



UNIVERSIDAD NACIONAL DE COLOMBIA

# Image Quality Assessment based on Full-referenced Perceptual Measures

**Julio César García Álvarez**

Universidad Nacional de Colombia  
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Eléctrica, Electrónica y de  
Computación  
Manizales, Colombia  
2011



# Image Quality Assessment based on Full-referenced Perceptual Measures

Julio César García Álvarez

Thesis submitted in partial fulfillment of the requirements for the degree of:  
**Ph. D. in Engineering**

Director:  
Prof. César Germán Castellanos Domínguez, Ph.D.

Research line of Automatics  
Research Group:  
Signal Processing and Recognition

Universidad Nacional de Colombia  
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Eléctrica, Electrónica y de  
Computación  
Manizales, Colombia  
2011



# Medida perceptual referenciada para la evaluación de la calidad en imágenes

Julio César García Álvarez

Tesis presentada como requisito parcial para optar al título de:  
**Doctor en Ingeniería - Línea Automática**

Director:  
Prof. Dr. César Germán Castellanos Domínguez

Línea de Investigación en Automática  
Grupo de Investigación en Procesamiento y Reconocimiento de Señales

Universidad Nacional de Colombia  
Facultad de Ingeniería y Arquitectura, Departamento de Ingeniería Eléctrica, Electrónica y de  
Computación  
Manizales, Colombia  
2011



To my mother Olga María, my wife Monica, my niece Stephanie and my sister Lina Marcela: thank you for praying (and suffering) for me.

# Acknowledgments

Author would like to thank to:

- Supporting research groups:
  - Signal Processing and Recognition Group, Universidad Nacional de Colombia: David Cárdenas, Paola Alexandra Castro, Hermes Alexander Fandiño.
  - Faculty of Mathematics, RWTH University: Prof. Hartmut Führ, Ph.D., Prof. Remco Duits, Ph.D.
  - Bioingenium Group, Universidad Nacional de Colombia: Prof. Eduardo Romero, Ph.D.
- Fellowships and projects:
  - DAAD–Alecol–MEN–Colciencias 2008–2009 Ph.D Fellowship;
  - Study Special Sponsorship Universidad Nacional de Colombia–Manizales 2006–2012;
  - Comparison of Texture Analysis Techniques to Assess Wear Labelling of Textile Floor Coverings (2011);
  - Remote monitoring service of cardiac activity for clinical screening in the telemedicine network of the Department of Caldas (2010);
  - High performance computing techniques in the automated interpretation of medical images and biosignals (2009);
- Academic (and motivational) support:
  - Prof. Eduardo Cano Plata, Ph.D., Doctorate Program Coordinator, Universidad Nacional de Colombia;
  - Prof. Camilo Younes Velosa, Ph.D., Dean of Engineering and Architecture Faculty, Universidad Nacional de Colombia;
  - Prof. César Arango Lemoine, Director of Electrical, Electronic, and Computer Science Engineering Department, Universidad Nacional de Colombia;
  - University professors, Universidad Nacional de Colombia: Mauricio Orozco, Armando Ustariz Farfán, Francisco Abel Roldán, Juan Bernardo Gómez, Gerard Olivar Tost, Simeon Casanova, Julio Fernando Suárez, Flavio Augusto Prieto (thank you for the database);
  - University workers: Maria Clemencia Rios Vargas, Sonia;
  - ColAachen RWTH student group: Liliana Tamayo, Diana Leyva, Cristina Müller, Catalina Mejía, Laura Florez, Jackson Rodríguez, Oscar Javier Duque;

- English and German language support: Ute Teske, Eve Celia MacDonald, Kristina Rehbein;
- Master’s education: Prof. Néstor Peña Traslaviña, Ph.D., Jorge Reynolds Pombo, Ph.D.
- Undergraduate education: Prof. Jesús Quintero Trujillo;
- High-school education: Prof. Jorge Eliécer García.

I also thank to my family, friends (national, latin-american and worldwide), coworkers, and students, for their patience and comprehension during the realization of this work. I will try to get more time for each one of you.



# Abstract

This work illustrates a proposed image quality assessment based on perceptual measures - an evaluation that can be used for different applications with limited computational loads. It aims to emulate the human vision for evaluation purposes by using perceptual measures. Examples of applications are medical image coding, machine fault identification, texture classification in the carpet industry, and facial recognition. Perceptual measures are calculated on the wavelet space, generating a scheme which can extract features such as frequency, position, and textures, all taken from one image at the same time. This scheme is called multi-resolution, where features of higher orders of complexity can be calculated, such as edges, transitions, and homogeneities. The proposed assessment is composed of the following elements: a filter, a set of measures, and a summation function. The filter is used to enhance the image features at perceptible frequencies to the human eye. This filter is known as Contrast Sensitivity Function. The measures are used to numerically describe how the features change through distortion. The summation function parameters are calculated by a regression strategy, whose adequate selection widens the range of measurement.

The proposed image quality assessment is called *Full-Wavelet Quality Index* (FW-QI), which is intended to be highly correlated with human vision models and also to reduce the number of processes involved in the calculation of perceptual measures, by using the same representation of the imaging process. However, due to this capability, the measuring accuracy must be addressed. This document will show that the accuracy and sensitivity of measurement are not affected, by using an appropriate selection of parameters.

Chapter 1 introduces the reader to the problem of image quality assessment, presenting a description of the research work for the development of the proposed methodology. Accordingly, a background summary on assessment methods and their parameters is presented. The work is shown and justified in further chapters, where the objectives and scope of this thesis are described.

Chapter 2 presents the theoretical aspects of Image Quality Assessment, in particular how quality is measured on an image that is encoded for transmission purposes. For instance, when a medical image is distorted when it passes through a lossy coding process to fit bandwidth constraints. This explanation aims to describe the elements required for a perceptual scoring. An overview on multi-resolution representations is shown, including the theoretical description of used representation.

Chapter 3 constitutes the core of the proposal and contribution. This chapter is dedicated to the description of the FW-QI assessment. The procedures for calculation of perceptual metrics and filters are based on the principles of functional analysis and human-eye inspired models. Although assessments are used for a wide range of distortions, the source of distortion considered here is that generated from coding - i.e. where the measured image is distorted after passing through a coding method. The measured distortion level is then turned into a performance assessment parameter for coders. At a specific transmission rate,

each coder outputs an image with different types and levels of distortion. Therefore, performance of such coders can be measured by quality assessment on output encoded image. As a result of the assessment, the image that gives the lowest distortion value implies that the related coder has the best performance. Thus, the proposal is motivated by the need for selecting an adequate representation such that more information can be obtained from the same representation of encoded image. Such an enrichment of the representation — or *generalization*— is particularly pertinent when there are representational limitations. A first attempt is described in Appendix A for Quad-tree Decomposition. Thereafter, wavelet representation is finally selected.

Chapter 4 describes the methodology for the adjustments made on FW-QI algorithm, which improves the measurement accuracy, and corrects the generalization problem. The adjustments are validated using the correlation and sensitivity with respect to human-based evaluations made on test images. It is experimentally demonstrated that the proposed assessment achieves a high correlation value with human-based models, despite the accuracy problem presented by generalization. This value is achieved mainly by the selection of an adequate regression strategy, where an experimental test is carried out to give the adjustments required for those parameters. These adjustments are made on the Contrast Sensitivity Function, on the Feature Measures and on the Linear Regression Strategy. Correlation is measured by coefficient of determination. The standardized *LIVE* database is used for adjustment, and for comparison purposes.

In Chapter 5, the FW-QI assessment is used for two purposes: first, to give a performance evaluation of a special type of encoders (see Appendix B), using medical images as database; second, to implement this assessment on a coding control scheme. The intention of this test is to show that FW-QI is suitable for use as a coder performance evaluator, such as parameter for a distortion control strategy within these coders. The test is arranged in the following manner: an image is encoded by several considering methods, resulting in a number of encoded images equal to the number of selected methods. Later, each encoded image is evaluated by FW-QI, and other assessments by three approaches: An assessment made on whole image, another made on an arbitrary selected region of interest, and the last made as a quality ratio between region of interest and background. Finally, the performance of several coding methods are evaluated based on the results of the assessment approaches. Finally, Chapter 6 concludes with future works on classification and pattern recognition applications.

**Keywords:** Quality Assessment, Wavelet Transforms, Distortion Measurement, Feature Extraction, Image Processing, Visual System

---

## Resumen

La presente tesis muestra un sistema de evaluación de la calidad en imágenes basado en medidas perceptuales. El uso de evaluaciones del tipo perceptual permiten que los sistemas, encargados del monitoreo de procesos industriales a través de imágenes, puedan dar evaluaciones que se aproximen a la de un experto. Esto cumple con los siguientes objetivos: Primero, proveer una alta correlación entre los valores medidos usando esta evaluación y los valores observados por seres humanos sobre la misma imagen. Por último, reducir el costo computacional usado para la evaluación, gracias a que las etapas de procesamiento al interior de su algoritmo usan la misma representación de procesamiento de la imagen. Este último aspecto conlleva a proponer la inclusión de este sistema de evaluación como parámetro para una estrategia de control de distorsión sobre algoritmos de codificación de imágenes.

El método de evaluación de la calidad en imágenes propuesto se llama *Índice de Calidad Completamente en Onditas* (IC-CO). Este método se basa en el uso de una combinación de medidas, las cuales se calculan en el espacio multi-resolución de onditas. El uso de este espacio busca obtener una alta correlación con el modelo de visión humano y reducir el número de procesos involucrados con el cálculo de medidas perceptuales. Este método utiliza la misma representación del proceso de imágenes, para evitar la extracción de cada característica usando una operación diferente, convirtiéndose esto en una carga adicional para la evaluación. Como entonces no se utilizan las funciones óptimas requeridas para extraer adecuadamente las características perceptuales, existe un problema de inexactitud en la medida. Sin embargo, se demostrará experimentalmente que la exactitud de la evaluación no se ve afectada, siempre y cuando se use una selección adecuada de parámetros. El Capítulo 1 introduce el problema relacionado con la evaluación de la calidad en imágenes. Aquí se desarrolla la metodología de evaluación de calidad en imágenes. Para ello, un resumen de los antecedentes se presenta en relación con los métodos de evaluación y sus parámetros. Además, se describen también los objetivos y el alcance de esta tesis.

El Capítulo 2 brinda los aspectos teóricos acerca de la evaluación de la calidad en imágenes, para describir cómo la calidad es medida sobre una imagen que está codificada para poder ser transmitida. En concreto, una imagen médica puede distorsionarse al pasar a través de un proceso de codificación con pérdidas para encajar con la limitación de ancho de banda de transmisión. Esta explicación tiene la intención de describir los elementos necesarios para dar un puntaje equivalente a la percepción humana. Además, este capítulo incluye una descripción teórica de la representación a ser usada, brindando una visión general de las representaciones multi-resolución.

El Capítulo 3 constituye el núcleo de nuestra propuesta y contribución. Este capítulo está dedicado a la descripción de la IC-CO. Los procedimientos para el cálculo de las medidas y filtros relacionados con esta evaluación se basan en los principios del análisis funcional y modelos inspirados en el ojo humano. Aunque las evaluaciones se utilizan para una amplia gama de distorsiones, la única distorsión considerada es la generada por el proceso de codifi-

cación de la imagen. Es decir, la imagen a evaluar está distorsionada debido a un proceso de codificación. Así, el nivel de distorsión medido se convierte en un parámetro de evaluación del rendimiento para estos procesos. Para un valor específico de tasa de transmisión, cada codificador entrega una imagen con diferentes tipos y niveles de distorsión. Por tanto, el desempeño de estos codificadores puede ser medido mediante la evaluación de la calidad sobre la imagen correspondiente. Como resultado de esta evaluación, la imagen que entregue el menor valor de distorsión implica que su codificador correspondiente tiene el mejor desempeño en comparación con los demás. Así, nuestra propuesta está motivada por la necesidad de seleccionar una representación, de tal manera que se obtenga un enriquecimiento o mejor uso de la información que se obtiene de la imagen codificada. Este enriquecimiento de la representación —a lo que llamaremos *generalización*— es particularmente pertinente cuando hay limitaciones en la representación. Sin embargo, esta generalización puede presentar una pérdida en la precisión de la medida. Una primera aproximación se describe en el Anexo A para Descomposición en Árboles de Cuadratura. Sin embargo, la representación en onditas es seleccionada finalmente.

El Capítulo 4 determina la metodología para realizar los ajustes sobre el algoritmo de IC–CO, de tal manera de corregir la pérdida de precisión debida al uso de la generalización. Estos ajustes son validados usando la correlación y la sensibilidad con respecto a evaluaciones humanas realizadas sobre las mismas imágenes de prueba. Con esto se demuestra experimentalmente que la evaluación propuesta logra resultados comparables con otros en correlación con modelos de evaluación humana, a pesar del problema producido por la generalización. Para tal caso, se selecciona una estrategia de regresión adecuada para obtener un incremento en la correlación con el modelo humano. Por lo tanto, este capítulo describe la prueba experimental para la realización de los ajustes sobre los parámetros. Estos parámetros son los siguientes: Función de Sensitividad de Contraste, Medidas Perceptuales y Estrategia de Regresión Lineal. Para los dos últimos, la correlación se mide por el coeficiente de determinación. Se utiliza una base de datos normalizada para el ajuste, y para propósitos de comparación.

En el capítulo 5, el *Índice de Calidad basado Completamente en Onditas* (IC–CO) se utiliza para dos propósitos: Primero, para dar una prueba de desempeño de una clase especial de codificadores, denominados *Codificadores de la Región de Interés*, usando imágenes médicas como base de datos; Segundo, para sugerir que la evaluación IC–CO puede implementarse como parámetro de control de calidad sobre los codificadores. Así, se propone que las evaluaciones basadas en medidas perceptuales son adecuadas como evaluadores del desempeño en codificadores, así como como parámetros de estrategia de control para los mismos. Finalmente, el Capítulo 6 concluye con los trabajos futuros en clasificación y reconocimiento de patrones.

**Palabras clave:** Evaluación de la Calidad, Transformada Onditas, Medición de Distorsión, Extracción de Características, Procesamiento de Imágenes, Sistema Visual.

---

## Published Works

1. Julio-César García-Alvarez, Germán Castellanos-Dominguez, and Hartmut Führ. Wavelet-based entropy measure for rate-distortion curve optimization. *International Journal of Electronics and Telecommunications (Former Kwartalnik Elektroniki I Telekomunikacji)*, 56(1):25–32, 2010.
2. Benhur Ortiz-Jaramillo, Julio-César García-Alvarez, and Germán Castellanos-Dominguez. Translating distortion measures to a multi-resolution space for image analysis (english abstract). *Ingenieria y Competitividad*, 12(1):43–55, 2010.
3. Sergio-Alejandro Orjuela-Vargas, Aleksandra Pizurica, Julio-César García-Álvarez, Benhur Ortiz-Jaramillo, Filip Rooms, and Simon De Meulemeester. Surface reconstruction of wear in carpets by using a wavelet edge detector. *Advanced Concepts For Intelligent Vision Systems*. 12th International Conference on Advanced Concepts for Intelligent Vision Systems, 2010.
4. Julio-César García-Álvarez, Hartmut Führ, Benhur Ortiz-Jaramillo, and Germán Castellanos-Domínguez. On improving calculation of wavelet contrast sensitivity function. 6th International Seminar on Medical Image Processing and Analysis SIPAIM, 2010.
5. Benhur Ortiz-Jaramillo, Julio-César García-Álvarez, Germán Castellanos-Domínguez, and Hernán Benitez Restrepo. Region of interest extraction based on multi-resolution analysis for infrared nondestructive testing. 10th International Conference on Quantitative InfraRed Thermography, 2010.
6. Benhur Ortiz-Jaramillo, Julio-César García-Álvarez, Diego Peluffo-Ordoñez, Germán Castellanos-Domínguez, and Jose-Luis Rodriguez-Sotelo. Region of interest extraction using redundant wavelet transform and unsupervised techniques on thermal imaging. 10th International Conference on Quantitative InfraRed Thermography, 2010.
7. Julio-César García-Álvarez and Germán Castellanos-Domínguez. Information quantity measure mapped in wavelet space. *International Conference on BroadBand Communications, Information Technology and Biomedical Applications BroadBandComm2009*, 2009.
8. Julio-César García-Álvarez and Germán Castellanos-Domínguez. Region of interest extraction method using wavelets. *International Conference on Communication Theory, Reliability, and Quality of Service CTRQ 2009*, 2009.

9. Julio-César García-Álvarez, Hartmut Führ, Benhur Ortiz-Jaramillo, and Germán Castellanos-Domínguez. Image distortion measure mappings in wavelet representation space. International Conference on BroadBand Communications, Information Technology and Biomedical Applications BroadBandComm2009, 2009.
10. Julio-César García-Álvarez and G. Castellanos-Dominguez. Estimación de tasas de compresión de imágenes mediante métricas de distorsión completamente referenciadas. *Ingeniería y Competitividad*, 10(1):73–83, 2008.
11. Julio-Cesar Garcia-Alvarez, Cesar-German Castellanos-Dominguez, and Benhur Ortiz-Jaramillo. Image information access using wedgelet filters. pages 1–4. IEEE International Symposium on Applied Sciences in Biomedical and Communication Technologies ISABEL, October 2008.
12. Julio-César Garcia-Alvarez and German Castellanos-Dominguez. Estimation of image compression rates by means of full-referenced distortion metrics (english abstract). *Ingeniería y Competitividad*, 10(1):73–83, 2008.
13. Julio-Cesar García-Álvarez and Jose Luis Rodríguez-Sotelo. Estimación de tasa de compresión para transmisión de fonocardiogramas. Congreso Anual de la Sociedad Española de Ingeniería Biomédica CASEIB, November 2007.
14. Julio-César García-Álvarez. Telemedicine integrated-services perspectives. *Revista del Congreso Internacional de Ingeniería Electrónica*, 1(1):1–4, 2007.
15. Julio-César García-Álvarez, Juan Pablo Tello, and Germán Castellanos-Domínguez. Análisis de calidad de imágenes médicas utilizando el modelo de sistema visual humano. XXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica CASEIB, 2006.

# Content

. <b>Agradecimientos</b>	<b>viii</b>
. <b>Abstract</b>	<b>xi</b>
. <b>Published Works</b>	<b>xv</b>
. <b>Acronym and Symbol List</b>	<b>xx</b>
<b>1. Introduction</b>	<b>1</b>
1.1. Preliminary . . . . .	1
1.2. Image Quality Assessment . . . . .	2
1.2.1. Perceptual Assessment . . . . .	3
1.2.2. Assessments as Coding Control Parameter . . . . .	4
1.3. Challenging Aspects Involved in Assessment of Image Coding Performance . . . . .	6
1.4. Proposed Fully-Wavelet Quality Index . . . . .	10
1.5. Objective . . . . .	11
<b>2. Image Quality Assessment Definitions</b>	<b>13</b>
2.1. Preliminary . . . . .	13
2.2. Image Constrains . . . . .	13
2.2.1. Wavelet Image Representation . . . . .	16
2.2.2. Image Feature Extraction . . . . .	19
2.3. Image Quality Assessment . . . . .	22
2.4. Perceptual Image Quality Assessment . . . . .	23
2.4.1. Contrast Sensitivity Function . . . . .	25
2.4.2. Feature Measures . . . . .	26
2.4.3. Regression Strategies . . . . .	27
2.5. Concluding Remarks . . . . .	29
<b>3. Full-Wavelet Quality Index Assessment for Image Coding</b>	<b>30</b>
3.1. Preliminary . . . . .	30
3.2. Proposed Image Quality Assessment . . . . .	31
3.2.1. Projection of Measures into Multi-resolution Domains . . . . .	32
3.2.2. Proposed Adaptive-Mask Wavelet-based Contrast Sensitivity Function . . . . .	32

3.2.3.	Proposed wavelet-based Blockiness Measure . . . . .	35
3.2.4.	Proposed Wavelet-based Edge Error Measure . . . . .	37
3.2.5.	Proposed Wavelet-based Visual Impairment Measure . . . . .	38
3.3.	Concluding Remarks . . . . .	39
<b>4.</b>	<b>Full-Wavelet Quality Index Adjustments</b>	<b>40</b>
4.1.	Preliminary . . . . .	40
4.2.	Image Database and Human Visual System Parameters . . . . .	41
4.3.	Correlation and Sensitivity Measurements . . . . .	41
4.3.1.	Correlation between Feature Measures and Differential Scores . . . . .	42
4.3.2.	Correlation between Assessments and Human Vision Model . . . . .	43
4.3.3.	Assessment Sensitivity . . . . .	44
4.3.4.	Evaluation of Contrast Sensitivity Functions. . . . .	45
4.4.	Results . . . . .	45
4.4.1.	Correlation between Feature Measures and DMOS . . . . .	45
4.4.2.	Correlation between Assessment and Human Vision Model . . . . .	46
4.4.3.	Assessment Sensitivity . . . . .	47
4.4.4.	Evaluation on Contrast Sensitivity Function . . . . .	49
4.5.	Discussion . . . . .	51
4.6.	Concluding Remarks on Experiment . . . . .	52
<b>5.</b>	<b>Evaluation of Coding Performance by Image Quality Assessment</b>	<b>53</b>
5.1.	Preliminary . . . . .	53
5.2.	Overview of Region Of Interest Coders . . . . .	54
5.3.	Implementation of Full-Wavelet Quality Index in PO-Shift Coder . . . . .	55
5.4.	Experimental Setup . . . . .	56
5.4.1.	Database and Preprocessing . . . . .	56
5.4.2.	Considered test evaluations . . . . .	60
5.4.3.	Decomposition Level-Assessment Value Plots . . . . .	60
5.4.4.	Numerical evaluation between ROI-coders . . . . .	61
5.5.	Results . . . . .	62
5.5.1.	Decomposition Level-Assessment Value Plots . . . . .	62
5.5.2.	Numerical ROI-Coding evaluation . . . . .	66
5.6.	Discussion . . . . .	67
5.7.	Concluding Remarks . . . . .	69
<b>6.</b>	<b>Conclusions and Further Work</b>	<b>70</b>
6.1.	Conclusion . . . . .	70
6.2.	Future Work . . . . .	71
6.2.1.	Evaluation on Face Recognition Process . . . . .	72
6.2.2.	Evaluation on Texture Wearing Scores . . . . .	73

---

<b>A. Findings about Feature Extraction using Quad-tree Decomposition</b>	<b>76</b>
A.1. Quad-tree Decomposition . . . . .	76
A.2. Blockiness calculation by Quad-tree Decomposition . . . . .	77
A.3. Considerations on Visual Impairment metric using Quad-tree Decomposition	78
A.4. Results on Blockiness and Visual Impairment calculated by Quad-tree De- composition . . . . .	79
A.5. Variation Coefficient-based Entropy . . . . .	80
A.6. Future works on Wedges . . . . .	81
<b>B. Region-of-Interest Coding</b>	<b>82</b>
B.1. Region-of-Interest Coding Methods . . . . .	82
<b>. Bibliography</b>	<b>85</b>

# Acronym and Symbol List

In this section, general symbols (Greek and Latin fonts), sub-indexes, super-indexes and acronyms are included. Each one of these lists is in alphabetical order, according to the first symbol character.

## Latin Font Symbols

Symbol	Term	Definition
$C$	Wavelet Representation	§2.2.1
$c_a$	Wavelet Approximation Coefficient	§2.2.1
$c_d$	Wavelet Decomposition Coefficient	§2.2.1
$\partial$	Partial Derivative	
$M_i$	Image $i$ -th dimension rank	Def. 2.2.1
$X$	Image Matrix	§2.2
$x$	Image element value (pixel)	Def. 2.2.1
$Y$	Encoded Image	§2.2.1
$B$	Block Area	§3.2.3
$j$	Wavelet Decomposition Level	§2.2.1
$J$	(Total) Number of Wavelet Decomposition Levels	§2.2.1
$k$	Wavelet Spatial Index	§2.2.1
$D$	Measuring Matrix	§2.3
$d$	Measure	§2.3
$m$	Image coordinate vector	Def. 2.2.1
$n$	Number of elements	
$\mathbb{R}$	Real Space Set	
$R^2$	Coefficient of Determination	
$\mathcal{T}$	Transformation	Def. 2.2.5

Symbol (cont.)	Term (cont.)	Definition (cont.)
$\mathbb{Z}$	Integer Space Set	

## Greek Font symbols

Symbol	Term	Definition
$\epsilon_b$	Quantization Exponent for sub-band $b$	Eq. (5-3)
$\eta$	Trial number for regression calculation	§4.3.2
$\lambda$	Linear Combination Weighting Coefficient	§4.3.2
$\theta$	Wavelet Orientation Index	§2.2.1
$\mu_b$	Quantization Mantissa in sub-band $b$	Eq. (5-3)
$\mu_B$	Expected value of elements inside Block	Eq. (3-11a)
$\mu_{DMOS}$	Expected value of DMOS evaluations	Eq. (4-3)
$\sigma$	Standard Deviation	
$\psi$	Wavelet Function	§2.2.1

## Subindexes

Subindex	Term
$b$	Wavelet Sub-band
$=$	Parallel or defined in same direction alongside
$\#$	Element belonging to an detected edge

## Superindexes

Superindex	Term
2	Quadratic expression

## Acronyms

Acronym	Term
ANNR	Artificial Neural Network Regression
AM-WCSF	Adaptive Mask Wavelet-based Contrast Sensitivity Function
BG	BackGround
CSF	Contrast Sensitivity Function
DCT	Discrete Cosine Transform
DMOS	Differential Mean Observed Score
DWT	Discrete Wavelet Transform
EBCOT	Embedded Block Coding with Optimized Truncation coding
FW-QI	Full-Wavelet Quality Index assessment
HVS	Human Visual System
HT	Hartley Transform
ITU	international Telecommunication Union
JPEG2000	Joint Photographic Experts Group 2000 Standard
LR	Linear Regression
MSB	Most Significant Bit
MSE	Mean Squared Error
PCAR	Principal Component Analysis Regression
PO	Perceptually Optimized
PRCD	Post-Compression Rate Distortion coding
PSNR	Peak Signal to Noise Ratio
QI	Quality Index assessment
QILV	Quality Image based on Local Variance asessment
QoS	Quality of Service
RF	Reflection Factor assessment
ROI	Region Of Interest
SCM	Selective Coefficient Mask
SPIHT	Set Partitioning in Hierarchical Trees
SSIM	Structural Similarity Index Metric assessment

---

<b>Acronym</b>	<b>Term</b>
SVD	Singular Value Decomposition
VIP	Visual Progressive Weighting
WCAP	Wavelet-based Compression scheme with adaptive prediction
WCSF	Wavelet-based Contrast Sensitivity Function

# 1. Introduction

## 1.1. Preliminary

Large images are coded for their adequate transmission under channels of reduced bandwidth. The growing demand for transmission of this kind of images has led to improving the performance of the coding process. However, due to the bandwidth restriction, some elements of the image are lost through the coding process, which leads to a quality diminishing of received counterpart. Hence, the quality of a transmitted image is mainly affected by the *Distortion*, which is due to phenomena like noise, fading, and delay, related to variations in the transmission channel. The *Noise* is the undesired random disturbance of the image, introduced before or after the coder. The *Fading* is the strength fluctuation of transmitted image elements, producing interference among successive elements. The *Delay* is predominant when the channel capacity is lower than the mean amount of transmitted data. Therefore, a reduction of this amount of data could help to reduce the generated distortion. This reduction can be achieved inside coders thanks to the high redundancy of images – Highly redundant data can be represented by a small amount of elements– which leads to rank image data properly for its subsequent reconstruction, where less redundant elements are estimated from those more redundant. For the sake of illustration, the following coders use respective integer or rounded representations of image for transmission: the *Hartley* coder, proposed in [58], uses linear quantization to reduce high–frequency components of an image; the JPEG coder [16] processes fixed–sized regular blocks on image to gather a considerable number of redundant elements; the Principal Component Analysis (PCA) coder, proposed in [45], extracts statistical patterns in images, putting most of these patterns to be represented by few parameters, thus reducing the number of representation elements of image. Currently, wavelet–based coders are used due to their ability to adequately represent spatial details, edges, and transitions, in one scheme at the same time, achieving better coding performance results. Two main methods inside this category are the *Set Partitioning In Hierarchical Trees* (SPIHT), which uses a hierarchical sorting of wavelet scheme [67] (some variations and improvements of this method are shown in [61, 84, 65]), and the *Embedded Block Coding with Optimized Truncation* (EBCOT), which uses a variable step quantization and is used for JPEG2000 encoder [73].

The coding methods can be also classified by the image quality handling during transmission: for *Lossless* coding, a full reconstruction of the transmitted image is achieved, because the image can be represented by a statistical model, making the reconstructed image equal to

its source counterpart. Though in this case distortion is avoided, a very low reduction rate is achieved, making this kind of coders unsuitable for necessary bandwidth constraints. In other words, a lossless scheme does not fix the delay problem. For instance, to accomplish a traditional bandwidth constrain of  $1 \times 10^6$  bits per second, an image of  $1024 \times 1024 \times 1 \times 12$  size (equivalent to  $12 \times 10^6$  bits) would require, at most, 12 seconds to be transmitted. Therefore, supposing an accepted delay of 0.5 seconds of time transmission, the image needs to be reduced in size up to 24 times smaller than its original, which is beyond of the most successful reduction rate for any lossless coding method. On the other hand, *Lossy* coding methods can avoid the delay problem, because they can achieve a higher transmission rate, through reducing data by losing irrelevant components of the image, which are supposedly imperceptible for the human eye. It means that the reconstructed image is an *estimated* version of the source. It gives lossy coded representations suitable to current bandwidth challenges, like Wireless Fidelity and Multimedia Internet technologies. Nonetheless, distortion is still observed for many transmission cases, due to quality loss on the image.

## 1.2. Image Quality Assessment

An adequate way to evaluate the performance of *lossy* coders is by measuring the loss of quality on reconstructed images. The value of these measurements is proportional to the performance level. Therefore, the coder achieving optimal measures is selected as the best performing one, being able to achieve the highest quality level on image for a desired bandwidth restriction. These measures can also bring information about the desired quality to be achieved for different bandwidth restrictions. Thus, a distortion control mechanism would be implemented in the coding scheme, like the distortion control for coding schemes described in [43], or the use of perceptual measurements for coding improvement in [21], both for the JPEG2000 scheme. This evaluation is termed as *Image Quality Assessment*. Various reviews present the different assessments used in several applications, including image coding evaluation: in [1], the different image quality assessments are analyzed for image processing applications, and the experimental approaches are surveyed to evaluate the image quality. Another literature survey is made in [75], where a general description of the various considerations regarding the development and implementation of image/video quality assessment systems is given. That study was complemented with the evaluation of measures made in [9] and [21], which summarizes and evaluates some of the existing methods of measuring and quantifying the quality of a digital image. Its aim is to determine if there is any performance difference between assessments. Results of a extensive subjective quality assessment study are presented in [68]. This is one of the largest subjective image quality assessment studies found in literature in terms of number of images, distortion types, and number of human judgments per image. This study is an suitable assessment comparison tool. Other reviews exist which focus on specific applications. For instance, a selection of measures used for image quality assessment on Steganalysis is described in [8] and an

evaluation of JPEG2000 encoder is investigated in [82], using the performance of human and model observers for clinically relevant visual tasks. From those reviews, evaluations can be grouped into the following categories: Fully-reference, Reduced-reference and No-reference. *Fully-referenced* measurements are commonly used because distortion is considered as a distance value between source and reconstructed image [38]. For this category, two methods are considered: *Objective* and *Perceptual* measurements. *Objective* measurement is based on distance metrics, commonly on image or spectral spaces, to obtain the quality values. Examples of used measures are the *Mean Squared Error* (MSE), *Peak Signal to Noise Ratio* (PSNR) and *Czekanowski* measures. Objective measurement is widely used, because its calculation is solely based on image domain; no additional operations are required. However, these measures are highly monotonic with respect to distance functions, therefore, they are less adaptable or correlated to human-based evaluation processes. Moreover, preservation of good perceptual quality of the reconstructed images is not guaranteed for coding methods using objective assessments as control parameter, which often lead to visually annoying artifacts [81]. Therefore, perceptual measurements are preferred over objective ones.

### 1.2.1. Perceptual Assessment

*Perceptual* measures use a set of functions to obtain features required for calculating quality value [78, 42]. The purpose of this kind of measurement relies on the fact that some kind of distortions have been perceptually characterized. Perceptual measures can classify and quantify the nature of distortion, according to human-based vision models. For instance, supposed rate restriction requires the coding stage to scale down image elements to be retrieved. If the coder uses *Discrete Cosine Transform* (DCT), elements to be retrieved are the local frequency representation of image. It means that a considerable number of frequency components can be scaled towards zero, resulting in shape distortion of the displayed image, namely *block artifact* distortion. In the same manner, if the coder uses *Discrete Wavelet Transform* (DWT), the down-scaling of retrieved elements leads to a *blurring* distortion. However, for an objective-based evaluation, it is possible that the same quality value is presented for two different kinds of distortion. This problem is avoided when a perceptual-based evaluation is used, because their features are discriminative respect to the distortion behavior.

Nonetheless, a perceptual evaluation can require more functions, depending on its complexity. For instance, a luminance function must be added to enhance features required in measuring. Thus, an additional number of calculations is required, building up to a concerning load respect to coding times. It is critical not only for coding of high size images, but also for distortion control inside video framing applications, because this load increases with the image size or frame rate. As described above, there would be some preference for using an objective method [32, 83], because it requires fewer operations than the perceptual one. However, a perceptual method is valued for parameter optimization for the follow-

ing reasons, among others: firstly, the sensitivity parametrization and adjustment of these measures is based on biological and psychological human–eye process approximations, like *Human Visual System* [6]; secondly, perceptual measurements are, by definition, intended to behave analogue to human–made perceptual evaluations on data quality, like *Observation Scores*. This scoring is stated as contrast comparison between reference and test images. This property makes a distance measurement related to human–made scoring.

