Multiscale Image Restoration in Nuclear Medicine

Vom Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades
eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Dissertation

Dipl.-Ing. **Ghada Jammal**
geboren am 26. Juli 1971
in Beirut, Libanon

Referent:	Prof. Dr.-Ing. H. Clausert
Korreferent:	Prof. Dr.-Ing. A. Bijaoui
Tag der Einreichung:	14. November 2000
Tag der mündlichen Prüfung:	04. Dezember 2000

D17
Darmstädter Dissertationen
2001

Acknowledgments

I would first like to thank my thesis advisor, Professor Dr.-Ing. H. Clausert for his confidence and his unstinting commitment to helping see this thesis to its final completion. I am very grateful to him for supporting me constructively the entire way and giving me the space to develop my own ideas.

A very unique thank goes to Professor Dr.-Ing. A. Bijaoui to whom I owe much of my development and growth over the last few years. I am extremely grateful to him for his equally generous and wise guidance during this thesis. His enthusiasm and his profound scientific intuition have been a constant inspiration.

Thanks are expressed to Professor Dr.-Ing. B. Dorsch, Professor Dr.rer.nat. Dr.h.c.mult. M. Glesner and Professor Dr.-Ing. T. Weiland for taking part in my thesis committee and for the interest they showed in my work.

I am indebted to Professor R. Bidet of the Hôpital Saint-Jacques in Besançon for introducing me to the field of nuclear medicine imaging.

I am very grateful to Dr.-Ing. Hervé Guillemin for the many fruitful discussions that initiated this work and for supplying the images used in this thesis.

I thank all my colleagues from the Institut für Nachrichtentechnik for providing a friendly and stimulating environment. Special thanks go to Dipl.-Ing. Olivier Ecabert and Dipl.-Ing. Michael Voigt. Lots of conversations with them on almost anything have made these years a worthwhile experience.

I am grateful to Dipl.-Ing. Laurent Bourdichon for devoting long hours to the design of the DeQuant logo.

Finally I would like to dedicate this work to my parents and my brother with all my affection, and to the one who has made with his love, support and unfailing encouragement this journey seem worthwhile after all, Rémy.

Darmstadt, February 2001 Ghada Jammal

Contents

Introduction and overview **1**

1. Physics in nuclear medicine and image model **5**

1.1. Radioactive decay 6

1.1.1. Binomial distribution 6

1.1.2. Poisson distribution 7

1.2. Interactions of particules with matter 7

1.2.1. Interactions of γ rays with matter 8

1.2.2. Interactions of electrons with matter 8

1.3. Gamma camera components 9

1.3.1. The collimator 9

1.3.2. The scintillation detector 12

1.3.3. Positioning and energy calculation electronics 16

1.4. Image degradation model 18

1.4.1. Interactions in the patient body 19

1.4.2. Degradation introduced by the camera 19

1.4.3. The degradation model 20

1.5. Conclusion 21

2. An introduction to wavelets 23

2.1. What are wavelets? 23

2.2. Continuous wavelet transform 25

2.3. Multiresolution analysis 27

2.3.1. One dimensional multiresolution analysis 27

2.3.2. Two dimensional multiresolution analysis 30

2.4. Discrete wavelet transform 32

2.4.1. One dimensional Mallat's algorithm 32

2.4.2. Two dimensional Mallat's algorithm 36

2.5. Shift-invariant wavelet transform 40

2.6. Conclusion 43

3. Some classical approaches to image restoration 45

3.1. General formulation of the image model 45

3.2. Restoration methods for stationary processes 47

3.2.1. The inverse filter and its variations 47

3.2.2. Least Mean Square Methods 49

3.3. Restoration methods for nonstationary processes 52

3.3.1. An example of adaptive method 52

3.3.2. Wavelet based methods 53

3.4. Conclusion 62

4. DeQuant: a flexible multiresolution restoration framework . 65

4.1. Denoising with DeQuant 67

4.1.1. Transformation 67

4.1.2. Selection 68
4.1.3. Regularization 73
4.1.4. Projection 75
4.1.5. Reconstruction 76
4.1.6. Iteration 77
4.2. Restoration with DeQuant 77
4.2.1. Regularization 78
4.2.2. Projection 79
4.2.3. Reconstruction 80
4.3. Shift-Invariant DeQuant 81
4.4. Edge-Preserving DeQuant 82
4.5. Conclusion 84

5. PDF of the wavelets coefficients of a Poisson process 87
5.1. The general case 87
5.1.1. PDF of the wavelet coefficient 88
5.1.2. Approximation for low densities of events 90
5.1.3. Approximation for high densities of events 90
5.2. Special case: the Haar wavelet 91
5.2.1. PDF of the wavelet coefficient 91
5.2.2. Approximation for low densities of events 93
5.2.3. Approximation for high densities of events 93
5.2.4. Numerical simulations 94
5.3. Conclusion 96

6. Application of DeQuant to nuclear medicine 97

6.1. Performance criteria 98

6.1.1. Global criteria 98

6.1.2. Local criteria 99

6.2. Performance of D-DeQuant 100

6.2.1. Contrast and size sensibility of D-DeQuant 101

6.2.2. Choice of D-DeQuant's parameters 103

6.3. Comparative studies on simulated data experiments 114

6.3.1. Introduction of the point spread function 114

6.3.2. Introduction of the edge preserving feature 116

6.3.3. Introduction of the shift invariant feature 119

6.3.4. Comparison of the eight different DeQuant algorithms . . 120

6.4. Thyroid phantom studies 129

6.5. Actual thyroid clinical images 131

6.6. Conclusion 135

7. DeQuant's Graphical User Interface 137

7.1. General structure of the GUI windows 137

7.2. Title Window 138

7.3. Nodule Generation Manager 139

7.3.1. Display 139

7.3.2. Uicontrols 139

7.3.3. Uimenus 141

7.4. PSF Generation Manager 142

7.5. Restoration Manager 142

7.5.1. Display 143
7.5.2. UiControls 143
7.5.3. Uimenus 145
7.6. Measure Lab 146
7.6.1. Display 146
7.6.2. Uicontrols 147
7.6.3. Uimenus 148
7.7. Conclusion 148

Summary and perspectives 149

A. DeQuant Algorithms 153
A.1. D-DeQuant 154
A.2. SI-D-DeQuant 155
A.3. EP-D-DeQuant 156
A.4. SI-EP-D-DeQuant 157
A.5. Simple derivation of the D-DeQuant variations 158
A.6. R-DeQuant 159
A.7. SI-R-DeQuant 160
A.8. EP-R-DeQuant 161
A.9. SI-EP-R-DeQuant 162
A.10. Simple derivation of the R-DeQuant variations 163

B. An algorithm to determine image edges 165
B.1. The continuous case 165
B.2. The discrete case 166

C. Proofs and tables of Chapter 5 . . . . . **169**

C.1. Proof of Proposition 1 169

C.2. Proof of Proposition 2 173

C.3. Tables 175

List of notations **183**

Publications **193**

Bibliography **195**

Abstract

This work develops, analyzes and validates a new multiscale restoration framework for denoising and deconvolution in photon limited imagery. Denoising means the estimation of the intensity of a Poisson process from a single observation of the counts, whereas deconvolution refers to the recovery of an object related through a linear system of equations to the intensity function of the Poisson data. The developed framework has been named DeQuant in analogy to **De**noising when the noise is of **Quant**um nature.

DeQuant works according to the following scheme. (1) It starts by testing the statistical significance of the wavelet coefficients of the Poisson process, based on the knowledge of their probability density function. (2) A regularization constraint assigns a new value to the non significant coefficients enabling therewith to reduce artifacts and incorporate realistic prior information into the estimation process. Finally, (3) the application of the inverse wavelet transform yields the restored object. The whole procedure is iterated before obtaining the final estimate.

The validation of DeQuant on nuclear medicine images showed excellent results. The obtained estimates enable a greater diagnostic confidence in clinical nuclear medicine since they give the physician the access to the diagnosis relevant information with a measure of the significance of the detected structures.

Zusammenfassung

Diese Arbeit entwickelt, untersucht und bewertet eine neue multiskalen Restaurationsmethode für die Rauschunterdrückung und Entfaltung im Bereich der bildgebenden Verfahren mit schwachen Photonenstrahlung. Hierbei bedeutet Rauschunterdrückung die Schätzung der Intensität eines Poisson Prozesses ausgehend von einer einzigen Realisierung. Die Entfaltung bezieht sich dabei auf die Zurückgewinnung eines Objektes, das über ein lineares Gleichungssystem mit der Intensität des Poisson-Prozesses verbunden ist. Die entwickelte Methode wurde DeQuant genannt in Anlehnung an das englishe Wort **De**noising, bei dem das **Quant**enrauschen unterdrückt wird.

DeQuant funktionniert nach folgendem Schema: (1) Im ersten Schritt findet eine statistische Auswertung der Wavelet Koeffizienten des Poisson-Prozesses statt, basierend auf deren Wahscheinlichkeitsdichte. (2) Danach werden im Rahmen eines Regularisierungsschrittes den nichsignifikanten Koeffizienten neue Werte zugewiesen. Dadurch werden Artefakte unterdrückt und a priori Information innerhalb des Schätzprozesses berücksichtigt. (3) Durch eine Abschließende Wavelet Transformation wird das restaurierte Bild gewonnen. Dieses Vorgang wird mehrmals durchlaufen bis man das entgültige Bild erhält.

Die Bewertung von DeQuant mit nuklearmedizinischen bildern lieferte sehr gute Ergebnisse. Die erzielten Resultate verbessern die Diagnostik-Möglichkeiten erheblich, da sie dem Arzt Zugang zu wertvollen Diagnoseinformationen ermöglichen. Gleichzeitig erhält er ein Maß für die Signifikanz der detektierten Strukturen.

Introduction and overview

Nuclear medicine procedures are widely used for diagnostic purposes. They are unique in the sense that they document the body's biochemistry or function, in contrast to traditional diagnostic techniques such as x-ray or ultrasound that produce images of the body's anatomy or structure. Since alteration of biochemical processes may occur before there is a change in anatomy, these procedures allow a disease to be detected and treated early in its course when there may be a more successful prognosis.

The purpose of nuclear medicine procedures is to obtain a picture of the distribution of a radioactive pharmaceutical after it has been administered to the patient. The raw nuclear medicine image is acquired by detecting gamma ray photons that are emitted from within the body as the radioactive pharmaceutical decays. The collected data follows a Poisson distribution whose intensity function is the pharmaceutical's distribution. Hence the recovery of the diagnostic relevant information involves the estimation of the intensity of a Poisson process from a single realization of the process.

Solving this estimation problem is especially challenging in low signal to noise ratio situations when the total number of counts observed is limited, as is the situation in nuclear medicine imaging modalities where the low count level acquired in typical studies is due to practical limitations on imaging time and the amount of radioactivity that can be administered safely to patients.

This work develops and analyzes a new multiscale restoration framework to Poisson intensity estimation in photon limited imagery. This framework has been named DeQuant in analogy to **De**noising when the noise is of **Quant**um nature.

DeQuant is based on a test of statistical significance of the wavelet coefficients of the Poisson process. This test enables to separate the significant coefficients containing information from the non significant coefficients for which the lack of local information is not rejected. A regularization constraint assigns a new value to the non significant coefficients. The application of the inverse wavelet transform yields the denoised intensity estimate.

The material in this work is organized into seven chapters. The first three chapters formulate the problem, present the mathematical tool and analyze the limits of existing restoration methods. Chapters 4 and 5 develop extensively DeQuant and constitute the theoretical part of our innovative contribution. Chapter 6 is dedicated to simulations and chapter 7 to the graphical user interface that has been built for DeQuant. A detailed outline is given below.

Chapter 1 formulates the restoration problem. It therefore presents the basic principles of the image formation process in nuclear medicine, discusses the many interacting factors affecting the quality of the image and introduces simplifying hypothesis leading to the image degradation model.

Chapter 2 deals with the mathematical tool used in DeQuant. It presents important concepts related to multiresolution analysis and formulates the transforms used for DeQuant, namely Mallat's wavelet transform [Mal89b] and Coifman and Donoho's shift invariant wavelet transform [CD95a] in the operator formalism. This latter is crucial for the simplicity of calculations in the development of the DeQuant algorithms.

Chapter 3 discusses some classical image restoration approaches and shows their inadequation to the particular restoration problem in nuclear medicine.

Chapter 4 develops in details the DeQuant framework. It shows that DeQuant can be used not only to estimate the intensity of a Poisson process from the observation of the counts but also to recover an object related through a linear system of equations to the intensity function of the Poisson process from the observation of the Poisson data.

Chapter 5 calculates the general expression of the probability density function (PDF) of the wavelet coefficients for a Poisson process. This PDF plays a fundamental role in the DeQuant framework.

Chapter 6 addresses the problem of the validation of the DeQuant algorithms, concentrating thereby on thyroid studies. The validation process starts with simulated data experiments that permit controlled evaluation over a wide range of conditions. The algorithms are then tested in a real world scanning environment by processing actual images of thyroid phantoms. The last stage of the evaluation process is the validation with real thyroid data acquired in vivo.

Chapter 7 presents the MATLAB-based graphical user interface (GUI) that has been built for DeQuant. This GUI increases the productivity of the user and provides an insight into the performances of DeQuant for people with little or no MATLAB programming experience.

Three appendices are included at the end of this work. Appendix A contains a detailed description of the eight DeQuant algorithms obtained by combining all possibilities presented in Chapter 4. Appendix B describes an algorithm for edge detection introduced by Bijaoui and Froesché [BF80] and used in Chapter 4. Appendix C contains outline proofs and derivations for the properties given in Chapter 5. It also contains the table of thresholds derived for the unnormalized Haar wavelet and used in the DeQuant framework.

1. Physics in nuclear medicine and image model

A patient undergoing a nuclear medicine exam receives orally or through injection a very small amount of an appropriate radiopharmaceutical. A radiopharmaceutical is a radionuclide labeled with a pharmaceutical. The function of the pharmaceutical is, ideally, to concentrate in the organ or tissues of interest. The role of the radionuclide is to signal the location of the pharmaceutical by the emission of gamma rays whose energy is high enough that most of them will escape from the patient. The radionuclide most commonly used is the metastable Technetium Tc^{99m} which emits 140 keV γ rays. These rays are collected by a gamma camera which produces an image of the pharmaceutical distribution within the body and gives thus information about the physiology of the organ under study.

This chapter presents the basic principles of the image formation process in nuclear medicine. It discusses the many interacting factors affecting the quality of the image. The purpose of the chapter is to elaborate the image degradation model that allows to formulate the adequate approach to the restoration problem.

The chapter is organized as follows. Section 1.1 discusses the mathematical aspects of radioactive decay. Section 1.2 briefly reviews the different types of interactions of particles and matter on which most of the imaging procedure is based. Section 1.3 describes the major components of a gamma camera together with the limitations they introduce in the imaging procedure. Section 1.4 gives the image degradation model. Section 1.5 offers some conclusions.

Most of the material concerning nuclear medicine has been taken from references [SST92] and [SP87].

1.1. Radioactive decay

Radioactive decay is the process by which an unstable nucleus transforms into a more stable daughter nucleus. Depending on the specific mode of decay, the transformation often involves particle and/or photon emissions and the release of nuclear energy. Tc^{99m} decays by emitting γ photons with 140keV energy.

This radioactive disintegration is a stochastic process governed statistical laws. It is impossible to predict which of the unstable nuclei in a sample will disintegrate in the next second. We will show in this section that radioactive decay follows generally a binomial distribution, and how this latter becomes a Poisson distribution under some restrictions.

1.1.1. Binomial distribution

Consider a very large set of objects consisting of two classes A and B. Let p represent the probability that any object selected at random will be of class A; then $1-p$ is the probability that it will be of class B. The probability Prob(n) that exactly n of N_0 objects selected from the set will be of class A can be shown to be:

$$\text{Prob}(n) = \binom{N_0}{n} p^n q^{N_0-n} \tag{1.1}$$

This is the binomial distribution [Sap90] with expectation $N_0 p$ and variance $N_0 pq$.

Consider now the radioactive decay in time T of a system containing N_0 radioactive atoms. These N_0 atoms can be divided into two groups, those that decay in time T and those that do not decay in this time. Let D be the random variable 'number of atoms that decay in time T'. The probability that a given atom does not decay is given by the so called decay factor q which is an exponential function of time,

$$q = e^{-\lambda_r T}$$

where λ_r is the decay constant for the species in questions. It follows also that the probability p for decay is

$$p = 1 - e^{-\lambda_r T}$$

Using Equation (1.1), we can write the probability Prob($D = n$) that n atoms decay in time T as:

$$\text{Prob}(D = n) = \binom{N_0}{n} (1 - e^{-\lambda_r T})^n (e^{-\lambda_r T})^{N_0-n} \tag{1.2}$$

Thus the average number of atoms decaying in time T is

$$m = N_0(1 - e^{-\lambda_r T})$$

and the variance is

$$\sigma^2 = N_0(1 - e^{-\lambda_r T})e^{-\lambda_r T}.$$

1.1.2. Poisson distribution

Under the restrictions that (i) the observation time is short compared with the half life (ii) a very small portion n of the N_0 radioactive atoms actually decays and (iii) N_0 is big:

$$\begin{array}{lcl} \lambda_r T & \ll & 1 \\ n & \ll & N_0 \\ N_0 & \gg & 1 \end{array}$$

The binomial distribution for radioactive decay with parameters N_0 and p can be approximated by the Poisson distribution [Sap90] of parameter $\lambda = N_0 p$:

$$\text{Prob}(D = n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

The average number of atoms decaying in time T and the variance are now:

$$m = \sigma^2 = \lambda = N_0(1 - e^{-\lambda_r T}).$$

1.2. Interactions of particules with matter

High energy photons such as γ rays transfer their energy to matter in complex interactions with atoms, nuclei and electrons. Some of the photons interactions result in the ejection of orbital electrons from atoms. These electrons in turn cause ionization effects, which are the basis for mechanisms by which high-energy photons are detected. We examine in this section the interactions of γ rays and electrons with matter.

1.2.1. Interactions of γ rays with matter

The interactions which are of significance in nuclear medicine are (1) coherent scattering, (2) photoelectric effect and (3) Compton scattering.

1. Coherent scattering is a type of scattering interaction that occurs between a photon and an atom as a whole. The photon is deflected into another direction loosing little energy and therefore with negligible change in wavelength. Coherent scattering generally occurs in low-energy radiation that does not carry enough energy to eject the orbital electrons out of the orbit or ionize the atom or molecule. This is the only interaction of the three between γ rays and matter that does not cause ionization.

2. The photoelectric effect is an atomic absorption process in which an atom absorbs totally the energy of an incident photon. The photon disappears and the energy absorbed is used to eject an orbital electron from the atom. The ejected electron is called photoelectron. It receives a kinetic energy equal to the difference between the incident photon energy and the binding energy of the electron shell from which it was ejected.

3. Compton scattering is a collision between a photon and a loosely bound outer shell orbital electron of an atom. In Compton scattering, because the incident photon energy greatly exceeds the binding energy of the electron to the atom, the interaction looks like a collision between the photon and a free electron. The photon does not disappear in Compton scattering. Instead, it is deflected through a scattering angle θ.

The probability for each process to occur depends among other things on the energy of the γ photon and the atomic number of the atom .

1.2.2. Interactions of electrons with matter

Electrons are the most important type of charged particles encountered in nuclear medicine. High energy electrons are generated when γ rays interact with matter and are responsible for the ultimate deposition of energy from these radiations in an absorbing medium. The interaction mechanisms that are of importance in nuclear medicine are (1) the ionization and (2) the excitation.

1. An ionization interaction looks like a collision between a charged particule and an orbital electron. The charged particule looses energy in the collision.

Part of this energy is used to overcome the binding energy of the electron to the atom, and the reminder is given to the ejected secondary electron as kinetic energy. The ejected electron may be sufficiently energetic to cause secondary ionizations on its own.

2. An atomic or molecular excitation occurs when a charged particle raises an orbital electron to an excited state. This type of interaction generally results in smaller energy losses than occur in ionization events. The energy transferred to an atom in an excitation interaction is dissipated in molecular vibrations, and atomic emissions of infrared, visible or UV radiations, etc.

1.3. Gamma camera components

The gamma camera is the most used imaging device in nuclear medicine. A gamma camera has three major components: the collimator, the scintillation detector and the processing unit, see Fig. 1.1. In this section, we present the basic principles and the performance limitations of each camera component.

1.3.1. The collimator

The collimator acts like the lens in an optical imaging system. However, instead of focusing the γ photons by refraction, the collimator uses the projection by absorption technique. It allows only those γ rays traveling along certain directions to reach the detector, see Fig. 1.2.

A parallel-hole collimator consists of an array of small parallel holes in a lead or another heavy metal absorber. The collimator is characterized by (1) the geometry of its holes, specifically, their shape, length l and diameter d, (2) by the septal thickness t, namely the thickness of the lead walls between the holes, and (3) by the linear attenuation coefficient of the collimator material μ.

How does it work? The function of the collimator is based on the photoelectric effect. Gamma rays striking the collimator with a solid angle that does not pass with the collimator geometry undergo a photoelectric interaction and disappear.

Performance characteristics Collimator efficiency and collimator resolution characterize the performance of the collimator.

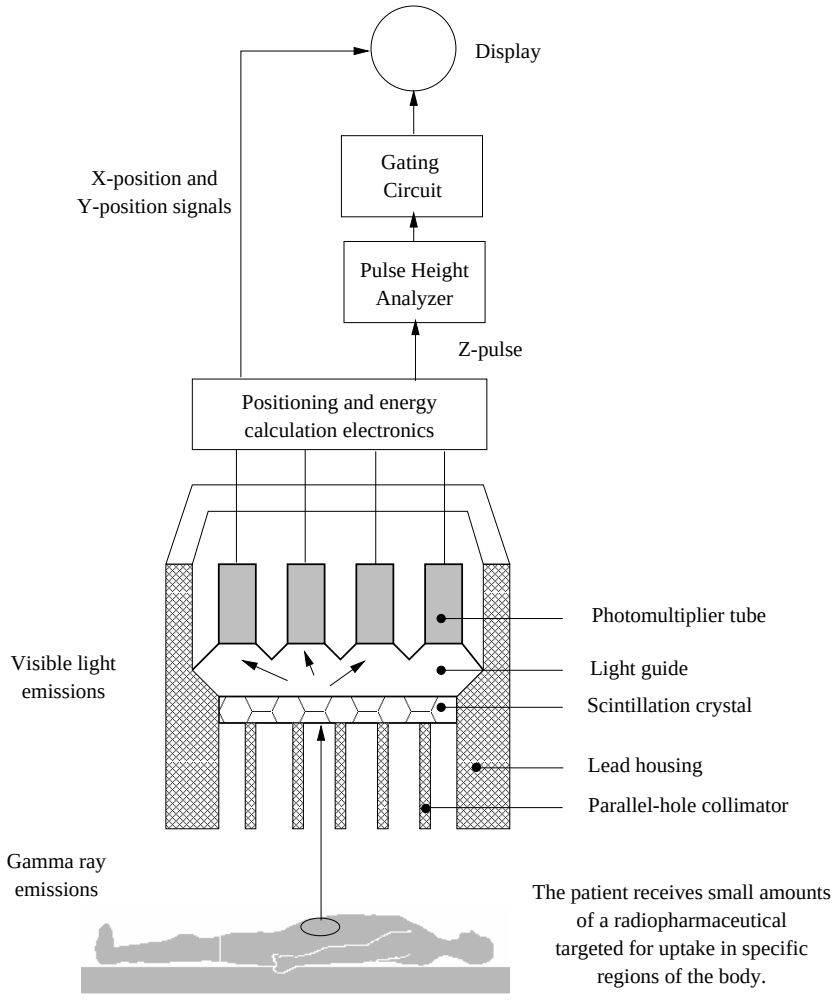


Figure 1.1.: Basic principles and components of the gamma camera

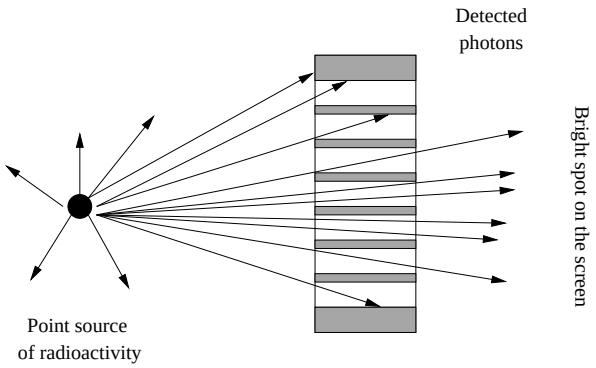


Figure 1.2.: Image of a point source of radioactivity given by the collimator.

- Collimator efficiency g: it is defined as the fraction of γ rays passing through the collimator per γ ray emitted by the source. It is given by:

$$g \backsimeq K^2 \left(\frac{d}{l_e}\right)^2 \left(\frac{d}{d+t}\right)^2 \tag{1.3}$$

  where K is a constant depending on the hole shape and $l_e = l - 2\mu^{-1}$ is the "effective length" of the collimator.

- Collimator resolution R_c: it refers to the sharpness or detail of the γ ray image projected onto the detector. It is defined as the full width at half minimum (fwhm) of the radiation profile from a point source of radiation projected by the collimator onto the detector. Collimator resolution is given by:

$$R_c \backsimeq \frac{d}{l_e}(l_e + b) \tag{1.4}$$

  where b is the distance from the radiation source to the collimator.

Several aspects of equations (1.3) and (1.4) should be noted. First, resolution improves as the ratio of hole diameter to effective length (d/l_e) is made smaller. Long narrow holes provide images with the best resolution; however, collimator efficiency decreases approximately as the square of the ratio of hole diameter to length $(d/l_e)^2$. Therefore, for a given septal thickness, collimator resolution is improved only at the expense of decreased collimator efficiency. Equation (1.3) also demonstrates the effect of septal thickness on efficiency. It is desirable that septal thickness be as small as possible so that the collimator septa obstruct the smallest possible area of detector surface and collimator efficiency is maximized; however, the collimator septa must be thick enough to ensure that the level of septal penetration by γ rays crossing from one collimator hole into another is reasonably small. Equation (1.4) indicates that collimator resolution becomes poorer as source-to-collimator distance b increases. Thus structures closest to the collimator are imaged with sharpest detail.

Some typical values The septal thickness required for low energy lead collimator (150 keV) having hole diameters 0.25 cm and lengths 2.5 cm in order to have a septal penetration of $\sim 5\%$ is $t \backsimeq 0.03$ cm. A low energy and high resolution collimator have typically an efficiency $g = 1.8\,10^{-4}$ and a resolution $R_c = 7.4$ mm.

1.3.2. The scintillation detector

The scintillation detector consists of (1) a scintillation crystal for converting high energy γ ray photons into visible light photons, (2) photomultiplier tubes (PMT) for converting light photons into a proportional pulse of electrical current, and (3) a pulse height analyzer (PHA) to analyze the energy of the absorbed photon.

How does it work? The most commonly used scintillation crystal is NaI(Tl) (thallium-activated sodium iodide) because of its favorable performance/cost ratio. Gamma photons transfer their energy to the crystal principally by photoelectric effect and Compton scattering. The resulting ejected orbital electrons cause ionization effects and atomic excitations. When the ionized or excited products undergo recombination or deexcitation, energy is released as visible light.

The visible radiations strike the front surface of a photomultiplier tube which is coated with a photoemissive substance. A photoemissive substance is one that ejects electrons when struck by photons of visible light. The photoemissive surface is called photocathode, and electrons ejected from it are called photoelectrons. A short distance from the photocathode is a metal plate called a dynode. The dynode is maintained at a positive voltage (typically 200-400 V) relative to the photocathode and attracts the photoelectrons ejected from it. The dynode is coated with a material having relatively high secondary emission characteristics. A high-speed electron striking the dynode surface ejects several secondary electrons from it. The electron multiplication factor depends on the energy of the photoelectron, which in turn is determined by the voltage difference between the dynode and the photocathode. Secondary electrons ejected from the first dynode are attracted to a second dynode, which is maintained at 50-150 V higher potential than the first dynode, and the electron multiplication process is repeated. This occurs through many additional dynode stages (typically 9-12 in all), until finally a shower of electrons is collected at the anode. Thus a relatively large pulse of current is produced when the tube is stimulated by even a relatively weak signal. The amount of current produced or equivalently the charge q appearing at the output of the photomultiplier tube is proportional to the energy E_n of the primary γ photon. Details on the proportionality factor are given by Fig. 1.5.

The PHA analyzes individual events for energy. To understand the utility of a pulse height analyzer, suppose that a monoenergetic γ ray source is placed in front of a scintillation detector. The amplitude of the voltage pulse from the PMT is proportional to the amount of energy $F_n E_\gamma$ deposited in the detector by the detected radiation event, see Fig. 1.5. While most of the photoelectric interaction result in full deposition of the γ ray energy in the detector, single and multiple Compton scattering result in transmission of only a part of the γ ray

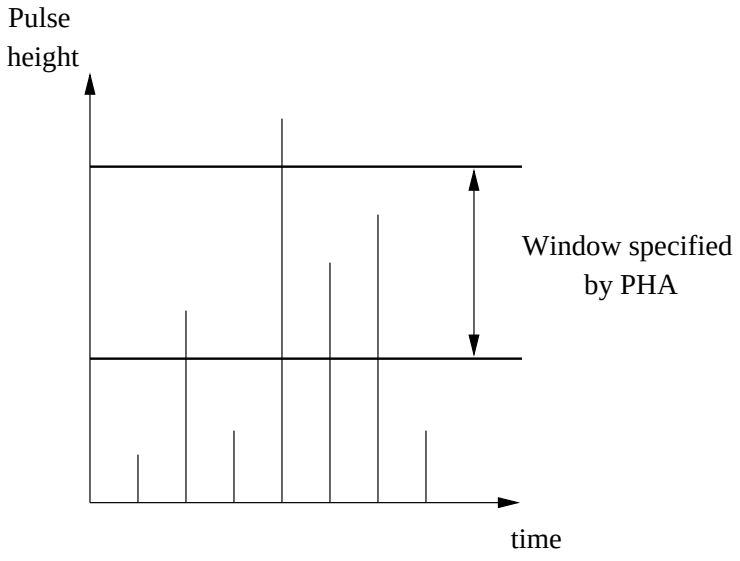


Figure 1.3.: Output from pulse height analyzer

energy to the detector. The ability of the PHA to select an energy window for photon detection is important for scatter rejection, see Fig. 1.3.

Limitation in performances Energy resolution and intrinsic resolution characterize the performance of the scintillation detector. In the following, we will only present the energy resolution. Intrinsic resolution will be discussed in Section 1.3.3 since it also depends on the positioning electronics.

Fig. 1.4 shows the number of pulses recorded per unit time at the output of the PHA versus energy, when the detector is exposed to a Tc^{99m} source. With an ideal radiation detector this would produce a single narrow line, called the photopeak, at a location corresponding to the γ ray energy E_γ. Compton interactions produce energies ranging from near zero up to the photopeak value. Sharp lines and sharp edges in the ideal spectrum become broadened lines and rounded edges in actual spectra. This is caused by electrical noise in the PMT and statistical variations in the proportionality factors C_{np}, S_m and M, namely, the number of scintillation light photons produced per keV of radiation energy deposited in the detector, the number of photoelectrons released from the photocathode and the electron multiplication factor of the dynode, see Fig. 1.5. Because of these factors, there are differences in amplitude of the signal from the detector for events in which precisely the same amount of radiation energy is deposited in the detector.

With NaI(Tl) detectors, the principal source of variation is the number of photoelectrons released from the photocathode. The average number is about three

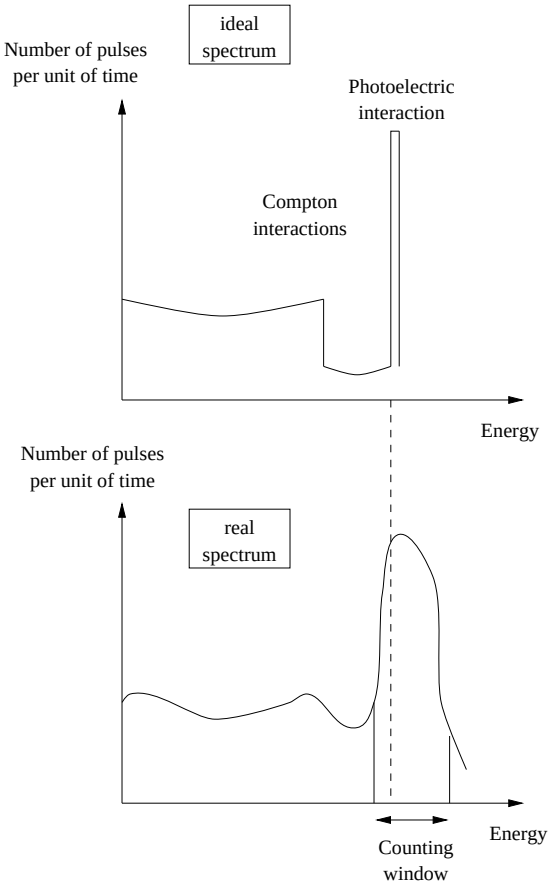


Figure 1.4.: Number of pulses recorded per unit time at the output of the PHA versus energy.

per keV of radiation energy absorbed in the NaI(Tl) crystal. Thus complete absorption of a 140 keV γ ray results in the release of about 400 photoelectrons from the photocathode on the average; however, the actual number varies from one γ ray to the next according to Poisson statistics, with a standard deviation of $\pm\sqrt{400} = 20$ photoelectrons or about $\pm 2\%$. The photopeak is thus a Poisson-shaped curve.

The spread or broadening of the energy peak at 140 keV is referred to as the energy resolution of the detector system. The energy resolution is usually defined as the full width at half maximum (fwhm) of the measured energy peak, ΔE. It is often expressed as a percentage of the photopeak energy E_γ, that is,

$$\text{energy resolution } (\%) = \frac{\Delta E}{E_\gamma} * 100$$

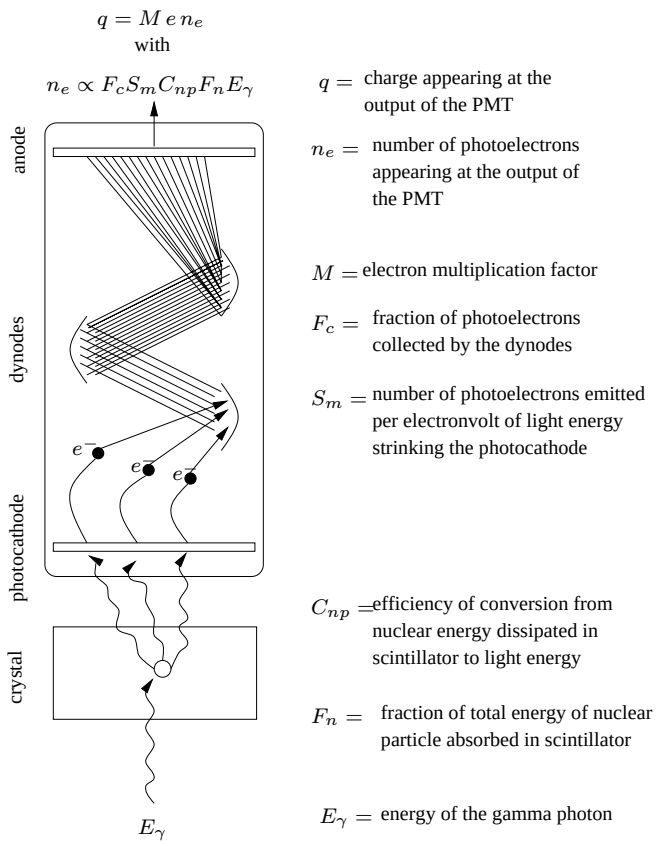


Figure 1.5.: Illustration the proportionality between the charge q appearing at the output of the photomultiplier tube and the energy E_γ of the primary gamma photon

The energy resolution affects the ability of a radiation detector to distinguish γ ray photons with similar energies. A good energy resolution is useful in rejecting scatter radiations.

Some typical values Typical electron multiplication factors are 3 to 6 per dynode. The total electron multiplication factor is very large, e.g. 6^{10} for a ten stage tube with average multiplication factor 6 at each dynode. A NaI(Tl) detector is usually 6 mm-1.25 cm thick and has 30-50 cm diameter. Most modern cameras employ 37, 61, 75 or 91 tubes arranged in a hexagonal pattern. With good quality PMT an energy resolution of 11-14 percent is achievable for the 140 keV γ rays of Tc^{99m}.

1.3.3. Positioning and energy calculation electronics

The positioning electronics determines the location of each scintillation event as it occurs in the crystal.

How does it work? Fig. 1.6 is a schematic drawing for a seven PMT version of the gamma camera and will be used to illustrate the principles of scintillation event localization.

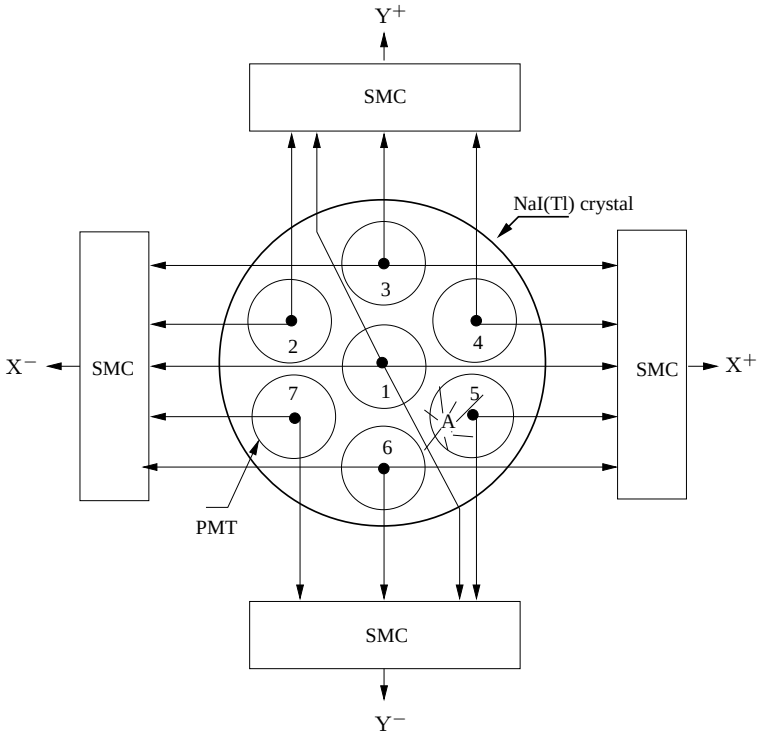


Figure 1.6.: Schematic representation of a seven PMT camera. Signals from individual PMT are combined in summing matrix circuits to obtain X^+, X^-, Y^+ and Y^- signals, which in turn are used to generate X and Y position signals. The abbreviation SMC stands for summing matrix circuits

Suppose that a scintillation event occurs at point A in the crystal. The PMT closest to the event will receive the greatest amount of light and thus will provide output signals of the greatest amplitude. Therefore, in the example shown in Fig. 1.6, the X^+ signal will be larger than the X^- signal because the event occurred in the right-hand half of the crystal, and the Y^- signal will be larger than the Y^+ signal because the event occurred in the lower half. The summing matrix circuits combine the signals from the individual PMT in such a way that the relative amplitudes of the X^- and X^+, and of the Y^+ and Y^- signals are proportional

to the distance of the scintillation event from the center line of the crystal. In a separate circuitry, the output of all seven PMTs are combined to form a Z signal. The Z signal is proportional in amplitude to the total amount of light produced by a scintillation event in the crystal and is used to determine its energy. The X^-, X^+, Y^- and Y^+ signals are then combined to obtain the X-position and Y-position signals. The X-position signal is the difference between the X^+ and X^- signals divided by the total light signal Z.

$$X = k\frac{X^+ - X^-}{Z} \tag{1.5}$$

and similarly for the Y position signal,

$$Y = k\frac{Y^+ - Y^-}{Z} \tag{1.6}$$

In Equations (1.5) and (1.6), k is a scaling factor. If the Z signal is accepted within the selected energy window of the PHA, the event is counted.

Limitation in performances The intrinsic resolution characterizes the performance of the positioning electronics among others.

Intrinsic resolution R_i is defined as the spatial resolution of the system without the collimator: sharp edges or small point objects produce blurred rather than sharply defined images. Part of the blurring arises from collimator resolution, discussed in Section 1.3.1 and part arises in the detector and positioning electronics. In this section, we will discuss the limitation factors due to the detector and the positioning electronics.

- The primary cause of limited intrinsic resolution is statistical fluctuation in the distribution of light photons between PMTs from one scintillation event to the next. As we saw in the paragraph concerning limitations in the detector performance, if a certain PMT records on the average N light photons from scintillation events occuring at a certain location in the crystal, the actual number detected from one event to the next varies with a standard deviation given by $\sqrt{N}$. Thus if a very fine beam of γ rays is directed at the detector, the light flashes appearing on the cathode ray tube screen are not all placed at precisely the same location but are distributed over a certain area, the size of this area depending on the magnitude of these statistical fluctuations.
- Intrinsic resolution also depends on the number of PMTs and the detector crystal thickness. With thicker detectors there is a greater spreading of scintillation light before it reaches the PMTs.

Some typical values One method for measuring intrinsic resolution is to obtain with the scintillation detector images of "bar patterns" which consist of strips of lead separated from each other by spaces equal to the width of the strips. Modern gamma cameras permit to resolve 3-5 mm bar patterns for Tc^{99m} (140 keV) radiation.

1.4. Image degradation model

The distribution of the pharmaceutical in a given organ is the diagnosis relevant information for the physician. It is measured by detecting radiation emissions from the decay of the radioactive substance labeling the pharmaceutical. The problem is that the interactions in the patient body and the performance limitations of the gamma camera lead to an image which is far from being a "perfect" representation of the radionuclide distribution in the organ. We are thus faced with a typical restoration problematic where an image that has been degraded must be recovered by using a priori knowledge about the degradation phenomenon.

We presented in the previous sections of this chapter some of the physical processes that cause the degradation mechanism. Since it would be extremely complex to take all of them into account, this section introduces some simplifying hypothesis that enable the formulation of the degradation model.

The following notations will be used in this section.

- $U(x, y, z)$ is the local uptake density. It is defined by the number of radionuclides taken at each point of the imaged organ.
- Q is the point spread function (PSF) corresponding to the interaction in the patient body. It also takes the imaging time into account.
- X is the image of the uptake density that would be given by an ideal imaging system.
- P is the operator associated with the point spread function of the imaging system. It can be seen as the probability that a photon that would be detected at a given position by an ideal imaging system is detected at another position by the real camera.
- $Y = PX$ is the distribution of photons after that has been spread by the PSF of the camera.

- $I(k_x, k_y)$ is the number of photon counts detected at position (k_x, k_y) of the image during the imaging time T.

1.4.1. Interactions in the patient body

Modeling the interactions in the patient body is equivalent to find the relationship between $U(x, y, z)$ and $X(x, y)$. We make the following hypotheses:

Independence: The disintegration process corresponds to independent events. It can be modelled by a Poisson law.

Linearity: We suppose that all interactions in the patient body can be considered as an independent selection of events. This means that even though there are some non linear processes the number of those is linearly proportional to the uptake and we neglect all processes that lead to induced phenomena. Thus, if we had two uptake distributions $U_1(x, y, z)$ and $U_2(x, y, z)$ that produce $X_1(x, y)$ and $X_2(x, y)$ the sum of the two distributions $U_1(x, y, z)$ + $U_2(x, y, z)$ would give the sum of the two images $X_1(x, y) + X_2(x, y)$. We can therefore write:

$$X(x, y) = \iiint Q(\alpha, \beta, \gamma, x, y)U(\alpha, \beta, \gamma)d\alpha d\beta d\gamma \tag{1.7}$$

Translation invariance: The interactions in the patient body can be seen as being translation invariant. Equation (1.7) reduces to:

$$X(x, y) = \iiint Q(x - \alpha, y - \beta, \gamma)U(\alpha, \beta, \gamma)d\alpha d\beta d\gamma \tag{1.8}$$

The translation invariance hypothesis is not always verified due to the fact that the environment of the organ can spatially vary. In this case we have to consider Equation (1.7).

Determining the exact formulation for Q is beyond the scope of this work. It is based on the biophysics of the interactions of gamma rays in the body.

1.4.2. Degradation introduced by the camera

We make the following hypotheses on the gamma camera and its imperfections.

- The gamma camera is considered as being a linear shift invariant system characterized by the operator P associated to the PSF.
- We have seen that the sharpness of the images recorded with a gamma camera is limited by the intrinsic resolution R_i of detector and electronics, and the collimator resolution R_c. The combined effect of these two factors produces the system resolution $R_s = \sqrt{R_i^2 + R_c^2}$ that is somewhat worse than either one alone. Since the collimator resolution depends on the source to collimator distance, see Fig. 1.4, the system resolution also depends on this parameter.

  We suppose that all points of the organ to be imaged are situated at an almost constant distance of the detector. Since for simplicity reasons, the fwhm of the PSF of the camera is assimilated to the system resolution, this hypothesis enables us to use a unique spatial PSF. Practically, this PSF will be approximated by a Gaussian profile which fwhm equals that of the PSF measured for the mean distance corresponding to the imaging procedure.
- We neglect the attenuation of the primary emitted photons due to the non detection of scattered photons or to photons absorbed by photoelectric interactions. This hypothesis can be justified by the fact that we are faced with a detection problem. We are not interested in the exact quantification of the emitted photons.
- The errors due to the detection of scattered photons are neglected.
- We do not take into account the energy resolution of the camera.

1.4.3. The degradation model

We model each point of the scintigram (i.e. image produced by the gamma camera) by the random variable $I(k_x, k_y)$ which represents the number of photon counts detected at position (k_x, k_y) of the image during the imaging time T. $I(k_x, k_y)$ obeys a Poisson distribution of parameter $Y(k_x, k_y)$:

$$\text{Prob}(I(k_x, k_y) = n) = \frac{Y^n(k_x, k_y)}{n!} e^{-Y(k_x, k_y)} \tag{1.9}$$

where

$$Y(k_x, k_y) = \int_d \rho(x, y) dx dy. \tag{1.10}$$

The noise free image or discrete intensity function $Y(k_x, k_y)$ is obtained by integrating the continuous intensity function $\rho(x, y)$ on a domain d corresponding to a pixel.

1.5. Conclusion

We described in this chapter the mechanisms causing the degradation of an image taken by a gamma camera. This enabled us to formulate an image formation model where the degraded image is the realization of a random Poisson process whose parameter is the original image.

Although nuclear medicine imaging suffers from relatively poor spatial resolution and high noise fluctuations due to the small number of counts in the images, it provides information about the physiological functions of the patient that is difficult or impossible to obtain from other imaging modalities such as X-ray or ultrasound techniques.

However, practical limitations on imaging time and the amount of radioactivity that can be administered safely to patients are serious impediments to substantial further improvements in nuclear medicine imaging. Hence improvements in image quality via optimized image processing represent a significant opportunity to advance the state of the art in this field.

We present in this work a novel multiscale image restoration framework which permits given one realization of the random Poisson process to estimate the parameter $Y(k_x, k_y)$.

2. An introduction to wavelets

We do not pretend in this chapter to give a detailed description of the subject area of wavelets. Our purpose is on the one hand to present the basic elements that enable to apprehend the multiresolution analysis philosophy and understand why the wavelet transform has been chosen in DeQuant's restoration framework and on the other hand to formulate the wavelet transforms used for DeQuant, namely Mallat's algorithm [Mal89b] and Coifman and Donoho's shift invariant wavelet transform [CD95a], in the operator formalism. This latter is crucial for the simplicity of calculations in the development of the DeQuant algorithms.

This chapter is organized as follows. Section 2.1 gives some elements of definition of a wavelet. Section 2.2 presents the continuous wavelet transform and some of its properties. Section 2.3 introduces the concept of multiresolution which enables to perform a fast and stable analysis and synthesis of signal and images using wavelets. Section 2.4 describes the one and two dimensional Mallat algorithms for computing the discrete orthogonal wavelet transform. Section 2.5 presents the shift invariant wavelet transform algorithm introduced by Coifman and Donoho for the special case of the unnormalized Haar wavelet. Section 2.6 offers some conclusions.

2.1. What are wavelets?

Given that the wavelet field keeps growing, the definition of a wavelet continuously changes. At this point, we would like to quote Wim Sweldens who wrote in [Swe96]

> One can compare the wavelet research with an infinite dimensional fractal, sometimes taking off in isolated directions but also many times folding back onto itself. Finding a definition of a wavelet is like approximation this fractal with a ball. The 'minimal' solution is a ball with a small radius which fits in the interior of the fractal. But this

> leads to a definition that contains only the very core material and leaves out most of the recent and very exciting developments. The 'maximal' solution is fitting the fractal in a ball. This results in a definition that includes almost any function.

Therefore we step away from the idea of giving a rigorous definition of a wavelet and introduce in the following some elements that capture essentially the reasons for calling a particular function a wavelet.

