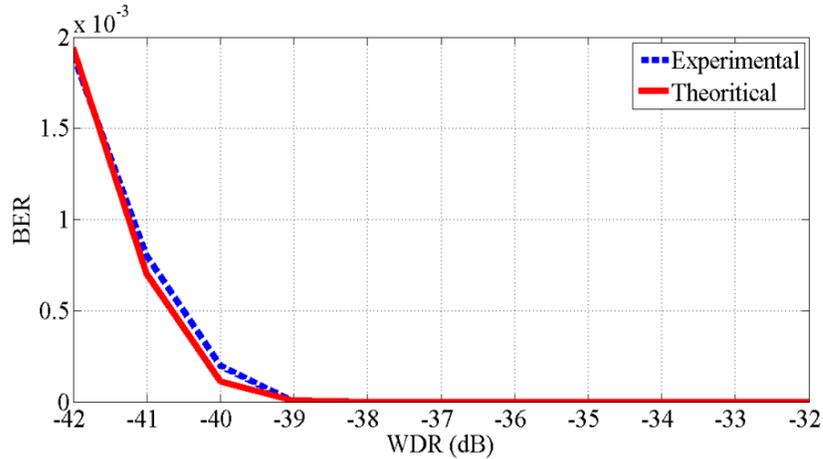


Figure 4.12. Theoretical and experimental BER of the proposed decoder averaged over a number of test images with message length of 128 bits for different WDR values.

The performance of the proposed watermark decoder in the wavelet domain by using the vector-based HMM is examined and compared it to that yielded by using the Cauchy [41] and GG [38], [39] decoders. For this purpose, we use the same framework as shown in Figure 4.3 for all the decoders employing the proposed vector-based HMM, Cauchy or GG distributions for the wavelet coefficients. Table 4.2 gives BER values obtained using the proposed decoder as well as that obtained using the Cauchy and GG-based decoders with message lengths of 64 and 128 bits and WDR = -42 dB test images, namely, *Lena*, *Baboon*, *Peppers*, *Barbara*, *Boat*, *Airplane*, *Man*, *Zelda*, *Elaine* and *Lake*, and the average over all these images. It is seen from this table that the proposed vector-based HMM decoder provides a BER that is lower than that provided by the other decoders.

TABLE 4.2: BER (%) obtained using various decoders for different test images, with message length 64 and 128 bits and WDR= -42 dB.

Image	VB-HMM	Cauchy	GG
Message length = 64 bits			
<i>Lena</i>	0.001281	0.001781	0.002375
<i>Baboon</i>	0.001141	0.001445	0.001563
<i>peppers</i>	0.001391	0.001578	0.001938
<i>Barbara</i>	0.000406	0.001148	0.001875
<i>Boat</i>	0.000563	0.001563	0.002094
<i>Airplane</i>	0.000664	0.001164	0.002023
<i>Man</i>	0.000578	0.000820	0.001609
<i>Zelda</i>	0.001707	0.001977	0.002477
<i>Elaine</i>	0.001195	0.001680	0.001961
<i>Lake</i>	0.001077	0.001531	0.001820
<i>Average</i>	0.001000	0.001468	0.001973
Message length = 128 bits			
<i>Lena</i>	0.002273	0.003609	0.004164
<i>Baboon</i>	0.001117	0.001852	0.002906
<i>peppers</i>	0.002406	0.002898	0.003945
<i>Barbara</i>	0.000906	0.003234	0.004305
<i>Boat</i>	0.001891	0.003383	0.004188
<i>Airplane</i>	0.002039	0.003047	0.004188
<i>Man</i>	0.001109	0.002516	0.003578
<i>Zelda</i>	0.002891	0.004063	0.003516
<i>Elaine</i>	0.002992	0.003875	0.004328
<i>Lake</i>	0.001758	0.002922	0.003922
<i>Average</i>	0.001938	0.003139	0.003485

We also study the robustness of the proposed watermark decoder using the vector-based HMM against common signal processing attacks such as JPEG compression, Gaussian noise, salt and pepper noise, median filtering, rotation, and gamma correction.

- ***JPEG Compression***

The results of BER when the test images, *Lena*, *Baboon*, *Peppers*, *Barbara* and *Boat*, are JPEG-compressed with quality factor changing from 5 to 80 are shown in Figures 4.13 to 4.17. The BER values averaged over 96 different test images is shown in Fig. 4.18. It can be seen from these figures that the proposed decoder is more robust against JPEG compression in comparison to the GG and Cauchy decoders. It is to be noted that for practical compression range of still images, i.e., $QF > 50$, the BER value approaches zero for the proposed decoder.

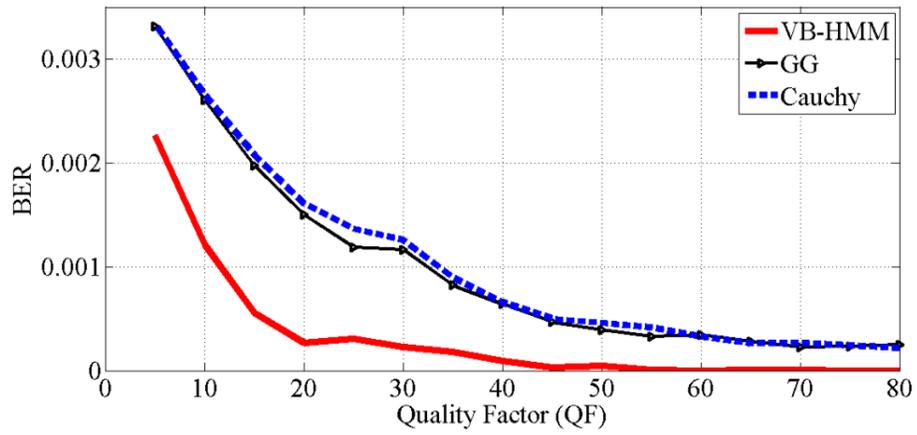


Figure 4.13. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when, *Lena* image is JPEG-compressed with different QFs.

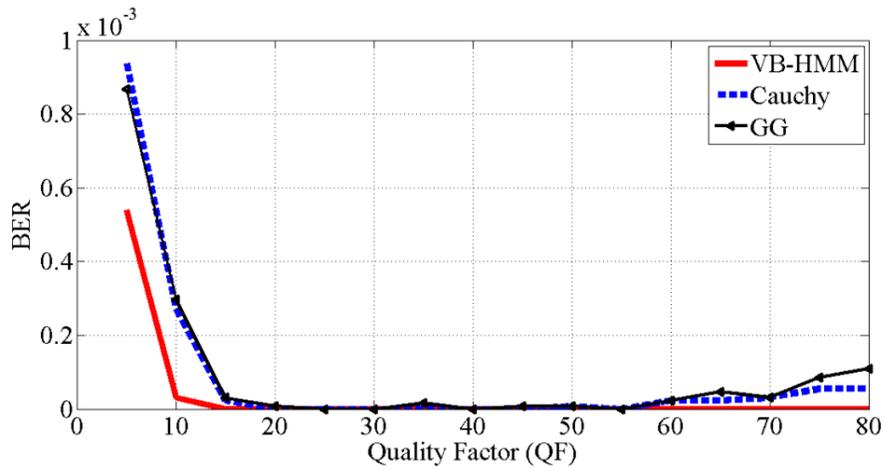


Figure 4.14. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when, *Baboon* image is JPEG compressed with different QFs.

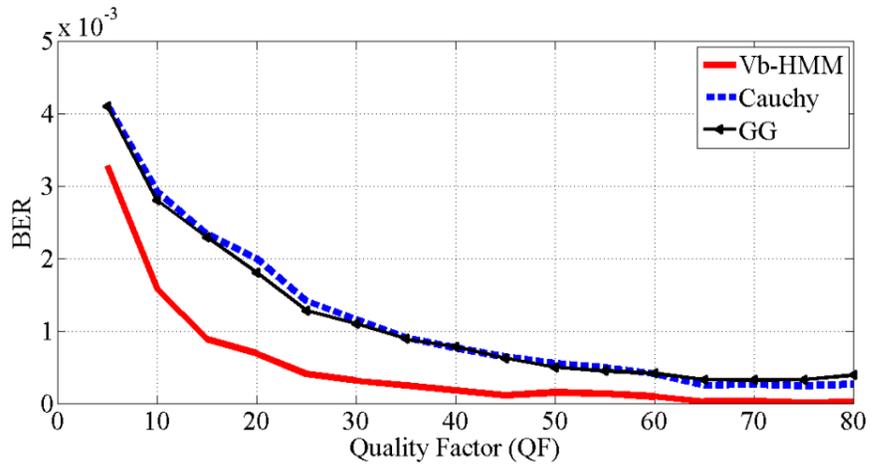


Figure 4.15. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Peppers* image is JPEG-compressed with different QFs.

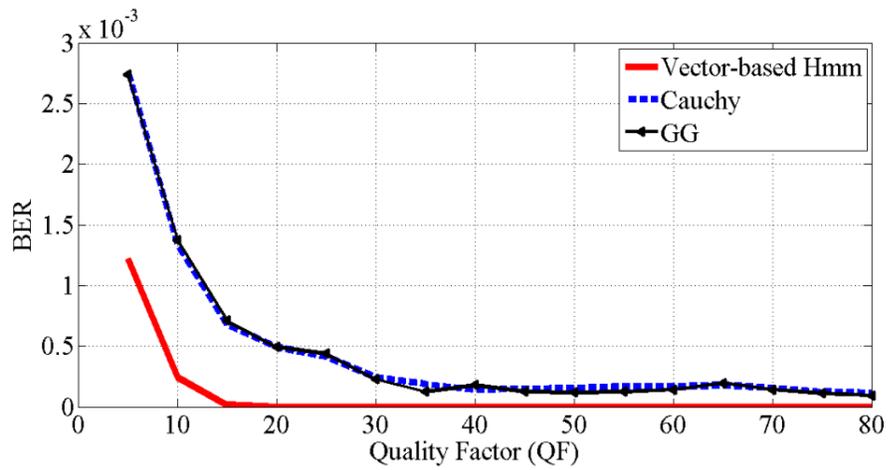


Figure 4.16. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Barbara* image JPEG-compressed with different QFs.