First of all, the main perceptual measure considered is, undoubtedly, the *Human–made Scoring*. This evaluation is used for *Quality of Service* (QoS) measures. These measures are related to a wide variety of disturbances on the transmission channel, like inter–symbol interference and packet dropping, contributing to the degradation of perceptual quality. Thus, it is important to note that QoS human–made assessment provides the main reference in terms of the distortion source. In this case, the evaluation methodology is made perceptually using human–made *Scoring*, which is subjectively adjusted by different psycho–perceptual analysis, vision–stimulus approximative models and experiments made on an image database with the participation of a wide population. It is established as standard in ITU recommendations [36]. An experiment of our own was made in [34], for a Video–digital library access system. Specifically for lossy cases, the level of distortion and delay are determined mainly by user evaluation. However, this evaluation cannot be implemented for a coding scheme. Therefore, a perceptual evaluation that brings some analogy with human vision models is useful. In literature, an evaluation of measures based on different Human Visual System models has been provided in [51] where the quality of coded images is assessed. Later, a relation between maps in the human visual cortex and low-dimensional maps of features used for perceptual evaluations is described in [52]. A newer paradigm in relation between human perception and quality evaluations is proposed in [77], following the assumption that human visual perception is highly adapted for extracting structural information of images. From this assumption, an alternative complementary framework is introduced for quality assessment based on the degradation of structural information.

### 1.2.2. Assessments as Coding Control Parameter

Image Quality Assessments are also used for distortion control strategies on coding schemes. Inside JPEG2000 coder, such strategies for coding methods consider block ranking of coefficients in order to increase redundancy [73]. For instance, *Post Compression Rate Distortion* (PRDC) convex–hull–based technique is used for establishing the priority of details by controlling rate allocation of bit–planes [11]. The required coding passes for this technique, such as *Significance Propagation*, *Magnitude Refinement* and *Cleanup*, use processed data blocks instead of the representation coefficients of image, impeding the directly evaluation of this technique using perceptual measures. Therefore, the implementation of an assessment on the coding scheme, before those mentioned passes, avoids the waste of calculations produced. *Visual Progressive weighting* (VIP) and visual masking have been provided as options in some

JPEG2000 implementations by setting the weights of a weighted distortion function based on the human visual system and some local visual masking effects. An encoding scheme based on a perceptual-based assessment has been developed in [43], to minimize the bit rate for a desired target quality. This scheme achieves the minimum bit-rate for a given desired distortion directly as part of the *Tier-1* coding stage. Moreover, the perceptual assessment exploits the luminance and contrast masking properties of the Human Visual System. It is particularly useful for a specific kind of coder, which is called the *Region of Interest* (ROI-) coder. This kind of coder is able to avoid the loss of quality on ROI areas of image [50], and it is particularly useful in medical images, primarily because the quality of selected areas on image is mandatory for avoiding some diagnosis misinterpretation [38][44][80]. Some applications using ROI-coder are breast cancer diagnosis assessment via mammography [85], and *store-and-forward* Tele-medicine applications. About the latter, techniques for an adequate ROI coding have been used for health services in emerging countries, as reported in [4] and [26].

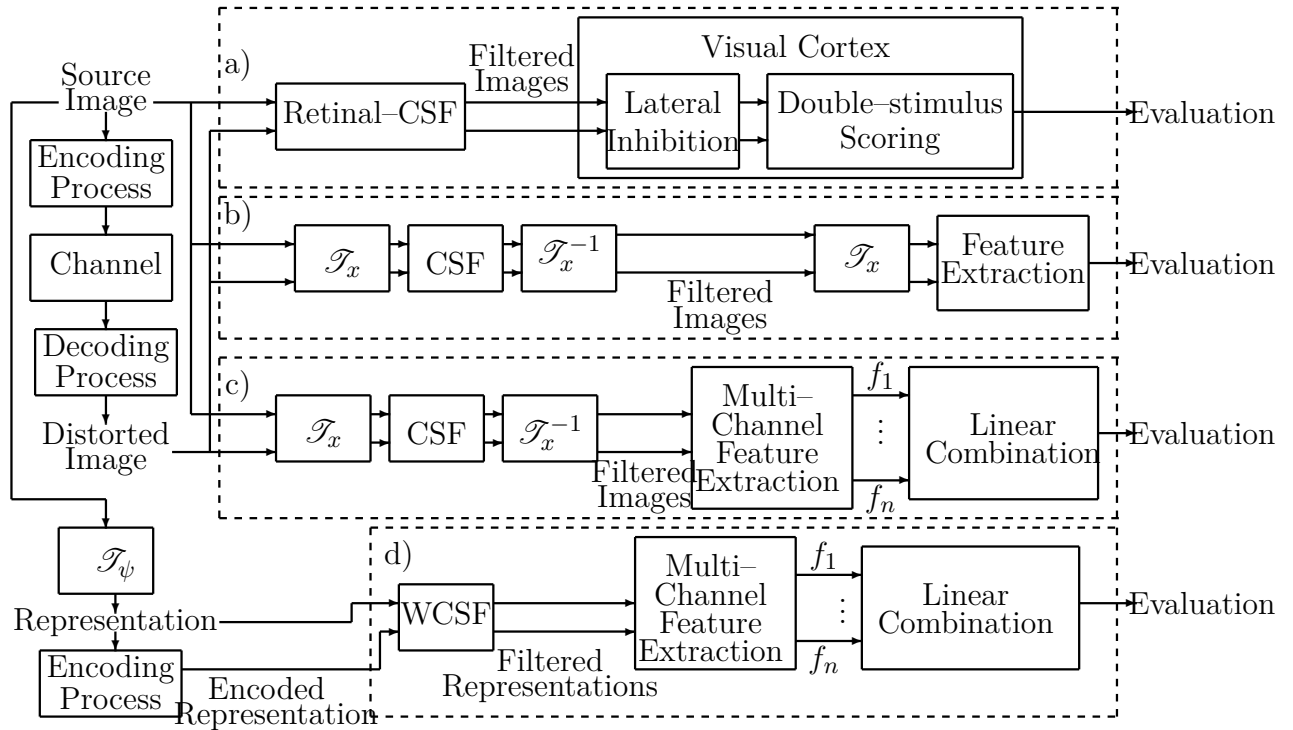
Two different strategies are used for quality control on ROI-coders [18]: Firstly, in *Max-Shift*-based strategy, ROI elements are binary-up-shifted to the higher bit-planes so that they attain a value above the maximal value of any coefficient of background (BG). Thus, only the scaling value is required to be transmitted with the encoded image, being this value the bit-plane number where the Most Significant Bit (MSB) of BG is located. It results in coefficients located at bit-planes higher than scaling value clearly identifiable as ROI at decoding stage. Lastly, for *General-scaling*-based strategy, ROI elements are also up-shifted but by a scaling value corresponding to an arbitrary-selected bit-plane. The main advantage of the *Max-Shift* strategy is that it requires neither the information about the ROI shape, nor the ROI position set, i.e., the ROI mask is implicit in coefficient set. Despite these advantages, it faces two main drawbacks, as stated in [71]: Firstly, there is no control over the relative importance of ROI and BG, and secondly, there is no possibility to encode a given ROI with different quality levels according to the user priorities over the image. Some coders, as the *Generalized Bitplane-by-Bitplane Shift* [76] and *Selective Coefficient Mask Shift* (SCM-Shift) [71], have been developed to overcome these drawbacks. These coders achieve both control over the relative importance of ROI and BG, and encoding of different regions of interest for a given set of arbitrary selected qualities. A better approach is the *Perceptually Optimized Bitplane Realignment* (PO-Shift) coder, which differs from *Max-Shift* in two aspects [81]: Firstly, it ranks the bit-planes according to a perceptual-valued order. Secondly, it uses two scaling values depending on ROI priority. These provide a numerical or functional control parameter to reduce the quantity of distortion in encoding framework. These coders for the bit-scaling of elements are shown in Figure B-1, where the *Max-Shift* coder prioritizes the encoding of ROI-related image information over the rest, whereas *PO-Shift* and *SCM-Shift* approaches interlace the bit-planes for coding ROI and BG information. Other coders have also been proposed for providing a quality control on ROI. In *General-scaling*-based coding methods, the ROI is also up-shifted, but by a scaling

value corresponding to an arbitrary-selected bit-plane. In this case, the scaling value is not enough for a clear identification of the ROI at decoding stage. Therefore, a mapping of locations of ROI is required to be transmitted with encoded image. This also applies for ROI-coders based on Set Partitioning in Hierarchical Trees (SPIHT), where the priority on elements of ROI must be transmitted by an additional bit-mapping [7]. As an improvement of this kind of methods, different estimators of the ROI mask for SPIHT coding algorithm have been proposed in [61] and [85]. Furthermore, a fractal estimator is suggested in [44], which avoids the need to transmit the whole information of the ROI mask. In [15], the *Wavelet-based Compression scheme with Adaptive Prediction* (WCAP) method is proposed, where a row-order scan and an adaptive arithmetic coding are used to encode the ROI mask. Then, the rest of coefficients are regarded as the predictor and response variables of a prediction equation. The generated prediction equations are then applied to predict most wavelet coefficients except the lowest-resolution ones. Finally, an adaptive arithmetic encoder is adopted to encode the differences between the original and corresponding predicted coefficients. All of the above methods increase coding complexity, but more scaling values are allowed, giving different quality priorities to different ROIs.

### 1.3. Challenging Aspects Involved in Assessment of Image Coding Performance

A perceptual assessment is intended to match the human vision model in order to give a perceptual quality control parameter on coding. There are two approximations found in literature for human vision model: On one hand, common evaluation models are based on the *Error Sensitivity* paradigm, which assumes that an image quality assessment can be thought as a sum of an undistorted reference signal and an error signal [77, Sec. II]. This widely adopted assumption takes the loss of perceptual quality as directly related to the visibility of the error signal. On the other hand, the Structural Similarity (SSIM) index, proposed in [77], is used to measure the perceived structural variation rather than the perceived error. Newer improvements are made to enhance the behavior of this evaluation, as in [79]. In Figure 1-1, the different approximations used for Error Sensitivity are shown, referring notation  $\mathcal{T}$  to the transform operator,  $f_n$  to the extracted feature, and  $n$  to the number of features. These approximations are explained as follows:

- Human-based evaluation (Figure 1-1a) is assumed to be modeled into three processes: The contrast sensitivity filter, the lateral inhibition and the double stimulus scoring. The *Contrast Sensitivity* filter resides in the eye retina, which acquires luminance and chromatic stimuli from image. This stimuli is *filtered* due to the packing density of the retinal photo-receptor cells. For instance, the human eye has 3.75 cycles/pixel of sensitivity peak, and sensitivity decreases at frequency values on either side of the peak. The lateral inhibition resides in the *Primary Visual Cortex*, which reconstructs



**Figure 1-1.:** Perceptual quality assessments between source and distorted images. a) Human-based evaluation. b) Single measurement. c) Multi-channel (multi-factor) image-based evaluation. d) Wavelet-based evaluation.

image in the brain by both, linking *hypercolumns* across the spatial visual field with a shared orientation preference, and allowing cells to integrate visual information from spatially separated receptive fields. A *hypercolumn* can be interpreted as a *visual pixel*, representing the optical world at a single location, neatly decomposed into a complete set of orientations. Thus, lateral inhibition increases the contrast and sharpness in visual response, in order to detect perceptual features. According to [66], non-random horizontal connections within the cortical circuitry, which have been identified in mammals, allow cells to integrate visual information from spatially separated receptive fields. The visual system not only constructs a score of local orientations, but also accounts for spatial context and alignment by excitation and inhibition *a priori* [62]. Finally, *Double Stimulus scoring* is the process in which the human being compares the features of two images or sequences, one of which is a reference sequence. However, the subject is not told which one is the reference and is asked to rate each picture independently. Results are marked on a continuous scale which is divided into five sections, marked with adjectives for guidance: *Excellent*, *good*, *fair*, *poor* and *bad*. The continuous quality scale allows greater precision in judgments to be made by the observers.

- Some fully-referenced perceptual assessments use an unique measure to obtain scores, as shown in Figure 1-1b. Among this kind of perceptual evaluations, the following are illustrated: The *Perceptual Mean Square Error* (PMSE), which is proposed in [81, ref. 8], is the normalized mean squared error function between distorted and source images, which are previously filtered by some perceptual function. In [3], the *Quality Index based on Local Variance* (QILV) is proposed, which measures the mean local variance using *Singular Value Decomposition* (SVD) method [46]. Another perceptual measure is proposed in [47] as a direct application of the SVD theorem. This measure is denoted as the *Reflection Factor* (RF), which is calculated from the SVD applied on the original image. Finally, in [5] the *Entropy-based distortion* measure is proposed, defined as the differential conditional entropy between original and encoded representations. This measure uses the perceptual homogenization criteria as mean-estimate of the difference between those representations.
  
- Other fully-referenced perceptual assessments are composed of a multivariate combination of numerical values which represent the level of presence of perceptual features related with distortion (Figure 1-1c). One of these is the *Quality Index* (QI) proposed in [30], which is a multi-factor approach to human-based evaluation. This approach consists of a weighted linear combination of several feature measures, inspired in luminance differential stimuli. The parameters of the linear combination are calculated by linear regression, where the independent variables are the feature measures on each image of a selected database, and the observed variable is the human-based evaluation on the same image. The database used for the regression is the half of the LIVE database (see [14] and [30, pg. 324] for more details), where each image is associated to its respective human-made evaluation [14]. Thus, the linear combination aims to give an assessment value according to the type of distortion, and a high correlation value with the human evaluation. The following aspects are to be improved: Firstly, considering that feature measures are calculated on image, additional Fourier-spectral direct and inverse transformations are required to filter image by some contrast filter. Secondly, correlation value between perceptual assessment and human evaluation has to be increased. It is used for evaluation of JPEG compression schemes, where the reduction of DCT coefficients produces strong transitions at the frontier of blocks, called *Blocking Artifacts*. QI calculates the degradation in blocks of regular size by the analysis of variation on transitions of adjacent blocks. The blocks must be arranged regularly with same size, restricting measurement to those transitions that are strictly in block positions. The distorted image is evaluated using processes in image-spectral domain. It is considered as the conventional multi-factor assessment. A description of the processes involved in this type of evaluation is described as follows:
  - Source and distorted images are Fourier-spectral transformed to be filtered by some contrast sensitivity function, giving filtered counterparts, if each outcome

is Fourier-spectral inverse transformed in order to extract features required for image-based feature measures.

- Afterwards, three features are extracted from filtered images, as described in [30], giving  $n = 3$  corresponding values.

Thus, for each reconstructed image in considered database, there should be a source image, a human-made evaluation, which is called *Differential Mean Observed Score* (DMOS), and distortion measured vector  $\{d_1, d_2, d_3\}$ . As described in Figure 1-1c, a weighted linear combination is used to calculate the evaluation. A linear regression strategy is often used to calculate the weighting coefficients. Therefore, for each strategy, a collection of distortion vectors and their respective DMOS are used as training input of regression process, giving as a result the regression coefficient set  $\{\lambda_1, \lambda_2, \lambda_3\}_\eta$  for a trial  $\eta$ . A number of trials are conducted, then the mean value of those regression coefficients is calculated. This mean value is finally used in linear combination equation, and QI is ready to measure the distorted image, using source image as reference.

Accordingly, these models have the following processes in common: firstly, a sensitivity filter that shrinks visual frequency range to the human eye parameters; secondly, a set of features that are extracted from the filtered image; finally, a multi-variable regression strategy, which combines the value of each feature in one numerical score. Depending on the relevant grade of each feature, this score aims to identify different types of distortion, whenever the feature is related with the distortion phenomena. For instance, a measure related with the edge strength feature is useful to detect a type of distortion degrading edges. This ability can be extended to perceptually evaluate the performance of coders, because a different type of coders produces a different type of distortion on image. Therefore, it is expected that those feature metrics allow one coder to be identified from another. In literature, regression strategies are deployed by *Pooling* algorithms. The processes involved with error sensitivity paradigm are explained as follows:

- A *Contrast Sensitivity Function* (CSF) is used to enhance the contrast of image by using a linear band-pass filter.
- Features are extracted from filtered image to give perceptual elements required for feature measures. A description of those features is illustrated for segmentation and pattern recognition purposes in [37]. Other features can be calculated from transforms. For instance, the *Hough* transform is used to identify positions of arbitrary shapes, most commonly circles or ellipses. In [57] an example of current improvements on this transform is shown. All these extraction techniques are stated in image domain. Accordingly, details in images are the primary aspect to be kept for transmission purposes.
- Regression strategies are considered for achieving a higher grade of correlation between

a linear combination function and a comparative model, which is the human-based model in this case.

Though the correlation and sensitivity in measurement can be partly overcome by a linear combination of perceptual measures, it is desirable to include measures on coding methods that use the same representation, in order to avoid the increase of computational load in coding schemes. For instance, it would be adequate to use a measure calculated in the same wavelet domain representation of JPEG2000 standard.

## 1.4. Proposed Fully-Wavelet Quality Index

The proposed *Fully-Wavelet Quality Index* (FW-QI) assessment uses a wavelet-based CSF filter, and thereafter a linear combination of wavelet-based feature measures, whose weighting coefficients follows a selected regression strategy, as shown in Figure 1-1d). This assessment is adequate for implementation on control mechanisms inside wavelet-based coders, because only the encoding representation is required for measurement, avoiding the additional step given by reconstruction of image. This proposed assessment overcomes the following challenging aspects: first, the assessment should provide a numerical outcome related with human-based evaluations; second, the processing involved with evaluation must provide an adequate implementation for distortion control. The features involved in this assessment are related with the *human-based* evaluation process shown in Figure 1-1a). Relationships between human vision process and wavelet transform has been studied, in order to give general wavelet representations for features like edges, transitions and lines. For the sake of illustration, this assessment is depicted with the respective number of operations as follows:

- A wavelet-based coding process encodes the source image. It gives the respective source and encoded representations.
- Thereafter, these representations are filtered by a wavelet-based CSF (WCSF in Figure 1-1d)), giving their filtered counterparts.
- On these filtered representations, three wavelet-based measures are calculated. It gives three quality values of different features on image. The assessment comes out from the weighted linear combination of those three values.

The aim of this assessment is to give an adequate correlation with a human-based scoring. To achieve this, a linear regression strategy is often used to calculate the parameters of that linear combination. The regression requires the DMOS as the data of observed variable, and the feature measure vector  $\{f_{1,\psi}, f_{2,\psi}, f_{3,\psi}\}$  as the data of independent variables. As described in Figure 1-1d, a weighted linear combination is used to calculate the evaluation. Therefore, for each strategy, a collection of distortion vectors and their respective DMOS are

used as training input of regression process, giving as a result the weighted coefficient set  $\{\lambda_1, \lambda_2, \lambda_3\}_\eta$  for a trial  $\eta$ . A number of trials are conducted, then the mean value of those regression coefficients is calculated. This mean value is finally used in linear combination equation, and FW-QI assessment is ready to measure the distorted image, using source image as reference.

## 1.5. Objective

This work proposes an image assessment, named *Full-Wavelet Quality Index* (FW-QI) assessment, which is intended to evaluate the loss of quality on an image, by using a selected set of measures. These measures are classified as perceptual full-reference, which are based on scoring of error sensitivity. To achieve this evaluation, the following steps are described: i) Wavelet coefficients of the image are taken; ii) Sensitivity range is increased by filtering wavelet coefficients using an adaptive mask function; iii) Features are extracted from filtered representations, then values of these features are arguments of three different measures; iv) Final assessment value comes out from the linear combination of measures. In order to achieve this objective, the following specific goals need to be met:

- Firstly, an adequate selection of features used for measures, with the usage of only the wavelet coefficients coming out from encoding image, without requiring additional functions for the whole evaluation process.
- Secondly, an increase of correlation level in which the quality evaluation approximates the human visual model and stability in the regression strategy through the improvement of the CSF filter used for perceptual measurements and modifications to the conventional measurement.
- Finally, the implementation of assessment on a ROI-coding control scheme. In this case, FW-QI assessment is used as control parameter for *PO-Shift* bit-plane ranking. This adjustment is named perceptual-Wavelet-ranked *PO-Shift* (*WPO-Shift*) method.

The FW-QI assessment is used specifically for the quality assessment of image coders, which use wavelet as representation. The term *Image Access* is often related with image coding method that uses a distortion control strategy inside its coding process. The main application of image access is the process in which a representation of image is tracked from repository database, through an interface, to some displaying screen. This work is focused on JPEG2000 standard, which uses wavelet as kernel. The concerning problem relies on the fact that the conventional feature measures required for evaluation use a representation different than wavelet. Moreover, the perceptual evaluation should be close to human-made evaluations. The following is the contribution list of this work, where the relevance in comparison with other techniques is illustrated:

- An image quality assessment that achieves comparative results on correlation with respect to human-based evaluations. The assessment is based on the Quality Index Assessment proposed in [30], and uses a linear combination of three proposed perceptual measures. The contribution relies on the fact that these measures are designed on human-based constrains and wavelet operators as well, discovering acceptable correlation results.
- In literature, several Contrast Sensitivity Functions (CSFs) have been employed for recreating the behavior of retinal human eye. Among them, different CSFs in wavelet domain have also been constructed, giving a closer approach to human vision. Inside the proposed assessment, a wavelet-based CSF called *Adaptive-Mask CSF* (AM-CSF) is proposed for increasing the correlation with respect to a human visual model. The contribution of this CSF is its property of generalization for any wavelet structure, because CSF parameters are scaled according to the center frequency of wavelet function used in representation.
- An experiment to improve the sensitivity to newer images is made by testing different regression strategies, which are required for the calculation of linear combination coefficients within the assessment process. The contribution relies on the fact that the selection for a Principal Component Analysis achieves a better sensitivity.
- Because the proposed assessment is arranged entirely in wavelet domain, it is possible its implementation onto distortion control process within coding systems.

## 2. Image Quality Assessment Definitions

### 2.1. Preliminary

The aim of a perceptual quality assessment is to reach some level of analogy with human evaluations. This level can be increased by the use of a sensitivity filter and an adequate selection of features. However, for a desired implementation of assessment into coding scheme as distortion control parameter, it is recommended to extract these features from encoded representation instead to use the measurements on reconstructed image. Moreover, it is intended to find related operations that do not represent a high computational load for the coding scheme, specially because the computing times of feature extraction processes are commonly proportional to the size of image (Section §2.2). Hence, two aspects worthy of consideration about image quality assessments are the features involved in evaluation process (Section §2.3), and the methods for improving the correlation and sensitivity of assessment with respect to human evaluations.

### 2.2. Image Constrains

In this Section, the problem of distortion during transmission of encoded medical images is explained. For a clear conception about image space to be measured, following definitions are given:

**Definition 2.2.1** *An Image is a  $n$ -dimensional array  $\mathbf{X} \in \mathbb{R}^n$ , whose elements are located in coordinate vector  $\mathbf{m} = (m_1, \dots, m_n)$ . A value  $x : \mathbb{R}^n \mapsto \mathbb{R}, f(m_1, \dots, m_n) = x$  is assigned to each element, and it is denoted as Pixel value. An image has a finite number of elements, depending on the dimension  $n$ .*

For instance, color images commonly have  $n = 3$  dimensions called Height, Width, and *Colormap*. Thus, *Size* of an image is the rank for each dimension. So, image is rearranged as  $\mathbf{X} \in \mathbb{R}^{M_1 \times M_2 \times M_3}$  that corresponds to its height, width, color map and depth limits, respectively.

**Definition 2.2.2** *A Medical Image is an image acquired for clinical purposes. Therefore, every medical image involves a score related to a diagnose or examination of human body. They must reach some size and resolution requirements [19], as shown in Table 2-1, where*

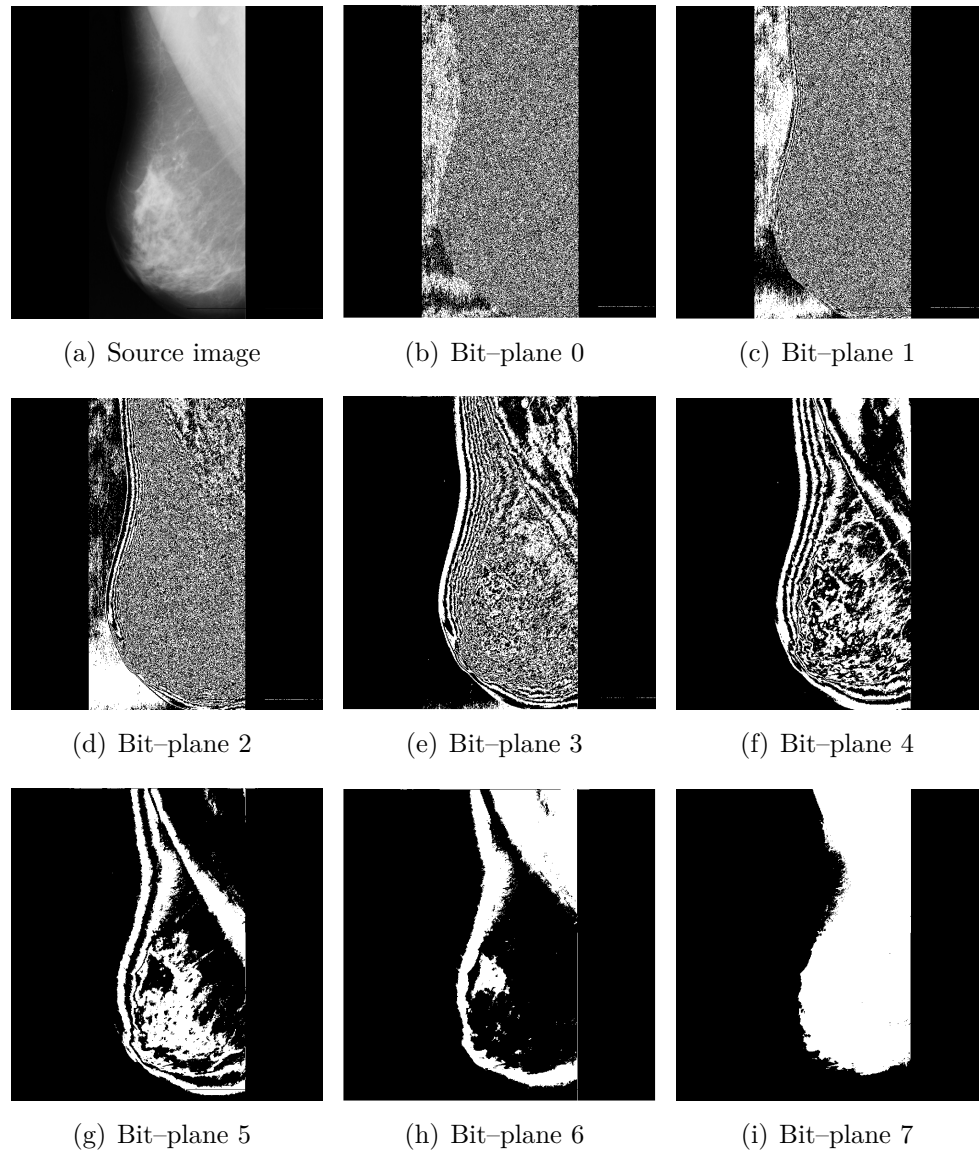
it is remarked that medical images require just one colormap, as shown in Table 2.1(a). Size of medical images, unless another definition is detailed, is  $M_1 \times M_2 \times 1$ .

**Table 2-1.:** Established parameters by Health and Social Care Department (Ministerio de la Protección social)[19].

Image set		Image features				Image size		
Category	Modality	Pixels		bits		source image (kbytes)	Comp. rate Typical	Comp. file (kbytes)
		Width	Height	Quant.	Storing			
Macro-photography	Topic photos (Dermatology)	1524	1120	24	24	5001	27	185
Micro-photography	Patology	800	600	24	24	1406	20	70
Radiology	Scanography	512	512	12	16	512	8	64
	Magnetic Resonance (MR)	256	256	12	16	512	8	64
	Ultrasound	512	512	6	8	256	10	26
	Nuclear Medical	128	128	8	8	16	10	2
	Digitized General x-rays	2048	2500	12	8	5000	10	500
	Frame Grabber	512	512	6	8	16	10	2
	Digitized mammo-gram	4096	4096	12	8	16384	10	1638

**Definition 2.2.3** A Bit-plane of a image is a binary matrix of the same size, where its bit elements having the same position in the respective binary numbers of the pixel value of image. The number of bit-planes in a image is defined as Image Depth  $M_b$ .

Figure 2-1 shows a medical image and its corresponding bit-planes. Incrementing a bit plane by one gives the final result half of a value of a previous bit-plane. Distortion is a



**Figure 2-1.:** Medical image and its bit-planes.

critical aspect in coding of medical images, because preservation of details is mandatory for areas on image where abnormalities are present. Thus, a distorted image can lead to a diagnosis misinterpretation, for instance, a false positive. Some levels of distortion affect the level of uncertainty that a transmitted medical image can present for diagnostic purposes. For this kind of images, a conventional transmission process, such as e-mail, file transfer protocol or sensor network protocol, suffers latency. It leads to an inefficient and discouraging

service for health care users (patients and experts). Moreover, a large transmission time at reduced bandwidths is often present. Hence, a common aspect comes into view: Distortion measurement is required for evaluation of image coders.

**Definition 2.2.4** *Image Coding is the process of transforming elements of an image into components of another representation, whose size is lower than the original. This process is composed of both the transform and the quantization of a source image, obtaining a representation susceptible to be transmitted. This representation is termed as Encoded, and the ratio between the information quantity of source image and the information quantity of encoded representation is the estimated Coding Rate. This rate is used as parameter to evaluate coding performance. An Image Coding System can be interpreted as the transmission process of an image, which is tracked from repository database, through an interface, to some displaying screen.*

Accordingly, a large image, such as a medical image, must be encoded in order to fit transmission bandwidth constrains. Thus, coding involves reduction of some components of image. Therefore, a large image is not suitable for transmission purposes. In other words, a large image can be transmitted only under coding arrangements, which is the case we are concerned with. In order to achieve a minimal margin of distortion and delay during transmission, the coder uses the redundancy property of images: an image often contains strongly correlated pixels, i.e., large regions whose pixel values are the same or almost the same. This redundancy can be explored in several ways, one of which is by predicting a pixel value based on the values of its neighboring pixels. The process of loss of details is called *Quantization*, indicating that a wider range of pixel values is normally mapped into a narrower range of output values. Thus, some non-redundant details can be lost, after which the losses must be estimated by the coder, as they lead to a quality decrease. It is the critical aspect during the assisted diagnosis stage, for instance, in tele-medicine applications [63]. Consequently, newer methods are intended to establish a reduction of data whether or not there is a low distortion margin.

### 2.2.1. Wavelet Image Representation

Therefore, for an adequate image coding process, a general representation is focused towards efficient treatment of data, which relies on (usually invertible) *transformations*. These have shown to be efficient to describe salient image features and to reduce the computational process, as well. It is accomplished because the transformations normally bring image to a representation with highly redundant elements, which can be neglected to be transmitted, providing that features are kept at lower redundant elements.