A wavelet or mother wavelet is a real or complex-value continuous function $\psi(x)$ with the following properties [RB98]:

- The function integrates to zero, or equivalently, its Fourier transform denoted $\psi(\omega)$ is zero at the origin:

$$\int_{-\infty}^{+\infty} \psi(x)dx = 0 \quad \Leftrightarrow \quad \psi(\omega)|_{\omega=0} = 0 \tag{2.1}$$

- It is square integrable, or equivalently, has finite energy:

$$\int_{-\infty}^{+\infty} |\psi(x)|^2 dx < \infty \tag{2.2}$$

- The Fourier transform of $\psi(x)$ must satisfy the admissibility condition [LMR97]

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(\omega)|^2}{|\omega|} d\omega < \infty \tag{2.3}$$

Equation (2.1) is suggestive of a function that is oscillatory or has a wavy appearance. Equation (2.2) implies that most of the energy in $\psi(x)$ is confined to a finite interval, or in other words that $\psi(x)$ has a good space localization. Ideally the function is exactly zero outside the finite interval: a so-called compactly supported function. In general, we want fast, e.g., inverse polynomial or exponential decay away from the center of mass of the function. Equation (2.3) is useful in formulating the inverse wavelet transform. For this equation to be satisfied, $\psi(\omega)$ must have a sufficient decay in frequency. This simply means that the Fourier transform of a wavelet is localized, i.e., a wavelet mostly contains frequencies from a certain frequency band. Since the Fourier transform is zero at the origin,

see Equation (2.1), and the spectrum decays at high frequencies, a wavelet has a bandpass characteristic.

A wavelet is thus a "small wave", that exhibits a good time-frequency localization. The Heisenberg uncertainty principle puts a lower bound on the product of space and frequency variance [Wic94].

A doubly indexed family of wavelets can be generated by dilating and translating the mother wavelet $\psi(x)$:

$$\psi_{(a,b)}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{2.4}$$

where $a \in \mathbb{R}^{*+}$ is the scale parameter and $b \in \mathbb{R}$ is the position parameter.

2.2. Continuous wavelet transform

The continuous wavelet transform (CWT) maps a one dimensional function $f(x)$ $\in L^2(\mathbb{R})$ into a two dimensional function of a and b. We present in this section the direct and inverse transforms and introduce some properties of the CWT.

Direct transform

The CWT consists of computing coefficients that are inner products of the signal and a family of wavelets $\psi_{(a,b)}(x)$, each wavelet corresponding to a scale a and a position b:

$$W_{f(x)}(a,b) =< f(x), \psi_{(a,b)}(x) >= \frac{1}{\sqrt{a}}\int f(x)\psi^*\left(\frac{x-b}{a}\right)dx \tag{2.5}$$

The variable x and the position parameter b vary continuously over $\mathbb{R}$. The scale parameter a is restricted to $\mathbb{R}^+$. $W_{f(x)}(a,b)$ are the wavelet coefficients of $f(x)$ with respect to the wavelet $\psi(x)$.

Inverse transform

Given the continuous wavelet transform coefficients $W_{f(x)}(a,b)$ of a function $f(x)$, the function can be recovered by the following reconstruction formula also called resolution of the identity [VK95]:

$$f(x) = \frac{1}{C_\psi} \int_{a=0}^{\infty} \int_{b=-\infty}^{+\infty} \frac{1}{a^2} W_{f(x)}(a,b)\, \psi_{(a,b)}(x)\, da\, db$$

Where C_ψ is given by Equation (2.3).

Properties of the CWT

In the following we give some properties of the continuous wavelet transform. A more detailed discussion can be found in [VK95] and [Gro89].

- Linearity: the CWT is given by the inner product which is a linear operation.

$$W_{cf_1(x)+df_2(x)} = cW_{f_1(x)} + dW_{f_2(x)}$$

- Energy conservation: the CWT has the energy conservation property

$$\int_{-\infty}^{+\infty} |f(x)|^2 dx = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} |W_{f(x)}(a,b)|^2 da db$$

  with C_ψ given by Equation (2.3)

- Shift invariance: shifting the input signal yields a shift in the wavelet transform

$$W_{f(x-\tau)} = W_{f(x)}(a, b-\tau)$$

- Scaling: a scaled input signal yields a scaling of the wavelet transform in both parameters.

$$W_{\frac{1}{\sqrt{\tau}} f(\frac{x}{\tau})} = W_{f(x)} \left(\frac{a}{\tau}, \frac{b}{\tau} \right)$$

2.3. Multiresolution analysis

A multiresolution analysis or approximation of a function $f(x) \in L^2(\mathbb{R})$ is a decomposition of $f(x)$ in a coarse part (or approximation) and a set of functions containing the difference of information (or the details) that has to be added to this coarse part in order to retrieve the original signal $f(x)$. This section constructs in one and two dimensions the appropriate mathematical environment that enables to formulate mathematically the qualitative notions of 'approximation' and 'details'.

2.3.1. One dimensional multiresolution analysis

We wish to build a multiresolution representation based on the difference of information available at two successive resolutions. This section shows that such a representation can be computed by decomposing the signal using a wavelet orthonormal basis.

The coarse approximation spaces V_{2^j}

A multiresolution analysis of $L^2(\mathbb{R})$ [Mal89a, Mal89b, Dau92] for the set of resolutions $\{2^j\}_{j\in\mathbb{Z}}$ consists of a sequence of embedded closed subspaces

$$\begin{array}{ccccc} \{0\} \in \ldots \in V_{2^{j+1}} \in & V_{2^j} & \in V_{2^{j-1}} \in \ldots \in L^2(\mathbb{R}) \\ \leftarrow \text{coarser} & & \text{finer} \rightarrow \end{array} \tag{2.6}$$

such that the following properties are satisfied.

1. Upward completeness

$$\lim_{j\to-\infty} V_{2^j} = \bigcup_{j=-\infty}^{+\infty} V_{2^j} \text{ is dense in } L^2(\mathbb{R}) \tag{2.7}$$

2. Downward completeness

$$\lim_{j\to+\infty} V_{2^j} = \bigcap_{j=-\infty}^{+\infty} V_{2^j} = \{0\} \tag{2.8}$$

3. Scale invariance

$$f(x) \in V_{2^j} \Leftrightarrow f(2^j x) \in V_{2^j} \tag{2.9}$$

4. Shift invariance

$$f(x) \in V_0 \Rightarrow f(x-k) \in V_0 \quad \forall k \in \mathbb{Z} \tag{2.10}$$

We denote by P_{2^j}, the orthogonal projection operator on V_{2^j} and $A_{2^j}(x)$ the approximation of $f(x)$ at scale 2^j:

$$\begin{array}{ccc} L^2(\mathbb{R}) & \xrightarrow{P_{2^j}} & V_{2^j} \\ f(x) & \longrightarrow & A_{2^j}(x) \end{array}$$

In order to characterize the projection operator, an orthonormal basis of V_{2^j} must be found. The following theorem [Mal89b] shows that such an orthonormal basis can be defined by dilating and translating a unique function $\phi(x)$.

Theorem 1 *Let $(V_{2^j})_{j\in\mathbb{Z}}$ be a multiresolution approximation of $L^2(\mathbb{R})$. There exists a unique function $\phi(x) \in L^2(\mathbb{R})$ such that $\left\{\phi_{(2^j,k)}(x) = \frac{1}{\sqrt{2^j}}\phi\left(\frac{x-k2^j}{2^j}\right)\right\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of V_{2^j}. $\phi(x)$ is called a scaling function.*

A direct consequence of the property $V_{2^0} \subset V_{2^{-1}}$ and Theorem 1 is the so called dilation equation:

$$\phi(x) = \sqrt{2} \sum_{k=-\infty}^{+\infty} h_0(k)\phi(2x-k) \tag{2.11}$$

This equation establishes the connection between two approximation spaces. It states that the weighted sum of the dilated and translated versions of the scaling function yields again $\phi(x)$. The coefficients h_0 are calculated as the inner products of the scaling function and its dilated and translated versions:

$$h_0(k) = < \phi(x), \sqrt{2}\phi(2x-k) > \tag{2.12}$$

The orthogonal projection of $f(x)$ on V_{2^j}, can be computed by decomposing the signal $f(x)$ on the orthonormal basis given by Theorem 1. Specifically, if we denote $F_{f(x)}(2^j,k)$ the approximation or scaling coefficients:

$$A_{2^j}(x) \stackrel{\triangle}{=} \sum_{k=-\infty}^{+\infty} F_{f(x)}(2^j,k)\frac{1}{\sqrt{2^j}}\phi\left(\frac{x-k2^j}{2^j}\right) \tag{2.13}$$

where

$$F_{f(x)}(2^j,k) \stackrel{\triangle}{=} < f(x), \frac{1}{\sqrt{2^j}}\phi\left(\frac{x-k2^j}{2^j}\right) > \tag{2.14}$$

The detail spaces O_{2^j}

The difference of information between the approximation of a function $f(x)$ at the resolutions 2^j and 2^{j-1} is called the detail signal at the resolution 2^j. It is given by the orthogonal projection of the original signal on the detail space O_{2^j} which is the orthogonal complement of V_{2^j} in $V_{2^{j-1}}$:

$$V_{2^{j-1}} = O_{2^j} \oplus V_{2^j}$$

In its turn V_{2^j} can be decomposed in:

$$V_{2^j} = O_{2^{j+1}} \oplus V_{2^{j+1}}$$

By repeating this process we obtain

$$V_{2^{j-1}} = O_{2^j} \oplus O_{2^{j+1}} \oplus O_{2^{j+2}} \oplus \ldots \oplus O_{2^J} \oplus V_{2^J}$$

where

$$V_{2^J} = \bigoplus_{j \geq J+1} O_{2^j}$$

A direct consequence of the property $O_{2^0} \subset V_{2^{-1}}$ is the so called wavelet equation:

$$\psi(x) = \sqrt{2} \sum_{k=-\infty}^{+\infty} h_1(k)\phi(2x-k) \tag{2.15}$$

where

$$h_1(k) =< \psi(x), \sqrt{2}\phi(2x-k) > \tag{2.16}$$

Again Theorem 2 [Mal89b] shows that an orthonormal basis of O_{2^j} can be built by scaling and translating a function $\psi(x)$.

Theorem 2 *Let $(V_{2^j})_{j\in\mathbb{Z}}$ be a multiresolution approximation of $L^2(\mathbb{R})$. There exists a function $\psi(x)$ such that $\left\{\psi_{(2^j,k)}(x) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{x-2^jk}{2^j}\right)\right\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of O_{2^j}. $\psi(x)$ is called an orthogonal wavelet.*

The detail signal of $f(x)$ at scale 2^j is the orthogonal projection on O_{2^j}. It can be computed by decomposing the signal $f(x)$ on the orthonormal basis given by Theorem 2. Specifically, if we denote $D_{2^j}(x)$ the detail signal and $W_{f(x)}(2^j,k)$ the detail or wavelet coefficients, we have:

$$D_{2^j}(x) \triangleq \sum_{k=-\infty}^{+\infty} W_{f(x)}(2^j,k)\frac{1}{\sqrt{2^j}}\psi\left(\frac{x-k2^j}{2^j}\right) \tag{2.17}$$

where

$$W_{f(x)}(2^j, k) \stackrel{\triangle}{=} < f(x), \frac{1}{\sqrt{2^j}} \psi \left(\frac{x - k2^j}{2^j} \right) > \tag{2.18}$$

Table 2.1 summarizes the relationships between the spaces V_{2^j} and O_{2^j} in $L^2(\mathbb{R})$ and gives their counterparts in terms of approximation and detail signals.

Signals		Spaces
Detail at scale 2^j	$D_{2^j}(x) = \sum_{k \in \mathbb{Z}} W_{f(x)}(2^j, k) \psi_{(2^j,k)}(x)$	O_{2^j}
The signal is the sum of its details	$f(x) = \sum_{j \in \mathbb{Z}} D_{2^j}(x)$	$L^2(\mathbb{R}) = \sum_{j \in \mathbb{Z}} O_{2^j}$
Approximation at scale 2^j	$A_{2^j}(x) = \sum_{j' > j} D_{2^{j'}}(x)$	V_{2^j}
Link between A_{2^j} and $A_{2^{j+1}}$	$A_{2^j}(x) = A_{2^{j+1}}(x) + D_{2^{j+1}}(x)$	$V_{2^j} = V_{2^{j+1}} \oplus O_{2^{j+1}}$
Several decompositions	$f(x) = A_{2^J}(x) + \sum_{j \leq J} D_{2^j}(x)$	$L^2(\mathbb{R}) = V_{2^J} \oplus \sum_{j \leq J} O_{2^j}$

Table 2.1.: Relationships between signals and spaces

2.3.2. Two dimensional multiresolution analysis

The multiresolution analysis can be easily generalized to any dimension $n > 0$. In this section, we study the two dimensional case for image applications [Mal89b]. The signal is now a finite energy function $f(x, y) \in L^2(\mathbb{R}^2)$.

The coarse approximation spaces V_{2^j}

A multiresolution approximation of $L^2(\mathbb{R}^2)$ is a sequence of subspaces of $L^2(\mathbb{R}^2)$ which satisfies a straightforward two dimensional extension of the properties (2.6) to (2.10). Let $\{V_{2^j}\}_{j \in \mathbb{Z}}$ be such a multiresolution approximation of $L^2(\mathbb{R}^2)$. The approximation of a signal $f(x, y)$ at scale 2^j is equal to its orthogonal projection on the vector space V_{2^j}. Theorem 1 is still valid in two dimensions, and one can show that there exists a unique scaling function $\phi(x, y)$ whose dilations and translations give an orthonormal basis of each space V_{2^j}.

We will restrict our presentation to the particular case of separable multiresolution approximations of $L^2(\mathbb{R}^2)$ [Mey86] for which each vector space V_{2^j} can be decomposed as a tensor product of two identical subspaces of $L^2(\mathbb{R})$

$$V_{2^j} = V_{2^j}^1 \otimes V_{2^j}^1.$$

The sequence of vector spaces $\{V_{2^j}\}_{j\in\mathbb{Z}}$ forms a multiresolution approximation of $L^2(\mathbb{R}^2)$ if and only if $\{V_{2^j}^1\}_{j\in\mathbb{Z}}$ is a multiresolution approximation of $L^2(\mathbb{R})$. One can then easily show that the scaling function $\phi(x,y)$ can be written as

$$\phi(x,y) = \phi(x)\phi(y)$$

where $\phi(t)$ is the one-dimensional scaling function of the multiresolution approximation $\{V_{2^j}^1\}_{j\in\mathbb{Z}}$. With a separable multiresolution approximation, extra importance is given to the horizontal and vertical direction in the image. The orthogonal basis of V_{2^j} is then given by

$$\left\{ \frac{1}{2^j}\phi(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}) \right\}_{(k_x,k_y)\in\mathbb{Z}^2} = \left\{ \frac{1}{2^j}\phi(\frac{x-2^j k_x}{2^j})\phi(\frac{y-2^j k_y}{2^j}) \right\}_{(k_x,k_y)\in\mathbb{Z}^2}$$

The approximation of a signal $f(x,y)$ at scale 2^j is therefore characterized by the set of inner products:

$$F_{f(x,y)}(2^j,k_x,k_y) \stackrel{\triangle}{=} < f(x,y), \frac{1}{2^j}\phi(\frac{x-2^j k_x}{2^j})\phi(\frac{y-2^j k_y}{2^j}) >$$

The detail spaces O_{2^j}

As in the one-dimensional case, the detail signal at scale 2^j is equal to the orthogonal projection of the signal on the orthogonal complement of V_{2^j} in $V_{2^{j-1}}$. Let O_{2^j} be this orthogonal complement. The following theorem [Mal89b] gives a simple extension of Theorem 2, and states that we can build an orthonormal basis of O_{2^j} by scaling and translating three wavelet functions, $\psi^h(x,y)$ $\psi^v(x,y)$ and $\psi^d(x,y)$. The superscipts h, v and d stand respectively for horizontal, vertical and diagonal.

Theorem 3 *Let $\{V_{2^j}\}_{j\in\mathbb{Z}}$ be a separable multiresolution approximation of $L^2(\mathbb{R}^2)$ and $\phi(x,y) = \phi(x)\phi(y)$ be the associated two-dimensional scaling function. Let $\psi(x)$ be the one dimensional wavelet associated with the scaling function $\phi(x)$. Then the three wavelets:*

$$\begin{aligned} \psi^h(x,y) &= \phi(x)\psi(y) \\ \psi^v(x,y) &= \psi(x)\phi(y) \\ \psi^d(x,y) &= \psi(x)\psi(y) \end{aligned}$$

are such that

$$\left\{ \begin{array}{l} \frac{1}{2^j}\psi^h\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right), \\ \frac{1}{2^j}\psi^v\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right), \\ \frac{1}{2^j}\psi^d\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right) \end{array} \right\}_{(k_x,k_y)\in\mathbb{Z}^2}$$

is an orthonormal basis of O_{2^j}.

Therefore the detail signals at scale 2^j are characterized by the set of inner products:

$$\begin{aligned} W_{h,f(x,y)}(2^j,k_x,k_y) &= < f(x,y), \psi^h_{(2^j,k_x,k_y)}(x,y) > \\ W_{v,f(x,y)}(2^j,k_x,k_y) &= < f(x,y), \psi^v_{(2^j,k_x,k_y)}(x,y) > \\ W_{d,f(x,y)}(2^j,k_x,k_y) &= < f(x,y), \psi^d_{(2^j,k_x,k_y)}(x,y) > \end{aligned}$$

where

$$\begin{aligned} \psi^h_{(2^j,k_x,k_y)}(x,y) &= \frac{1}{2^j}\psi^h\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right) \\ \psi^v_{(2^j,k_x,k_y)}(x,y) &= \frac{1}{2^j}\psi^v\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right) \\ \psi^d_{(2^j,k_x,k_y)}(x,y) &= \frac{1}{2^j}\psi^d\left(\frac{x-2^j k_x}{2^j}, \frac{y-2^j k_y}{2^j}\right) \end{aligned} \tag{2.19}$$

2.4. Discrete wavelet transform

In Equation (2.5) both a and b are continuous variables and there is redundancy in the CWT representation of $f(x)$. The discrete wavelet transform (DWT) removes this redundancy by discretizing the parameters (a, b) to a dyadic grid where $a = 2^j$, $b = k2^j$ and $j,\ k \in \mathbb{Z}$.

An attractive feature of the DWT is that the underlying multiresolution structure leads to an efficient discrete-time algorithm based on a filter bank implementation. Since this connection was pointed out by Mallat in [Mal89b], the computational procedure is referred to as Mallat's algorithm. In the following we present the Mallat algorithms for both the one and two dimensional cases.

2.4.1. One dimensional Mallat's algorithm

The algorithm can be summarized as follows:

- It is assumed that the initial, discrete data, which is a sequence $\{f(k)\}_{k\in\mathbb{Z}}$ already represents an approximation of $f(x)$ at a certain scale. By convention this scale is fixed at $j = 0$.

$$f(k) \overset{\text{by convention}}{=} F_{f(x)}(2^0, k) =< f(x), \phi(x-k) > \tag{2.20}$$

  Thus the starting point of our algorithm is the space V_{2^0}.

- The analysis proceeds by calculating the projections of $f(x)$ on the approximation spaces and on the details spaces at a bigger scale, that is to say on the spaces $\{V_{2^j}, O_{2^j}\}_{j>0}$

- From the identity $V_{2^0} = O_{2^1} \oplus O_{2^2} \oplus \ldots \oplus O_{2^J} \oplus V_{2^J}$, one can reconstruct $f(k)$ exactly by starting with its approximation at scale J and adding the sequence of details.

Analysis

In order to calculate the approximation coefficient at scale 2^j as a function of the approximation coefficient scale 2^{j-1}, we use the dilation equation (2.11) to establish the relationship between the the scaling function at scale 2^j and the its counterpart at scale 2^{j-1}:

$$\begin{aligned} \phi_{(2^j,k)}(x) \overset{\triangle}{=} \frac{1}{\sqrt{2^j}}\phi\left(\frac{x-2^j k}{2^j}\right) &= \frac{1}{\sqrt{2^j}}\frac{1}{\sqrt{2^{-1}}}\sum_n h_0(n)\phi\left(\frac{\frac{x-2^j k}{2^j} - 2^{-1}n}{2^{-1}}\right) \\ &= \sum_n h_0(n-2k)\phi_{(2^{j-1},n)}(x) \end{aligned} \tag{2.21}$$

The approximation coefficient at scale 2^j can now be calculated as follows:

$$\begin{aligned} F_{f(x)}(2^j, k) &\overset{\triangle}{=} < f(x), \phi_{(2^j,k)}(x) > \\ &= < f(x), \sum_n h_0(n-2k)\phi_{(2^{j-1},n)}(x) > \\ &= \sum_n h_0^*(n-2k) < f(x), \phi_{(2^{j-1},n)}(x) > \\ &= \sum_n h_0^*(n-2k) F_{f(x)}(2^{j-1}, n) \end{aligned} \tag{2.22}$$

Equivalently, the wavelet equation (2.15) is used to derive the relationship:

$$\psi_{(2^j,k)}(x) \overset{\triangle}{=} \frac{1}{\sqrt{2^j}}\psi\left(\frac{x-2^j k}{2^j}\right) = \sum_n h_1(n-2k)\phi_{(2^{j-1},n)}(x) \tag{2.23}$$

and the wavelet coefficient at scale 2^j is given by:

$$\begin{aligned} W_{f(x)}(2^j,k) &\stackrel{\triangle}{=} < f(x), \psi_{(2^j,k)}(x) > \\ &= \sum_n h_1^*(n-2k) F_{f(x)}(2^{j-1},n) \end{aligned} \tag{2.24}$$

Synthesis

As one would expect, a reconstruction of the original small scale coefficient of the signal can be made from a combination of the approximation and wavelet coefficients at a larger scale. This is derived by considering a signal $f(x) \in V_{2^j}$. This signal can be decomposed on the orthonormal basis $\{\phi_{(2^j,k)}(x)\}$ of V_{2^j}:

$$f(x) = \sum_{k=-\infty}^{+\infty} F_{f(x)}(2^j,k)\phi_{(2^j,k)}(x) \tag{2.25}$$

Since $O_{2^{j+1}}$ is the orthogonal complement of $V_{2^{j+1}}$ in V_{2^j}, $\{\phi_{(2^{j+1},k)}(x), \psi_{(2^{j+1},k)}(x)\}$ is also an orthonormal basis of V_{2^j} on which $f(x)$ can be decomposed:

$$f(x) = \sum_{k=-\infty}^{+\infty} F_{f(x)}(2^{j+1},k)\phi_{(2^{j+1},k)}(x) + W_{f(x)}(2^{j+1},k)\psi_{(2^{j+1},k)}(x) \tag{2.26}$$

Substituting Equations (2.21) and (2.23) in (2.26) gives:

$$\begin{aligned} f(x) &= \sum_{k=-\infty}^{+\infty} F_{f(x)}(2^{j+1},k) \sum_n h_0(n-2k)\phi_{(2^j,n)}(x) \\ &+ \sum_{k=-\infty}^{+\infty} W_{f(x)}(2^{j+1},k) \sum_n h_1(n-2k)\phi_{(2^j,n)}(x) \end{aligned} \tag{2.27}$$

Because all of these functions are orthonormal, multiplying Equations (2.25) and (2.27) by $\phi_{(2^j,k')}(x)$ and integrating gives:

$$F_{f(x)}(2^j,k) = \sum_{n=-\infty}^{+\infty} h_0(k-2n)F_{f(x)}(2^{j+1},n) + \sum_{n=-\infty}^{+\infty} h_1(k-2n)W_{f(x)}(2^{j+1},n) \tag{2.28}$$

Implementation as a filter bank

Equations (2.22) and (2.24) show that the approximation and wavelet coefficients at scale 2^j can be obtained by convolving the approximation coefficients at scale 2^{j-1} by the lowpass filter $h(n)$ and the highpass filter $g(n)$ defined by:

$$h(n) = h_0^*(-n) \tag{2.29}$$

$$g(n) = h_1^*(-n) \tag{2.30}$$

then downsampling or decimating (taking every other term, the even term). In other words, the scale 2^j coefficients are filtered by two FIR digital filters after which downsampling gives the next approximation and wavelet coefficients. This filtering and decimation can then be repeated on the approximation coefficients. We introduce the analysis operators H and G that correspond to a convolution with $h(n)$ and $g(n)$ respectively followed by a subsampling by a factor 2. The analysis Equations (2.22) and (2.24) can be written as:

$$F(2^{j+1}) = HF(2^j) \tag{2.31}$$
$$W(2^{j+1}) = GF(2^j) \tag{2.32}$$

Equation (2.28) is evaluated by upsampling the 2^{j+1} scale approximation coefficient sequence (this is done by inserting one zero between each of the original terms), then colvolving with $h_0(n)$. The same is done for the 2^{j+1} scale wavelet coefficients and the results are added to give the 2^j scale approximation coefficient. We introduce the synthesis operators $\widetilde{H}$ and $\widetilde{G}$ that correspond to an upsampling by a factor of 2 followed by a convolution with $\tilde{h}$ and $\tilde{g}$ respectively. These latter are defined by:

$$\tilde{h}(n) = h_0(n) \tag{2.33}$$
$$\tilde{g}(n) = h_1(n) \tag{2.34}$$

The synthesis Equation (2.28) can be written as:

$$F(2^{j+1}) = \widetilde{H}F(2^j) + \widetilde{G}W(2^j) \tag{2.35}$$

The combining process can be continued to any level by combining the appropriate approximation and wavelet coefficients. The implementation of the analysis and synthesis equations is illustrated in Fig. 2.1.

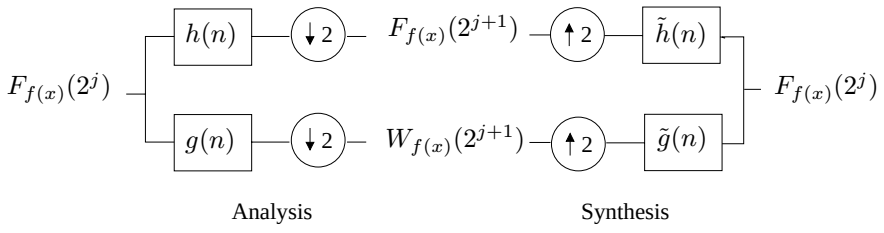


Figure 2.1.: 1D analysis and synthesis filter bank. The down-pointing arrow $\downarrow 2$ means "keep one sample out of two". The up-pointing arrow $\uparrow 2$ means "put one zero between consecutive samples". The boxes denote convolution with the filter which is inside the box.

2.4.2. Two dimensional Mallat's algorithm

In two dimensions, the DWT can be computed with a pyramidal algorithm similar to the one dimensional algorithm described in Section 2.4.1. It is again assumed that the initial, discrete data, which is a sequence $\{f(k_x, k_y)\}_{(k_x,k_y)\in\mathbb{Z}^2}$ already represents an approximation of $f(x, y)$ at scale 2^0.

$$f(k_x, k_y) \overset{\text{by convention}}{=} F_{f(x,y)}(2^0, k_x, k_y) =< f(x, y), \phi(x - k_x)\phi(y - k_y) > \tag{2.36}$$

Analysis

As in the one dimensional case, the approximation coefficient at scale 2^j can be calculated as follows:

$$F_{f(x,y)}(2^j, k_x, k_y) = \sum_{n_x,n_y} h_0^*(n_x - 2k_x)h_0^*(n_y - 2k_y)F_{f(x,y)}(2^{j-1}, k_x, k_y) \tag{2.37}$$

and the wavelet coefficients are given by:

$$W_{h,f(x,y)}(2^j, k_x, k_y) = \sum_{n_x,n_y} h_0^*(n_x - 2k_x)h_1^*(n_y - 2k_y)F_{f(x,y)}(2^{j-1}, k_x, k_y) \tag{2.38}$$

$$W_{v,f(x,y)}(2^j, k_x, k_y) = \sum_{n_x,n_y} h_1^*(n_x - 2k_x)h_0^*(n_y - 2k_y)F_{f(x,y)}(2^{j-1}, k_x, k_y) \tag{2.39}$$

$$W_{d,f(x,y)}(2^j, k_x, k_y) = \sum_{n_x,n_y} h_1^*(n_x - 2k_x)h_1^*(n_y - 2k_y)F_{f(x,y)}(2^{j-1}, k_x, k_y) \tag{2.40}$$

Synthesis

The reconstruction of the original small scale coefficient can be made from a combination of the approximation and wavelet coefficients at a larger scale:

$$\begin{aligned} F_{f(x,y)}(2^j, k_x, k_y) \quad = \sum_{n_x,n_y} \quad & h_0(k_x - 2n_x)h_0(k_y - 2n_y)F_{f(x,y)}(2^{j+1}, k_x, k_y) \\ + \quad & h_0(k_x - 2n_x)h_1(k_y - 2n_y)W_{h,f(x,y)}(2^{j+1}, k_x, k_y) \\ + \quad & h_1(k_x - 2n_x)h_0(k_y - 2n_y)W_{v,f(x,y)}(2^{j+1}, k_x, k_y) \\ + \quad & h_1(k_x - 2n_x)h_1(k_y - 2n_y)W_{d,f(x,y)}(2^{j+1}, k_x, k_y) \end{aligned} \tag{2.41}$$

Implementation as a filter bank

In two dimensions, the wavelet transform can be seen as two dimensional wavelet transforms, one along the x axis and one along the y axis. It can be computed with a separable two dimensional extension of the one dimensional decomposition algorithm described in Section 2.4.1.

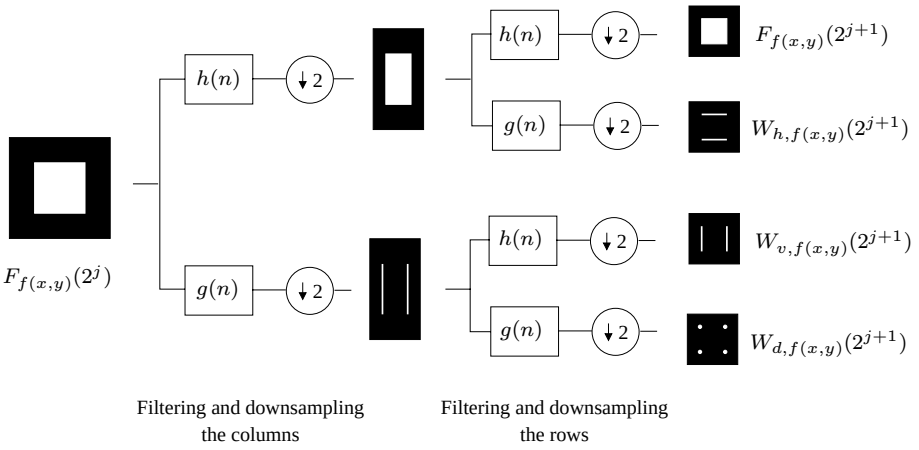


Figure 2.2.: Decomposition of an image $F_{f(x,y)}(2^j)$ into $F_{f(x,y)}(2^{j+1})$, $W_{h,f(x,y)}(2^{j+1}, k_x, k_y)$, $W_{v,f(x,y)}(2^{j+1}, k_x, k_y)$ and $W_{d,f(x,y)}(2^{j+1}, k_x, k_y)$ with the 2D Mallat transform.

For the analysis, we first convolve the rows of $F_{f(x,y)}(2^{j-1}, k_x, k_y)$ with a one dimensional filter, retain every other row, convolve the columns of the resulting signals with another one dimensional filter and retain every other column. The filters $h(n)$ and $g(n)$ used in this decomposition are defined by Equations (2.29) and (2.30). If we introduce the following operators:

$$
\begin{aligned}
H_2 &= \text{1D conv. of rows with } h(n) + \text{downsampling of rows by 2} \\
&+ \text{1D conv. of columns with } h(n) + \text{downsampling of columns by 2} \\
\\
G_h &= \text{1D conv. of rows with } h(n) + \text{downsampling of rows by 2} \\
&+ \text{1D conv. of columns with } g(n) + \text{downsampling of columns by 2} \\
\\
G_v &= \text{1D conv. of rows with } g(n) + \text{downsampling of rows by 2} \\
&+ \text{1D conv. of columns with } h(n) + \text{downsampling of columns by 2} \\
\\
G_d &= \text{1D conv. of rows with } g(n) + \text{downsampling of rows by 2} \\
&+ \text{1D conv. of columns with } g(n) + \text{downsampling of columns by 2}
\end{aligned}
$$

The analysis Equations (2.37) to (2.40) can be written as:

$$\begin{array}{lcll} F_{f(x,y)}(2^j) & \stackrel{\triangle}{=} & F_{F_{f(x,y)}(2^{j-1})} & = H_2 F_{f(x,y)}(2^{j-1}) \\ W_{h,f(x,y)}(2^j) & \stackrel{\triangle}{=} & W_{h,F_{f(x,y)}(2^{j-1})} & = G_h F_{f(x,y)}(2^{j-1}) \\ W_{v,f(x,y)}(2^j) & \stackrel{\triangle}{=} & W_{v,F_{f(x,y)}(2^{j-1})} & = G_v F_{f(x,y)}(2^{j-1}) \\ W_{d,f(x,y)}(2^j) & \stackrel{\triangle}{=} & W_{d,F_{f(x,y)}(2^{j-1})} & = G_d F_{f(x,y)}(2^{j-1}) \end{array} \tag{2.42}$$

For the synthesis, we add between each column of the images $F_{f(x,y)}(2^{j+1}, k_x, k_y)$, $W_{h,f(x,y)}(2^{j+1}, k_x, k_y)$, $W_{v,f(x,y)}(2^{j+1}, k_x, k_y)$ and $W_{d,f(x,y)}(2^{j+1}, k_x, k_y)$ a column of zeros, convolve the rows with a one dimensional filter, add a row of zeros between each row of the resulting image, and convolve with another one dimensional filter. The filters $\tilde{h}(n)$ and $\tilde{g}(n)$ used in the reconstruction are defined by Equations (2.33) and (2.34). If we introduce the following operators:

$$\begin{array}{lcl} \widetilde{H}_2 & = & \text{upsampling of columns by 2 + 1D conv. of columns with } \tilde{h}(n) \\ & + & \text{upsampling of rows by 2 + 1D conv. of columns with } \tilde{h}(n) \end{array}$$

$$\begin{array}{lcl} \widetilde{G}_h & = & \text{upsampling of columns by 2 + 1D conv. of columns with } \tilde{g}(n) \\ & + & \text{upsampling of rows by 2 + 1D conv. of columns with } \tilde{h}(n) \end{array}$$

$$\begin{array}{lcl} \widetilde{G}_v & = & \text{upsampling of columns by 2 + 1D conv. of columns with } \tilde{h}(n) \\ & + & \text{upsampling of rows by 2 + 1D conv. of columns with } \tilde{g}(n) \end{array}$$

$$\begin{array}{lcl} \widetilde{G}_d & = & \text{upsampling of columns by 2 + 1D conv. of columns with } \tilde{g}(n) \\ & + & \text{upsampling of rows by 2 + 1D conv. of columns with } \tilde{g}(n) \end{array}$$

The synthesis Equation (2.41) can be written as:

$$\begin{array}{lcl} F_{f(x,y)}(2^j) & = & \widetilde{H}_2 F_{f(x,y)}(2^{j+1}) + \widetilde{G}_h W_{h,f(x,y)}(2^{j+1}) \\ & + & \widetilde{G}_v W_{v,f(x,y)}(2^{j+1}) + \widetilde{G}_d W_{d,f(x,y)}(2^{j+1}) \end{array} \tag{2.43}$$

If the initial image $f(k_x, k_y)$ is a NxN matrix, it is transformed into four matrices of dimension $N/2$x$N/2$. These latter can be stored again in the original matrix and only the approximation image is decomposed further. Figs. 2.2 and 2.3 illustrate the 2D Mallat wavelet decomposition and reconstruction on the example of a white square on a black background. Fig. 2.4 represents the disposition of the wavelet and approximation images for the first two decomposition levels of the 2D Mallat wavelet transform.

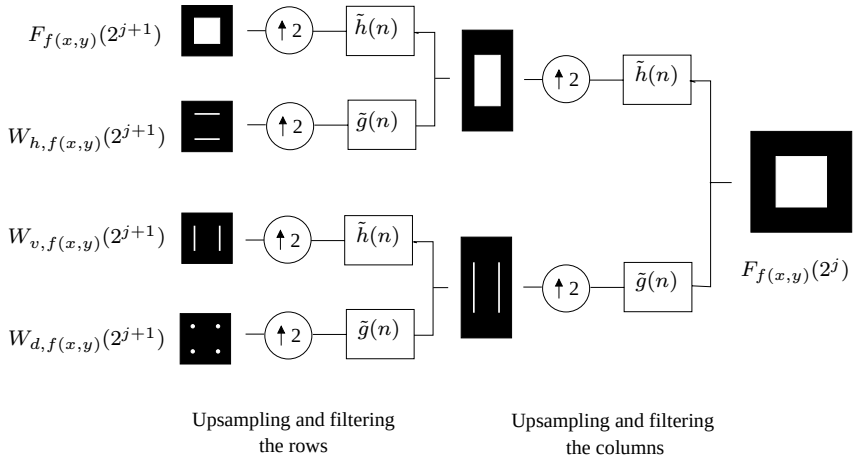


Figure 2.3.: Reconstruction of an image $F_{f(x,y)}(2^j)$ from $F_{f(x,y)}(2^{j+1})$, $W_{h,f(x,y)}(2^{j+1}, k_x, k_y)$, $W_{v,f(x,y)}(2^{j+1}, k_x, k_y)$ and $W_{d,f(x,y)}(2^{j+1}, k_x, k_y)$ with the 2D Mallat transform.

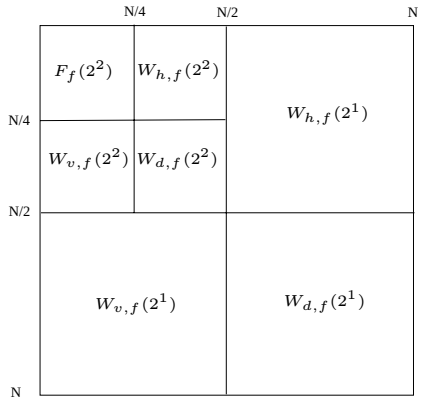


Figure 2.4.: Disposition of the wavelet and approximation images for the first two decomposition steps of the 2D Mallat wavelet transform.

For the unnormalized Haar wavelet used for DeQuant, the analysis and synthesis filters will be taken equal to:

$$h = [1,\ 1] \qquad g = [1,\ -1] \tag{2.44}$$

$$\tilde{h} = [1,\ 1] \qquad \tilde{g} = [-1,\ 1] \tag{2.45}$$

This means that the analysis equations reduce to:

$$\begin{aligned} F_{f(x,y)}(2^{j+1},k_x,k_y) &= \gamma_1+\gamma_2+\gamma_3+\gamma_4 \\ W_{h,f(x,y)}(2^{j+1},k_x,k_y) &= \gamma_1-\gamma_2+\gamma_3-\gamma_4 \\ W_{v,f(x,y)}(2^{j+1},k_x,k_y) &= \gamma_1+\gamma_2-\gamma_3-\gamma_4 \\ W_{d,f(x,y)}(2^{j+1},k_x,k_y) &= \gamma_1-\gamma_2-\gamma_3+\gamma_4 \end{aligned} \tag{2.46}$$

where

$$\begin{aligned} \gamma_1 &= F_{f(x,y)}(2^j,k_x,k_y) & \gamma_2 &= F_{f(x,y)}(2^j,k_x,k_y+1) \\ \gamma_3 &= F_{f(x,y)}(2^j,k_x+1,k_y) & \gamma_4 &= F_{f(x,y)}(2^j,k_x+1,k_y+1) \end{aligned}$$

Since the filter coefficients are unnormalized, the additional multiplicative factor $1/4$ has to be taken into account in the synthesis equation (2.43)

$$\begin{aligned} F_{f(x,y)}(2^j) = \big[\, &\widetilde{H}_2 F_{f(x,y)}(2^{j+1}) + \widetilde{G}_h W_{h,f(x,y)}(2^{j+1}) \\ &+ \widetilde{G}_v W_{v,f(x,y)}(2^{j+1}) + \widetilde{G}_d W_{d,f(x,y)}(2^{j+1}) \,\big] \,/\, 4 \end{aligned} \tag{2.47}$$

2.5. Shift-invariant wavelet transform

Mallat's algorithm computes the wavelet expansion on a discrete grid corresponding to scales $a = 2^j$ and shifts $b = k2^j$. The associated wavelets form an orthonormal basis but the transform is not shift invariant.

We derive in this section for the special choice of the unnormalized Haar wavelet, the analysis and synthesis equations of the shift invariant wavelet transform introduced by Coifman and Donoho [CD95a]. This undecimated wavelet transform where scales are still restricted to powers of two, but shifts are now arbitrary integers, can be seen as a special case of the so-called *à trous* algorithm introduced by Holshneider et al. [MHT89]. It is a special case because (i) instead of using one wavelet filter, it associates three wavelet directional filters with the analysis and (ii) the synthesis is done by averaging out four possible reconstructions [SMB98].

Analysis equations

Let $F^T_{f(x,y)}(2^j,k_x,k_y)$ be the scaling coefficient of $f(x,y)$ at scale 2^j and location (k_x,k_y) and $W^T_{h,f(x,y)}(2^j,k_x,k_y)$, $W^T_{v,f(x,y)}(2^j,k_x,k_y)$ and $W^T_{d,f(x,y)}(2^j,k_x,k_y)$

its horizontal, vertical and diagonal wavelet coefficients. The superscript T refers to the undecimated wavelet transform of type à **T**rous. The analysis equations are the answer to the question: how does the undecimated wavelet transform with the unnormalized Haar wavelet calculate the scaling and wavelet coefficients at the next scale 2^{j+1} and the same position (k_x, k_y)? In order to answer this question we introduce the following notations for the scaling and wavelet coefficients at scale 2^j and location $(k_x + m2^j, k_y + n2^j)$

$$\begin{aligned} \gamma^j_{mn} &= F^T_{f(x,y)}(2^j, k_x + m2^j, k_y + n2^j) \\ wh^j_{mn} &= W^T_{h,f(x,y)}(2^j, k_x + m2^j, k_y + n2^j) \\ wv^j_{mn} &= W^T_{v,f(x,y)}(2^j, k_x + m2^j, k_y + n2^j) \\ wd^j_{mn} &= W^T_{d,f(x,y)}(2^j, k_x + m2^j, k_y + n2^j) \end{aligned}$$

The analysis equations are then given by:

$$\begin{aligned} \gamma^{j+1}_{00} &= \gamma^j_{00} + \gamma^j_{10} + \gamma^j_{01} + \gamma^j_{11} \\ wh^{j+1}_{00} &= \gamma^j_{00} + \gamma^j_{10} - \gamma^j_{01} - \gamma^j_{11} \\ wv^{j+1}_{00} &= \gamma^j_{00} - \gamma^j_{10} + \gamma^j_{01} - \gamma^j_{11} \\ wd^{j+1}_{00} &= \gamma^j_{00} - \gamma^j_{10} - \gamma^j_{01} + \gamma^j_{11} \end{aligned} \tag{2.48}$$

In order to write Equations (2.48) with the operator formalism, we introduce the analysis operators $H_2(2^j)$, $G_h(2^j)$, $G_v(2^j)$ and $G_d(2^j)$, where the operator $R(2^j)$ denotes convolution by the two dimensional filter $r(2^j)$. This latter is represented by a square matrix of size $2^j + 1$ obtained by inserting $2^j - 1$ columns of zeros between the columns of $r(2^0)$, and $2^j - 1$ rows of zeros between the rows of $r(2^0)$. The matrices $r(2^0)$ are defined by Table 2.2.

$h_2(2^0)$	$g_h(2^0)$	$g_v(2^0)$	$g_d(2^0)$
$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

Table 2.2.: Two dimensional analysis filters for the shift invariant wavelet transform.

The analysis equations (2.48) can then be written as:

$$\begin{aligned} F^T_{f(x,y)}(2^{j+1}) &\triangleq F^T_{F^T_{f(x,y)}(2^j)} &= H_2(2^j)F^T_{f(x,y)}(2^j) \\ W^T_{h,f(x,y)}(2^{j+1}) &\triangleq W^T_{h,F^T_{f(x,y)}(2^j)} &= G_h(2^j)F^T_{f(x,y)}(2^j) \\ W^T_{v,f(x,y)}(2^{j+1}) &\triangleq W^T_{v,F^T_{f(x,y)}(2^j)} &= G_v(2^j)F^T_{f(x,y)}(2^j) \\ W^T_{d,f(x,y)}(2^{j+1}) &\triangleq W^T_{d,F^T_{f(x,y)}(2^j)} &= G_d(2^j)F^T_{f(x,y)}(2^j) \end{aligned} \tag{2.49}$$

Synthesis equation

The synthesis exploits the redundancy of the analysis by averaging out all possible reconstructions of one pixel. To make our meaning clear, we will demonstrate this on the reconstruction of γ^j_{00} which is the scaling coefficient at scale 2^j and position (k_x, k_y).

This scaling coefficient is used in four different sets of analysis equations see Fig. 2.5. The first set given by Equation (2.48) enables to calculate (γ^{j+1}_{00}, wh^{j+1}_{00}, wv^{j+1}_{00}, wd^{j+1}_{00}). The other three sets give (γ^{j+1}_{mn} , wh^{j+1}_{mn}, wv^{j+1}_{mn}, wd^{j+1}_{mn}), where $mn \in \{(-1)0, 0(-1), (-1)(-1)\}$. Hence reconstructing γ^j_{00} can be done by four different ways by calculating S^{j+1}_{00}, $S^{j+1}_{(-1)0}$, $S^{j+1}_{0(-1)}$ or $S^{j+1}_{(-1)(-1)}$ where S^{j+1}_{mn} is defined by:

$$\begin{aligned} S^{j+1}_{mn} = \frac{1}{4} \big[& \gamma^{j+1}_{mn} + (-1)^n wh^{j+1}_{mn} \\ & + (-1)^m wv^{j+1}_{mn} + (-1)^{m+n} wd^{j+1}_{mn} \big] \end{aligned} \tag{2.50}$$

We take advantage of this redundancy to average out those four possibilities. The computation of γ^j_{00} is thus given by:

$$\gamma^j_{00} = \frac{S^{j+1}_{00} + S^{j+1}_{(-1)0} + S^{j+1}_{0(-1)} + S^{j+1}_{(-1)(-1)}}{4} \tag{2.51}$$

At this point we introduce the synthesis operators $\widetilde{H_2}(2^j)$, $\widetilde{G_h}(2^j)$, $\widetilde{G_v}(2^j)$ and $\widetilde{G_d}(2^j)$, where the operator $\widetilde{R}(2^j)$ denotes convolution with the two dimensional filter $\widetilde{r}(2^j)$. The synthesis filters $\widetilde{r}(2^j)$ are defined like the analysis filters by "dilating" the initial matrices given by Table 2.3.

The synthesis equation (2.51) can then be written as:

$$\begin{aligned} F^T_{f(x,y)}(2^j) = \big[& \widetilde{H_2}(2^j)F^T_{f(x,y)}(2^{j+1}) + \widetilde{G_h}(2^j)W^T_{h,f(x,y)}(2^j) \\ & + \widetilde{G_v}(2^j)W^T_{v,f(x,y)}(2^j) + \widetilde{G_d}(2^j)W^T_{d,f(x,y)}(2^j) \big] \,/\, 16 \end{aligned} \tag{2.52}$$

$\widetilde{h}_2(2^0)$	$\widetilde{g}_h(2^0)$	$\widetilde{g}_v(2^0)$	$\widetilde{g}_d(2^0)$
$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

Table 2.3.: Two dimensional synthesis filters for the shift invariant wavelet transform

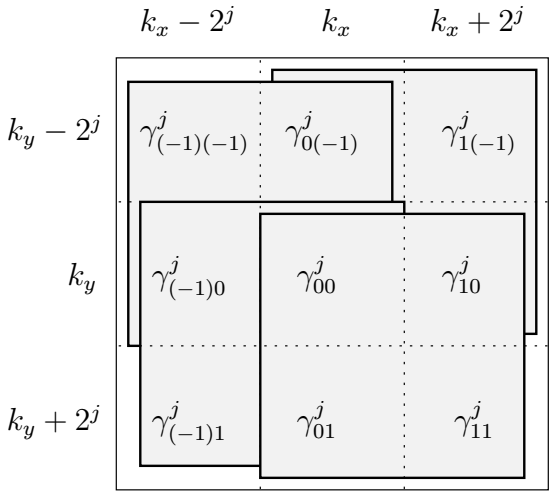


Figure 2.5.: The scaling coefficient γ^j_{00} at scale 2^j and position (k_x, k_y) appears in four different sets of analyzing equations.

2.6. Conclusion

After giving some elements of definition of a wavelet and presenting the continuous wavelet transform and some of its properties, this chapter established the link between multiresolution analysis and the discrete wavelet transform. It described using the operator formalism the wavelet transform algorithms that are used in the DeQuant framework, namely Mallat's algorithm and the shift invariant algorithm introduced by Coifman and Donoho and examined the particular case of the unnormalized Haar wavelet.