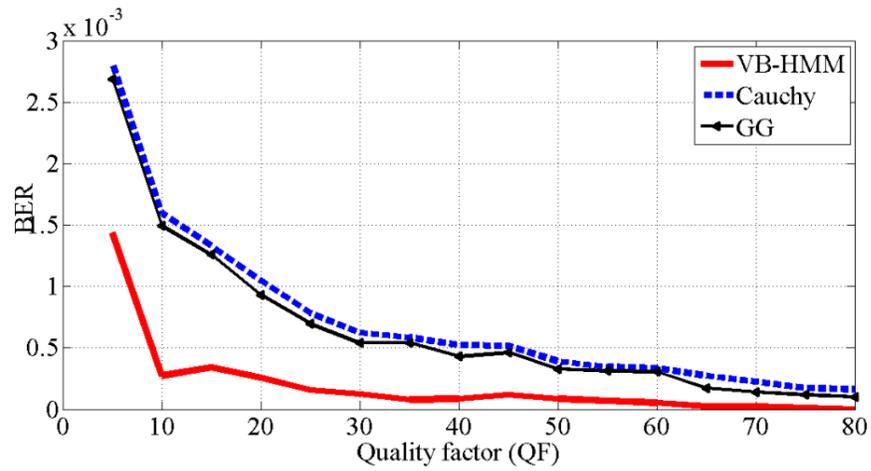


Figure 4.17. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Boat* image is JPEG-compressed with different QFs.

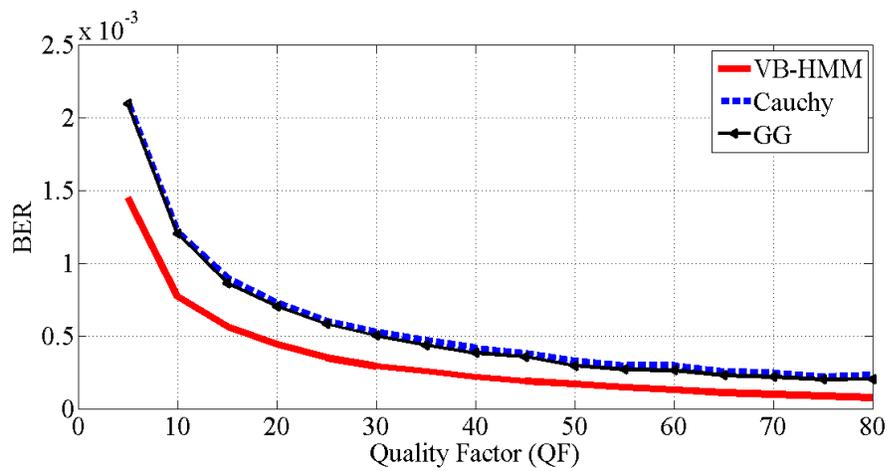


Figure 4.18. BER values of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy decoders when the images are JPEG-compressed with different QFs.

- *Additive Gaussian Noise*

The results of BER when the test images, *Lena*, *Baboon*, *Peppers*, *Barbara* and *Boat*, are contaminated by the additive Gaussian noise with noise standard deviation σ_n varying from 5 to 35 are shown in Figs. 4.19 to 4.23. The BER values averaged over 96 different test images when the images are contaminated by different levels of additive Gaussian noise are shown in Fig. 4.24. It can be seen from these figures that the proposed watermarking scheme using the vector-based HMM exhibits a better performance in presence of Gaussian noise compared to that provided by the decoders based on the GG and Cauchy distributions, especially for higher noise levels, except for the case of the image *Peppers* .

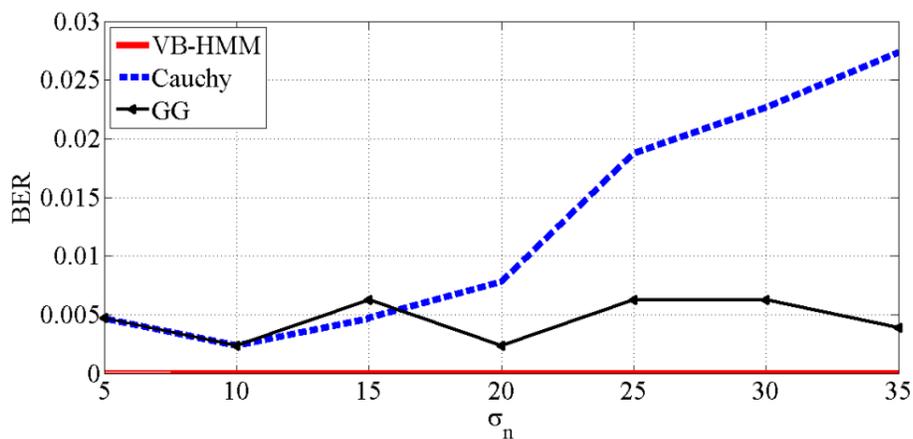


Figure 4.19. BER of the extracted watermark using the proposed VB-HMM, GG and Cauchy decoders when *Lena* image is corrupted by the additive Gaussian noise with different σ_n values.

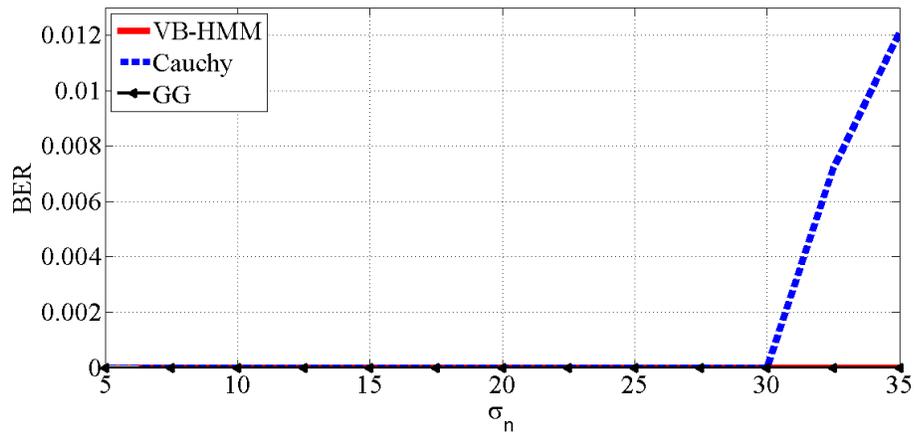


Figure 4.20. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders when *Baboon* image is corrupted by the additive Gaussian noise with different σ_n values.

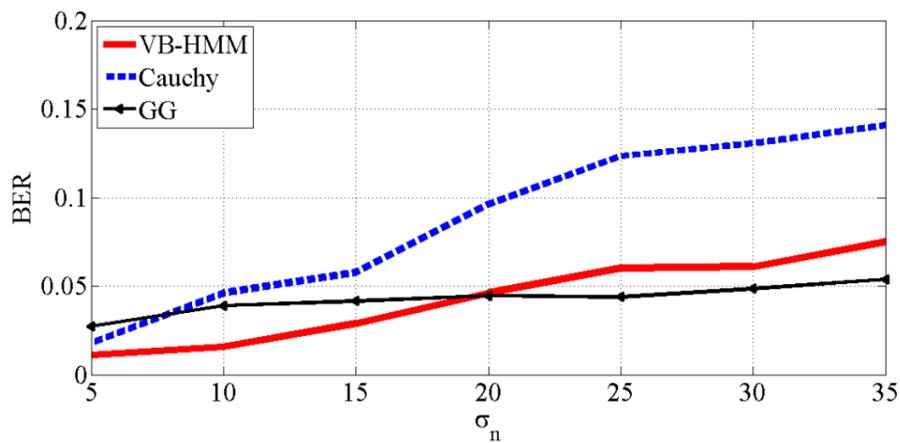


Figure 4.21. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Peppers* image is corrupted by additive Gaussian noise with different σ_n values.

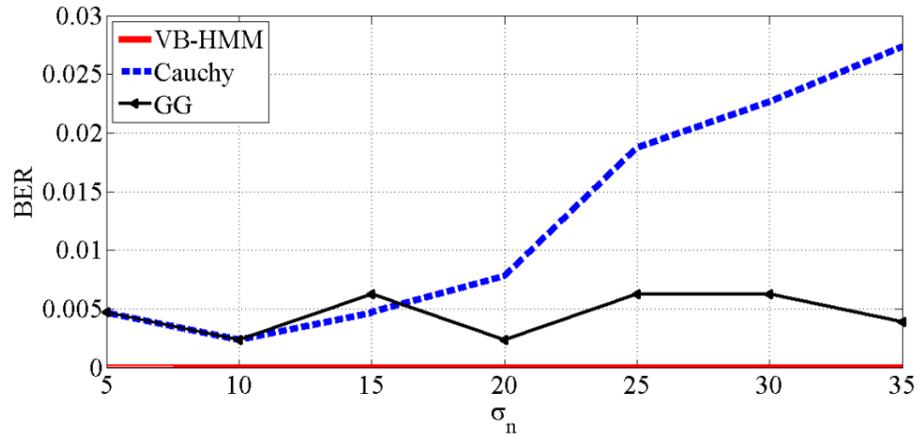


Figure 4.22. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Barbara* image is corrupted by the additive Gaussian noise with different σ_n values.