**Definition 2.2.5** *A Transformation  $\mathcal{T}\{\cdot\}$  is an operator which maps image function into another representation space, containing features which are otherwise difficult to extract directly from the image. For instance, luminance feature can be obtained from Radon represen-*

*tation, frequency feature can be obtained from Fourier transform, frequency and orientation features can be obtained from Wavelet transform, and so on.*

These transformations have shown to be efficient in both abilities, to describe image features and to reduce the computational process. Commonly, a representation space is involved with spectral features of image, like gradients, in the same manner as frequency feature is considered for one-dimensional signals. Several representations are used to manage the distortion level on the encoded image, achieving higher reduction rates with acceptable distortion scoring, as described in the following examples: an image coder, proposed in [58], uses the *Hartley Transform* (HT), which is a Discrete-real-Fourier Transform; other image coder, proposed in [16], uses *Discrete Cosine Transform* (DCT); the *PCA Coder* [45] uses the Karhunen-Loève Transform; the *Set Partitioning In Hierarchical Trees* (SPIHT) [67] and *Embedded Block Coding with Optimized Truncation* (EBCOT) [73] methods use Wavelet Transform. Currently, wavelet representation is used, due to its properties to provide an optimal signal reconstruction, reducing the loss of quality by rounding coefficient through coding. The benefits of such representation have led current coding schemes to use wavelet transform thanks to its capabilities of complete reconstruction. This can be seen in the use of integer-based wavelet functions for the estimation of wavelet coefficients, like the *LeGall 5/3* wavelet filter [41]. The use of wavelet coefficients for feature extraction facilitates the implementation of accurate measurement, avoiding an inverse transformation for extracting features from the image. An image can be represented by using the wavelet defined as follows:

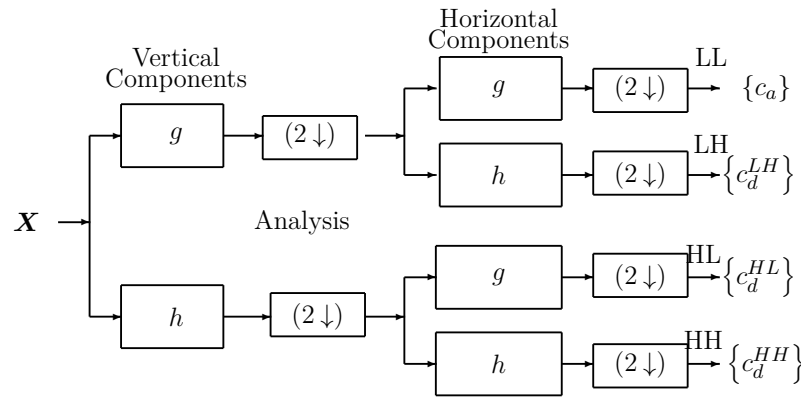
**Definition 2.2.6** *A compactly supported wavelet  $\psi(\mathbf{k}, j)$  is a function having its energy grouped into short intervals in considered spatial position  $\mathbf{k} \in \mathbb{N}^2$ , and decomposition level  $j \in \mathbb{N}$ . These functions are generated from a scaling and shifting of a source function, namely Mother Wavelet  $\psi(\mathbf{m})$ .*

Each mother wavelet is scaled and translated using the Equation 2-1, in order to obtain representation coefficients of each element of image:

$$\psi(\mathbf{k}, j) = \frac{1}{2^{j/2}} \psi\left(\frac{\mathbf{m} - \mathbf{k}}{2^j}\right) \quad (2-1)$$

Let  $\mathbf{C} \in \mathbb{R}^{M_1 \times M_2 \times J}$ ,  $\forall j = 1, \dots, J$  be the set of coefficients  $c \in \mathbf{C}$ , generated by the convolution between image and wavelet functions. This set of coefficients  $\{c_{\mathbf{k}, j, \theta}\} \in \mathbf{C}$  is located at index  $(\mathbf{k}, j)$ . It can be interpreted as a sub-sample of the convolution of a function  $\mathbf{X}$  with a wavelet family  $\Psi_j$ . Though index  $j$  is termed as the decomposition level, the term *sub-band* is used to determine the location of a sub-set of wavelet coefficients. For example,  $c_{\mathbf{k}, j}$  and  $c_{\mathbf{k}, j+1}$  coefficients are located at the same spatial coordinate  $\mathbf{k}$ , but in two different sub-bands. A *dyadic* separation of  $2^{-j}$  between each sub-band is stated such that coefficients

can be scaled between sub-bands in factor of  $2^{-j}$ . This is a scaling of the *bit-plane*  $b$ , so that a wavelet coefficient can be represented as a bit-ranked sequence. Although wavelet functions accomplish the orthonormal condition in space  $\mathbb{L}_2(\mathbb{R})$ , i.e.,  $\|\psi(s, \tau, t)\|_{L^2} = \|\psi\|_{L^2} = 1$ , the main disadvantage of orthogonal wavelets is that they are highly constrained. For instance, wavelet coders which use linear phase filters have improved performance. However, there are no nontrivial orthogonal wavelet filters with linear phase. *Biorthogonal* wavelets are a more general class of filters which do not retain the strict orthogonality conditions of orthogonal wavelets. Thus, bi-orthogonal functions are considered for image coders, because a better coding performance can be achieved. A well known algorithm is proposed by [49] to calculate the wavelet coefficients from the *Lifting Scheme* by using a filter bank analysis-synthesis system of two channels, as shown in Figure 2-2. For each sub-band, the



**Figure 2-2.:** Two-dimensional wavelet filter bank. The  $(2 \downarrow)$  term denotes dyadic down-sampling

coefficients pass through filters  $h(\cdot)$  and  $g(\cdot)$  at directions  $\theta = HL, LH$ , where  $HL$  indicates that the filter is applied on vertical components, and  $LH$  that the filter is applied on horizontal components. It results in the following coefficient subsets: approximation (LL)  $\{c_a\}$ , detail vertical  $\{c_d^{HL}\}$ , detail horizontal  $\{c_d^{LH}\}$  and detail diagonal  $\{c_d^{HH}\}$ . Therefore, the Finite-Impulsive-Response low-pass filter  $h(\mathbf{k}, \theta)$  belongs to the coarse sub-band scaling function and the Finite-Impulsive-Response low-pass filter  $g(\mathbf{k}, \theta)$  belongs to the fine sub-band scaling function. Assuming that the coefficient set  $\{c_j\}$  is located in sub-band space  $V_j$ , coefficients in sub-band  $j + 1$  are calculated by Equation (2-2):

$$c(j+1, \mathbf{k}) = \sum_{-\infty}^{\infty} x 2^{(j+1)/2} \psi(2^{j+1} \mathbf{m} - \mathbf{k}) dt \quad (2-2)$$

So, coefficients  $c_a(j+1, \mathbf{k})$  pass through another two-channel filter bank to the next sub-band as shown in Figure 2-2, calculating the scaling and shifting coefficients at decomposition level  $j$  for the coarse and fine channels. Accordingly, detail coefficients belong to the finest sub-bands and approximation coefficients belong to the coarsest channel. Coarser

sub-bands contain elements that can be reconstructed at finer sub-band. The coarsest (approximation) sub-band is highly redundant, and the finest sub-band contains the lowest redundant elements.

$$c_a(j, k) = \sum_l h(l - 2\mathbf{k}, \theta) c_a(j + 1, l) \quad \text{and} \quad c_d(j, k) = \sum_l g(l - 2\mathbf{k}, \theta) c_a(j + 1, l) \quad (2-3)$$

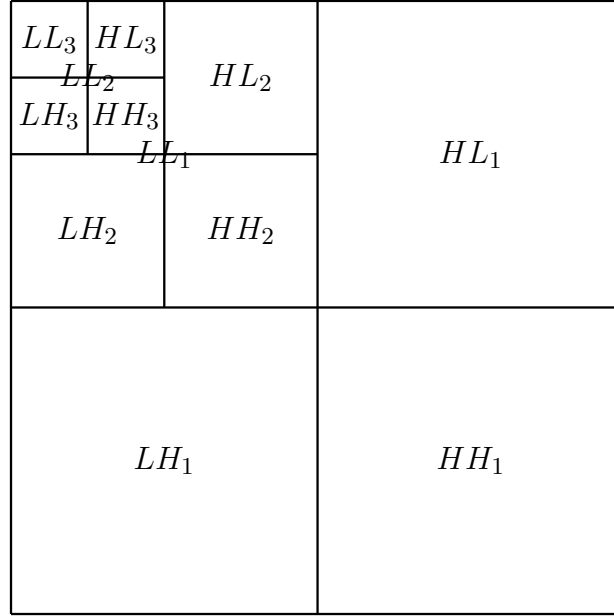
However, though finer sub-bands are sometimes not transmitted to fit rate constrains, coarser sub-band allow reconstruction of these lost sub-bands at synthesis scheme, when approximation and detail coefficients belonging to some sub-band are over-sampled and filtered to give the scaling coefficients at its coarser sub-band:

$$c_a(j + 1, \mathbf{k}) = \sum_l [h(\mathbf{k} - 2l, \theta) c_a(j, l) + g(l - 2\mathbf{k}) c_d(j, l)] \quad (2-4)$$

Equations (2-3) and (2-4) define the forward and inverse discrete wavelet transform. The algorithm defined in [49] establishes a relationship between scaling approximation coefficients  $c_a(\mathbf{k}, j - 1)$  at finer index, and detail coefficients ( $c_d(\mathbf{k}, j)$ ) at coarser index  $j$ . When the approximation coefficient  $c_a(\mathbf{k}, j)$  pass through the analysis scheme from the filter bank, new approximation and detail coefficients are generated at the following coarser index  $j + 1$ . Thus, image reconstruction can be stated from approximation coefficients  $c_a(\mathbf{k}, J)$  in sub-band space  $V_J$ , by inverse filtering of Equation (2-3). For this purpose, approximation coefficients are  $(J - N)$  times lifted up using a synthesis filter bank structure located at  $(J - N)$  finer sub-bands, producing wavelet coefficients at finer sub-bands  $N \leq j < J$ , thus making a reconstruction of image. The details of the wavelet representation for image are shown in Figure 2-3. The preference of image coding schemes for wavelet transform concerns the fact that this representation is an acceptable approximation to the Human Visual System model [48]. Briefly explained, wavelet transform, as human-eye, splits an image into several frequency bands which can be processed independently [23], and it is a multi-resolution transform that permits the location of image features such as smooth areas, edges or textured areas [24]. It is critical for Tele-medicine applications, where the preservation of image details is required for low transmission rates. This preservation implies the use of techniques which measure the quality of image, a critical aspect during the assisted diagnosis stage.

### 2.2.2. Image Feature Extraction

The details in the image are the primary aspect to be measured for quality evaluations. Each detail is conformed by different *Features*. Recent studies have demonstrated that the human eye takes features from reflected light and reconstructs these as brain images [10]. Several features used for measuring purposes are also used for segmentation and pattern recognition purposes [37]. Commonly, these features are extracted in image domain using spatial transformations. For instance, the *Hough* transform is used to identify positions of arbitrary shapes, most commonly circles or ellipses [57]. In this case, *Edge* and *Transition* are



**Figure 2-3.:** Two-dimensional wavelet representation using 3 decomposition levels

selected as general features for perceptual distortion measurement. A model approximation of these features is described in [22], where a Gaussian filter is used. Let  $G_s(m_1, m_2)$  be the Gaussian function at an arbitrary parameter  $\sigma$ :

$$G_\sigma(m_1, m_2) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{1}{2\sigma^2(m_1^2 + m_2^2)}\right)$$

A *Transition* function is defined as:

$$T_\sigma(m_1, m_2) = \frac{A}{2} \left(1 + \operatorname{erf}\left(\frac{x}{\sigma\sqrt{2}}\right)\right),$$

where  $\operatorname{erf}(\cdot)$  is the error function, and  $A$  and  $\sigma$  are the respective amplitude and smoothing size of the transition. An *Edge* function is defined as:

$$L_\sigma(m_1, m_2) = 2\pi\sigma^2 A G_\sigma(m_1, 0),$$

where  $A$  and  $\sigma$  are the amplitude and thickness of the edge, respectively. Thus, a variation of feature parameters, like  $A$  and  $\sigma$ , indicates a variation of the quality of detail. Two main approaches are found for feature extraction in wavelet domain:

- Firstly, the features can be detected and classified by complex Gaussian derivative wavelet function defined in [22]. This representation provides an acceptable separability criterion between transitions and edges. However, this method requires the strict use of the complex wavelet Gaussian filter for adjusting by parameter scaling in order to

define the candidates, but filter bank decomposition scheme is not supported for this scheme, so the computational cost becomes inconvenient.

- Secondly, in *Orientation Score* representation, features are obtained from a general wavelet-based space, which obtains close approximations to findings of human perception and vision [23]. This representation has gained acceptance since its results on reconstruction of details, for instance, in Microscopy applications. However, computational cost and the addition of data related with orientation map of an image are aspects that could not be suitable for Image Coding purposes.

To extract the considered edge, line and transition features, which are required for the proposed measures, following basic operators are enumerated:

- Firstly, *Maximal Module* consists of a mapping of detail wavelet coefficients which achieve the maximal gradient. These gradients are calculated by taking respective wavelet coefficients  $c_d(j, \mathbf{k}, \theta)$  as:

$$\angle c_d(j, \mathbf{k}) = \begin{cases} \arctan(c_d(j, \mathbf{k}, LH)/c_d(j, \mathbf{k}, HL)), & c_d(j, \mathbf{k}, HL) \geq 0 \\ \pi - \arctan(c_d(j, \mathbf{k}, LH)/c_d(j, \mathbf{k}, HL)), & \text{otherwise} \end{cases}$$

Then, maximal modules are calculated for wavelet coefficients wavelet as:

$$M^j(\mathbf{k}) \geq M^j(\mathbf{k} \pm \mathbf{v}_j) \quad (2-5)$$

where  $M^j(\mathbf{k}) = \sqrt{\sum_{\theta=HL, LH, HH} |c_d(j, \mathbf{k}, \theta)|^2}$ ,  $\forall k \in \mathbb{Z}^2$ , is the wavelet module. A value from the set  $\mathbf{v}_j \in \{(1,0), (0,1), (1,1), (1,-1)\}$  is chosen, so that it is located near the vector with coordinates  $(-\sin(\angle c_d(j, \mathbf{k})), \cos(\angle c_d(j, \mathbf{k})))$ , orthogonal to the direction of local edge. After calculating the edge point, position  $2^j \mathbf{k}$ , locates the maximal module. Thus,  $M^j(\mathbf{k})$  is the magnitude and  $\angle c_d(j, \mathbf{k})$  is the direction of edge at sub-band  $j$ , respectively.

- Secondly, *Detail variance* is the local variance between two blocks between sub-bands, and it is calculated as:

$$\sigma_{c,c'} = 2^{-2J} \sum_{j=1}^J \sum_{\theta=HL, LH, HH} \sum_{\mathbf{k} \in \mathbf{B}_\psi, \mathbf{k}' \in \mathbf{B}'_\psi} c_d(\mathbf{k}, j, \theta) c_d(\mathbf{k}', j, \theta), \quad (2-6)$$

where  $\mathbf{B}_\psi$  and  $\mathbf{B}'_\psi$  are dyadic squares of size  $M_1 2^{-J} \times M_2 2^{-J}$ . the position index  $\mathbf{k}$  ranges over all sub-bands.

- Finally, *Local mean* is also used for feature extraction purposes. Whereas blocks  $\mathbf{B}$  are dyadic squares, *Haar* wavelets can be used to compute these blocks. This corresponds to the set of coarsest sub-bands, specifically the approximation coefficient set  $\{c_a\} \in$

$\mathbf{B}_\psi$ . Assuming that coarser coefficients come from a number of  $2J$  low-band filtering processes, then mean value of any block  $\mathbf{B}$  of size  $m_B \times m_B$  on image can be obtained from approximation coefficients of this block at decomposition level  $J = \lfloor \log_2 m_B \rfloor$ . Thus, calculation of mean becomes

$$\mu_{\mathbf{B}} = 2^{-J} c_a(\mathbf{k}), \quad (2-7)$$

where for each position  $\mathbf{m}$  of  $x \in \mathbf{B}$ , the statement  $\lfloor 2^{-J} \mathbf{m} \rfloor = \mathbf{k}$  mets.

Let  $J_0$  be the number of decomposition levels required to achieve desired mean value and  $J$  the number of decomposition levels of wavelet structure. If  $J_0 > J$ , then the mean is computed from another successive analysis process, unless maximal decomposition value is achieved. In latter case, mean cannot be calculated for block of this size. If  $J_0 < J$ , then the number of decomposition levels are reduced from inverse synthesis process until  $J_0 = J$ , and the mean is calculated. Due to the properties of pyramid structure underlying the wavelet transform, the partial inversion algorithm will use strictly less operations than direct computation of the mean values would necessitate.

## 2.3. Image Quality Assessment

A general methodology for measuring the performance of a coding method uses a plot of the distortion measured against each different bandwidth constraint, determining the distortion level that the evaluated coder produces on an image for a given rate restriction. This plot is named *Rate-Distortion Curve*, which is sketched by using some distortion measure evaluating the coder performance. This curve describes a monotonic function of the rate variable. Therefore, a minimal – or maximal – value can be found in this curve, where distortion on encoded image fits some rate restriction and grant QoS requirements as well.

**Definition 2.3.1** *An Image Quality Assessment is the procedure giving a numerical value of the amount of distortion on image. Thus, the main idea behind the assessment is to give a numerical approximation of the distortion detected on those altered elements of image. It can be stated as a single quality measure or a pool of several measures.*

An image quality assessment is therefore composed of the following steps:

- A source image is encoded to achieve a rate  $R$ , obtaining a representation  $\mathbf{Y} = \mathcal{Q}\{\mathcal{T}\{\mathbf{X}\}\}$  susceptible to be transmitted, where  $\mathcal{Q}\{\cdot\}$  represents the rounding operator or *quantization* process to reduce the number of representation elements, and  $\mathcal{T}\{\cdot\}$  is the transformation used for the representation. This representation is named *Encoded*.
- Commonly, encoded representation differs from source, since elements of the former are truncated versions of the latter, with the purpose of achieving bandwidth restrictions.

This difference is measured, either between source  $\mathbf{X}$  and reconstructed  $\tilde{\mathbf{X}} = \mathcal{T}^{-1}\{\mathbf{Y}\}$  images or between source  $\mathcal{T}\{\mathbf{X}\}$  and encoded  $\mathbf{Y}$  representations, depending on the assessment.

When a large image travels from a repository data to a remote destination through channels with reduced bandwidth, an undesired latency effect at destination is perceived. Therefore, a transformed and then quantized image is transmitted to avoid this effect for a fitted rate restriction  $R$ . However, the encoded image is prone to distortion when it is transmitted, because it is corrupted primarily by quantization distortion. So, if at least pertinent elements are transmitted with some grade of quality, remaining image elements can be transmitted at subsequent time intervals, without apparent annoyance. Distortion measure  $D$  belongs to measuring space  $(\mathbb{R}^{M_1 \times M_2}, \mathbf{X}, D)$ . However, other measuring spaces, such as the conventional human-made evaluation and the wavelet space  $(\mathbb{R}^{M_1 \times M_2 \times J}, \mathbf{C}, D)$  are also considered. The conventional measuring space for determination of image quality is based on the QoS, which is a human-based observation that indicates perceptually the grade of distortion level that an image achieves. It indicates that a perceptual evaluation is an adequate way of constructing an assessment close to this human assessment. Specifically for a perceptual evaluation, a pool of measures is required. Thus, a *Perceptual Measure* obtains a numerical value upon the distance between source and distorted images based in perceptual features. In order to obtain these features, the introduction of a transform  $\mathcal{T}\{\cdot\}$  is usually required. So, perceptual measure is defined as distance function between two transformed images  $d_{\mathcal{T}} = d(\mathcal{T}\{\mathbf{X}\}, \mathcal{T}\{\tilde{\mathbf{X}}\})$ , where  $\mathbf{X}$  and  $\tilde{\mathbf{X}}$  are the respective source and reconstructed images. Therefore, the behavior of distortion can also be described by using some perceptual measure. The aspect worthy of consideration in selection of perceptual measurements for coders is that perceptual measures normally use features that are taken from image domain, because parameter adjustment is provided from previous perceptual evaluations.

## 2.4. Perceptual Image Quality Assessment

The goal of this kind of assessment is to assign a numerical score to a perceptual-measured phenomenon. It requires two main steps: The first one is the *Contrast Sensitivity Function*, which enhances perceptual features through a filtering process. Specifically, this filter is intended to enhance the features related with human-made perception, such as edges, transitions and textures. The second one is a linear combination of perceptual measures, based on those enhanced perceptual features. Since this score must be in accordance with perceptual behavior, the evaluation must be accurate enough to correspond to a good interpretation of scores, and achieve enough sensitivity to measure any distortion variation.

**Definition 2.4.1** *Perceptual Evaluation is the determination or estimation of human-based scores, which can be compared by assigning a numerical value in terms of a unit of measure-*

ment. Thus, quantitative attributes are those possible to measure, at least in principle. Thus, numbers are assigned based on correspondences or similarities between the structure of number systems and the structure of human-based qualitative systems. A property is quantitative if such structural similarities can be established.

For perceptual assessments, the features are extracted from filtered elements and thereafter perceptual measurement is made upon operations on these features. Perceptual methods become more worthy of parameter optimization for the following reasons, among others: the sensitivity parametrization and adjustment of this measures is based on biological and psychological human-eye process approximations, like *Human Visual System* [6]. Also, perceptual measurements are *Observation Scores*, intended to behave analogue to human-made perceptual evaluations on data quality. This is the characteristic which relates a distance measurement to human-made scoring. This scoring is thus analogue to a contrast comparison between source image  $\mathbf{X} \in \mathbb{R}^{M_1 \times M_2}$  of size  $M_1 \times M_2$  and distorted image  $\tilde{\mathbf{X}} \in \mathbb{R}^{M_1 \times M_2}$  of the same size. Therefore, a loss of quality detected in the distorted image should be enclosed with an monotonic quality measure value. Two perceptual assessments are further considered: The *Quality Index based on Local Variance* (QILV) that is proposed in [3], and the the *Reflection Factor* (RF) that is proposed in [47]. QILV consists of the linear combination of local variance estimator functions, calculated as:

$$QILV = \frac{2\mu_{V_{\mathbf{X}}}\mu_{V_{\tilde{\mathbf{X}}}}}{\mu_{V_{\mathbf{X}}} + \mu_{V_{\tilde{\mathbf{X}}}}} \cdot \frac{2\sigma_{V_{\mathbf{X}}}\sigma_{V_{\tilde{\mathbf{X}}}}}{\sigma_{V_{\mathbf{X}}}^2 + \sigma_{V_{\tilde{\mathbf{X}}}}^2} \cdot \frac{\sigma_{V_{\mathbf{X}}V_{\tilde{\mathbf{X}}}}}{\sigma_{V_{\mathbf{X}}} + \sigma_{V_{\tilde{\mathbf{X}}}}^2},$$

where  $\mu_{V_{\mathbf{X}}}$  and  $\sigma_{V_{\mathbf{X}}}$  are the respective expected value and standard deviation of the local variance calculated for each element of source image  $\mathbf{X}$ , and  $\sigma_{V_{\mathbf{X}}V_{\tilde{\mathbf{X}}}}$  is the covariance between the local variances of the source  $\mathbf{X}$  and distorted  $\tilde{\mathbf{X}}$  images. A QILV outcome ranges between  $[0, 1]$ , tending towards zero when quality decreases, like PSNR. The RF is calculated as:

$$RF = \frac{\sum_{i=1}^m |d_i| w_i}{\sum_{i=1}^m s_i},$$

where  $s_i$  is each one of the  $m$  nonzero singular values from Singular Value Decomposition (SVD) method on source image  $\mathbf{X}$ ,  $d_i$  is each one of the elements of the vector, which is calculated from the reliable difference factor between the original and the distorted image, and  $w_i$  is each one of the elements of the normalized vector of singular values on source image. RF values are normally below one, but tending towards zero if quality in distorted image increases.

In order to improve the correlation between data obtained from perceptual measurement and human-based scoring values, the following adjustments are suggested in [30]: firstly, a Sensitivity Filter, which increases the details used for feature extraction; secondly, an adequate selection of features, providing that they are closely related with human-eye process; and finally, an adequate selection of weighting coefficients to give a normalized scoring.

### 2.4.1. Contrast Sensitivity Function

**Definition 2.4.2** *The Contrast Sensitivity Function (CSF) is an operator which enhances contrast of image by using a linear band-pass filter, increasing the sensitivity at middle frequencies.*

In order to allow the detection of features related with perceptual distortion measurement, CSF operator is described by the spectral transfer function  $\mathbf{S}$ , which filters an image as follows:

$$\mathbf{X}' = \mathcal{T}^{-1} \{ \mathbf{S} \cdot \mathcal{T} \{ \mathbf{X} \} \} , \quad (2-8)$$

where  $\mathbf{X}'$  is the filtered image, and  $\mathcal{T}\{\cdot\}$  and  $\mathcal{T}^{-1}\{\cdot\}$  are the spectral (direct) and image (inverse) transforms, respectively. Thus, the transfer function becomes a mask  $S : (u, v) \mapsto \mathbb{R}$  with spectral arguments  $(u, v)$ , which is given as follows:

$$S(u, v) = \left( \frac{3}{2} \exp \left( \frac{-\sigma^2 \omega^2}{2} \right) - \exp \left( -2\sigma^2 \omega^2 \right) \right) \frac{1 + \exp(\beta(\omega - \omega_0)) \cos^2(2\theta)}{1 + \exp(\beta(\omega - \omega_0))} , \quad (2-9)$$

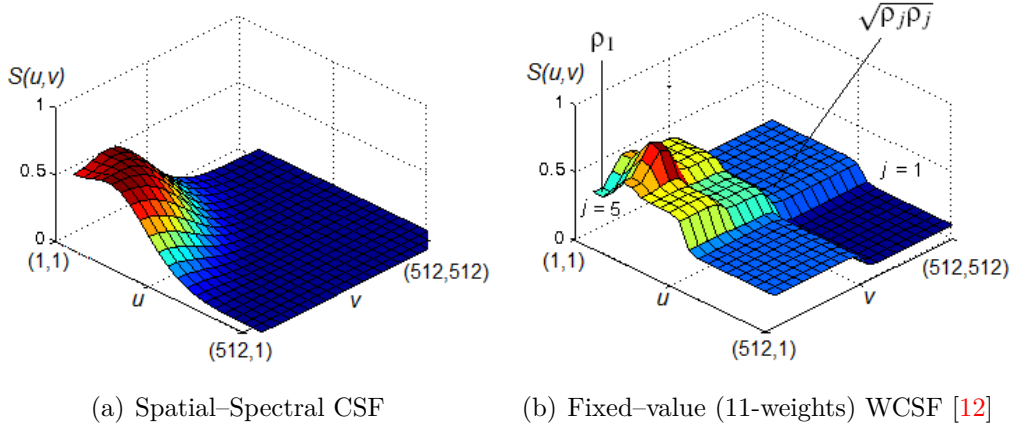
where  $\omega = \sqrt{u^2 + v^2}$  is the normalized contrast frequency,  $\theta = \arctan(v/u)$ , and  $\sigma$ ,  $\beta$  and  $\omega_0$  are the selected values which accomplish the observability condition. Values of  $\sigma = 4$  cycles/degree,  $\beta = 8$  and  $\omega_0 = 11.13$  cycles/degree, as recommended in [25]. In order to lighten the exponential operations in Equation (2-9), the following CSF operator is recommended [42]:

$$S(u, v) = 2.6(0.0192 + 0.114\omega) \exp(-(0.114\omega)^{1.1}) . \quad (2-10)$$

In the same manner, a wavelet-based CSF can be constructed using a transform map given as a function of index  $\{j, \mathbf{k}, \theta\}$ . Each wavelet coefficient is multiplied by each element of map, providing the following aspects: First of all, for approximation coefficients  $c_a(\mathbf{k})$ , a low-pass filter is used, after which, for detail coefficients  $c_d(\mathbf{k}, j, \theta)$ , a directional band-pass filter is used, either in vertical, horizontal or diagonal (according to the value of  $\theta = HL, LH, HH$ ), as detailed in [49]. Therefore, given both source  $\mathbf{C}$  and encoded  $\mathbf{Y}_\psi$  representations, processed wavelet coefficients  $\mathbf{C}'$  and  $\mathbf{Y}'_\psi$  can be obtained by a simple scale-dependent re-weighting. However, the main problem that concerns CSF function of Equations (2-9) and (2-10) is that both equations have no analytical impulse response function, thus filtering must be implemented in Fourier-spectral domain, increasing number of transformations.

By using wavelet-based representation, the inverse transformation is not required, because components required for filtering are already ordered in spectral sub-bands. For this reason, the use of wavelet-based Contrast Sensitivity Function (WCSF) for feature enhancing is considered. This function weights the wavelet coefficients relative to their perceptual importance. Accordingly, for a first approximation, the linear filter  $\mathbf{S}$  is calculated in the

wavelet domain, in order to weight the wavelet coefficients which belong to selected frequencies. From this assumption, the fixed-value *11-weights Mask Weighting* WCSF method is proposed in [12], where a set of CSF mask values  $\{\rho_j, \delta_j\}$ ,  $j = 1, \dots, 5$  is introduced. As shown in Figure 2.4(b), each  $\rho_j$  value corresponds to a peak value in detail sub-band  $j$ ; whereas  $\delta_j$  is related to a peak value of approximation sub-band  $j = 5$ . All these values are normalized, i.e.,  $\min\{\rho_1, \dots, \rho_5, \delta_1, \dots, \delta_5\} = 1$ , so that the admissibility condition holds. In addition, detail coefficients  $\{c_d(j)\}$  are multiplied by  $\sqrt{\delta_j \rho_j}$  value at sub-band  $j$ , while approximation coefficients  $\{c_a\}$  are multiplied by  $\delta_1$  value. However, the present advantage of the perceptual distortion measures is faced up with the additional (and perhaps different) transform operator on the image for feature extraction, the inverse transform required to suit encoded image into image-spatial domain, and the transform required for the perceptual CSF enhancement. Another CSF mask considers the noticeable difference between luminance [43].



**Figure 2-4.:** Simulation of CSF masking by considered methods, over a considered spectra of  $512 \times 512$  pixel size. For the WCSF method, 5 decomposition levels and biorthogonal 9/7 wavelet are used. Dark-red areas illustrate the maximal value of the CSF mask assigned to pixels

### 2.4.2. Feature Measures

As FW-QI is an improvement of the *Quality Index* (QI) assessment proposed in [30], which is depicted in following equation:

$$QI = 5 \cdot \left[ 1 - k \cdot \frac{w_1 d_1 + w_2 d_2 + w_3 d_3}{w_1 + w_2 + w_3} \right], \quad (2-11)$$

the measures involved in this assessment are described as follows. QI is a linear combination of three perceptual measures  $d_1$ ,  $d_2$  and  $d_3$ , respectively: the *Blockiness*, *Edge Error* and

*Visual Impairment* perceptual measures.  $w_1 = 1.5$ ,  $w_2 = 2$ ,  $w_3 = 5$ , and  $k = 4$  are the weighted factors resulting from the evaluation of physiological studies of perception of distortion. For the sake of explanation, the term *Feature measures* is used for these selected measures, the same as those defined in [52], distinguishing them from main QI assessment.

Thereafter, feature measures are *Blockiness*  $d_1$ , *Edge Error*  $d_2$  and *Visual Impairment*  $d_3$ , and they are explained as follows:

**Definition 2.4.3** *Blockiness measure is commonly used for block-based image coders. It measures the descent of details inside each block of regular size, resulting in homogeneous blocks at the highest compression ratio, and the appearance of horizontal and vertical borders in correspondence with each block transition.*

**Definition 2.4.4** *Edge Error measure evaluates the edge similarity between source and distorted image. It is the sum of squared errors between edges into source and distorted images, weighted by the corresponding edge weight value at source image. Edge values are calculated by any edge estimator, like Sobel, Canny or Roberts edge detectors. They are widely accepted to extract edges, transitions and lines. Several improvements on these detectors are still in process, as presented in [53].*

**Definition 2.4.5** *Visual Impairment measures the effect of each distorted pixel element over global distortion on image, by averaging block similarity and luminance distortion features over the whole image.*

### 2.4.3. Regression Strategies

Concerning variability of the measure when a linear regression is used, the Linear, PCA and Artificial Neural Network (ANN) regression strategies are considered as being included in the proposed FW-QI. Due to the fact that a linear combination of feature measures requires a set of weighting coefficients, the behavior of regression strategies is meaningful, because the calculated regression coefficients can provide a stronger weight to the feature measure that is most relevant for the orientation score. Therefore, any suboptimal characterization of feature measure or inaccuracy in sensitivity filter can vary the value of estimated regression coefficients in a significant way. The following selected regression strategies are considered for several works in order to achieve a higher grade of correlation between evaluations and a theoretical model, which is the human-based model in this case.

- First of all, the multi-variate linear regression (LR) strategy, where the regression coefficients are calculated linearly on a training set that is composed of both the DMOS values and the quality measured for each image. The linear regression is stated as  $QI = \lambda_0 + \lambda_1 d_1 + \lambda_2 d_2 + \lambda_3 d_3$ , where  $\{\lambda_i : i = 0, 1, 2, 3\} \in \mathbb{R}^1$  are the multi-variate linear regression coefficients. Coefficient  $\lambda_0$  can be neglected, because DMOS values are differentials.

- Secondly, the Artificial Neural Network regression (ANNR) strategy which is studied in [70], and where DMOS data is used as output and calculated measures for each image are used as inputs of an *Number of selected measures–3–DMOS value* layer perceptron Back–propagation Network is used. Assuming  $\boldsymbol{\lambda}_h = \{\lambda_h(i, j)\}$  as the set of weights of the neural connection between the  $i$ –th input and the  $j$ –th hidden–layered perceptron and  $\boldsymbol{\lambda}_o = \{\lambda_o(j, k)\}$  as the weight set of the neural connection between the  $j$ –th hidden–layered perceptron and the  $k$ –th output perceptron, the estimated DMOS (pDMOS) becomes:

$$pDMOS = f_o(\langle \mathbf{f}_h(\langle \mathbf{d}, \boldsymbol{\lambda}_h \rangle), \boldsymbol{\lambda}_o \rangle) ,$$

where  $\langle \cdot, \cdot \rangle$  denotes feature product operator,  $\mathbf{d}$  is the vector of input values of considered measures,  $\mathbf{f}_h$  is the vector of grouping functions at the hidden layer, and  $f_o$  is the grouping function at the output layer. From [70],  $\mathbf{f}_h$  is composed of tangential sigmoid functions and  $f_o$  is a linear function.

- Thirdly, a *Principal Component Analysis*, as recommended in [13], is considered because it selects a low dimensionality sub–space which restricts the set of possible regression functions. Moreover, feature measures can be considered for a representation by using PCA space. It can be considered as the Karhunen–Loève transform of considered measures, opening the possibility for a PCA regression for representation of measures. Thus, an optimization for the *Partial Least Squares* (PLS) algorithm for linear regression, proposed in [13], is used to translate the measures into a Principal Component Analysis regression (PCAR) strategy. This strategy gives an increase in correlation, without requiring any additional spectral transform. Let  $\mathbf{D}\mathbf{w} = \mathbf{d}$  be the equation of representation of measures  $\mathbf{D}$  by using some projection on basis  $\mathbf{w}$ . Assuming that equation can be solved by *Least Squares* method, the basis is calculated as:

$$\mathbf{w} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{d} .$$

However, in the case considered here, measures  $\mathbf{d} = (d_{1,x}, d_{2,x}, d_{3,x})$  are highly redundant themselves, bringing the system  $(\mathbf{D}^T \mathbf{D})^{-1}$ , with  $T$  as transpose operator, to be ill–conditioned. Therefore, PCA regression seems an appropriate candidate for preserving the parameters of system when the distortion source changes. This regression strategy is stated as follows: Let  $\mathbf{D} \in \mathbb{R}^{m \times n}$  be the data matrix of  $n$  values measured on every  $m$ –th image in considered database, and  $\boldsymbol{\lambda} \in \mathbb{R}^{n \times 1}$  be the associated weighted vector such as  $\mathbf{D}\boldsymbol{\lambda} = \mathbf{d}_{dmos}$ , where  $\mathbf{d}_{dmos} \in \mathbb{R}^{m \times 1}$  is the vector of DMOS values calculated for every image of the underlying database. Moreover, the data matrix can be also represented as  $\mathbf{D} = \widetilde{\mathbf{D}}\mathbf{V}$ , where  $\widetilde{\mathbf{D}} \in \mathbb{R}^{m \times n}$  is the set represented by some orthogonal space generated from principal component analysis and  $\mathbf{V} \in \mathbb{R}^{n \times n}$  is the eigen–vector set associated to the singular value decomposition (SVD) of matrix

$\mathbf{D}^T \mathbf{D} / (m - 1)$ . In this sense,  $\mathbf{d}_{dmos} \approx \tilde{\mathbf{D}} \boldsymbol{\lambda}$ . Thereafter, by least squares we obtain regression coefficients with:

$$\boldsymbol{\lambda} \approx (\tilde{\mathbf{D}}^T \tilde{\mathbf{D}})^{-1} \tilde{\mathbf{D}}^T \mathbf{d}_{dmos}$$

Though the correlation and sensitivity in measurement can be partially overcome by a linear combination of perceptual measures, like Equation (2-11), an undesirable inclusion of additional processing is attained, especially for coding methods that use different representation, e.g., use of wavelet domain representation in JPEG2000 standard.