3. Some classical approaches to image restoration

Image restoration is a process that attempts to reconstruct or recover an image that has been degraded by using some a priori knowledge of the degradation phenomenon. The effectiveness of image restoration filters depends on the extent and the accuracy of the knowledge of the degradation process, the filter design criterion as well as the adequation to the considered stochastic process. Stochastic processes can be divided into two broad classes: stationary and nonstationary. While Fourier based filters are adapted to filtering linear stationary processes, nonstationary processes need adaptive, space frequency methods.

This chapter reviews some classical approaches to image restoration and examines their degree of adequation to our particular restoration problem in nuclear medicine. It concludes by pointing out the need for a new restoration method and describes its new features.

This chapter is organized as follows. Section 3.1 gives different formulations of the image formation model that will be used in the rest of the chapter. Section 3.2 presents some classical restoration methods for stationary processes. We discuss on the one hand the inverse filter and two of its variations and on the other hand two least mean square restoration methods. Section 3.3 treats the restoration of nonstationary processes. We focus thereby on wavelet methods that have been developed for Gaussian and Poisson noise. Section 3.4 concludes by pointing out the need for a new restoration method and enumerates its new features.

3.1. General formulation of the image model

The degradation process is modeled as a linear, shift invariant operator, which together with an additive noise term $n(k_x, k_y)$ operates on an input image $f(k_x, k_y)$ to produce a degraded image $g(k_x, k_y)$ [GW93].

The effect of a linear, shift invariant operator characterized by its point spread function $h(k_x, k_y)$ on the input signal is given by:

$$g(k_x, k_y) = \sum_{\alpha=0}^{N-1} \sum_{\beta=0}^{N-1} h(k_x - \alpha, k_y - \beta) f(k_x, k_y) + n(k_x, k_y) \tag{3.1}$$

Digital image restoration may be viewed as the process of obtaining an approximation to $f(k_x, k_y)$, given $g(k_x, k_y)$ and a knowledge of the degradation in the form of the point spread function $h(k_x, k_y)$. We assume that knowledge of $n(k_x, k_y)$ is limited to information of a statistical nature.

Fourier domain formulation of Equation (3.1): Assuming that the images are repeated periodically in all directions, Equation (3.1) can be written in the Fourier domain using the result of the convolution theorem:

$$G(u, v) = H(u, v)F(u, v) + N(u, v) \tag{3.2}$$

where $G(u, v)$, $H(u, v)$, $F(u, v)$ and $N(u, v)$ are the discrete Fourier transforms of g, h, f and n respectively.

Matrix formulation of Equation (3.1): The effect of a linear operator on an image can also be expressed with the help of matrices. Let $\mathbf{g}$, $\mathbf{f}$ and $\mathbf{n}$ represent N^2 columns vectors formed by stacking the rows of the N x N functions $g(k_x, k_y)$, $f(k_x, k_y)$ and $n(k_x, k_y)$. The first N elements of $\mathbf{f}$, for example, are the elements in the first row of $f(k_x, k_y)$, the next N elements are from the second row, and so on for all N rows of $f(k_x, k_y)$. This convention allows Equation (3.1) to be expressed in vector-matrix form:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n} \tag{3.3}$$

where $\mathbf{g}$, $\mathbf{f}$ and $\mathbf{n}$ are of dimensions N^2x1 and $\mathbf{H}$ is of dimensions N^2xN^2. This matrix consists of N^2 partitions, each partition being of size NxN and ordered according to:

$$\begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_{N-1} & \mathbf{H}_{N-2} & \dots & \mathbf{H}_1 \\ \mathbf{H}_1 & \mathbf{H}_0 & \mathbf{H}_{N-1} & \dots & \mathbf{H}_2 \\ \mathbf{H}_2 & \mathbf{H}_1 & \mathbf{H}_0 & \dots & \mathbf{H}_3 \\ \vdots & & & & \\ \mathbf{H}_{N-1} & \mathbf{H}_{N-2} & \mathbf{H}_{N-3} & \dots & \mathbf{H}_0 \end{bmatrix} \tag{3.4}$$

Each partition $\mathbf{H}_j$ is constructed from the jth row of the function $h(k_x, k_y)$ as follows:

$$\begin{bmatrix} h(j,0) & h(j,N-1) & h(j,N-2) & \dots & h(j,1) \\ h(j,1) & h(j,0) & h(j,N-1) & \dots & h(j,2) \\ h(j,2) & h(j,1) & h(j,0) & \dots & h(j,3) \\ \vdots & & & & \\ h(j,N-1) & h(j,N-2) & h(j,N-3) & \dots & h(j,0) \end{bmatrix} \tag{3.5}$$

The circular behavior of $\mathbf{H}_j$ is a direct consequence of the assumed periodicity of $h(k_x, k_y)$. Moreover, the blocks of $\mathbf{H}$ are subscripted in a circulant manner. For these reasons the matrix $\mathbf{H}$ is often called a block circulant matrix.

3.2. Restoration methods for stationary processes

A stochastic process $X(k)$ is called strict-sense stationary if its statistical properties are invariant to a shift of the origin. This means that the processes $X(k)$ and $X(k + k_0)$ have the same statistics for any k_0 [Pap84]. We present in the following some restoration methods for stationary processes.

3.2.1. The inverse filter and its variations

Inverse filtering is the process of recovering the input of a system from its output. In this section, we present the inverse filter as well as two derived implementations, Janson - Van Cittert and Metz, based on another writing of the filter.

Inverse filter

When the point spread function of the degradation is assumed to be known, one approach to the image restoration problem is inverse filtering [GW93]. Using Equation (3.2), the estimate $\hat{F}(u, v)$ of the object is simply obtained by dividing the image by the point spread function in the Fourier domain:

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)} = F(u,v) + \frac{N(u,v)}{H(u,v)} \tag{3.6}$$

This straightforward approach produces unacceptably poor results. Computational difficulties are encountered in the restoration process if $H(u, v)$ has any zeros. Moreover, in many cases $H(u, v)$ often drops off rapidly as a function of

distance from the origin of the (u, v) plane. The noise term, however, usually falls off at a much slower rate. In such situations, the term $N(u, v)/H(u, v)$ could dominate the restoration result, which means that the noise term may be enormously amplified.

Janson - Van Cittert

Van Cittert [VC31] developed another method of implementing the inverse filter by using an iterative procedure where an estimate of the object is updated in each iteration. It is based on writing $1/H(u, v)$ as follows:

$$\frac{1}{H(u,v)} = \frac{1}{1-(1-H(u,v))} \tag{3.7}$$

Provided that $|1-H(u, v)| < 1$, Equation (3.7) can be approximated at the order k by the expression

$$\sum_{i=0}^{k}(1-H(u,v))^i \tag{3.8}$$

Let $\hat{F}^{(k)}(u, v)$ denote the estimate of the Fourier transform of the object at iteration k. We have

$$\begin{aligned}
\hat{F}^{(k)}(u,v) &= \left[\sum_{i=0}^{k}(1-H(u,v))^i\right] G(u,v) \\
&= \left[1+(1-H(u,v))\sum_{i=0}^{k-1}(1-H(u,v))^i\right] G(u,v) \\
&= G(u,v)+(1-H(u,v))\hat{F}^{(k-1)}(u,v) \\
&= \hat{F}^{(k-1)}(u,v)+(G(u,v)-H(u,v)\hat{F}^{(k-1)}(u,v))
\end{aligned}$$

Janson integrated a parameter ξ in the Van Cittert iteration. This parameter is used to control the convergence behavior of the iterative procedure:

$$\hat{F}^{(k)}(u,v) = \hat{F}^{(k-1)}(u,v)+\xi(G(u,v)-H(u,v)\hat{F}^{(k-1)}(u,v)) \tag{3.9}$$

If $\hat{F}^{(k-1)}(u, v)$ is a good estimate of $F(u, v)$, $H(u, v)\hat{F}^{(k-1)}(u, v)$ will be close to $G(u, v)$. The Fourier transform of the estimated object after k iterations is obtained by adding to $\hat{F}^{(k-1)}(u, v)$ a correction term which consists of a scaling constant ξ multiplied to the difference between $G(u, v)$ and $H(u, v)\hat{F}^{(k-1)}(u, v)$.

For the points where $H(u, v) = 0$, it can be seen from Equation (3.9) that the correction term equals $\xi G(u, v)$. This will cause the iteration procedure to diverge. This can be avoided by stopping the procedure after a few iterations. The result after a finite number of iterations is not in general the same as that of inverse filtering, but it is usually less sensitive to the presence of noise.

Metz filter

The Metz filter [IN67] is a variation of the inverse filter based on writing $1/H(u,v)$ as follows:

$$\frac{1}{H(u,v)} = \frac{H^*(u,v)}{|H(u,v)|^2} = \frac{H^*(u,v)}{1-(1-|H(u,v)|^2)} = H^*(u,v)\sum_{i=0}^{\infty}(1-|H(u,v)|^2)^i \tag{3.10}$$

He then approximated the infinite sum in Equation (3.10) by a finite sum of p terms:

$$H^*(u,v)\sum_{i=0}^{p-1}(1-|H(u,v)|^2)^i = \frac{1-(1-|H(u,v)|^2)^p}{H(u,v)} \tag{3.11}$$

For $|H(u,v)| \to 1$ the Metz filter becomes nearly equal to the inverse filter. For $|H(u,v)| \to 0$ the Metz filter tends to zero:

$$\frac{1-(1-|H(u,v)|^2)^p}{H(u,v)} \longrightarrow \begin{cases} \frac{1}{H(u,v)} & \text{for} \quad |H(u,v)| \to 1 \\ pH(u,v)^* & \text{for} \quad |H(u,v)| \to 0 \end{cases} \tag{3.12}$$

Since the blurring process is usually a low-pass filter, Equation (3.12) can be interpreted as a low-pass filter

The behaviour of the Metz filter is similar to that of the Wiener filter since it achieves a compromise between the low-pass smoothing filter and the high-pass inverse filter resulting in a band pass filter. The order p determines the cut-off frequency and the maximum of the filter.

3.2.2. Least Mean Square Methods

This section treats the restoration problem through the minimization of a predefined criterion of performance. Because of their simplicity, we focus here on least squares criterion functions and present the Gradient method and the Wiener filter.

Landweber or Gradient method

From Equation (3.3), the noise term in the degradation model is:

$$\mathbf{n} = \mathbf{g} - \mathbf{Hf} \tag{3.13}$$

In the absence of any knowledge about $\mathbf{n}$, a meaningful criterion function is to seek an $\hat{\mathbf{f}}$ such that $\mathbf{H}\hat{\mathbf{f}}$ approximates $\mathbf{g}$ in a least square sense by assuming that the norm of the noise term is as small as possible. In other words, we want to find $\hat{\mathbf{f}}$ such that

$$\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \tag{3.14}$$

is minimum. This is equivalent to minimizing the criterion function

$$J(\hat{\mathbf{f}}) = \|\mathbf{g} - \mathbf{H}\hat{\mathbf{f}}\|^2 \tag{3.15}$$

with respect to $\hat{f}$. After differentiating $J(f)$ with respect to f and setting the result equal to the zero vector, we obtain:

$$\mathbf{H}^t\mathbf{H}\hat{\mathbf{f}} = \mathbf{H}^t\mathbf{g} \tag{3.16}$$

Writing Equation (3.16) in the Fourier domain gives:

$$H^*(u,v)H(u,v)\hat{F}(u,v) = H^*(u,v)G(u,v) \tag{3.17}$$

At points (u,v) where $H(u,v) = 0$, $H^*(u,v)$ will also be zero, the ratio

$$H^*(u,v)/H^*(u,v)H(u,v)$$

will be $0/0$, i.e undetermined. All this means that for the particular frequencies (u,v) where $H(u,v) = 0$, the frequency content of the original image cannot be recovered. One can overcome this problem by using the Janson - Van Cittert iteration procedure on Equation (3.17). We obtain:

$$\hat{F}^{(k)}(u,v) = \hat{F}^{(k-1)}(u,v) + \xi H^*(u,v)(G(u,v) - H(u,v)\hat{F}^{(k-1)}(u,v)) \tag{3.18}$$

This method named after Landweber [Lan51] is more robust to noise than the Janson - Van Cittert method [VC31]. Indeed, at the points where $H(u,v) = 0$, it can be seen from Equation (3.18) that the correction term equals now $\xi G(u,v)H^*(u,v)$ which also equals zero. Since the correction term added to $\hat{F}^{(k-1)}(u,v)$ is zero, the procedure will not diverge.

Wiener-Helstrom

Let $f(k_x,k_y)$ and $g(k_x,k_y)$ be random sequences. It is desired to obtain an estimate, $\hat{f}(k_x,k_y)$, of $f(k_x,k_y)$ such that the mean square error

$$e^2 = E[\{f(k_x,k_y) - \hat{f}(k_x,k_y)\}^2]$$

is minimized. The best estimate $\hat{f}(k_x, k_y)$ is known to be the conditional mean of $f(k_x, k_y)$ given $g(k_x, k_y)$, that is

$$\hat{f}(k_x, k_y) = E[f(k_x, k_y)/g(k_x, k_y)] \tag{3.19}$$

Equation (3.19) is quite difficult to solve in general [Jai89]. Therefore one generally settles for the best linear estimate. This latter can be written in the Fourier domain as

$$\hat{F}(u, v) = T(u, v)G(u, v)$$

The Wiener filter $T(u, v)$ is given by [GW93]:

$$T(u, v) = \frac{P_{ff}(u, v)}{P_{ff}(u, v) + P_{nn}(u, v)}$$

$P_{ff}(u, v)$ and $P_{nn}(u, v)$ are the power spectra of the signal and the noise. It is well known that if $f(k_x, k_y)$ and $g(k_x, k_y)$ are jointly Gaussian sequences, then the solution of (3.19) is linear [Jai89].

In the absence of blur, the Wiener filter depends only on the signal to noise ratio $SNR = \frac{P_{ff}}{P_{nn}}$. For frequencies where $SNR \gg 1$, $T(u, v)$ becomes nearly equal to unity which means that all these frequency components are in the passband. When $SNR \ll 1$, $T(u, v) \approx SNR$. This means that all frequency components are attenuated proportionally to their signal to noise ratio. For images, SNR is usually high at lower spatial frequencies. Therefore, the noise smoothing filter is a low-pass filter. In the presence of blur, the expression of the filter is [Hel67]:

$$T_h(u, v) = \frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + \frac{P_{nn}(u,v)}{P_{ff}(u,v)}} \tag{3.20}$$

The expression (3.20) is obtained under the following assumptions:

- $g(k_x, k_y)$, $f(k_x, k_y)$ and $n(k_x, k_y)$ are supposed to be ergodic with respect to the mean and the autocorrelation function. This implies that they are also supposed to be mean value and covariance stationary.
- the noise $n(k_x, k_y)$ and the true image $f(k_x, k_y)$ are uncorrelated
- at least one of the noise or the true image $f(k_x, k_y)$ has zero mean

Equation (3.20) shows that the Wiener filter simply determines a correction factor with which the inverse transfer function of the degradation process has to be multiplied before it is used as a filter, so that the effect of noise is taken care

of. It achieves a compromise between the low-pass noise smoothing filter and the high-pass inverse filter resulting in a band-pass filter.

All filters presented in this section were developed under the assumption that the characteristics of the image and noise do not change over the different regions of the image. This has resulted in space invariant filters. In a typical nuclear medicine image, we have to deal with a double nonstationarity. On the one hand, the image characteristics differ considerably from one region to another since the repartition of the nodules follows biological processes. And on the other hand, the noise varies from one region to another since it depends on the signal. It is reasonable, then, to adapt the processing to the changing characteristics of the image and degradation.

3.3. Restoration methods for nonstationary processes

Nonstationary processes necessitate restoration methods that adapt to their varying statistical properties. Whereas a lot of the adaptive procedures proposed in the literature can be assimilated to a kind of "recipe" lacking generality, the wavelet transform provides a general framework that enables the filtering procedure to adapt to the local variations in the signal frequency content and thereby balances the trade off between noise removal and excessive smoothing. Its multiscale and localizing or concentrating properties makes it particularly effective for restoration problems.

In the following, we first discuss an example of an adaptive image restoration system, the so-called "Noise Cheating enhancement", chosen from among the many proposed in the literature. We then present some wavelet-based restoration methods and explain why these methods are inadequate for our particular restoration problem.

3.3.1. An example of adaptive method

The Noise Cheating technique developed by Zweig and al. [ZBH75] adapts the resolution to the signal to noise ratio. It is based on the fact that in the case of a Poisson process the number of photons detected at a given pixel must be superior to a threshold T to insure that the signal to noise ratio at this pixel is big enough.

Their method can be described as follows. Let n be the number of photons detected at position (k_x, k_y), i.e, if we use the notations of Equation (3.1) $g(k_x, k_y) = n$. Let A_i denote the pixel (k_x, k_y) together with its "immediate" neighborhood: A_1 corresponds to (k_x, k_y) plus the four surrounding pixels, A_2 to (k_x, k_y) plus the 8 surrounding pixels . . . .

- if $n > T$ then

$$f(k_x, k_y) = g(k_x, k_y)$$

- if $n \leq T$, calculate the sum S_i of photon counts in the areas A_i starting with the smallest one until $S_i > T$. If i_0 denotes the first value of i for which this latter condition is verified then

$$f(k_x, k_y) = \frac{S_{i_0}}{A_{i_0}}$$

3.3.2. Wavelet based methods

We present in the following some wavelet based adaptive restoration methods. Thereby, we consider separately the methods developed for Gaussian noise and those for Poisson noise. For simplicity, all methods will be described in the one dimensional case. An extension to two dimensional images is straightforward.

Wavelet methods for Gaussian noise

Consider the finite length signal $g(k)$ of observations of the signal $f(k)$ corrupted by independent identically distributed (iid) zero mean, white Gaussian noise $n(k)$ with standard deviation σ.

$$g(k) = f(k) + n(k), \quad k = 1, \ldots, N \tag{3.21}$$

We present in the following some wavelet based restoration methods that yield the signal estimate $f(k)$.

Starck and Bijaoui: Wiener-like filtering in the wavelet space [SP96]. In addition to the general hypothesis above, we consider here that $f(k)$ is the realization of a random variable following a Gaussian distribution. The idea is to find the wavelet coefficients $W_{\hat{f}}(2^j, k)$ that minimize the mean square error

$$e^2 = \mathrm{E}[\{W_f(2^j, k) - W_{\hat{f}}(2^j, k)\}^2]$$

The best estimate $W_{\hat{f}}(2^j,k)$ is known to be the conditional mean of $W_f(2^j,k)$ given $W_g(2^j,k)$.

$$W_{\hat{f}}(2^j,k) = \mathrm{E}[W_f(2^j,k)/W_g(2^j,k)] \tag{3.22}$$

In order to determine (3.22), Starck et al. calculated $\mathrm{Prob}(W_f(2^j,k)/W_g(2^j,k))$ using the Bayes theorem. For simplicity the position parameter k has been omitted in the following equation.

$$\mathrm{Prob}\left(\, W_f(2^j)/W_g(2^j)\,\right) = \frac{\mathrm{Prob}\,(\, W_g(2^j)/W_f(2^j)\,)\,\mathrm{Prob}\,(\, W_f(2^j)\,)}{\mathrm{Prob}\,(\, W_g(2^j)\,)}$$

Using the linearity property of the wavelet transform, Equation (3.21) becomes

$$W_g(2^j,k) = W_f(2^j,k) + W_n(2^j,k)$$

where $W_f(2^j,k)$, $W_n(2^j,k)$ and $W_g(2^j,k)$ follow Gaussian distributions of variances $\sigma^2_{W_f}(2^j,k)$, $\sigma^2_{W_n}(2^j)$ and $\sigma^2_{W_g}(2^j) = \sigma^2_{W_f}(2^j) + \sigma^2_{W_n}(2^j)$ respectively. This implies that $\mathrm{Prob}\,(\, W_f(2^j)/W_g(2^j)\,)$ follows also a Gaussian distribution of conditional mean:

$$\mathrm{E}[W_f(2^j,k)/W_g(2^j,k)] = \frac{\sigma^2_{W_f}(2^j)}{\sigma^2_{W_f}(2^j) + \sigma^2_{W_n}(2^j)} W_g(2^j,k)$$

The filtering algorithm can be described as follows:

- Compute the wavelet transform of $g(k)$. Starck et al. used the B-spline wavelet.
- Calculate $\hat{\sigma}^2_{W_n}(2^1)$ from the histogram of $W_g(2^1,k)$, since the values of the wavelet coefficients at scale 2^1 are mainly due to noise. The histogram shows a Gaussian peak around zero. The standard deviation at scale 2^1 is computed with a 3-sigma clipping, rejecting the pixels where the signal could be significant. The standard deviation at scale 2^j, can then be written as $\hat{\sigma}^2_{W_n}(2^j) = \kappa(2^j)\hat{\sigma}^2_{W_n}(2^1)$, where $\kappa(2^j)$ depends on the used wavelet transform algorithm. In the case of the Mallat algorithm with orthonormal wavelets is $\kappa(2^j) = 1$. For the à trous algorithm, $\kappa(2^j)$ depends on the value of the coefficients of the analysis and synthesis filters, see [Bob97].
- Estimate $\hat{\sigma}^2_{W_g}(2^j)$ and calculate $\hat{\sigma}^2_{W_f}(2^j) = \hat{\sigma}^2_{W_g}(2^j) - \hat{\sigma}^2_{W_n}(2^j)$.
- Modify the wavelet coefficients according to:

$$W_{\hat{f}}(2^j,k) = \frac{\hat{\sigma}^2_{W_f}(2^j)}{\hat{\sigma}^2_{W_f}(2^j) + \hat{\sigma}^2_{W_n}(2^j)}\, W_g(2^j,k)$$

- Apply the inverse wavelet transform and reconstruct the estimate $\hat{f}(k)$.

Donoho and Johnstone: Wavelet shrinkage [DJ94] The idea of using wavelets for denoising was initiated by Donoho and Johnstone [DJ94, Don75, Don93b, DJKP95]. The solution they proposed, called wavelet shrinkage, has proven to be quite successful in a number of applications. It can be described as follows.

Suppose that the observations contain only noise. Under this null hypothesis the wavelet coefficients $W_g(2^j, k)$ will be iid, Gaussian variables with zero mean and standard deviation σ. The threshold $t_N^G = \sigma\sqrt{2\log N}$ is a probabilistic upper bound on the coefficients in the sense that:

$$\text{Prob}\left(\max W(2^j, k) \leq t_N^G\right) \to 1 \qquad \text{as } N \to \infty, \tag{3.23}$$

see Leadbetter et al. [LLR83], page 14. Combined with the fact that the value $-t_N^G$ is in turn a probabilistic lower bound on the coefficients, this result suggests using the value t_N^G as a single threshold on the absolute value of the wavelet coefficients. For data in which the underlying signal may be well compressed by a wavelet transform, this threshold should serve to separate the "signal" and the "noise" into sets of "large" and "small" coefficients, respectively. Application of the appropriate inverse wavelet transform to the thresholded coefficients yields a denoised estimate of the object underlying the data. The algorithm can then be described as follows:

- Take the orthogonal discrete wavelet transform of $g(k)$.
- Apply one of the following non-linear transformations to the wavelet coefficients:
  - Hard thresholding

$$W_{\hat{f}}(2^j, k) = \begin{cases} W_g(2^j, k) & \text{if } |W_g(2^j, k)| > t \\ 0 & \text{if } |W_g(2^j, k)| \leq t \end{cases}$$

  - Soft thresholding

$$W_{\hat{f}}(2^j, k) = \begin{cases} W_g(2^j, k) - t & \text{if } W_g(2^j, k) > t \\ 0 & \text{if } |W_g(2^j, k)| \leq t \\ W_g(2^j, k) + t & \text{if } W_g(2^j, k) < -t \end{cases}$$

- Apply the inverse wavelet transform and reconstruct the estimate $\hat{f}(k)$.

It has been shown that also hard thresholding yields better results in terms of the l_2 error, the estimated function does not have the desired smoothness properties.

Whereas if one employs soft thresholding the estimate is with high probability at least as smooth as the original function. Moreover, it came out that denoising with the traditional (orthogonal maximally decimated) wavelet transform typically exhibits undesired, spurious oscillations that can be attributed to the lack of translation invariance of the wavelet basis. One solution to suppress such artifacts was proposed by Coifman and Donoho [CD95b] with the translation invariant denoising.

Wavelet shrinkage can be seen as a thresholding with a constant artifacts number, this explains why the threshold depends on the length of the signal or equivalently on the size of the image. This dependence means that if we consider a noisy image containing one significant structure we would obtain two different denoising thresholds depending on the size of the surrounding region that is taken into account around the structure. This latter fact can be confusing since the level of noise in the significant structure does not depend on the area of the local analysis region surrounding it.

Starck, Murtagh and Bijaoui: wavelet domain adaptive filtering [SP96, SMB95]. This method is based on a test of statistical significance on the wavelet coefficients of the observed signal. This test enables to separate between the significant coefficients due to information and the non significant coefficients for which the lack of local information is not rejected. After stating the two test hypotheses

$$\begin{aligned} H_0 &= f(k) \text{ locally constant} \\ H_1 &= f(k) \text{ locally varies,} \end{aligned}$$

the evidence against H_0 is assessed by finding the probability of getting an outcome at least as far as the actually observed statistic from what we would get when H_0 is true. This probability is calculated as follows:

- If $W_g(2^j, k) > 0$

$$P = \text{Prob}\left(\mathcal{W}_g(2^j, k) > W_g(2^j, k)\right)$$

- If $W_g(2^j, k) < 0$

$$P = \text{Prob}\left(\mathcal{W}_g(2^j, k) < W_g(2^j, k)\right)$$

Given a fixed level of evidence ϵ we consider as decisive, If $P > \epsilon$, we accept H_0 and consider that the wavelet coefficient is non significant. Otherwise, we say that we detected a significant coefficient. At each scale the significant level ϵ leads to a threshold that we will note $t(2^j, k)$.

Due to the linearity of the wavelet transform, the wavelet coefficients of $g(k)$ follow a Gaussian distribution under the hypothesis H_0. Its standard deviation, which can be written as $\sigma_{W_g}(2^j) = \kappa(2^j)\sigma$, depends on the used wavelet transform [Bob97].

$$\text{Prob}(W_g(2^j,k)/H_0) = \frac{1}{\sigma_{W_g}(2^j)\sqrt{2\pi}} e^{-\frac{w(2^j,k)}{2\sigma^2_{W_g}(2^j)}}$$

The algorithm can be described as follows.

- Take the wavelet transform of the sequence $g(k)$. Starck et al. used the B-spline wavelet.
- Choose a value of ϵ. Estimate the standard deviation of the noise in $g(k)$. This can be done using the iterative procedure proposed by Starck and al. in [SP96]. Determine the thresholds $t(2^j,k)$.
- Modify the wavelet coefficients of $g(k)$ according to:

$$W_{\hat{f}}(2^j,k) = \begin{cases} W_g(2^j,k) & \text{if } |W_g(2^j,k)| \geq t(2^j,k) \\ 0 & \text{if } |W_g(2^j,k)| < t(2^j,k) \end{cases}$$

- Apply the inverse wavelet transform and reconstruct the estimate $\hat{f}(k)$.

Different variations of this algorithm that take into account the link between successive scales have also been proposed [SMB98, SP96]. The authors have shown that zeroing out the non significant coefficients leads to artifacts.

Bobichon and Bijaoui: regularized wavelet domain filtering [BB97]. In order to reduce the artifacts introduced by setting the non significant coefficients to zero, Bobichon and al. modified the preceding algorithm by introducing the regularization theory [PTK85, KE90]. They considered that the problem of inverting the wavelet transform is a typical ill-posed problem as defined by Hadamard [Had23], since there exists an infinity of values that can be given to the non significant wavelet coefficients. They restricted the set of solutions by taking advantage of the a priori knowledge on the expected result of the restoration. Among all possible values that can be assigned to the non significant wavelet coefficients they chose those that will give the smoothest image when applying the inverse wavelet transform. This was carried out by means of the Tikhonov constraint [Bob97]. The solution which verifies this constraint minimizes the energy of the gradient at each scale of the reconstructed image. The restoration algorithm can be described as follows:

- Take the wavelet transform of the noisy image $g(k)$. Bobichon et al. used the B-spline wavelet and the Haar wavelet.
- Analyze the statistical significance of the wavelet coefficients.
- Calculate iteratively the wavelet coefficients of $\hat{f}$.

  Initialize:

  $$F_{\hat{f}}^{(1)}(2^J) = F_{\hat{g}}(2^J)$$

  for j=J down to 1

  $$W_{\hat{f}}^{(1)}(2^j) = 0$$

  Iterate

  $$W_{\hat{f}}^{(n+1)}(2^j) = W_{\hat{f}}^{(n)}(2^j) - \xi_L \, GLF_{\hat{f}}^{(n)}(2^j)$$

  n is the iteration step, $L = [-1\ 2\ -1]$ is the Laplacian operator, G is the 1D Mallat wavelet analysis operator defined by $W(2^j) = GF(2^{j-1})$ (see Chapter 2) and ξ_L is a constant.
- Apply some additional non linear constraints to $W_{\hat{f}}$. These constraints enable to avoid contradictions between the significance of W_g and that of the new defined $W_{\hat{f}}$.
- Take the inverse wavelet transform of $W_{\hat{f}}$ to obtain $\hat{f}(k)$.

Wavelet methods for Poisson noise

In the section above, we have seen restoration methods for data which are adequately modeled as observations of a signal plus additive, Gaussian noise. Deviations from this model are often dealt with through judicious use of transformations. For example, in the case where the data have Poisson noise, it is common to preprocess the data first using some version of a square-root transformation. For example, Anscombe's [Ans48] transformation $g_A(k) = 2\sqrt{g(k) + \frac{3}{8}}$ yields data that are more nearly Gaussian than the original and possess a relatively constant noise level of 1. Following this preprocessing, one typically proceeds as if the data really did have Gaussian noise, and the same threshold is used, see [Don93a, SMB98, Bob97]. However the effectiveness with which the preprocessing creates Gaussian noise of a constant level decreases with decreasing intensity levels in the data. Therefore the success with which this approach is able to

remove the Poissonian noise while maintaining the integrity of the underlying structure varies with intensity as well. In particular, whatever degree of success the method may achieve at higher intensity levels, it will become less and less successful as the intensity decreases. Moreover this method has been criticized for often smoothing away more structure than is tolerable.

Therefore new approaches have been developed with the aim of providing an alternative that is more finely attuned to the characteristics of Poissonian signals. For all three methods presented in the following, the observed noisy data are supposed to be a realization of a random sequence $C(k)$ following a Poisson distribution of parameter $\lambda(k)$:

$$\text{Prob}(C(k) = c) = \frac{\lambda^c(k)}{c!} e^{-\lambda(k)}, \quad k = 1, \ldots, N \tag{3.24}$$

The aim of the restoration is the estimation of $\lambda(k)$.

Kolaczyk: the wavelet shrinkage paradigm adapted to Poisson data [Kol97, Kol99]. The method introduces "corrected" versions of the usual Gaussian based shrinkage thresholds proposed by Donoho and al. [DJ94]. In moving from the Gaussian noise problem to that of Poisson processes, the analogous null hypothesis of "no signal" is $H_0 : \lambda(t) = \lambda_0$. Under this condition the wavelet coefficients have zero mean and common variance as in the Gaussian settings, but they also have non negligible skewness and kurtosis. These facts motivated Kolaczyk to derive a new set of scale dependent thresholds, calibrated so that:

$$\text{Prob}\left(\max_{0 \le k \le 2^j - 1} W(2^j, k) \le t_{2^j}\right)$$

approaches 1 at a rate similar to that in (3.23), as j increases. Kolaczyk determined explicitly the thresholds for the Haar wavelet coefficients calculated with the translation invariant wavelet transform algorithm proposed by Donoho and Coifman [CD95b]. The translation invariant algorithm serves to avoid the staircase-like artifacts that usually accompany the use of the Haar wavelet. The resulting algorithm called TIPSH, for Translation Invariant Poisson Smoothing using Haar wavelets, can be described as follows.

- Compute the TI Haar wavelet transform of the signal $c(k)$.
- Apply level-dependant thesholds t_{2^j} (hard or soft) to the wavelet coefficients:

  $$t_{2^j} = 2^{\frac{J-j}{2}} \{\log 2^j + [\log^2(2^j) + 2\log(2^j)\lambda_j]^{\frac{1}{2}}\}$$

  where $J = \log_2(N)$, and $\lambda_j = \lambda_0 / 2^j$.

- Compute the inverse TI wavelet coefficients of the new defined wavelet coefficients to obtain $\hat{\lambda}(k)$.

Notice that the wavelet coefficients that are smaller than the threshold are set to zero, which leads to artifacts.

Nowak and Baraniuk: a wavelet-domain Wiener filter adapted to Poisson processes [NB97]. The filter proposed by Nowak and Baraniuk is a PRESS-optimal (it minimizes the PRedictive Sum of Squares), data adaptive filter designed using the method of cross-validation [Now96]. The method can be described as follows:

- Compute the wavelet transform of the signal $c(k)$. Nowak et al. used the Haar wavelet.
- Modify the wavelet coefficients according to:

  $$W_{\hat{\lambda}(k)}(2^j, k) = h^{PRESS}(2^j, k) W_c(2^j, k)$$

  The filter $h^{PRESS}(2^j, k)$ is defined as:

  $$h^{PRESS}(2^j, k) = \left(\frac{\hat{\sigma}^2_{W_c}}{\hat{\sigma}^2_{W_c} + \hat{\sigma}^2_{W_\lambda}} \right)_+ \tag{3.25}$$

  with

  $$\hat{\sigma}^2_{W_c} = \sum_k \psi^2(2^j, k) c(k) \tag{3.26}$$

  an unbiased estimate of the noise power in the wavelet coefficient $W(2^j, k)$, and

  $$\hat{\sigma}^2_{W_\lambda} = W_c^2(2^j, k) - \hat{\sigma}^2_{W_c}$$

  an unbiased estimate of the signal power in the wavelet coefficient $W_c(2^j, k)$. Thus each noisy wavelet coefficient is simply weighted by a factor equal to the estimated signal power divided by the estimated signal-plus-noise power. The $(.)_+$ operation thresholds the weight to zero if the estimated signal to signal-plus-noise ratio is negative.
- Invert the modified wavelet transform to obtain the signal estimate $\hat{\lambda}(k)$.

Here again the non significant wavelet coefficients are zeroed out, which leads to artifacts.

Timmermann and Nowak: a bayesian multiscale approach to Poisson intensity estimation based on the non normalized Haar wavelet transform [TN99]. In this method λ is regarded as an unknown realization of a random sequence Λ with prior density f_Λ. The optimal estimate of λ with respect to the squared error loss is:

$$\hat{\lambda} = E[\Lambda/C = c] = \int \lambda f(\lambda/c) d\lambda \tag{3.27}$$

where E[.] denotes the expectation operator and $f(\lambda/c)$ is the posterior density function $f_{\Lambda/C}(\lambda/c)$. Applying Bayes theorem to (3.27) they obtained:

$$\hat{\lambda} = \frac{\int \lambda \mathrm{P}(c/\lambda) f(\lambda) d\lambda}{\int \mathrm{P}(c/\lambda) f(\lambda) d\lambda} \tag{3.28}$$

where $\mathrm{P}(c/\lambda)$ is the likelihood that $C = c$ given that $\Lambda = \lambda$. Instead of specifying directly an a prior probability model for the unknown intensity, i.e f_Λ, Timmermann et al. modeled each Haar wavelet coefficient as independant perturbation of its corresponding scaling coefficient by introducing the so called perturbation variables $\Delta(2^j, k)$:

$$\mathcal{W}_\lambda(2^j, k) = \Delta(2^j, k)\, \mathcal{F}_\lambda(2^j, k)$$

for which they specified a prior density f_Λ of the form:

$$f(\lambda) = \sum_{i=1}^{M} p_i \frac{(1-\delta^2)^{s_i - 1}}{B(s_i, s_i) 2^{2s_i - 1}}$$

for $-1 \leq \delta \leq 1$, where B is the Euler beta function [Sap90], $0 \leq p_i \leq 1$ is the weight of the i-th beta density $\frac{(1-\delta^2)^{s_i-1}}{B(s_i,s_i)2^{2s_i-1}}$ with parameter $s_i \geq 1$, and $\sum_{i=1}^{M} p_i = 1$. This enabled to calculate the minimum mean square error optimal estimate of the innovation coefficient:

$$\hat{\delta}(2^j, k) = W_c(2^j, k) \frac{\sum_i p_i \frac{B(s_i + F_c(2^{j-1}, 2k),\, s_i + F_c(2^{j-1}, 2k+1))}{B(s_i,\, s_i)(2s_i + F_c(2^j, k))}}{\sum_i p_i \frac{B(s_i + F_c(2^{j-1}, 2k),\, s_i + F_c(2^{j-1}, 2k+1))}{B(s_i,\, s_i)}} \tag{3.29}$$

The authors have chosen a three componenents prior beta mixture model, i.e. $M = 3$. They suggest the shape parameter $s_1 = 1$, $s_2 = 100$ and $s_3 = 10000$ with weights $p_1 = 0.001$, $p_3 = 1 - p_2 - p_1$ and p_2 adapted at each scale. The overall algorithm is described below.

- Estimate coarsest scale coefficient

$$F_{\hat{\lambda}}(2^J, 0) = F_c(2^J, 0)$$

- For $j = J$ down to 1 and $k = 0$ to $N/2^j - 1$

  Compute:

  $$\hat{\delta}(2^j, k) \text{ according to (3.29)}$$
  $$W_{\hat{\lambda}}(2^j, k) = \hat{\delta}(2^j, k) F_{\hat{\lambda}}(2^j, k)$$

  Refine

  $$\begin{aligned} F_{\hat{\lambda}}(2^{j-1}, 2k) &= \tfrac{1}{2}\left(F_{\hat{\lambda}}(2^j, k) + W_{\hat{\lambda}}(2^j, k)\right) \\ F_{\hat{\lambda}}(2^{j-1}, 2k+1) &= \tfrac{1}{2}\left(F_{\hat{\lambda}}(2^j, k) - W_{\hat{\lambda}}(2^j, k)\right) \end{aligned}$$

3.4. Conclusion

The problem we are trying to solve in nuclear medicine imaging is the estimation of the intensity of a Poisson process from a single observation of the process.

The restoration methods for stationary processes are not adapted to our problem since the observed process in nuclear medicine exhibit a double nonstationarity. The intensity is nonstationary because it is mainly due to biological uptake phenomena and the noise is nonstationary since it depends on the signal intensity.

Moreover, practical limitations on imaging time and the amount of radioactivity that can be administered safely to patients leads to nuclear medicine images with a small number of counts. For such photon-limited imaging cases, we have seen that the adaptive wavelet based restoration methods developed for Gaussian noise can not be successfully adapted to Poissonian data through the use of the Anscombe's square root transform.

The wavelet based methods for Poisson processes we discussed in this chapter present the following drawbacks:

Kolaczyk [Kol97, Kol99]: sets the non significant wavelet coefficients to zero which leads to artifacts.

Nowak and Baraniuk [NB97]: the estimation of the standard deviation of the wavelet coefficients of the counts given by Equation (3.26) is not correct for low counts. This leads to a filter h^{PRESS} with a large variance, see Equation (3.25). Here again, zeroing out the non significant coefficients leads to artifacts.

Timmermann and Nowak [TN99]: the Bayesian perspective adopted in this paper leads to a detection process where the values assigned to the wavelet coefficients do not depend on the local characteristics of the image but on the global a prior density $f(\lambda)$ chosen for the intensity.

Moreover none of these methods gives information about the significance of the restored structures.

The purpose of this PhD thesis is to develop and analyze a new wavelet based regularized restoration method for noise removal in photon-limited imagery. This method, called DeQuant, is based on the work of Bobichon et al. [BB97] presented in Section 3.3.2 for restoration of astronomical images. This latter presents the two main following advantages: (1) it assigns a new value to the non-significant wavelet coefficients which reduces the artifacts and enables the incorporation of realistic prior information into the estimation process, and (2) it is based on a local detection process, with a measure of the significance of the detected structure. DeQuant introduces seven major contributions which are listed in a comparative way in Table 3.1.

	Bobichon [BB97, Bob97]	DeQuant
Noise type	Gaussian	Poisson
Transform	Mallat	Mallat and à trous
Wavelet	Haar and B-spline	unnormalized Haar
Selection	determination of the PDF of the wavelet coefficients for a Gaussian noise	determination of the PDF of the wavelet coefficients for a Poisson noise
Regularization	use of the Laplacian operator	use of the Laplacian operator and investigation of a new regularization constraint
Deconvolution	started after completing the denoising operation	incorporated in the denoising framework
Application field	Astronomy	Nuclear medicine

Table 3.1.: Original contributions introduced by DeQuant. Comparison with the work of Bobichon et al.

4. DeQuant: a flexible multiresolution restoration framework

This chapter develops and analyzes a new wavelet based regularized restoration method for noise removal in photon limited imagery. This method has been named DeQuant in analogy to **De**noising where the noise is of **Quant**um nature. It presents the following main advantages: (1) it assigns a new value to the non significant wavelet coefficients which reduces artifacts and enables the incorporation of realistic prior information into the estimation process, and (2) it is based on a local detection process with a measure of the significance of the detected structures. DeQuant follows basically the following scheme:

- The images we are working with in nuclear medicine exhibit a hierarchical structure of localized variations. This kind of images can be accurately and economically represented by a wavelet transform which will concentrate most of the information of the image among a small subset of the wavelet coefficients. Thus the first step of DeQuant will be a **transformation** step which leads to a representation of the noisy image in the wavelet domain.

- DeQuant then analyzes the statistical significance of the obtained wavelet coefficients in order to separate between the significant coefficients containing information and the non significant coefficients for which the hypothesis of a lack of local information was not rejected. A significance measure has been introduced enabling the physician to be aware of the significance of the structures in the restored image. This is the **selection** step.

- Restoring the image using only the significant coefficients and zeroing out the non significant ones leads to artifacts. Therefore a rule has to be found to assign a new value to the non significant coefficients. The choice is constrained by the third step of DeQuant: the **regularization** step.

- The value assigned to the non significant coefficients after the previous step is submitted to some non linear constraints. These constraints take into account the admissible range of variation of the wavelet coefficients and the positivity of the reconstructed image. This is the **projection** step.
- The image is reconstructed using the significant coefficients and the projected non significant ones. This is the **reconstruction** step.
- The solution obtained is not directly optimal and the whole procedure has to be iterated. This is the **iteration** step.

The image degradation model we are working with has been introduced in Chapter 1. We recall here the main notations that will be used in the following. Each point of the scintigram (i.e. image produced by the gamma camera) is modeled by the random variable $I(k_x, k_y)$ which represents the number of photon counts detected at position (k_x, k_y) of the image during the imaging time T. $I(k_x, k_y)$ obeys a Poisson distribution of parameter $Y(k_x, k_y)$:

$$\text{Prob}(I(k_x, k_y) = n) = \frac{Y^n(k_x, k_y)}{n!} e^{-Y(k_x, k_y)} \tag{4.1}$$

where

$$Y = PX = \int_d \rho(x, y) dx dy \tag{4.2}$$

X is the object; it corresponds to the image of the local uptake density that would be given by an ideal imaging system. Y will be referred to as the noise free image or the discretized intensity function since it is obtained by integrating the continuous intensity function $\rho(x, y)$ on the pixel domain d; it corresponds to the distribution of photons after it has been spread by the point spread function P. $I(k_x, k_y)$ is the noisy image.

Our aim is, given one realization of the Poisson process (which is the noisy image $I(k_x, k_y)$) to estimate on the one hand the discretized intensity function $Y(k_x, k_y)$ and to remove on the other hand the effect of the point spread function by estimating $X(k_x, k_y)$.

This chapter is organized as follows. Section 4.1 explains how DeQuant can be used for denoising. The goal is to estimate the intensity $Y(k_x, k_y)$ of a Poisson process from a single observation $I(k_x, k_y)$ of the process. Section 4.2 shows how the introduction of the point spread function as additional constraint in the restoration procedure enables to improve the denoising procedure, the objective being here to estimate the object X given $I(k_x, k_y)$ and P. Section 4.3 investigates the benefits of integrating the Coifman and Donoho's translation invariant

wavelet transform presented in Chapter 2 in the general multiresolution framework. Section 4.4 presents the edge preserving version of DeQuant. We conclude in section 4.5.

4.1. Denoising with DeQuant

The objective of this section is to show how DeQuant estimates the intensity $Y(k_x, k_y)$ of a Poisson process from a single observation $I(k_x, k_y)$ of the process. This section is organized as follows. After justifying the choice of the wavelet transform, we discuss the significant wavelet coefficient extraction based on the multiscale behavior of the noise. We then describe the regularized reconstruction. The algorithm can be found in Appendix A Section A.1.

4.1.1. Transformation

The wavelet transform has been chosen for its ability to provide an accurate and economical representation of data that exhibit hierarchically localized variations, such as the nuclear medicine images we are dealing with. Indeed, wavelets have certain properties that make them very effective in analyzing the class of locally varying signals, see Chapter 2. We list two fundamental properties here with a brief discussion of their meaning.

Multiscale: The wavelet transform represents the data as a nested set of scales. The idea is simple: separate the information to be analyzed into a principal part and a residual part. In applications to signal processing, the principal part should be thought of as primarily low-pass (or approximation at a given scale) and the residual part as primarily high pass (or details at a given scale). The process of decomposition can be applied again to one or both of the parts. If it is repeatedly applied to the low-pass part, this process is the one introduced by Mallat [Mal89b].

Locality: Each wavelet coefficient represents the image contents localized in spatial location and frequency. In the wavelet notation k indexes the spatial location of analysis. For a wavelet $\psi(t)$ centered at time zero and frequency f_0, the wavelet coefficient $W(2^j, k)$ measures the signal content around time $2^j k$ and around frequency $2^{-j} f_0$. Thus wavelets exhibit simultaneous spatial and frequency localization.

Together the locality and multiscale property enable the wavelet transform to efficiently match images organized into levels or scales of localized variations. Thus the first step of DeQuant will be to transform the noisy image $I(k_x, k_y)$ in the wavelet domain. We obtain a sparse representation with a few large dominating coefficients.

4.1.2. Selection

A large part of the wavelet coefficients does not carry out a significant information. We select the significant ones by the procedure explained in this section.

The methodology we use for estimating the underlying intensity function is based on a statistical inference test where the evidence provided by the data in favor of a statement must be assessed: we hope to show that the intensity function locally varies, to do this we suppose that the intensity function is locally constant and then look for evidence against the supposition we made. This reasoning, referred to as significance test [Moo91], leads in terms of the discrete variable Y to the following hypotheses:

$$\begin{array}{ll} H_0\text{:} & Y \text{ locally constant} \\ H_a\text{:} & Y \text{ locally varies} \end{array}$$

Choice of the Test Parameter

In order to measure how well the data conform to H_0, a test parameter has to be chosen. This test parameter must be able to characterize variation and locality. The wavelet coefficients fulfill these conditions [Mal97]: indeed, due to the fact that wavelets are zero mean functions, they will associate to every locally constant signal a zero-valued wavelet coefficient. Moreover the wavelet analysis is a multiresolution method that enables to perform a local analysis at different scales. Stating the hypothesis in terms of the wavelet coefficients of Y, we obtain:

$$\begin{array}{ll} H_0\text{:} & \mathcal{W}_{a,Y}(2^j, k_x, k_y) = 0 \\ H_a\text{:} & \mathcal{W}_{a,Y}(2^j, k_x, k_y) \neq 0 \end{array}$$

The parameter a stands for the letters h, v and d. Since we do not know Y, we approximate its wavelet coefficients by the wavelet coefficients of the image I. These latter are unbiased estimates of the true coefficients of the intensity function:

$$\mathrm{E}\left[\mathcal{W}_{a,I}(2^j, k_x, k_y)\right] = \mathcal{W}_{a,Y}(2^j, k_x, k_y)$$

Additionally the variance of each coefficient, which is a useful indicator of the overall random error, is influenced as might be expected by the value of Y in the region of support of the corresponding wavelet function

$$\text{Var}\left[\mathcal{W}_{a,I}(2^j, k_x, k_y)\right] = \sum \left[\psi^a_{(2^j,k_x,k_y)}(x,y)\right]^2 Y(k_x, k_y)$$

The functions $\psi^a_{(2^j,k_x,k_y)}(x,y)$ are defined by Equation (2.19). We can now state the hypothesis in terms of the wavelet coefficients of the noisy image I:

$$\begin{array}{ll} H_0: & \mathcal{W}_{a,I}(2^j, k_x, k_y) = 0 \\ H_a: & \mathcal{W}_{a,I}(2^j, k_x, k_y) \neq 0 \end{array}$$

In order to assess the evidence against H_0, we have to find the probability of getting an outcome at least as far as the actually observed statistic from what we would expect when H_0 is true. This probability is called the P-value and calculated as follows:

- If $W_I(2^j, k_x, k_y) > 0$

$$\text{P-value} = \text{Prob}(\mathcal{W}_{a,I}(2^j, k_x, k_y) > W_{a,I}(2^j, k_x, k_y)) \tag{4.3}$$

- If $W_I(2^j, k_x, k_y) < 0$

$$\text{P-value} = \text{Prob}(\mathcal{W}_{a,I}(2^j, k_x, k_y) < W_{a,I}(2^j, k_x, k_y)) \tag{4.4}$$

where $W_{a,I}(2^j, k_x, k_y)$ are the observed wavelet coefficients of the noisy image and $W_{a,I}(2^j, k_x, k_y)$ denote the random variables. The evidence we obtain is then compared with the significance level ϵ which is a fixed level of evidence that we regard as decisive. If the P-value associated with $W_{a,I}(2^j, k_x, k_y)$ is as small or smaller than ϵ, we cannot consider at the level of decision ϵ that the value of the coefficient is only due to noise. We reject H_0 in favor of H_a which means that the image is not considered being constant over a neighborhood of size proportional to 2^j centered in (k_x, k_y). We say that we detected a statistically significant wavelet coefficient at level ϵ. If the P-value is bigger than ϵ, although non zero, we consider that the wavelet coefficient is caused by a chance fluctuation of the underlying random process and accept H_0. In this case we say that we detected a non significant wavelet coefficient.