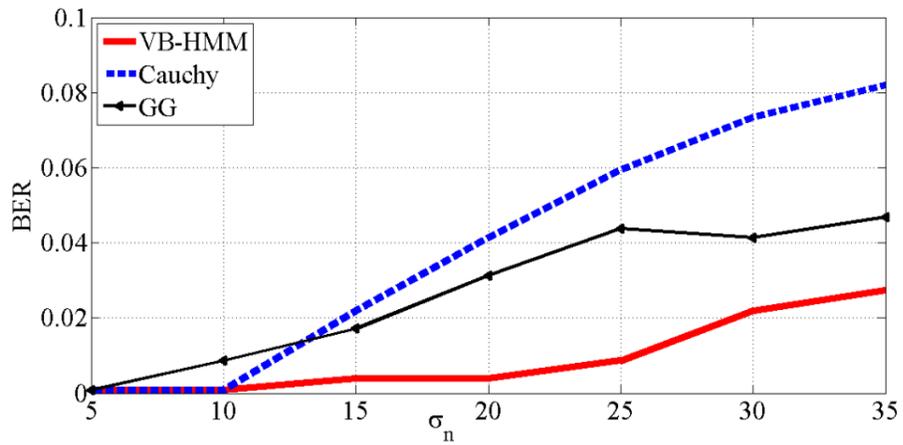


Figure 4.23. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Boat* image is corrupted by the additive Gaussian noise with different σ_n values.

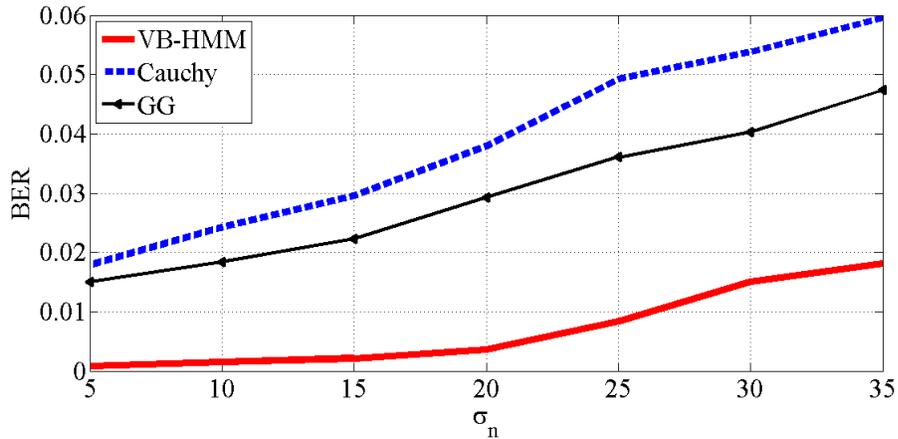


Figure 4.24. BER values of the extracted watermark averaged over 96 test images using the proposed VB-HMM, GG and Cauchy decoders, when the images are corrupted by additive Gaussian noise with different σ_n values.

- ***Salt and Pepper Noise***

Salt and pepper noise is the most commonly used long-tailed noise in image processing. The results of BER, when the different test images corrupted by salt and pepper noise, are shown in Table 4.3. It can be seen from this table that the proposed watermarking scheme is more robust against salt and pepper noise in comparison to that yielded by the GG and Cauchy-based decoders. It can also be seen from this table that the decoders can perfectly extract the watermark bits in case of the *Baboon* image.

TABLE 4.3: BER (%) of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy decoders when various test images are corrupted by salt and pepper noise with different noise density.

Image	VB-HMM	Cauchy	GG
	$p = 0.05$		
<i>Lena</i>	0.0011	0.0430	0.0203
<i>Baboon</i>	0	0	0
<i>peppers</i>	0.0007	0.0313	0.0133
<i>Barbara</i>	0	0.0016	0.0007
<i>Boat</i>	0.0013	0.0375	0.0219
<i>Airplane</i>	0.0023	0.0453	0.0156
<i>Man</i>	0	0.0086	0.0016
<i>Zelda</i>	0.0070	0.1086	0.0508
<i>Elaine</i>	0.0082	0.1391	0.0563
<i>Lake</i>	0	0.0375	0.0094
<i>Average</i>	0.0020	0.0452	0.0189
	$p = 0.1$		
<i>Lena</i>	0.0076	0.1156	0.0867
<i>Baboon</i>	0	0	0
<i>peppers</i>	0.0078	0.1211	0.0961
<i>Barbara</i>	0.0015	0.0227	0.0109
<i>Boat</i>	0.0044	0.0820	0.0531
<i>Airplane</i>	0.0069	0.1109	0.0875
<i>Man</i>	0.0013	0.0109	0.0094
<i>Zelda</i>	0.0165	0.1852	0.2102
<i>Elaine</i>	0.0214	0.2156	0.2016
<i>Lake</i>	0.0009	0.0297	0.0133
<i>Average</i>	0.0068	0.0893	0.0767

- **Median Filtering**

Robustness of watermark decoder against median filtering, a non-linear filter, is a challenging task since it might destroy the watermark severely. The results of BER when the test images, *Lena*, *Baboon*, *Peppers*, *Barbara* and *Boat*, undergo median filtering with window sizes 3×3 , 5×5 , 7×7 and 9×9 are shown Figures 4.25 to 4.29. The BER values averaged over 96 different test images when the images are median-filtered are shown in Figure 4.30. It can be seen from these figures that the proposed scheme is more robust against median filtering in comparison to the GG and Cauchy-based schemes especially when the window size is bigger than 3×3 .

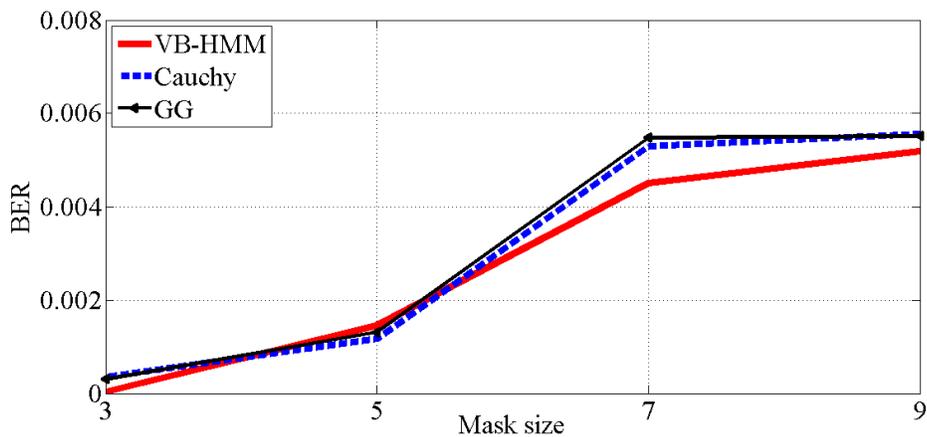


Figure 4.25. BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when *Lena* image undergoes median filtering with different window sizes.

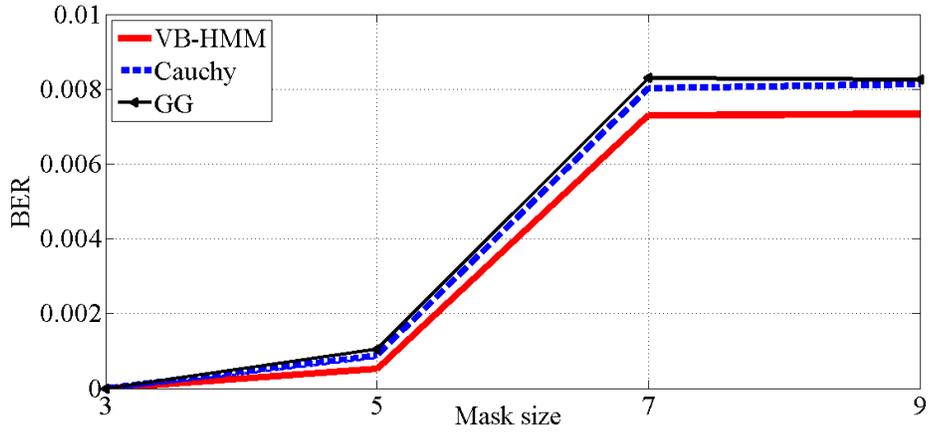


Figure 4.26. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy-based decoders when *Baboon* image undergoes median filtering with different window sizes.

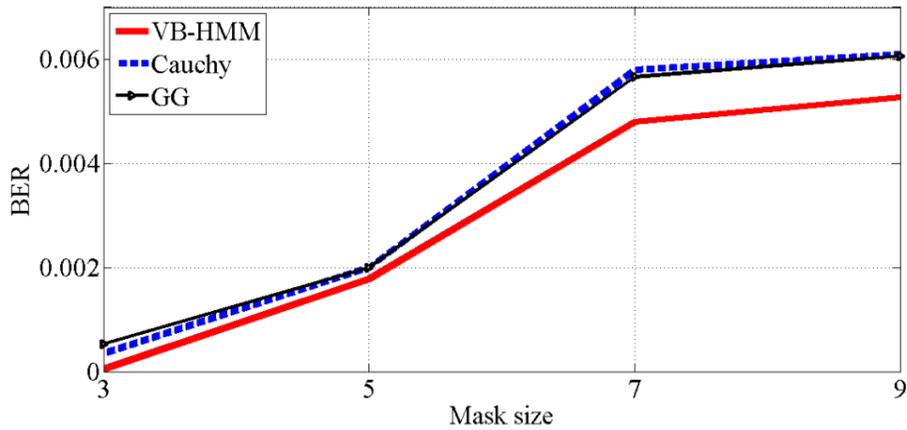


Figure 4.27. BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy decoders, when *Peppers* image undergoes median filtering with different window sizes.

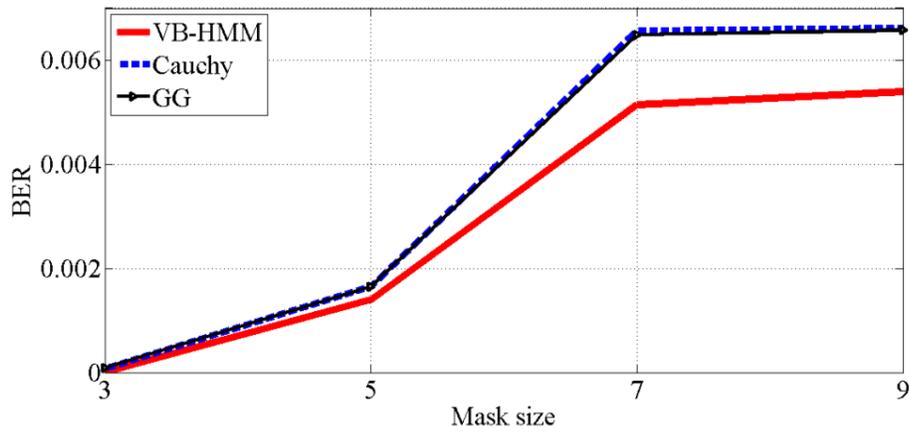


Figure 4.28. BER of the extracted watermark obtained using the proposed vector-based HMM, GG and Cauchy-based decoders when *Barbara* image undergoes median filtering with different window sizes.