## 2.5. Concluding Remarks

In conclusion, a saving of computational load is achieved when the measures are calculated from encoded wavelet coefficients. The main challenges involving the evaluation of image coding are twofold: Firstly, an adequate selection of features is mandatory; secondly, perceptual measurement must achieve an accepted grade of correlation with human scoring and a certain grade of sensitivity. Let  $N$  be the number of transforms required to achieve desired representation. For the QI evaluation, test image is perceptually measured using measures in image-spectral domain. First of all, the inverse transform is used to reconstruct the image. Later, this image is Fourier-spectral transformed to be filtered by CSF, due to representation problems discussed in Section 2.4.1. After filtering, the outcome is Fourier-spectral inverse transformed in order to extract features required for distortion measurements, as in JPEG compression schemes. As result, the number of transforms is  $N = 4$ , which means that evaluation requires as minimal  $O(4 \times M_1 M_2 \log M_1 M_2)$  transform operations. On the other hand, for the FW-QI, an encoded representation is perceptually measured in wavelet domain. It requires only one scheme, because all considered processes are stated in the same domain, which means that the evaluation requires the number of  $O(M_1 M_2)$  transform operations. Therefore,  $O(M_1 M_2) < O(N M_1 M_2 \log M_1 M_2)$  for  $N > 1$ . In the following chapter, these improvements provide high correlation with observed scores, thus an increase in sensitivity of measurement. The proposed *Fully-Wavelet Quality Index* (FW-QI), which is an improvement of Quality Index evaluation, offers, as principal improvement, to avoid the use of additional schemes, using proposed feature measures in wavelet domain. Moreover, the use of perceptual measures related with features aims to reach a closer approximation with vision stimuli, as it is demonstrated by several studies on neural vision.

# 3. Full-Wavelet Quality Index Assessment for Image Coding

## 3.1. Preliminary

A critical aspect in image quality assessments is the adequate selection of features. The selection should lead to an improvement in evaluation. In this chapter, the proposed assessment is discussed, providing high correlation with observed scores, thus an increase of sensitivity of measurement. As shown in Figure 3-1, the *Full-Wavelet Quality Index* (FW-QI) assessment should be interpreted as the *wavelet transform* of the Quality Index (QI) one. For instance, the initial proposal on each wavelet-based measure  $d_{i,\psi}$ ,  $i = 1, 2, 3$  is to be the wavelet transform of its image-based measure counterpart  $d_{i,x}$ . Accordingly, FW-QI assessment includes the following steps: i) Source image is transformed into wavelet representation and then quantized, obtaining an encoded representation  $\mathbf{Y}_\psi = \mathcal{Q}\{\mathbf{C}\}$  susceptible to be transmitted. ii) Encoded representation differs from source, since elements of the former are truncated versions of the latter. This difference is the measured distortion on encoded  $\mathbf{Y}_\psi$  representation, using source representation as reference; iii) Representations  $\mathbf{C}$  and  $\mathbf{Y}_\psi$  are filtered using *Adaptive-masking* Wavelet-based Contrast Sensitivity Function (AM-WCSF block, described in Section §3.2.2) in order to enhance perceptual features; iv) The features are extracted from filtered elements in *Feature Extraction* block after which the perceptual measurement (Section §3.2), namely *Wavelet-based Quality Index* (WQI), is made based on operations on these features; v) As far as perceptual measured values are calculated, the *Linear Combination* block gathers them, giving a numerical result of quality evaluation.

Therefore, with these improvements, the FW-QI assessment is introduced with the selection of wavelet-based measures. This aims to give a numerical value corresponding to the assessment brought to the human vision model, represented by DMOS values. In order to achieve high correlation with human made scores, the contrast sensitivity masking is made adaptable to any wavelet structure, and weights in Equation (2-11) are recalculated by a regression strategy, selected from those described in Section §2.4.3. These improvements are tested by using coefficient of determination between assessment and DMOS, using the database of distorted images proposed in [68] for testing of image quality assessments.

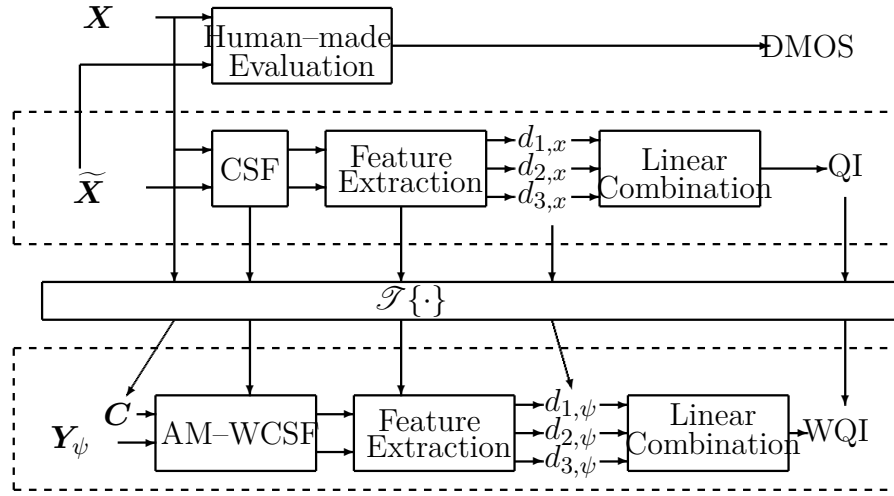


Figure 3-1.: DMOS, QI and FW-QI Assessments.

## 3.2. Proposed Image Quality Assessment

This section shows the realized improvements made on the *Full-Wavelet Quality Index* (FW-QI) assessment, which is a full-referenced assessment, since an original *reference* image is required together with the test one. Moreover, the FW-QI is *perceptual*, because it requires the features which give an analogy with human-based evaluations. An adequate selection of these features increases the correlation between FW-QI assessment and DMOS. It consists of the following steps, as shown in Figure 3-1: i) AM-WCSF described in Section §3.2.2, ii) Calculation of proposed wavelet-based *Blockiness*, *Edge Error* and *Visual Impairment* feature measures, respectively described in Sections §3.2.3, §3.2.4 and §3.2.5, and iii) Linear Combination of feature measures based on regression strategies depicted in Section §2.4.3. Accordingly, the usage of feature measures into wavelet-based representation is proposed, letting the extraction of perceptual features in the same scheme. First of all, QI of Equation (2-11) is generalized by comprising a polynomial function of normalized weighted summation of another perceptual measures, as follows:

$$QI = \lambda_0 + \lambda_1 d_1 + \lambda_2 d_2 + \lambda_3 d_3 \quad (3-1)$$

where  $\{d_i : i = 1, 2, 3\} \in \mathbb{R}^1(0, 1)$  is the set of selected feature measures and  $\{\lambda_i : i = 0, 1, 2, 3\}$  is the weighted coefficient set. The proposed assessment uses Equation (3-1) but feature measures are calculated in wavelet domain, by using features explained in Section §2.2.2. In order to reduce cluttered notation, for the  $i$ -th measure,  $d_{i,x} = d(\mathbf{X}, \tilde{\mathbf{X}})$  denotes which measure is calculated over image, and  $d_{i,\psi} = d(\mathbf{C}, \mathbf{Y}_\psi)$  denotes which measure is calculated over wavelet representation. Thus, features required for perceptual measurements are extracted completely from wavelet coefficients. Differently from QI proposed in [30], weighted coefficients in WQI are calculated using a regression strategy depicted in Section §2.4.3. Summarizing, as shown in Figure 3-1, FW-QI assessment requires the following steps: i)

Wavelet-based feature extraction, ii) Calculation of feature measures, and iii) Linear Combination of feature measures, according with following equation:

$$FW - QI = \lambda_0 + \lambda_1 d_{1,\psi} + \lambda_2 d_{2,\psi} + \lambda_3 d_{3,\psi} \quad (3-2)$$

### 3.2.1. Projection of Measures into Multi-resolution Domains

The projection of the feature measures into the same multi-resolution space is suggested for the linear combination of Equation (3-2). To give an example, the wavelet-based *Blockiness* can be interpreted as the wavelet transformation of its image-based counterpart,  $d_{1,\psi} = \mathcal{T}_\psi\{d_{1,x}\}$ , as shown in Figure 3-1. Prior to selection of wavelet domain, diverse choices are considered for different multi-resolution domains. For instance, wedgelet representation is proposed in [28] and Quad-tree Decomposition (QD) structure in [60]. Quad-tree Decomposition is particularly favored, thanks to its ability to find local homogeneities. Nonetheless, the selection of the wavelet transformation is made for various reasons, including: firstly, wavelet scheme gives less complexity than other representation; secondly, *Edge Error* feature measure cannot be modeled by Quad-tree Decomposition arguments [27]; thirdly, correlation values for measures in QD-domain are highly dependent on standard parameters of the model that represents each image, which can lead to erroneous measurements if the test image contains a high quantity of details. In any case, the QD-based attempt is detailed in Appendix A. To summarize, stating  $d_1$  as the Blockiness,  $d_2$  as the Edge Error, and  $d_3$  as the Visual Impairment measures, the notation  $d_{\circ,\circ}$  refers to measure which is projected in the corresponding space, e.g.,  $d_{2,\psi}$  is the expression for the *Edge Error* measure projected in wavelet space.

### 3.2.2. Proposed Adaptive-Mask Wavelet-based Contrast Sensitivity Function

WCSF described in Section 2.4.1 is calculated only for fixed-values of wavelet decomposition sub-bands given a fixed wavelet family, such as Daubechies 9/7. It indicates that the encoded representation must be rearranged to the wavelet basis used for the WCSF weights to fit with orthogonality constraints. However, for measurement purposes, the number of sub-bands (and even filters) varies in order to enhance the features required for measures. For instance, in Maximal Module feature case, there are a number of finer sub-bands, where this feature can be depicted accurately. Therefore, calculation of WCSF weights is required to be dependent on the number of sub-bands. An improvement of this method is proposed, termed *Adaptive Masking* (AM-WCSF) method, where the arguments of CSF filter in Equation (2-8) are adjusted in order to be located properly at frequency argument of its maximal value. This method aims to fit the frequency arguments of CSF with the wavelet scheme, because these frequency arguments are different to the scale arguments  $\mathbf{k}$  of the wavelet scheme. In other words, the generation of wavelet index  $\mathbf{k}$  comes from the number of sub-bands, which

is a value different from spectral frequency. To adjust the frequency scaling, the maximal value of spectral function is re-located into wavelet structure, making the mask the same for any number of decomposition levels or type of wavelet filter. This adjustment depends on the following parameters among others as discussed in [83], coming from perceptual evaluations: Let  $\zeta_r$  be the perceptual contrast parameter, which depends on characterized *eye-to-screen* distance  $\zeta_{view}$  and screen resolution  $\zeta_{res}$ :

$$\zeta_s = \frac{2}{0.0254} \zeta_{view} \zeta_{res} \tan\left(\frac{1}{2}\right). \quad (3-3)$$

Frequency scaling of spectral mask is *adjusted* at the sample frequency  $f_s = 2\zeta_s$ . Frequency arguments are adjusted by a factor  $\Delta_j$  calculated for each sub-band  $j$  as:

$$\Delta_j = f_s \frac{f_\psi}{2^{-j}},$$

where  $2^{-j}$  is the wavelet sub-band scaling argument for each  $j = 1, \dots, J$  and  $f_\psi$  is the argument of maximal value of module of Fourier Transform of wavelet basis  $\psi$ , which in the following is called the *Center Frequency of Wavelet* [2].

**Definition 3.2.1** *The Spectral Mask is defined as the square matrix  $\mathbf{X}_s \in \mathbb{R}^{\mathbf{Q} \times \mathbf{Q}}$  whose size  $\mathbf{Q} \in \mathbb{N}$  accomplishes with following conditions: first,  $\mathbf{Q} = \min\{q \in \mathbb{N} : q \geq \max\{M_1, M_2\}\}$ ; second, a number  $J_{\mathbf{Q}} \in \mathbb{N}$  must exist such that  $\mathbf{Q} = 2^{J_{\mathbf{Q}}}$ . The latter condition assures that the lowest frequency arguments are located into approximation sub-bands, regardless of the number of sub-bands used for wavelet scheme.*

**Definition 3.2.2** *The number  $J_{\mathbf{Q}}$  is named the Adjusted Number of Decomposition Levels. Thus, the spectral mask has size  $2^{J_{\mathbf{Q}}} \times 2^{J_{\mathbf{Q}}}$ . It is recommended that total number of sub-bands representing the image in wavelet space  $J$  are inside the interval  $[1, J_{\mathbf{Q}}] \in \mathbb{N}$ .*

Afterwards, the CSF weighted matrix is scaled to the desired wavelet structure just calculating the projection of frequency scaling on wavelet space, allowing to estimate the values of  $(u, v)$  required for Equation (2-10) to calculate the CSF using wavelet-based arguments. At this line, adjusted CSF becomes *Adapted* mask. Therefore, with  $u_0 \in \mathbf{u}_0 = \{1, 2, \dots, 2^{J_{\mathbf{Q}}}\}$ ,  $v_0 \in \mathbf{v}_0 = \{1, 2, \dots, 2^{J_{\mathbf{Q}}}\}$  as the correspondent vertical and horizontal grid drawn over wavelet structure, Thereafter, scaled arguments become  $u_s = u_0 * \Delta_j$  and  $v_s = v_0 * \Delta_j$ , while  $2^{j-1} \leq u_0, v_0 \leq 2^j$ . Thus,  $2^{-j}$  argument provides the required flexibility to adaptively arrange the WCSF mask independent of the number of decomposition levels, and the  $f_\psi$  argument allows the wavelet representation to be fitted to sensitivity mask, thus increasing correlation. Therefore, WCSF mask coefficient values do not need to be calculated any time when wavelet mother is changed, or any of its vanishing moments. It is a relevant aspect, because normally WCSF mask coefficients calculation requires a previous additional wavelet transform of the CSF. Instead, in the proposed method the Fourier transform of wavelet function is calculated. Thus, the AM-WCSF mask is assumed as the *scaled*

wavelet transform of CSF of Equation (2-10), because wavelet index  $\mathbf{k} = \{u_s, v_s\}$  is scaled to frequency argument  $\omega$  using following scale value:

$$\Delta_{u, \mathbf{Q}, j} = (M_u \cdot 2^{1-j}) - 1$$

$$\Delta_{v, \mathbf{Q}, j} = (M_v \cdot 2^{1-j}) - 1$$

Thus, for approximation coefficients, scaled values are:

$$\mathbf{u}_a = \left\{ 0, \frac{\Delta_{J_\psi}}{\Delta_{u, \mathbf{Q}, J_\psi}}, 2 \frac{\Delta_{J_\psi}}{\Delta_{u, \mathbf{Q}, J_\psi}}, \dots, M_u \frac{\Delta_{J_\psi}}{\Delta_{u, \mathbf{Q}, J_\psi}} \right\}$$

$$\mathbf{v}_a = \left\{ 0, \frac{\Delta_{J_\psi}}{\Delta_{v, \mathbf{Q}, J_\psi}}, 2 \frac{\Delta_{J_\psi}}{\Delta_{v, \mathbf{Q}, J_\psi}}, \dots, M_v \frac{\Delta_{J_\psi}}{\Delta_{v, \mathbf{Q}, J_\psi}} \right\}$$

Values are updated to  $\mathbf{u} = \mathbf{u}_a$  and  $\mathbf{v} = \mathbf{v}_a$ . Finally, for each sub-band, scaling values become:

$$\mathbf{u} = \left\{ \mathbf{u}, \frac{\Delta_j}{\Delta_{u, \mathbf{Q}, j}}, 2 \frac{\Delta_j}{\Delta_{u, \mathbf{Q}, j}}, \dots, M_u \frac{\Delta_j}{\Delta_{u, \mathbf{Q}, j}} \right\}$$

$$\mathbf{v} = \left\{ \mathbf{v}, \frac{\Delta_j}{\Delta_{v, \mathbf{Q}, j}}, 2 \frac{\Delta_j}{\Delta_{v, \mathbf{Q}, j}}, \dots, M_v \frac{\Delta_j}{\Delta_{v, \mathbf{Q}, j}} \right\}$$

The positions  $(u_c, v_c)$  are calculated where the argument of maximal value of CSF is located.

$$(u_c, v_c) = \arg \min |(\mathbf{u}, \mathbf{v}) - \arg \max_{\omega} S(\omega)|$$

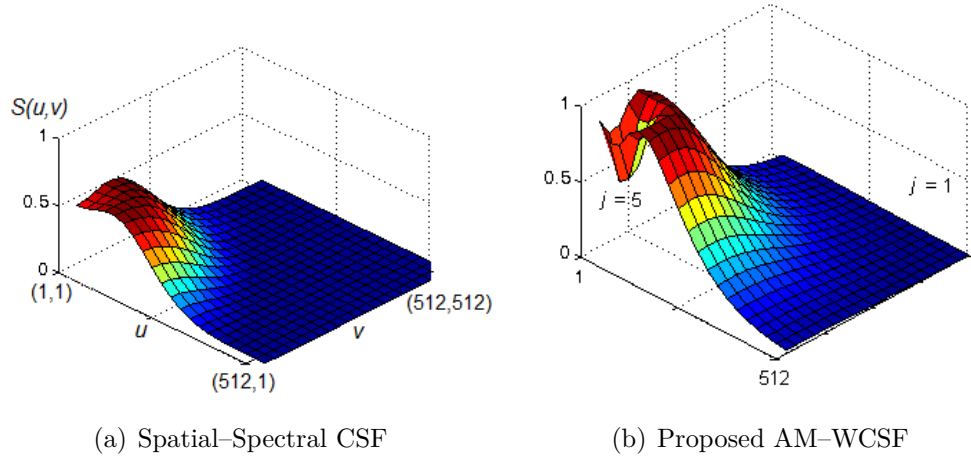
The wavelet frequency-scale grid is then the sequence:

$$(\Delta \mathbf{u}, \Delta \mathbf{v}) = \{-u_c, 1 - u_c, 2 - u_c \dots, M_u - 1 - u_c, -v_c, 1 - v_c, 2 - v_c \dots, M_v - 1 - v_c\}$$

Finally, the CSF weighted matrix  $\mathbf{S}$  is calculated using frequency grid in Equation (3-3), with normalized frequency axis  $\omega = \sqrt{(\Delta \mathbf{u})^2 + (\Delta \mathbf{v})^2}$ . Then, having the wavelet representation  $\mathbf{C}$ , filtered counterpart is calculated as:

$$\mathbf{C}' = \mathbf{S} \cdot \mathcal{T}_\psi\{\mathbf{X}\} \quad (3-4)$$

Figure 2-4 shows the following aspects: firstly, AM-WCSF method gives a close approximation of the conventional CSF, though this method calculates the mask in wavelet space; secondly, the contour of WCSF mask calculated by proposed method is distributed accordingly along the vertical, horizontal and diagonal detail wavelet coefficients. The AM-WCSF mask is contrasted against the WCSF mask calculated by fixed-value method in Figure 2.4(b), where its contour is centered over the diagonal detail coefficients. Figure 3-2 depicts the CSF and AM-CSF masks, where the maximal frequency argument of sensitivity function is located always in spatial-scale sub-bands demanded for human-vision model constrains. It permits the increase of correlation, since the shape similarity with theoretical CSF shown in Figure 3.2(a). Thereafter,  $\mathbf{C}'$  and  $\mathbf{Y}'_\psi$  filtered representations calculated in Equation (3-4) are susceptible to be perceptually measured. Following sections §3.2.3, §3.2.4 and §3.2.5 describe the calculation of respective wavelet-based *Blockiness*, *Edge Error* and *Visual Impairment* feature measures.



**Figure 3-2.:** Simulation of CSF masking by considered methods, over a considered spectra of  $512 \times 512$  pixel size. For the AM-WCSF method, 5 decomposition levels and biorthogonal 9/7 wavelet are used. Dark-red areas illustrate the maximal value of the CSF mask assigned to pixels

### 3.2.3. Proposed wavelet-based Blockiness Measure

Block degradation perceptually appears as strong transitions over the image. These transitions are generated by the strong quantization of elements in several blocks. As discussed in Section 2.2.2, a transition is represented by a cutting edge with homogeneous areas alongside. This transition can then be found when following conditions are met: when a straight edge is found, when elemental variance inside areas becomes low, and finally, when this variance is lower than the variance of elements along the edge. Therefore, required features are involved with the covariance of transition and local variance of each block in the vicinity of the transition, which can be calculated using Equation (2-6). The proposed measure supposes that the transitions are located in blocks at indistinct positions. However, due to the nature of transition, its directions must be horizontal and vertical. Thus, the direction of the cutting edge selected to detect the block degradation is defined by the angle of lifting-based wavelet filter  $\theta = HL, LH$ , where  $HL$  and  $LH$  correspond to the vertical and horizontal orientations, respectively. Defining  $\mathbf{D}_e = \mathbf{C} - \mathbf{Y}_\psi$  as the difference between source and encoded representations, a measurable distortion takes place when a transition exists in this difference  $\mathbf{D}_e$ . It prevents the detection of transitions that would otherwise come from the original image. Let  $C_\mu$  be the sub-set of Maximal Modulus wavelet coefficients. As shown in Figure 3.3(b), a transition is found when the following conditions hold:

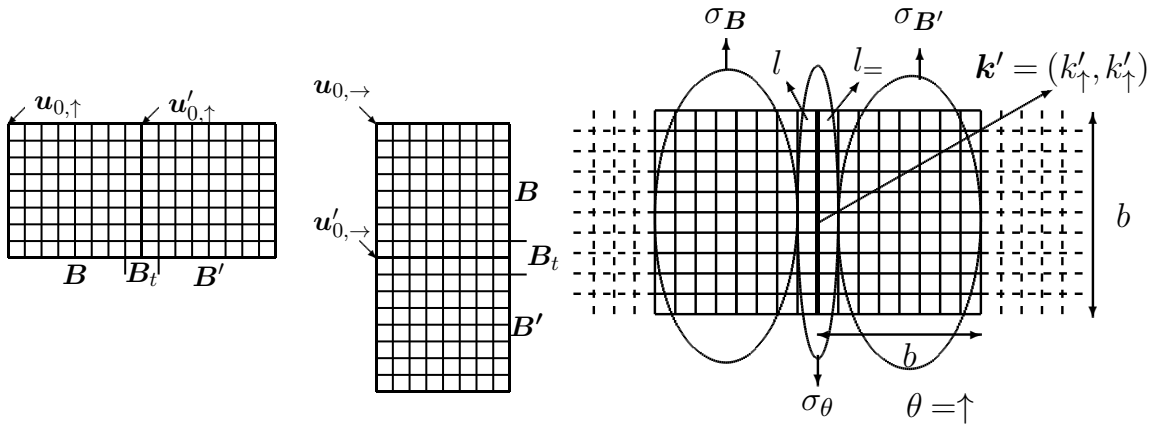
- **B1. Cutting edge condition.** Let  $\mathbf{k}' = (k'_{HL}, k'_{HL})$  be the spatial position of a wavelet coefficient detected as maximal module of Equation (2-5). Thus, a sequence

of  $b$  Maximal Modulus elements expressed by a set

$$\{c_\mu(\mathbf{k}')\}_{\mathbf{k}'=(k'_{HL},k'_{LH})}^{(k'_{HL}+b,k'_{LH})} \in \mathbf{Y}_{\psi,\mu},$$

must draw a line  $l$  of direction  $\theta = HL, LH$  and length  $b$ , where  $b$  is an arbitrary length. Then, this line meets with the condition of *Cutting Edge*.

- **B2. Local variance threshold condition.** Once cutting edge is detected, two blocks  $\mathbf{B}, \mathbf{B}'$  of size  $b \times b$  each, are located adjacent to line  $l$  at orientation  $\theta$ . So, standard deviation of each block  $\sigma_{\mathbf{B}}, \sigma_{\mathbf{B}'}$  and cutting edge  $\sigma_\theta$  are calculated. Thus, restrictions  $\sigma_\theta > \sigma_{\mathbf{B}}$  and  $\sigma_\theta > \sigma_{\mathbf{B}'}$  must be met.
- **B3. Mean Squared Error condition.** Along the co-linear direction of line  $l$ , another line  $l_=\mathbf{}$  is drawn. Mean Squared Error between  $\mathbf{D}_e$  elements on  $l$  and  $l_=\mathbf{}$  lines is calculated. Thus,  $\text{MSE}_{HL}$  and  $\text{MSE}_{LH}$  are the respective calculations for vertical and horizontal directions. Once again, the condition  $\text{MSE}_\theta > \sigma_\theta$  must be met. Figure 3-3 shows the comparison between conventional calculation of transitions in image domain and its proposed counterpart.



(a) Image domain, regular-sized blocks [30] (b) Wavelet domain, transition selected from Maximal Module elements

**Figure 3-3.:** Illustration of transitions considered for extraction of features required for Wavelet-based *Blockiness* measure

A blockiness calculation is stated for features satisfying statements **B1**, **B2** and **B3**, after which the measure is calculated as follows: Let  $n_{HL}$  and  $n_{LH}$  be the number of detected block transition features along the respective horizontal and vertical directions. The error

component for each direction is:

$$e_{HL} = \sum_{i=1}^{n_{HL}} \left( \frac{MSE_{HL}}{MSE_{HL} + \sigma_{B,HL} + \sigma_{B',HL}} \right)^2 \quad (3-5a)$$

$$e_{LH} = \sum_{i=1}^{n_{LH}} \left( \frac{MSE_{LH}}{MSE_{LH} + \sigma_{B,LH} + \sigma_{B',LH}} \right)^2 \quad (3-5b)$$

Using Equations (3-5a) and (3-5b), the blockiness measure is calculated as:

$$d_{1,\psi} = \sqrt{\frac{e_{HL} + e_{LH}}{n_{HL} + n_{LH}}}. \quad (3-6)$$

### 3.2.4. Proposed Wavelet-based Edge Error Measure

Edge detection is one of the classic wavelet applications. In the edge detection application of [49, Section 6.3], the detection criteria is directly related with the scale resolution. A high resolution implies a low scale, resulting in the detection of noise as sharp edges or the fragmentation of detected edges. In turn, a low resolution implies a high scale, making sharp edges undetectable. In order to address this issue, the wavelet-based detectors are adjusted by scale, in contrast with the classical edge detectors [40]. Then, principal components of each edge are localized in finer wavelet sub-bands. In that way, the non-relevant features and noise can be located and excluded. The fact that wavelet family  $\psi_{j,\theta}$  has vanishing moments in horizontal, vertical and diagonal orientations  $\theta$ , allows for viewing  $\psi_{j,\theta}$  as partial derivative of another finitely supported function  $\Theta_{j,\theta}$  [49]. Thus,  $\psi_{j,\theta} = \partial\Theta_{j,\theta}/\partial\theta$ , where  $\theta = HL, LH$ . Hence,

$$c_\theta = c_d(j, \mathbf{k}, \theta) = \frac{\partial}{\partial\theta} (x * \Theta^*(j, \mathbf{k}, \theta)), \quad (3-7)$$

where  $\Theta^*(j, \mathbf{k}, \theta) = \Theta(j, -\mathbf{k}, \theta)$ . The relationship (3-7) reveals that wavelet coefficients are partial derivatives of a smoothed version of the image, and it can be therefore interpreted as a first step in edge detection. A procedure is proposed in [40] in order to increase detector sensitivity, by using an adequate selection of wavelet filter and/or a thresholding technique for the selected maximal modules. Concretely, thresholding assures that only the relevant maximal modules are selected, e.g., a criteria based on Lipschitz exponent for relevant edge detection [40]. The change of the edge detector operator is made, and the measure proposed in [30] can be calculated from the maximal module as edge magnitude and from  $\angle d_j(\mathbf{k})$  as gradient direction. From [30], features required for *Edge Error* measure are edge strength and phase of source image, distorted image and error image  $\mathbf{X}_e = \mathbf{X} - \tilde{\mathbf{X}}$ . Accordingly, equivalent features are found for source representation, encoded representation and error representation  $\mathbf{C}_e = \mathbf{C} - \mathbf{Y}_\psi$ . Hence, wavelet-based *Edge Error* measure can be found directly using a wavelet-based edge detector. Let  $\theta(j, \mathbf{k})$  and  $\tilde{\theta}(j, \mathbf{k})$  be the respectively direction of edges located in source and encoded representations at index  $\{j, \mathbf{k}\}$ , and  $|c_\#|(\mathbf{k}, j)$  and  $|\tilde{c}_\#|(\mathbf{k}, j)$  be

the respectively strengths of edges located in source and encoded representations, for each position  $(\mathbf{k}, j)$ . Using  $\Delta\theta(\mathbf{k}, j) = |\theta(\mathbf{k}, j) - \tilde{\theta}(\mathbf{k}, j)|$ , and  $\Delta|c|(\mathbf{k}, j) = ||c_{\#}|(\mathbf{k}, j) - |\tilde{c}_{\#}|(\mathbf{k}, j)|$ , the total edge error function is defined as:

$$F_e(\mathbf{k}, j) = \frac{\Delta\theta(\mathbf{k}, j)}{\pi/2} + \frac{\Delta|c|(\mathbf{k}, j)}{|c_{\#}|(\mathbf{k}, j) + |\tilde{c}_{\#}|(\mathbf{k}, j)}. \quad (3-8)$$

Using Equation (3-8), the edge error measure becomes:

$$d_{2,\psi} = \sqrt{\frac{1}{n_c} \sum_{j=1}^J \sum_{\mathbf{k}} c_{\#}(\mathbf{k}, j) \cdot F_e(\mathbf{k}, j)}, \quad (3-9)$$

where  $n_c$  is the number of wavelet coefficients.

### 3.2.5. Proposed Wavelet-based Visual Impairment Measure

The Visual Impairment measure detects changes in visual orientations and spatial frequencies, as modeled for human visual cortex [52, 54]. Every change is statistically characterized by searching for local homogeneities [30]. The first proposed approximation resorts to the Quad-tree Decomposition (QD) approximation, which calculates local variance of inhomogeneous blocks. As mentioned earlier, since blocks  $\mathbf{B}$  are calculated from the QD partitions, they allow identification of the relevant visual impairments, without the use of regular blocks over the whole image. Considerations regarding this treatment are in Section A.3. Relevant visual impairments are identified in most cases on edges and non-homogeneous regions. In addition, the various block (co-)variances are conveniently computed at wavelet coefficient level, using Equation (2-6), and the mean value using Equation (2-7). Hence, the following features are used for calculation of wavelet-based Visual Impairment measure  $d_{3,\psi}$ , defined as follows:

$$d_{3,\psi} = \sqrt{\frac{1}{b} \sum_{\mathbf{k}} \varepsilon(\mathbf{k})^2}, \quad (3-10)$$

where  $\varepsilon(\mathbf{k})$  is calculated as:

$$\varepsilon(\mathbf{k}) = \frac{d_{sim}(\mathbf{k}) + e_{lum}(\mathbf{k})}{2}.$$

The luminance error  $e_{lum}$  and similarity distortion  $d_{sim,\psi}$  features are calculated in the following Equations (3-11a) and (3-11b):

$$e_{lum,\psi}(\mathbf{k}) = \frac{|\mu_{\mathbf{B}}(\mathbf{k}) - \mu_{\tilde{\mathbf{B}}}(\mathbf{k})|}{2^J} \quad (3-11a)$$

$$d_{sim,\psi}(\mathbf{k}, J) = \frac{\sigma_{c_e, c_e}}{\sigma_{c,c} + \sigma_{\tilde{c}, \tilde{c}}} \cdot \frac{1 - \rho(\mathbf{k})}{2} \quad (3-11b)$$

where  $\mathbf{B} \in \mathbf{C}$ ,  $\widetilde{\mathbf{B}} \in \mathbf{Y}_\psi$ , and

$$\rho(\mathbf{k}) = \frac{\sigma_{c,\tilde{c}}}{\sqrt{\sigma_{c,c}\sigma_{\tilde{c},\tilde{c}}}}.$$

The wavelet-based estimation of visual impairment holds the advantage of using blocks from the dyadic structure. By contrast, this measure in image domain requires the additional process of block calculation on the reconstructed image. In other words, visual impairment translates directly to an easily computed wavelet-based quantity, if one employs Haar wavelets. For other wavelets, the derived formulation for mean values and block variances can still be evaluated, but with results that are not directly interpretable as mean values, variances etc. Still, the resulting quality index can be expected to be meaningful.

### 3.3. Concluding Remarks

In this chapter, the *Full Wavelet-Quality Index* (FW-QI) assessment has been formally introduced. The outcome of this assessment is the quality level of a distorted image. This assessment is put forward in order to achieve the following goals: firstly, to increase the correlation between these outcomes and human-based score data; secondly, to implement a control strategy inside coding schemes. To achieve the former goal, the FW-QI assessment is composed of a linear combination of three metrics: *Blockiness*, *Edge Error* and *Visual Impairment*. To achieve the latter goal, feature metrics are calculated directly from the encoded coefficients, avoiding the measurement on the reconstructed image. In the next Chapter, an experimental test is arranged to give the adjustments required for the FW-QI parameters.

# 4. Full–Wavelet Quality Index Adjustments

## 4.1. Preliminary

The *Coefficient of Determination* is used to verify the grade of correlation of an assessment. A coefficient of determination indicates how close the values given by test regression strategy fit with *Differential Mean Observed Score* (DMOS). The aim of this experiment is to determine, both for feature measure and for the assessment, the correlation level needed in order to approach the human vision model.

**Definition 4.1.1** *Correlation is the ability of a measurement to match the actual value of the quantity being measured. Sensitivity is the quality or condition of measurement of being responsive to external conditions or stimulation.*

A fully–referenced multi–channel perceptual assessment is composed of three procedures: contrast sensitivity filtering, perceptual measures and linear combination. An adjustment of parameters of assessment leads to a variation of the correlation. The following adjustments are given for proposed Fully–Wavelet Quality Index (FW–QI) assessment: i) evaluation of feature measures through correlation; ii) correlation of regression strategies, that is the correlation between the assessment and DMOS, for a considered regression strategy; iii) sensibility of assessment, by calculation of variability of assessment using considered regression strategies; iv) Evaluation of different CSFs using correlation of assessments.