In order to calculate the P-value, we need to compute the probability density function (PDF) of the wavelet coefficient associated to the noisy image I under the hypothesis H_0, that is, for a locally uniform Poisson distribution. This law depends on the analyzing wavelet. We have chosen to work with the unnormalized Haar wavelet. The reasons for this choice and the resulting PDF are given in the following.

Choice of the wavelet

The choice of the unnormalized Haar wavelet was motivated by two reasons: (1) there is a natural match between the Haar wavelet and the image formation process, and (2) the unnormalized Haar wavelet enables an analytical computation of the PDF of the wavelet coefficients of a Poisson process.

A natural match Decomposition coefficients in the orthogonal wavelet transform case are computed with the fast Mallat algorithm that cascades discrete convolutions with a lowpass filter h and a highpass filter g and subsamples the output, see Chapter 2. This algorithm assumes that the initial discrete data, which is the sequence $I(k_x, k_y)$ already represents an approximation of $I(x, y)$ at a certain scale, which is by convention fixed at $j = 0$, see Equation (2.36):

$$I(k_x, k_y) =< I(x, y), \phi(x - k_x)\phi(y - k_y) > .$$

In our case, the discrete input image $I(k_x, k_y)$ is obtained by a finite resolution device, the gamma camera, that averages and samples the analog input image. We have seen in Chapter 1 that in order to obtain an image with the gamma-camera, it is necessary to project γ-rays from the source distribution onto the camera detector. Gamma rays cannot be focused; thus a "lens" principle similar to that used in photography cannot be applied [SP87]. Therefore most practical γ-rays imaging systems employ the principle of absorptive collimation for image formation. An absorptive collimator projects an image of the source distribution onto the detector by allowing only those γ-rays traveling along certain directions to reach the detector. Gamma rays not traveling in the proper direction are absorbed by the collimator before they reach the detector, see Chapter 1.

Hence, the averaging kernel of our physical device is the function which takes the value one on the area corresponding to the hole of the collimator and zero otherwise. This scaling function corresponds to the Haar wavelet.

An easy computation of the PDF The filters usually associated with this wavelet are $[1/\sqrt{2},\ 1/\sqrt{2}]$ for the lowpass filter and $[1/\sqrt{2},\ -1/\sqrt{2}]$ for the highpass filter, but we have chosen to work with the unnormalized version of the Haar wavelet which corresponds to the filters:

$$h = [1,\ 1] \quad g = [1,\ -1]$$

Looking at the resulting analysis equations given by (2.46), we notice that every scaling coefficient is the sum of four finer scale scaling coefficients, and conse-

quently, due to the reproducing property of the Poisson distribution [SHW93], every scaling coefficient is Poisson distributed.

$$C_i \sim \text{Poisson}(\lambda_i),\ C_i \text{ independent} \Rightarrow \sum C_i \sim \ \text{Poisson}(\sum \lambda_i)$$

Furthermore this choice of the wavelet enables an analytical computation of the PDF of the wavelet coefficients of a Poisson process as we will seen in short in the following section and in details in Chapter 5. This fact is crucial in the development of the proposed intensity estimate.

Similar attributes (reproducibility and simple PDF) do not hold for more general multiscale analyzes of Poisson processes, based on other wavelet systems for example. Hence the unnormalized Haar transform can be viewed as the "canonical" multiscale analysis tool for Poisson processes.

Probability distribution of the wavelet coefficients

Once the wavelet has been chosen, we can calculate the PDF of the wavelet coefficients of the noisy image under the hypothesis H_0.

H_0 states that the intensity function is locally constant. The locality is measured by the size of the wavelet at the scale of the analysis. Thus for the Haar wavelet at scale 2^j, H_0 can be written as $\rho(x, y) = \rho_0$ on an interval of size 2^{2j}. We note:

$$\alpha = 2^{2j}\rho_0.$$

Equation (2.46) shows that the wavelet coefficients $\mathcal{W}_{a,I}$ where $a = h,\ v,$ or d are calculated as the difference of two sums of Poisson variables:

$$\mathcal{W}_{a,I}(2^j, k_x, k_y) = S_1 - S_2$$

where S_1 and S_2 are Poisson variables of parameter $\frac{\alpha}{2}$ since they are taken on a domain whose size equals half that of the wavelet. The PDF of the wavelet coefficients can be computed as follows (see [BJ00] or Chapter 5 for more details):

$$\begin{aligned} p_{\mathcal{W}}(\alpha, n) &\triangleq \text{Prob}(\mathcal{W}_{a,I}(2^j, k_x, k_y) = n) && (4.5)\\ &= \sum_{m=0}^{+\infty} \text{Prob}(S1 = n + m)\,\text{Prob}(S2 = m) && (4.6)\\ &= e^{-\alpha} \sum_{m=0}^{+\infty} \frac{\left(\frac{\alpha}{2}\right)^{n+2m}}{(n+m)!m!} && (4.7)\\ &= e^{-\alpha} I_{|n|}(\alpha) && (4.8) \end{aligned}$$

We obtain a symmetrical PDF, where $I_{|n|}(\alpha)$ is the modified Bessel function of integer order $n \in \mathbb{Z}$ [GM52]. In Equation (4.8), we approximate α by its discrete counterpart $F_Y(2^j, k_x, k_y)$ which corresponds to the scaling coefficient of Y at scale 2^j. We obtain:

$$\text{Prob}\big(\mathcal{W}_{a,I}(2^j, k_x, k_y) = W_{a,I}(2^j, k_x, k_y)\big) = e^{-F_Y(2^j,k_x,k_y)} I_{|W_{a,I}(2^j,k_x,k_y)|}(F_Y(2^j, k_x, k_y)) \tag{4.9}$$

Note that the PDF given by Equation (4.9) depends on the variable Y we want to estimate. This leads to an iterative denoising algorithm.

Separation between signal and noise

At each scale, the significant level ϵ defines a threshold that we denote $t(2^j, k_x, k_y)$. This threshold is the smallest positive integer such that the integral of the PDF from this integer to infinity is smaller than ϵ:

$$t(2^j, k_x, k_y) = \min_x\{\int_x^\infty p_{\mathcal{W}}(\alpha, u)du\}$$

Based on this threshold, we define the mask $M^s_{a,I}(2^j, k_x, k_y)$ of significant wavelet coefficients i.e. those whose value is due to a real variation of the underlying intensity:

$$M^s_{a,I}(2^j, k_x, k_y) = \begin{cases} 1 & \text{if } |W_{a,I}(2^j, k_x, k_y)| > t(2^j, k_x, k_y), \\ 0 & \text{otherwise.} \end{cases}$$

with $a = h$, v and d. The non-significant wavelet coefficients, i.e. those whose value could be due to a pure statistical fluctuation, will then correspond to the complementary mask $\overline{M}^s_{a,I}(2^j, k_x, k_y)$:

$$\overline{M}^s_{a,I}(2^j, k_x, k_y) = \begin{cases} 0 & \text{if } |W_{a,I}(2^j, k_x, k_y)| > t(2^j, k_x, k_y), \\ 1 & \text{otherwise.} \end{cases}$$

These masks enable to separate the contribution of the signal from that of the noise and obtain a description of the image in terms of significant, i.e. $W^s_{a,I}(2^j, k_x, k_y)$, and non-significant, i.e. $W^{ns}_{a,I}(2^j, k_x, k_y)$, wavelet coefficients:

$$\begin{aligned} W^s_{a,I}(2^j, k_x, k_y) &= W_{a,I}(2^j, k_x, k_y) * M^s_{a,I}(2^j, k_x, k_y) \\ W^{ns}_{a,I}(2^j, k_x, k_y) &= W_{a,I}(2^j, k_x, k_y) * \overline{M}^s_{a,I}(2^j, k_x, k_y) \end{aligned}$$

We define W_I^s and W_I^{ns} as the sets of significant respectively non-significant wavelet coefficients

$$\begin{aligned} W_I^s &= \{W_{a,I}^s(2^j, k_x, k_y), \quad a = h, v \text{ and } d, \quad j = 1, \dots, J\} \\ W_I^{ns} &= \{W_{a,I}^{ns}(2^j, k_x, k_y), \quad a = h, v \text{ and } d, \quad j = 1, \dots, J\} \end{aligned}$$

J is the coarsest scale of the analysis.

4.1.3. Regularization

Our objective is to obtain an estimate of the noise free image Y using the raw data image $I(k_x, k_y)$.

The solution that consists in reconstructing the estimate from W_I^s taking $W_I^{ns} = 0$ is not acceptable due to the block effects that it causes: when a wavelet coefficient is lost by thresholding, information on the local gradient of the image is lost and the signal is reconstructed using the local mean. That is why some rule has to be found to assign a new value to the non significant coefficients.

The problem of inverting the wavelet transform is a typical ill-posed problem as defined by Hadamard [Had23], since there exists an infinity of values that can be given to the non significant wavelet coefficients. The unicity of the solution can be forced by a regularization constraint based on a priori knowledge on Y.

In this regularized approach that we adopt for denoising, the restored image takes into account the information left in the significant coefficients (fidelity to the data) but its non-significant coefficients are defined by the a priori on the solution (regularization constraint).

Smoothness constraint on the image

For nuclear medicine applications, we want the restored free of noise image to be smooth. Several regularization constraints have been proposed for this purpose. The most popular are certainly the Maximum Entropy Method [Jay57] and the Tikhonov's regularization [Tik63].

The Maximum Entropy Method (MEM) for image restoration was initially proposed by Jaynes [Jay57]. This algorithm is based on the idea that a given realization of a random variable carries a certain amount of information quantifiable by the entropy [Jay57] [Sha48] [SM99]. When trying to invert an ill-posed problem the entropy is used as a regularizing functional to constraint the solution and

give the simplest (in the sense of the amount of information contained) possible solution which is compatible with the data. The MEM leads to a component-wise modification of the wavelet coefficients W_I by weighting them with a coefficient h_w: $\widetilde{W_I} = h_w W_I$.

Since the MEM is a point operation, it is not optimally adapted for denoising purposes where an operation relating neighborhood pixels such as Tikhonov's regularization performs better. Note that in the case of restoration, this role is taken by the PSF.

The Tikhonov constraint

Tikhonov's regularization expresses smoothness of the restored image. This latter was obtained this by constraining the gradient of the scaling images associated with Y to be minimum at each scale 2^j:

$$\|D_x F_Y(2^j)\|^2 + \|D_y F_Y(2^j)\|^2 \quad \text{minimum} \tag{4.10}$$

D_x and D_y are the horizontal and vertical gradient operators. This condition is called the multiresolution minimum gradient constraint [BB97]. By deriving Equation (4.10) with respect to $F_Y(2^j)$ and setting the result to zero we obtain:

$$L_2 F_Y(2^j) = 0 \tag{4.11}$$

where L_2 is the Laplacian operator. It corresponds to the convolution with the filter[1]:

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{4.12}$$

Using the Van-Cittert algorithm [VC31], see Chapter 3 Section 3.2.1, to resolve equation (4.11) we obtain the iterative form:

$$F_Y^{(n+1)}(2^j) = F_Y^{(n)}(2^j) - \xi_{L_2}\, L_2 F_Y^{(n)}(2^j) \tag{4.13}$$

For an optimal convergence we take $\xi_{L_2} = 0.2$ [Bob97]. Remember that we are looking for the wavelet coefficients of $F_Y^{(n)}(2^j)$ that enables Equation (4.10) to be verified. We therefore apply on both parts of Equation (4.13) the analysis

[1] Although the analytical Laplacian $\bigtriangledown^2 f(x,y)$ takes positive values for hollows and negative values for humps, we have chosen to work with a discrete approximation usually used in image processing applications which compared to the analytical Laplacian gives opposite signs for hollows and humps.

operators of the Mallat transform. These operators denoted G_h, G_v and G_d are defined Chapter 2, Section 2.4.2. We get:

$$\begin{aligned} W_{h,F_Y^{(n+1)}(2^j)} &= W_{h,F_Y^{(n)}(2^j)} - \xi_{L_2} W_{h,L_2F_Y^{(n)}(2^j)} \\ W_{v,F_Y^{(n+1)}(2^j)} &= W_{v,F_Y^{(n)}(2^j)} - \xi_{L_2} W_{v,L_2F_Y^{(n)}(2^j)} \\ W_{d,F_Y^{(n+1)}(2^j)} &= W_{d,F_Y^{(n)}(2^j)} - \xi_{L_2} W_{d,L_2F_Y^{(n)}(2^j)} \end{aligned} \tag{4.14}$$

Thus if the non significant wavelet coefficients of Y are calculated according to the set of Equations (4.14), then the solution Y satisfies the multiresolution minimum gradient constraint defined by Equation (4.10).

4.1.4. Projection

The non-linear constraints enable to avoid contradictions between the significance of the wavelet coefficients of the noisy image and the significance of the wavelet coefficients of the reconstructed image.

To make our mean clear, let us consider the following example. If for a given scale 2^{j_0} and at a given position (k_{x_0}, k_{y_0}) the wavelet coefficient of the data $W_{a,I}(2^{j_0}, k_{x_0}, k_{y_0})$ has been found to be non significant, then the wavelet coefficient of the reconstructed image $W_{a,F_Y^{(n)}(2^{j_0-1})}(k_{x_0}, k_{y_0})$ at the same scale 2^{j_0} and the same position (k_{x_0}, k_{y_0}) must be also non significant. Note that $W_{a,F_Y^{(n)}(2^{j_0-1})}$ corresponds to the wavelet coefficient of Y at scale 2^{j_0}, see Equations (2.42). This means that, if the calculated value of $W_{a,F_Y^{(n)}(2^{j_0-1})}(k_{x_0}, k_{y_0})$ is outside the interval $[-t(2^{j_0}, k_{x_0}, k_{y_0})\ ;\ t(2^{j_0}, k_{x_0}, k_{y_0})]$, it must be artificially constrained to be in the interval.

In order to apply the previously cited rule, the significance of the calculated wavelet coefficients $W_{a,F_Y^{(n)}(2^{j-1})}(k_x, k_y)$ at each scale 2^j and each position (k_x, k_y) is compared at each iteration n with the significance of the wavelet coefficients of the noisy image, $W_{a,I}(2^j, k_x, k_y)$. The result of this comparison determines the final value to assign to the calculated wavelet coefficients.

Let NS denote the set of points where the wavelet coefficients of the data have been found to be non significant:

$$NS = \{\mathbf{k}_{ns} = (k_x, k_y) \text{ where } \overline{M}^s_{a,I}(2^j, k_x, k_y) = 1\}$$

And S denote the set of points where the wavelet coefficients of the data have been found to be significant:

$$S = \{\mathbf{k}_s = (k_x, k_y) \text{ where } M^s_{a,I}(2^j, k_x, k_y) = 1\}$$

In the following we omit the subscript a in order to make the notation easier.

For all locations $\mathbf{k}_{ns} \in NS$ where the wavelet coefficients of the data $W_I(2^j, \mathbf{k}_{ns})$ have been found to be non significant, $W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns})$ may only vary in the range $[-t(2^j, \mathbf{k}_{ns})\ ;\ t(2^j, \mathbf{k}_{ns})]$:

$$W^{ns}_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns}) = \begin{cases} -t(2^j, \mathbf{k}_{ns}) & \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns}) < -t(2^j, \mathbf{k}_{ns}) \\ W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns}) & \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns}) \in [-t(2^j, \mathbf{k}_{ns}); \\ & \qquad\qquad t(2^j, \mathbf{k}_{ns})] \\ t(2^j, \mathbf{k}_{ns}) & \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_{ns}) > t(2^j, \mathbf{k}_{ns}) \end{cases}$$

For all locations $\mathbf{k}_s \in S$ where the wavelet coefficients of the data $W_I(2^j, \mathbf{k}_s)$ have been found to be significant, the action of the regularization constraint will be restricted to the interval $[-\delta\, t(2^j, \mathbf{k}_s)\ ;\ \delta\, t(2^j, \mathbf{k}_s)]$ centered around the initial value $W_I(2^j, \mathbf{k}_s)$:

$$W^{s}_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_s) = \begin{cases} W_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s) \\ \quad \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_s) < W_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s) \\ W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_s) \\ \quad \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_s) \in [W_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s); \\ \qquad\qquad W_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s)] \\ W_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s) \\ \quad \forall\, W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k}_s) > W_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s) \end{cases}$$

The parameter δ is a tolerance parameter that has been introduced to take into account the noise that affects the significant wavelet coefficients. The bigger δ is, the smoother will be the reconstructed image. The choice of the value of δ is explained in Chapter 6.

4.1.5. Reconstruction

Once the wavelet coefficients of $F_Y^{(n)}$ at scale 2^j have been calculated, the approximation image of Y at a smaller scale can be computed using the synthesis equation of the Mallat wavelet transform (2.47):

$$\begin{aligned} F_Y^{(n)}(2^{j-1}) \;=\; & \big[\, \widetilde{H}_2 F_Y^{(n)}(2^j) + \widetilde{G}_h W_{h, F_Y^{(n)}(2^{j-1})} \\ & + \widetilde{G}_v W_{v, F_Y^{(n)}(2^{j-1})} + \widetilde{G}_d W_{d, F_Y^{(n)}(2^{j-1})} \,\big]\, /4 \end{aligned} \tag{4.15}$$

Poisson processes being inherently non negative, the restored scaling image is then projected on the space of positive functions:

$$F_Y^{(n)}(2^{j-1}) = \begin{cases} F_Y^{(n)}(2^{j-1}) & \forall\ F_Y^{(n)}(2^{j-1}) \geq 0 \\ 0 & \forall\ F_Y^{(n)}(2^{j-1}) < 0 \end{cases}$$

Once all the scales have been gone through, we obtain the reconstructed image $Y = F_Y^{(n)}(2^0)$.

4.1.6. Iteration

Since the probability density function of the wavelet coefficients of the noisy image I, depends on the noise free image Y we want to estimate, see Equation (4.9), the algorithm has to be iterative.

We suppose for the first iteration that the image Y is constant and equals the mean value of the noisy image I:

$$\forall (k_x, k_y), \qquad Y(k_x, k_y) = \left(\sum I\right) / N_I$$

where N_I is the number of pixels of the image I. A constant value has been chosen because we did not want to introduce in the initialization prior structures which would be difficult to get rid of through iterations.

The algorithm corresponding to the denoising method is given in Appendix A Section A.1.

4.2. Restoration with DeQuant

In the previous section, we presented a denoising method (that will be referred to as the D-method) which enables to restore a free of noise estimation of the observed image.

When the PSF of the imaging system is known, it represents an important source of information that can be integrated as supplementary constraint in the process of restoration of the non significant wavelet coefficients. Moreover, the regularization constraint in the D-method was based on a multiresolution criteria of smoothness of the *image* Y whereas it is the smoothness of the *object* X that has to be aimed at. Finally, the D-method restored the *image* whereas we would like to restore the *object*.

The purpose of this section is to show how we improved the filtering algorithm by introducing the PSF of the imaging system as additional constraint in the restoration of the wavelet coefficients and choosing the regularization constraint in the right space, namely the object space. The new restoration method where denoising and deconvolution are coupled in a multiresolution frame will be referred to as the R-method.

As for the D-method, the R-method starts by studying the statistical significance of the wavelet coefficients of the noisy image. It then separates the contribution of signal from that of the noise, see Section 4.1.2. Differences between the two methods appear at the regularization, projection and reconstruction steps. These differences are presented in this section. The resulting algorithm is given in Appendix A Section A.6.

4.2.1. Regularization

Our objective now is to obtain an estimation of the object X. This problem will be resolved iteratively by estimating the parameter of the Poisson distribution using the information given by the significant wavelet coefficients of I. We are again faced with an ill posed problem [Dem89] that needs to be regularized by the introduction of an additional constraint that concerns now the smoothness of the object X. An iterative deconvolution algorithm will be used to calculate X from Y.

We want to find the wavelet coefficients associated with Y so that the reconstruction of Y with these coefficients minimizes the gradient of the object X. This results in:

$$\|D_x X\|^2 + \|D_y X\|^2 \quad \text{minimum},$$

where D_x and D_y are the horizontal and vertical gradient operators. This latter equation is equivalent to

$$L_2 X = 0 \tag{4.16}$$

where L_2 corresponds to the Laplacian obtained by convolution with the filter given by Equation (4.12). Using the Van-Cittert algorithm to resolve Equation (4.16) we obtain the iterative form:

$$X^{(n+1)} = X^{(n)} - \xi_{L_2}\, L_2 X^{(n)} \tag{4.17}$$

We first apply the PSF operator P on both parts of Equation (4.17):

$$Y^{(n+1)} = Y^{(n)} - \xi_{L_2}\, P L_2 X^{(n)} \tag{4.18}$$

We then apply j times the analysis operators G_h, G_v and G_d defined Chapter 2, Section 2.4.2. The result is

$$\begin{aligned} W_{h,Z^{(n+1)}}(2^j) &= W_{h,Y^{(n)}}(2^j) - \xi_{L_2} W_{h,PL_2X^{(n)}}(2^j) \\ W_{v,Z^{(n+1)}}(2^j) &= W_{v,Y^{(n)}}(2^j) - \xi_{L_2} W_{v,PL_2X^{(n)}}(2^j) \\ W_{d,Z^{(n+1)}}(2^j) &= W_{d,Y^{(n)}}(2^j) - \xi_{L_2} W_{d,PL_2X^{(n)}}(2^j) \end{aligned} \tag{4.19}$$

Note that the notation $Z^{(n+1)}$ has been introduced in the left part of equation (4.19) to replace $Y^{(n+1)}$. This has been done to emphasize the difference between the result of the convolution of P with $X^{(n+1)}$, when $X^{(n+1)}$ is known, that we denote $Y^{(n+1)}$, and the image which is calculated using the wavelet coefficients of $PX^{(n)}$ and $PL_2X^{(n)}$. This latter will be referred to as $Z^{(n+1)}$.

4.2.2. Projection

The non-linear constraints oncern now the wavelet coefficients of $Z^{(n)}$. At each iteration, the significance of the calculated wavelet coefficients, $W_{Z^{(n)}}(2^j, k_x, k_y)$, is compared with the significance of the wavelet coefficients of the noisy image, $W_I(2^j, k_x, k_y)$. The result of this comparison determines the final value to assign to the calculated wavelet coefficients.

For all locations $\mathbf{k}_{ns} \in NS$ where the wavelet coefficients of the data $W_I(2^j, \mathbf{k}_{ns})$ have been found to be non significant, $W_{Z^{(n)}}(2^j, k_x, k_y)$ may only vary in the range $[-t(2^j, \mathbf{k}_{ns})\ ;\ t(2^j, \mathbf{k}_{ns})]$:

$$W^{ns}_{Z^{(n)}}(2^j, k_x, k_y) = \begin{cases} -t(2^j, \mathbf{k}_{ns}) & \forall\ W^{ns}_{Z^{(n)}}(2^j, \mathbf{k}_{ns}) < -t(2^j, \mathbf{k}_{ns}) \\ W^{ns}_{Z^{(n)}}(2^j, \mathbf{k}_{ns}) & \forall\ W^{ns}_{Z^{(n)}}(2^j \mathbf{k}_{ns}) \in [-t(2^j, \mathbf{k}_{ns})\ ;\ t(2^j, \mathbf{k}_{ns})] \\ t(2^j, \mathbf{k}_{ns}) & \forall\ W^{ns}_{Z^{(n)}}(2^j, \mathbf{k}_{ns}) > t(2^j, \mathbf{k}_{ns}) \end{cases}$$

For all locations $\mathbf{k}_s \in S$ where the wavelet coefficients of the data $W_I(2^j, \mathbf{k}_s)$ have been found to be significant, the action of the regularization constraint will

be restricted to the interval $[-\delta\, t(2^j, \mathbf{k}_s)\ ;\ \delta\, t(2^j, \mathbf{k}_s)]$ centered around the initial value $W_I(2^j, \mathbf{k}_s)$:

$$W^s_{Z^{(n)}}(2^j, k_x, k_y) = \begin{cases} W^s_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s) \\ \quad \forall\ W^s_{Z^{(n)}}(2^j, \mathbf{k}_s) < W^s_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s) \\ W^s_{Z^{(n)}}(2^j, \mathbf{k}_s) \\ \quad \forall\ W^s_{Z^{(n)}}(2^j, \mathbf{k}_s) \in [W^s_I(2^j, \mathbf{k}_s) - \delta\, t(2^j, \mathbf{k}_s); \\ \qquad\qquad W^s_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s)] \\ W^s_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s) \\ \quad \forall\ W^s_{Z^{(n)}}(2^j, \mathbf{k}_s) > W^s_I(2^j, \mathbf{k}_s) + \delta\, t(2^j, \mathbf{k}_s) \end{cases}$$

4.2.3. Reconstruction

Once the wavelet coefficients of $Z^{(n)}$ at scale 2^j have been calculated, the approximation of $Z^{(n)}$ at a smaller scale can be computed using the synthesis equation of the Mallat wavelet transform (2.47):

$$\begin{aligned} F_{Z^{(n)}}(2^{j-1}) \quad = \quad & \big[\ \widetilde{H_2} F_{Z^{(n)}}(2^j) + \widetilde{G_h} W_{h,Z^{(n)}}(2^j) \\ & + \widetilde{G_v} W_{v,Z^{(n)}}(2^j) + \widetilde{G_d} W_{d,Z^{(n)}}(2^j)\ \big]\ /4 \end{aligned} \tag{4.20}$$

We then project the restored approximation on the space of positive functions:

$$F_{Z^{(n)}}(2^j) = \begin{cases} F_{Z^{(n)}}(2^j) & \forall\ F_{Z^{(n)}}(2^j) \geq 0 \\ 0 & \forall\ F_{Z^{(n)}}(2^j) < 0 \end{cases}$$

Once all the scales have been gone through the reconstructed image $Z^{(n)}$ is obtained.

Recall that the final objective was to obtain an estimate of the object X. Having P, the point spread function, and $Z^{(n)}$, the estimation of the Poisson parameter, we are faced with a deconvolution problem. The Richardson-Lucy method [Luc74] is used to calculate the next estimate of the desired object. This method is an iterative deconvolution procedure derived from Bayes' theorem on conditional probabilities:

$$X^{(n)} = X^{(n-1)} \left[P^t \frac{Z^{(n)}}{P X^{(n-1)}} \right]$$

P^t represents the transpose of the point spread function. The calculated estimate is always positive.

4.3. Shift-Invariant DeQuant

The D-method and the R-method have been successfully applied for noise reduction in nuclear medicine images, but the restored images exhibited two kinds of artifacts.

- The first kind has to do with behavior near discontinuities, where we observe pseudo-Gibbs phenomena, alternating undershoot and overshoot of a specific target level.
- The second kind concerns some residual blocking artifacts which are obviously due to the use of the Haar wavelet. The Haar wavelet has only one vanishing moment [Mal97]. It is therefore a priori not well adapted to approximating smooth functions. Although this drawback has been overcome by the introduction of a smoothing regularization constraint which tends to smooth the image, we still observe some residual blocking artifacts. A possible fix would be to change the wavelet, but this makes the selection of the significant wavelet coefficients much harder. Indeed this latter is based on the knowledge of the probability density function of the wavelet coefficients for a locally uniform Poisson distribution. The particular choice of the unnormalized Haar wavelet leads to a simple analytical expression, see Equation (4.8), but for any other wavelet the problem is not trivial.

Coifman and Donoho [CD95a] attributed the first kind of artifacts to the lack of translation invariance of the traditional (orthogonal, maximally-decimated [Mal97]) wavelet transform. They noticed that these artifacts are connected in some way with mis-alignments between features in the signal and features in the wavelet. Their approach was to forcibly apply a range of shifts to the signal and average over the results so obtained. Their method, termed "cycle-spinning" [CD95a] consists in "averaging out" the translation dependence of the wavelet transform, see Chapter 2.

Following the idea of Coifman and Donoho we replaced the so far used maximally decimated wavelet transform by a non-decimated shift invariant one described in Section 2.5 in the general framework of DeQuant. It comes out that this new approach does not only improves the Gibbs phenomenon but also shows benefits in suppressing the residual blocking artifacts.

Thanks to the operator notation introduced in Chapter 2, the introduction of the SI feature in the DeQuant framework does not alter the general structure of the algorithms. It only modifies the transform analysis and synthesis equations as can be seen in the algorithms descriptions given in Appendix A, Sections A.2 and A.7 and in the tables of Sections A.5 and A.10.

4.4. Edge-Preserving DeQuant

Recall that the regularization constraint was the condition which, on the one hand garantied the unicity of the restoration problem, and on the other hand enabled to introduce a priori information on the desired solution. In the case where an a priori of smoothness was made on images, we have chosen the regularization constraint to be of type minimum gradient, which lead to the equation

$$L_2 F_Y(2^j) = 0 \tag{4.21}$$

for the denoising method, and

$$L_2 X = 0 \tag{4.22}$$

for the restoration method. Equations (4.21) and (4.22) define edges in the images $F_Y(2^j)$ and X. But the edges defined by these equations do not correspond to the human visual perception, since the eye associates the limit of an object to zero-crossings of the second directional derivative along the gradient direction. This fact motivated the introduction of another 2-dimensional edge detector associated with inflexion points along the maximum slope line.

Regularization

Image edges occur in places of significant intensity changes in the image. The usual aim of edge detection is to locate edges corresponding to the human visual perception of the limits of an object. Since the human eye is sensible to inflexion points, the limits of an object could be defined, as in the 1D case, by identifying zero-crossings of the second derivative. This leads, for a two dimensional intensity profile I to an equation of the type

$$L_2 I = 0 \tag{4.23}$$

where L_2 is the two dimensional Laplacian operator. The problem is that the solutions of Equation (4.23) do not correspond to the inflexion points along the maximum slope line.

In order to demonstrate this, let us consider the example of a Gaussian intensity profile. In this case the inflexion points in the gradient direction are localized at positions $\pm\sigma$ for the one dimensional case, and, due to the isotropy of the profile, on a circle of radius σ for the two dimensional case. Let us now have a look at the zero crossings of the Laplacian. For the one dimensional case we obtain

$$g(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \qquad \frac{d^2 g(x)}{dx^2} = 0 \Rightarrow x = \pm\sigma \tag{4.24}$$

The solution corresponds to the inflexion points. In the two dimensional case, we have

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad \frac{\partial^2 g(x,y)}{\partial x^2} + \frac{\partial^2 g(x,y)}{\partial y^2} = 0 \Rightarrow x^2 + y^2 = 2\sigma^2 \quad (4.25)$$

The solution here is a circle of radius $\sqrt{2}\sigma$, which does not correspond to the inflexion points along the maximum slope line.

Another algorithm to determine image edges

Bijaoui and Froeschlé showed in [SMB95] that for inflexion points location, we have to search along the gradient direction for zero-crossings of the second directional derivative. The algorithm they developed to determine image edges leads to an equation of the type

$$MI = 0$$

where the operator M is defined by:

$$M = \begin{bmatrix} 1 - 3\sin\theta\cos\theta & -2\sin^2\theta + \cos^2\theta & 1 + 3\sin\theta\cos\theta \\ \sin^2\theta - 2\cos^2\theta & -2 & \sin^2\theta - 2\cos^2\theta \\ 1 + 3\sin\theta\cos\theta & -2\sin^2\theta + \cos^2\theta & 1 - 3\sin\theta\cos\theta \end{bmatrix} \quad (4.26)$$

The angle θ corresponds to the normal to the gradient direction. For more details and proofs, see Appendix B.

A new regularization constraint

We replace the Laplacian based edge detectors given by Equations (4.21) and (4.22) by the new edge detector M. Equation (4.21) thus becomes

$$MF_Y(2^j) = 0 \quad (4.27)$$

and Equation (4.22) becomes

$$MX = 0 \quad (4.28)$$

Using the Van Cittert algorithm to resolve equation (4.27), we obtain instead of Equation (4.13) the new iterative form:

$$F_Y^{(n+1)}(2^j) = F_Y^{(n)}(2^j) - \xi_M \, MF_Y^{(n)}(2^j) \quad (4.29)$$

and instead of Equation (4.17)

$$X^{(n+1)} = X^{(n)} - \xi_M \, MX^{(n)} \tag{4.30}$$

The parameter ξ_M is chosen equal to 0.001 for an optimal convergence of the Van Cittert algorithm with the new operator M.

The introduction of the edge preserving feature in the DeQuant framework does not alter the general structure of the algorithms. It only modifies the regularization equation as can be seen in the algorithms descriptions given in Appendix A, Sections A.3 and A.8 and in the tables of Sections A.5 and A.10.

4.5. Conclusion

As conclusion we would like to comment the term "flexible" that appears in the title of this chapter to qualify the multiresolution restoration framework we developed.

Recall that DeQuant is based on the six following functional steps: Transformation, Selection, Regularization, Projection, Reconstruction and Iteration. For the special case of images corrupted by Poisson noise, we made some choices for the following four steps.

- Transformation: we presented the algorithm with both the Mallat orthogonal maximally decimated wavelet transform (MD-WT) and the Coifman and Donoho's shift invariant wavelet transform (SI-WT). But we could also have used the wavelet packet decomposition [BGG98] which gives a rich structure that allows adaptation to a larger spectrum of object forms. The cost of the shift invariance and of a richer structure is a larger computational complexity. The choice of the Haar wavelet was motivated by the natural match between this wavelet and the image formation process and by the Poisson nature of the data.

- Selection: the flexibility of the selection step lies in the choice of the threshold. We have chosen to work with a unique threshold and apply a hard thresholding, but we could also use two or more thresholds and apply a soft thresholding.

- Regularization: we presented two regularization operators, the Laplacian L_2 and the operator M. After establishing the relation between the regularization constraint and edge detection, we showed that the Laplacian is not optimal for defining image edges. This was the reason that motivated

the introduction of the operator M which corresponds to the human visual perception of the limits of an object.

- Projection: some constraints have been adopted to avoid contradictions between the significance of the wavelet coefficients of the noisy image and the significance of the wavelet coefficients of the reconstructed image. Another choice of constraints could have been made. An additional flexibility is introduced by the parameter δ that appears in the modification of the significant wavelet coefficients.

Hence, each of the method steps can be chosen in such a way that the method adapts to the characteristics of the images it has to restore. If we combine all possibilities we presented in this chapter, we obtain eight different algorithms having the same global structure. The denomination of these algorithms and their descriptions are detailed in Appendix A.

5. PDF of the wavelets coefficients of a Poisson process

The major limitation of nuclear medicine images is the low count level acquired in typical studies. We have seen in Chapter 1 that this fact results in very noisy images that need to be processed by taking into account the statistics of photon events, i.e., Poisson noise. Chapter 3 showed that most of the proposed restoration wavelet methods derive from a Gaussian approach through judicious use of Anscombe's square root transformation. Contrary to this latter approach which does not process perfectly the information, generating bias and artifacts, DeQuant offers a new wavelet restoration method based on the determination of the exact statistic of the wavelet coefficients of the Poisson process.

We calculate in this chapter the general expression of the PDF of the wavelet coefficients of a Poisson process. We then examine the special case of the Haar wavelet for which we give threshold tables.

This chapter is organized as follows. Section 5.1 determines an expression of the PDF of the wavelet coefficients of a Poisson process which is valid for any wavelet function and gives approximations of this PDF for low and high densities of events. Section 5.2 examines the special case of the Haar wavelet. This latter leads to some simplifications in the expression of the PDF and enables to calculate the threshold tables given in Appendix C Section C.3. Section C.3 offers some conclusions.

5.1. The general case

The PDF of the wavelet coefficient of a Poisson process depends on the choice of the wavelet. In this section we derive an expression of the PDF as a function of the wavelet. This expression holds for any choice of the wavelet.

Let $\rho(x) \geq 0$ be a 1-dimensional function taking positive or zero values. The wavelet coefficient of $\rho(x)$ at a given position b and scale a is defined by [Mal97]:

$$W_{\rho(x)}(a,b) = \int_{-\infty}^{+\infty} \rho(x)\psi(\frac{x-b}{a})dx \tag{5.1}$$

We decided not to introduce the normalization factor $\frac{1}{\sqrt{a}}$ in the definition of the wavelet coefficient (see Equation (5.1)) in order to simplify the expression of the PDF. Moreover the wavelet ψ is supposed to take real values. If the function $\rho(x)$ is corrupted by a Poisson noise, we calculate the corresponding wavelet coefficient as follows. We start by dividing the x axis into intervals of size Δx. We define n_k as the number of events counted in the interval located at position $k\Delta x$. The wavelet coefficient of the function $\rho(x)$ corrupted by Poisson noise is then given by the discrete sum [SMB98]

$$W_{\rho(x)}(a,b) \approx \sum_{k=-\infty}^{+\infty} n_k \psi(\frac{k\Delta x - b}{a}) \tag{5.2}$$

n_k is the realization of a random variable N_k following a Poisson distribution [Pap84] of parameter λ_k

$$Prob(N_k = n_k) = \frac{\lambda_k^{n_k}}{n_k!} e^{-\lambda_k} \tag{5.3}$$

where

$$\lambda_k = \int_{k\Delta x}^{(k+1)\Delta x} \rho(x)dx. \tag{5.4}$$

5.1.1. PDF of the wavelet coefficient

The PDF of the wavelet coefficient will be determined by using its characteristic function. Indeed, characteristic function [Sap90] and PDF form a Fourier transform pair so that having one of them enables to obtain the other. Proposition 1 (for proof see Appendix C Section C.1) gives two equivalent expressions for the characteristic function of the wavelet coefficient of a Poisson process.

Property 1 *Consider a wavelet function ψ having a compact support D of length 1. If the function $\rho(x)$ is constant and equal to ρ_0 on an interval of length a, where a is the scale parameter, then the characteristic function of the wavelet coefficient is given by one of the following equivalent expressions :*

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha(H_\psi(\nu)-1)} \tag{5.5}$$

or

$$\phi_{\mathcal{W}}(\nu) = e^{a\rho_0 \int_D (-1+e^{-i2\pi\nu\psi(s)})ds} \tag{5.6}$$

where the parameter $\alpha = \rho_0 a$ corresponds to the mean number of events in the interval a and $H_\psi(\nu)$ is the Fourier transform of the normalized histogram of the wavelet ψ.

The characteristic function given by Equation (5.5) corresponds to the PDF $p_{\mathcal{W}}(\alpha, w)$ given by Proposition 2 (for proof see Appendix C Section C.2)

Property 2 *The probability density function corresponding to the characteristic function*

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha(H_\psi(\nu)-1)}$$

is given by

$$p_{\mathcal{W}}(\alpha, w) = e^{-\alpha}\delta(w) + e^{-\alpha}\sum_{n=1}^{\infty}\frac{\alpha^n}{n!}H_\psi^{(*n)}(w) \tag{5.7}$$

where $H_\psi^{(*n)}(w)$ denotes n autoconvolutions of $H_\psi(w)$

$$H_\psi^{(*n)}(w) = \underbrace{H_\psi * \ldots * H_\psi}_{n \text{ times}}(w)$$

The characteristic function given by Equation (5.6) corresponds to the PDF $p_{\mathcal{W}}(\beta, u)$ given by Proposition 3 (for proof see Appendix C Section C.1). The parameter β is defined as $\frac{1}{\sqrt{\alpha}}$ and the variable u is the centered and normalized version of w:

$$u = \frac{w - \mathrm{E}(\mathcal{W})}{\sqrt{\mathrm{Var}(\mathcal{W})}}$$

Property 3 *The probability density function $p_{\mathcal{W}}(\beta, u)$ corresponding to the characteristic function*

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha \int_D (-1+e^{-i2\pi\nu\psi(s)})ds}$$

is given by Table C.1.

The Hermite polynomials [Bec73] [SO] [AS64] introduced in Table C.1 are generated by successive differentiation of the function $e^{-\frac{u^2}{2}}$[1]:

$$P_n(u) = e^{\frac{u^2}{2}} \frac{d^n}{du^n} e^{-\frac{u^2}{2}} \tag{5.8}$$

explicitly, we can write:

$$P_n(u) = \sum_{k=0}^{2k \leq n} (-1)^{n-k} \frac{n!}{2^k k!(n-2k)!} u^{n-2k} \tag{5.9}$$

The first eighteen polynomials are given by Table C.5.

The expressions of the PDF given by Equation (5.7) and Table C.1 are valid for all densities of events, i.e., for all values of α and β. These expressions have the drawback to contain an infinity of terms. The objective of the next two sections is to obtain expressions made of finite sums by limiting the validity of the Equation (5.7) and Table C.1 to low, respectively high, densities of events.

5.1.2. Approximation for low densities of events

For small values of α, Equation (5.7) can be limited to terms of the first order in α:

$$\begin{aligned} p_{\mathcal{W}}(\alpha, w) = e^{-\alpha}\delta(w) + \alpha e^{-\alpha} H_{\psi}^{(*1)}(w) \\ + (1 - e^{-\alpha} - \alpha e^{-\alpha}) H_{\psi}^{(*2)}(w) + o(\alpha) \end{aligned} \tag{5.10}$$

Equation (5.10) has been derived by taking the first two terms of Equation (5.7) and adding a corrective term that insures that the PDF sums to 1. This corrective term is (1 - term in α^0 - term in α^1)$H_{\psi}^{(*2)}(w)$.

5.1.3. Approximation for high densities of events

If we limit the validity of the formula of Table C.1 to large values of α or equivalently to small values of β, we obtain a finite sum that represents an approximation at a given order of the PDF. The approximation to the 6th order in β is given by:

$$p_{\mathcal{W}}(\beta, u) = g(u)[1 + a\beta + b\beta^2 + c\beta^3 + d\beta^4 + e\beta^5 + f\beta^6] + o(\beta^6) \tag{5.11}$$

[1] In some theoretical investigations, Hermite polynomials are defined by the differentiation of the function e^{-u^2}. However, in practical applications, particularly those associated with probability theory, it is more advantageous to use the slightly different polynomials associated to the function $e^{-\frac{u^2}{2}}$

with

$$g(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

The coefficients a, b, c, d, e, f are given by Table C.6. The expression (5.11) shows that if β is small enough the PDF converges to a Gaussian distribution.

To conclude this section, we can say that three separate domains can be distinguished for the calculation of the PDF of the wavelet coefficient for a Poisson process.

- For low density of events ($\alpha \leq 2^{-10}$), the approximation given by Equation (5.10) holds. The error is of the order of α.
- For high densities of events ($\alpha \geq 2^{10}$), the PDF tends towards a Gaussian distribution and the approximation given by Equation (5.11) can be used. The error is of the order of $1/\alpha^3$.
- For an intermediate density, the histogram of the wavelet has to be estimated and the PDF is computed by the inverse Fourier transform of Equation (5.5).

5.2. Special case: the Haar wavelet

In this section, we examine the special case of the Haar wavelet [Mal97] which is defined on the interval $[0, 1]$. Its value equals 1 for $x \leq \frac{1}{2}$ and -1 on the other part of the definition interval. This wavelet leads to some simplifications in the expression of the PDF.

5.2.1. PDF of the wavelet coefficient

We will show that the PDF of the wavelet coefficient of a Poisson process in the case of the Haar wavelet can either be calculated by applying the general Equation (5.5) to the particular case of the Haar wavelet or directly derived from the definition of the wavelet coefficient.

Derivation from Equation (5.5)

In order to use the general Equation (5.5) in the particular case of the Haar wavelet, we first determine the histogram of this wavelet which is obviously:

$$H_\psi(w) = \frac{1}{2}[\delta(w-1) + \delta(w+1)] \tag{5.12}$$

After substituting the Fourier transform of Equation (5.12) in Equation (5.5), we obtain the formula for the characteristic function of the wavelet coefficient of a Poisson process for the Haar wavelet.

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha(\cos(2\pi\nu)-1)} \tag{5.13}$$

$\phi_{\mathcal{W}}(\nu)$ being periodic, it corresponds in the time domain to a discrete PDF [OS89], $p_{\mathcal{W}}(\alpha, n)$. Moreover since $\phi_{\mathcal{W}}(\nu)$ is an even function we can write the discrete time Fourier transform as:

$$\begin{aligned} p_{\mathcal{W}}(\alpha, n) &= 2 \cdot \frac{1}{2\pi} \int_0^\pi e^{\alpha(\cos(\omega)-1)} \cos(n\omega) d\omega \qquad \text{with } \omega = 2\pi\nu & (5.14)\\ &= e^{-\alpha} \frac{1}{\pi} \int_0^\pi e^{\alpha\cos(\omega)} \cos(n\omega) d\omega & (5.15)\\ &= e^{-\alpha} I_n(\alpha) & (5.16) \end{aligned}$$

Where $I_n(\alpha)$ is the modified Bessel function [GM52] corresponding to Equation 9.6.19 in [AS64]:

$$I_n(z) = \frac{1}{\pi} \int_0^\pi e^{z\cos\theta} \cos(n\theta) d\theta \tag{5.17}$$

Direct derivation

The probability density function could also be found directly by noticing that in the case of the unnormalized Haar wavelet, a wavelet coefficient is obtained by the difference of two counts: $\mathcal{W} = C_1 - C_2$. Each of those two counts taken on an interval of size $\frac{1}{2}$ is distributed following a Poisson law of parameter $\frac{\alpha}{2}$. The PDF of the wavelet coefficient can be computed as follows. For $n \geq 0$, we have

$$\begin{aligned} p_{\mathcal{W}}(\alpha, w = n) &= \sum_{m=0}^{+\infty} p(C_1 = n+m) p(C_2 = m) \\ &= e^{-\alpha} \sum_{m=0}^{+\infty} \frac{(\frac{\alpha}{2})^{n+2m}}{(n+m)!m!} \\ &= e^{-\alpha} I_n(\alpha) \end{aligned}$$

For negative values of n, $p_{\mathcal{W}}(\alpha, w = n)$ can be computed by symmetry. Notice that we obtain again the result given by Equation (5.16).

5.2.2. Approximation for low densities of events

For low densities of events, Equation (5.10) can be used in the particular case of the Haar wavelet to give an approximation of the PDF in the first order in α. But in the case of the Haar wavelet, due to the simplicity of the expression of the histogram we can go a little bit further and estimate more precisely the value taken by the PDF for a given value n taken by the wavelet coefficient. We therefore replace $H_\psi(w)$ by its value for the Haar wavelet in Equation (5.7):

$$\begin{aligned} p_{\mathcal{W}}(\alpha, w) &= e^{-\alpha}\delta(w) \\ &+ \tfrac{\alpha^1}{1!2}\, e^{-\alpha}[\delta(w-1) + \delta(w+1)] \\ &+ \tfrac{\alpha^2}{2!2^2}\, e^{-\alpha}[\delta(w-2) + 2\delta(w) + \delta(w+2)] \\ &+ \tfrac{\alpha^3}{3!2^3}\, e^{-\alpha}[\delta(w-3) + 3\delta(w-1) + 3\delta(w+1) + \delta(w+3)] \\ &+ \ldots \end{aligned}$$

Looking at the above equation, we can see that for small values of α, the probability that $\mathcal{W} = n$ can be approximated by:

$$p_{\mathcal{W}}(\alpha, w = n) = \frac{\alpha^n}{n!2^n}\, e^{-\alpha} \tag{5.18}$$

Due to the fact that for faint values of α, $e^{-\alpha} \backsimeq 1$, equation (5.18) can be replaced by:

$$p_{\mathcal{W}}(\alpha, w = n) = \frac{\alpha^n}{n!2^n} \tag{5.19}$$

The repartition function is approximated by the value of the PDF:

$$F(\alpha, w = n) = \frac{\alpha^n}{n!2^n} \tag{5.20}$$

5.2.3. Approximation for high densities of events

We will use Equation (5.11) and Table C.6 to derive the expression of the PDF for large values of α in the case of the Haar wavelet. In order to calculate the coefficients a, b, c, d, e, f we need to compute for the Haar wavelet the value of the integrals I_n defined in Equation (C.14). It comes out that:

$$\begin{aligned} I_n &= 0 \quad \text{for } n \text{ odd} \\ I_n &= 1 \quad \text{for } n \text{ even} \end{aligned}$$

After substituting those values in Equation (5.11) we obtain:

$$p_{\mathcal{W}}(\beta, u) = g(u)[1 + b\beta^2 + d\beta^4 + f\beta^6] + o(\beta^6) \tag{5.21}$$

with :

$$\begin{aligned} b &= \tfrac{1}{24}P_4(u) \\ d &= \tfrac{1}{720}P_6(u) + \tfrac{1}{1152}P_8(u) \\ f &= \tfrac{1}{40320}P_8(u) + \tfrac{1}{17280}P_{10}(u) + \tfrac{1}{82944}P_{12}(u) \end{aligned}$$

The repartition function $F(\beta, u)$ can be computed by integrating $p_{\mathcal{W}}(\beta, u)$. Due to the fact that we have a discrete variable, a systematic bias will appear in the numerical integration. This will be corrected by adding half of the value of the PDF at the coefficient value. While integrating Equation (5.21), we take into account the property of the Hermite polynomials [SO] stating that the primitive of $P_n(u)g(u)$ is $P_{n-1}(u)g(u)$. This property derives directly from the definition $P_n(u)$, see Equation (5.8). We obtain for the repartition function:

$$F(\beta, u) = \frac{1}{2}erfc(\frac{u}{\sqrt{2}}) + g(u)[k\beta^2 + l\beta^4 + m\beta^6] + \frac{1}{2}p_{\mathcal{W}}(\beta, u) + o(\beta^6) \tag{5.22}$$

with

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}dt$$

and

$$\begin{aligned} k &= \tfrac{1}{24}P_3(u) \\ l &= \tfrac{1}{720}P_5(u) + \tfrac{1}{1152}P_7(u) \\ m &= \tfrac{1}{40320}P_7(u) + \tfrac{1}{17280}P_9(u) + \tfrac{1}{82944}P_{11}(u) \end{aligned}$$

5.2.4. Numerical simulations

A host of applications including compression, denoising and restoration of data corrupted by Poisson noise, are based on the determination of the statistical significance of the wavelet coefficients of a Poisson process. The statistical significance of a coefficient $W(a, b)$ is measured by (i) first calculating the probability of getting an outcome at least as far as the actually observed coefficient from what we would expect if the coefficient is supposed non significant, this probability is called the P-value, and (ii) then comparing the obtained value with a fixed level of evidence ϵ that we regard as decisive:

- If $W(a,b) > 0$

$$\text{P-value=Prob}(\mathcal{W}(a,b) > W(a,b)) < \epsilon$$

- If $W(a,b) < 0$

$$\text{P-value=Prob}(\mathcal{W}(a,b) < W(a,b)) < \epsilon$$

If the P-value is smaller than ϵ, we say that we detected a significant wavelet coefficient at level ϵ.