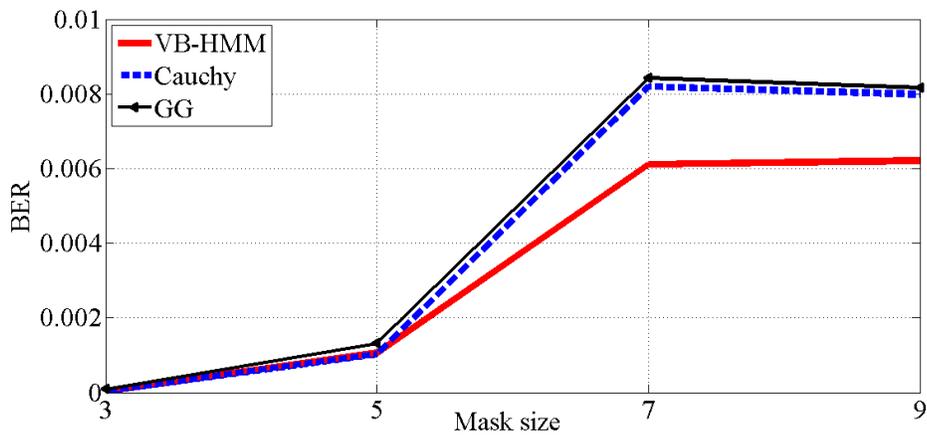


Figure 4.29. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy-based decoders, when *Boat* image undergoes median filtering with different window sizes.

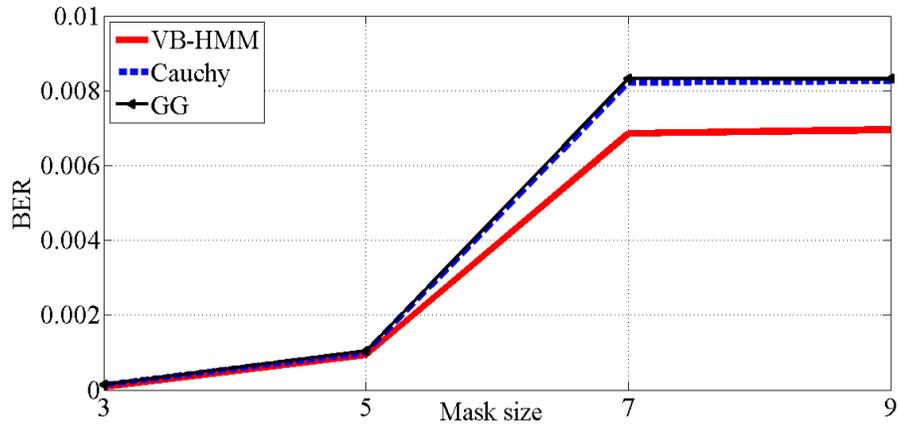


Figure 4.30. BER values of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy-based decoders when the images undergo median filtering with different window sizes.

- **Rotation**

We then investigate the robustness of the proposed watermarking scheme using the vector-based HMM decoder against rotation attack and compare it to schemes using GG and Cauchy-based decoders. The results of BER when the test images, *Lena*, *Baboon*, *Peppers*, *Barbara* and *Boat*, are rotated with different angles are shown in Figures 31-35. The BER values averaged over 96 different test images when the images are rotated by various angles are shown in Figure 4.36. It can be seen from these figures that the proposed scheme is more robust against rotation as compared to the GG and Cauchy-based schemes.

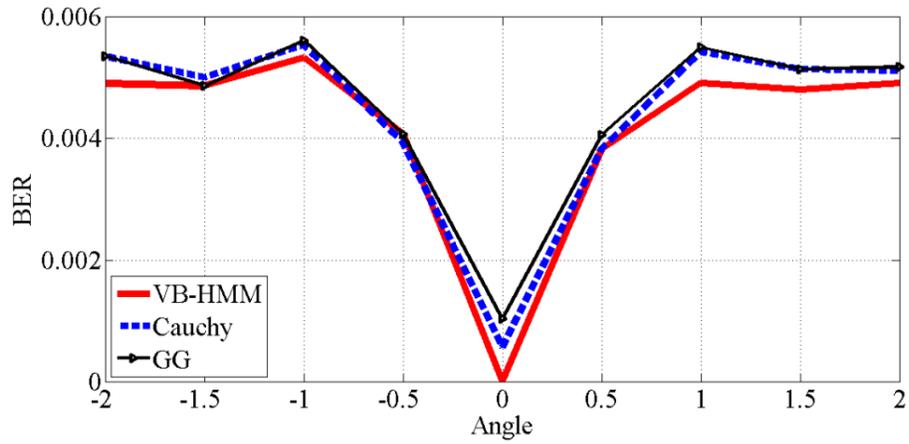


Figure 4.31. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy-based decoders, when *Lena* image is rotated with different angles.

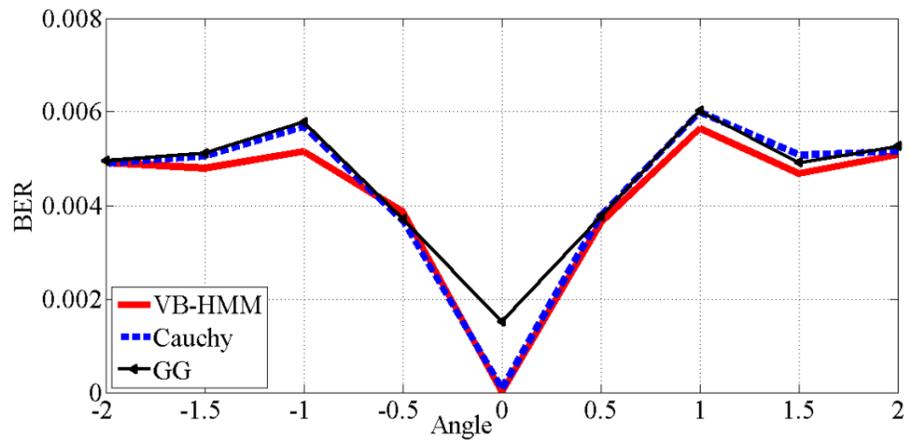


Figure 4.32. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Baboon* image is rotated with different angles.

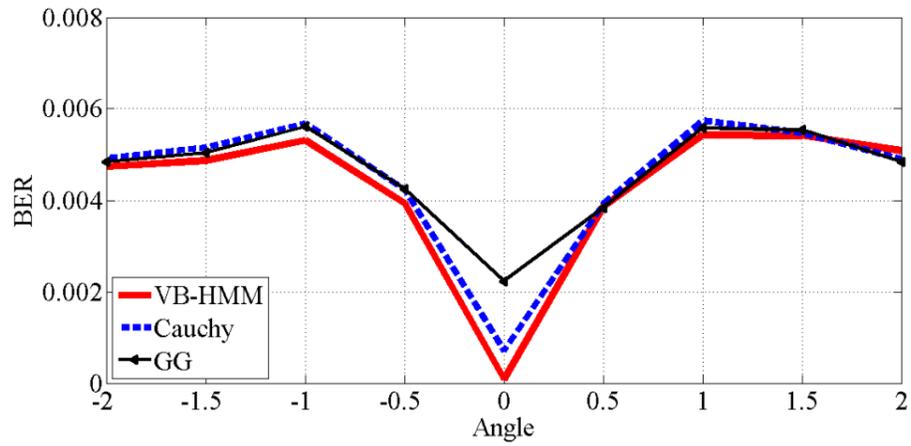


Figure 4.33. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Peppers* image is rotated with different angles.

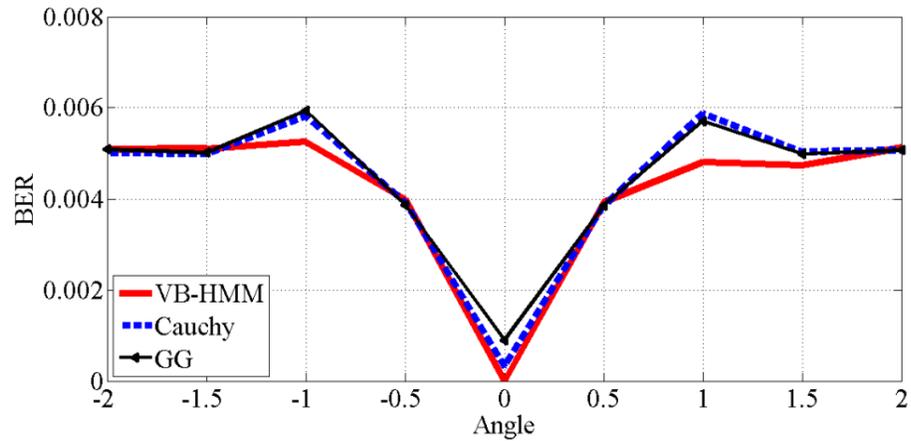


Figure 4.34. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Barbara* image is rotated with different angles.