This chapter is divided as follows: Human Visual Parameters for the CSF filter and Images Database are enumerated in Section §4.2. The measures that are used to determine the correlation and sensitivity grade of assessment with respect to human evaluation are explained in Section §4.3. A selective evaluation of different CSF filters, measures and regression strategies through correlation is illustrated in Section §4.1. The above lead to selecting the adequate parameters for the proposed assessment.

## 4.2. Image Database and Human Visual System Parameters

A certain number of image databases with human evaluations have been found in literature to be used for providing a validation of a perceptual assessments with respect to a human-based evaluation. The correlation and sensitivity in measurement for this perceptual evaluation are two parameters to consider for testing assessment. In order to become a remarked evaluation, a huge number of human-made scores for different distortions types and levels are located in LIVE Image Quality Assessment Database [14], Release 2 – 2005. The database contains a collage of reference and distorted images, where the latter are ranked by the following type of distortion:

- 174 images, degraded by Fast Fading distortion, labeled as *fastfad*.
- 174 images, degraded by Blurring distortion, labeled as *gblur*.
- 227 images, degraded by JPEG2000 compression, labeled as *j2k*.
- 233 images, degraded by jpeg compression, labeled as *jpeg*.
- 174 images, degraded by White Gaussian Noise, labeled as *wn*.

In [68], the scores of a subjective rating are presented, particularly the DMOS, for each image in the cited database. Other distortions considered for evaluation are Fast Fading distortion, JPEG2000 compression, and White Gaussian noise addition. For this experiment, all images of the LIVE database are used, widening the range of considered images tested in [30] where only *j2k* and *jpeg* databases were used. For this setup, 50% of the considered database is used as training (estimation of parameter vector), and the rest is used as validation. Finally, the following parameters are suggested in [42] for AM-WCSF masking: CSF central value of  $\arg \max_{\omega} \mathbf{S} = 3.75$  cycles/pixel, and perceptual values of  $\zeta_{view} = 0.3$  meters and  $\zeta_{res} = 75$  pixels/inch.

## 4.3. Correlation and Sensitivity Measurements

The methodology used to verify the performance of assessment is based on the *Coefficient of Determination*.

**Definition 4.3.1** *The Coefficient of Determination  $R^2$  is a statistical measure which indicates the correlation grade for an estimated set of data in relation with the real data points.*

Therefore, for each input trial  $i$  in LIVE database [14], let  $a(i)$  be a value point of the set of data measured on experiment, and  $b(i)$  a value point from a observed model or real data.

The coefficient of determination becomes:

$$R_{a,b}^2 = 1 - \frac{\sum_i (b(i) - a(i))^2}{\sum_i (b(i) - \mu_b)^2}, \quad (4-1)$$

where  $\mu_b$  is the mean value of the observed values. A detailed evaluation methodology is described in [68], where a coefficient of determination indicates how close the values given by test assessment fit with DMOSs. The aim of this experiment is to determine, for both a feature measure and the assessment, the correlation level to achieve the human vision model. The methodology for calculating the coefficient of determination is described as follows:

- Firstly, the 70% of DMOS values are used for training and the other 30% for validation, as used in [29]. Then, the linear regression is stated as  $DMOSP = \lambda_0 + \lambda_1 d_1 + \lambda_2 d_2 + \lambda_3 d_3$ , where  $\{\lambda_i : i = 0, 1, 2, 3\} \in \mathbb{R}^1$  are the multi-variate linear regression coefficients as explained earlier. Coefficient  $\lambda_0$  can be neglected, because scores are differential.
- Secondly, the correlation index between DMOS and assessment values of validation set is calculated for each type of distortion. After that, the correlation index for the validation set of the whole database is calculated. The training set samples are permuted again and the correlation coefficient is calculated 10 times. The outcome is a sequence of 10 samples of correlation for each type of distortion, another sequence of 10 samples of correlation for the whole database, and the mean and standard deviation of the 10 trials.
- Finally, the following two parameters are analyzed for evaluation of different coding methods: first, the correlation level, which answers numerically if evaluation fits with human vision observations; later, the sensitivity in measurement, which indicates how variable the proposed evaluation can behave against images out of training database.

### 4.3.1. Correlation between Feature Measures and Differential Scores

In this case, the coefficient of determination is calculated between each feature measure and DMOS. Its aim is to evaluate whether a feature measure can achieve an acceptable correlation level to be included in FW-QI assessment, which would determine whether the quality of image can be measured by proposed wavelet-based feature measures. An optimal result could lead to a possible reduction of the less correlated measure at distortion measuring stage. It must be quoted that all measures in question are evaluated with images that pass through CSF filter.

Coefficient of determination between assessment values and DMOS is calculated for different samples of validation data, in order to give the correlation level that each CSF calculation method can produce to reach the human vision model, and the level of generalization that the regression strategy can offer. In other words, the higher the coefficient of determination the

more accurate the measurement. A description of the procedures involved in the calculation of this type of correlation is depicted as follows:

- Source and reconstructed images (or source and encoded representations where it applies) are filtered by selected CSF. When source and reconstructed images are used, the CSF filter of Equation (2-10) is used, giving filtered  $\mathbf{X}'$  and  $\widetilde{\mathbf{X}}'$  counterparts. On the other hand, when source and encoded representations are used in measurement, they are filtered by a WCSF, giving filtered  $\mathbf{C}'$  and  $\widetilde{\mathbf{C}}'$  counterparts.
- After filtering, the features are extracted from filtered images or representations. For the case of the image-based measures  $\{d_{1,x}, d_{2,x}, d_{3,x}\}_i$ , filtered images are used. For the case of the wavelet-based measures  $\{d_{1,\psi}, d_{2,\psi}, d_{3,\psi}\}_i$  (Equations (3-6), (3-9) and (3-10)), filtered representations are selected to feature extraction and measurement. As a result, the number of operations caused by transforms is  $O(3 \times M_1 M_2 \log M_1 M_2)$ .

Thus, each test image  $\widetilde{\mathbf{X}}_i$  is related with following elements: a source image  $\mathbf{X}$ , a human-made Differential Mean Observed Score  $DMOS_i$ , and a vector of quality values  $\{d_{1,\psi}, d_{2,\psi}, d_{3,\psi}\}_i$ . Then, a *trial set*  $\{d_1(i), d_2(i), d_3(i), DMOS(i)\}$  is composed of measured distortion values  $d(i)$  on each  $i$ -th image, followed by the DMOS value. This set feeds the coefficient of determination equation  $R^2$  of Equation (4-1) as follows:

$$R_{d,DMOS}^2 = 1 - \frac{\sum_i (DMOS(i) - d(i))^2}{\sum_i (DMOS(i) - \mu_{DMOS})^2}, \quad (4-2)$$

where  $\mu_{DMOS}$  denotes, in this instance, the mean value of the observed DMOS values from the database. The training set samples are permuted again and the coefficient of determination is calculated 10 times. The outcome is a sequence of 10 samples of coefficients of determination, and the mean and standard deviation of those 10 trials. After that, the coefficient of determination is calculated for each permuted validation set, resulting in another sequence of 10 samples of coefficients, and the mean and standard deviation for the respective trials. However, results show that variability is negligible for coefficients of determination, thus only mean results are shown.

### 4.3.2. Correlation between Assessments and Human Vision Model

This test determines whether the correlation increases when a linear combination of measures is used in assessment rather than a single measure. The considered regression strategies are used to estimate the weighting coefficients. In this case, the strategy which produces better results of coefficient of determination leads to the best strategy to be used for FW-QI assessment. The term  $pDMOS$  is used to name the resulting output quality value, regardless of the assessment used. Thus, for each considered regression strategy used within assessment, 50% of the considered database is used as training (estimation of parameter

vector), and the rest is used as validation. The training set samples are permuted again and the coefficient of determination is calculated 10 times. The outcome is a sequence of 10 samples of coefficients of determination, and the mean and standard deviation of those 10 trials. After that, coefficient of determination is calculated for each permuted validation set, resulting in another sequence of 10 samples of coefficients, and the mean and standard deviation for the respective trials. However, results shown that variability is negligible for coefficients of determination, thus only mean results are shown. In this case, LR, PCAR and ANNR strategies of Section §2.4.3 are used to calculate regression coefficients. For each strategy, a collection of distortion vectors and their respective DMOS values is used as training input of regression process, giving as result the regression coefficient set  $\{\lambda_1, \lambda_2, \lambda_3\}_\eta$  for a trial  $\eta$ . A number of 10 trials is made, then mean value of those regression coefficients is calculated, which is finally used in assessment, giving the  $pDMOS$  values. Thus, the coefficient of determination is calculated as:

$$R_{DMOS,pDMOS}^2 = 1 - \frac{\sum_i (DMOS(i) - pDMOS(i))^2}{\sum_i (DMOS(i) - \mu_{DMOS})^2}, \quad (4-3)$$

where  $\mu_{DMOS}$  denotes, for this instance, the mean value of the observed DMOS values from the database. Each regression strategy calculates the parameter vector  $\boldsymbol{\lambda} = \{\lambda_i\}$  with feature measures as input variables and the assessment  $pDMOS$  as output variable.

### 4.3.3. Assessment Sensitivity

This test determines which considered regression strategy is the more sensible for FW-QI assessment. The variability of the regression coefficients determines the grade of sensitivity of such strategy. For instance, highly variable regression coefficients indicate that strategy cannot be generalized to other databases, leading the strategy to be considered inadequate. The evaluation of the regression strategy is achieved with the analysis of variability of the weighting vector calculated by the regression strategy. In this case, the feature measures are the input variables. Thus, the adequate regression strategy for calculation of QI is determined by the the lowest value of the variability in the parameter vector, and this variability determines the sensitivity, at least for the images of the considered database.

Input samples for calculation of parameter vector are the set of feature measures calculated for each image in the validation set. The training set samples are permuted again and the parameter vector is calculated 10 times. The outcome is a sequence of 10 samples of parameter vectors, and the mean and standard deviation of those 10 trials. In the same manner, a sequence of 10 samples of eigen-vector matrices  $\mathbf{V}$  calculated by the PCAR strategy, and the mean and standard deviation of those trials are calculated. For each considered regression strategy, the variability is measured by the standard deviation of the sequence of weighting vectors. This variability illustrates how CSF calculation method can show dramatical variation in the performance of the regression strategy, and in turn, whether the regression

strategy is adequate for the proposed assessment. Assuming that  $\sigma_{\lambda_i}$  is the standard deviation calculated for  $i$ -th element of parameter vector  $\{\lambda_i\}$ , the following equation is used to determine the grade of sensitivity in measurement:

$$\bar{\sigma}_\lambda = \frac{1}{N} \sum_{i=0}^{N-1} \sigma_{\lambda_i}, \quad (4-4)$$

where  $N = 4$  is the number of elements of the weighted vector. Thus, a lower value of  $\bar{\sigma}_\lambda$  denotes a more sensitive regression strategy.

#### 4.3.4. Evaluation of Contrast Sensitivity Functions.

This test determines how much a CSF method changes the grade of correlation on assessments. Correlation is tested for CSF, WCSF and AM-WCSF filters, providing that each one is considered only in the representation which are analyzed. For instance, CSF is only analyzed for image-based feature measures, and AM-WCSF for wavelet-based feature measures. Again, a description of the procedures and the transforms involved in this type of evaluation is depicted as follows:

- Source and encoded representations are filtered using either CSF, WCSF, or AM-WCSF. This gives the filtered images and when CSF is used, and filtered representations when either WCSF or AM-WCSF is used.
- Afterwards, *Blockiness*, *Edge Error* and *Visual Impairment* feature measures are calculated using filtered images, when CSF is used, giving  $d_{1,x}, d_{2,x}$  and  $d_{3,x}$  measures required for QI of Equation (3-1). Likewise, measures  $d_{1,\psi}, d_{2,\psi}$  and  $d_{3,\psi}$  required for FW-QI assessment of Equation (3-2) are calculated only when WCSF or AM-WCSF are used. The number of operations is approximately  $O(4 \times M_1 M_2 \log M_1 M_2)$ .

Thus, for each test distorted image  $\tilde{\mathbf{X}}_i$  in the database in question, there are a source image  $\mathbf{X}_i$ , a DMOS $_i$ , and a measure vector  $\{d_{1,x}, d_{2,x}, d_{3,x}\}_i$ . Again, LR, PCAR and ANNR strategies of Section 2.4.3 are used to calculate regression coefficients. Therefore, for each strategy, a collection of distortion vectors and their respective DMOS values are used as training input of regression process, giving as a result the regression coefficient set  $\{\lambda_1, \lambda_2, \lambda_3\}_\eta$  for a trial  $\eta$ . A number of trials are conducted, then the mean value of those regression coefficients is calculated. Thus, a positive value leads to an improvement of correlation due to the selected contrast filter.

## 4.4. Results

### 4.4.1. Correlation between Feature Measures and DMOS

For this aspect, the comparison is made by calculating the coefficient of determination between each feature measure and DMOS, using as input values those that come from mea-

asuring validation image database. The measures  $d_{1,x}$ ,  $d_{2,x}$  and  $d_{3,x}$  are those proposed in [30], where conventional CSF and image-domain feature measures are used. Coefficient of determination results for these values are located at three first rows of Table 4-1: 0.74 for *Blockiness*, 0.89 for *Edge Error* and 0.73 for *Visual Impairment*. It means that those values are used for comparative purposes. Therefore, proposed wavelet-based metrics are expected to achieve close values of correlation. As a result, it is noted that the proposed feature measures fit with human vision model at comparable levels. For instance, using AM-CSF filter, the proposed wavelet-based *Blockiness* achieves a increase of 8% in comparison with its image-based counterpart; wavelet-based *Visual Impairment* improves the correlation level by 9%. However, wavelet-based *Edge Error* scores worse, with a decrease of 12%.

**Table 4-1.:** Mean coefficient of determination between feature measures and DMOS values.

Applied Sensitivity Filter	Feature Metric	$R^2$
CSF	$d_{1,x}$	0.74
	$d_{2,x}$	0.89
	$d_{3,x}$	0.73
WCSF	$d_{1,\psi}$	0.80
	$d_{2,\psi}$	0.75
	$d_{3,\psi}$	0.85
AM-WCSF	$d_{1,\psi}$	0.82
	$d_{2,\psi}$	0.76
	$d_{3,\psi}$	0.84

#### 4.4.2. Correlation between Assessment and Human Vision Model

The mean coefficient of determination (Equation (4-2)) between feature measures used for different assessments and DMOS values is calculated for different contrast sensitivity filters. The performance results of regression strategy are achieved by calculating the percentage of coefficient of determination (Equation 4-3) difference. The correlation value used as a reference is the QI, which is obtained from CSF filtering, image-based feature measures and LR strategy. Therefore, the determination coefficient value of 0.86 in Table 4-2 must at least be met, ideally improved on. The following highlights are observed:

- Firstly, the correlation using either PCAR or ANNR is increased. For instance, a evaluation that uses image-based feature measures and PCAR strategy gives an increase

of 7%, and an evaluation that uses image-based feature measure and ANNR strategy gives an increase of 6%. This indicates that an adequate regression strategy is critical in obtaining a high correlation with human-based model evaluations.

- Secondly, objectives of this work are achieved when WCSF filter is used in FW-QI assessment (See highlighted values in Table 4-2). It is thanks to Daubechies 9/7 function, used both in coding and for WCSF filter. These results confirm the fact that the influence of sub-optimal wavelet-based Edge detector plays a fundamental role in correlation. Though best results are achieved when AM-WCSF filter and image-based feature measures are used, improvement in coefficient of determination are also achieved when WCSF filter is used in an assessment that uses wavelet-based feature measures, as in the case of proposed assessment. With regards to AM-WCSF, the coefficient of determination of assessment is maintained when using the wavelet-based feature measures and LR strategy.

To sum up, the proposed assessment achieves comparable results about correlation with human-based evaluations, using PCAR strategy and AM-CSF filter.

**Table 4-2.:** Mean coefficient of determination of QI and FW-QI assessments.

Assessment	Filter	Regression Strategy	$R_{DMOS,pDMOS}^2$
IQ	CSF	LR	0.86
		PCAR	<b>0.93</b>
		ANNR	0.92
FW-IQ	WCSF	LR	<b>0.87</b>
		PCAR	<b>0.88</b>
		ANNR	<b>0.89</b>
	AM-WCSF	LR	0.86
		PCAR	0.86
		ANNR	0.88

### 4.4.3. Assessment Sensitivity

Results on sensitivity consists in describing the magnitude of standard deviation of calculated regression coefficients for all 10 trials. As for results depicted above, the aim is to achieve,

or ideally improve the reference value. In this case, the result of QI values using CSF filter and LR strategy (0.6822, CSF-LR row, QI column in Table 4-3) is used as reference. As described in Figure 4-3, this leads to the following remarks:

- Firstly, PCAR appears as the more stable strategy, where variations are stated at the third decimal cypher, whereas LR show variations at the first and second decimal cypher. Results of the ANNR strategy are not shown, as this strategy is proved to be unstable because variations at the integer and first decimal digit are found.
- Secondly, FW-QI assessment is more stable when weighting coefficients are calculated by PCAR.
- Lastly, the most stable measurement is achieved with a value of  $5 \times 10^{-3}$  of averaged standard deviation, using AM-WCSF filter, wavelet feature measures and PCAR strategy, thus achieving the specific goal. It corresponds to FW-QI assessment, with selected features obtaining adequate results on correlation.

Derived from this result, using PCAR strategy, Equation (3-1) obtains the following regression parameters:  $\lambda_0 = -0.01$ ,  $\lambda_1 = 0.07$ ,  $\lambda_2 = 3.19$ , and  $\lambda_3 = 0.73$  (see [27, pg.4] for more details). As additional remark, a particular analysis on eigen-vector matrix  $\mathbf{V}$  used for

**Table 4-3.:** Averaged standard deviations of regression coefficients  $\bar{\sigma}_\lambda$ (Equation (4-4)), which are used for calculation of QI and WQI assessments by LR and PCAR strategies.

Filter	Regression Strategy	Assessment	
		QI	FW-QI
CSF	LR	0.6822	<b>0.1493</b>
	PCAR	0.0370	<b>0.0172</b>
WCSF	LR	0.6721	<b>0.0568</b>
	PCAR	0.0238	<b>0.0082</b>
AM-WCSF	LR	0.7738	<b>0.0666</b>
	PCAR	0.0875	<b>0.0050</b>

PCAR strategy is shown in Table 4-4, which aims to illustrate the level of variation of this matrix when contrast filter or measure is changed. For this results, mean values of the 10 eigen-vector matrix trials are calculated for each feature measure and each contrast filter. In this case, the standard deviation of  $\mathbf{V}$  is suppressed as the values are lower than the

fourth decimal digit. Again, a reference is used. In this case, eigen-vector matrix calculated for the QI that uses CSF filter and image-based feature measures is used as reference (CSF rows,  $d_{1,x}, d_{2,x}, d_{3,x}$  columns, producing a  $3 \times 3$  eigen-vector matrix). Let  $\mathbf{V}_{ref}$  and  $\mathbf{V}_i$  be the respective reference and test eigen-vector matrices, assuming that the test assessment is using PCAR strategy for any CSF filter. The perturbation value  $d_v = \|\mathbf{V}_{ref} - \mathbf{V}_i\|$  is used to determine the level of change of eigen-vector matrix, when filter or measure is changed. AM-WCSF filter depicts the minimal perturbation  $d_v$ , and is considered an acceptable filter for increasing feature characteristics and for allowing PCAR strategy to increase the correlation between, either FW-QI or QI assessments, and DMOS. As well as this, AM-WCSF filter preserves eigen-vector matrix, giving a more stable measurement

**Table 4-4.:** Mean eigen-vector matrix  $\mathbf{V}$  calculated for considered measures at the validation image database. Each column of each matrix corresponds to eigen-vectors calculated for each feature measure.

Filter	Feature measure					
	$d_{1,x}$	$d_{2,x}$	$d_{3,x}$	$d_{1,\psi}$	$d_{2,\psi}$	$d_{3,\psi}$
CSF	0.5904	0.3525	0.7260	0.6337	0.3241	0.7023
	0.5543	-0.8309	-0.0473	0.4374	-0.8989	0.0229
	0.5866	0.4303	-0.6859	0.6380	0.2944	-0.7114
	$d_v = 0.00$			$d_v = 0.22$		
WCSF	0.5844	0.3911	0.7108	0.5874	0.4291	0.6859
	0.5634	-0.8259	-0.0153	0.5524	-0.8319	0.0477
	0.5840	0.4056	-0.7030	0.5914	0.3509	-0.7257
	$d_v = 0.06$			$d_v = 0.16$		
AM-WCSF	0.5912	0.3663	0.7185	0.5857	0.3632	0.7238
	0.5513	-0.8337	-0.0286	0.5627	-0.8248	-0.0430
	0.5887	0.4130	-0.6948	0.5834	0.4311	-0.6875
	$d_v = \mathbf{0.03}$			$d_v = \mathbf{0.02}$		

#### 4.4.4. Evaluation on Contrast Sensitivity Function

The following results demonstrate the performance of the CSF filters and their contribution to the correlation with human vision model. As shown in Table 4-5, the following is

depicted for image-based feature measures: firstly, the usage of WCSF improves the correlation between image-domain based *Blockiness* and DMOS by 4% and the correlation between image-domain *Visual Impairment* and DMOS by 7%, but this usage decreases the correlation between image-domain *Edge Error* and DMOS by 2%; secondly, the usage of AM-WCSF maintain the correlation between image-domain based *Edge Error* and DMOS, but it decreases the correlation between *Blockiness* and DMOS by 5% and the correlation between image-domain *Visual Impairment* and DMOS by 4%. This suggests that the usage of CSF filters can be stated in wavelet-domain, providing the forward and the inverse functions of the transform use a filter with the same properties. This problem is noted for the proposed AM-WCSF, because the weighting function is equivalent to using a mother wavelet, which has different properties to the wavelet function used in source representation, making the reconstruction susceptible to degrading details for image-based extraction. Therefore, AM-WCSF uses a non-inversive representation. However, the increase of correlation to DMOS model by WCSF filtering can be justified by the fact that the masking coefficients are specifically calculated for the related representation. Thus, for another wavelet family, another masking coefficient set must be calculated off-line. In contrast, the proposed AM-WCSF does not require this step. The edge detector becomes an acceptable feature extraction technique for image-based measures, even when using the AM-WCSF filtering.

**Table 4-5.:** Mean coefficient of determination between image-domain feature measures and DMOS values using different contrast sensitivity filters.

	Filter	$d_{1,x}$	$d_{2,x}$	$d_{3,x}$
DMOS	CSF	0.74	<b>0.89</b>	0.73
	WCSF	<b>0.78</b>	0.87	<b>0.80</b>
	AM-WCSF	0.69	<b>0.89</b>	0.69

Remarkable results are achieved when the AM-WCSF filter is used inside QI assessment (Table 4-6). The usage of AM-CSF filter, image-based feature measures and LR strategy gives an increase of 5% of correlation between this assessment and DMOS, and the usage of AM-WCSF filter, image-based feature measures and ANNR strategy gives an increase of 8% of correlation between this assessment and DMOS, which confirms that the usage of AM-WCSF filter, together with either PCAR or ANNR strategy, increases the correlation. It is worth noting that the highest coefficient of determination is achieved (0.94) when AM-WCSF filter and ANNR strategy is used.

**Table 4-6.:** Mean coefficient of determination between QI assessment and DMOS values.

Filter	Regression Strategy	$R_{DMOS,pDMOS}^2$
CSF	LR	0.86
	PCAR	<b>0.93</b>
	ANNR	0.92
WCSF	LR	0.87
	PCAR	0.87
	ANNR	0.88
AM-WCSF	LR	<b>0.91</b>
	PCAR	0.92
	ANNR	<b>0.94</b>

## 4.5. Discussion

From the results of Table 4-1, the following points are discussed: First of all, wavelet-based *Edge Error* decreases correlation results with respect to the image-based counterpart. For instance, a decrease of 13%, 12% and 13% of correlation for this measure is noted using CSF, WCSF and AM-WCSF filters respectively. This indicates that wavelet-based Edge Detector is suboptimal in comparison with image-based edge detector. On the other hand, the coefficient of determination increases when wavelet-based *Blockiness* and *Visual Impairment* measures are used: Using WCSF the respective increase is of 6% and 12%, and using AM-WCSF the respective increase is of 8% and 11%. This indicates that a greater improvement in correlation can be achieved, providing that WCSF filters are used in assessment. Nonetheless, an improvement is still noted as the inverse wavelet transform is not required, avoiding some distortion caused by synthesis filter. These results demonstrate that WCSF and AM-WCSF filters work with the improvement of the coefficient of determination, except for wavelet-based *Edge Error* measure. Though the selection of feature measures leads to comparative correlation results, wavelet-based Edge Detector is not optimal in contrast with image-based *Canny* detector, and important revisions on the former must be made.

The results obtained in Table 4-6 show that AM-WCSF filter contributes to an improvement on assessments, regardless of the regression strategy used. An increase of 5% in correlation is noted when AM-WCSF is used in comparison with CSF counterpart, in LR strategy. Likewise, an increase of 2% of correlation is noted for the ANNR strategy, indicating that

CSF filtering is suitable only for image-based feature measures, which makes AM-WCSF more adequate for wavelet-based feature measures. It is also shown that PCAR and ANNR strategies allow an increase in correlation value, leading to proposed FW-QI assessment being adjusted to using the AM-WCSF filter, the wavelet-based measures and the PCAR strategy.

## 4.6. Concluding Remarks on Experiment

As stated in this Chapter, an adjustment on FW-QI assessment consists in the selection of the AM-WCSF filter that enhances perceptual features, and the linear combination of three wavelet-based feature measures, whose weighting factors are calculated by PCA regression. As a main advantage, FW-QI assessment uses the same wavelet scheme of representation. This particular aspect improves the processes involved with perceptual features, avoiding additional processing and enhancing measurement performance, specially for wavelet-based coders. From the results mentioned above, the following are worth highlighting: first of all, AM-WCSF filtering improves correlation between measures and human-made observation scores; secondly, FW-QI assessment can be used for distortion evaluation of image coders, providing the wavelet-based Edge Detector used for *Edge Error* feature measure must be improved; thirdly, higher correlation and sensitivity in measurement is achieved if PCA regression strategy is used for calculation of the parameter vector; lastly, adjusted FW-QI assessment gives comparable results of correlation and sensitivity against QI, with the advantage of the measurement being made on the encoded representation rather than on the reconstructed image.

# 5. Evaluation of Coding Performance by Image Quality Assessment

## 5.1. Preliminary

Quality evaluations on Medical images are expected to be as close as possible to human observations. In this case, the quality level is required for the expert in order to make an accurate diagnosis. Thus, a specific type of medical images is used to test performance of coders. The Mammogram (Table 2-1) is used for this evaluation, as quality for this kind of image is critical to aid the detection of abnormalities related with breast cancer. Breast cancer is one of the most frequent causes of medical consultation, hospitalization, morbidity and mortality, and the number of cases grows annually [35]. Therefore, the preservation of details is critical for areas where abnormalities are present in image, because some distortion can lead to a diagnosis misinterpretation, for instance, a false positive. An image quality assessment determines the level of uncertainty that a transmitted medical image can present for diagnosis purposes. With this in mind, the following considerations for coder evaluation must met:

- Firstly, they must not inflict an unaccepted load on the encoding process, because the number of processes in perceptual evaluations conventionally depends on the image size. It is an aspect of high concern, because medical images are characterized by their large size. FW-QI assessment aims to reduce the number of required operations for evaluation, thanks to the assessment being entirely wavelet-based.
- Also, the assessments must have a close approximation to the human-based scoring. In Chapter 4, FW-QI's ability to give a perceptual score was demonstrated, with an acceptable correlation value with human-based evaluations. Moreover, weighting coefficients have been calculated in order to achieve a high correlation between FW-QI assessment and human-based evaluations. For this purpose, half the LIVE database [14] is used as training set ([27, pg.4]). For the proposed assessment, weighting coefficients are calculated by regression using *Principal Component Analysis*, as recommended in [13], using 70% of the LIVE database as training set, where the DMOS values are outputs, whereas the distortion measures are the respective inputs for each image, as recommended in [29]. As a result, the coefficients stated in Equation (3-2) become the following suggested values:  $\lambda_0 = -0.01$ ,  $\lambda_1 = 0.07$ ,  $\lambda_2 = 3.19$ , and  $\lambda_3 = 0.73$ .

This test aims to illustrate the following special aspects that FW–QI assessment can provide for the perceptual distortion evaluation of coders: FW–QI assessment gives a perceptual score which behaves analogue with conventional evaluations used as distortion control parameters, such as PSNR. Also, the FW–QI assessment makes it suitable for widely standard coding schemes, like *JPEG2000*. These aspects will determine whether proposed assessment should be used as a parameter for distortion control within coding process. This Chapter is divided as follows: In Section §5.4, the experimental set–up is described, where the FW–QI assessment is used to evaluate the considered coders on Medical Images. Section §5.5 is a discussion on the results of the evaluation and on highlights of the proposed FW–QI assessment.

## 5.2. Overview of Region Of Interest Coders

This evaluation focuses on Region of Interest–coders, defined as follows:

**Definition 5.2.1** *The Region–Of–Interest–Coding is a kind of coding process preserving the quality of certain areas on the image, named Region of Interest (ROI), over the background (BG).*

ROI–coders are used for applications such as detail preservation, face detection (pattern recognition area) and content search (content description area). Fast and accurate ROI–coders have been used in emerging countries for *store-and-forward* telemedicine applications [4]. Here, a high–resolution image travels from repository data source to a remote destination through channels with reduced bandwidth, producing undesired effects at destination such as latency and/or loss of information. In other words, the received image is distorted by bandwidth constrains. Thus, if at least the ROI is transmitted with acceptable quality, the remaining information can be transmitted at subsequent time intervals, without relevant loss of quality. With this in mind, a number of ROI–coders has been proposed, mainly based on the progressive transmission of a set of remarked image details under some arbitrary bandwidth restriction.

A number of performance evaluations of ROI–coders have been conducted using this kind of assessments in [44], [81], and [11]. These evaluations focus on distortion measurement for three approaches: The whole encoded image, the ROI, and the distortion ratio between ROI and BG. The parameters used for performance evaluation of ROI–coders are usually those based on objective features (like *Peak Signal to Noise Ratio* –PSNR measure).

### 5.3. Implementation of Full-Wavelet Quality Index in PO-Shift Coder

The inclusion of the FW-QI assessment into *PO-Shift* distortion control strategy, or perceptual-Wavelet-ranked *PO-Shift* (*WPO-Shift*) coder is considered here. This implementation aims to prove that the proposed assessment can act as control distortion parameter within coders. As is illustrated in Figure 5-1, *PO-Shift* coder ranks the bit-planes by two scaling values  $s_1$  and  $s_2$ , using the *Perceptual Mean Squared Error* (PMSE) [81, Eq. 1]:

$$PMSE = \frac{\sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \|F(x(m_1, m_2) - \tilde{x}(m_1, m_2))\|^2}{\sum_{m_1=1}^{M_1} \sum_{m_2=1}^{M_2} \|x(m_1, m_2)\|^2}, \quad (5-1)$$

where  $F(\cdot)$  is a perceptual measure, and  $x(m_1, m_2)$  and  $\tilde{x}(m_1, m_2)$  are the elements of source and distorted images, respectively. The replacement of PSME of Equation (5-1) by FW-QI

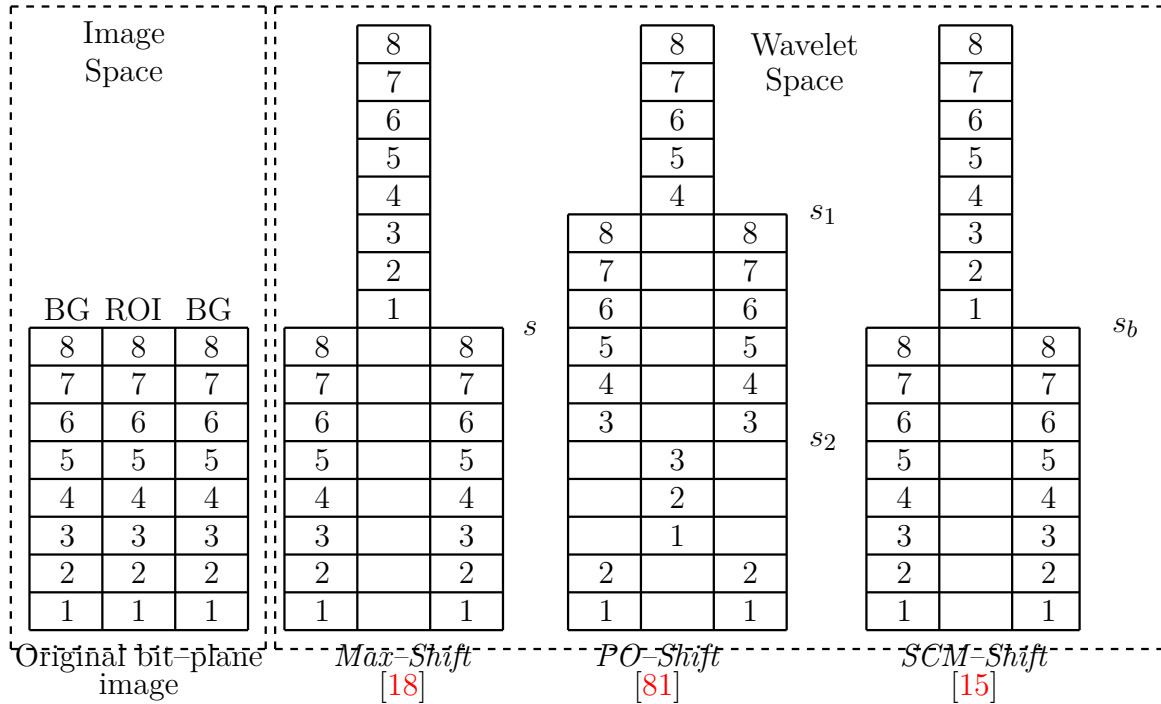


Figure 5-1.: Bit-plane scaling for considered *Max-Shift*-based ROI-coders.

assessment of Equation (3-2) converts *PO-Shift* into *WPO-Shift* coder, by displacing the bits in every bit-plane with zeros. Then, for every image the FW-QI assessment is computed. Finally, FW-QI values are sorted by size so that the bit-plane realigned ranking is obtained. Thereafter, the latter coder follows same methodology of former coder. As described in [81], the rest of the encoder is given as follows: All the bit-planes are realigned based on the different priority of the ROI and BG before coding. Then the whole bit-planes of ROI

coefficients are divided into two parts: The *Most Significant Bit-planes* (MSB) and the *Less Significant Bit-planes* (LSB). Also the whole bit-planes of BG coefficients are divided into two parts: The *Important Background Bit-planes* (IBB) and the *Less Background Bit-planes* (LBB). As shown in Figure 5-1, the LBB are shifted down by scaling value  $s_2$ , the MSB are shifted up by scaling value  $s_1$ , and the LSB and LBB are retained in their original place (they are not shifted). At the decoder, all the MSB coefficients can be identified in the same way as *Max-Shift* coder by shifting back. All bits lower than the original LSB are shifted up. The scaling values  $s_1$  and  $s_2$  can be changed according to the desired ROI visual quality. The rest of considered ROI-coder process used to preserve the quality of some regions of medical images are described in Section §B.