In order to compute the P-value, we need the PDF $p_{\mathcal{W}}(\alpha, w)$ of the wavelet coefficient for a locally uniform Poisson process. For each value of α, the significant level ϵ corresponds to a threshold depending on α and ϵ.

Through numerical simulations of the PDF, we calculated the thresholds corresponding to the following choices of α and ϵ:

$$\begin{array}{rcl} \alpha & \in & [2^{-30}\ ;\ 2^{30}] \\ \epsilon & \in & [10^{-2},\ 10^{-3},\ 10^{-4},\ 10^{-5},\ 10^{-6}] \end{array}$$

These thresholds, given by Table C.7, correspond to the smallest positive integer such that the integral of the PDF from this integer to infinity is smaller than ϵ:

$$\text{threshold}(\epsilon, \alpha) = \min_x \{ \int_x^{\infty} p_{\mathcal{W}}(\alpha, u) du \leq \epsilon \} \tag{5.23}$$

Figure 5.1 shows for all values of ϵ the normalized thresholds corresponding to

$$\text{normalized threshold} = \frac{\text{threshold}(\epsilon, \alpha)}{\sqrt{\alpha}} \tag{5.24}$$

Table C.7 shows that for large values of α, namely for $\alpha \geq 2^{16}$, the normalized thresholds tend towards the values that would be obtained for a Gaussian distribution:

	ϵ				
	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
normalized thresholds for a Gaussian distribution	4.75	4.26	3.72	3.09	2.33

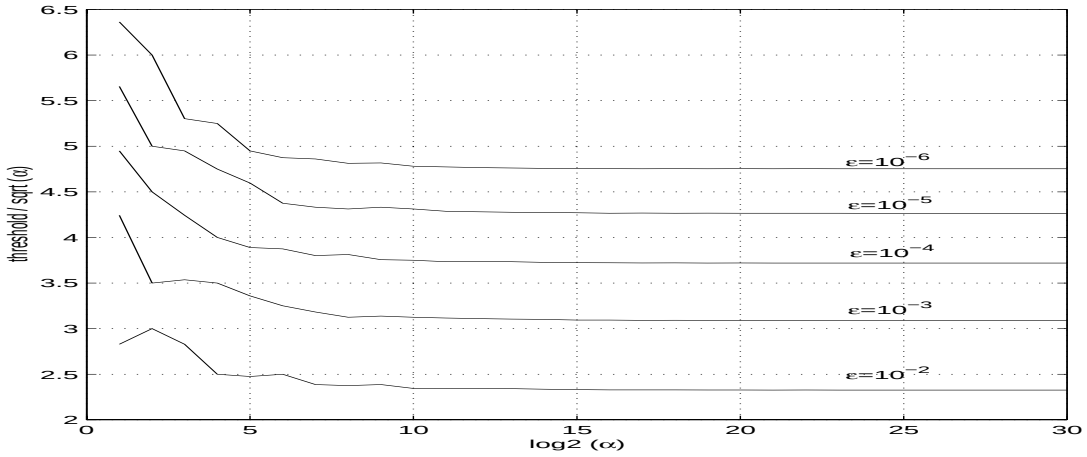


Figure 5.1.: Normalized thresholds for $\alpha \in [2^0 ; 2^{30}]$ and $\epsilon \in [10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}]$

5.3. Conclusion

We determined in this chapter the general expression of the PDF of the wavelet coefficients of a Poisson process. We then gave approximations of this PDF for large and faint densities of events and showed that for large densities of events the PDF tends to a Gaussian distribution. We considered the special case of the Haar wavelet for which we calculated tables of thresholds through numerical simulations. These tables have been used in DeQuant to assess the statistical significance of the wavelet coefficients, see Chapter 4 and [JB99b] [JB99a].

6. Application of DeQuant to nuclear medicine

After conception of any new medical image processing algorithm, validation is an important step to insure that the procedure fulfills all requirements set forth at the initial design stage. We will concentrate for the validation of DeQuant on thyroid studies where nuclear medicine imaging is of particular importance in the detection of the so called nodular disease. This latter corresponds to the development by the thyroid of one or more localized swellings called nodules.

The thyroid is a butterfly shaped gland which wraps around the front part of the windpipe just below the Adam's apple, see Fig. 6.1. It produces hormones that regulate the body's metabolism and organ function. A nodule is an area of the gland that appears to be different than the rest by being raised or having a different texture. Nodules are very common: it is estimated that about 50% of the population will develop a thyroid nodule at some time in their life. Nodules can be very small, less than a millimeter, or as large as a few centimeters. They can occur alone or with others of different sizes in the same gland and may be cancerous.

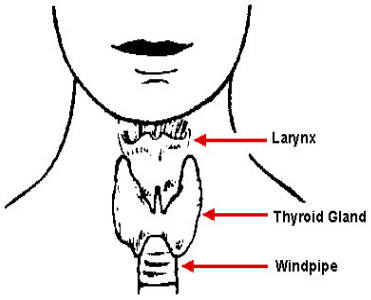


Figure 6.1.: The thyroid gland

Thyroid nodules do not function like normal thyroid tissue. They are either underactive (cold) if they fail to produce enough hormone or overactive (hot) if

they produce more thyroid hormone than is needed. When the patient receives the radioactive solution prior to the scan, cold nodules take up less and hot nodules more radioactive material than the normal thyroid tissue. This behavior enables to differenciate them in the thyroid scan.

However the noise that corrupts nuclear medicine images makes it difficult for a physician examining a thyroid scan to evaluate the nodules, particularly if they are small and not good differenciated. Evaluation means here, identifying the nodules, recognizing their function and estimation their form and size. The objective of this chapter is to demonstrate the potential of DeQuant in increasing diagnostic confidence in nuclear medicine.

This chapter is organized as follows. Section 6.1 presents the criteria that have been adopted to measure the accuracy of the restoration algorithms. Section 6.2 analyzes the contrast and size sensibility of D-DeQuant and gives some insights in the choice of the algorithms parameter. Section 6.3 illustrates individually the effect of the introduction of the point spread function, the edge preserving regularization constraint and the shift invariant feature in the general framework of DeQuant. It then compares the performances of the eight DeQuant algorithms on a simulated data experiment. Section 6.4 examines the behavior of DeQuant on actual images of thyroid phantoms. Section 6.5 presents the last stage of the evaluation process, namely the algorithm's validation on real data acquired in vivo. Section 6.6 offers some conclusions.

6.1. Performance criteria

The accuracy of the restoration is measured using a bunch of performance criteria, that we introduce in the following section. We distinguish thereby between the global criteria that give an overall idea of the accuracy of the restoration process and the local criteria that render more precisely the quality of the restoration of a given region of interest.

6.1.1. Global criteria

We considered two global criteria, the signal to noise ratio and the compression ratio.

Signal to noise ratio (SNR): it is defined by

$$\mathrm{SNR}_{dB}(I_0, I_1) = 10\ \log_{10}\left(\frac{\sum_{k_x}\sum_{k_y} I_0^2(k_x,k_y)}{\sum_{k_x}\sum_{k_y}(I_0(k_x,k_y)-I_1(k_x,k_y))^2}\right)$$

I_0 is usually the original image. I_1 can represent the noisy image or the image restored with one of the algorithms.

Compression ratio (CR): this criterion is only defined for DeQuant when used with the maximally decimated wavelet transform. Since the sole knowledge of the significant wavelet coefficients and of the last approximation image permits, using the reconstruction algorithm, to obtain the estimated image, it suffices to store the significant wavelet coefficients and the last approximation image instead of storing the whole reconstructed image. This results in an important gain of storage place, which is quite convenient from an archival point of view. The compression ratio (CR) is defined as follows:

$$\mathrm{CR} = \frac{N_I}{N_{SWC} + N_{F_I(2^J)}}$$

where N_{SWC} is the number of detected significant wavelet coefficients, $N_{F_I(2^J)}$ is the number of pixels of the last approximation image and N_I is the number of pixels of the original image I.

6.1.2. Local criteria

In nuclear medicine images, the relevant information is often very localized and a global criterion, such as the SNR, will not render the medical quality of the image. A restoration algorithm for medical images should conserve intensity related, position related and morphological information. It was thus necessary to introduce some criteria that are directly associated with these properties.

Definition of the region of interest

The definition of these local criteria requires the definition of an object or a region of interest. This has been done using either an interactive or an automatic method.

While the interactive method enables the user to select a polygonal region of interest using the mouse by clicking on points within the image, the automatic method operates by estimating the background and the noise levels in the image

after the elimination of the significant structures. We obtain two thresholds that define a mask isolating the objects from the background. We used for objects having a Gaussian distributions the so called $k\sigma$ clipping algorithm which can be described as follows:

1. Extract the mean value, m, and the standard deviation σ of the image.
2. Throw away all pixels whose values are outside the interval $[m-3\sigma\,;\,m+3\sigma]$. The remaining pixels make up the new set of values.
3. Return to 1 with the new set of values until the limit thresholds are reached.

Definition of the local properties

Once the region of interest has been defined, the local criteria can be calculated. They give information on the position, intensity and morphology of the object.

Position: is given by the coordinates (x_G, y_G) of the object's center of gravity.

Size: is represented by the full width at half maximum of the object in the x and y directions, (fwhm_x, fwhm_y). For a Gaussian intensity profile, these latter are related to the standard deviation of the object by $\text{fwhm} = 2\sigma\sqrt{2\ln 2}$.

Perimeter: is given by the number of pixels that are part of the perimeter of the object.

Area: corresponds to the total number of pixels associated with the object.

Flux: equals the sum of the intensities of the pixels associated with the object.

Relative Amplitude (RA): corresponds to the amplitude of the center of gravity with respect to the background.

Theta: equals the angular deviation of the object with respect to the horizontal axis.

6.2. Performance of D-DeQuant

This section assesses the performance of D-DeQuant on simulated data experiments. We start by investigating the contrast and size sensibility of D-DeQuant as a function of the image's noise. We then gives some insights in the choice of the algorithms' parameters for an optimal result in the restoration process.

6.2.1. Contrast and size sensibility of D-DeQuant

The detectability of a nodule in a noisy image depends on the one hand on the size and contrast of the nodule, and on the other hand on the noise level of the image. The size of the nodule is given by its fwhm and the contrast is given by its RA. We define the noise level (NL) of an image $I(k_x, k_y)$ corrupted by a Poisson noise as:

$$NL = \sqrt{\frac{\sum_{N_I} I(k_x, k_y)}{N_I}} \tag{6.1}$$

where N_I is the number of pixels of the image I. In order to investigate the interdependance of RA, fwhm and NL, we consider simulated images of size 64x64 consisting of a background of constant intensity where a hot nodule is represented by a Gaussian intensity profile. These images are corrupted by a Poisson noise and then restored using D-DeQuant with the parameters $J = J_{Max} = \log_2(64) = 6$, $\epsilon = 10^{-5}$ and $\delta = 0.5$.

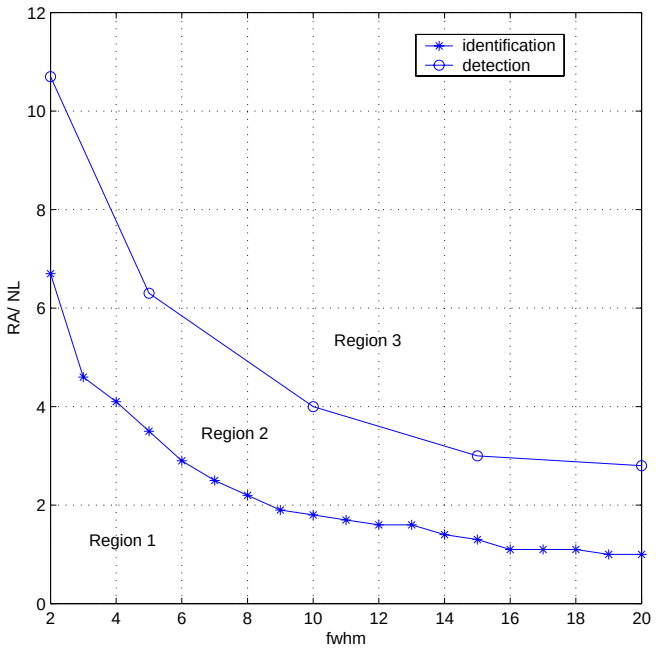
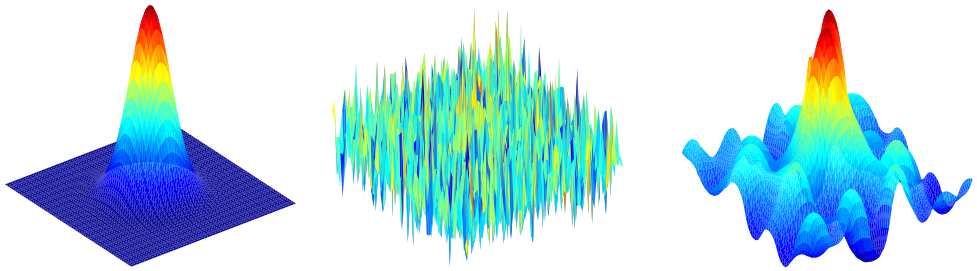


Figure 6.2.: Identification and detection curves of D-DeQuant

We were first interested in finding the minimum value of RA for which an identification of the nodule in the restored image is beyond doubt possible. This value will be referred to as RA^{Id}_{min}. We therefore made the fwhm of the nodule vary

between 2 and 20 pixels and retained the value of RA under which no nodule was detected in the restored image. For the determination of RA^{Id}_{min}, we did not pay a lot of importance to the form of the detected nodule and to the artifacts of the background as long as we where able to say beyond doubt that there is one and only one nodule in the restored image. Fig. 6.3 illustrates the quality criteria adopted for RA^{Id}_{min}. The obtained identification curve given by Fig. 6.2 decreases to the asymptote $RA^{Id}_{min} = NL$. The smaller the fwhm, the bigger the minimum value of RA that enables a detection.

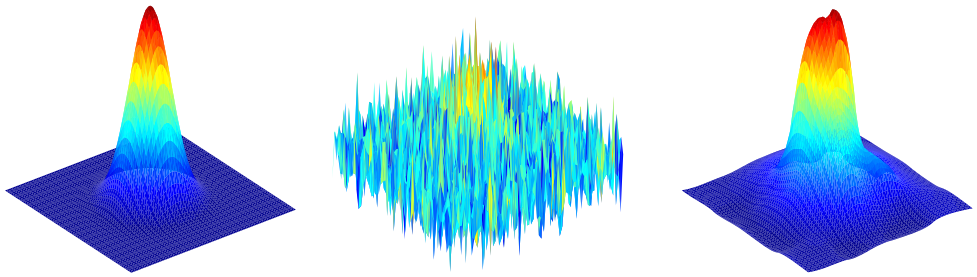


(a) Original image (b) Noisy image (c) Result of D-DeQuant

Figure 6.3.: Denoising with D-DeQuant of an structure which characteristics are on the identification curve; fwhm = 15, $RA = 1.3 * NL$.

We then wanted to find the minimum value of RA for which a good denoising of the nodule was possible. This value will be referred to as RA^{D}_{min}. With good denoising we mean that position related and morphological information remain faithful to the original data. Fig. 6.4 illustrates the quality criteria adopted for RA^{D}_{min}. The same experience as previously has been performed and the the detection curve in Fig. 6.2 came out. Notice now that the detection curve seems to decrease to the asymptote $RA^{D}_{min} = 2.8 * NL$. Again the smaller the fwhm, the bigger the minimum value of RA that enables a good denoising.

The identification and detection curves divide the plane in three regions, see Fig. 6.2. If a structure is in region 1, it will not be detected by D-DeQuant. If it is in region 2, the structure is restored with some small imperfections regarding its form and size. In region 3, we have an optimal denoising in the sense that intensity-related, position related and morphological information remain faithful to the original data.



(a) Original image (b) Noisy image (c) Result of D-DeQuant

Figure 6.4.: Denoising with D-DeQuant of an structure which characteristics are on the detection curve; fwhm = 15, $RA = 3 * NL$.

6.2.2. Choice of D-DeQuant's parameters

The D-DeQuant algorithm depends on three parameters: the scale J to which the wavelet transformation is calculated, the level of evidence ϵ and the tolerance parameter δ that enables to take into account the noise in the significant wavelet coefficients. The following simulated data experiments help understanding the role of each parameter, and give insights in their choice.

Choice of J

The parameter J represents the scale to which the wavelet transform is calculated. One would a priori assume that J must be taken equal to its maximum value $J_{Max} = \log_2(\sqrt{N_I})$, where N_I is the number of pixels in the image, in order to have the possibility to detect the significant wavelet coefficients corresponding the all object sizes in the image. The SWC corresponding to the small objects would appear at low scales and those associated to big objects at high scales. The following experiment investigates the validity of this assumption and gives some additional insights for the choice of J.

We consider three images of size 64x64 consisting of a background of constant intensity where a hot nodule of fwhm=15, and RA=0.9*NL ($\in$ Region 1), 2*NL ($\in$ Region 2) and 4*NL ($\in$ Region 3) respectively, is represented by a Gaussian intensity profile. After introducing Poisson noise, all three images are restored with D-DeQuant, where the parameters $\epsilon = 10^{-5}$ and $\delta = 0.5$ have been fixed and

the scale J varies from 3 to the maximum scale $J_{Max} = 6$. The restored images are given in Fig. 6.5 together with their corresponding last smooth images. Fig. 6.6 shows the the total number of detected significant coefficients and the variation of the SNR as a function of the scale. ΔSNR in % is defined as:

$$\Delta\text{SNR}(\%) = \frac{\text{SNR(orig. im., restored im.)-SNR(orig. im., noisy im.)}}{\text{SNR(orig. im., noisy im.)}} * 100 \tag{6.2}$$

Let us first have a look at the image with RA=0.9*NL. Since this image belongs to Region 1, see Fig. 6.2, we are expecting to obtain a constant image for $J = 6$. This means that no significant coefficients have been detected, see Fig. 6.6. In this case, the restored image is reconstructed using only the information in the last smooth image. This fact can be easily verified by noticing the similarities between the restored images and their corresponding last smooth images. Thus the restored image is nothing but a local average of the noisy image. This local averaging smoothes naturally the noise, and if it is not pushed too far as it is the case for J=5 and 6, we can distinguish a structure in the restored image which is due to nodule (J=3 and 4). The ΔSNR values confirm the visual impression, see Fig. 6.6.

If we now consider the images with RA=2*NL and RA=4*NL, we first notice that although ΔSNR is maximum for $J = 3$ some residual blocks due to the last smooth images are still present in the restored images. Comparing the restored images at scale $J = 5$ and $J = 6$, we notice that they are visually equivalent and exhibit the same number of significant wavelets coefficients and the same ΔSNR value.

These observations yield the following conclusion. It is not necessary to use the full scale J_{Max} since $J = J_{Max} - 1$ provides the same results. However for structures which characteristics are under the identification curve, it may be interesting to compute the restoration at $J = J_{Max}/2$ where they might appear, surrounded by artifacts.

Choice of ϵ

The parameter ϵ corresponds to the significance level of the hypothesis test. It represents the probability that the null hypothesis is rejected when it is in fact true, that is H_0 is wrongly rejected. ϵ equals thus the probability that a coefficient be classified as significant when it in fact lacks enough local information. The following experience illustrates the influence of the choice of ϵ on the restored image.

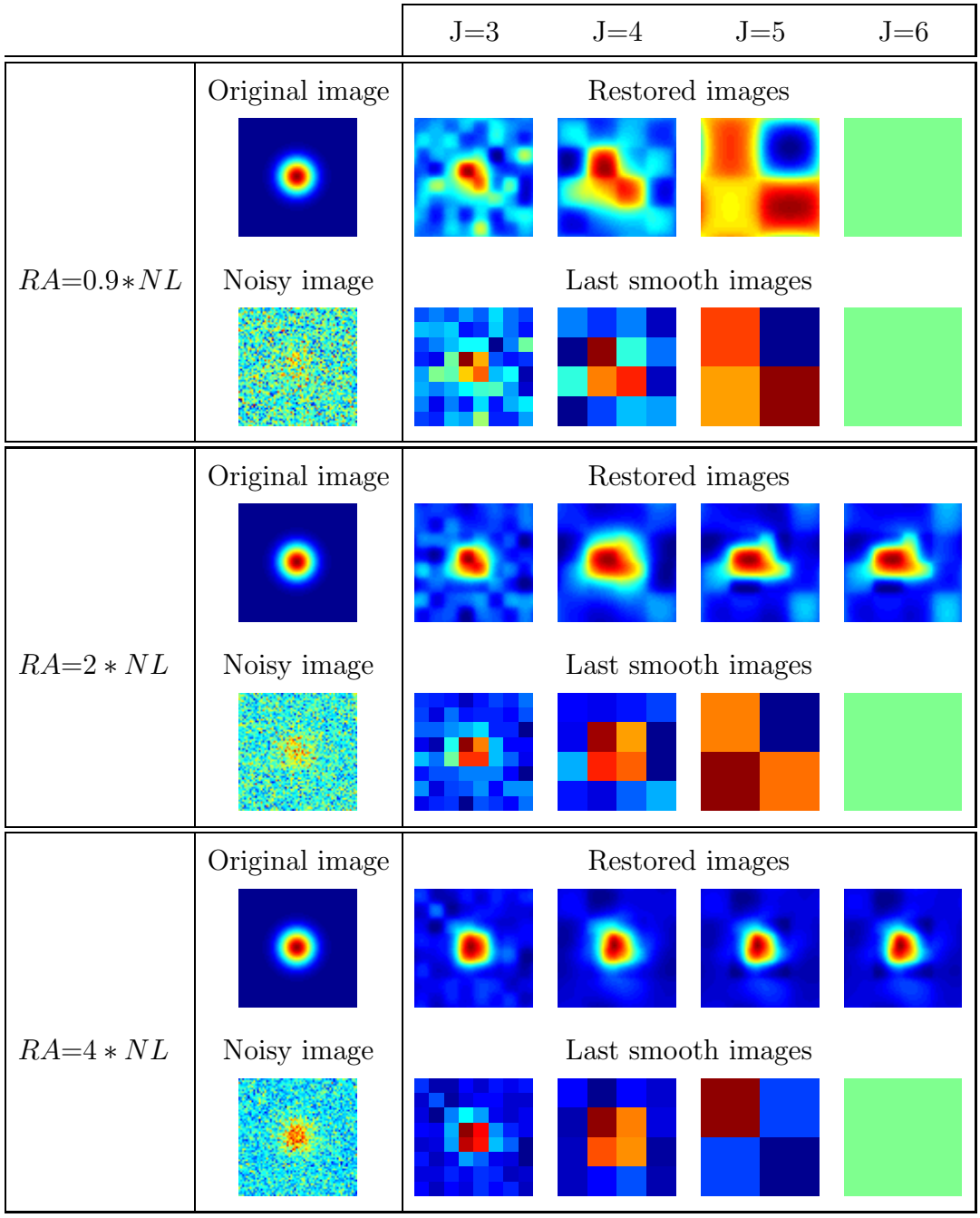
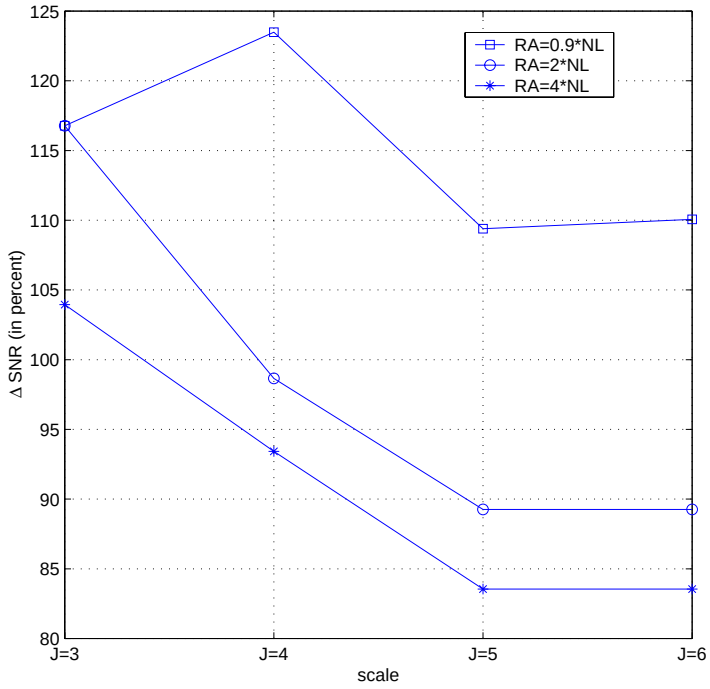


Figure 6.5.: Denoising with D-DeQuant of an image of fwhm=15



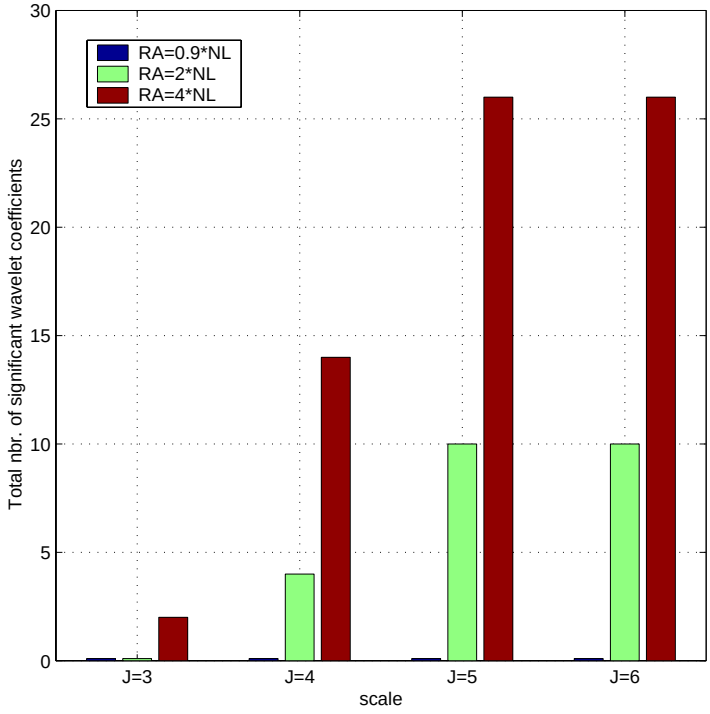


Figure 6.6.: ΔSNR (for definition see Eq. (6.2)) and total number of significant coefficients of the restored images of Fig. 6.5

We restore the noisy image of RA=4*NL and fwhm=15 introduced in Fig. 6.5, using D-DeQuant with the following parameters: $J = 5$, $\delta = 0.5$ and

$$\epsilon \in [10^{-2},\ 10^{-3},\ 10^{-4},\ 10^{-5},\ 10^{-6}].$$

The restored images can be seen in Fig. 6.7. The total number of significant coefficients detected in the restoration process and a ΔSNR curve are given by Fig. 6.8.

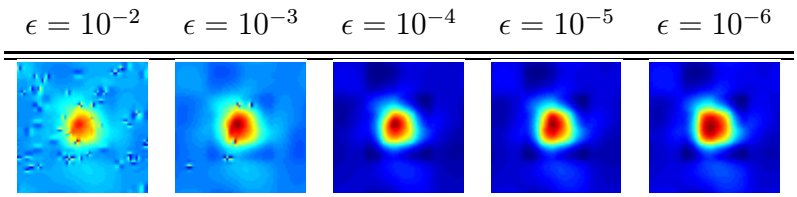


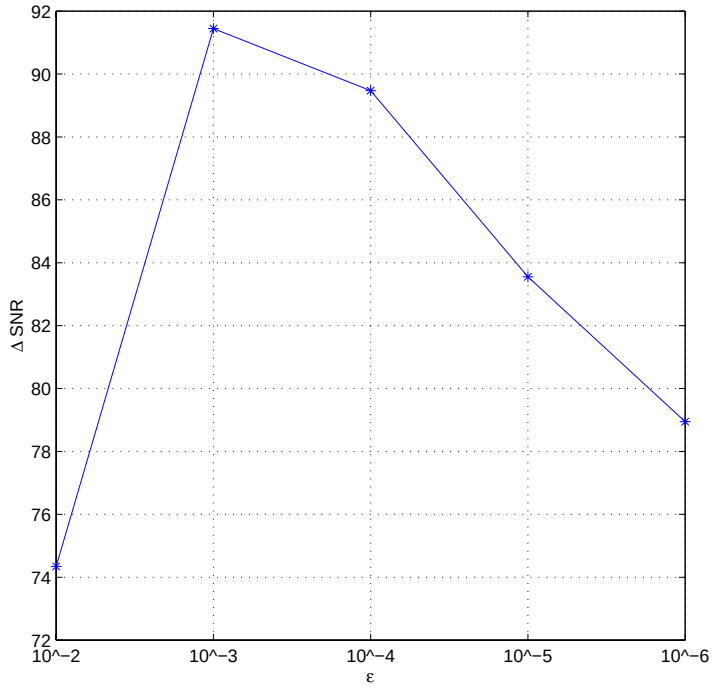
Figure 6.7.: Influence of the parameter ϵ on the result of the restoration with D-DeQuant of the noisy image with the nodule of RA=4*NL, given by Fig. 6.5.

Looking at Fig. 6.7, we notice that the restoration with $\epsilon = 10^{-2}$ and 10^{-3} exhibits small artifacts due to the detection of significant wavelet coefficients at scale $j = 1$ and 2, see Fig. 6.8. Given the size of the nodule, wavelet coefficients due to the nodule do not appear under the scales $j = 3/4$.

If we admit that we can approximate the probability by a frequency, ϵ would represent the proportion of coefficients classified as significant when they are actually due to noise. This enables us to make the following calculation. The image we considered corresponds to $64^2 = 4096$ wavelet coefficients. Taking $\epsilon = 10^{-2}$ means that we made $10^{-2} * 4096 \approx 40$ times the wrong decision in the wavelet classification process. Since for $\epsilon = 10^{-2}$ a total of 102 significant wavelet coefficients was detected, see Fig. 6.8, the number of erroneously classified wavelet coefficients is 40/102, that is 39%. If we apply the same reasoning to the other values of ϵ, we obtain 11%, 1.5%, 0.15% and 0.02% for $\epsilon = 10^{-3}$, 10^{-4}, 10^{-5} and 10^{-6} respectively.

At this point, one might conclude that ϵ should always be taken equal to the smallest possible value, since it garanties the minimum number of erroneous classifications. The problem is that the smaller the value of ϵ is, the higher the risk one runs to eliminate truly significant wavelet coefficients.

Conclusion: values of ϵ between 10^{-4} and 10^{-5} are a good compromise between eliminating the artifacts and keeping the truly significant wavelet coefficients.



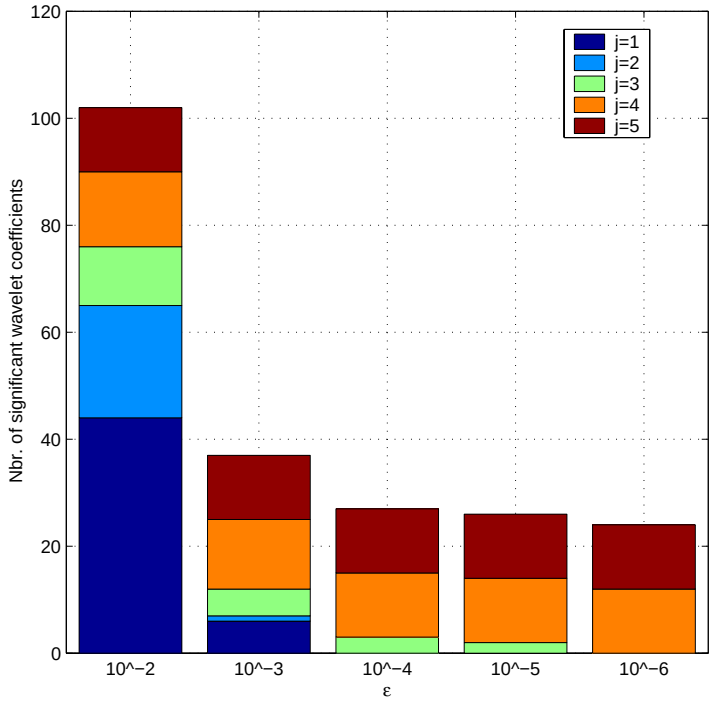


Figure 6.8.: ΔSNR (for definition see Eq. (6.2)) and number of significant coefficients of the restored images of Fig. 6.7

Choice of δ

The parameter δ is a tolerance parameter that has been introduced to take into account the noise that affects the significant wavelet coefficients. As we have seen in Chapter 4 Section 4.1.4, the calculated wavelet coefficient $W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k_s})$ is compared to the wavelet coefficient of the data $W_I(2^j, \mathbf{k_s})$ for all locations $\mathbf{k_s}$ where this latter has been found to be significant. If $W_{F_Y^{(n)}(2^{j-1})}(\mathbf{k_s})$ is in the so called regularization interval $[-\delta t(2^j, \mathbf{k_s}\ ;\ \delta t(2^j, \mathbf{k_s})]$ centered around $W_I(2^j, \mathbf{k_s})$, the calculated value is kept. Otherwise, it is changed to $W_I(2^j, \mathbf{k_s}) \pm \delta t(2^j, \mathbf{k_s})$ depending on the position of the calculated value with respect to the interval. $t(2^j, \mathbf{k_s})$ is the threshold that separates between significant and non significant wavelet coefficients. The following experiences give some insights in the choice of δ.

We restore the noisy image of RA=4*NL and fwhm=15 introduced in Fig. 6.5, using D-DeQuant with the following parameters: $J = 5$, $\epsilon = 10^{-2}$ and 10^{-5}, and

$$\delta \in [0,\ 0.5,\ 1,\ 1.5,\ 2,\ 3,\ 4,\ 5]$$

The restored images can be seen in Fig. 6.10 and 6.9. The ΔSNR curves are given by Figs. 6.13 and 6.11. The number of significant coefficients detected, into and out of the regularization interval during the restoration process are given by Figs. 6.14 and 6.12.

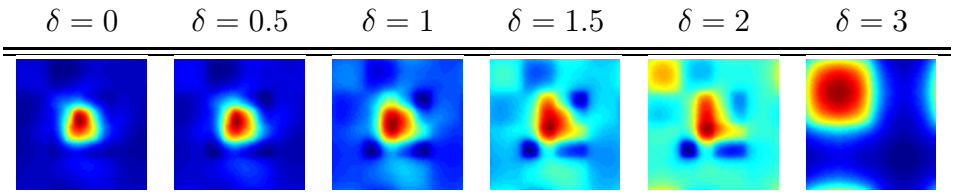


Figure 6.9.: Influence of the choice of δ on the restoration with D-DeQuant, when $\epsilon = 10^{-5}$. The processed noisy image is shown in Fig. 6.5. The nodule has a RA=4*NL and a fwhm=15.

Let us first have a look at the evolution of the restored images for $\epsilon = 10^{-5}$, see Fig. 6.9. We notice that except for the two first values of δ between which there is no perceptible difference, the visual appearance of the restored images worsens with the increasing value of δ. This fact is confirmed by the ΔSNR curve of Fig. 6.11. The key to this behavior is given by the location of the significant wavelet coefficient with respect to the regularization interval as shown in Fig. 6.12. For $\delta = 0$, the SWC of the data are not modified. For $\delta = 0.5$, all SWC lie out of the regularization interval. This implies that they will be modified

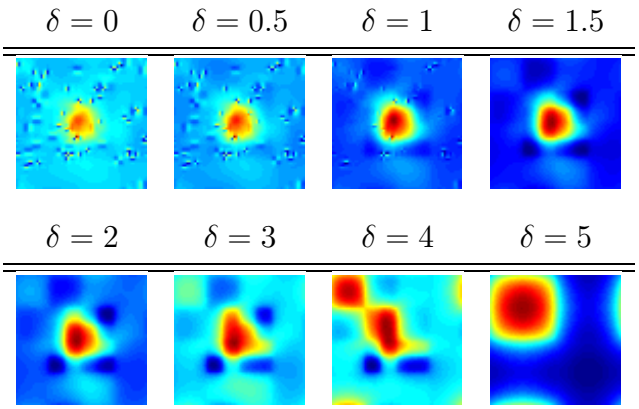


Figure 6.10.: Influence of the choice of δ on the restoration with D-DeQuant, when $\epsilon = 10^{-2}$. The processed noisy image is shown in Fig. 6.5. The nodule has a RA=4*NL and a fwhm=15.

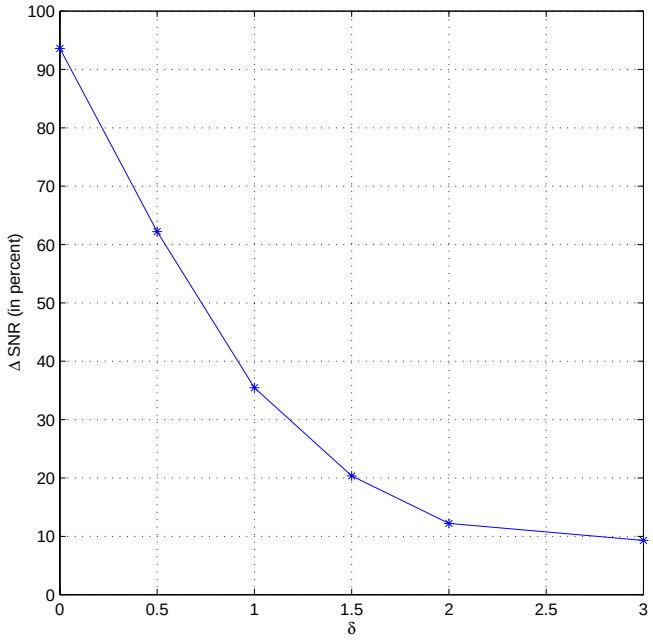


Figure 6.11.: ΔSNR of the restored images of Fig. 6.9.

by translating the wavelet coefficients of the data by an amount proportional to δ. Although this seems to decrease the SNR performance of the algorithm, it has no visual incidence on the restored image. But, as soon as the SWC are replaced by the calculated ones stemming from the Tikhonov constraint, the aspect of the restored image gets worse. This happens for $\delta \geq 1$ where the SWC begin

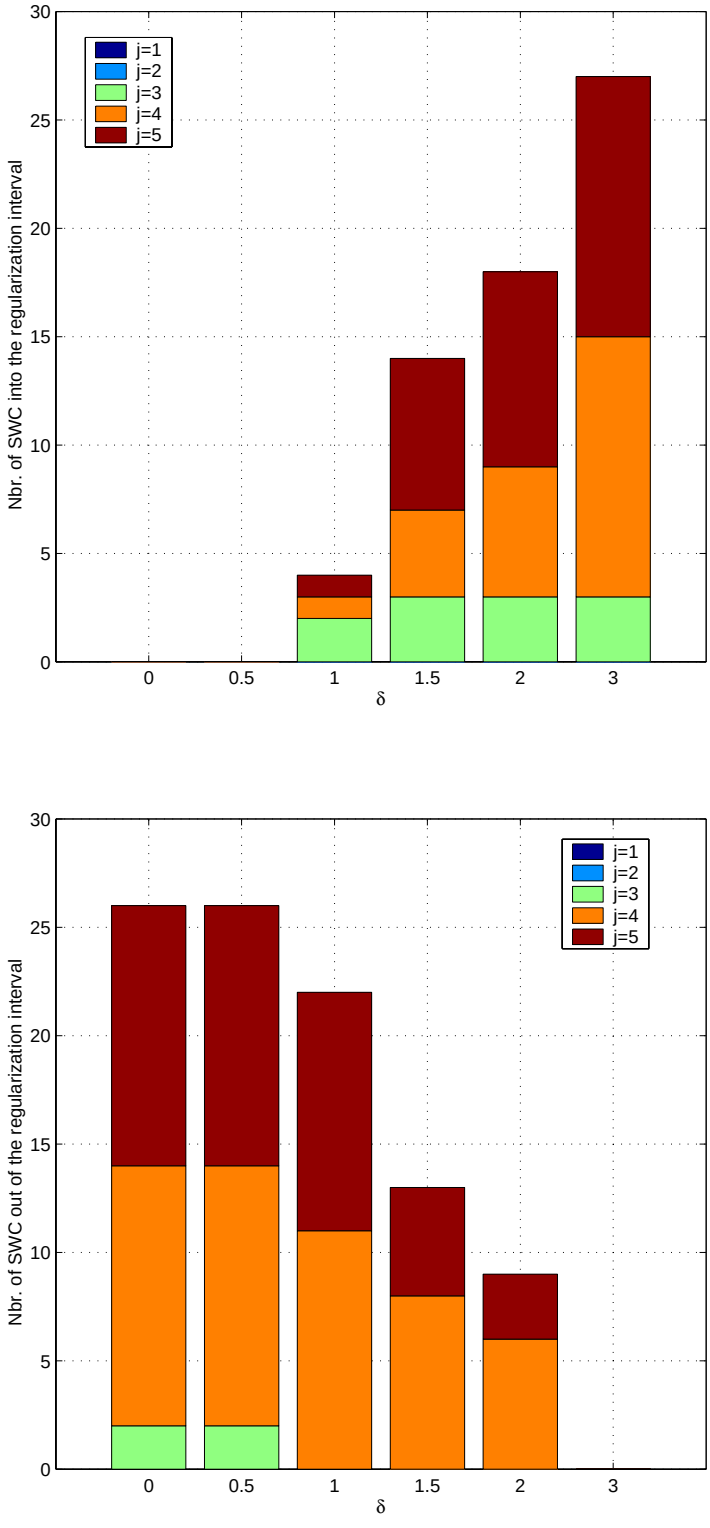


Figure 6.12.: Number of significant wavelet coefficients into and out of the regularization interval for the restored images of Fig. 6.9 where $\epsilon = 10^{-5}$.

progressively to come into the regularization interval, until becoming all into it for $\delta = 3$.

Before drawing a conclusion, let us have a look at the restoration results for $\epsilon = 10^{-2}$. We observe with growing values of δ an improvement followed by a deterioration of the aspect of the restored images. This fact is again confirmed by the ΔSNR curve of Fig. 6.13. If we analyze the location of the SWC given by Fig. 6.14, the following comes out. The improvement we noticed is due to the replacement of the SWC due to the artifacts that appear at scales j=1 and 2, by the values calculated from the Tikhonov contraint that tend to give a smooth image. For $\delta = 1.5$ all SWC values at scales j=1 and 2 has been changed. The deterioration coincides with the replacement of the SWC due to the nodule by the calculated ones. This happens mainly for $\delta > 1.5$, when these coefficients come into the regularization interval. Starting $\delta > 1.5$, we observe the same phenomena as for $\epsilon = 10^{-5}$. Notice the parallel between the images of Fig. 6.9 and those of 6.10 for $\delta > 1.5$.

Conclusion: for $\epsilon \geq 10^{-3}$ the number of SWC due to noise is small enough to choose δ in the range $[0, 0.5]$. For values of $\epsilon \leq 10^{-2}$, a choice of δ in $[1,\ 1.5]$ smoothes the artifacts that could not be eliminated due to the value of ϵ.

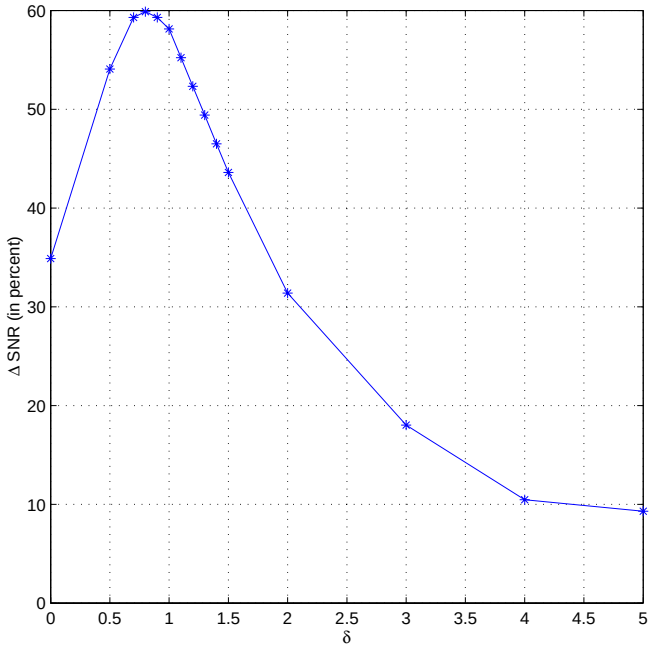


Figure 6.13.: ΔSNR of the restored images of Fig. 6.10.

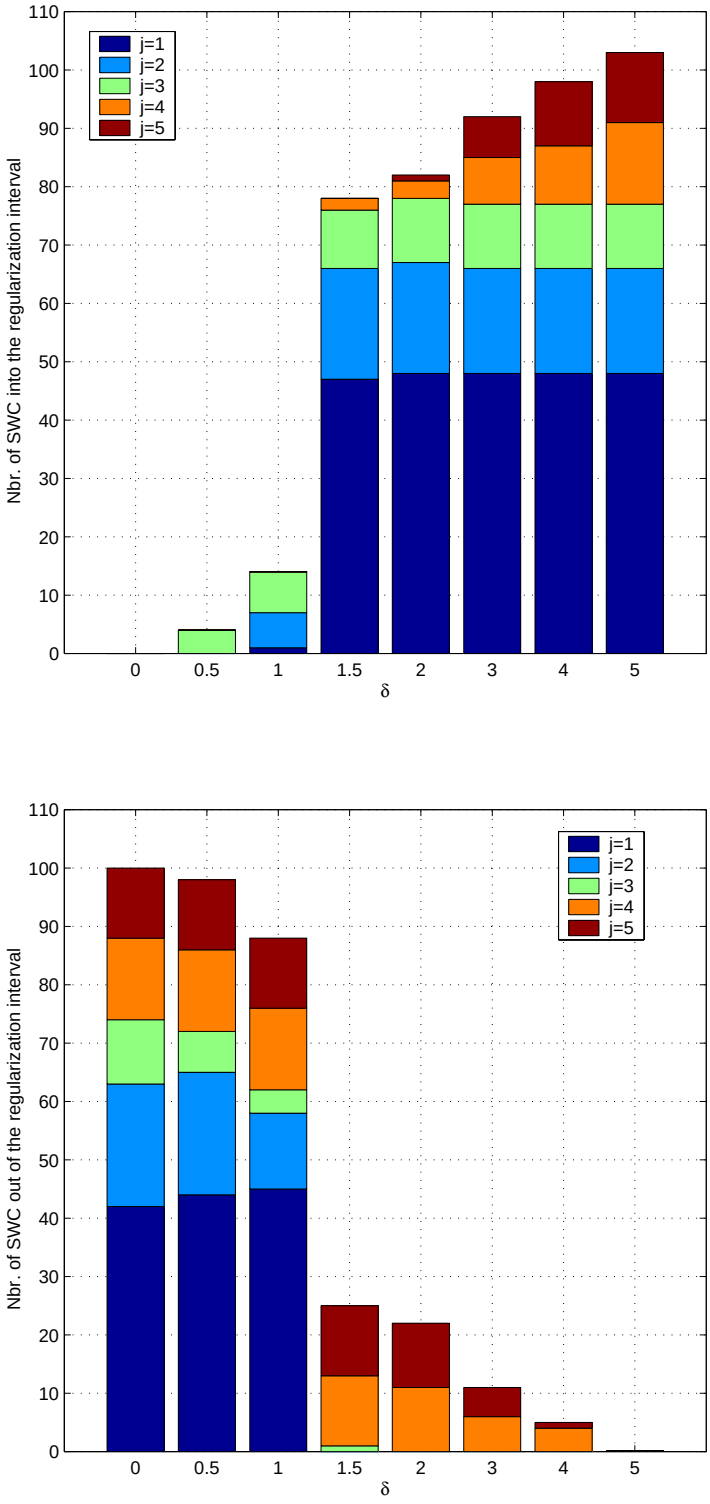


Figure 6.14.: Number of significant wavelet coefficients into and out of the regularization interval for the restored images of Fig. 6.10 where $\epsilon = 10^{-2}$.