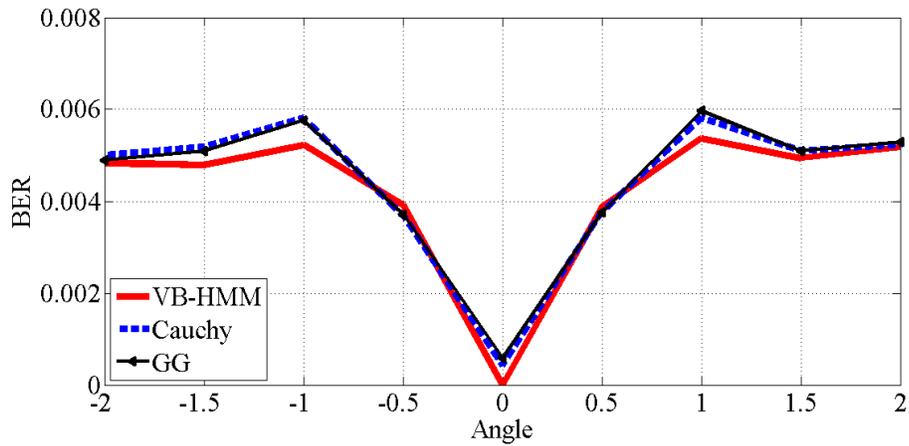


Figure 4.35. BER of the extracted watermark obtained using the proposed VB-HMM, GG and Cauchy decoders, when *Barbara* image is rotated with different angles.

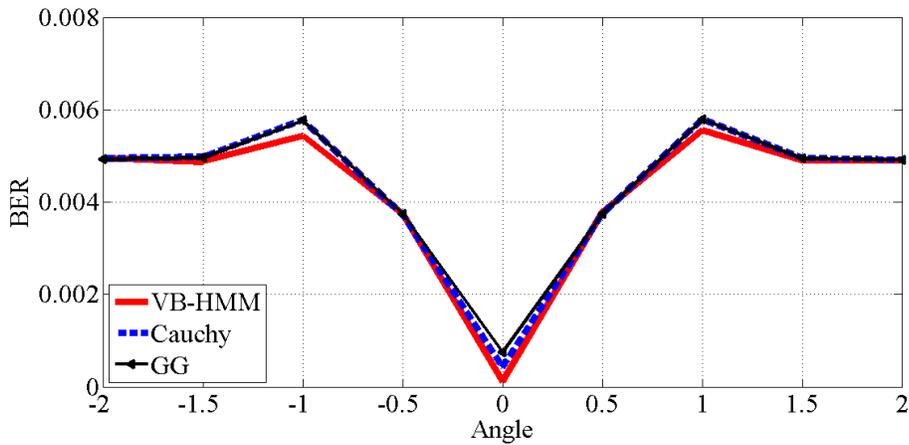


Figure 4.36. BER values of the extracted watermark averaged over 96 test images obtained using the proposed VB-HMM, GG and Cauchy-based decoders, when the images are rotated by different angles.

- **Gamma Correction**

The performance of the proposed decoder is then investigated and compared to the Cauchy and GG-based decoders against the gamma correction attack. Table 4.4 gives BERs when the test images, *Lena*, *Baboon*, *peppers*, *Barbara* and *Boat*, undergo gamma correction with different gamma values 2, 1.5, 0.9 and 0.75. It is seen from this table that the proposed vector-based HMM decoder is more robust against gamma correction as compared to the Cauchy and GG-based decoders.

In order to further investigate the performance of the proposed multiplicative watermarking scheme using the vector-based HMM decoder, we now compare its performance with that of the existing methods including the works in [33], [35], [37], [61] and [53]. In order to make a fair comparison, for a given message length, we have set the PSNR values of the watermarked images in our proposed method to be the same as the values reported in the other works. In other words, the watermark strength is the considered to be the same for various methods.

Table 4.5 gives BER values of the proposed decoder for additive and multiplicative embedded message of 256 bits against different attacks, namely, JPEG compression with $QF = 11$, additive Gaussian noise with $\sigma_n = 10$, and median filtering with a 3×3 window size for the test images, *Barbara*, *Baboon*, *Peppers* and *Lena*. The corresponding BERs for the methods in [33], [37] and [35] are also listed in this table. It can be seen from this table that the proposed multiplicative watermark decoder is more robust than the others against these attacks as indicated by the lower values of BER. It should be noted that due to the reasons mentioned in section 4.1, the proposed watermark decoder for a multiplicative embedding approach provides more robustness than the additive one.

TABLE 4.4: BER values obtained using different decoders when the images undergo gamma correction with different values of gamma.

γ	2	1.5	0.9	0.75
	<i>Lena</i>			
VB-HMM	0.1030	0.1164	0.1194	0.1144
Cauchy	0.1189	0.2867	0.2891	0.2617
GG	0.3117	0.3133	0.3156	0.2961
	<i>Baboon</i>			
VB-HMM	0.0345	0.0354	0.0358	0.0358
Cauchy	0.2344	0.2250	0.2289	0.2266
GG	0.2641	0.2625	0.2500	0.2648
	<i>Peppers</i>			
VB-HMM	0.1359	0.1409	0.1351	0.1385
Cauchy	0.2695	0.2805	0.2555	0.2734
GG	0.2938	0.3016	0.2906	0.2984
	<i>Barbara</i>			
VB-HMM	0	0	0	0
Cauchy	0.0055	0.0547	0.1148	0.1414
GG	0.0055	0.0258	0.0437	0.0219
	<i>Boat</i>			
VB-HMM	0	0.0023	0.0023	0.0023
Cauchy	0.0219	0.1133	0.1220	0.1422
GG	0.0187	0.0367	0.0773	0.0641

TABLE 4.5: BERs (%) obtained using the proposed additive and multiplicative watermarking scheme as well as that obtained using the schemes in [33], [37] and [35] under various attacks for the different test images. (Message length = 256 bits, PSNR = 42 dB)

	Multiplicative VB-HMM	Additive VB-HMM	[33]	[37]	[35]
<i>Barbara</i>					
JPEG (QF=11)	0	4.34	16.45	(0.43)	9.64
AWGN $\sigma_n = 10$	0	(1.15)	1.45	0	1.40
Median filtering 3×3	0	(0.89)	24.95	5.03	1.10
<i>Baboon</i>					
JPEG (QF=11)	0.15	3.81	16.95	0.73	9.86
AWGN $\sigma_n = 10$	0	0	1.30	0	(1.28)
Median filtering 3×3	0	0.87	31.65	1.60	5.03
<i>Peppers</i>					
JPEG (QF=11)	0.33	4.05	26.10	(0.55)	10.68
AWGN $\sigma_n = 10$	0	1.19	1.25	(0.07)	1.32
Median filtering 3×3	0	0	29.35	0.16	1.17
<i>Lena</i>					
JPEG (QF=11)	0.28	(7.93)	29.80	N/A	8.64
AWGN $\sigma_n = 10$	0	(1.24)	1.45	N/A	1.85
Median filtering 3×3	0	0	30.80	N/A	0

In Table 4.6, we compare the robustness of the proposed decoder for additive and multiplicative for an embedded message of 128 bits with that of the works in [53] and [35], when the watermarked *Barbara* and *Baboon* images undergo JPEG compression with QF = 20, additive noise with $\sigma_n = 20$ and salt and pepper noise with $p = 0.05$. It is seen from this table that the proposed multiplicative decoder provides lower BERs than the other decoders do, indicating its higher robustness.

Table 4.7 gives BER values for the proposed multiplicative decoder for an embedded message of 128 bits as well as that of the methods in [36] and [35], when the *Lena* image is contaminated by the additive Gaussian noise for various values of the noise standard deviation and is JPEG-compressed with different values of quality factor. It is seen from this table that the proposed vector-based HMM decoder outperforms those in [36] and [35] in terms of providing lower BERs.

Table 4.8 gives BER values for the proposed multiplicative decoder for an embedded message of 64 bits against different attacks, namely, JPEG compression with QF = 5 and 20, additive median filtering with window sizes 5×5 , 7×7 and 9×9 , salt and pepper noise with $p = 0.08$, and rotation of 0.5° , for the test images, *Peppers*, *Baboon* and *Lena*. The corresponding BERs for methods in [37], [35] and [61], are also listed in this table. It can be seen from this table that the proposed watermark decoder is more robust than the other decoders against these attacks.

TABLE 4.6: BER (%) obtained using the proposed additive and multiplicative watermarking schemes as well as that of the schemes in [53] and [35], when watermarked images are under various attacks for the different images. (Message length = 128 bits)

	Multiplicative VB-HMM	Additive VB-HMM	[53]	[35]
<i>Barbara</i> , PSNR=36 dB				
JPEG (QF=20)	0	(0.03)	0.4	0
AWGN $\sigma_n = 20$	(0.3)	0.43	0.1	1.07
Salt & pepper (p=0.05)	0	(0.13)	1.48	0.43
<i>Baboon</i> , PSNR=39 dB				
JPEG (QF=20)	0	0	(1.89)	0
AWGN $\sigma_n = 20$	0.13	(0.27)	0.30	1.48
Salt & pepper (p=0.05)	0	(0.18)	2.89	0.89

TABLE 4.7: BER (%) obtained using the proposed multiplicative watermarking scheme as well as that obtained using the schemes in [36] and [35], when watermarked *Lena* image is under various attacks. (Message length=128 bits, PSNR = 45 dB)

	Multiplicative VB-HMM	[36]	[35]
σ_n	AWGN		
5	0	0	0
20	2.15	10.16	2.34
35	8.17	13.44	20.31
QF	JPEG		
4	0.21	37.5	32.03
10	0.12	3.91	6.25
16	0	0	0
20	0	0	0

TABLE 4.8: BERs obtained using the proposed multiplicative watermarking scheme as well as that obtained using the schemes in [37], [35], [61] and [51], when watermarked images are under various attacks. (Message length = 64 bits, PSNR = 42 dB)