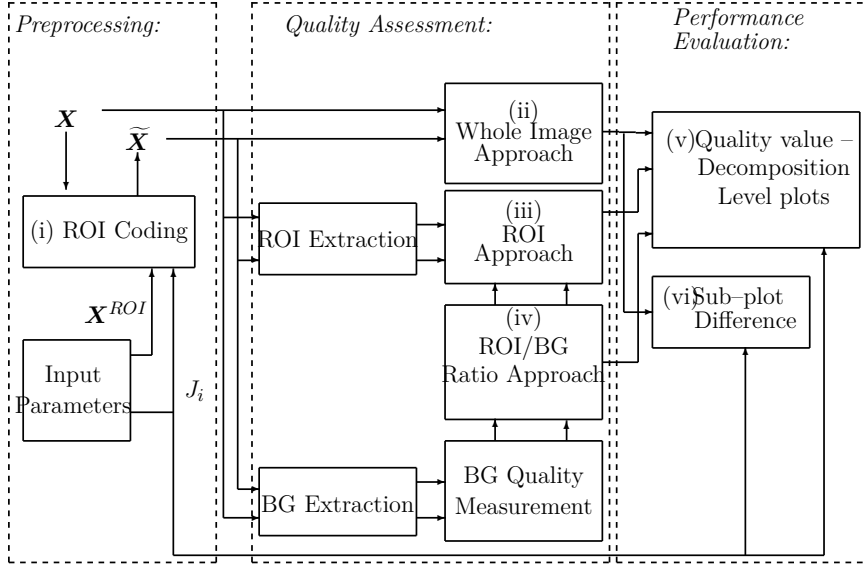
## 5.4. Experimental Setup

The evaluation of ROI-coders by perceptual quality assessments is carried out in the following stages, as illustrated in Figure 5-2: i) A source image is encoded, resulting in the encoded image for the number of sub-bands  $J_i$  used as input of encoding process; ii) Both source and encoded images are used in the assessment process, giving as result the *Whole Image* assessment approach; iii) The source and encoded ROIs are extracted from respective images, after which these two ROIs are used in another assessment process, giving as result the *ROI* assessment approach; iv) The source and encoded BGs are extracted from the respective images, then these two BGs are used for the BG assessment. Thereafter, the *Ratio between ROI and BG* assessments is calculated as the third assessment approach; v) Each Image and ROI Quality values and ROI/BG ratio is laid out on its respective *Quality-Decomposition Level plot*, whose axes are the calculated assessment values and the number of sub-bands  $\{J_i\}$ . This plot describes the relationship between the rate restriction used for encoding and the resulting loss of quality on the encoded image. For each image there are three Quality-Decomposition level plots: *Whole Image*, *ROI* and *ROI/BG Ratio*; vi) the difference between plots of different ROI-Coders is calculated.

Testing is carried out with an Intel® Core(TM)2 2.13 GHz processor, with 3.7 GB of RAM, under Matlab environment, where ROI-coders are implemented on the routine as proposed in [31].

### 5.4.1. Database and Preprocessing

The input test space consists of images from the mini-MIAS database of mammograms [69]. This database is a repository of mammograms of  $1024 \times 1024 \times 1$  size, which are classified according to a labeled set of found abnormalities. This set is composed by the following parameters: Character of background tissue, Class of present abnormality, Severity of abnormality, and  $(x, y)$  Position of abnormality in image. For this experiment, only the **Severity of abnormality = Malignant** cases are considered, due to the grade of relevance



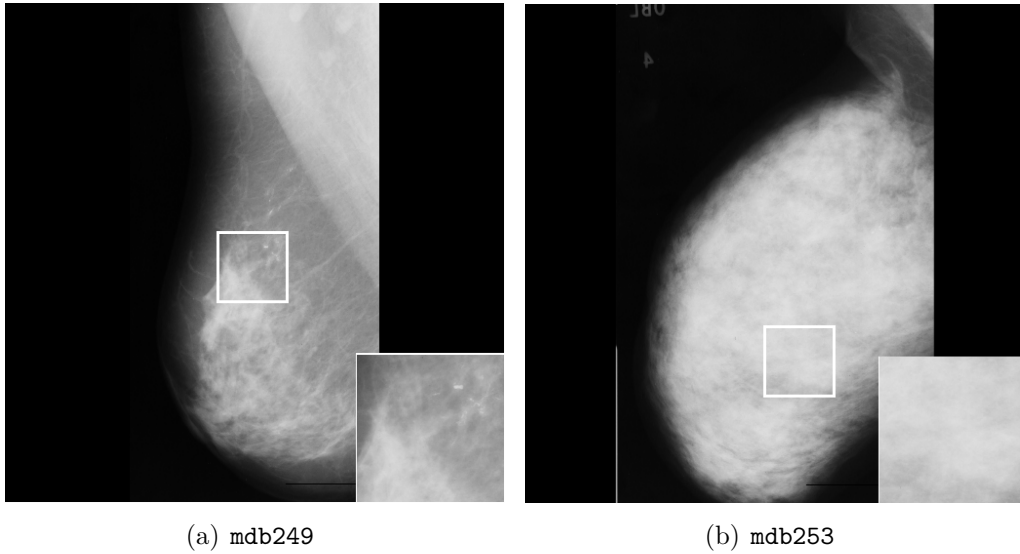
**Figure 5-2.:** Schematic representation of an assessment for ROI-coders.

for Breast Cancer detection techniques. Three samples are shown in Figure 5-3 with its respective ROI. The position of the center of ROI is the marked Position of abnormality and radii of ROI is a value where most of the clustered malignant detected abnormalities can be covered. In this case, a radii value of 75 pixels is arbitrarily selected. Moreover, the shape of ROI-Mask used has been fixed empirically as **Square**. The shape is selected in compliance with requirements in JPEG2000 standard for ROI-coders, especially, *Max-Shift*-based ones. The shape and position of ROI remain unchanged for each image through each coding process, though concrete polygon placing is accomplished manually. Thereafter, encoding is stated by *Max-shift*, *SCM-Shift*, *PO-Shift*, and *WPO-Shift* coders, where the quality of the image is controlled by loss of a number of less-significant bit-planes to fit rate restriction. A sequence of defined rates should be used as input parameter for each ROI-coder, and the output is then the corresponding sequence of measured distortion values. However, as considered for JPEG2000 standard, the coding rate parameter is defined only when *Tier-1* and *Tier-2* subsequent encoding processes are achieved. Different parameters must be used for the encoder, as follows: In *Tier-1* coder, every wavelet coefficient  $c_b \in \mathcal{C}$  found in sub-band  $b$  is quantized into coefficient  $q_b$  by using quantization parameter  $\mathcal{Q}_b(\cdot)$  [72]:

$$q_b = \mathcal{Q}_b(c_b) = \text{sgn}(c_b) \left\lfloor \frac{|c_b|}{\Delta_b} \right\rfloor,$$

where  $\Delta_b$  is the quantization step at bit-plane  $b$  reached by the decomposition level  $j_b$ , and defined as:

$$\Delta_b = 2^{M_b - \epsilon_b} \left( 1 + \frac{\mu_b}{2^{11}} \right), \quad (5-2)$$



(a) mdb249

(b) mdb253

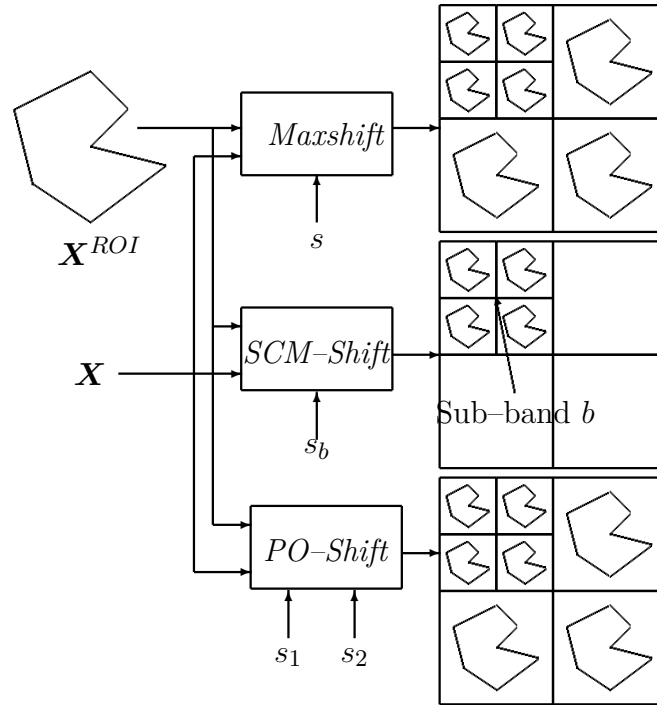
**Figure 5-3.:** Test images labeled M (Malignant) in mini-MIAS database for ROI-Coding Evaluation purposes. The white square corresponds to ROI area, which is amplified in the bottom-right corner. Type of image: 8-bit gray-scale Mammogram. Size of each image:  $1024 \times 1024 \times 1$ . Images are fitted from their original size.

where  $M_b$  is the number of bit-planes in the original image, i.e., the source image bit-depth, assuming that the gain expected is equal to 1 for all sub-bands in source representation. Besides,  $\epsilon_b$  and  $\mu_b$  are the respective exponent and mantissa of the quantization step at wavelet sub-band  $b$ . The pair  $(\epsilon_b, \mu_b)$  from Eq. (5-2) is calculated by [31]:

$$(\epsilon_b, \mu_b) = (\epsilon_{LL} - J + j_b, \mu_{LL}) , \quad (5-3)$$

where the subscript LL indicates the coarsest (lower) sub-band,  $j_b$  is the number of decomposition levels required to achieve sub-band  $b$ , and  $\epsilon_{LL}$  is the exponent parameter at the LL sub-band. Considering that exponent  $\epsilon_{LL}$  and mantissa  $\mu_{LL}$  are fixed parameters, the quantization width depends only on the number of decomposition levels  $J$  in Equation (5-3). Thus, it is assumed as input value for quality control in encoder. According to [31], the proper value for mantissa is  $\mu_b = \mu_{LL} = 8$ , and for exponent is  $\epsilon_{LL} = 8.5$ , in order to get a lossless recovered image when  $J = 1$  ( $\Delta_b \approx 1$ ). The higher the total decomposition level number – the wider quantization step at finer sub-bands, and the less-significant number of bit-planes are lost. The adjustment of scaling values are described for each ROI-coder as follows: In *Max-shift* coder,  $s$  corresponds to the bit-plane number where the maximal value of BG resides, i.e., the maximal bit-plane of BG. The value  $s = 8$  is assumed to cover the maximal number of bit-planes of BG. For *SCM-Shift* coder, wavelet coefficients are scaled-up by the same factor  $s_b = 8$ , but ROI-shifting is made from coarsest (including ap-

proximation coefficients) to  $b$  sub-bands, as shown in Figure 5-4. The parameters of this coder are intentionally adjusted, to give a lossless ROI quality when  $J \leq b$ , and a ROI quality loss when  $J > b$ . Thus, the arbitrary selected sub-band value is  $b = 3$ . For instance,



**Figure 5-4.:** Mappings for considered *Max-Shift*-based ROI-coders.

an image is encoded at  $J = 1$  decomposition levels, a lossless image – and therefore lossless ROI – should be appreciated. Otherwise, if the image is encoded at  $J > 3$ , an incremental loss of ROI quality will show in the encoded image. Finally, for the *PO-Shift* and *WPO-Shift* coders, scaling factors  $s_1$  and  $s_2$  are selected by previous perceptual analysis of bit-planes in images to give quality priorities for MSB, LSB, IBB and LBB. In [81], the scaling values of  $s_1 = 5$  and  $s_2 = 3$  are suggested. Note that  $s_1 + s_2 = 8$  is used to cover the same maximal number of bit-planes as for *Max-Shift*.

Concerning test images, pre-processed versions of source and encoded images are used for wavelet-based *Blockiness*, *Edge Error* and *Visual Impairment* feature metrics, in order to take into account luminance masking and contrast sensitivity effects. Pre-processing is then conducted by passing the image through AM-WCSF filter, as discussed in Section §3.2. The distortion is measured on each ROI-encoded image. After wavelet coefficients are extracted from the image, they are quantized to increase occurrence probability for a selected set of elements, thus fitting transmission to rate constraints. Therefore, inside JPEG2000 coder, an image processed by a given ROI-coders suffers distortion caused by quantization either on ROI or BG [50].

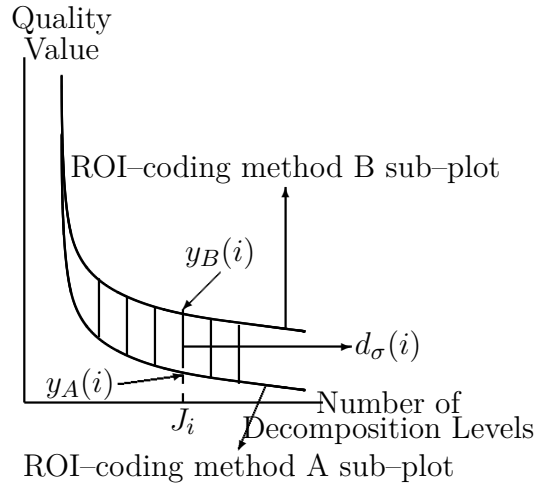
### 5.4.2. Considered test evaluations

For the selection of the ROI-coder achieving the lowest distortion values whether by preserving quality of ROI elements or by maintaining relevance between ROI and BG, FW-QI assessment is compared with other evaluations: For the objective-based evaluation, *Peak Signal to Noise Ratio* (PSNR) is considered. For those that are perceptual-based, *Quality Index based on Local Variance* (QILV) [3] and *Reflection Factor* (RF) [47] assessments are considered. For the proposed evaluation, the number of wavelet decomposition levels required should be adjusted to 3 for optimal extraction of perceptual features. Coarser sub-bands are synthesized by filter bank, until the required number of decomposition levels is reached, fitting the encoded wavelet structure to level restriction. Finally, for PSNR, QILV and RF evaluations, original source and encoded images are evaluated without the use of any CSF pre-processing on image. Testing is carried out using an Intel® Core(TM)2 2.13 GHz processor, with 3.7 GB of RAM, under Matlab environment, where ROI-coders are implemented following the routine given in [31].

### 5.4.3. Decomposition Level-Assessment Value Plots

An *Assessment Plot* is the plot considered for the evaluation of the considered ROI-coders. Axes of this plot are the measured quality assessment values and a rank of the number of decomposition levels  $\{J_i\}$  used as ROI-coding input parameter (Fig. 5-5). To take these quality values, the following full-referenced assessments are used: *Peak Signal to Noise Ratio* (PSNR), *Quality Index based on Local Variance* (QILV) [3], *Reflection Factor* (RF) [47], and FW-QI assessment. The rank of number of decomposition levels is arbitrarily selected as  $i = [1, 7]$ . Starting with the first lossy condition at  $J_{i=1} = 2$  decomposition levels, as the lossless condition  $J = 1$  is discarded. Thereafter,  $J_i$  rank is  $\{J_i\} = [2, 8]$  for  $i = 1, \dots, 7$ . Each source image from the database is encoded at  $J_i$  number of decomposition levels, giving a encoded image. The distortion of the image is then measured by each of the three assessment approaches, as shown in Fig. 5-2: *Whole Image*, *ROI* and *ROI/BG Quality Ratio*. The results of this measurement are the respective quality values given for each image, which are then averaged, giving the quality value  $y(i)$  shown in Figure 5-5, and sketched on the plot. These approaches are depicted as follows:

- The *Whole Image* approach provides a general ROI-coder performance evaluation. Each test image is coded by some ROI-coder, using as distortion parameter the number of decomposition levels  $J_i, i = [1, 7]$ . For the number of  $n$  images, a number of  $7n$  distorted images are given. Thereafter, an assessment is made on the whole image, obtaining  $7n$  values. These values are averaged with respect to the number of images. As a result, 7 average values per assessment are given. The same procedure is conducted for the rest of ROI-coders. It gives, for *Max-Shift*, *SCM-Shift*, *PO-Shift* and *WPO-Shift* coders, a number of 4 sub-plots for each plot. Finally, 4 plots are displayed,



**Figure 5-5.:** Number of Decomposition Levels vs. Assessment Value Plot used for evaluation of ROI-Coders.

one for each respective assessment. For the sake of illustration, a description of these sub-plots is shown in Figure 5-5, where *A* ROI-coder plot is gathered with the *B* counterpart. Thus, considered coder performances can be compared.

- The *ROI* approach is used to give an evaluation on preservation of ROI quality. It uses the same procedure as for *Whole Image* approach, but assessment is made only on ROI, requiring ROI extraction for source and distorted image (Fig. 5-2).
- The *ROI/BG Ratio* approach is used to evaluate ROI-coder's ability to manage the preservation of ROI quality in comparison with BG quality. Assessments of both ROI and BG are required (Fig. 5-2), and their respective areas are to be extracted separately. ROI quality value is then divided by BG quality value. The following points are considered: first, a ROI/BG ratio higher than one means that ROI quality is preserved by compromising only BG elements in quantization process; second, when the ratio is around one, both ROI and BG elements are compromised in quantization process. Therefore, this approach allows measurement of the distortion on encoded images, when either a region or the whole image is quantized.

#### 5.4.4. Numerical evaluation between ROI-coders

In order to give a numerical evaluation of ROI-coder performance, the difference between sub-plots is measured for each number of decomposition level  $J_i$ . As each assessment is very different in scale to each other, the normalized error between plots is required for comparison purposes. Let  $y_A(i)$  and  $y_B(i)$  be the average quality assessment values given for the respective ROI-coding sub-plots *A* and *B*, at decomposition level  $J_i$ , as shown in

Figure 5-5. Normalized error  $d_\sigma$  for each decomposition level is calculated as:

$$d_\sigma(i) = \frac{y_A(i) - y_B(i)}{y_A(i)},$$

where  $y_A(i)$  is used as reference value. *Max-Shift* coder is arbitrarily selected as reference sub-plot; if the value  $d_\sigma(i) > 0$ , ROI-coder evaluation  $B$  gives better results of quality preservation in comparison with reference  $A$  coder evaluation. Otherwise, coder  $A$  gives a better preservation of quality. This evaluation is stated only for the *Whole Image* approach.

## 5.5. Results

The following results illustrate the performance of FW-QI assessment in comparison with other considered evaluations. It aims to show the particular characteristics that the proposed assessment offers for image quality evaluation, when they are encoded by some considered ROI-coder. Accordingly, the  $x$ -axis of each *Assessment* plot is composed of an increased ranking of decomposition level values, which are used as coder input parameter.

### 5.5.1. Decomposition Level-Assessment Value Plots

Three approaches provide the corresponding assessment plots: Whole image, ROI and ROI/BG Ratio. For the *Whole-Image* approach, which is shown in Figure 5-6, a description of resulting quality values is given, where images are encoded by the coders in question with measurement given on the whole image. For PSNR (Figure 5.6(a)), *Max-Shift* coder shows the highest assessment value for the considered images, at any number of decomposition levels. However, for FW-QI assessment (Figure 5.6(d)), the *Max-Shift* coder gives lower quality levels than for *PO-Shift* or *WPO-Shift*, indicating that FW-QI assessment detects perceptual features that are relevant for performance of ROI-coders. It is also noted that QILV and RF assessments (Figures 5.6(b) and 5.6(c), respectively), also give a denoted separation between objective-based distortion plots (*Max-Shift* and *SCM-Shift*) and perceptual-based plots (*PO-Shift* and *WPO-Shift*). The results obtained in Figure 5-7 illustrate the average quality level of images encoded by considered ROI-coders, using the *ROI* approach. In this case, the *Max-Shift* sub-plot shows the highest quality values using PSNR assessment (Figure 5.7(a)) in comparison with the other sub-plots. In the same way, the QILV, RF and FW-QI assessments illustrate in their respective ROI Quality plots (Figures 5.7(b), 5.7(c) and 5.7(d)) that *Max-Shift* sub-plot gives the best quality index values. The decrease of quality that is aimed for *SCM-Shift* coder is not detected for the QILV assessment. However, the PSNR, RF and WQI assessments confirms in their respective ROI Quality plots the decrease of quality in ROI, when the number of decomposition levels used for encoding is more than 5. A more detailed behavior is sketched by using PSNR and WQI assessments, as shown in respective Figures 5.7(d) and 5.7(a), where *SCM-Shift* sub-plot

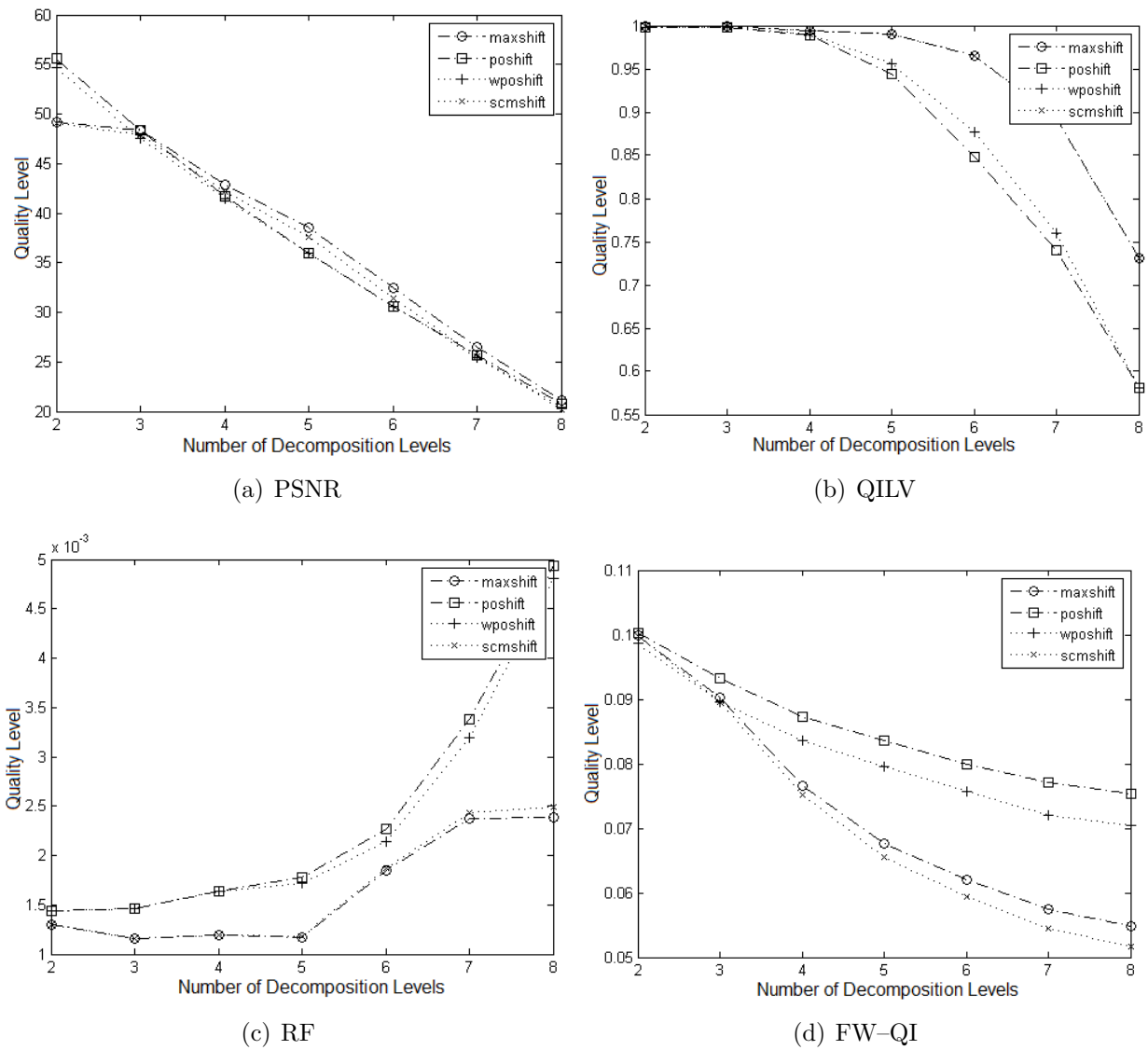
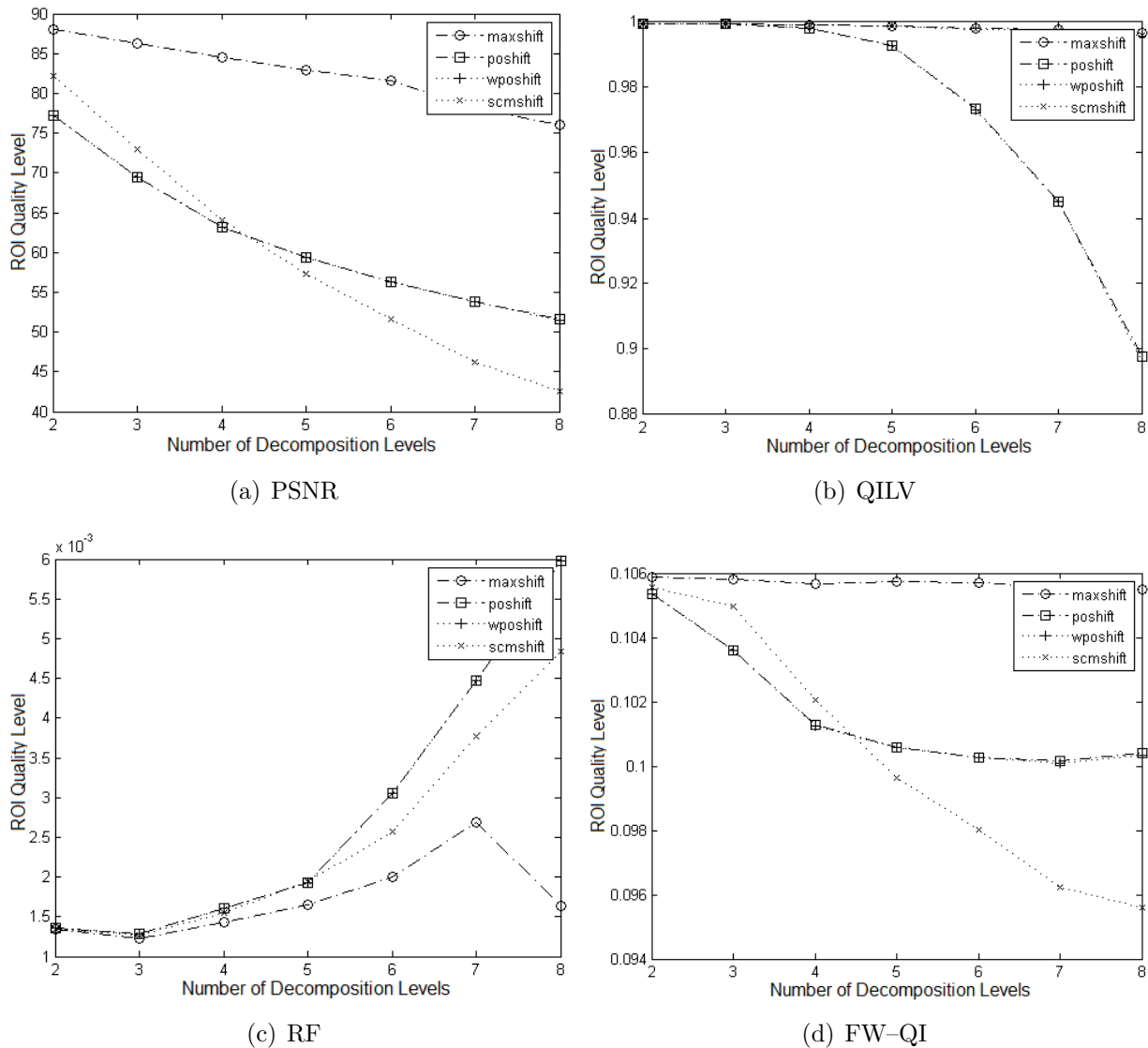


Figure 5-6.: ROI-Coder evaluation, using *Whole Image* approach.

seems clearly separable from its *Max-Shift* counterpart. Finally, the QILV assessment fails to detect quality decrease of *SCM-Shift* coder, as shown in its respective ROI approach of Figures 5.7(b), where the *SCM-Shift* sub-plot is equal to *Max-Shift* counterpart. In addition, it is observed for the RF assessment (Figures 5.7(c)), that the *Max-Shift* sub-plot shows a low increase with respect to the number of decomposition levels used in encoding. This result is expected, as ROI quality is always preserved by *Max-Shift* coder for any number of decomposition levels used in encoding. Finally, the *ROI/BG ratio* approach aims



**Figure 5-7.:** ROI-Coder evaluation, using *ROI* approach.

to give a measure of how the BG quality is reduced with respect to ROI. As described in Section 5.4.3, a ROI/BG ratio value above one means that the encoded image becomes a

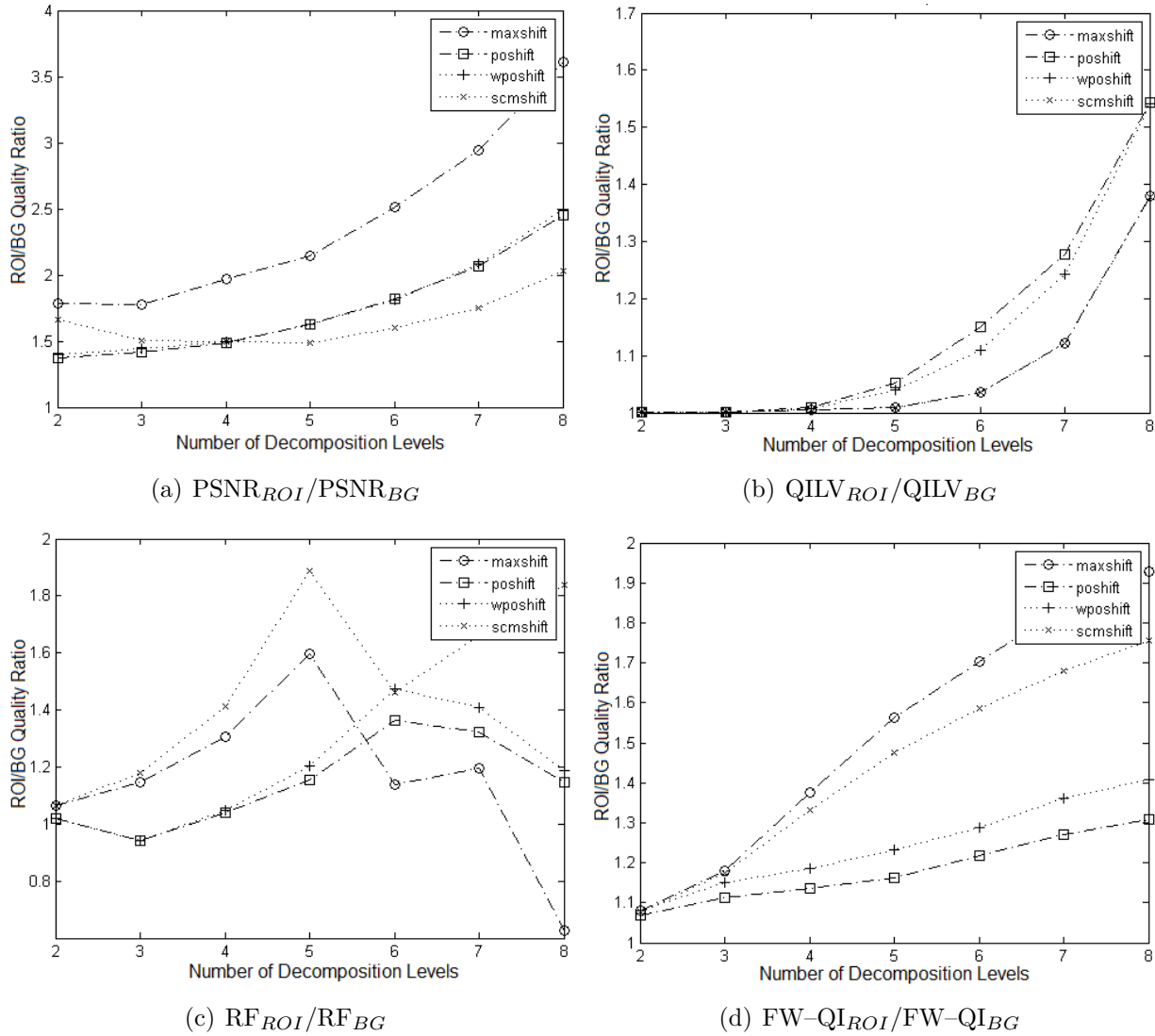
high ROI quality value, whereas BG quality has been drastically reduced. Otherwise, both ROI and BG qualities has been balanced. However, the results show that it is true only for assessments whose values are greater than one. For instance, results for ROI/BG ratio using QILV and RF assessments give different results, as expected. For QILV assessment (Figure 5.8(b)) it is noted that ROI-coding sub-plot values increase only from the number of 5 decomposition levels, even for *Max-Shift* plot, where an increment of ROI/BG ratio to above one for any number of decomposition levels is expected. The above results explain the following: in Figure 5.7(b), it is noted that *Max-Shift* sub-plot achieves values close to one, whereas both *PO-Shift* and *WPO-Shift* plots get values below that. For *Max-Shift* coder, the ratio  $\text{QILV}_{ROI}/\text{QILV}_{BG}$  is approximately equal to  $1/\text{QILV}_{BG}$ , which is around one, as  $\text{QILV}_{BG} \approx 1$  for any considered decomposition level. For the other coders,  $\text{QILV}_{BG} \leq 1$ , which means that ratio  $\text{QILV}_{ROI}/\text{QILV}_{BG}$  is greater than one only if  $\text{QILV}_{ROI} > \text{QILV}_{BG}$ . This result is shown in Figure 5.8(b), concluding that ROI quality is preserved in a higher level than the BG. QILV shows the decrement of quality in ROI for *SCM-Shift* coder when the number of decomposition levels used is higher than 5. Likewise, for the RF assessment (Figure 5.8(c)), a decrease of the ROI/BG ratio is shown for *Max-Shift* and *SCM-Shift* sub-plots from 5 decomposition levels. From evaluations shown in Figures 5.7(c) and 5.6(c), this decrease is caused because RF values are very close to zero. Therefore, the following cases come out for the ROI/BG ratio:

$$\frac{\text{RF}_{ROI}}{\text{RF}_{BG}} = \begin{cases} > 1 & \text{RF}_{ROI} > \text{RF}_{BG} \\ = 1 & \text{RF}_{ROI} = \text{RF}_{BG} \\ < 1 & \text{RF}_{ROI} < \text{RF}_{BG} \end{cases}$$

The RF evaluation can give results of ROI/BG ratio below one when distortion on BG is higher than ROI. Figure 5.8(c) confirms that *Max-Shift* coder gives a higher preservation of ROI, especially when a number higher than 5 decomposition levels is used. For these cases, RF values at BG are higher than ROI, giving the decreasing ratio values in Figure 5.8(c). The same occurs for *SCM-Shift* coder, where ROI/BG ratio decays from 5 decomposition levels. Besides, *PO-Shift* and *WPO-Shift* plots achieve ROI/BG ratio values inside interval  $[0.8, 1, 6]$ , which suggests that these ROI-coders can perceptually preserve the quality of ROI without a relevant degradation on BG areas.