6.3. Comparative studies on simulated data experiments

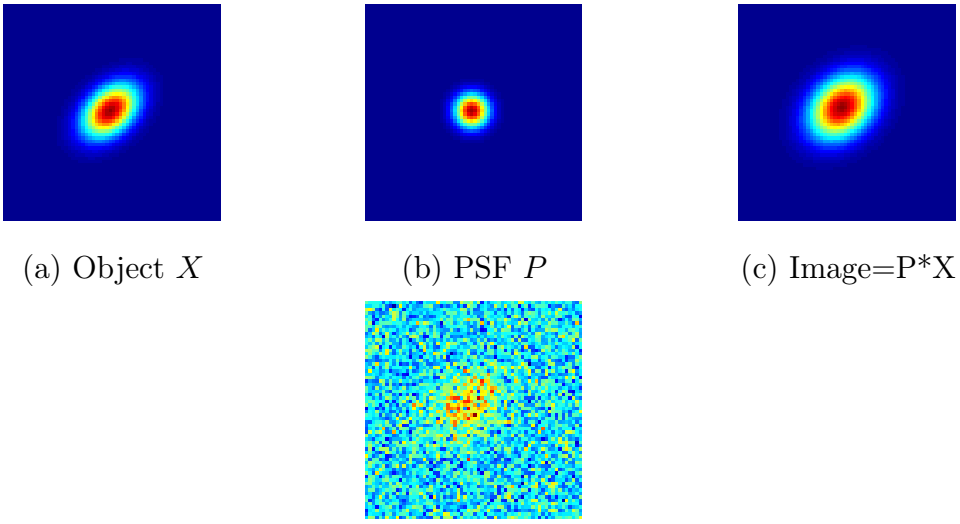
In this section we illustrate individually the effect of the introduction of the point spread function, the edge preserving regularization constraint and the shift invariant feature in the general framework of DeQuant. We then compare the performances of the eight DeQuant algorithms on a simulated data experiment.

6.3.1. Introduction of the point spread function

The following experiment compares the performance of D-Quant with that of R-DeQuant.

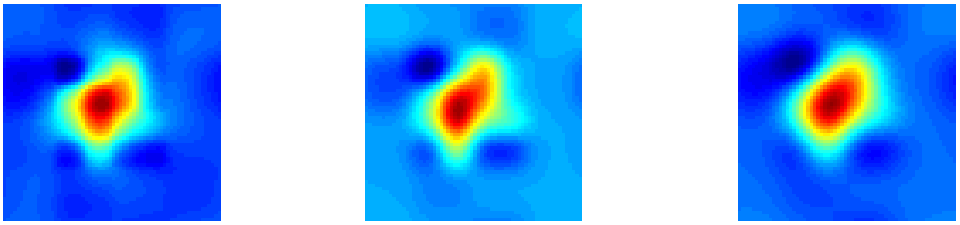
Experiment description: the simulated object of size 64x64, Fig. 6.15(a), consists of a background of constant intensity where a hot elliptical nodule is represented by a Gaussian intensity profile. The nodule which characteristics are in Region 3, has a RA of 4*NL and is rotated through an angle of 45 degrees about its center Fig. 6.15(c) represents the result of the convolution between X and the PSF with a Gaussian profile given by Fig. 6.15(b). The form the PSF was designed to be similar to those encountered in nuclear medicine imaging. Fig. 6.15(d) shows the noisy image obtained from Y by the introduction of Poisson noise. This latter, which represents the raw image that would be observed at the output of the camera without any processing, has been processed with D-DeQuant and R-DeQuant with the parameters $J = 5$, $\epsilon = 10^{-5}$ and $\delta = 0$. Three to four iterations were necessary to obtain the results given by Figs. 6.15(e) and 6.15(f).

Results interpretation: the qualitative visual comparison of D-DeQuant, Fig. 6.15(e), and the convolution of the PSF with the result of R-DeQuant, Fig. 6.15(g) shows that the form of the nodule in P*R-DeQuant is closer to the image PX as to the result of D-DeQuant. However the introduction of the PSF did not lead in this example to an improvement in the SNR value: indeed we obtained SNR(P*X, D-DeQuant)=31.4dB and SNR(P*X, P*R-DeQuant)=29.0dB, while SNR(P*X, Noisy Image)=14.9dB. A compression ratio of 216 was reached in this example (15 significant wavelet coefficients).



(a) Object X (b) PSF P (c) Image=P*X

(d) Noisy image = Poisson(P*X)



(e) D-DeQuant (f) R-DeQuant (g) P*R-DeQuant

Figure 6.15.: Comparing denoising and restoration with DeQuant for an image which charateristics are in Region 3. Fig. (e) is the result of the application of D-DeQuant on Fig (d). Fig. (g) is obtained by convoluting the result R-DeQuant, depicted in Fig (f), with the PSF represented in Fig. (b).

The calculated local criteria for the images D-DeQuant and P*R-DeQuant are represented Fig. 6.16. Notice that the numbers in this figure correspond to the difference to the original value expressed in percent, for example:

$$\Delta\text{flux}(\%) = \frac{\text{flux(nodule in restored im.) - flux(nodule in original im.)}}{\text{flux(nodule in original im.)}} * 100 \tag{6.3}$$

Thus the smaller the absolute value of the obtained number for a given parameter, the better the algorithm performs regarding this parameter. Examining Fig. 6.16 we notice that P*R-DeQuant performed better than D-DeQuant, for the Flux, area, perimeter and fwhm$_x$. It was as good as D-DeQuant for restoring the angular deviation theta. For the the position of the center of gravity, the RA

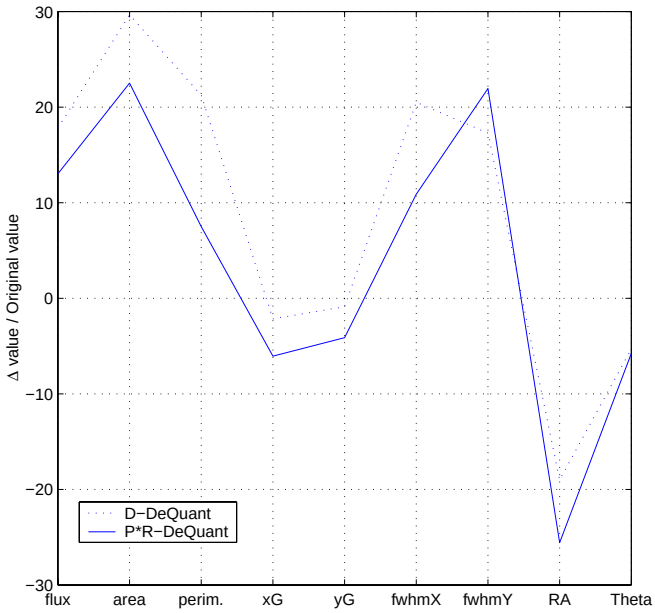


Figure 6.16.: Local performances of D-DeQuant and P*R-DeQuant. The numbers on the y-axis correspond to the difference to the original value expressed in percent of the original value, see Eq. (6.3).

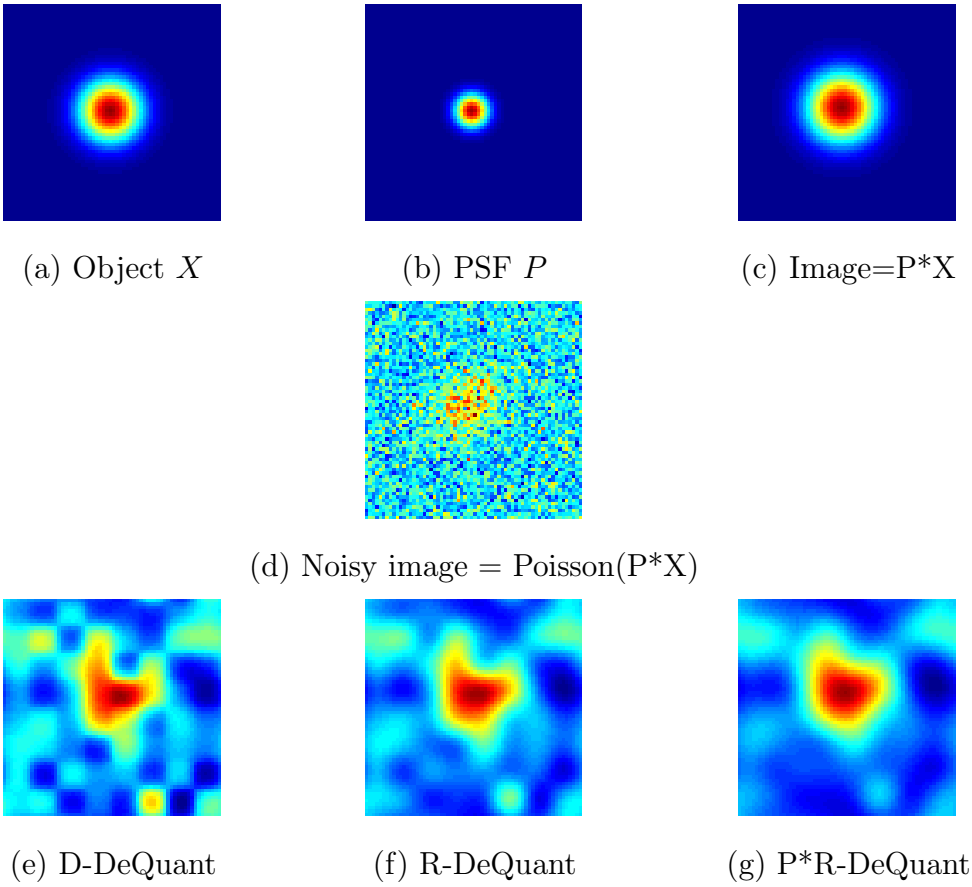
and the fwhm_y, D-DeQuant performed better. The small translation of position introduced by P*R-DeQuant is due to the convolution with the PSF.

The evolution of the SNR is not the same for an image which characteristics are under the identification curve. This is demonstrated by the following example. We consider an image in Region 1 and restore it with both algorithms using the parameters $J = 3$, $\epsilon = 10^{-5}$ and $\delta = 0$, see Fig. 6.17. We notice now that the block effects that can be seen in the result of the restoration with D-DeQuant have been dramatically reduced in P*R-DeQuant, thus improving the detectability of the nodule. For this example we have: SNR(P*X,D-DeQuant)=32.8dB and SNR(P*X,P*R-DeQuant)=37.0dB, while SNR(P*X,Noisy Image)=14.7dB.

6.3.2. Introduction of the edge preserving feature

The following experiment illustrates the effect of the introduction of the edge preserving operator in D-DeQuant.

Experiment description: we consider a 64x64 image consisting of a constant background where a two dimensional step edge is represented by a square intensity



(a) Object X (b) PSF P (c) Image=P*X

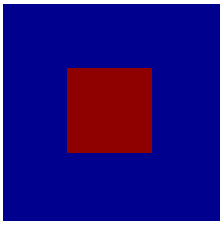
(d) Noisy image = Poisson(P*X)

(e) D-DeQuant (f) R-DeQuant (g) P*R-DeQuant

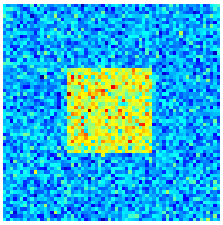
Figure 6.17.: Comparing denoising and restoration with DeQuant for an image which charateristics are in Region 1. Fig. (e) is the result of the application of D-DeQuant on Fig (d). Fig. (g) is obtained by convoluting the result R-DeQuant, depicted in Fig (f), with the PSF represented in Fig. (b).

profile, see Fig. 6.18(a). The square profile has a side length of 25 pixels and its RA=3.6*NL, which is in Region 3 . Fig. 6.18(b) shows the noisy image obtained from Fig. 6.18(a) by the introduction of Poisson noise. This noisy image has been first processed using D-DeQuant, the results are given by Figs. 6.18(c) and 6.18(d). We then processed the noisy image using EP-D-DeQuant, the results are given by Figs. 6.18(e) and 6.18(f). The parameters of the algorithms have been chosen as follows: $J = 5$, $\epsilon = 10^{-5}$ and $\delta = 0.5$. The algorithms converged after 5 iterations.

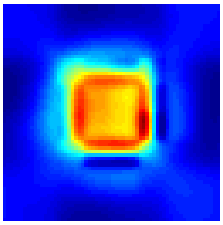
Results interpretation: comparing Figs. 6.18(c) and 6.18(d) with 6.18(a) we can clearly see that while transition regions in the image restored with D-DeQuant are wider than those in the original image, EP-D-DeQuant has preserved sharp



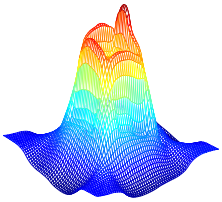
(a) Object X



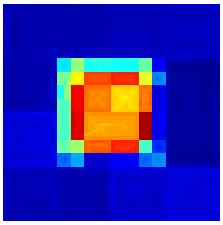
(b) Noisy object = Poisson(X)



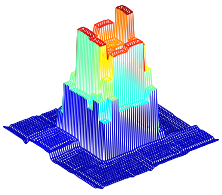
(c) D-DeQuant 2D



(d) D-DeQuant 3D



(e) EP-D-DeQuant 2D



(f) EP-D-DeQuant 3D

Figure 6.18.: Comparing denoising with and without the edge preserving feature. Fig. (c) and its 3D counterpart are the result of the application of D-DeQuant on Fig. (b). Similarly Figs. (e) and (f) are obtained by applying EP-D-Dequant Fig. (b).

discontinuities. The result given by D-DeQuant is too smooth and the one given by EP-D-DeQuant is too blocky.

The SNR are almost the same for both algorithms. We obtained SNR(X,D-DeQuant)=20.2dB and SNR(X,EP-D-DeQuant)=20.1dB, while SNR(P*X,Noisy Image)=15.3dB. A compression ratio of 95 was reached in this example (39 significant wavelet coefficients).

The local performances of D-DeQuant and EP-D-DeQuant are almost the same as can be seen in Fig. 6.19. We notice that the flux, area and RA are not very faithful to the original data.

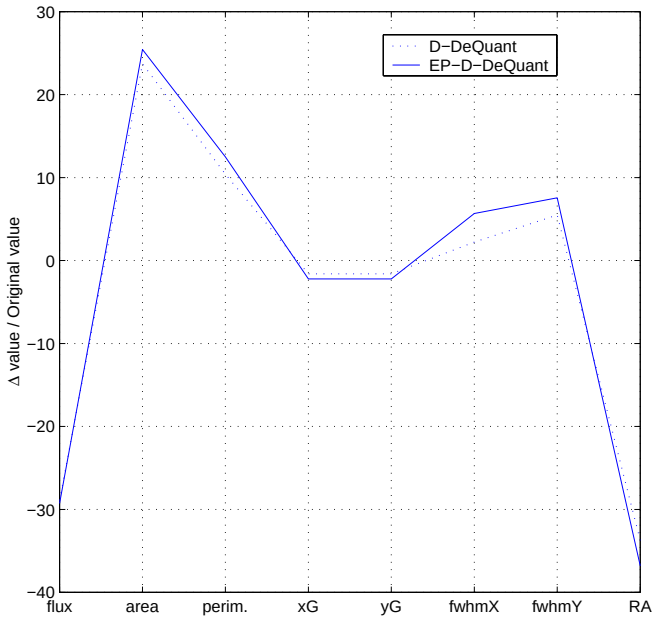


Figure 6.19.: Local performances of D-DeQuant and EP-D-DeQuant. The numbers on the y-axis correspond to the difference to the original value expressed in percent, see Eq. (6.3).

6.3.3. Introduction of the shift invariant feature

The following experiment illustrates the effect of the introduction of the shift invariance feature in D-DeQuant.

Experiment description: we processed the three noisy images of Fig. 6.5 with D-DeQuant and SI-D-DeQuant. Recall that these images consist of a background of constant intensity where one hot nodule of RA=0.9*NL, RA=2*NL and RA=4*NL respectively is represented by a Gaussian intensity profile. The restored images with both algorithms are given by Fig. 6.20. The parameters of the algorithms have been chosen as indicated Table 6.1. The algorithms needed 3 to 6 iterations to converge.

Results interpretation: we notice first that the blocking effects that appear in the restoration with D-DeQuant of the nodule of RA=0.9*NL are considerably reduced by SI-D-DeQuant, see Fig. 6.20 leading to a much better visual appearance. Moreover for all values of RA, the form of the nodule seems to be much more faithful to the original in the image restored with SI-D-DeQuant as in the one restored with D-DeQuant.

The numerical performance of the two algorithms in terms of SNR is summarized in Table 6.1. The smaller RA, the larger is the relative improvement in the value of the SNR between D-DeQuant and SI-D-DeQuant. It ranges from $(35.7 - 32.3)/15 = 22.7\%$ for RA=0.9*NL down to $(29.1 - 27.9)/15.3 = 7.8\%$ for RA=4*NL. The Compression ratios and the number of significant coefficients are given by Table 6.1.

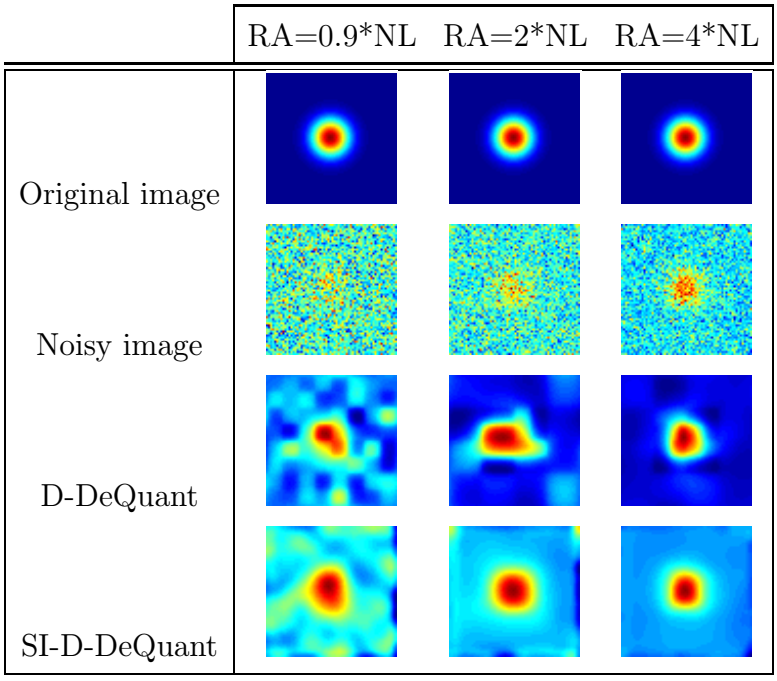


Figure 6.20.: Comparing D-DeQuant and SI-D-DeQuant for three values of the RA.

The local performances of the algorithms are given by Fig. 6.21. For RA=0.9*NL, we notice that SI-D-DeQuant performs better than or as good as D-DeQuant for the area, perimeter, position of the center of gravity and fwhm_y. The flux, fwhm_x and RA are better restored by D-DeQuant. For RA=2*NL and RA=4*NL, all parameters are better restored by SI-D-DeQuant. Notice that for all values of RA, the restored parameters flux and RA are far away from their original value.

6.3.4. Comparison of the eight different DeQuant algorithms

The following experiment compares the performances of the eight DeQuant algorithms based on a simulated image.

	RA=0.9*NL	RA=2*NL	RA=4*NL
J	3	5	5
ϵ	10^{-5}	10^{-5}	10^{-5}
δ	0.5	0.5	0.5
Number of SWC	0	10	26
CR	64	293	137
SNR(X,Noisy image)	15.0	14.8	15.3
SNR(X,D-DeQuant)	32.3	28.1	27.9
SNR(X,SI-D-DeQuant)	35.7	31.3	29.1

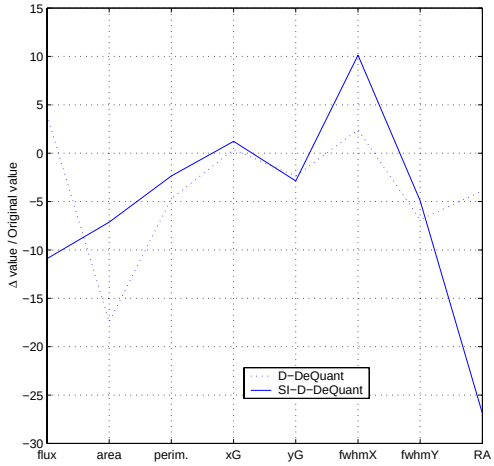
Table 6.1.: Parameters and global performances of D-DeQuant and SI-D-DeQuant for the images of Fig. 6.20

Experiment description: The simulated intensity image of size 64x64 consists of a background of constant intensity with three nodules. Each of these nodules is in a region of the detection plane and exhibit a particular intensity profile, see Fig. 6.22(a) and Table 6.2.

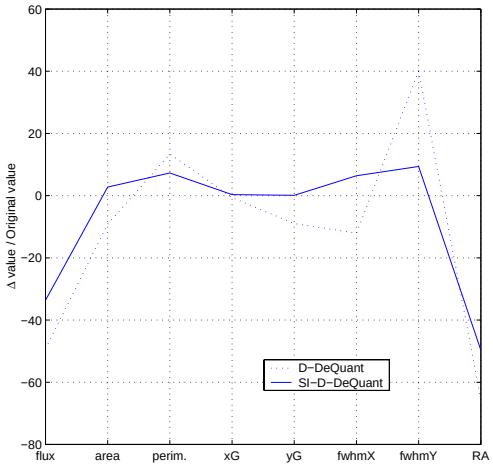
	Nodule 1	Nodule 2	Nodule 3
intensity profile	Gaussian	cylindrical	cubical
type	cold	hot	hot
Region	Region 1	Region 2	Region 3
fwhm(in pixels)	8	14	18
RA	-2*NL	2.3*NL	3.5* NL

Table 6.2.: Characteristics of the nodules of the image of Fig. 6.22

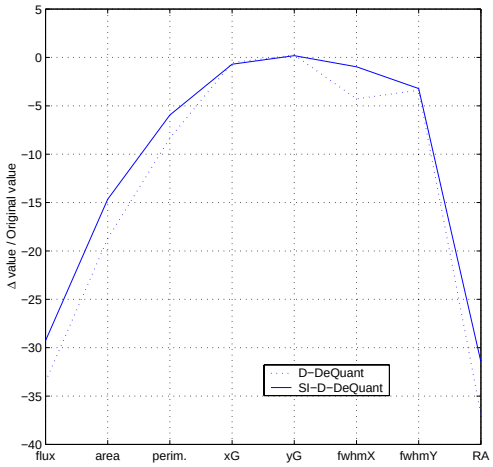
The point spread function P associated with the image has been assumed to have a Gaussian intensity profile. As can be seen in Fig. 6.22(b), the statistical fluctuation in the noisy image handicaps the recognition of form and size of the nodules, particularly for the cold nodule which seems to be lost in the middle of the noise. Figs. 6.22 shows the results of the processing of the noisy image with the eight different DeQuant algorithms. The algorithms' parameters have



(a) $RA = 0.9 * NL$



(b) $RA = 2 * NL$



$RA = 4 * NL$

Figure 6.21.: Local performances of D-DeQuant and SI-D-DeQuant. The numbers on the y-axis correspond to the difference to the original value expressed in percent of the original value, see Eq. (6.3)

been chosen as follows: $J = 4$, $\epsilon = 10^{-5}$ and $\delta = 0.3$. Notice that the value of J has been taken equal to 4 because one of the nodules is under the identification curve, and could disappear in the restoration if the scale to which the wavelet transform is calculated is too high, see Section 6.2.2. Five to six iterations were necessary to obtain the final estimates.

Results interpretation: Qualitatively, we can say that the images denoised or restored with the shift invariance feature, see Figs. 6.22(g), (h), (i) and (j), look

better than the others. The EP-D-DeQuant algorithm exhibits block effects, see Fig. 6.22(e). D-DeQuant, P*R-DeQuant and EP-R-DeQuant seem to distort the nodules, see Figs. 6.22(c), (d), and (f).

The calculated ΔSNR given in Fig. 6.23 confirm the visual impression. It shows an oscillating curve with a general upward trend. The falls coincide with the occurence of EP-D-DeQuant, P-R-DeQuant and EP-R-DeQuant. A compression ratio of 108 was reached in this example (22 significant wavelet coefficients).

In order to facilitate the comparison of the algorithms, we brought some of the local performances together. We thus define, three new performance parameters which are: the variation in the position of the center of gravity

$$\Delta\text{Position (in \%)} = \sqrt{\frac{(x_G - x_{G_0})^2 + (y_G - y_{G_0})^2}{x_{G_0}^2 + y_{G_0}^2}} * 100, \tag{6.4}$$

the distortion

$$\text{Distortion (in \%)} = \frac{\left(\frac{\text{fwhm}_x}{\text{fwhm}_y}\right) - \left(\frac{\text{fwhm}_{x_0}}{\text{fwhm}_{y_0}}\right)}{\left(\frac{\text{fwhm}_{x_0}}{\text{fwhm}_{y_0}}\right)} * 100, \tag{6.5}$$

and the variation in the flux per unit area

$$\Delta\text{Flux-per-Unit-Area (in \%)} = \frac{\frac{\text{flux}}{\text{area}} - \frac{\text{flux}_0}{\text{area}_0}}{\frac{\text{flux}_0}{\text{area}_0}} * 100. \tag{6.6}$$

The subscript 0 refers to the initial value of the parameters. The previously defined local performances are given together with ΔArea and ΔAmplitude by Figs. 6.24, 6.25, 6.26, 6.27 and 6.28.

ΔPosition: we can globally say that the bigger the RA is, the better the restored position of the center of gravity will be. While the introduction of the PSF results in an additional translation, the SI feature seems to improve the restored position.

ΔArea and Distortion: it is interesting to observe those parameters together, in order to be able to see whether a deformation of the nodule is associated with a variation in its area or not. The positivity (respectively negativity) of the distortion parameter shows a deformation of the nodule in the vertical (respectively horizontal) direction. Figs. 6.26 and 6.27 lead to the following observations. The cubical nodule is not distorted by the restoration and denoising processes but its area decreases on average of about 15%.

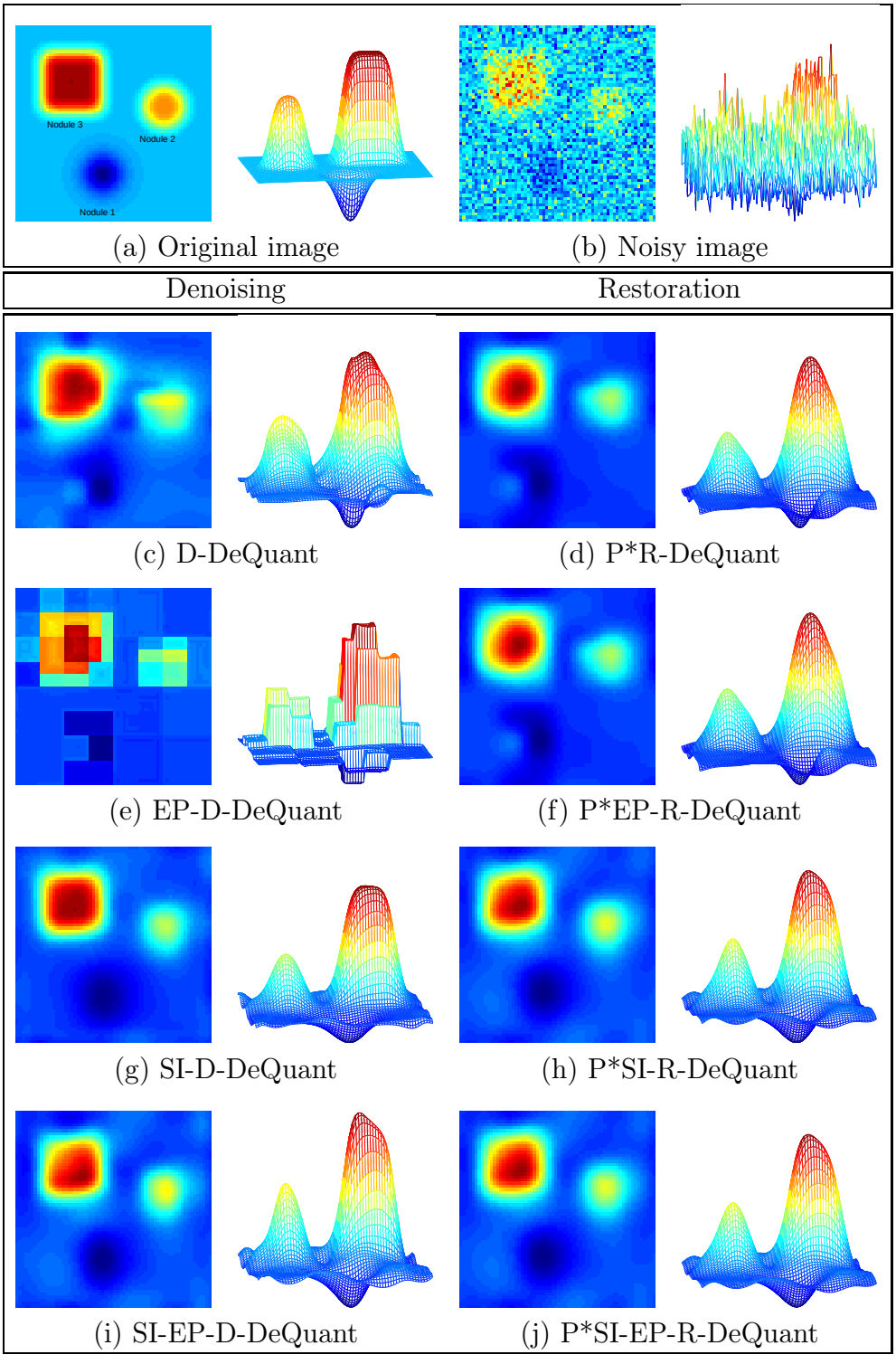


Figure 6.22.: Denoising and Restoration of the noisy image of Fig. (b) with the eight DeQuant algorithms. Fig. (c), (e), (g) and (i) are obtained with the denoising algorithms. Fig. (d), (f), (h) and (j) are obtained by convoluting the result of the restoration algorithms with the simulated point spread function P.

The cylindrical nodule is mostly distorted by the denoising algorithms. We notice a horizontal deformation for D-DeQuant and EP-D-DeQuant and a vertical distortion for SI-D-DeQuant and SI-EP-D-DeQuant together with an average decrease of 10% of the area. The Gaussian nodule is vertically distorted with decrease or increase of the area. The deformation is particularly strong for D-DeQuant and P*R-DeQuant. Globally, the bigger the RA is, the smaller is the distortion of the restored nodule.

ΔFlux-per-Unit-Area: this criteria is globally improved by the introduction of the PSF.

ΔAmplitude: this parameter is better restored when the RA is big enough with respect to the NL. EP-D-DeQuant gives particularly bad results with respect to the amplitude.

Although the number of local characteristics have been reduced from eight to five, it is still not an easy task to analyze individually all curves for the three types of nodules. It therefore might be interesting to average these curves to a unique one which will give an additional indication on the "average local performance" of the algorithms. We define the average local performance as follows:

$$\text{average local performance} = \frac{1}{8*3}\sum_{k=1}^{8}\sum_{l=1}^{3}|\Delta\text{parameter}(k,l)| \tag{6.7}$$

where k enables to take the sum over the eight local parameters: flux, RA, area, perimeter, x_G, y_G, fwhm_x and fwhm_y, and l takes into account the three different intensity profiles: the Gaussian, the cylindrical and the cubical ones. Examining the curve of Fig. 6.29, we can see that the SI-R-DeQuant and the SI-EP-R-algorithms have the better average local performances.

Conclusion: which algorithm should be used to obtain optimal restoration results for an image exhibiting nodules with different RA with respect to the NL and possibly different intensity profiles? the ΔSNR curve, the previous observations on the individual local performances and the average local performance curve plead for the algorithms that incorporates the PSF and the SI feature, namely SI-R-DeQuant and SI-EP-R-DeQuant. In the cases where the PSF is not known, SI-EP-D-DeQuant gives on the average good results.

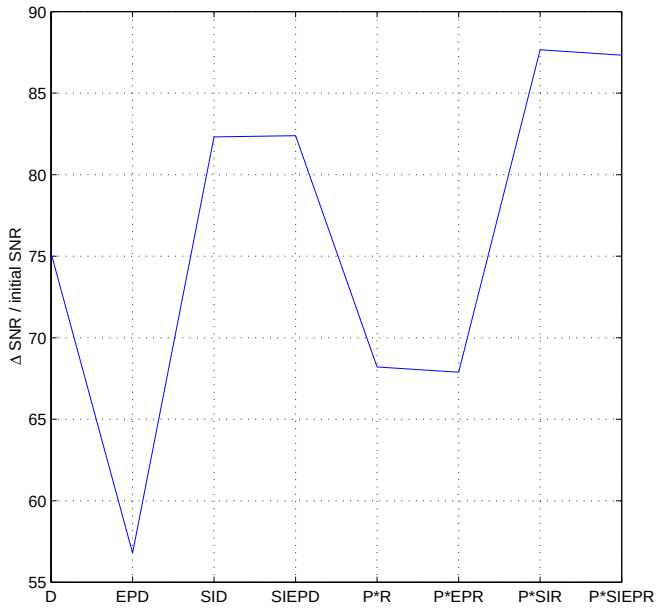


Figure 6.23.: Signal to noise ratio for the eight DeQuant algorithms applied on the example of Fig. 6.22. The initial SNR value is SNR(Original image, Noisy image)=15.7dB. The numbers on the y-axis correspond to the difference to the original value, expressed in percent, see Eq. (6.3).

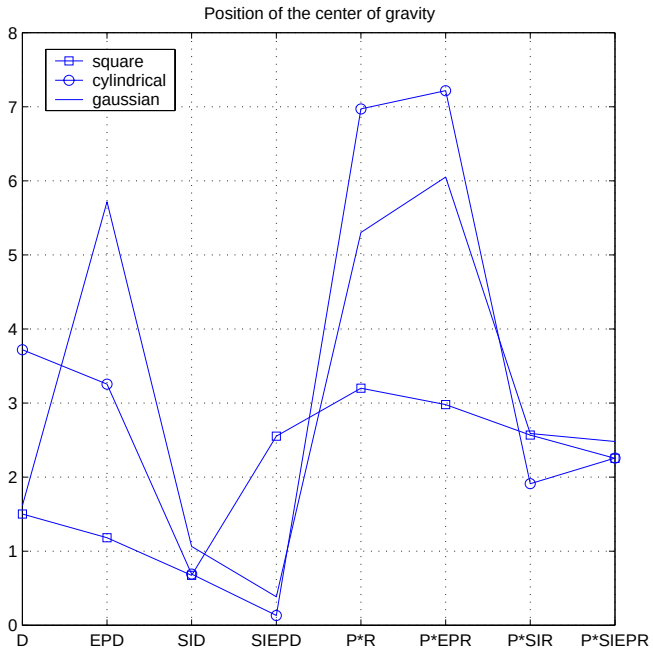


Figure 6.24.: Variation in the position of the center of gravity, see Eq. (6.4).

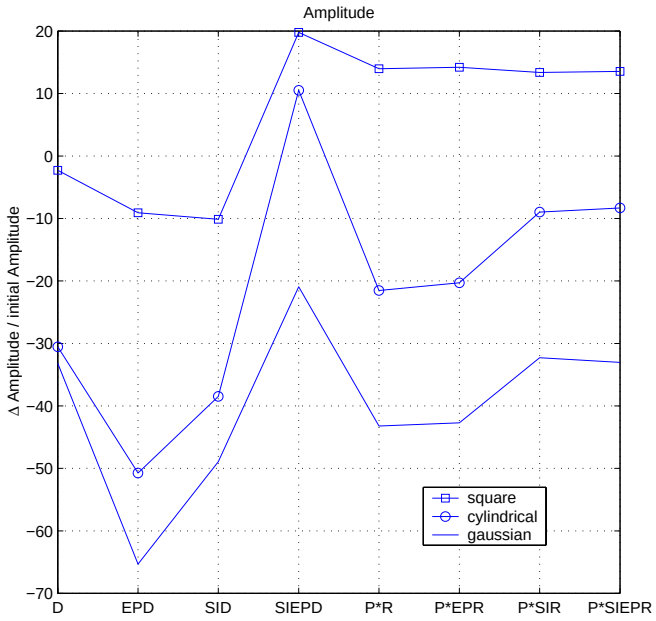


Figure 6.25.: Variation in the amplitude of the restored nodule.

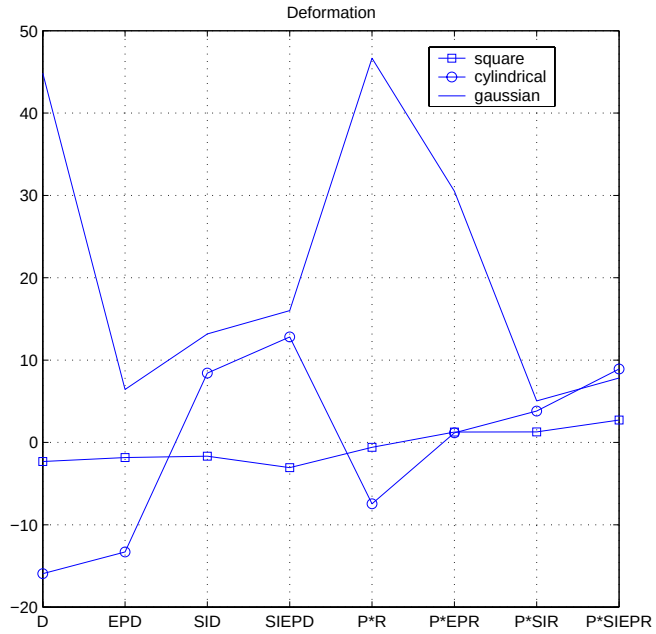


Figure 6.26.: Distortion of the restored nodule, see Eq. (6.5).

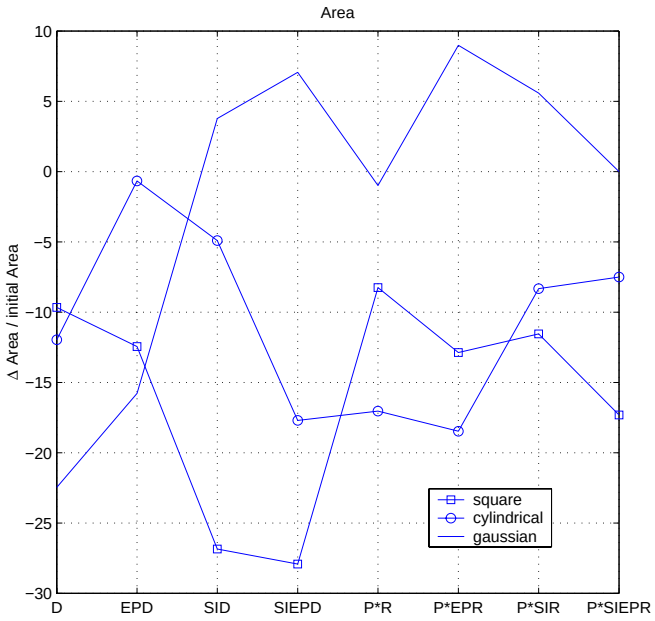


Figure 6.27.: Variation in the area of the nodule.

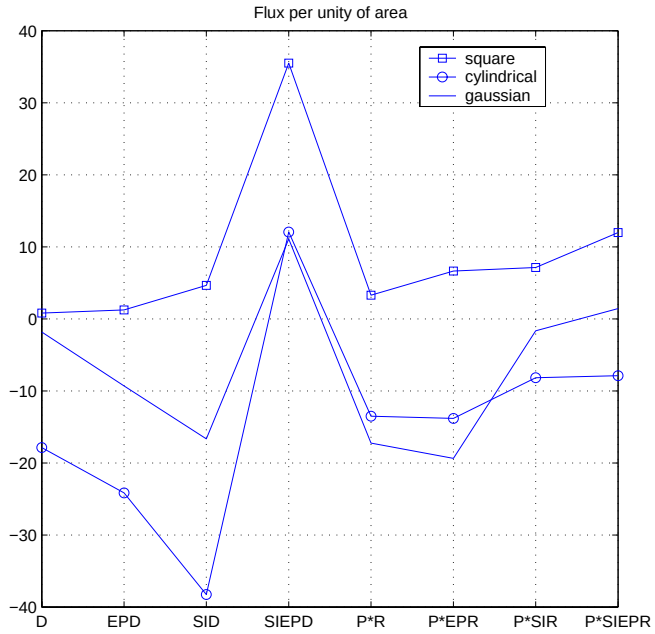


Figure 6.28.: Variation in the flux per unit area, see Eq. (6.6).

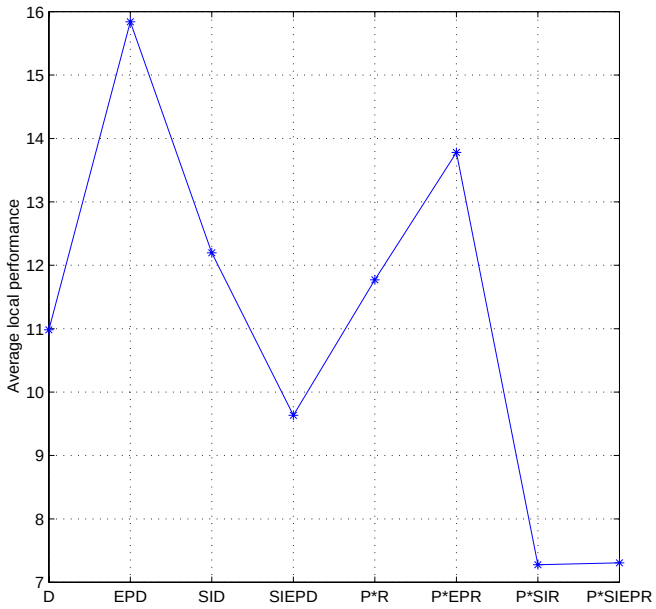


Figure 6.29.: Average local performance of the eight DeQuant algorithms, see Eq. 6.7.

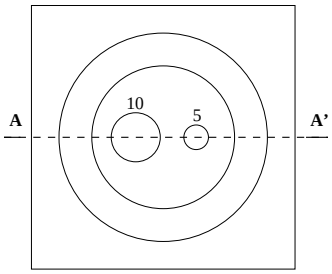
6.4. Thyroid phantom studies

After investigating the performances of the DeQuant algorithms on simulated data experiments, we examine their behavior on actual images of thyroid phantoms.

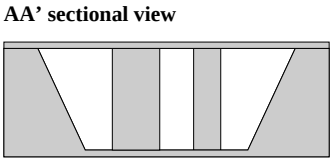
A thyroid phantom is a hollow object generally constructed of plexiglas that permits precise duplication of clinical conditions. The phantom can be filled with a radioactive solution and includes artifacts typical of abnormal thyroids: hot and cold nodules and regions of varied activity. Cold nodules are simulated by adding to the phantom massive plexiglas objects, hot nodules are hollow cavities within the phantom, the activity is modulated by varying phantom's depth. An image of the filled phantom is taken by the gamma camera. Physical phantoms permit extensive testing in a real-world scanning environment since they provide perfectly known reference data.

Experiment description: We present in the following the result of the restoration of three phantoms with SI-EP-R-DeQuant. The parameters of the algorithm are $J = 4$, $\delta = 0$ and $\epsilon = 10^{-4}$. Five to eight iterations were necessary to obtain the final estimates.

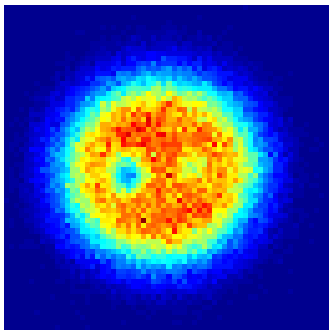
- Phantom 1: the first phantom is conically shaped and contains two cylindrical cold nodules of size 10 mm and 5 mm, see Fig. 6.30(a) and (b). The raw image has a size of 64x64.
- Phantom 2: the second phantom has approximately the butterfly form of a thyroid, see Fig. 6.31(a). The lobes have different depths so that the intensity on the right side is 50% of the left side, see Fig. 6.31(b). The hot lesion (bottom right) has a 11.5 mm diameter, and the cold lesions have 11.5, 8.5, and 5.5 mm diameters. The raw image has a size of 128x128.
- Phantom 3: the third phantom is also butterfly shaped. It contains cold nodules having 11.5, 8.5, 7, 6, 5, 4 and 3 mm diameters, see Fig. 6.32(a) and (b). The image of this phantom has been intentionally degraded by placing it in a cylinder full of water in order to increase the proportion of scattered photons. The raw image has a size of 256x256.



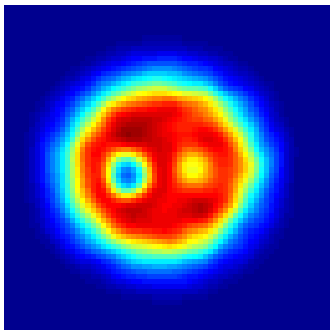
(a) top view of phantom 1



(b) sectional view of phantom 1



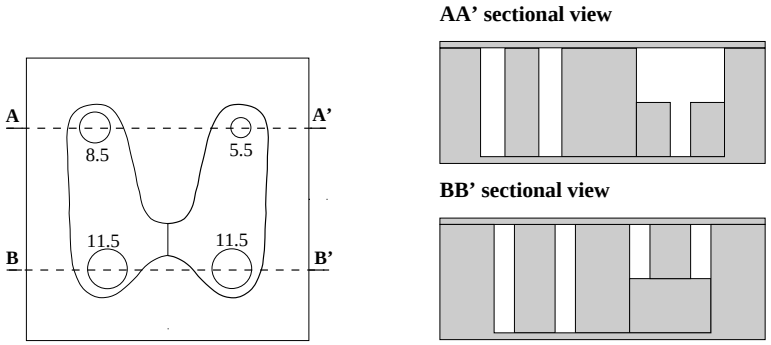
(c) raw image of phantom 1



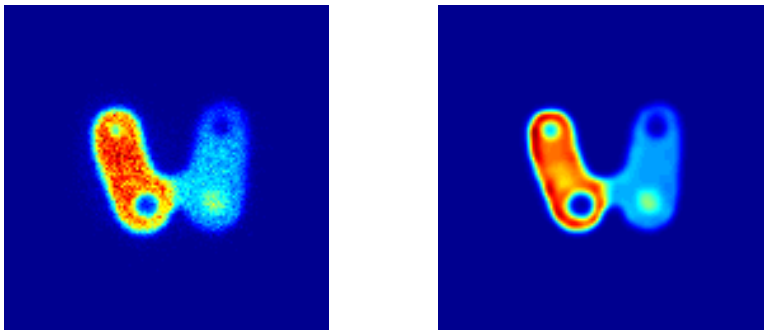
(d) SI-EP-R-DeQuant

Figure 6.30.: Noisy image of Phantom 1 (c) and its restored counterpart (d). A top and cutaway view of or phantom 1 are given by Figs (a) and (b) respectively.

Results interpretation: The results of the restoration process are given by Figs. 6.30, 6.31 and 6.32. The algorithm did a good job of reducing the background



(a) top view of phantom 2 (b) sectional view of phantom 2



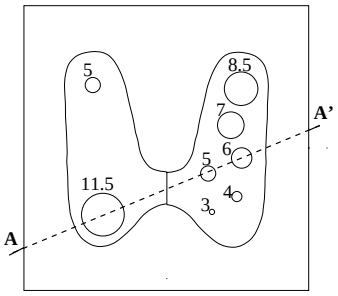
(c) raw image of phantom 2 (d) SI-EP-R-DeQuant

Figure 6.31.: Noisy image of phantom 2 and its restored counterpart. A top and cutaway view of or phantom 2 are given by Figs (a) and (b) respectively.

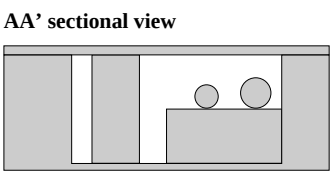
noise. The position-related and morphological information remained faithful to the original data. In the third phantom, the small nodules in the right lobe (bottom right) seem to be under the identification curve. Although they have not been clearly restored, we can guess their existence.