	Proposed multiplicative	[37]	[35]	[61]	[51]
<i>Peppers</i>					
JPEG (QF=5)	0.32	0.78	N/A	6.25	N/A
JPEG (QF=20)	0	0	0.06	0	N/A
Median filter 5×5	0.15	0	1.56	7.81	5.31
Median filter 7×7	0.04	0	0	9.36	17.18
Median filter 9×9	0.05	3	4.62	51.56	28.75
S&P (p=0.08)	0.03	N/A	0.40	2.51	N/A
Rotation ($\theta=0.5$)	0.41	0	22.87	40.63	N/A
<i>Baboon</i>					
JPEG (QF=5)	0.05	0	N/A	4.69	N/A
JPEG (QF=20)	0	0	0	0	N/A
Median filter 5×5	0.13	1.55	0.50	12.50	20.93
Median filter 7×7	0.42	4.88	3.81	12.50	30.62
Median filter 9×9	0.49	0.88	12.31	78.13	35.00
S&P (p=0.08)	0	N/A	0.4	3.34	N/A
Rotation ($\theta=0.5$)	0.53	3.33	20.81	45.31	N/A
<i>Lena</i>					
JPEG (QF=5)	0.02	N/A	N/A	N/A	N/A
JPEG (QF=20)	0	N/A	0	0	N/A
Median filter 5×5	0	N/A	0	9.38	N/A
Median filter 7×7	0	N/A	0.65	12.50	N/A
Median filter 9×9	3.5	N/A	3.84	51.56	N/A
S&P (p=0.08)	0.13	N/A	0.28	2.67	N/A
Rotation ($\theta=0.5$)	0.44	N/A	20.87	43.75	5.46

In this section, a meaningful message, e.g., a logo, is chosen as a watermark. In this experiment, a binary logo of size 32×32 pixels is inserted in the original image, multiplicatively. In order to compare the extracted watermark \hat{b} with the original watermark logo b , the normalized correlation (NC) given by

$$NC = \frac{\sum_{k=1}^{N_B} b_k \hat{b}_k}{\sqrt{\sum_{k=1}^{N_B} b_k^2} \sqrt{\sum_{k=1}^{N_B} \hat{b}_k^2}} \quad (4.25)$$

is used [125], [126]. Table 4.9 shows the original watermark as well as the extracted ones when the watermarked *Lena* image undergoes JPEG compression with $QF = 10$, additive Gaussian noise corruption with $\sigma_n = 20$, salt and pepper noise contamination with $p = 0.05$, median filtering 3×3 , rotation with $\theta = 0.5^\circ$ and gamma correction with $\gamma = 0.9$. The NC values are also compared in this figure. It is obvious from the results of this figure that the proposed decoder has a good performance in extracting watermark logo in presence of various attacks.

TABLE 4.9: Extracted watermark logo for watermarked *Lena* image of size 512×512 with multiplicative decoder in presence of different attacks when message length = 1024 bits.

	
(a) Original watermark	(b) With no attack NC=1
	
(c) JPEG $QF=10$ NC= 0.9970	(d) AWGN $\sigma_n = 20$ NC= 0.9627
	
(e) Salt and pepper $p=0.05$ NC= 0.9985	(f) Median filtering 3×3 NC=1
	
(g) Rotation $\theta = 0.5^\circ$ NC=0.9940	(h) Gamma correction $\gamma = 0.9$ NC=0.9296

4.5. Summary

In this chapter, robust blind multiplicative watermark decoder and detector have been proposed using the vector-based HMM for the image wavelet coefficients. The watermark bits have been first embedded in the wavelet transformed image using the multiplicative embedding approach. The detector has been formulated by employing a binary hypothesis test for the cases when there does and does not exist a watermark in the received image. This test has been reduced to a log-likelihood ratio test exploiting the statistical properties of the image coefficients. Closed-form test statistic leading to the receiver operating characteristic curves has been derived. The watermark decoder has been developed based on the vector-based HMM using the maximum likelihood criterion. Theoretical closed-form expression for the watermark decoder has been derived and validated experimentally through Monte Carlo simulations. A closed-form expression for the bit error rate of the decoder has also been derived. The performance of the proposed watermark detector and decoder have been investigated in detail by conducting several experiments on a large number of test images and comparing the results with that of the other existing methods. It has been shown that the proposed watermark detector for a multiplicative embedding of the watermark is superior to other existing detectors, including its additive counterpart, by providing higher watermark detection rates with or without the imposed of any distortions. It has also been shown that the proposed multiplicative watermark decoder yields a superior performance, by providing bit error rate that is lower than that provided by other decoders. The robustness of the proposed multiplicative watermarking scheme using the vector-based HMM decoder against different kinds of attacks has been studied and shown to be more robust than of the others.

CHAPTER 5

Conclusion

5.1 Concluding Remarks

This thesis has been concerned with digital image watermarking problem in the wavelet domain by developing watermark detection and extraction techniques using the statistical properties of the wavelet image coefficients. The modeling of images in the wavelet domain has been first investigated. It has been shown that the vector-based hidden Markov model can fit more accurately the distributions of the wavelet subband coefficients of natural images. This is due to the fact that this distribution can not only capture the non-Gaussian behavior of the wavelet coefficients of images but also take into account their inter-scale and inter-orientation dependencies.

In the context of digital image watermarking as a possible solution for copyright protection and secure communication, new watermark detectors and decoders based on the statistical properties of the wavelet coefficients of images have been designed. To this end, motivated by the capability of the vector-based hidden Markov model in modeling the wavelet coefficients of images, first, new blind additive and multiplicative watermark detectors for grayscale images have been designed. The proposed detectors have been developed based on the Bayesian log-likelihood ratio criterion for the watermark detection. The performances of the designed detectors have been evaluated through extensive experiments. It has been shown that these detectors are capable of providing rates of detection for a given probability of false alarm higher than that provided by other

existing detectors. The robustness of the proposed detectors against various possible distortions of the watermarked images has been studied and shown to be superior to that of other detectors.

A blind additive and multiplicative watermark decoder in the wavelet domain for watermarking of greyscale images have also been proposed. By employing the hidden Markov model, the statistical watermark decoders have been designed based on the maximum likelihood criterion and closed-form expressions for the watermark decoders have been derived. It has been shown that the performance of the proposed watermark decoders are superior to that of the other decoders in being able to extract the watermark bits with a lower error both with and without attacks on the watermarked images.

5.2 Scope for Future Work

While the research work undertaken in this thesis has focused on developing efficient and cost-effective techniques for various estimators and detectors, there are a number of additional studies that can be undertaken along the ideas developed in this thesis. Some of the possible studies are as follows:

- The vector-based hidden Markov model employed in this thesis can be also applied to RGB color images, where their inter-channel dependencies can be effectively captured in the wavelet domain.
- The proposed watermarking schemes for images can be extended to videos to protect them from their illegal use and unauthorized duplication. In intra-mode video processing, the hidden Markov model can be applied to model the video

frames in designing watermark detection and extraction techniques. In addition to using HMM for intra-mode modeling, this model can be used advantageously to capture inter-frame dependencies of the video frames, i.e., for estimating the motion vectors.

REFERENCES

- [1] Z. Azimifar, P. Fieguth, and E. Jernigan, “Textures and wavelet-domain joint statistics”, *Image Analysis and Recognition*, vol. 3212, pp. 331-338, 2004.
- [2] R. Chandramouli and N. D. Memon. “On sequential watermark detection”, *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 1034-1044, April 2003.
- [3] M. Mihcak, I. Kozintsev, and K. Ramchandran, “Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising”, in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3253-3256, March 2009.
- [4] J. Ho and W. Hwang, “Wavelet Bayesian network image denoising”, *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1277-1290, 2013.
- [5] M. Amini, M.O. Ahmad, and M.N.S. Swamy, “A new Map estimator for wavelet domain image denoising using vector-based hidden Markov model”, in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 445-448, May 2015.
- [6] M. Amini, M.O. Ahmad and M.N.S. Swamy, “Image denoising in wavelet domain using the vector-Based hidden Markov model”, in *Proc. IEEE International conference on New Circuits and Systems (NEWCAS)*, pp. 29-32, June 2014.
- [7] M. Amini, M.O. Ahmad and M.N.S. Swamy, “SAR image despeckling using vector-based hidden Markov model in wavelet domain”, in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pp. 1-4, May 2016.

- [8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, “Image coding using wavelet transform”, *IEEE Transactions on image processing*, vol. 1, no. 2, pp. 205-220, April 1999.
- [9] S. Lo-Presto, K. Ramchandran, and T. Orchard, “Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework”, in *Proc. Data Compression Conference (DCC)*, pp. 221–230, 1997.
- [10] R. W. Buccigrossi and E. P. Simoncelli, “Image compression via joint statistical characterization in the wavelet domain”, *IEEE Transactions on Image Processing*, vol. 8, no. 12, pp. 1688-1701, December 1999.
- [11] H. Choi and R. Baraniuk, “Multiscale texture segmentation using wavelet-domain hidden Markov models”, in *Proc. Asilomar Conference on Signals, Systems, and Computers*, pp. 1692-1697, 1998.
- [12] J. Li, R. M. Gray, and R. A. Olshen, “Multiresolution image classification by hierarchical modeling with two-dimensional hidden Markov models,” *IEEE Transactions on Information Theory*, pp. 1826-1841, August 2000.
- [13] M. Do and M. Vetterli, “Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance,” *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 146-158, February 2002.
- [14] M. Barni, F. Bartolini, and A. Piva, “Improved wavelet-based watermarking through pixel-wise masking”, *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 783-791, May 2001.

- [15] M. M. Rahman, M. O. Ahmad and M. N. S. Swamy, "A new statistical detector for DWT-based additive image watermarking using the Gauss-Hermite expansion", *IEEE Transactions on Image Processing*, vol. 18, no. 8, pp. 1782-1796, August 2009.
- [16] T. M. Ng. and H. K. Garg, "Maximum-likelihood detection in DWT domain image watermarking using Laplacian modeling", *IEEE Signal Processing Letters*, vol. 12, no. 4, pp. 285-288, April 2005.
- [17] X. Huang and B. Zhang, "Robust detection of additive watermarks in transform domains", *IEE Proceedings-Information Security*, vol. 153, no. 3, pp. 97-106, September 2006.
- [18] Y. Bian and S. Liang, "Locally optimal detection of image watermarks in the wavelet domain using Bessel K-form distribution", *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2372-2384, June 2013.
- [19] Y. Bian and S. Liang, "Image watermark detection in the wavelet domain using Bessel K densities", *IET Image Processing*, vol. 2, no. 1, pp.1-13, 2013.
- [20] X. Huang, B. Zhang, "Statistically robust detection of multiplicative spread-spectrum watermarks", *IEEE Transactions on Information Forensics Security*, vol. 88, no. 1, pp. 117-130, 2007.
- [21] Q. Cheng and T. S. Huang, "Optimum detection of multiplicative watermarks using locally optimum decision rule", in *Proc. IEEE International Conference on Multimedia Expo*, pp. 425-428, 2001.