For the rest of the assessments, results are given as expected: using PSNR assessment (Figure 5.8(a)), the *Max-Shift* sub-plot increases from values from 2 to 4, because ROI quality becomes the highest priority, then BG quality becomes lower than the ROI. In addition, respective *SCM-Shift*, *PO-Shift* and *WPO-Shift* subplots are located between 1.5 and 2.5, indicating that ROI quality is diminished in comparison with BG, but ROI quality is close or higher than the BG counterpart. The change of quality is appreciated for *SCM-Shift* sub-plot, when the number of decomposition levels used is higher than 5. Finally, from WQI perspective (Figure 5.8(d)), *Max-Shift* sub-plot is located between 1.1 and 2, and *PO-Shift* and *WPO-Shift* sub-plots rise increasingly from 1.1 to 1.4. Finally, a noticeable difference

is observed for *SCM-Shift* sub-plot, where its values increase form 1.1 to 1.7, indicating that ROI quality has a priority higher than BG, as expected for this coder.



**Figure 5-8.:** ROI-Coder evaluation, using  $ROI/BG$  ratio approach.

### 5.5.2. Numerical ROI-Coding evaluation

Results of the evaluation of the ROI-coders in question are shown in Table 5-1 using Image Quality assessment, where *Max-Shift* coder gives higher PSNR-Image Quality assessment values, whereas *PO-Shift* gives higher values. By evaluating the difference between *Max-Shift* and the other ROI-coding sub-plots using PSNR assessment, the *SCM-Shift* sub-plot illustrates a decrease of quality in up to 4.6% of images encoded by this coder. In turn,

the *PO-shift* and *WPO-Shift* sub-plots show a decrease in quality up to 6.8% and 6.7%, respectively. Furthermore, by evaluating the difference between *Max-Shift* and the other ROI-coding sub-plots using WQI assessment, the *PO-shift* and *WPO-Shift* sub-plots give an increased quality up to 37.2% and 28.6%, respectively, indicating that FW-QI assessment has the ability to describe a perceptual assessment, whose details cannot be observed in an objective-based assessment. The following issues are found for QILV and RF assessments by numerical evaluation of ROI-coders: for the QILV assessment, values of *Max-Shift* sub-plot are very close to one, as described graphically in Figure 5.6(b). However, there is no difference between *Max-Shift* and *SCM-Shift* sub-plots (Table 5-1), indicating that the intended degradation of quality ROI from 5 decomposition levels is not detected by QILV assessment. On the other hand, the *PO-Shift* and *WPO-Shift* sub-plots achieve a wider loss of quality on the whole image approach, in comparison with *Max-Shift*. For the RF assessment, values of *Max-Shift* sub-plot are of the order of  $10^{-3}$ , as illustrated in Fig. 5.6(c). It is compared with results given in [47], where the order of assessment values is the same  $10^{-3}$  for images distorted by JPEG2000. However, percentage of sub-plot differences are shown in Table 5-1, showing an acceptable discrimination value between sub-plots. The *Max-Shift* coder gives the lower distortion values, and in this case, the intended distortion arranged for *SCM-Shift* is detected. Finally, comparing *PO-Shift* and *WPO-Shift* sub-plots from Table 5-1, the ROI-coder that uses FW-QI assessment as quality control parameter gives comparable performance results, making implementation of FW-QI assessment inside ROI-coders to be possible.

## 5.6. Discussion

The results illustrated in Figure 5-6 suggest that a perceptual evaluation of considered ROI-coders (*Max-Shift*, *SCM-Shift*, *PO-Shift* and *WPO-Shift*) can be adequately illustrated using FW-QI assessment. In this case, *Whole Image*, *ROI* and *ROI/BG* quality plots show the preservation of ROI quality used for the *Max-Shift* coder, the decay of the ROI quality for the *SCM-Shift* coder from a certain number of decomposition levels, and the balance between ROI and BG quality achieved for *PO-Shift* and *WPO-Shift* coders.

However, from the results illustrated in Figures 5-7 and 5-8, there are some conditions that assessments must achieve to give a correct interpretation of results. For instance, the QILV assessment does not adequately detect some relevant features for evaluation. In this case, the quality values for images encoded by *Max-Shift* coders are close to one at *ROI* approach. Moreover, RF and QILV assessments do not show a coherent behavior of quality values at *ROI/BG ratio* approach, as shown in Fig. 5-8. On the contrary, FW-QI and RF assessments fit with considered assessments, suggesting that the use of a linear combination of measures improves the detection rank of features and improves its correlation with human-based models. In other words, a better perceptual evaluation can be stated using a weighted sum of perceptual parameters, specially for some distortion behavior aspects that

**Table 5-1.:** Averaged Image Quality Assessment difference percentages between sub-plots by *Whole Image* approach

Coder	Assessment	Sub-plot differences (%)							
		$J_i$	2	3	4	5	6	7	8
<i>SCM-Shift</i>	<i>PSNR</i>	-0.32	-0.91	-1.7	-2.5	-3.0	-3.8	-4.6	-2.4
	<i>QILV</i>	-0.00	-0.00	-0.00	-0.00	-0.01	-0.01	0.00	-0.003
	<i>RF</i>	-0.02	0.73	0.92	1.90	0.93	3.53	5.42	1.9
	<i>WQI</i>	-0.07	-0.6	-1.8	-3.1	-4.1	-5.4	-5.9	-3.0
<i>PO-Shift</i>	<i>PSNR</i>	-0.35	-0.79	-1.8	-6.8	-5.8	-3.2	-1.3	-2.9
	<i>QILV</i>	-0.08	-0.09	-0.50	-4.65	-12.2	-17.2	-20.7	-7.9
	<i>RF</i>	19.2	33.9	41.1	56.1	35.3	52.5	104.6	49.0
	<i>WQI</i>	0.41	3.3	13.8	23.7	29.0	33.9	37.2	20.2

should be imperceptible for other assessments. Moreover, FW-QI and RF assessments give a close description of perceptual features involved with loss of quality.

The results of Table 5-1 show that the ROI quality is one of the comparison parameters for ROI-coders. The ROI-coders based in a balanced quantization between ROI and BG provide better perceptual evaluations than the coders which highly prioritize ROI quality over BG.

Another point to note from these results is that *Max-Shift* sub-plot shows the highest PSNR whole-image and ROI quality approaches. It confirms the fact that this coder keeps ROI coefficients from any relevant quantization step. On the other hand, the other coders take some sub-band (*SCM-Shift* case) or set of bit-planes (*PO-Shift* and *WPO-Shift* cases) of ROI coefficients through the quantization process, making ROI distorted. However, from results on the ROI/BG ratio quality plot for *Max-Shift* coder, the quality balance between ROI and BG quality exceeds the expected value around one, which is successfully achieved for *PO-Shift* and *WPO-Shift* coders. Using FW-QI assessment, *PO-Shift* sub-plot shows the highest whole-image values, maintaining the ROI/BG ratio quality values around one as well.

Regarding *SCM-Shift* coder, the results given from *Whole Image* and *ROI* approaches present a decrease of quality values for PSNR and FW-QI assessments when the number of decomposition levels used is higher than 5 (Figures 5.7(a) and 5.7(d), respectively), illustrating the effect caused when ROI elements are shifted only for selected sub-bands. This is

observed only for PSNR and FW-QI assessments, demonstrating that proposed assessment, despite being a perceptual assessment, gives results comparable to others.

For most evaluations, *WPO-Shift* coder achieves comparable results with respect to *PO-Shift* counterpart. For the whole image approach, *WPO-Shift* achieves the same scores as for *PO-Shift*, as shown in Figure 5-6, which indicates that such FW-QI assessment achieves same correlation with respect to perceptual evaluations. This indicates that FW-IQ can be considered as distortion control strategy. However, the wavelet-based Edge Detector must be improved in order to achieve a better performance of evaluation.

## 5.7. Concluding Remarks

The Full-Wavelet Quality Index assessment provides a perceptual evaluation of ROI-coding performance, where perceptual adjustment of bit-planes for image elements in the ROI has improved perceptual quality in coded image. This note on the assessment relies on the fact that some aspects similar to PNSR evaluation are detected, especially making a measurable comparison between objective and perceptual ROI-coders. The importance of this result relies on the ability to perceptually evaluate and differentiate the performance of ROI-coders by using FW-QI assessment. Perceptual degradation of ROI can be demonstrated, for instance, by comparison between *Max-Shift* and *SCM-Shift* coders. However, in this case, the later coder is purposefully arranged to achieve a degradation of the ROI quality at an arbitrary sub-band. The challenging goal for these assessments is to provide a score, which is used for ROI-coders as control parameter. The control scheme is based on a novel approach for ROI masking and bit-plane-scaling stages that are formulated in the framework of common ROI coders. As a result, wavelet coefficients are sorted into foreground, or ROI, and background coefficients, and quantized accordingly. In conclusion, the proposed assessment provides a support for the preservation of ROI for large images of large size, where the acquisition of ROI implies the processing of a large quantity of elements, reducing the amount of computational load for on-line implementations used in segmentation or texture classification, without imposing additional charge for the coding process.

# 6. Conclusions and Further Work

## 6.1. Conclusion

A novel image quality assessment for evaluation of coder performance was presented. This *Full-Wavelet Quality Index* (FW-QI) assessment is composed of the following stages:

- An improved Contrast Sensitivity Function, whose masking function is calculated in wavelet domain.
- A linear combination of measures calculated in wavelet space, whose features are related with human vision models.
- An optimized set of weighting coefficients, which increases the correlation between assessment and human-based scores.

This assessment uses perceptual distortion measures in wavelet domain, whose features are increased by using an adaptive masking sensitivity filter. This method comprises the following improvements:

- A reduction of the number of additional processes in evaluation is achieved using only one multi-resolution representation for the whole evaluation process. Note that the proposed assessment requires only the source and encoded wavelet representations.
- In Section §3.2.2, a novel *Adaptive Masking Wavelet Contrast Sensitivity Function* (AM-WCSF) filter is a generalized filter which adapts the wavelet image representation to a generalized contrast map, making it flexible to any wavelet scheme. This contributes to increase the correlation between perceptual assessments and human-based evaluations, thanks to its adaptive masking attribute. The use of the AM-WCSF as perceptual enhancing filter allows the CSF mask of Equation (2-10) to be adapted any wavelet structure, by the arrangement of scaling arguments of wavelet structure to the desired CSF mask, whose calculation depends solely of the wavelet filter used for either image representation or coder. This filter gives a generalized representation by re-sampling of the filter arguments.
- In Section §3.2, the selection in wavelet domain for calculation of features allows the assessment to be conducted on encoded image instead of reconstructed image. The calculation of perceptual feature measures in wavelet domain assures a measurement

on encoded representation rather than on reconstructed image. The use of wavelet-based *Blockiness* (Section §3.2.3), *Edge Error* (Section §3.2.4) and *Visual Impairment* (Section §3.2.5) feature measures permits the use of perceptual measurement without additional filtering or transform, as features required for measures are located in the same wavelet representation.

- In Section §4, the weighting coefficients of linear combination of proposed *Wavelet-based Quality Index* are adjusted. For this purpose, a PCA-based regression strategy was selected to calculate these coefficients. The PCAR strategy gives remarkable results in correlation and stability in measurement, as shown in Tables 4-2. A PCA-based regression strategy is selected for calculation of weights for the weighted linear combination, which gives a more accurate and sensible measurement. However, during the realization of this work, other image databases has been analyzed by human-based evaluations [33, 39, 64]. Thus, it is important to re-evaluate the adjustments based on these newer databases, in order to discover if the proposed assessment should achieve similar results of correlation with respect to human evaluations.
- In Section §5.4, an evaluation of Medical Image Coding process is made by using proposed FW-QI. The evaluation is focused on Region-of-Interest Coding methods, where Mammograms are used as test images. As a result, FW-QI gives not only a perceptual characterization of ROI-Coding Methods, but its feature measures can be used for perceptual-based ROI-Coding methods, such as *PO-Shift*.

In conclusion, the proposed FW-QI assessment leads to a direct measurement on encoded wavelet coefficients and a remarkable correlation and sensitivity in measurement. An increase of correlation between assessment and human evaluation and stability in measurement is attained to evaluation, thanks to the adaptive properties of AM-WCSF filter and the selection of a PCA-based regression strategy.

## 6.2. Future Work

The following issues must be considered for future works:

- Wavelet-based Edge Detector must be improved, due its sub-optimal results in detection of edges required for *Edge Error* function. This is demonstrated in the results of Table 4-1, where a decrease of correlation between this measure and human evaluations is shown.
- It is suggested that a *Contourlet* or *Bandelet* directional filter is more suitable in order to extract features. These filters must be evaluated to participate in feature measures.

- Other regression/pooling strategies can be considered for the perceptual measure. For instance, fractal and fuzzy strategies, as in [78], are presented as candidates. This might give a increase in correlation between assessment and human vision models.
- A particularly relevant aspect to consider in terms of evaluation is the capacity to classify different types of distortion. In this work, only distortion caused by coding process is considered. Thus, a new experimental set-up must be stated to prove whether FW-QI can separate different distortion types. Other perceptual measures should be considered for inclusion in WQI linear combination equation. For instance, those measures based on structural similarity can be considered.
- The inclusion of FW-QI as parameter into distortion control process in JPEG2000 Standard.

Beyond medical applications, *Full-Wavelet Quality Index* (FW-QI) assessment has been tested in the following image processing applications: i) Face recognition classification, ii) Wearing textures characterization in carpets, and iii) Motor fault identification. These applications are still in testing, because in some cases only the feature measures are used separately rather FW-QI, and testing with all elements of evaluation is still in consideration. Nonetheless, the above enumerated experiments are described as follows.

### 6.2.1. Evaluation on Face Recognition Process

Feature measures are tested to measure distortion over a face recognition classifier proposed in [74], based on *Principal Component Analysis* (PCA) vector space. First of all, it is noted that regression coefficient estimation setup is stated using images distorted by JPEG, JPEG2000, Gaussian Blur, Fast Fading and white noise processes [56], [60]. Therefore, the intention of this experiment is to verify whether measures can distinguish face images that are distorted by the above commented processes for recognition purposes, in other words, to determine if an image can be adequately classified when it fits with WQI-based distortion criteria. Experimental set-up is described in Figure 6-1.

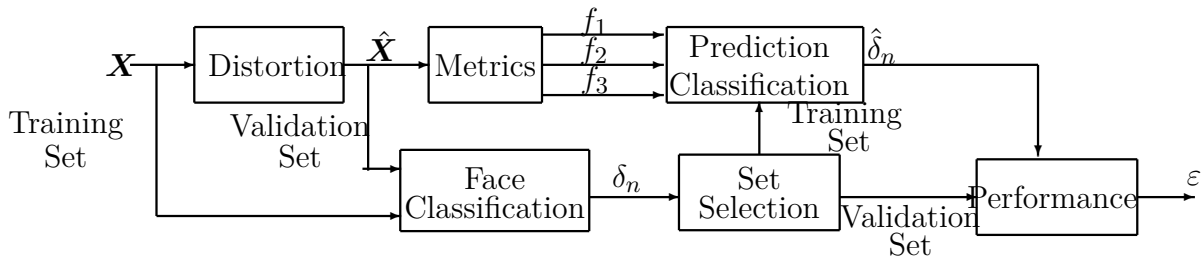


Figure 6-1.: Evaluation scheme

- Let  $\{\mathbf{X}\}_{i=1}^N$  be the set of  $N$  images that belong to the *Carnegie Mellon University* (CMU) Face Images Data Set. This database can be obtained in [55].
- In the distortion stage, all images inside  $\{\mathbf{X}\}$  are distorted using a selected process for different distortion levels, which is adjusted perceptually in order to give three observations: Imperceptible, perceptible and disturbing distortion. The result is the set  $\{\widetilde{\mathbf{X}}\}$ .
- In the classification stage, the classifier is trained using  $\{\mathbf{X}\}$ . Thereafter,  $\{\widetilde{\mathbf{X}}\}$  is used as validation sub-set. The outcome of this classification stage is  $\delta_i = \{0, 1\}$ , where  $\delta_i = 1$  means that  $i$ -th image is suitable to be classified (TRUE statement), and  $\delta_i = 0$  means the opposite (FALSE statement).  $\delta_i = 0$  is expected to correspond to each image  $\widetilde{\mathbf{X}}_i$  that stems from a disturbing distortion. This is not, however, true for all images and distortions.
- In the *Measure* stage, wavelet-based feature measures  $d_{1,i}, d_{2,i}, d_{3,i}$  are used to measure the distortion of each image  $\widetilde{\mathbf{X}}_i$ , using its corresponding original image  $\mathbf{X}_i$  as reference.
- In the *prediction classification* process, a classical Back-propagation Neural Network is employed, using feature measures as training input vector  $\mathbf{d}_i = \{d_{1,i}, d_{2,i}, d_{3,i}\}$  and  $\delta_i$  statements as training outputs. However, a percentage of total number of input vectors is used as training sub-set and the rest is used as validation sub-set. For the validation vector  $\mathbf{d}_i$ , expected output is  $\widetilde{\delta}_i$ . Images are scrambled 10 times to get different training and validation sub-sets.
- Finally, prediction error  $\varepsilon$  is calculated by the percentage of events when  $\widetilde{\delta}_i \neq \delta_i$ . As a total of 10 trials are used, mean and standard deviation values of 10 calculated prediction errors for each distortion are obtained.

Results show that classification error is around 8% for distortion processes in question (Table 6-1). This provides some confidence in terms of the use of perceptual measures for classification purposes, as their capability to give information related with features. A similar experiment, using feature measures as predictor, is suggested for visualizing whether some improvement is shown. This suggestion relies on the fact that predictor performance varies with test data, and more sensitivity in measurement can be achieved with the usage of measures. However, an increase in classification error is cautioned.

### 6.2.2. Evaluation on Texture Wearing Scores

One of the most important parameters for the classification of the quality of floor coverings is the conservation of appearance after an accelerated wear process [59]. To assess their aspect preservation, textile floor coverings are compared with previously rated reference samples.

**Table 6-1.:** Classification error of face images distorted by different processes. Eigen-faces classifier proposed in [74] and Back-propagation Neural Network predictor are used. Database can be obtained in [55]

Distortion Process	$\varepsilon$	
	Mean	Std. Dev.
Fast Fading	0.08	0.01
White Noise	0.08	0.01
JPEG	0.08	0.01
JPEG2000	0.08	0.02
Blurring	<b>0.07</b>	<b>0.01</b>

As the physical features of these reference samples may change over time, a normalization committee decided to propose certified photographs by international committees instead. In the assessment process, the carpets are first subjected to accelerated mechanical wear to simulate traffic exposure. Consequently, a group of trained experts compare the original carpet with the carpet subjected to an accelerated mechanical wear. Experts evaluate the wear level from color appearance and 3D structure, attributing values between 1 and 5 to each carpet. Value 1 corresponds to a total loss of structure and value 5 is attributed to a carpet that did not undergo a structural change. This visual assessment lacks in reproducibility and the method requires at least three experts, which is not always possible, especially within small companies. The human evaluation assessment is also somewhat subjective and industry is very interested in converting these traditional standards to automated objective standards. However, no automated system exists yet to enable the labeling process. In order to bring an automatic tool, Maximal Module feature is inserted in *Edge Wav* method, which consist in the location of the highest position of the edges. The need for a multi-resolution evaluation relies on the fact that edges are optimally detected by using different scales, whereas classical edge detectors only use a single scale parameter for the whole image. Thus, the basic single scale edge detection combines information on the behavior of the edges across a set of different scales. The number of decomposition levels required for edge detection is determined by applying a Gaussian low pass filter with increasing standard deviation, termed  $\sigma$ , from 1 to 5 until separate fibers can no longer be distinguished in the image. Hence, a value of  $\sigma = 3$  is found experimentally. The edges are detected in same manner as is depicted in Section §3.2.4. In order to fit selected number of scales, lower resolution sub-bands are merged by using a Sobel filter in order to preserve the position of the edge in the image. To be consistent with the visual analysis, features involved with scoring, termed as  $\kappa$ , are

---

expected to change monotonically as well as to be clearly distinguished between consecutive scorings. In order to evaluate whether the EdgeWav method offers a more monotonic and higher discriminant representation of the  $\kappa$  values related to wear labels, the following tests are conducted: firstly, Spearman linear-rank correlation between the wear labels and the mean values of the associated  $\kappa$  are calculated; secondly, the number of Consecutive Wear labels that can be Statistically Discriminated (CWSD) is checked. This evaluation is stated using ANalysis Of VAriance tests (ANOVA). The results in [59] show that using the *EdgeWav* method to construct depth images improves the discrimination value between consecutive wear labels as well as the linear rank between  $\kappa$  values and wear labels significantly. The results also show that the method works well for seven of the eight carpet types evaluated. A Gabor-based, *Contourlet* or *Bandelet* filter is considered for improvement of *EdgeWav* method, in order to achieve good results for the each of the eight carpets.

# A. Findings about Feature Extraction using Quad-tree Decomposition

This Chapter is dedicated to some incursions to the calculation of metrics by using Quad-tree Decomposition. Though metrics did not achieve expected results, the usage of this decomposition in wavelet-based techniques as *Wedgelets* motivates to bring future works for the improvement of the calculation of this metrics.

## A.1. Quad-tree Decomposition

A Quad-tree decomposition allows the detection of features by a reduced number of block calculations, instead the regular block segmentation. *Quad-Tree Decomposition* is a technique that splits an image in successive homogeneous regions (Figure A-1), as it is tested in [17], whose homogeneity criteria is arbitrary defined by some threshold value. This process is resumed in Algorithm 1, where  $H$  is the criteria function for the local homogeneity of blocks. Following three criteria are considered: Min-Max difference, Variance and Coefficient of Variation.

---

**Algorithm 1** Quad-tree Decomposition

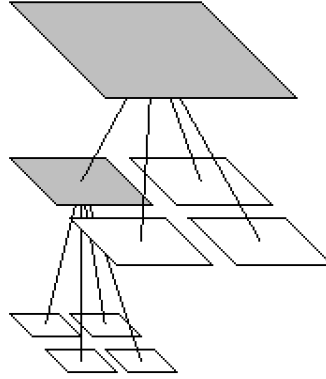
---

**Require:** Threshold value  $T$ , minimal dyadic block size  $\min_b\{\mathbf{B}\}$ , image size  $M_1M_2$

- 1:  $j = 0$
  - 2:  $b_j = M_1M_2$
  - 3: **while**  $H(\mathbf{B}_j) > T$  &  $b_j > \min_b\{\mathbf{B}\}$  **do**
  - 4:   Split  $\mathbf{B}_j$  in four children blocks  $\mathbf{B}_{j+1}$  of size  $b_{j+1} = b_j/2$
  - 5:   Calculate  $H(\mathbf{B}_{j+1})$
  - 6:    $j = j + 1$
  - 7: **end while**
- 

Let  $\mathbf{B}_i$  be some dyadic square, and let  $T$  be the splitting threshold value. Thus, square  $\mathbf{B}_i$  is split when one of the following conditions met, depending of selected criteria:

- For Min-Max difference case:  $|\min_x\{\mathbf{B}\} - \max_x\{\mathbf{B}\}| > T$ ,
- For Variance case:  $\sigma_{\mathbf{B}}^2 > T$ ,



**Figure A-1.:** Quad-tree Decomposition

- For Coefficient of Variation case:  $c_v = \sigma_{\mathbf{B}}/|\mu_{\mathbf{B}}| > T$ ,

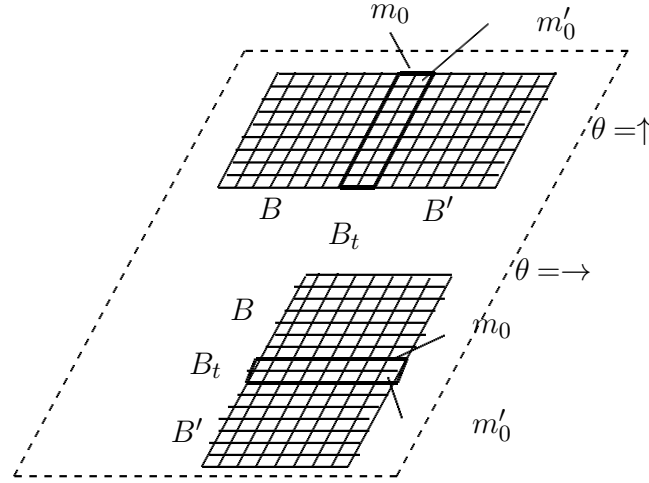
where  $\mu_{\mathbf{B}}, \sigma_{\mathbf{B}}, \sigma_{\mathbf{B}}^2$  denote respectively the non-zero mean, standard deviation and variance of elements into dyadic square  $\mathbf{B}$ . The advantage that the Coefficient of Variation case has over the other conditions, is that the calculated value is a ratio relationship. It leads to generalized threshold values, independent of the statistical parameters of evaluated image. For instance, values higher than one express a high-variable square. Therefore, normally a threshold value of  $T = 0.5$  is stated for a block that achieves a good homogeneity condition. For the other cases, threshold value selection is highly dependent of statistical features of each image. Besides, a size limit  $\min_b \mathbf{B}$  is put for stop criteria.

## A.2. Blockiness calculation by Quad-tree Decomposition

The main component required for this metric is a structured block division of image. It makes Quad-tree Decomposition very suitable for calculation, because each Dyadic square  $\mathbf{B}_i$  is defined by position  $\mathbf{m}_i$  and size  $b_i$  inside Quad-tree structure. Therefore, transition conditions  $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$  that are defined in Section §3.2.3 can be evaluated for each block inside Quad-tree Decomposition, as it is shown in Figure A-2. However, in the Quad-tree case, block size can be variable, depending on the splitting level  $j$ . It can be also interpreted as the block *sub-band*, accordingly with the wavelet-based dyadic partition. Thus, each block required for evaluation is located at position  $\mathbf{m}_i$ , with size  $b_j = 2^{-j} M_1 M_2$ .

Assuming that *Blockiness* is conventionally calculated for  $8 \times 8$ -sized regular blocks [30], this value is used as minimal size  $\min_b \{\mathbf{B}\}$  for the stop criteria. Thus, the estimated image  $\widetilde{\mathbf{X}}$  is Quad-tree decomposed, resulting in a set of dyadic blocks  $\mathbf{B}_i$  of size  $b_{i,j}$  for each Dyadic sub-band  $j$ . Therefore, *Blockiness* is computed as follows:

- Calculate the Quad-tree Decomposition structure for reconstructed image  $\widetilde{\mathbf{X}}$ , using



**Figure A-2.:** Block transition and its parameters. Two transition blocks of size  $b = 8$ , at respective vertical  $\uparrow$  and horizontal  $\rightarrow$  directions are shown.

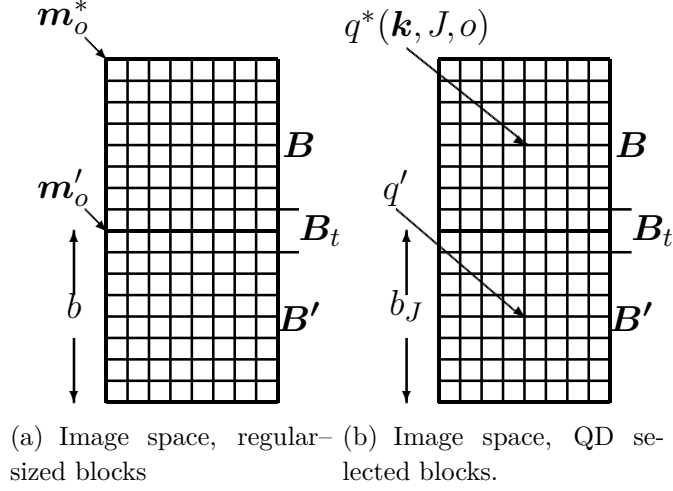
considered threshold criteria. It gives the set of dyadic blocks  $\{\mathbf{B}_i\}$ , with index  $(m_i, b_j)$ , where  $j$  is the dyadic sub-band.

- For each Dyadic square  $\mathbf{B}$ , a square of same size  $\mathbf{B}'$  is constructed at index  $(m_i + b_j, b_j)$ , for each orientation index  $\theta$ .
- Calculate  $\sigma_{\mathbf{B}_t}, \sigma_{\mathbf{B}, \mathbf{B}'}$ .
- Calculate Equations (3-5a), (3-5b) for each dyadic block, providing that transition conditions  $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$  met.
- Continue with the remaining blocks in order to find the metric  $d_{1,x}$  of Equation (3-6).

### A.3. Considerations on Visual Impairment metric using Quad-tree Decomposition

Corresponding with Blockiness calculation, Quad-tree Decomposition is also used for Visual impairment feature metric, avoiding the use of regular blocks. Moreover, a value of  $\min_b \{\mathbf{B}\} = 8 \times 8$  size is used as the stop parameter. Quad-tree Decomposition is then stated for each CSF-filtered image  $\mathbf{X}'$  and  $\tilde{\mathbf{X}}'$ , source and reconstructed respectively, giving respective dyadic blocks  $\{\mathbf{B}_i\}, \{\tilde{\mathbf{B}}_i\}$  of index  $(\mathbf{m}_i, b_j)$ . However, index  $(m_i, b_j)$  is derived only from source image, and index for reconstructed image must be adjusted accordingly. Assuming  $\mu_{\mathbf{B}}(i)$  and  $\mu_{\tilde{\mathbf{B}}}(i)$  as the respective local mean of source and reconstructed dyadic block at source index  $i$ , the luminance error for each block is calculated as:

$$e_{lum}(i) = \frac{|\mu_{\mathbf{B}}(i) - \mu_{\tilde{\mathbf{B}}}(i)|}{\text{gvrangle}}, \quad (\text{A-1})$$



**Figure A-3.:** Equivalence of parameters considered for Quad-tree Decomposition-based *Blockiness* metric

where  $g_{\text{vrange}}$  denotes a normalization factor, which is usually the maximal range of Grey values. The correlation factor for each dyadic block is defined as:

$$\rho(i) = \frac{\sigma_{\mathbf{B}, \tilde{\mathbf{B}}}(i)}{\sqrt{\sigma_{\mathbf{B}, \mathbf{B}}(i) \sigma_{\tilde{\mathbf{B}}, \tilde{\mathbf{B}}}(i)}},$$

and the similarity distortion factor is defined as

$$d_{sim}(i) = \frac{\sigma_{\mathbf{B}_e, \mathbf{B}_e}}{\sigma_{\mathbf{B}, \mathbf{B}}(i) + \sigma_{\tilde{\mathbf{B}}, \tilde{\mathbf{B}}}(i)} \cdot \frac{1 - \rho(i)}{2}, \quad (\text{A-2})$$

where  $\mathbf{B}_e(i) = \mathbf{B}_i - \tilde{\mathbf{B}}_i$ . The linear combination of Equations (A-1) and (A-2), is used as:

$$\varepsilon(i) = \frac{d_{sim}(i) + e_{lum}(i)}{2}$$

Finally, Quad-tree Decomposition-based Visual Impairment metric is calculated as expressed in Equation (3-10).

## A.4. Results on Blockiness and Visual Impairment calculated by Quad-tree Decomposition

The coefficient of determination between metrics and DMOS of Section §4.3 is used to visualize the correlation that the evaluation can achieve to human vision model by QD-based metrics. Though the improvements in  $R_{d, DMOS}^2$  is also evidenced for feature metrics in QD structure, when a WCSF method is used, their values are very low in comparison with

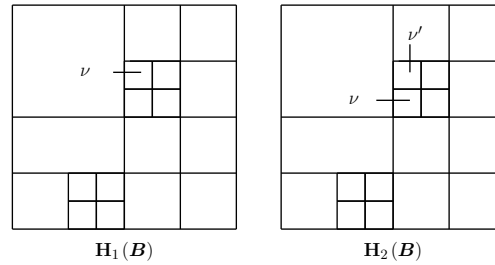
the feature metrics when the spatial-spectral CSF method is used (Table A-1). It evidences that feature metrics in QD space are sub-optimal for evaluation purposes. Hence, feature metrics involved with QD-QI are excluded from following evaluations.