6.5. Actual thyroid clinical images

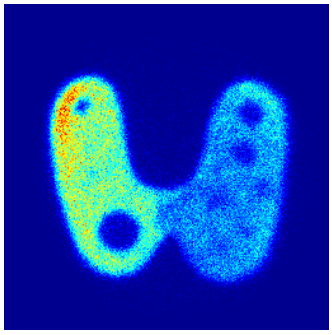
The last stage in our evaluation process is the validation of the algorithms using real data acquired in vivo. Since there is no direct information available for the imaged object to corroborate the results of the various restorations, our analysis is restricted to the interpretation of the restored images and the comparison of the different algorithms.



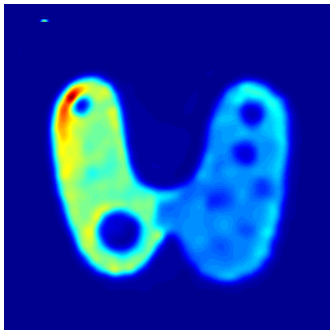
(a) top view of phantom 3



(b) sectional view of phantom 3



(c) raw image of phantom 3



(d) SI-EP-R-DeQuant

Figure 6.32.: Noisy image of phantom 3 (c) and its restored counterpart (d). A top and cutaway view of or phantom 3 are given by Figs (a) and (b) respectively.

The real noisy thyroid images are given by Figs. 6.33 and 6.34 together with the results of denoising and restoration with the DeQuant algorithms. These figures yield the following observations.

Globally, we notice that the restored images (with R-DeQuant and derivates) are smoother than the denoised images (with D-DeQuant and derivates). The restoration algorithms seem to give narrower and better delimited objects than the denoising ones. EP-D-DeQuant is obviously not adapted to the type of image encountered in nuclear medicine applications due to the blocky aspect of the denoised images.

The restored object with SI-EP-R-DeQuant, which is, based on the previous studies the closest to the truth, shows a significant decrease of activity in the bottom part of the left lobe for the thyroid of Fig. 6.33 and an underactivity of the bottom part of the right lobe for the thyroid of Fig. 6.34.

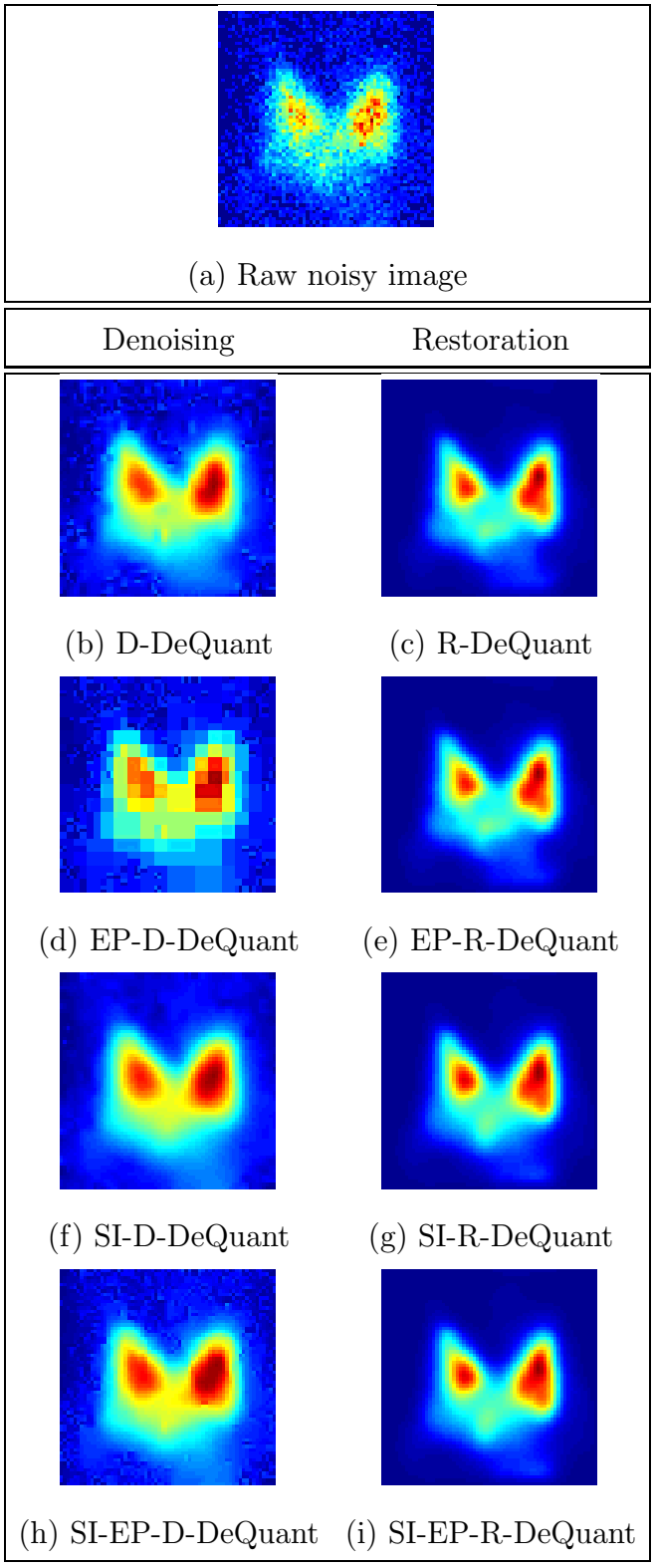


Figure 6.33.: Denoising and Restoration with the eight DeQuant algorithms. Fig. (b), (d), (f) and (h) are obtained with the denoising algorithms. Fig. (c), (e), (g) and (i) are obtained with the restoration algorithms.

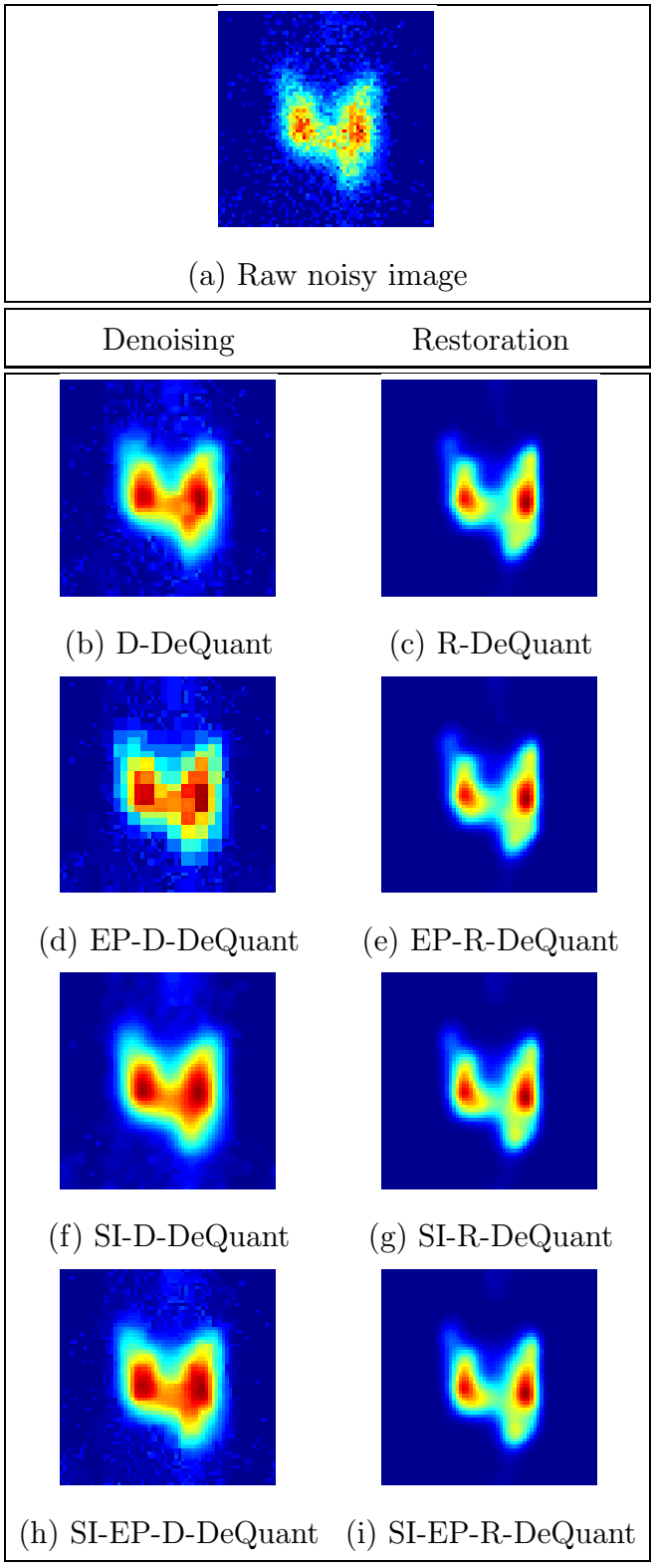


Figure 6.34.: Denoising and Restoration with the eight DeQuant algorithms. Fig. (b), (d), (f) and (h) are obtained with the denoising algorithms. Fig. (c), (e), (g) and (i) are obtained with the restoration algorithms.

6.6. Conclusion

We addressed in this chapter the problem of the validation of the DeQuant algorithms, concentrating thereby on thyroid studies. We started the validation process by experiments with simulated data that permit controlled evaluation over a wide range of conditions. We then tested the algorithms in a real-world scanning environment by processing actual images of thyroid phantoms. The last stage of the evaluation process was the validation of the algorithms using real data acquired in vivo.

The simulated experiments enabled to determine the contrast and size sensibility of the DeQuant algorithms and gain insight into the choice of the algorithm's parameters. They also allowed to compare the eight different DeQuant algorithms and analyze individually the effect of the choice of the wavelet transform (Mallat /à trous), the choice of the regularization operator (Laplacian / operator M) and the introduction of the PSF on the restoration process.

The processing of phantom images and real data confirmed the results of the investigations on simulated data experiments, namely that the estimates provided by the DeQuant algorithms would enable a greater diagnostic confidence in clinical nuclear medicine since they give the physician an easier access to the diagnosis-relevant information.

7. DeQuant's Graphical User Interface

A MATLAB-based Graphical User Interface (GUI) has been built for DeQuant [Eca99] [Bou00]. This GUI, which provides a means through which individuals can communicate with the computer without programming commands, increases the productivity of the user and provides an insight into the performances of DeQuant for people with little or no MATLAB programming experience [Mar96].

The DeQuant GUI is made up of five windows: the Title Window, the Nodule Generation Manager, the Psf Generation Manager, the Restoration Manager, and the Measure Lab. A comprehensive hypertext documentation that can be loaded in the web browser has also been written to help the user find its way through the GUI.

This chapter is organized as follows. Section 7.1 presents the general structure of the GUI's windows and introduces the new terminology that will be used in this chapter. Sections 7.2 to 7.6 describe the function and illustrate the appearance of each GUI window. Section 7.7 offers some conclusions.

7.1. General structure of the GUI windows

Each of the DeQuant GUI windows is based on the 3-parts structure shown in Fig. 7.1. Part 1 is devoted to the display of images and results of measures. Part 2 is the interactive part, in which the parameters of the different procedures can be set by the user. Part 3 insures the display management of the window.

The components of each window come in two classes: user interface controls (uicontrols) and user interface menus (uimenus). The uicontrols elements are created with the purpose of performing an action or setting up the options for a future action. The uimenus are pull-down menus residing in the menu bar at the top of a figure window.

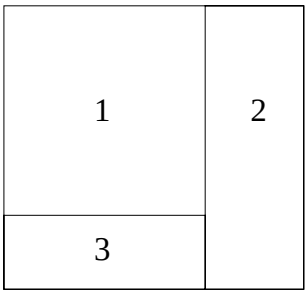


Figure 7.1.: General structure of the GUI windows

The uicontrols come in five styles.

The push button (PB): is used to indicate that a desired action should take place immediately.

The editable text (ET): is used in situations that require the user to enter strings of characters or numbers.

The radio button (RB): is a useful means of representing two states of an option you may want to provide a user.

The pop-up menu (PUP): is usually used in situations where multiple choices need to be available to the user.

Sliders (SLI): are useful in representing to a user that he or she has a fixed range of values to choose from.

It is important to become familiar with the introduced new terminology, so that the topics discussed in the following sections are clear and easy to follow.

7.2. Title Window

The title window is the entrance door to the functional GUI, see Fig. 7.2. It contains three push buttons whose functions are listed below.

Restoration (PB): opens the Restoration Manager which is the main functional window of the GUI.

Help (PB): opens the default internet browser of the computer and displays the hypertext help desk that has been written as DeQuant GUI support.

Close (PB): closes the title window and the application.

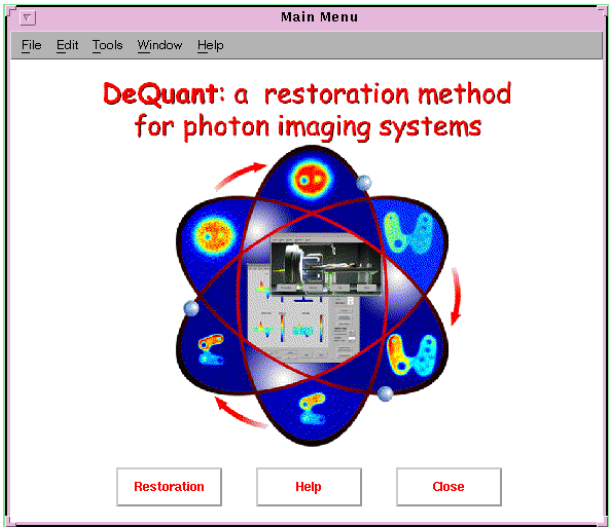


Figure 7.2.: Title window of the application

7.3. Nodule Generation Manager

The Nodule Generation Manager window, see Fig. 7.3, enables the user to create his own simulated uptake profiles. After setting the background intensity of the image, he can introduce nodules (which may have Gaussian, cylindrical or cubical profiles) either by clicking and dragging the mouse pointer in the axes or by setting their parameters' value in the appropriate editable box. The interface enables also the automatic generation of nodule grids that may be useful for contrast-detail studies. A detailed description of the window's functions is given below.

7.3.1. Display

Area where the user can draw the nodules. This may be done by clicking the position of the center of gravity of the object and dragging the mouse pointer until the desired size is reached. The user can also define the nodules by setting their parameters' value in the corresponding editable text uicontrols.

7.3.2. Uicontrols

Image properties

Image size (ET): size of the image in pixels.

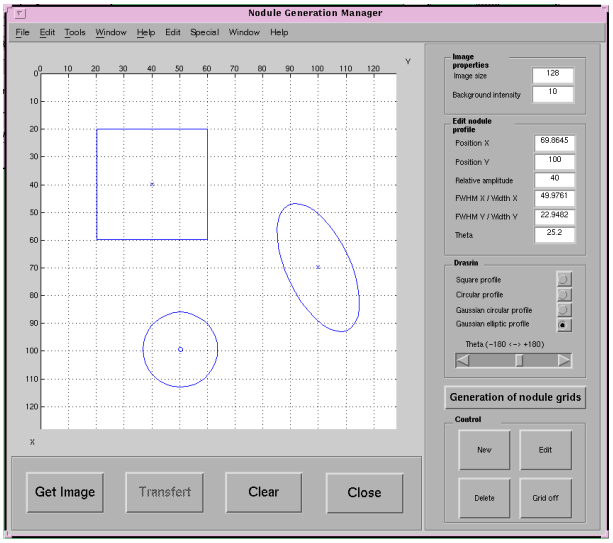


Figure 7.3.: Nodule Generation Manager window

Background intensity (ET):
: background intensity of the image. It is defined as the intensity of the image when there are no nodules.

Nodule properties

Position X and Position Y (ET):
: position of the center of gravity of the nodule.

Relative Amplitude (ET):
: Relative amplitude of the nodule with respect to the background of the image.

FWHM X and FWHM Y (ET):
: Full Width at Half Maximum of the nodules along X and Y. It corresponds to the side length for a square profile and to the diameter for a cylindrical profile.

Theta (ET/SLI):
: angle of rotation of the nodule around its main axes. This angle can also be set by the slider Theta.

Square/Circle/Gaussian circular/Gaussian elliptic (RB):
: These buttons enable to choose the shape of the intensity profile of the nodule.

Nodule grids

Generation of nodule grids (PB):
this option allows the user to generate automatically a nodule grid, which is an image containing a set of nodules of different size and contrast.

Display and window management

New (PB): initializes the editable uicontrols to draw a new nodule.

Edit (PB): selects the nodule to edit with the mouse pointer.

Delete (PB): deletes the nodule indicated by the mouse pointer.

Grid on/off (PB): sets/removes the grid.

Get Image (PB): switches from the axis representation to the preview of the intensity image.

Transfer (PB): transfers the image from the Nodule Generation Manager to the Restoration Manager window.

Clear (PB): clears the axes and the edit uicontrols.

Close (PB): closes the Nodule Generation Manager window.

7.3.3. Uimenus

Edit / Font Size: changes the size of the characters of the window.

Special / Save Data / Commented:
saves the nodules' parameters in a text file with an elegant format.

Special / Save Data / Raw:
stores a raw version of the nodules' parameters.

Special / Save Image / To Postscript:
saves the current image as a postscript file.

Special / Save Image / To .mat file:
stores the image in a .mat file.

Window: opens another DeQuant GUI window.

7.4. PSF Generation Manager

The purpose of this window is to generate a simulated PSF. Since the PSF in nuclear medicine can be assimilated to a Gaussian intensity profile, the structure of the Psf Generation Manager window is similar to that of the Nodule Generation Manager window. The superfluous options have been removed, see Fig. 7.4.

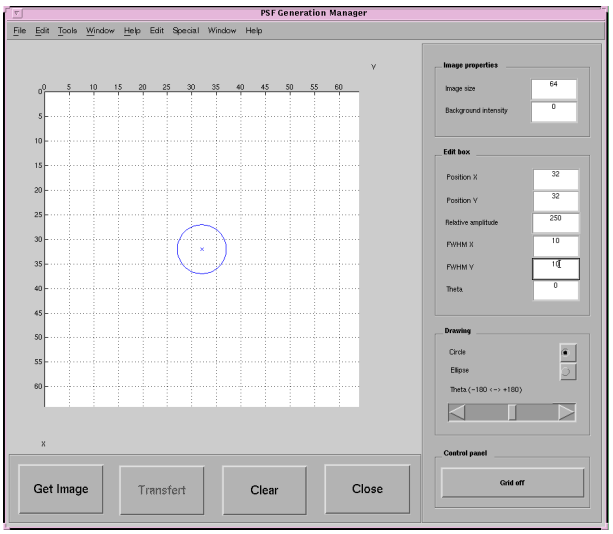


Figure 7.4.: Title window of the application

7.5. Restoration Manager

The Restoration Manager is the main window of the application, see Fig 7.5. In the case of a simulated data experiment, the Restoration Manager enables to convolve the uptake profile generated by the Nodule Generation Manager with the PSF generated by the Nodule Generation Manager. The resulting image can be then corrupted by Poisson noise. The window enables then to choose one of the eight DeQuant algorithms, set the parameters J, ϵ and δ, and perform the restoration. It also enables to compute the restoration with the Metz and Wiener filters with the aim of comparison. When working on phantom images or clinical data, the noisy image and the measured PSF can be loaded into the Restoration Manager in order to be processed. Finally, the Measure Lab window can be called from the Restoration Manager in order to quantitatively evaluate the quality of the restoration. More details on the functions and uicontrols of the Restoration Manager window are given below.

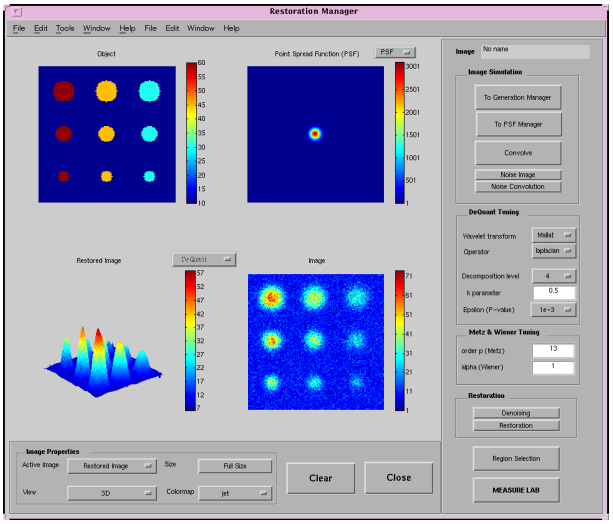


Figure 7.5.: Restoration Manager window

7.5.1. Display

Area where the original uptake image, the result of the convolution, the noisy image and the restored images are displayed. A color bar beside each image makes it easier for the observer to associate colors with the surface values. Every image can be rotated interactively in three dimensions using the mouse. There are two pop-up menus in the axes area:

PSF / Convolution (PUP):
: enables to switch between the the display of the PSF and that of the result of the convolution of the PSF with the object.

Popup button Object / Convolution (PUP):
: enables to switch between the display of the restored object, the restored object convoluted with the PSF, the Wiener restored object and the Metz restored object.

7.5.2. UiControls

Image Simulation

To Generation Manager (PB):
: opens the Nodule Generation Manager window.

To PSF Manager (PB):
: opens the Psf Generation Manager window.

Convolve (PB): convolves the object with the PSF and displays the result in the top right axes.

Noise Image and Noise Convolution (PB):
allows the introduction of Poisson noise. This can be done either on the object (Noise Image) or on the result of the convolution of the object with the PSF (Noise convolution).

DeQuant tuning

Wavelet transform (PUP):
enables to choose between the Mallat wavelet transform and the à trous wavelet transform.

Operator (PUP): enables to choose between the laplacian operator L and the M operator.

Decomposition level (PUP):
allows to choose the parameter J, in $[1, \ldots, J_{Max}]$. Recall that J is the scale to which the wavelet transform is calculated and $J_{Max} = \log_2(\sqrt{N_I})$, where N_I is the number of pixels of the image.

k paramer (ET): permits to set the value of the parameter δ which is the tolerance parameter that has been introduced to take into account the noise in the significant wavelet coefficients.

Epsilon (PUP): enables to choose the significance level of the hypothesis test in $[10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}]$.

Metz and Wiener tuning

Order p (ET): enables to set the order p of the Metz filter.

alpha (ET): allows to set the parameter α of the spectral power division of the Wiener filter

Denoising / Restoration

Denoising (PB): enables to denoise the image (processing with D-DeQuant and its derivates).

Restoration (PB): enables to restore the image (processing with R-DeQuant and its derivates)

Assesment of the restoration results

To Measure Lab (PB):
opens the Mesure Lab window.

Display and window management

Active Image (PB): enables to select one of the images in the display area and sets it as active.

View (PB): permits to switch between the 2D and 3D view of the active image.

Full Size (PB): allows to view the active image in full size in a separate window.

Colormap (PB): changes the color map of all images displayed in the window.

Clear (PB): clears all images and uicontrols.

Close (PB): closes the Restoration Manager window.

7.5.3. Uimenus

File / Load / Object:
loads a .mat file as object.

File / Load / Noisy Image:
loads a .mat file as noisy image.

File / Load / PSF / From mat file:
loads a .mat file as PSF.

File / Load / PSF / From text file:
loads a .txt file as PSF. The data must be stored one per line and in the same order as the PSF parameters in the Psf Generation Manager window (i.e. position x, position y, relative amplitude, $FWHM_X$, $FWHM_Y$, theta).

Edit / Font Size: changes the size of the characters in the window.

Window: opens another window of the application.

7.6. Measure Lab

The Measure Lab window enables to evaluate the performances of the DeQuant algorithms by computing local and global criteria related to the restored image. It also permits to visualize the number of detected horizontal, vertical and diagonal significant wavelet coefficients, as a function of the scale as well as the last smooth image of the analysis. An overview of the uicontrols of the window and their functions is given in the following.

7.6.1. Display

Once the Measure Lab window is opened, one can choose by selecting a push button on the right of the window, between the visualization of the local and global performance criteria related to the restored image on the one hand, and the display of the significant wavelet coefficients detected during the restoration process as a function of the scale.

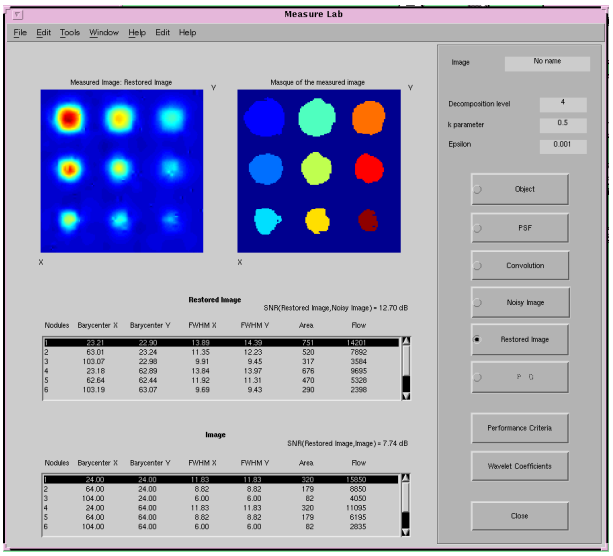


Figure 7.6.: Restoration Manager window - Performance criteria

- Performance criteria: by pushing this button, two images and two slide boxes appear in the display area, see Fig. 7.6. The figure on the top right represents the restored image. The figure on the top left represents the mask resulting from the segmentation of the restored image by the kσ clipping algorithm. The local parameters of the nodules detected by this segmentation are displayed in the upper slide box, while the local parameters of the

nodules in the original image are the lower slide box. The calculated SNR values are displayed at the top of each slide box.

- Wavelet coefficients: when this option is selected, four images and an information block appear in the display area, see Fig 7.7. The figure on the top left represents the last approximation image of the analysis. The other three figures are the histograms of the horizontal, respectively vertical and diagonal detected significant wavelet coefficients. These histograms enable to visualize the number of SWC as a function of the analyzing scale. The information block at the bottom of the display area, gives the total number of detected SWC, the size of the last approximation image and the compression ratio (CR).

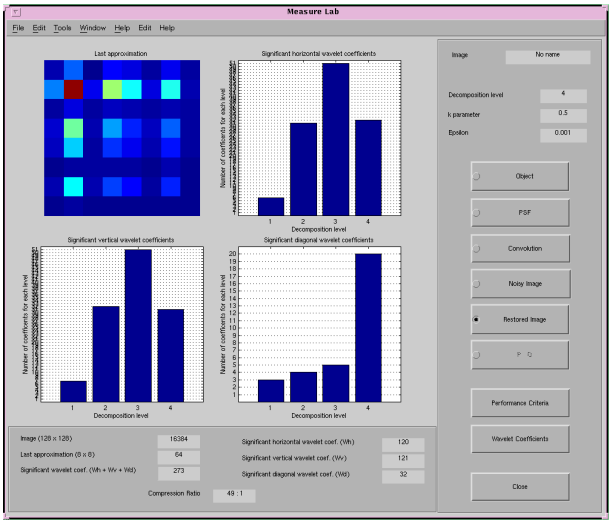


Figure 7.7.: Restoration Manager window - Wavelet coefficients

7.6.2. Uicontrols

Images to measure (RB):
set of radio buttons that enable to choose the image to be measured.

Performance criteria (PB):
displays the performance criteria.

Wavelet coefficients (PB):
displays the number of detected SWC.

Close (PB): closes the Restoration Manager window.

7.6.3. Uimenus

Edit / Font Size: changes the size of the characters in the window.

7.7. Conclusion

This chapter can be considered as a guide that helps the user find his way through the DeQuant GUI. It gives detailed information on the function of each window of the GUI and the different uicontrols it contains. The flexible modular way in which the GUI has been built, enables to integrate easily any new feature in the program.

Summary and perspectives

This work developed, analyzed and validated a new multiscale restoration framework to Poisson intensity estimation in photon limited imagery. This framework, named DeQuant, follows a six steps scheme:

- The **transformation** step leads to a representation of the noisy image in the wavelet domain. This domain has been chosen for its ability to provide an accurate and economical representation of data exhibiting hierarchically localized variations.

  We used Mallat's orthogonal maximally decimated transform [Mal89b] and Coifman and Donoho's shift invariant transform [CD95a] with the unnormalized Haar wavelet. The choice of the Haar wavelet was motivated by the Poisson nature of the data. Although the shift invariant transform results in a larger computational complexity, it showed benefits in suppressing residual blocking artifacts.

- The **selection** step analyzes the statistical significance of the obtained wavelet coefficients in order to separate the "large" and the "small" ones into sets of significant and non significant coefficients respectively. The significant coefficients contain information. The non significant coefficients are those for which the hypothesis of lack of information is not rejected.

  The selection step prerequisites the knowledge of the probability density function of the wavelet coefficients of a Poisson process. This PDF depends on the choice of the wavelet. We calculated in Chapter 5 a general expression valid for any wavelet. We then considered the particular case of the unnormalized Haar wavelet for which we computed the threshold tables given in Appendix A.

- The **regularization** step assigns a new value to the non significant coefficients by taking into account a priori information on the object of interest.

  We showed that this step can be seen as an edge detecting operation and investigated two edge detectors: the Laplacian L_2 and an operator M introduced by Bijaoui and Froeschlé [BF80].

- The **projection** step enables to avoid contradictions between the significance of the wavelet coefficients of the noisy image and the significance of the wavelet coefficients of the restored image.
- The **reconstruction** step calculates the inverse wavelet transform using the significant wavelet coefficients and the projected non significant ones.
- The **iteration** step repeats the whole procedure. This is due to the fact that the PDF of the selection step is a function of the object that has to be estimated.

We showed that DeQuant can be used for denoising and restoration purposes. By denoising we mean the estimation of the intensity of a Poisson process from the observation of the counts. By restoration we mean the recovery of an object related through a linear system of equations to the intensity function of the Poisson process from the observation of the Poisson data. The restoration problem arises in nuclear medicine when trying to remove the effect of the point spread function of the imaging device.

Combining all possibilities that have been addressed in this work, namely denoising or restoration for the purpose, Mallat or Coifman for the wavelet transform, and L_2 or M for the regularization operator, we obtained the eight different algorithms described in Appendix A. The many possible algorithms based on the same global structure are a strength of DeQuant.

The DeQuant algorithms depend on three parameters: the scale J to which the wavelet transform is calculated, the level of evidence ϵ that represents the probability that a coefficient be classified as significant when it in fact lacks enough local information, and the tolerance parameter δ that enables to take into account the noise in the significant wavelet coefficients.

The validation of the algorithms started with simulated experiments which enabled to analyze the contrast and size sensibility of DeQuant, gain some insights in the choice of the algorithms' parameters J, ϵ and δ and compare the eight different algorithms using global and local performance criteria. The algorithms have then been tested in the real world scanning environment by processing actual images of thyroid phantoms. The last stage of the evaluation process was the validation with real thyroid data acquired in vivo.

The evaluation showed that DeQuant fulfills the requirement set forth at the initial design stage, namely the reduction of noise in photon limited imagery. The estimates provided by DeQuant would enable a greater diagnostic confidence in clinical nuclear medicine since they give the physician the access to the diagnosis relevant information with a measure of the significance of the detected structures.

DeQuant has been applied to conventional radionuclide imaging using projection imaging techniques (planar scintigraphy). Future work could extend the framework to computed emission tomographic imaging such as the single photon emission tomography (SPECT).

DeQuant can be used in any problem in science and engineering involving the recovery of an object of interest from collected Poisson data. Nuclear medicine imaging and astronomical imaging are just two examples thereof.

A. DeQuant Algorithms

If we combine all possibilities presented in Chapter 4, we obtain eight different DeQuant algorithms having the same global structure. The names of these algorithms are listed in Table A.1, where the prefix D- refers to denoising, the prefix R- to restoration, the prefix SI- to shift invariance and the prefix EP- to edge preserving.

	Transformation		Regularization		Objective	
	MD-WT	SI-WT	L_2	M	X	Y
D-DeQuant	x		x		x	
SI-D-DeQuant		x	x		x	
EP-D-DeQuant	x			x	x	
SI-EP-D-DeQuant		x		x	x	
R-DeQuant	x		x			x
SI-R-DeQuant		x	x			x
EP-R-DeQuant	x			x		x
SI-EP-R-DeQuant		x		x		x

Table A.1.: Eight different DeQuant algorithms

In all algorithms, a coarse to fine processing is applied to the wavelet coefficients of the raw data image $W_{h,I}$, $W_{v,I}$ and $W_{d,I}$. The restoration starts with the biggest scale 2^J until it reaches the scale 2^0 corresponding to the restored image. The noise free image is first supposed to be constant, it is then progressively modified and takes the value $Y^{(i)}(k_x, k_y) = F^{(n)}_{Y^{(i)}}(2^0)$. A detailed description of the algorithms is given in the following pages. ϵ_0 and ϵ_1 are convergence parameters.

A.1. D-DeQuant

The D-DeQuant algorithm can be described as follows:

1. Compute the wavelet coefficients $W_{h,I}$, $W_{v,I}$, and $W_{d,I}$ of the noisy image I until the desired scale 2^J is reached. This will be done using the equations for the orthogonal maximally decimated Mallat wavelet transform, given by (2.42).
2. Initialize i to 0.
3. Start with the constant Poisson parameter
$$Y^{(0)}(k_x, k_y) = \sum I / (\text{number of pixels of } I)$$
4. $F^{(1)}_{Y^{(1)}}(2^J) = F_I(2^J)$
5. **Repeat**
   A. $i \leftarrow i + 1$
   B. **For** $j = J$ to $j = 1$ do
      B-1. $n = 0$
      B-2. $W_{a,F^{(1)}_{Y^{(i)}}(2^{j-1})} = 0$ for $a = h$, v, and d
      B-3. **Repeat**
         B-3-1. $n \leftarrow n + 1$
         B-3-2. Compute $F^{(n)}_{Y^{(i)}}(2^{j-1})$ using (4.15)
$$\begin{aligned} F^{(n)}_{Y^{(i)}}(2^{j-1}) &= \big[\, \widetilde{H}_2 F^{(n)}_{Y^{(i)}}(2^j) + \widetilde{G}_h W_{h,F^{(n)}_{Y^{(i)}}(2^{j-1})} \\ &\quad + \widetilde{G}_v W_{v,F^{(n)}_{Y^{(i)}}(2^{j-1})} + \widetilde{G}_d W_{d,F^{(n)}_{Y^{(i)}}(2^{j-1})} \,\big] / 4 \end{aligned}$$
         B-3-3. Compute the wavelet coefficients of $F^{(n+1)}_{Y^{(i)}}(2^{j-1})$ using (4.14)
$$\begin{aligned} W_{h,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{h,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_{L_2} W_{h,L_2F^{(n)}_{Y^{(i)}}(2^{j-1})} \\ W_{v,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{v,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_{L_2} W_{v,L_2F^{(n)}_{Y^{(i)}}(2^{j-1})} \\ W_{d,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{d,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_{L_2} W_{d,L_2F^{(n)}_{Y^{(i)}}(2^{j-1})} \end{aligned}$$
         B-3-4. Apply the non-linear constraints to the wavelet coefficients.
         until $\| W_{a,F^{(n)}_{Y^{(i)}}(2^{j-1})} - W_{a,F^{(n-1)}_{Y^{(i)}}(2^{j-1})} \|^2 < \epsilon_0$ for $a = h$, v, and d
      B-4. Apply the positivity constraint on $F^{(n)}_{Y^{(i)}}(2^{j-1})$
      EndFor
   C. $Y^{(i)}(k_x, k_y) = F^{(n)}_{Y^{(i)}}(2^0)$

   until $\| Y^{(i)}(k_x, k_y) - Y^{(i-1)}(k_x, k_y) \|^2 < \epsilon_1$

The restored image is $Y^{(i)}(k_x, k_y)$.

A.2. SI-D-DeQuant

1. Compute the wavelet coefficients $W_{h,I}^T$, $W_{v,I}^T$, and $W_{d,I}^T$ of the noisy image I until the desired scale 2^J is reached. This will be done using the equations for Coifman and Donoho's undecimated shift-invariant wavelet transform, given by (2.49).
2. Initialize i to 0.
3. Start with the constant Poisson parameter

$$Y^{(0)}(k_x, k_y) = \sum I/(\text{number of pixels of } I)$$

4. $F_{Y^{(1)}}^{T(1)}(2^J) = F_I^T(2^J)$
5. **Repeat**

   A. $i \leftarrow i + 1$

   B. **For** $j = J$ to $j = 1$ do

      B-1. $n = 0$

      B-2. $W^T_{a,F_{Y^{(i)}}^{T(1)}(2^{j-1})} = 0$ for $a = h$, v, and d

      B-3. **Repeat**

         B-3-1. $n \leftarrow n + 1$

         B-3-2. Compute $F_{Y^{(i)}}^{T(n)}(2^{j-1})$ using

$$\begin{aligned} F_{Y^{(i)}}^{T(n)}(2^{j-1}) &= \big[\, \widetilde{H_2}(2^{j-1}) F_{Y^{(i)}}^{T(n)}(2^j) + \widetilde{G_h}(2^{j-1}) W^T_{h,F_{Y^{(i)}}^{T(n)}(2^{j-1})} \\ &+ \widetilde{G_v}(2^{j-1}) W^T_{v,F_{Y^{(i)}}^{T(n)}(2^{j-1})} + \widetilde{G_d}(2^{j-1}) W^T_{d,F_{Y^{(i)}}^{T(n)}(2^{j-1})} \,\big]/16 \end{aligned}$$

         B-3-3. Compute the wavelet coefficients of $F_{Y^{(i)}}^{T(n+1)}(2^{j-1})$

$$\begin{aligned} W^T_{h,F_{Y^{(i)}}^{T(n+1)}(2^{j-1})} &= W^T_{h,F_{Y^{(i)}}^{T(n)}(2^{j-1})} - \xi_{L_2} W^T_{h,L_2F_{Y^{(i)}}^{T(n)}(2^{j-1})} \\ W^T_{v,F_{Y^{(i)}}^{T(n+1)}(2^{j-1})} &= W^T_{v,F_{Y^{(i)}}^{T(n)}(2^{j-1})} - \xi_{L_2} W^T_{v,L_2F_{Y^{(i)}}^{T(n)}(2^{j-1})} \\ W^T_{d,F_{Y^{(i)}}^{T(n+1)}(2^{j-1})} &= W^T_{d,F_{Y^{(i)}}^{T(n)}(2^{j-1})} - \xi_{L_2} W^T_{d,L_2F_{Y^{(i)}}^{T(n)}(2^{j-1})} \end{aligned}$$

         B-3-4 Apply the non-linear constraints to the wavelet coefficients.

      until $\|W^T_{a,F_{Y^{(i)}}^{T(n)}(2^{j-1})} - W^T_{a,F_{Y^{(i)}}^{T(n-1)}(2^{j-1})}\|^2 < \epsilon_0$ for $a = h$, v, and d

      B-4 Apply the positivity constraint on $F_{Y^{(i)}}^{T(n+1)}(2^{j-1})$

      EndFor

   C. $Y^{(i)}(k_x, k_y) = F_{Y^{(i)}}^{T(n)}(2^0)$

   until $\|Y^{(i)}(k_x, k_y) - Y^{(i-1)}(k_x, k_y)\|^2 < \epsilon_1$

The restored image is $Y^{(i)}(k_x, k_y)$.

A.3. EP-D-DeQuant

1. Compute the wavelet coefficients $W_{h,I}$, $W_{v,I}$, and $W_{d,I}$ of the noisy image I until the desired scale 2^J is reached. This will be done using the equations for the orthogonal maximally decimated wavelet transform, given by (2.42).
2. Initialize i to 0.
3. Start with the constant Poisson parameter
$$Y^{(0)}(k_x, k_y) = \sum I / (\text{number of pixels of } I)$$
4. $F^{(1)}_{Y^{(1)}}(2^J) = F_I(2^J)$
5. **Repeat**
   A. $i \leftarrow i + 1$
   B. **For** $j = J$ to $j = 1$ do
      B-1. $n = 0$
      B-2. $W_{a,F^{(1)}_{Y^{(i)}}(2^{j-1})} = 0$ for $a = h$, v, and d
      B-3. **Repeat**
         B-3-1. $n \leftarrow n + 1$
         B-3-2. Compute $F^{(n)}_{Y^{(i)}}(2^{j-1})$ using
$$\begin{aligned} F^{(n)}_{Y^{(i)}}(2^{j-1}) &= \big[\, \widetilde{H}_2 F^{(n)}_{Y^{(i)}}(2^j) + \widetilde{G}_h W_{h,F^{(n)}_{Y^{(i)}}(2^{j-1})} \\ &+ \widetilde{G}_v W_{v,F^{(n)}_{Y^{(i)}}(2^{j-1})} + \widetilde{G}_d W_{d,F^{(n)}_{Y^{(i)}}(2^{j-1})} \,\big] / 4 \end{aligned}$$
         B-3-3. Compute the wavelet coefficients of $F^{(n+1)}_{Y}(2^{j-1})$
$$\begin{aligned} W_{h,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{h,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W_{h,MF^{(n)}_{Y^{(i)}}(2^{j-1})} \\ W_{v,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{v,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W_{v,MF^{(n)}_{Y^{(i)}}(2^{j-1})} \\ W_{d,F^{(n+1)}_{Y^{(i)}}(2^{j-1})} &= W_{d,F^{(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W_{d,MF^{(n)}_{Y^{(i)}}(2^{j-1})} \end{aligned}$$
         B-3-4. Apply the non-linear constraints to the wavelet coefficients.
         until $\| W_{a,F^{(n)}_{Y^{(i)}}(2^{j-1})} - W_{a,F^{(n-1)}_{Y^{(i)}}(2^{j-1})} \|^2 < \epsilon_0$ for $a = h$, v, and d
      B-4. Apply the positivity constraint on $F^{(n)}_{Y^{(i)}}(2^{j-1})$
      EndFor
   C. $Y^{(i)}(k_x, k_y) = F^{(n)}_{Y^{(i)}}(2^0)$

   until $\| Y^{(i)}(k_x, k_y) - Y^{(i-1)}(k_x, k_y) \|^2 < \epsilon_1$

The restored image is $Y^{(i)}(k_x, k_y)$.

A.4. SI-EP-D-DeQuant

1. Compute the wavelet coefficients $W^T_{h,I}$, $W^T_{v,I}$, and $W^T_{d,I}$ of the noisy image I until the desired scale 2^J is reached. This will be done using the equations for Coifman and Donoho's undecimated shift-invariant wavelet transform, given by (2.49).

2. Initialize i to 0.

3. Start with the constant Poisson parameter

$$Y^{(0)}(k_x, k_y) = \sum I/(\text{number of pixels of } I)$$

4. $F^{T(1)}_{Y^{(1)}}(2^J) = F^T_I(2^J)$

5. **Repeat**

   A. $i \leftarrow i + 1$

   B. **For** $j = J$ to $j = 1$ do

      B-1. $n = 0$

      B-2. $W^T_{a,F^{T(1)}_{Y^{(i)}}(2^{j-1})} = 0$ for $a = h$, v, and d

      B-3. **Repeat**

         B-3-1. $n \leftarrow n + 1$

         B-3-2. Compute $F^{T(n)}_{Y^{(i)}}(2^{j-1})$ using

$$F^{T(n)}_{Y^{(i)}}(2^{j-1}) = \big[\, \widetilde{H_2}(2^{j-1}) F^{T(n)}_{Y^{(i)}}(2^j) + \widetilde{G_h}(2^{j-1}) W^T_{h,F^{T(n)}_{Y^{(i)}}(2^{j-1})} + \widetilde{G_v}(2^{j-1}) W^T_{v,F^{T(n)}_{Y^{(i)}}(2^{j-1})} + \widetilde{G_d}(2^{j-1}) W^T_{d,F^{T(n)}_{Y^{(i)}}(2^{j-1})} \,\big] /16$$

         B-3-3. Compute the wavelet coefficients of $F^{T(n+1)}_{Y^{(i)}}(2^{j-1})$

$$\begin{aligned} W^T_{h,F^{T(n+1)}_{Y^{(i)}}(2^{j-1})} &= W^T_{h,F^{T(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W^T_{h,MF^{T(n)}_{Y^{(i)}}(2^{j-1})} \\ W^T_{v,F^{T(n+1)}_{Y^{(i)}}(2^{j-1})} &= W^T_{v,F^{T(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W^T_{v,MF^{T(n)}_{Y^{(i)}}(2^{j-1})} \\ W^T_{d,F^{T(n+1)}_{Y^{(i)}}(2^{j-1})} &= W^T_{d,F^{T(n)}_{Y^{(i)}}(2^{j-1})} - \xi_M W^T_{d,MF^{T(n)}_{Y^{(i)}}(2^{j-1})} \end{aligned}$$

         B-3-4 Apply the non-linear constraints to the wavelet coefficients.

      until $\|W^T_{a,F^{T(n)}_{Y^{(i)}}(2^{j-1})} - W^T_{a,F^{T(n-1)}_{Y^{(i)}}(2^{j-1})}\|^2 < \epsilon_0$ for $a = h$, v, and d

      B-4 Apply the positivity constraint on $F^{T(n+1)}_{Y^{(i)}}(2^{j-1})$

      EndFor

   C. $Y^{(i)}(k_x, k_y) = F^{T(n)}_{Y^{(i)}}(2^0)$

until $\|Y^{(i)}(k_x, k_y) - Y^{(i-1)}(k_x, k_y)\|^2 < \epsilon_1$

A.5. Simple derivation of the D-DeQuant variations

Sections A.2, A.3 and A.4 gave an detailed description of the SI-D-DeQuant, EP-D-DeQuant and SI-EP-D-DeQuant algorithms to make the programing of these algorithms easier.

But since the global structure of the D-DeQuant remains unchanged, we can describe in a more compact way the derivation of these algorithms by simply pointing out the points of D-DeQuant that have to be modified, see Table A.2.