- [22] Q. Cheng and T. S. Huang, "An additive approach to transform-domain information hiding and optimum detection structure", *IEEE Transactions on Multimedia*, vol. 3, no. 3, pp. 273-284, September 2001.
- [23] K. Mairgiotis, P. Nikolas and Y. Yang, "New additive watermark detectors based on a hierarchical spatially adaptive image model", *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 29-37, March 2008.
- [24] W. Zeng and B. Liu, "A statistical watermark detection technique without using original images for resolving rightful ownerships of digital images", *IEEE Transactions on Image Processing*, vol. 8, no. 11, pp. 1534-1548, 1999.
- [25] I. J. Cox, J. Kilian, F. T. Leighton and T. Shamoan, "Secure spread spectrum watermarking for multimedia", *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673-1687, 1997.
- [26] JR. Kwitt, P. Meerwald and A. Uhl, "Lightweight detection of additive watermarking in the DWT domain", *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 474-484, February 2011.
- [27] X. Zhang, Z. Wang and X. Wang, "Correlation-and-bit-aware additive spread spectrum data hiding for Laplacian distributed host image signals", *Signal Processing: Image Communication*, vol. 29, no. 10, pp. 1171-1180, November 2014.
- [28] H. W. Kim, D. Choi, H. Choi and T. Kim, "Selective correlation detector for additive spread spectrum watermarking in transform domain", *Signal Processing: Image Communication*, vol. 90, no. 8, pp. 2605-2610, August 2010.

- [29] M. Amini, M.O. Ahmad, and M.N.S. Swamy, "A new blind wavelet domain watermark detector using hidden Markov model", in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2285-2288, 2014.
- [30] M. Amini, M.O. Ahmad and M.N.S. Swamy, "A new locally optimum watermark detection using vector-based hidden Markov model in wavelet", under review in *Signal processing*, 2016.
- [31] M. Amini, M.O. Ahmad and M.N.S. Swamy, "A robust multibit multiplicative watermark decoder using vector-based hidden Markov model in wavelet Domain", *IEEE Transactions on Circuit and System for Video Technology*, pp. 1-12, September 2016.
- [32] M. Amini, M.O. Ahmad and M.N.S. Swamy, "Digital watermark extraction in wavelet domain using hidden Markov model", *Multimedia Tools and Applications*, pp. 1-19, September 2016.
- [33] Y. Wang, J. F. Doherty, and R. E. Van Dyck, "A wavelet-based watermarking algorithm for ownership verification of digital images", *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 77-88, February 2002.
- [34] S.-H. Wang and Y.-P. Lin, "Wavelet tree quantization for copyright protection watermarking", *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 154-165, February 2004.
- [35] E. Nezhadarya, Z. J. Wang, and R. K. Ward, "Robust image watermarking based on multiscale gradient direction quantization", *IEEE Transactions on Information Forensics and Security*, no. 4, vol. 6, pp. 1200-1213, December 2011.

- [36] N. K. Kalantari and S. M. Ahadi, "A logarithmic quantization index modulation for perceptually better data hiding", *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1504-1517, June 2010.
- [37] M.A. Akhaee, S. Sahraeian, B. Sankur, and F. Marvasti, "Robust scaling-based image watermarking using maximum-likelihood decoder with optimum strength factor", *IEEE Transactions on Multimedia*, vol. 11, no. 5, pp. 822-833, August 2009.
- [38] K. Zebbiche and F. Khelifi, "Efficient wavelet-based perceptual watermark masking for robust fingerprint image watermarking", *IET Image Processing*, vol. 8, no. 1, pp. 23-32, January 2014.
- [39] J. R. Hernandez, M. Amado and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: detector performance analysis and a new structure", *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 55-68, 2000.
- [40] A. Briassouli and M.G. Strintzis, "Locally optimum nonlinearities for DCT watermark detection", *IEEE Transactions on Image Processing*, vol. 13, no. 12, pp. 1604-1617, December 2004.
- [41] A. Briassouli, P. Tsakalides and A. Stouraitis, "Hidden message in heavy-tails: DCT-domain watermark detection using alpha-stable models", *IEEE Transactions on Multimedia*, vol. 7, no. 4, pp. 700-715, August 2005.

- [42] J.M. Fadili and L. Boubchir, "Analytical form for a Bayesian wavelet estimator of images using the Bessel K-form densities", *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 231-240, February 2005.
- [43] M.S. Crouse, R. D. Nowak and R.G. Baraniuk, "Wavelet-based statistical signal processing using hidden Markov models", *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 886-902, April 1998.
- [44] J.K. Romberg, C. Hyeokho and R.G. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models", *IEEE Transactions on Image Processing*, vol. 10, no. 7, pp. 1056-1068, July 2001.
- [45] G. Fan and X. Xia, "Improved hidden Markov models in the wavelet-domain", *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 115-120, January 2001.
- [46] M.N. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models", *IEEE Transactions on Multimedia*, vol. 4, no. 4, pp. 517-527, December 2002.
- [47] J. Sun, D. Gu, Y. Chen, and S. Zhang, "A multiscale edge detection algorithm based on wavelet domain vector hidden Markov tree model", *Pattern Recognition*, vol. 37, no. 7, pp. 1315-1324, July 2004.
- [48] M.R. Luetzgen, W.C. Karl, A. S. Willsky and R. R. Tenney, "Multiscale representations of Markov random fields", *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3377-3396, December 1993.
- [49] N. Nikolaidis and I. Pitas, "Robust image watermarking in spatial domain", *Signal Processing*, vol. 66, pp. 385-403, May 1998.

- [50] H. Sadreazami, and A. Amini, "A robust spread spectrum based image watermarking in ridgelet domain", *International Journal of Electronics and Communications*, vol. 66, no. 5, pp. 364-371, May 2012.
- [51] N. K. Kalantari, S. M. Ahadi, and M. Vafadust, "A robust image watermarking in the ridgelet domain using universally optimum decoder", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 3, pp. 396-406, March 2010.
- [52] H. Sadreazami, M.O. Ahmad, and M.N.S. Swamy, "A study of multiplicative watermark detection in the contourlet domain using alpha-stable distributions", *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4348-4360, October 2014.
- [53] M. A. Akhaee, S. Sahraeian, and F. Marvasti, "Contourlet-based image watermarking using optimum detector in a noisy environment", *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 967-980, April 2010.
- [54] L. Pérez-Freire and F. Pérez-González, "Spread-spectrum watermarking security", *IEEE Transactions on Information Forensics and Security*, Vol. 4, no. 1, pp. 2-24, March 2009.
- [55] Q. Cheng and T. S. Huang, "Robust optimum detection of transform domain multiplicative watermarks", *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 906-924, April 2003.
- [56] B. Chen and G. W. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking an information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423-1443, May 2001.

- [57] S. Pei, J. Chen, “Robustness enhancement for noncentric quantization-based image watermarking”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 12, pp. 1507-1518, December 2006.
- [58] V. R. Doncel, N. Nikolaidis and I. Pitas, “An optimal detector structure for the Fourier descriptors domain watermarking of 2-D vector graphics”, *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 5, pp. 851-863, Sept-Oct 2007.
- [59] M. Barni, F. Bartolini, V. Cappellini, A. Piva, “A DCT-domain system for robust image watermarking”, *Signal Processing*, vol. 66, pp. 357-372, May 1998.
- [60] A. Nikolaidis and I. Pitas, “Asymptotically optimal detection for additive watermarking in DCT and DWT domains”, *IEEE Transactions on Image Processing*, vol. 12, no. 5, pp. 563-571, May 2003.
- [61] N. Bi, Q. Sun, D. Huang, Z. Yang, and J. Huang, “Robust image watermarking based on multiband wavelets and empirical mode decomposition”, *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 1956-1966, August 2007.
- [62] M. Barni, F. Bartolini, A. DeRosa and A. Piva, “Optimum decoding and detection of multiplicative watermarks”, *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 1118-1123, April 2003.
- [63] J. Zhong and S. Huang, “An Enhanced multiplicative Spread Spectrum Watermarking Scheme”, *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 16, no. 12, pp. 1491-1506, December 2006.

- [64] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, "A new decoder for the optimum recovery of non-additive watermarks", *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 755-766, May 2001.
- [65] M. Hamghalam, S. Mirzakuchaki and M. A. Akhaee, "Robust image watermarking using dihedral angle based maximum likelihood detector", *IET Image Processing*, vol. 7, no. 5, pp. 451-463, July 2013.
- [66] M. Hamghalam, S. Mirzakuchaki and M. A. Akhaee, "Geometric modeling of the wavelet coefficients for image watermarking using optimum detector", *IET Image Processing*, vol. 8, no. 3, pp. 162-172, March 2014.
- [67] X. Zhu, J. Ding, H. Dong, K. Hu, and X. Zhang, "Normalized correlation-based quantization modulation for robust watermarking", *IEEE Transactions on Multimedia*, vol. 16, no. 7 pp. 1888-1904, November 2014.
- [68] M. Hamghalam, S. Mirzakuchaki and M. A. Akhaee, "Vertex angle image watermarking with optimal detector", *Multimedia Tools and Applications*, vol. 74, Issue 9, pp. 3077-3098, May 2015.
- [69] F. Balado, "New geometric analysis of spread-spectrum data hiding with repetition coding, with implications for side-informed schemes", *Digital Watermarking*, vol. 3710, pp. 336-350, 2005.
- [70] S. A. Kassam, "Signal Detection in Non-Gaussian Noise", *New York, USA: Springer-Verlag*, 1988.