	Method	$d_1$	$d_3$
DMOS	CSF	0.37	<b>0.31</b>
	WCSF	<b>0.55</b>	0.28
	AM-WCSF	0.29	0.21

**Table A-1.:** Mean coefficient of determination  $R_{d,DMOS}^2$  for metrics calculated using Quad-tree Decomposition method. Images are pre-processed by using either CSF, WCSF or AM-WCSF filters.

## A.5. Variation Coefficient-based Entropy

Quad-tree Decomposition is also used for the calculation of entropy measure. In this case coefficient of variation criteria is used.



**Figure A-4.:** Calculation of occurrence elements by Quadtree Decomposition

The variation coefficient values of same sized blocks  $B$  and  $B'$ , denoted as  $C_\nu$  and  $C'_\nu$ , respectively, can be generated by QD as shown in Figure A-4). The advantage of using those dyadic squares relies on the fact that the number of occurrences are reduced, because only the occurrences that fit the homogeneity condition  $c_\nu > 1$  are considered. It locates the neighbor nearby the element, reducing the number of neighbor occurrences by 2 for each pixel. Moreover, because the size of  $B$  is higher than one (size of pixel element), the number of occurrences per image also decreases, e.g., instead of computing four times the single pixel element occurrence the 4 pixel element calculation is once accomplished. Thus, for those

cases when it holds that  $C_v = \nu$  and  $C'_v = \nu'$  respective values of  $\mathbf{H}_h$ , are estimated as follows:

$$\mathbf{H}_1(\mathbf{B}) = \sum_{\nu=0}^1 P_\nu \log_2 P_\nu$$

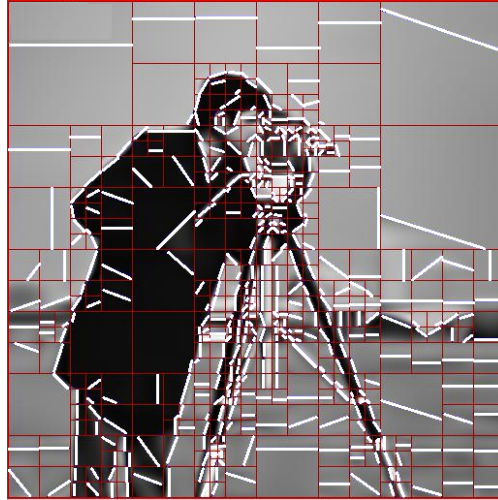
$$\mathbf{H}_2(\mathbf{B}) = \sum_{\nu=0}^1 P_{\nu'} \sum_{\nu'=0}^1 P(\nu|\nu') \log_2 P(\nu|\nu')$$

## A.6. Future works on Wedges

The future motivation for a finding of perceptual features using Quad-tree Decomposition is based on a line recognition technique that uses Wedgelets, which is the collection of functions that expresses all ways of splitting a square  $\mathbf{B}$  into two pieces or splits, including the special case of not splitting at all [20]. Assuming that each *Dyadic* square  $\mathbf{B} \subset \Omega$  is a set of the form with  $0 \leq n \leq N$  and  $0 \leq i, j < 2^{N-n}$ , a digital line or edge through the dyadic partition in direction  $\alpha$  is then defined for some Wedgelet as:

$$L_\alpha^n := \left\{ x \in Z^2 : \left( n - \frac{1}{2} \right) \delta_\alpha < \langle x, v_\alpha^\perp \rangle \leq \left( n + \frac{1}{2} \right) \delta_\alpha \right\} \quad (\text{A-3})$$

Then, using a *Sobel* filter, image can be segmented by dyadic squares containing a wedge. Figure A-5 shows an example on cameraman image, with  $8 \times 8$  pixel threshold block.



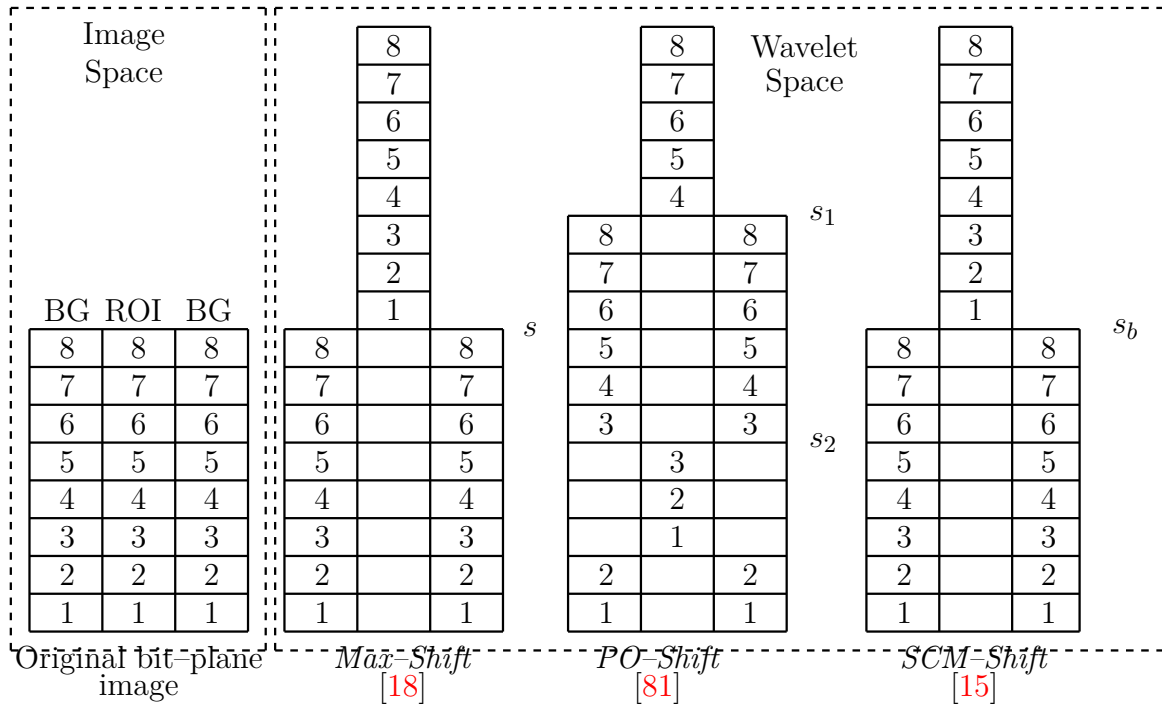
**Figure A-5.:** *Wedges* detected for cameraman image, highlighted in white lines.

## B. Region-of-Interest Coding

### B.1. Region-of-Interest Coding Methods

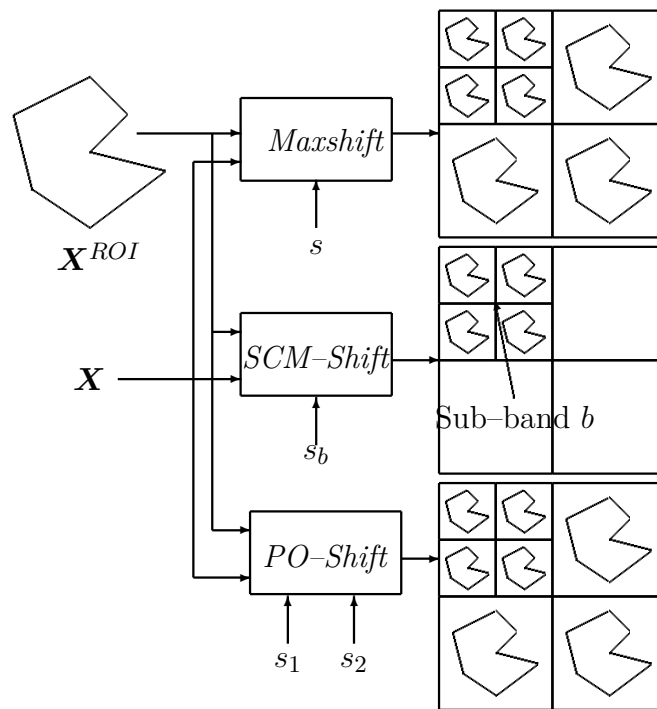
To avoid the distortion on ROI, two different strategies are used to make it relevant in order to be correctly identified further at decoding stage [18]: Firstly, in *Max-Shift*-based strategy, ROI elements are binary-up-shifted to the higher bit-planes so that they become a value above the maximal one of any coefficient of background (BG). Thus, only scaling value, i.e., the value of the bit-plane where the Most Significant Bit (MSB) of BG is located, is required at decoding stage. It makes coefficients that are located at bit-planes higher than scaling value clearly to be identified as ROI at decoding stage. Lastly, for *General-scaling*-based strategy, ROI elements are also up-shifted but by scaling value that corresponds to an arbitrary-selected bit-plane. Thus, the quality priority for ROI elements is stated by using an explicit remark, named *Scaling Value*  $s$ . Thus, ROI elements are shifted  $s$  bit-planes. For instance, for *Max-Shift* method, ROI elements are shifted by the maximal bit-plane of BG, as it is shown in Fig. B-1. In *Generalized Bitplane-by-Bitplane Shift* [76] and *Selective Coefficient Mask Shift* (SCM-Shift) coding methods [71], only ROI elements of the coarsest  $b$  sub-bands are preserved, thus shifting the rest of coefficients by  $s_b$  value, as it is shown in Fig. B-2. Those methods achieve both control over the relative importance of ROI and BG, and encoding of different regions of interest for a given set of arbitrary selected qualities. A better approach is the *Perceptually Optimized Bitplane Realignment* (PO-Shift) coding method that differs from *Max-Shift* in two aspects [81]: Firstly, it ranks the bit-planes by some perceptual-optimized order. Secondly, it uses two scaling values ( $s_1, s_2$ ) depending on ROI priority. These ones provide a numerical of functional control parameter to reduce the quantity of distortion in encoding framework. These methods for the bit-scaling of elements are shown in Fig. B-1, where *Max-Shift* method makes ROI-related image information to be encoded with priority over the rest, whereas *PO-Shift* and *SCM-Shift* approaches interlace the bit-planes for coding ROI and BG information. It is illustrated in Fig. B-1, where the central column represents the part of the data that have been identified as relevant for ROI coding. In such cases, the ROI input is required.

Intentionally, *SCM-Shift* algorithm is modified to diminish the quality of ROI from certain number of decomposition levels. Therefore, for this method, from  $J = 2$  to  $J = 4$  decomposition levels, ROI elements are not quantized for all sub-bands. Besides, from  $J = 5$  to  $J = 8$  decomposition levels, ROI elements are not quantized for the 3 coarsest sub-bands, letting conventional quantization of ROI elements at the finer ones. The intention of this experi-



**Figure B-1.:** Bit-plane scaling for considered *Max-Shift*-based ROI-coding methods.

ment is demonstrate that FW-QI can give a perceptual evaluation that might be in capacity to discriminate the coding methods that affect ROI form those that does not. Specifically, numerical results of FW-QI would show a great distance between *Max-Shift* and *PO-Shift* coding methods, because the later is a perceptual-based coder. Thus, different scaling factors  $s$  are used for the number of bit-planes required to scale (up or down) each element. For instance, in *Max-Shift* method,  $s$  corresponds to the base-2-logarithm of the maximal value of  $\mathbf{C}^{BG}$ , i.e., the maximal bit-plane of BG ( $s = 8$  in this case). For the *PO-Shift* method, the bit-planes are ranked from some perceptual evaluation, and scaling factors  $s_1$  and  $s_2$  are selected for respective  $\mathbf{C}^{ROI}$  and  $\mathbf{C}^{BG}$  by the same evaluation ( $s_1 = 5$  and  $s_2 = 3$  in this case). For *SCM-Shift* method, wavelet coefficients are scaled-up by same factor  $s_b$  at sub-bands from coarser to  $b$ . Thus, distortion on each encoded image is measured by each considered metric.



**Figure B-2.:** Mappings for considered *Max-Shift*-based ROI-coding methods.

# Bibliography

- [1] ABDON, Ikram E. ; DUSAUSOY, Nicolas J.: Survey of image quality measurements., ACM Fall Joint Computer Conference, 1986. – ISBN 0–8186–4743–4
- [2] Kap. Ondelettes et turbulence In: ABRY, P.: *Multirésolutions, algorithmes de décomposition, invariance d'échelles*. Diderot, 1997
- [3] AJA-FERNÁNDEZ, Santiago ; SAN-JOSÉ, Raúl ; ALBEROLA-LÓPEZ, Carlos ; WESTIN, Carl-Fredrick: Image Quality Assesment based on Local Variance, 28th IEEE EMBS Annual International Conference, August 2006
- [4] AKHTAR, Pervez ; BHATTI, Muhammad I. ; ALI, Tariq J. ; MUQEET, Muhammad A.: Significance of ROI Coding using MAXSHIFT Scaling applied on MRI Images in Teleradiology- Telemedicine. En: *SciRP Journal of Biomedical Science and Engineering* 1 (2008), p. 110–115
- [5] ANDRÉ, Thomas ; ANTONINI, Marc ; BARLAUD, Michael ; GRAY, Robert M.: Entropy-based distortion measure and bit allocation for wavelet image compression. En: *IEEE Transactions on Image Processing* 16 (2007), Nr. 12, p. 3058–3064
- [6] ASHIDA, Hiroshi: Action-specific extrapolation of target motion in human visual system. En: *Elsevier Neuropsychologia* (2004), Nr. 42, p. 1515–1524
- [7] ATSUMI, E. ; FARVADIN, N.: Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees, IEEE International Conference on Image Processing ICIP, 1998
- [8] AVCIBAS, Ismail ; MEMON, Nasir ; SANKUR, Bülent: Steganalysis Using Image Quality Metrics. En: *IEEE Transactions on Image Processing* 12 (2003), Nr. 2, p. 221–229
- [9] AVCIBAS, Ismail ; SANKUR, Bülent ; SAYOOD, Khalid: Statistical evaluation of image quality measures. En: *SPIE Journal of Electronic Imaging* 11 (2002), Nr. 2, p. 206–223
- [10] BALAS, Benjamin: Texture synthesis and perception: Using computational models to study texture representations in the human visual system. En: *Elsevier Vision Research* (2006), Nr. 46, p. 299–309

- 
- [11] BARTRINA-RAPESTA, Joan ; SERRA-SAGRISTA, Joan ; AULI-LLINAS, Francesc ; MUNOZ GOMEZ, Juan: JPEG2000 ROI coding method with perfect fine-grain accuracy and lossless recovery. En: *Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers* , p. 558–562. ISBN 978–1–4244–5825–7
- [12] BEEGAN, A. P. ; IYER, L. R. ; BELL, A. E. ; MAHER, V. R. ; ROSS, M. A.: Design and evaluation of perceptual masks for wavelet image compression, IEEE 10th Digital Signal Processing Workshop, 2002
- [13] BENNETT, K. P. ; EMBRECHTS, M. J.: An Optimization Perspective on Kernel Partial Least Squares Regression. En: *Advances in Learning Theory: Methods, Models and Applications*. 190 (2003), p. 227–250
- [14] BOVIK, Alan C. *Laboratory for Image & Video Engineering*. 2009
- [15] CHEN, Yao-Tien ; TSENG, Din-Chang ; CHANG, Pao-Chi: Wavelet-Based Image Compression with Polygon-Shaped Region of Interest. En: *Lecture Notes in Computer Science, Advances in Image and Video Technology* (2006)
- [16] CHEN, Ying ; HAO, Pengwei: Integer reversible transformation to make JPEG Lossless, International Conference on Signal Processing ICSP, 2004
- [17] CHINGA, G.: A quadtree decomposition approach for surface assessment. En: *Springer Journal on Pattern Analysis and Applications* 9 (2006), Nr. 1
- [18] CHRISTOPOULOS, C.A. ; EBRAHIMI, T. ; SKODAS, A.N.: JPEG2000: The new Still Picture Compression Standard, ACM workshops on Multimedia, 2000
- [19] REPÚBLICA DE COLOMBIA, Imprenta N.: Diario Oficial / Ministerio de la Protección Social. 2004. – Informe de Investigación. – ISSN 0122–2112
- [20] DONOHO, David L.: Wedgelets: Nearly-Minimax Estimation of Edges / Stanford University, U.C. Berkeley. 1997. – Informe de Investigación
- [21] DOSSELMANN, Richard ; YANG, Xue D.: Existing and emerging Image Quality Metrics. En: *IEEE Canadian Conference on Electrical and Computer Engineering* IEEE, 2005, p. 1906–1913
- [22] DUCOTTET, C. ; FOURNEL, T. ; BARAT, C.: Scale-adaptive detection and local characterization of edges based on wavelet transform. En: *Elsevier Journal on Signal Processing* 1 (2004), Nr. 84, p. 2115–2137
- [23] DUIJS, Remco ; FELSBURG, Michael ; GRANLUND, Gösta ; HAAR-ROMENY, Bartter: Image Analysis and Reconstruction using a Wavelet Transform Constructed from a Reducible Representation of the Euclidean Motion Group. En: *Springer International Journal of Computer Vision* 72 (2007), Nr. 1, p. 79–102

- [24] ELLINAS, John N. ; KENTERLIS, Panagiotis: A Wavelet-Based Watermarking method Exploiting the Contrast Sensitivity Function, INSTICC International Conference on Computer Vision Theory and Applications VISAPP, 2007
- [25] GADDIPATI, A. ; MACHIRAJU, R. ; YAGEL, R.: Steering Image Generation with Wavelet Based Perceptual Metric. En: *Eurographics Journal* (1997)
- [26] GARCÍA-ÁLVAREZ, Julio-César: Telemedicine integrated-services perspectives. En: *Revista del Congreso Internacional de Ingeniería Electrónica* 1 (2007), Nr. 1, p. 1–4. ISBN 978–958–708–309–5
- [27] GARCÍA-ÁLVAREZ, Julio-César ; FÜHR, Hartmut ; ORTIZ-JARAMILLO, Benhur ; CASTELLANOS-DOMÍNGUEZ, Germán: Image Distortion Measure Mappings in Wavelet Representation Space, International Conference on BroadBand Communications, Information Technology and Biomedical Applications BroadBandComm2009, 2009
- [28] GARCIA-ALVAREZ, Julio-Cesar ; CASTELLANOS-DOMINGUEZ, Cesar-German ; ORTIZ-JARAMILLO, Benhur: Image information access using Wedgelet filters, IEEE International Symposium on Applied Sciences in Biomedical and Communication Technologies ISABEL, October 2008, p. 1–4
- [29] GARCIA-ALVAREZ, Julio-César ; CASTELLANOS-DOMINGUEZ, German: Estimation of image compression rates by means of full-referenced distortion metrics (english abstract). En: *Ingenieria y Competitividad* 10 (2008), Nr. 1, p. 73–83
- [30] GINESU, Giaime ; MASSIDA, Francesco ; GIUSTO, Daniele: A multi-factors approach for image quality assessment based on a human visual system model. En: *Elsevier Journal on Signal Processing: Image Communication* (2006), Nr. 21, p. 316–333
- [31] GONZALES, Rafael C. ; WOODS, Richard E. ; EDDINS, Steven L.: *Digital Image Processing Using MATLAB*. Upper Saddle River, NJ, USA : Prentice-Hall, Inc., 2003. – ISBN 0130085197
- [32] HEKSTRA, A.P ; BEERENDS, J.G ; LEDERMANN, D. ; DE CALUWE, F.E ; KOHLER, S. ; KOENEN, R.H. ; RIHS, S. ; M.EHR SAM ; SCHLAUSS, D.: PVQM - A perceptual video quality measure. En: *Elsevier Journal on Signal Processing: Image Communication* (2002), Nr. 17, p. 781–798
- [33] HORITA, Yuukou. *Media Information and Communication Technology Laboratory*. 2009
- [34] IBARRA, Oscar ; ORTIZ JARAMILLO, Benhur ; REINOSA, J. ; HENAO, Vladimir ; GARCÍA-ALVAREZ, Julio-César: Evaluación de la Calidad del Servicio en un Sistema Integrado Duo-Play de Ludoteca Multimedia, IEEE Colombia Conference on Communications, 2007

- 
- [35] INSTITUTE, Colombia National C.: Anuario Estadístico 2005 / Instituto Nacional de Cancerología E.S.E. 2005-2007. – Informe de Investigación
- [36] ITU-R: Recommendation BT.500-11 Methodology for the subjective assessment of the quality of television pictures. (2002)
- [37] KAUFHOLD, John ; HOOGS, Anthony: Learning to Segment Images Using Region-Based Perceptual Features, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004
- [38] KOSHELEVA, Olga M. ; CABRERA, Sergio D.: Application of task-specific metric in JPEG2000 ROI compression, IEEE Southwest Symposium on Image Analysis and Interpretation, 2002
- [39] LE CALLET, Patrick ; AUTRUSSEAU, Florent: *Subjective quality assessment IRC-CyN/IVC database*. 2005. – <http://www.irccyn.ec-nantes.fr/ivcdb/>
- [40] LI, Jun: *A Wavelet Approach to Edge Detection*, Sam Houston State University, Tesis de Grado, 2003
- [41] LI, Xin: Image Resolution Enhancement via Data-Driven Parametric Models in the Wavelet Space, EURASIP Journal on Image and Video Processing, 2007
- [42] LI, Xuelong ; TAO, Dacheng ; GAO, Xinbo ; LU, Wen: A natural image quality evaluation metric. En: *Elsevier Journal on Signal Processing* (2009), Nr. 89, p. 548–555
- [43] LIU, Zhen ; KARAM, Lina J. ; WATSON, Andrew B.: JPEG2000 encoding with perceptual distortion control. En: *IEEE Transactions On Image Processing* , Juli, Nr. 7, p. 1763–78. – ISSN 1057–7149
- [44] LO, Chien-Shun ; CHUNG, Pau-Choo ; LEE, Sn-Kan ; HSU, Giu-Cheng: Fractal Based JPEG2000 ROI Coding, IEEE International Conference on Systems, Man, and Cybernetics, 2006, p. 553–555
- [45] Kap. A Universal PCA for Image Compression In: LV, Chuanfeng ; ZHAO, Qiangfu: *Lecture Notes in Computer Science*. Vol. 3824: *Embedded and Ubiquitous Computing - EUC 2005 Proceedings*. First. Springer, 2005, p. 910–919. – ISBN 978–3540308072
- [46] LV, Chuanfeng ; ZHAO, Qiangfu: A Simplified MPC for Image Compression., IEEE International Conference on Computer and Information Technology CIT, 2005, p. 580–584

- [47] MAHMOUDI-AZNAVEH, Ahmad ; MANSOURI, Azadeh ; TORKAMANI-AZAR, Farah ; ESLAMI, Mohammad: Image Quality Measurement Besides Distortion Type Classifying. En: *Springer Optical Review Journal* 16 (2009), Nr. 1, p. 30–34
- [48] MALLAT, S.: Wavelets for a vision. En: *Proceedings of the IEEE* 84 (1996), apr, Nr. 4, p. 604–614. – ISSN 0018–9219
- [49] MALLAT, S.: *A Wavelet Tour of Signal Processing*. Academic Press, 1998
- [50] MARCELLIN, Michael W. ; LEPLEY, Margaret A. ; BILGIN, Ali ; FLOHR, Thomas J. ; CHINEN, Troy T. ; KASNER, James H.: An overview of quantization in JPEG 2000. En: *Elsevier Journal on Signal Processing: Image Communication* (2002), Nr. 17, p. 73–84
- [51] MAYACHE, A. ; EUDE, T. ; CHERIFI, H.: A Comparison of Image Quality Models and Metrics Based on Human Visual Sensitivity, IEEE International Conference on Image Processing, 1998, p. 409–413
- [52] MAYER, Norbert ; HERRMANN, J. M. ; GEISEL, Theo: Curved feature metrics in models of visual cortex. En: *Elsevier Neurocomputing* 44 (2002), Nr. 46, p. 533–539
- [53] MEDINA-CARNICER, R. ; MUÑOZ-SALINAS, R. ; YEGUAS-BOLIVAR, E. ; DIAZ-MAS, L.: A novel method to look for the hysteresis thresholds for the Canny edge detector. En: *Elsevier Pattern Recognition* (2011), December, Nr. 44, p. 1201–1211
- [54] MÜHLICH, Mathias ; AACH, Til: Analysis of Multiple Orientations. En: *IEEE Transactions on Image Processing* 18 (2009), Nr. 7, p. 1424–1437
- [55] MITCHELL, Tom. *CMU Face Images*. 1994
- [56] MITCHELL, Tom: *Machine Learning*. McGraw Hill, 1997
- [57] MOCHIZUKI, Yoshihiko ; TORII, Akihiko ; IMIYA, Atsushi: N-Point Hough transform for line detection. En: *Elsevier Journal of Visual Communication and Image Representation* (2009), Nr. 20, p. 242–253
- [58] NABIL, Ahmed ; DUNN, Mark ; NUNKESSER, Marc ; SIDHOM, Sahbi: Image Compression using Hartley transform / Institute of Computer Aided Automation, Vienna University of Technology. 2003. – Informe de Investigación
- [59] ORJUELA-VARGAS, Sergio-Alejandro ; PIZURICA, Aleksandra ; GARCÍA-ÁLVAREZ, Julio-César ; ORTIZ-JARAMILLO, Benhur ; ROOMS, Filip ; DE MEULEMEESTER, Simon: Surface Reconstruction of Wear in Carpets by Using a Wavelet Edge Detector, Advanced Concepts For Intelligent Vision Systems. 12th International Conference on Advanced Concepts for Intelligent Vision Systems, 2010

- 
- [60] ORTIZ-JARAMILLO, Benhur ; GARCÍA-ALVAREZ, Julio-César ; CASTELLANOS-DOMINGUEZ, Germán: Translating distortion measures to a multi-resolution space for image analysis (english abstract). En: *Ingenieria y Competitividad* 12 (2010), Nr. 1, p. 43–55
- [61] PARK, Keun-Hyeong ; PARK, Hyun-Wook: Region-of-Interest Coding Based on Set Partitioning in Hierarchical Trees. En: *IEEE Transactions on Circuits and Systems for Video Technology* 12 (2002), Nr. 2, p. 106–113
- [62] PETKOV, N. ; A.WESTENBERG, M.: Suppression of contour perception by badlimited noise and its realltion to nonclassical receptive field inhibition. En: *Biological Cybernetics* 88 (2003), p. 236–246
- [63] PISANO, E. ; COLE, E. ; HEMMINGER, B. ; YAFFE, M. ; AYLWARD, S. ; KOPANS, D. ; BROWN, M. ; PIZER, S.: Image Processing Algorithms for Digital Mammography: A Pictorial Essay 1. En: *RSNA RadioGraphics* (2009), 02
- [64] PONOMARENKO, N. ; LUKIN, V. ; ZELENSKY, A. ; EGIAZARIAN, K. ; CARLI, M. ; BATTISTI, F.: TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics. En: *Advances of Modern Radioelectronics* 10 (2009), p. 30–45
- [65] RAMAKRISHNAN, B. ; SRIRAAM, N.: Compression of DICOM Images based on Wavelets and SPIHT for Telemedicine Applications, United Arab Emirates International Conference on Biological and Medical Physics, March 2005
- [66] RAMDSEN, B. M. ; HUNG, C. P. ; ROE, A. W.: Real and illusory contour processing in area v1 of the primate - a cortical balancing act. En: *Cerebral Cortex* 11 (2001), p. 648–665
- [67] SAMET, H.: Quadtree and Related Hierarchical Data Structures. En: *ACM Computers Survey* 16 (1984), Nr. 2, p. 187–260
- [68] SHEIKH, Hamid R. ; SABIR, Muhammad F. ; BOVIK, Alan C.: A Statistical Evaluation of Recent Full Reference Image Quality Aessment Algorithms. En: *IEEE Transactions on Image Processing* 15 (2006), Nr. 11, p. 3441–3452
- [69] SUCKLING, J. ; PARKER, J. ; DANCE, D. ; ASTLEY, S. ; HUTT, I. ; BOGGIS, C. ; RICKETTS, I. ; STAMATAKIS, E. ; CERNEAZ, N. ; KOK, S. ; TAYLOR, P. ; BETAL, D. ; SAVAGE, J.: The Mammographic Image Analysis Society Digital Mammogram Database. En: *Exerpta Medica, International Congress Series* 1069 (1994), p. 375–378
- [70] SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH ; SIGNAL AND INFORMATION PROCESSING LABORATORY: Polynomial Regression and Neural Networks / Swiss Federal Institute of Technology Zurich and Signal and Information Processing Laboratory. 2007. – Informe de Investigación

- [71] TAHOSES, Pablo G. ; VARELA, J. R. ; LADO, María J. ; SOTO, Miguel: Image compression: Maxshift ROI encoding options in JPEG2000. En: *Elsevier Journal on Computer Vision and Image Understanding* 109 (2008), Nr. 1, p. 139–145
- [72] TAUBMAN, David: High Performance Scalable Image Compression with ECBOT. En: *IEEE Transactions On Image Processing* 9 (2000), Nr. 7, p. 1158–1170
- [73] TAUBMAN, David ; PRANDOLINI, Robert: Architecture, philosophy and performance of JPIP: Internet protocol standard for JPEG2000, SPIE International Symposium on Visual Communications and Image Processing VCIP, July 2003, p. 649–663
- [74] TURK, Matthew ; PENTLAND, Alex: Face recognition using eigenfaces, IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR, July 1991, p. 586 – 591
- [75] WANG, Zhou: Objective Image/Video Quality Measurement - A Literature Survey / University of Texas. 1998. – Informe de Investigación
- [76] WANG, Zhou ; BANERJEE, Serene ; EVANS, Brian L. ; BOVIK, Alan C.: Generalized Bitplane-by-Bitplane Shift Method for JPEG2000 ROI Coding, IEEE International Conference on Image Processing, September 2002, p. 81 – 84
- [77] WANG, Zhou ; BOVIK, Alan C. ; SHEIKH, Hamid R. ; SIMINCELLI, Eero P.: Image Quality Assessment: From Error Visibility to Structural Similarity. En: *IEEE Transactions On Image Processing* 13 (2004), Nr. 4, p. 600–613
- [78] DER WEKEN, Dietrich V. ; NACHTHEAGEL, Mike ; KERRE, Etienne: Combining neighbourhood-based and histogram similarity measures for the design of image quality measures. En: *Elsevier Journal On Image and Vision Computing* (2007), Nr. 25, p. 184–195
- [79] YI, Yaohua ; YU, Xiaoqing ; WANG, Leiguang ; YANG, Zhigao: Image Quality Assessment Based on Structural Distortion and Image Definition. En: *International Conference on Computer Science and Software Engineering* (2008)
- [80] YOU, Kang S. ; LEE, Han J. ; KWAK, Hoon S.: Multiple-ROI Image Coding Method Using Maxshift over Low-Bandwidth. En: *Lecture Notes in Control and Information Sciences* 345 (2006), p. 334–342
- [81] ZHANG, Tan ; GU, Hai-Ming: Region-of-Interest Image Coding Based on Perceptually Optimized Bitplane Realignment. En: *IEEE International Conference on Electronic Computer Technology* IEEE, 2009, p. 495–498

- 
- [82] ZHANG, Yani ; PHAM, Binh ; ECKSTEIN, Miguel: Evaluation of JPEG2000 Encoder Options: Human and Model Observer Detection of Variable Signals in X-Ray Coronary Angiograms. En: *IEEE Transactions on Medical Imaging* 23 (2004), Nr. 5, p. 613–632
- [83] ZHU, Haijun ; WU, Huayi: New Paradigm for Compressed Image Quality Metric: Exploring Band Similarity with CSF and Mutual Information. En: *IEEE International Geoscience and Remote Sensing Symposium IGARSS* IEEE, 2005, p. 696–699
- [84] ZIEGLER, Gernot ; LENSCH, Hendrik P. ; MAGNOR, Marcus ; SEIDEL, Hans-Peter: Multi-Video Compression in Texture Space using 4D SPIHT. En: *IEEE Workshop on Multimedia Signal Processing* (2004), September, p. 39–42
- [85] ZYOUT, Imad ; ABDEL-QADER, Ikhlas ; AL-OTUM, Hazem: Embedded ROI Coding of Mammograms via Combined SPIHT and Integer Wavelet Transforms, *IEEE International Conference on Electro/Information Technology*, May 2007, p. 173–177

# Index

- Assessment
  - Image Quality, 22
- Coding
  - Lossless, 1
  - Lossy, 2
- Contrast Sensitivity Function, 9, 25
- Correlation, 40
  - Coefficient of Determination, 40, 41
- Distortion, 1
  - Block Artifact, 3
  - Blurring, 3
  - Delay, 1
  - Fading, 1
  - Noise, 1
- Evaluation Models
  - Error Sensitivity, 6
- Feature, 19
  - Detail Variance, 21
  - Edge, 19
  - Local Mean, 21
  - Maximal Module, 21
  - Transition, 19
- Image, 13
  - Bit-plane, 14
  - Medical, 13
  - Size, 13
- Image Coding, 16
- Measure
  - Full-referenced, 3
  - Objective, 3
    - Czekanowski, 3
    - Mean Squared Error, 3
    - Peak Signal to Noise Ratio, 3
  - Perceptual, 3
    - Blockiness, 27
    - Differential Mean Observed Score, 40
    - Edge Error, 27, 37
    - Entropy-based distortion, 8
    - Feature Measure, 27
    - Perceptual Mean Square Error, 8
    - Quality Index, 26
    - Quality Index based on Local Variance, 8
    - Reflection Factor, 8
    - Score, 4
    - Visual Impairment, 27, 38
  - Quality-of-Service, 4
- Quad-tree Decomposition, 76
- Quantization, 16
- Sensitivity, 40
- Transformations, 16
- Wavelet
  - Biorthogonal, 18
  - Contrast Sensitivity Function, 25
  - Measure, 38
  - Sub-band, 17
  - Wedgelet, 81