- [71] A. Abrardo, and M. Barni, “Informed watermarking by means of orthogonal and quasi-orthogonal dirty paper coding”, *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 824-833, February 2005.
- [72] M. Vetterli, and J. Kovacevic. “*Wavelets and subband coding*”, Prentice Hall PTR, Englewood Cliffs, New Jersey, 1995.
- [73] I. Daubechies, “*Ten Lectures on Wavelets*”, 1st ed. Philadelphia, PA: SIAM, 1992.
- [74] S. Mallat, “*A Wavelet Tour of Signal Processing*”, 2nd ed. San Diego, CA: Academic Press, 1999.
- [75] J. C. Goswami and A. K. Chan, “*Fundamentals of Wavelets. Theory, Algorithms, and Application*”, 1st ed. NY: John Wiley & Sons, 1999.
- [76] S. Mallat, “A theory for multiresolution signal decomposition: The wavelet representation”, *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 11, no. 7, pp. 674-693, July 1989.
- [77] G. Strang and T. Nguyen, “*Wavelets and Filter Banks*”, 1st ed. Wellesley, MA: Wellesley-Cambridge Press, 1996.
- [78] R. C. Gonzalez and R. E. Woods, “*Digital Image Processing*”, 3rd ed. Addison Wesley, 2008.
- [79] A. Srivastava, A. B. Lee., E. P. Simoncelli, and S. C. Zhu, “On advances in statistical modeling of natural images”, *Journal of Mathematical Imaging and Vision*, vol. 18, pp. 17-33. 2003.

- [80] A. Lumini and D. Maio, "A wavelet-based image watermarking scheme", in *Proc. IEEE International Conference on Information Technology: Coding and Computing*, pp. 122-127, March 2000.
- [81] P. Meerwald and A. Uhl, "A survey of wavelet-domain watermarking algorithms", in *Proc. Electronic Imaging, Security and Watermarking of Multimedia Contents III*, vol. 4314, pp. 505-516, January 2001.
- [82] S. Agreste, G. Andaloro, D. Prestipino, and L. Puccio, "An image adaptive, wavelet-based watermarking of digital images", *Journal of Computational and Applied Mathematics*, vol. 210, no. 1, pp. 13-21, December 2007.
- [83] C. Christopoulos, A. Skodras, and T. Ebrahimi, "The JPEG-2000 still image coding system: An overview", *IEEE Transactions on Consumer Electronics*, vol. 46, pp.1103-1127, November 2000.
- [84] T. P. O'Rourke, R. Stevenson, "Human visual system based wavelet decomposition for image compression", *Journal of Visual Communication and Image Representation*, vol. 6, no. 2, pp. 109-121, June 1995.
- [85] H.A. Chipman, E.D. Kolaczyk, and R.E. McCulloch, "Adaptive Bayesian wavelet shrinkage", *Journal of the American Statistical Association*, vol. 92, no. 440, pp. 1413-1421, 1997.
- [86] E.P. Simoncelli, "Statistical models for images: Compression, restoration and synthesis," in *Proc. Asilomar Conference on Signals, Systems and Computers*, pp. 673-678, 1997.

- [87] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 710-732, July 1992.
- [88] J. Shapiro, "Embedded image coding using zero-trees of wavelet coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445-3462, December 1993.
- [89] M.T. Orchard and K. Ramchandran, "An investigation of wavelet-based image coding using an entropy-constrained quantization framework," in *Proc. Data Compression Conference*, pp. 341-350, 1994.
- [90] S. Mallat and W. Hwang, "Singularity detection and processing with wavelets", *IEEE Transactions on Information Theory*, vol. 38, pp. 617-643, March 1992.
- [91] J. Huang and D. Mumford, "Statistics of natural images and models", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 541-547, 1999.
- [92] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter and intra-scale statistical model for Bayesian wavelet based image denoising", *IEEE Transactions on Image Processing*, vol. 11, no.5, pp. 545-557, May 2002.
- [93] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using a generalized Gaussian and complexity priors", *IEEE Transactions on Information Theory*, vol. 45, no. 3, pp. 909-919, April 1999.
- [94] M. Malfait and D. Roose, "Wavelet-based image denoising using a Markov random field a priori model", *IEEE Transactions on Image Processing*, vol. 6, no. 4, pp. 549-565, April 1997.

- [95] J. Portilla, V. Strela, M. J. Wainwright, and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain", *IEEE Transactions on Image Processing*, vol. 12, no. 11, pp.1338-1351, November 2003.
- [96] M.J. Wainwright, E.P. Simoncelli, and A.S. Willsky, "Random cascades on wavelet trees and their use in analyzing and modeling natural images", *Applied and Computational Harmonic Analysis*, vol. 11, no. 1, pp. 89-123, July 2001.
- [97] Z. Azimifar, P. Fieguth, and E. Jernigan, "Hierarchical Markov models for wavelet domain statistics", in *Proc. IEEE Statistical Signal Processing Workshop*, pp. 258-261, 2003.
- [98] Z. Azimifar, P. Fieguth, and E. Jernigan, "Towards random field modeling of wavelet statistics", in *Proc. IEEE International Conference on Image Processing*, vol. 1, pp. 361-364, 2002.
- [99] J. Besag, "Spatial interaction and the statistical analysis of lattice systems", *Journal of the Royal Statistical Society. Series B*, pp.192-236, 1974.
- [100] G. Fan and X. Xia, "Image denoising using a local contextual hidden Markov model in the wavelet domain", *IEEE Signal Processing Letters*, vol. 8, no. 5, pp.125-128, May 2001.
- [101] G. Fan and X. Xia, "Wavelet-based texture analysis and synthesis using hidden Markov models", *IEEE Transactions on Circuits and Systems*, vol. 50, no. 1, pp.106-120, January 2003.

- [102] D. Donoho, “De-noising by soft thresholding”, *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613-627, May 1995.
- [103] D. Donoho, “Wedgelets: nearly minimax estimation of edges”, *Ann. Statist.*, vol. 27, no. 3, pp. 859-897, 1999.
- [104] P. Flandrin, “Wavelet analysis and synthesis of fractional Brownian motion”, *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910-916, March 1992.
- [105] J. Liu and P. Moulin, “Information-theoretic analysis of inter-scale and intra-scale dependencies between image wavelet coefficients”, *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp.1647–1658, November 2001.
- [106] D. Po, M.N. DO, “Directional multiscale modeling of images using the contourlet transform”, *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1610-1620, June 2006.
- [107] P. Dempster, N.M. Laird and D.B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm”, *Journal of the Royal Statistical Society, Series B*, vol. 39, pp. 1-38, 1977.
- [108] G. McLachlan and T. Krishnan, “*The EM algorithm and extensions*”, John Wiley & Sons, 2007.
- [109] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, “Multiplicative watermark decoder in contourlet domain using the normal inverse Gaussian distribution,” *IEEE Transactions on Multimedia*, vol. 18, no. 2, pp. 196-207, Feb. 2016.

- [110] W.W. Daniel, “Kolmogorov-Smirnov one-sample test”, *Applied Nonparametric Statistics*, pp. 319-330, 1990.
- [111] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, “A study on image denoising in contourlet domain using the alpha-stable family of distributions,” *Signal Processing*, 2016.
- [112] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, “Contourlet domain image modeling by using the alpha-stable family of distributions,” in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1288-1291, June 2014.
- [113] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, “Contourlet domain image denoising using the alpha-stable distribution,” in *Proc. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 141-144, Aug 2014.
- [114] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, “Despeckling of the SAR images in the contourlet domain using alpha-Stable distribution,” in *Proc. IEEE International Symposium on Circuits & Systems (ISCAS)*, pp. 121-124, May 2015.
- [115] A. C. Bovik, “*Handbook of Image and Video Processing*”, Ed. NY: Academic Press, 2000.
- [116] G.C. Langelaar, I. Setyawaan, and R.L. Lagendijk, “Watermarking digital image and video data: A state-of-the-art overview”, *IEEE Signal Processing Magazine*, vol. 17, no. 5, pp. 20-46, September 2000.
- [117] M. Amini and H. Sadreazami, “Binary image watermarking in ridgelet domain,” *10th International Conference on Signal Processing (ICSP)*, pp. 1813-1816, 2010.

- [118] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, "Multiplicative watermark detector for color images in sparse domain," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 62, no. 12, pp. 1159-1163, Dec 2015.
- [119] H. Sadreazami, M.O. Ahmad and M.N.S. Swamy, "Optimum multiplicative watermark detector in contourlet domain using the normal inverse Gaussian distribution," *IEEE International Symposium on Circuits & Systems (ISCAS)*, pp. 1050-1053, 2015.
- [120] P. Moulin and R. Koetter, "Data-hiding codes", *Proceedings of the IEEE*, vol. 93, no. 12, pp. 2083-2126, December 2005.
- [121] J.R. Hernandez and F.P. Gonzalez, "Statistical analysis of watermarking schemes for copyright protection of images", *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1142-1166, July 1999.
- [122] G. Guo, M. Mandal and Y. Jing, "A Robust detector of known signal in non-Gaussian noise using threshold systems", *Signal Processing*, vol. 92, no. 11, pp. 2676-2688, November 2012.
- [123] S. M. Kay, "*Fundamental of Statistical Signal Processing, volume II, Detection theory*", Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.
- [124] K. Mairgiotis, P. Nikolas and Y. Yang, "New additive watermark detectors based on a hierarchical spatially adaptive image model", *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 29-37, March 2008.

- [125] M. Amini, K. Yaghmaie and H. Sadreazami, "A new scheme for dual watermarking using DWT-PCA technique," *International Conference on Imaging Theory and Applications (IMAGAPP)*, pp. 43-46, 2010.
- [126] H. Sadreazami and M. Amini, "Dual wavelet watermarking using principal component analysis", *10th International Conference on Signal Processing (ICSP)*, pp. 1821-1824, 2010.
- [127] [Online]. Available: <http://decsai.ugr.es/cvg/dbimágenes/index.php>.
- [128] A. Valizadeh, Z.J. Wang "Efficient blind decoders for additive spread spectrum embedding based data hiding", *EURASIP Journal on Advances in Signal Processing*, no. 1, pp. 1-21, April 2012.