Points of D-DeQuant to be modified	SI-D-DeQuant	SI-EP-D-DeQuant	EP-D-DeQuant
1	Compute the wavelet coefficients $W^T_{h,I}$, $W^T_{v,I}$, and $W^T_{d,I}$ of the noisy image I using the equations for Coifman and Donohos's undecimated SI wavelet transform, given by (2.49).	-	
B-3-2	Compute $F^{T(n)}_{Y^{(i)}}(2^{j-1})$ using Equation (2.52)	-	
B-3-3	-	Compute the wavelet coefficients of $F^{(n+1)}_{Y^{(i)}}(2^{j-1})$ or $F^{T(n+1)}_{Y^{(i)}}(2^{j-1})$ using a modified version of (4.14) where L_2 is replaced by the operator M and the parameter ξ_{L_2} is replaced by ξ_M	

Table A.2.: List of the points to modify in D-DeQuant in order to obtain EP-D-DeQuant, SI-D-DeQuant or SI-EP-D-DeQuant

A.6. R-DeQuant

The R-DeQuant algorithm can be described as follows:

1. Compute the wavelet coefficients $W_{h,I}$, $W_{v,I}$, $W_{d,I}$ of the raw data image I until the desired scale 2^J is reached. This will be done using the equations for the orthogonal maximally decimated Mallat wavelet transform, given by (2.42).
2. $X^{(0)} = \sum I/(\text{number of pixels of } I)$
3. $Y^{(0)} = PX^{(0)}$
4. $n = 0$
5. **Repeat**
   A. $n \leftarrow n + 1$
   B. $F_{Z^{(n)}}(2^J) = F_I(2^J)$
   C. **For** $j = J$ to $j = 1$ do
      C-1. Compute the wavelet coefficients of $Z^{(n)}$ using (4.19)

$$
\begin{aligned}
W_{h,Z^{(n)}}(2^j) &= W_{h,Y^{(n-1)}}(2^j) - \xi_{L_2} W_{h,PL_2X^{(n-1)}}(2^j) \\
W_{v,Z^{(n)}}(2^j) &= W_{v,Y^{(n-1)}}(2^j) - \xi_{L_2} W_{v,PL_2X^{(n-1)}}(2^j) \\
W_{d,Z^{(n)}}(2^j) &= W_{d,Y^{(n-1)}}(2^j) - \xi_{L_2} W_{d,PL_2X^{(n-1)}}(2^j)
\end{aligned}
$$

      C-2. Apply the non-linear constraints to the wavelet coefficients
      C-3. Compute $F_{Z^{(n)}}(2^{j-1})$, the approximation image of $Z^{(n)}$ using (4.20)

$$
\begin{aligned}
F_{Z^{(n)}}(2^{j-1}) \quad = \quad & \big[\, \widetilde{H_2} F_{Z^{(n)}}(2^j) + \widetilde{G_h} W_{h,Z^{(n)}}(2^j) \\
& + \widetilde{G_v} W_{v,Z^{(n)}}(2^j) + \widetilde{G_d} W_{d,Z^{(n)}}(2^j) \,\big] \,/4
\end{aligned}
$$

      C-4. Apply the positivity constraint on $F_{Z^{(n)}}(2^{j-1})$
      EndFor
   D. $Z^{(n)} = F_{Z^{(n)}}(2^0)$
   E. $X^{(n)} = X^{(n-1)} \left[\tilde{P} \frac{Z^{(n)}}{PX^{(n-1)}} \right]$
   F. $Y^{(n)} = PX^{(n)}$

   until $\|Y^{(n)} - Y^{(n-1)}\| < \epsilon_0$

The restored image is $Z^{(n)}$

A.7. SI-R-DeQuant

1. Compute the wavelet coefficients $W^T_{h,I}$, $W^T_{v,I}$, $W^T_{d,I}$ of the raw data image I until the desired scale 2^J is reached. This will be done using the equations for Coifman and Donoho's undecimated shift-invariant wavelet transform, given by (2.49).
2. $X^{(0)} = \sum I/(\text{number of pixels of } I)$
3. $Y^{(0)} = PX^{(0)}$
4. $n = 0$
5. **Repeat**

   A. $n \leftarrow n + 1$

   B. $F^T_{Z^{(n)}}(2^J) = F^T_I(2^J)$

   C. **For** $j = J$ to $j = 1$ do

      C-1. Compute the wavelet coefficients of $Z^{(n)}$

$$
\begin{aligned}
W^T_{h,Z^{(n)}}(2^j) &= W^T_{h,Y^{(n-1)}}(2^j) - \xi_{L_2} W^T_{h,PL_2X^{(n-1)}}(2^j) \\
W^T_{v,Z^{(n)}}(2^j) &= W^T_{v,Y^{(n-1)}}(2^j) - \xi_{L_2} W^T_{v,PL_2X^{(n-1)}}(2^j) \\
W^T_{d,Z^{(n)}}(2^j) &= W^T_{d,Y^{(n-1)}}(2^j) - \xi_{L_2} W^T_{d,PL_2X^{(n-1)}}(2^j)
\end{aligned}
$$

      C-2. Apply the non-linear constraints to the wavelet coefficients

      C-3. Compute $F^T_{Z^{(n)}}(2^{j-1})$, the approximation image of $Z^{(n)}$

$$
\begin{aligned}
F^T_{Z^{(n)}}(2^{j-1}) \;=\; & \big[\, \widetilde{H_2}(2^{j-1}) F^T_{Z^{(n)}}(2^j) + \widetilde{G_h}(2^{j-1}) W^T_{h,Z^{(n)}}(2^j) \\
& + \widetilde{G_v}(2^{j-1}) W^T_{v,Z^{(n)}}(2^j) + \widetilde{G_d}(2^{j-1}) W^T_{d,Z^{(n)}}(2^j) \,\big] \,/16
\end{aligned}
$$

      C-4. Apply the positivity constraint on $F^T_{Z^{(n)}}(2^{j-1})$

      EndFor

   D. $Z^{(n)} = F^T_{Z^{(n)}}(2^0)$

   E. $X^{(n)} = X^{(n-1)} \left[\tilde{P} \frac{Z^{(n)}}{PX^{(n-1)}} \right]$

   F. $Y^{(n)} = PX^{(n)}$

   until $\|Y^{(n)} - Y^{(n-1)}\| < \epsilon_0$

The restored image is $Z^{(n)}$

A.8. EP-R-DeQuant

1. Compute the wavelet coefficients $W_{h,I}$, $W_{v,I}$, $W_{d,I}$ of the raw data image I until the desired scale 2^J is reached. This will be done using the equations for the orthogonal maximally decimated Mallat wavelet transform, given by (2.42).
2. $X^{(0)} = \sum I/(\text{number of pixels of } I)$
3. $Y^{(0)} = PX^{(0)}$
4. $n = 0$
5. **Repeat**
   A. $n \leftarrow n + 1$
   B. $F_{Z^{(n)}}(2^J) = F_I(2^J)$
   C. **For** $j = J$ to $j = 1$ do
      C-1. Compute the wavelet coefficients of $Z^{(n)}$

$$
\begin{aligned}
W_{h,Z^{(n)}}(2^j) &= W_{h,Y^{(n-1)}}(2^j) - \xi_M W_{h,PMX^{(n-1)}}(2^j) \\
W_{v,Z^{(n)}}(2^j) &= W_{v,Y^{(n-1)}}(2^j) - \xi_M W_{v,PMX^{(n-1)}}(2^j) \\
W_{d,Z^{(n)}}(2^j) &= W_{d,Y^{(n-1)}}(2^j) - \xi_M W_{d,PMX^{(n-1)}}(2^j)
\end{aligned}
$$

      C-2. Apply the non-linear constraints to the wavelet coefficients
      C-3. Compute $F_{Z^{(n)}}(2^{j-1})$, the approximation image of $Z^{(n)}$

$$
\begin{aligned}
F_{Z^{(n)}}(2^{j-1}) \quad = \quad & \big[\, \widetilde{H_2} F_{Z^{(n)}}(2^j) + \widetilde{G_h} W_{h,Z^{(n)}}(2^j) \\
& + \widetilde{G_v} W_{v,Z^{(n)}}(2^j) + \widetilde{G_d} W_{d,Z^{(n)}}(2^j) \,\big] \,/4
\end{aligned}
$$

      C-4. Apply the positivity constraint on $F_{Z^{(n)}}(2^{j-1})$

      EndFor

   D. $Z^{(n)} = F_{Z^{(n)}}(2^0)$
   E. $X^{(n)} = X^{(n-1)} \left[\tilde{P} \frac{Z^{(n)}}{PX^{(n-1)}} \right]$
   F. $Y^{(n)} = PX^{(n)}$

   until $\|Y^{(n)} - Y^{(n-1)}\| < \epsilon_0$

The restored image is $Z^{(n)}$

A.9. SI-EP-R-DeQuant

1. Compute the wavelet coefficients $W^T_{h,I}$, $W^T_{v,I}$, $W^T_{d,I}$ of the raw data image I until the desired scale 2^J is reached. This will be done using the equations for Coifman and Donoho's undecimated shift-invariant wavelet transform, given by (2.49).
2. $X^{(0)} = \sum I/(\text{number of pixels of } I)$
3. $Y^{(0)} = PX^{(0)}$
4. $n = 0$
5. **Repeat**

   A. $n \leftarrow n + 1$

   B. $F^T_{Z^{(n)}}(2^J) = F^T_I(2^J)$

   C. **For** $j = J$ to $j = 1$ do

      C-1. Compute the wavelet coefficients of $Z^{(n)}$

$$
\begin{aligned}
W^T_{h,Z^{(n)}}(2^j) &= W^T_{h,Y^{(n-1)}}(2^j) - \xi_M W^T_{h,PMX^{(n-1)}}(2^j) \\
W^T_{v,Z^{(n)}}(2^j) &= W^T_{v,Y^{(n-1)}}(2^j) - \xi_M W^T_{v,PMX^{(n-1)}}(2^j) \\
W^T_{d,Z^{(n)}}(2^j) &= W^T_{d,Y^{(n-1)}}(2^j) - \xi_M W^T_{d,PMX^{(n-1)}}(2^j)
\end{aligned}
$$

      C-2. Apply the non-linear constraints to the wavelet coefficients

      C-3. Compute $F^T_{Z^{(n)}}(2^{j-1})$, the approximation image of $Z^{(n)}$

$$
\begin{aligned}
F^T_{Z^{(n)}}(2^{j-1}) \;=\; & \big[\, \widetilde{H_2}(2^{j-1}) F^T_{Z^{(n)}}(2^j) + \widetilde{G_h}(2^{j-1}) W^T_{h,Z^{(n)}}(2^j) \\
& + \widetilde{G_v}(2^{j-1}) W^T_{v,Z^{(n)}}(2^j) + \widetilde{G_d}(2^{j-1}) W^T_{d,Z^{(n)}}(2^j) \,\big] \,/16
\end{aligned}
$$

      C-4. Apply the positivity constraint on $F^T_{Z^{(n)}}(2^{j-1})$

      EndFor

   D. $Z^{(n)} = F^T_{Z^{(n)}}(2^0)$

   E. $X^{(n)} = X^{(n-1)} \left[\tilde{P} \frac{Z^{(n)}}{PX^{(n-1)}} \right]$

   F. $Y^{(n)} = PX^{(n)}$

   until $\|Y^{(n)} - Y^{(n-1)}\| < \epsilon_0$

The restored image is $Z^{(n)}$

A.10. Simple derivation of the R-DeQuant variations

Sections A.7, A.8 and A.9 gave an detailed description of the SI-R-DeQuant, EP-R-DeQuant and SI-EP-R-DeQuant algorithms to make the programing of these algorithms easier.

But since the global structure of the R-DeQuant remains unchanged, we can describe in a more compact way the derivation of these algorithms by simply pointing out the points of R-DeQuant that have to be modified, see Table A.3.

Points of R-DeQuant to be modified	SI-R-DeQuant	SI-EP-R-DeQuant	EP-R-DeQuant
1	Compute the wavelet coefficients $W^T_{h,I}$, $W^T_{v,I}$, and $W^T_{d,I}$ of the noisy image I using the equations for Coifman and Donohos's undecimated SI wavelet transform, given by (2.49).	-	
C-1	-	Compute the wavelet coefficients of $Z^{(n)}$ using a modified version of (4.19) where L_2 is replaced by the operator M and the parameter ξ_{L_2} is replaced by ξ_M	
C-3	Compute $F^T_{Z^{(n)}}(2^{j-1})$ using Equation (2.52)	-	

Table A.3.: List of the points to modify in R-DeQuant in order to obtain EP-R-DeQuant, SI-R-DeQuant or SI-EP-R-DeQuant

B. An algorithm to determine image edges

This appendix presents the algorithm proposed by Bijaoui and Froeschlé in [SP96]. The authors showed that in order to find the location of the inflexion points, we have to search along the gradient direction for zero-crossings of the second directional derivative.

B.1. The continuous case

For a continuous image $I(x, y)$, they found the expression of this derivative by considering a point (x_0, y_0) of the image belonging to the isophote I_0. Since the intensity is constant along an isophote we have $dI = 0$. This gives at the first order:

$$\frac{\partial I(x,y)}{\partial x}dx + \frac{\partial I(x,y)}{\partial y}dy = 0 \qquad \Rightarrow \qquad \tan\theta = \frac{dy}{dx} = -\frac{\frac{\partial I(x,y)}{\partial x}}{\frac{\partial I(x,y)}{\partial y}} \tag{B.1}$$

They considered then the new set of axis going through (x_0, y_0) and having the direction θ of the tangent to the isophote at point (x_0, y_0), see Figure B.1. We have:

$$\begin{array}{rcrcl} X & = & x\cos\theta & + & y\sin\theta \\ Y & = & -x\sin\theta & + & y\cos\theta \end{array}$$

where

$$\sin\theta = -\frac{\frac{\partial I(x,y)}{\partial x}}{\sqrt{\frac{\partial I(x,y)}{\partial x} + \frac{\partial I(x,y)}{\partial y}}} \qquad \cos\theta = \frac{\frac{\partial I(x,y)}{\partial y}}{\sqrt{\frac{\partial I(x,y)}{\partial x} + \frac{\partial I(x,y)}{\partial y}}}$$

The variation of the intensity in the gradient direction is given by:

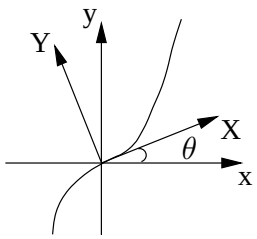


Figure B.1.: Set of axes associated with the point (x_0, y_0)

$$\frac{\partial I(x,y)}{\partial Y} = -\frac{\partial I(x,y)}{\partial x}\sin\theta + \frac{\partial I(x,y)}{\partial y}\cos\theta$$

The second derivative in the Y direction is given by:

$$\frac{\partial^2 I(x,y)}{\partial Y^2} = \frac{\partial^2 I(x,y)}{\partial x^2}\ \sin^2\theta + \frac{\partial^2 I(x,y)}{\partial y^2}\ \cos^2\theta - 2\,\frac{\partial^2 I(x,y)}{\partial x \partial y}\ \sin\theta\cos\theta \quad \text{(B.2)}$$

Setting Equation (B.2) to zero gives the inflexion points in the gradient direction, that is the direction which is normal to the isophote.

B.2. The discrete case

For a discrete image $I(k_x, k_y)$, the second directional derivative $\frac{\partial^2}{\partial Y^2}$(B.2) can be associated with the mask M defined by:

$$M = D_{20}\sin^2\theta + D_{02}\cos^2\theta - 2D_{11}\sin\theta\cos\theta \quad \text{(B.3)}$$

where D_{20}, D_{02} and D_{11} are the discrete approximations of the partial second derivatives $\frac{\partial^2}{\partial x^2}$, $\frac{\partial^2}{\partial y^2}$ and $\frac{\partial^2}{\partial x \partial y}$. $\cos\theta$ and $\sin\theta$ are given by:

$$\cos\theta = \frac{D_{01}I}{\sqrt{D_{10}I + D_{01}I}} \text{ and } \sin\theta = -\frac{D_{10}I}{\sqrt{D_{10}I + D_{01}I}}$$

where D_{10} and D_{01} approximate the first partial derivatives $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$. The masks D_{kl} are calculated by approximating the intensity function $I(x, y)$ on a 3 x 3 pixel field around the point (k_{x_0}, k_{y_0}) by a homogeneous polynomial of the second degree.

Computation of the masks

Let $P(k_{x_i}, k_{y_j})$ be an homogeneous polynomial of the second degree which approximates the intensity function $I(x, y)$ on a 3 x 3 pixel field around the point (k_{x_0}, k_{y_0}):

$$P(k_{x_i}, k_{y_j}) = A_{00} + A_{10}i + A_{01}j + A_{20}i^2 + A_{02}j^2 + A_{11}ij$$

with

$$(i, j) \in [-1,\ 0,\ 1] \text{ x } [-1,\ 0,\ 1]$$

and

$$\begin{matrix} (k_{x_{-1}}, k_{y_1}) & (k_{x_0}, k_{y_1}) & (k_{x_1}, k_{y_1}) \\ (k_{x_{-1}}, k_{y_0}) & (k_{x_0}, k_{y_0}) & (k_{x_1}, k_{y_0}) \\ (k_{x_{-1}}, k_{y_{-1}}) & (k_{x_0}, k_{y_{-1}}) & (k_{x_1}, k_{y_{-1}}) \end{matrix}$$

The coefficients A_{lm} of the polynomial P are computed with the least mean square estimator:

$$A_{lm} \,/ \sum_{i=-1,0,1} \sum_{j=-1,0,1} \left(I(k_{x_i}, k_{y_j}) - \sum_{l=0,1,2} \sum_{m=0,1,2} A_{lm} i^l j^m \right)^2 \quad \text{minimum}$$

If we set

$$A_{kl} = D_{kl} I(k_x, k_y)$$

we get for D_{kl}:

$$D_{00}{=}\tfrac{1}{9}\begin{bmatrix} -1 & 2 & -1 \\ 2 & 5 & 2 \\ -1 & 2 & -1 \end{bmatrix} \quad D_{10}{=}\tfrac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad D_{01}{=}\tfrac{1}{6}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

$$D_{20}{=}\tfrac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix} \quad D_{02}{=}\tfrac{1}{6}\begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{bmatrix} \quad D_{11}{=}\tfrac{1}{4}\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{B.4}$$

Replacing the values of the masks given by Equation (B.4) in Equation (B.3) we obtain the expression of M given by Equation (4.26).

C. Proofs and tables of Chapter 5

This appendix contains outline proofs and derivations for the properties given in Chapter 5. It also contains all tables referred to in this chapter.

C.1. Proof of Proposition 1

The proof of Proposition 1 is done in three steps: (i) Proof of Equation (5.5) (ii) Proof of Equation (5.6) and (iii) Proof of the equivalence of Equations (5.5) and (5.6).

Proof of Equation (5.5)

For Δx tending towards 0, the parameter λ_k of the Poisson process tends also towards zero, see Equation (5.4), which means that the number of events n_k will either be zero or one. Equation (5.2) becomes:

$$W_{\rho(x)}(a,b) = \sum_m \psi(\frac{x_m - b}{a}) \tag{C.1}$$

where x_m corresponds to the position of event m. If the wavelet function ψ has a compact support, the number of events that contribute to the estimation of $W_{\rho(x)}(a,b)$ is limited. Let n denote this number.

Moreover, if the function $\rho(x)$ is constant and equal to ρ_0 on the interval a where $\psi(\frac{x-b}{a})$ is not equal to 0, then the events are randomly distributed in this interval. Each event contributes to $W_{\rho(x)}(a,b)$ for a value which is distributed following the histogram H_ψ of the wavelet ψ. This means that the wavelet coefficient can be considered as a realization of the sum of independent random variables having the same distribution as the values of the wavelet function. The distribution of one event in the wavelet space is given by the histogram H_ψ of the wavelet ψ,

which must be normalized in order to represent a probability density function. Since independent events are considered, the distribution of a wavelet coefficient related to n events is given by n autoconvolutions of H_ψ:

$$p_{\mathcal{W}/N}(w/n) = \underbrace{H_\psi * \ldots * H_\psi}_{n \text{ times}}(w) \tag{C.2}$$

In the following we will use the notation:

$$p_{\mathcal{W}/N}(w/n) = H_\psi^{(*n)}(w)$$

Using the total probability theorem [Sap90] we can write

$$p_{\mathcal{W}}(\alpha, w) = \sum_{n=0}^{\infty} p_{\mathcal{W}/N}(w/n) p_N(n) \tag{C.3}$$

with

$$p_N(n) = \frac{\alpha^n}{n!} e^{-\alpha} \tag{C.4}$$

and

$$p_{\mathcal{W}/N}(w/n) = \begin{cases} \delta(w) & \text{if } n = 0 \\ H_\psi^{(*n)}(w) & \text{if } n \neq 0 \end{cases} \tag{C.5}$$

The parameter α corresponds to the mean number of events in the interval where the wavelet function is not equal to 0. It is related to the parameter ρ_0 by the relation $\alpha = \rho_0 a$, where a is the scale parameter and correspond, since $D = 1$ to the size of the wavelet support. Substituting Equations (C.4) and (C.5) in Equation (C.3), we obtain for the probability distribution of the wavelet coefficient:

$$p_{\mathcal{W}}(\alpha, w) = e^{-\alpha}\delta(w) + \frac{\alpha^1}{1!} e^{-\alpha} H_\psi^{(*1)}(w) + \frac{\alpha^2}{2!} e^{-\alpha} H_\psi^{(*2)}(w) + \frac{\alpha^3}{3!} e^{-\alpha} H_\psi^{(*3)}(w) + \ldots \tag{C.6}$$

The characteristic function $\phi_{\mathcal{W}}(\nu)$ is defined as the Fourier transform $\psi(\nu)$ of the probability density function [Sap90]. This gives in our case:

$$\begin{aligned} \phi_{\mathcal{W}}(\nu) &= e^{-\alpha} + \tfrac{\alpha^1}{1!} e^{-\alpha} H_\psi^1(\nu) + \tfrac{\alpha^2}{2!} e^{-\alpha} H_\psi^2(\nu) + \tfrac{\alpha^3}{3!} e^{-\alpha} H_\psi^3(\nu) + \ldots \\ &= e^{-\alpha}[\delta(w) + \tfrac{\alpha^1}{1!} H_\psi^1(\nu) + \tfrac{\alpha^2}{2!} H_\psi^2(\nu) + \tfrac{\alpha^3}{3!} H_\psi^3(\nu) + \ldots] \\ &= e^{-\alpha} e^{\alpha H_\psi(\nu)} \end{aligned}$$

Which corresponds to

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha\,(\,H_\psi(\nu) - 1\,)}$$

Proof of Equation (5.6)

The characteristic function of the Poisson variable N_k (see Equation (5.3)) of parameter λ_k is given by [Sap90]:

$$\phi_{N_k}(\nu) = e^{\lambda_k(-1+e^{-i2\pi\nu})} \tag{C.7}$$

Multiplying N_k by $\psi(\frac{k\Delta x-b}{a})$ results in

$$\phi_{N_k\psi(\frac{k\Delta x-b}{a})}(\nu) = e^{\lambda_k(-1+e^{-i2\pi\nu\psi(\frac{k\Delta x-b}{a})})}$$

because the characteristic function of a random variable Y defined by $Y = dX$, where d is a nonzero real constant, is $\phi_Y(\nu) = \phi_X(d\nu)$. Summing $N_k\psi(\frac{k\Delta x-b}{a})$ over k gives for the characteristic function

$$\phi_{\sum\limits_k N_k\psi(\frac{k\Delta x-b}{a})}(\nu) = e^{\sum\limits_k \lambda_k(-1+e^{-i2\pi\nu\psi(\frac{k\Delta x-b}{a})})} \tag{C.8}$$

We have used here the fact that the characteristic function of Y_n defined as the sum of n mutually independent random variables X_k of characteristic function ϕ_{X_k} is $\phi_{Y_n}(\nu) = \prod_{k=1}^{n}\phi_{X_k}(\nu)$. For a locally constant $\rho(x) = \rho_0$, we have $\lambda_k = \rho_0\Delta x$. Equation (C.8) becomes:

$$\phi_{\mathcal{W}}(\nu) = e^{\rho_0\sum\limits_k(-1+e^{-i2\pi\nu\psi(\frac{k\Delta x-b}{a})})\Delta x}$$

For Δx tending towards zero, Δx transforms in a dx and the discrete sum in an integral sum:

$$\phi_{\mathcal{W}}(\nu) = e^{\rho_0\int\limits_{aD}(-1+e^{-i2\pi\nu\psi(\frac{x-b}{a})})dx}$$

where $D = 1$ is the length of the support of the wavelet $\psi(x)$. With the introduction of the reduced variable $s = \frac{x-b}{a}$ we obtain

$$\phi_{\mathcal{W}}(\nu) = e^{a\rho_0\int\limits_D(-1+e^{-i2\pi\nu\psi(s)})ds}$$

Proof of the equivalence of Equations (5.5) and (5.6)

We are going to prove the equivalence of Equations (5.5) and (5.6) by showing that the following equality holds:

$$\int\limits_D(-1+e^{-i2\pi\nu\psi(s)})ds = H_\psi(\nu) - 1 \tag{C.9}$$

We therefore start by dividing the s axis into intervals of size Δs, the left part of Equation (C.9) can then be written in the form

$$E1 = \int_D (-1 + e^{-i2\pi\nu\psi(s)})ds \approx \sum_s (-1 + e^{-i2\pi\nu\psi(s)})\Delta s \tag{C.10}$$

We define $N(\psi)$ as the number of values of $\psi(s)$ falling in the interval $[\psi,\ \psi+\Delta\psi]$. With this definition we can rewrite Equation(C.10) in the form

$$E1 = \sum_\psi (-1 + e^{-i2\pi\nu\psi(s)})N(\psi)\Delta s \tag{C.11}$$

Having $\sum\limits_\psi N(\psi) = \frac{1}{\Delta s}$, we define the frequency of occurence of a value by

$$f(\psi) = \frac{N(\psi)}{\frac{1}{\Delta s}} = N(\psi)\Delta s \tag{C.12}$$

We replace $N(\psi)$ in Equation (C.11) by its expression as a function of the frequency calculated from Equation (C.12) :

$$E1 = \sum_\psi (-1 + e^{-i2\pi\nu\psi(s)})\, f(\psi) \tag{C.13}$$

The frequency $f(\psi)$ tends towards a probability when $\Delta\psi$ tends towards zero. We define the histogram $H(\psi)$ as being the limit of the expression $\frac{f(\psi)}{\Delta\psi}$ if $\Delta\psi$ tends towards zero:

$$H(\psi) = \lim_{\Delta\psi\to 0} \frac{f(\psi)}{\Delta\psi}$$

we have

$$\int H(\psi)d\psi = 1$$

In order to introduce the histogram in Equation (C.13), we multiply and divide it by $\Delta\psi$ and let $\Delta\psi$ tend towards zero:

$$\begin{aligned}
\lim_{\Delta\psi\to 0} E1 &= \lim_{\Delta\psi\to 0} \sum_\psi (-1 + e^{-i2\pi\nu\psi(s)})\, \tfrac{f(\psi)}{\Delta\psi}\Delta\psi \\
&= \int_D (-1 + e^{-i2\pi\nu\psi(s)})H(\psi)d\psi \\
&= -\int_D H(\psi)d\psi + \int_D e^{-i2\pi\nu\psi(s)}H(\psi)d\psi \\
&= H_\psi(\nu) - 1
\end{aligned}$$

C.2. Proof of Proposition 2

The characteristic function

$$\phi_{\mathcal{W}}(\nu) = e^{\alpha \int\limits_D (-1+e^{-i2\pi\nu\psi(s)})ds}$$

can be written as:

$$\begin{aligned}\phi_{\mathcal{W}}(\nu) &= e^{\alpha \int\limits_D \left(-1+1+(-i2\pi\nu\psi(s))+\ldots+\frac{(-i2\pi\nu\psi(s))^n}{n!}+\ldots\right)ds} \\ &= e^{\alpha\left[(-i2\pi\nu)\int\limits_D \psi(s)ds+\ldots+\frac{(-i2\pi\nu)^n}{n!}\int\limits_D \psi^n(s)ds+\ldots\right]}\end{aligned}$$

Let I_n be the integrals of the functions $\psi^n(s)$:

$$I_n = \int\limits_D \psi^n(s)ds \tag{C.14}$$

Having $I_1 = 0$ for any admissible wavelet, we obtain:

$$\begin{aligned}\phi_{\mathcal{W}}(\nu) &= e^{\alpha \sum\limits_{n=2}^{\infty}(-i2\pi\nu)^n \frac{I_n}{n!}} \\ &= e^{\alpha(-i2\pi\nu)^2\frac{I_2}{2!}}\, e^{\alpha \sum\limits_{n=3}^{\infty}(-i2\pi\nu)^n \frac{I_n}{n!}} \\ &= e^{-2\pi^2\nu^2\alpha I_2}\left[1+T+\frac{T^2}{2!}+\ldots\right]\end{aligned}$$

with

$$T = \alpha\sum_{n=3}^{\infty}(-i2\pi\nu)^n\frac{I_n}{n!}$$

The function $e^{-2\pi^2\nu^2\alpha I_2}$ can be interpreted as the characteristic function of a Gaussian variable with a zero mean and a variance $v = \alpha I_2$:

$$\phi_g(\nu) = e^{-2\pi^2\nu^2\alpha I_2} \quad \longleftrightarrow \quad g(w) = \tfrac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2\alpha I_2}}$$

We thus have:

$$\phi_{\mathcal{W}}(\nu) = \phi_g(\nu)\left[1+T+\frac{T^2}{2!}+\ldots\right] \tag{C.15}$$

Recall that our goal is to find the probability density function of the wavelet coefficient. We must therefore calculate the inverse Fourier transform of $\phi_{\mathcal{W}}(\nu)$. If

we can express $\phi_{\mathcal{W}}(\nu)$ as a function of terms of the form $(i2\pi\nu)^n\phi_g(\nu)$, the inverse Fourier transform would be easily computed due to the following property:

$$(i2\pi\nu)^n\phi_g(\nu) \quad \longleftrightarrow \quad g^{(n)}(w) \tag{C.16}$$

where $g^{(n)}(w)$ corresponds to the nth derivative of $g(w)$. Applying the results of Table C.2 on the development of terms of the form

$$\left(\alpha\sum_{n=3}^{\infty}(-i2\pi\nu)^n\frac{I_n}{n!}\right)^j \tag{C.17}$$

we obtain a new expression of the characteristic function given by Table C.3. This table is to be read by summing the expressions in each row.

The probability density function $p_{\mathcal{W}}(\alpha, w)$ is the inverse Fourier transform of $\phi_{\mathcal{W}}(\nu)$. Using the correspondence given by Equation (C.16) between the Fourier space and the direct space, Table C.3 transforms in Table C.4.

The idea now is to replace the variable w by a the centered and normalized variable u defined as

$$u = \frac{w - \mathrm{E}[\mathcal{W}]}{\sqrt{\mathrm{Var}[\mathcal{W}]}} \tag{C.18}$$

We therefore need to calculate the moments of the random variable $\mathcal{W}$. This can be done using the moment generating property [Sap90] of the characteristic function:

$$\begin{aligned} \mathrm{E}[\mathcal{W}] &= \alpha I_1 = 0 \\ E[\mathcal{W}^2] &= \alpha I_2 \end{aligned}$$

The mean being equal to zero, the variance $\mathrm{Var}[\mathcal{W}]$ of the random variable $\mathcal{W}$ equals its 2nd order moment:

$$\mathrm{Var}[\mathcal{W}] = \alpha I_2 \tag{C.19}$$

The final expression for u can be written as:

$$u = \frac{w}{\sqrt{\mathrm{Var}[\mathcal{W}]}} = \frac{w}{\sqrt{\alpha I_2}} = \frac{\beta w}{\sqrt{I_2}} \tag{C.20}$$

with

$$\beta = \frac{1}{\sqrt{\alpha}} \tag{C.21}$$

The nth derivative of the function $g(w)$ is related to that of the function $g(u) = e^{-\frac{u^2}{2}}$ by:

$$g^{(n)}(w) = \frac{1}{\sqrt{(\mathrm{Var}[\mathcal{W}])^n}} g^{(n)}(u) = \frac{\beta^n}{\sqrt{I_2^n}} g^{(n)}(u)$$

Since (see Equation (5.8))

$$g^{(n)}(u) = P_n(u)g(u)$$

we can write $g^{(n)}(w)$ in the form

$$g^{(n)}(w) = \frac{\beta^n}{\sqrt{I_2^n}} P_n(u)g(u). \tag{C.22}$$

By replacing in Table C.4 the parameter α by $\frac{1}{\beta^2}$, see Equation (C.21), and $g^{(n)}(w)$ by the expression given by Equation (C.22), we obtain the PDF given by Table C.1 as a function of u and β.

C.3. Tables

j		$p_{\mathcal{W}}(\beta,u)^{\#}$
0		$g(u)$
1	+	$\sum_{n=3}^{\infty}(-1)^n\beta^{n-2}P_n(u)g(u)\frac{I_n}{n!\sqrt{I_2^n}}$
2	+	$\frac{1}{2!}\sum_{n=3}^{\infty}(-1)^{2n}\beta^{2n-4}P_{2n}(u)g(u)\frac{I_n^2}{(n!)^2\sqrt{I_2^{2n}}}$
	+	$\frac{2}{2!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-1)^{n+l}\beta^{n+l-4}P_{n+l}(u)g(u)\frac{I_nI_l}{n!\ l!\sqrt{I_2^{n+l}}}$
3	+	$\frac{1}{3!}\sum_{n=3}^{\infty}(-1)^{3n}\beta^{3n-6}P_{3n}(u)g(u)\frac{I_n^3}{(n!)^3\sqrt{I_2^{3n}}}$
	+	$\frac{3}{3!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-1)^{2n+l}\beta^{2n+l-6}P_{2n+l}(u)g(u)\frac{I_n^2I_l}{(n!)^2\ l!\sqrt{I_2^{2n+l}}}$
	+	$\frac{6}{3!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}(-1)^{n+l+m}\beta^{n+l+m-6}P_{n+l+m}(u)g(u)\frac{I_nI_lI_m}{n!\ l!\ m!\sqrt{I_2^{n+l+m}}}$
4	+	$\frac{1}{4!}\sum_{n=3}^{\infty}(-1)^{4n}\beta^{4n-8}P_{4n}(u)g(u)\frac{I_n^4}{(n!)^4\sqrt{I_2^{4n}}}$
	+	$\frac{4}{4!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-1)^{3n+l}\beta^{3n+l-8}P_{3n+l}(u)g(u)\frac{I_n^3I_l}{(n!)^3\ l!\sqrt{I_2^{3n+l}}}$
	+	$\frac{6}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-1)^{2(n+l)}\beta^{2(n+l)-8}P_{2(n+l)}(u)g(u)\frac{I_n^2I_l^2}{(n!)^2\ (l!)^2\sqrt{I_2^{2(n+l)}}}$
	+	$\frac{12}{4!}\sum_{n=3}^{\infty}\sum_{l\neq n}^{\infty}\sum_{m>l,m\neq n}^{\infty}(-1)^{2n+l+m}\beta^{2n+l+m-8}P_{2n+l+m}(u)g(u)\frac{I_n^2I_lI_m}{(n!)^2\ l!\ m!\sqrt{I_2^{2n+l+m}}}$
	+	$\frac{24}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}\sum_{p>m}^{\infty}(-1)^{n+l+m+p}\beta^{n+l+m+p-8}P_{n+l+m+p}(u)g(u)\frac{I_nI_lI_mI_p}{n!\ l!\ m!\ p!\sqrt{I_2^{n+l+m+p}}}$
≥ 5	+	...

$^{\#}$ I_n are the integrals of the function $\psi^n(s)$, which represents the wavelet function to the power of n:

$$I_n = \int_D \psi^n(s)ds.$$

The Gaussian function $g(u)$ is given by:

$$g(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}$$

$P_n(u)$ are the Hermite polynomials defined by Equation (5.9) and given by Table C.5.

Table C.1.: Probability density function $p_{\mathcal{W}}\beta,u)$.

j	$(\alpha \sum_{n=3}^{\infty}(-i2\pi\nu)^n \frac{I_n}{n!})^j$
2	$\sum_{n=3}^{\infty}(-i2\pi\nu)^{2n}\alpha^2\frac{I_n^2}{(n!)^2}$
	$+ \quad 2\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-i2\pi\nu)^{n+l}\alpha^2\frac{I_nI_l}{n!\,l!}$
3	$\sum_{n=3}^{\infty}(-i2\pi\nu)^{3n}\alpha^3\frac{I_n^3}{(n!)^3}$
	$+ \quad 3\sum_{n=3}^{\infty}\sum_{l=n,l\neq n}^{\infty}(-i2\pi\nu)^{2n+l}\alpha^3\frac{I_n^2I_l}{(n!)^2\,l!}$
	$+ \quad 6\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}(-i2\pi\nu)^{n+l+m}\alpha^3\frac{I_nI_lI_m}{n!\,l!\,m!}$
4	$\sum_{n=3}^{\infty}(-i2\pi\nu)^{4n}\alpha^4\frac{I_n^4}{(n!)^4}$
	$+ \quad 4\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-i2\pi\nu)^{3n+l}\alpha^4\frac{I_n^3I_l}{(n!)^3\,l!}$
	$+ \quad 6\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-i2\pi\nu)^{2(n+l)}\alpha^4\frac{I_n^2I_l^2}{(n!)^2\,(l!)^2}$
	$+ \quad 12\sum_{n=3}^{\infty}\sum_{l\neq n}^{\infty}\sum_{m>l,m\neq n}^{\infty}(-i2\pi\nu)^{2n+l+m}\alpha^4\frac{I_n^2I_lI_m}{(n!)^2\,l!\,m!}$
	$+ \quad 24\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}\sum_{p>m}^{\infty}(-i2\pi\nu)^{n+l+m+p}\alpha^4\frac{I_nI_lI_mI_p}{n!\,l!\,m!\,p!}$

Table C.2.: Development of $(\alpha \sum_{n=3}^{\infty}(-i2\pi\nu)^n \frac{I_n}{n!})^j$ for j=2,3 and 4

p		$\phi_{\mathcal{W}}(\nu)^{\#}$
0		$\phi_g(\nu)$
1	+	$\sum_{n=3}^{\infty}(-i2\pi\nu)^n\phi_g(\nu)\alpha\frac{I_n}{(n!)}$
2	+	$\frac{1}{2!}\sum_{n=3}^{\infty}(-i2\pi\nu)^{2n}\phi_g(\nu)\alpha^2\frac{I_n^2}{(n!)^2}$
	+	$\frac{2}{2!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-i2\pi\nu)^{n+l}\phi_g(\nu)\alpha^2\frac{I_nI_l}{n!\,l!}$
3	+	$\frac{1}{3!}\sum_{n=3}^{\infty}(-i2\pi\nu)^{3n}\phi_g(\nu)\alpha^3\frac{I_n^3}{(n!)^3}$
	+	$\frac{3}{3!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-i2\pi\nu)^{2n+l}\phi_g(\nu)\alpha^3\frac{I_n^2I_l}{(n!)^2\,l!}$
	+	$\frac{6}{3!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}(-i2\pi\nu)^{n+l+m}\phi_g(\nu)\alpha^3\frac{I_nI_lI_m}{n!\,l!\,m!}$
4	+	$\frac{1}{4!}\sum_{n=3}^{\infty}(-i2\pi\nu)^{4n}\phi_g(\nu)\alpha^4\frac{I_n^4}{(n!)^4}$
	+	$\frac{4}{4!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-i2\pi\nu)^{3n+l}\phi_g(\nu)\alpha^4\frac{I_n^3I_l}{(n!)^3\,l!}$
	+	$\frac{6}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-i2\pi\nu)^{2(n+l)}\phi_g(\nu)\alpha^4\frac{I_n^2I_l^2}{(n!)^2\,(l!)^2}$
	+	$\frac{12}{4!}\sum_{n=3}^{\infty}\sum_{l\neq n}^{\infty}\sum_{m>l,m\neq n}^{\infty}(-i2\pi\nu)^{2n+l+m}\phi_g(\nu)\alpha^4\frac{I_n^2I_lI_m}{(n!)^2\,l!\,m!}$
	+	$\frac{24}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}\sum_{p>m}^{\infty}(-i2\pi\nu)^{n+l+m+p}\phi_g(\nu)\alpha^4\frac{I_nI_lI_mI_p}{n!\,l!\,m!\,p!}$
≥ 5	+	...

$^{\#}$ I_n are the integrals of the function $\psi^n(s)$, which represents the wavelet function to the power of n:

$$I_n = \int_D \psi^n(s)ds.$$

$\phi_g(\nu)$ is the characteristic function associated to the Gaussian function

$$g(w) = \frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2\alpha I_2}}$$

Table C.3.: Characteristic function $\phi_{\mathcal{W}}(\nu)$ of the wavelet coefficient.

j		$p_{\mathcal{W}}(\alpha, w)^{\#}$
0		$g(w)$
1	+	$\sum_{n=3}^{\infty}(-1)^n g^{(n)}(w)\alpha\frac{I_n}{n!}$
2	+	$\frac{1}{2!}\sum_{n=3}^{\infty}(-1)^{2n} g^{(2n)}(w)\alpha^2\frac{I_n^2}{(n!)^2}$
	+	$\frac{2}{2!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-1)^{n+l} g^{(n+l)}(w)\alpha^2\frac{I_n I_l}{n!\,l!}$
3	+	$\frac{1}{3!}\sum_{n=3}^{\infty}(-1)^{3n} g^{(3n)}(w)\alpha^3\frac{I_n^3}{(n!)^3}$
	+	$\frac{3}{3!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-1)^{2n+l} g^{(2n+l)}(w)\alpha^3\frac{I_n^2 I_l}{(n!)^2\,l!}$
	+	$\frac{6}{3!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}(-1)^{n+l+m} g^{(n+l+m)}(w)\alpha^3\frac{I_n I_l I_m}{n!\,l!\,m!}$
4	+	$\frac{1}{4!}\sum_{n=3}^{\infty}(-1)^{4n} g^{(4n)}(w)\alpha^4\frac{I_n^4}{(n!)^4}$
	+	$\frac{4}{4!}\sum_{n=3}^{\infty}\sum_{l=3,l\neq n}^{\infty}(-1)^{3n+l} g^{(3n+l)}(w)\alpha^4\frac{I_n^3 I_l}{(n!)^3\,l!}$
	+	$\frac{6}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}(-1)^{2(n+l)} g^{(2(n+l))}(w)\alpha^4\frac{I_n^2 I_l^2}{(n!)^2\,(l!)^2}$
	+	$\frac{12}{4!}\sum_{n=3}^{\infty}\sum_{l\neq n}^{\infty}\sum_{m>l,m\neq n}^{\infty}(-1)^{2n+l+m} g^{(2n+l+m)}(w)\alpha^4\frac{I_n^2 I_l I_m}{(n!)^2\,l!\,m!}$
	+	$\frac{24}{4!}\sum_{n=3}^{\infty}\sum_{l>n}^{\infty}\sum_{m>l}^{\infty}\sum_{p>m}^{\infty}(-1)^{n+l+m+p} g^{(n+l+m+p)}(w)\alpha^4\frac{I_n I_l I_m I_p}{n!\,l!\,m!\,p!}$
≥ 5	+	$\ldots$

$^{\#}$ I_n are the integrals of the function $\psi^n(s)$, which represents the wavelet function to the power of n:

$$I_n = \int_D \psi^n(s)ds.$$

$g^{(n)}(w)$ is the nth derivative of the Gaussian function

$$g(w) = \frac{1}{\sqrt{2\pi}}e^{-\frac{w^2}{2\alpha I_2}}$$

Table C.4.: Probability density function $p_{\mathcal{W}}(\alpha, w)$.

	u^0	u^1	u^2	u^3	u^4	u^5	u^6	u^7
P_0	1							
P_1		-1						
P_2	-1		1					
P_3		3		-1				
P_4	3		-6		1			
P_5		-15		10		-1		
P_6	-15		45		-15		1	
P_7		105		-105		21		-1
P_8	105		-420		210		-28	
P_9		-945		1260		-378		36
P_{10}	-945		4725		-3150		630	
P_{11}		10395		-17325		6930		-990
P_{12}	10395		-62370		51975		-13860	
P_{13}		-135135		270270		-135135		25740
P_{14}	-135135		945945		-945945		315315	
P_{15}		2027025		-4729725		2837835		-675675
P_{16}	2027025		-16216200		18918900		-7567560	
P_{17}		-34459425		91891800		-64324260		18378360
P_{18}	-34459425		310134825		-413513100		192972780	

	u^8	u^9	u^{10}	u^{11}	u^{12}	u^{13}	u^{14}	u^{15}	u^{16}	u^{17}	u^{18}
P_0											
P_1											
P_2											
P_3											
P_4											
P_5											
P_6											
P_7											
P_8	1										
P_9		-1									
P_{10}	-45		1								
P_{11}		55		-1							
P_{12}	1485		-66		1						
P_{13}		-2145		78		-1					
P_{14}	-45045		3003		-91		1				
P_{15}		75075		-4095		105		-1			
P_{16}	1351350		-120120		5460		-120		1		
P_{17}		-2552550		185640		-7140		136		-1	
P_{18}	-41351310		4594590		-278460		9180		-153		1

Table C.5.: Hermite Polynomials $P_n(u)$ defined by Equation (5.9)

	a	b	c	d	e	f
$P_3(u)$	$-\frac{I_3}{6\sqrt{I_2^3}}$					
$P_4(u)$		$\frac{I_4}{24I_2^2}$				
$P_5(u)$			$-\frac{I_5}{120\sqrt{I_2^5}}$			
$P_6(u)$		$\frac{I_3^2}{72I_2^3}$		$\frac{I_6}{720I_2^3}$		
$P_7(u)$			$-\frac{I_3I_4}{144\sqrt{I_2^7}}$		$-\frac{I_7}{5040\sqrt{I_2^7}}$	
$P_8(u)$				$\frac{I_4^2}{1152I_2^4}+\frac{I_3I_5}{720I_2^4}$		$\frac{I_8}{40320I_2^4}$
$P_9(u)$			$-\frac{I_3^3}{1296\sqrt{I_2^9}}$		$-\frac{I_3I_6}{4320\sqrt{I_2^9}}-\frac{I_4I_5}{2880\sqrt{I_2^9}}$	
$P_{10}(u)$				$\frac{I_3^2I_4}{1728I_2^5}$		$\frac{I_5^2}{28800I_2^5}+\frac{I_3I_7}{30240I_2^5}+\frac{I_4I_6}{17280I_2^5}$
$P_{11}(u)$					$-\frac{I_3^2I_5}{8640\sqrt{I_2^{11}}}-\frac{I_4^2I_3}{6912\sqrt{I_2^{11}}}$	
$P_{12}(u)$				$\frac{I_3^4}{31104I_2^6}$		$\frac{I_4^3}{82944I_2^6}+\frac{I_3^2I_6}{51840I_2^6}+\frac{I_3I_4I_5}{17280I_2^6}$
$P_{13}(u)$					$-\frac{I_3^3I_4}{31104\sqrt{I_2^{13}}}$	
$P_{14}(u)$						$\frac{I_3^3I_5}{155520I_2^7}+\frac{I_3^2I_4^2}{82944I_2^7}$
$P_{15}(u)$					$-\frac{I_3^5}{933120\sqrt{I_2^{15}}}$	
$P_{16}(u)$						$\frac{I_3^4I_4}{746496I_2^8}$
$P_{17}(u)$						
$P_{18}(u)$						$\frac{I_3^6}{33592320I_2^9}$

Table C.6.: Coefficients of Equation (5.11). This table is to be read columnwise as shown on the following example: $b = \frac{I_4}{24I_2^2}P_4(u) + \frac{I_3^2}{72I_2^3}P_6(u)$

α	ϵ				
	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
Thresholds$^{\#}$					
$\alpha \leq 2^{-19}$	1	1	1	1	1
2^{-18}	2	1	1	1	1
2^{-17}	2	1	1	1	1
2^{-16}	2	1	1	1	1
2^{-15}	2	2	1	1	1
2^{-14}	2	2	1	1	1
2^{-13}	2	2	1	1	1
2^{-12}	2	2	2	1	1
2^{-11}	2	2	2	1	1
2^{-10}	2	2	2	1	1
2^{-9}	2	2	2	1	1
2^{-8}	3	2	2	2	1
2^{-7}	3	2	2	2	1
2^{-6}	3	3	2	2	1
2^{-5}	3	3	3	2	2
2^{-4}	4	3	3	2	2
2^{-3}	4	4	3	3	2
2^{-2}	5	4	4	3	2
2^{-1}	6	5	5	4	3
Normalized Thresholds*					
1	7	6	6	5	3
2^{1}	6,36	5,66	4,95	4,24	2,83
2^{2}	6	5	4,5	3,5	3
2^{3}	5,30	4,95	4,24	3,53	2,83
2^{4}	5,25	4,75	4	3,5	2,5
2^{5}	4,95	4,60	3,89	3,36	2,47
2^{6}	4,87	4,37	3,87	3,25	2,5
2^{7}	4,86	4,33	3,80	3,18	2,39
2^{8}	4,81	4,31	3,81	3,12	2,37
2^{9}	4,82	4,33	3,76	3,14	2,39
2^{10}	4,78	4,31	3,75	3,12	2,34
2^{11}	4,77	4,29	3,73	3,12	2,34
2^{12}	4,77	4,28	3,73	3,11	2,34
2^{13}	4,76	4,28	3,73	3,10	2,34
2^{14}	4,76	4,27	3,73	3,10	2,34
2^{15}	4,76	4,27	3,72	3,10	2,33
$\alpha \geq 2^{16}$	4,75	4,26	3,72	3,09	2,33

$^{\#}$ defined by Equation (5.23)
* defined by Equation (5.24)

Table C.7.: Thresholds corresponding to $\alpha \in [2^{-30}; 2^{30}]$

List of notations

Symbols

$\triangleq$	definition symbol
x^*	complex conjugate of x
$\hat{x}$, $\hat{X}$	estimate of x respectively X
$M^{(n)}$	matrix M at iteration n
$f^{(n)}$	nth derivate of the function f
f^n, M^n	function f respectively matrix M to the power of n
$\mathrm{E}[X]$	expectation of the random variable X
$\mathrm{Var}[X]$	variance of the random variable X
M^t	transpose of the matrix M

Abbreviations

CR	compression ratio
CWT	continuous wavelet transform
D-method	denoising method
DWT	discrete wavelet transform
EP	edge preserving
ET	editable text
FIR	finite impulse response

fwhm	full width at half maximum
GUI	graphical user interface
iid	independent identically distributed
MEM	maximum entropy method
MD-WT	maximally decimated wavelet transform - used here for the Mallat algorithm
NaI(Tl)	thallium-activated sodium iodide
NL	noise level
NS	set of points where the wavelet coefficients of the data have been found to be non significant
PB	push button
PMT	photomultiplier tube
PHA	pulse height analyzer
Prob	probability
PRESS	predictive sum of squares
PSF	point spread function
PUP	pop-up menu
RA	relative amplitude
RB	radio button
R-method	restoration method
S	set of points where the wavelet coefficients of the data have been found to be significant
SI-WT	shift invariant wavelet transform - used here for Coifman and Donoho's algorithm
SLI	slider
SMC	summing matrix circuits
SNR	signal to noise ratio

SWC	significant wavelet coefficients
Tc^{99m}	metastable Technetium
TI	Translation invariant
TIPSH	translation invariant Poisson smoothing using Haar wavelets
SI	shift invariant

DeQuant algorithms

D-DeQuant	denoising DeQuant
EP-D-DeQuant	edge preserving denoising DeQuant
SI-D-DeQuant	shift invariant denoising DeQuant
SI-EP-D-DeQuant	shift invariant edge preserving denoising DeQuant
R-DeQuant	restoration DeQuant
EP-R-DeQuant	edge preserving restoration DeQuant
SI-R-DeQuant	shift invariant denoising DeQuant
SI-EP-R-DeQuant	shift invariant edge preserving restoration DeQuant

Image degradation model

$I(k_x, k_y)$	number of photon counts detected at position (k_x, k_y) of the image during the imaging time T.
P	operator associated with the point spread function of the imaging system. It can be seen as the probability that a photon that would be detected at a given position by an ideal imaging system is detected at an other position by the real camera.
Q	PSF corresponding to the interaction in the patient body. It also takes the imaging time into account.
T	imaging time
$U(x, y, z)$	local uptake density. It is defined by the number of radionuclides taken at each point of the imaged organ.

X	image of the uptake density that would be given by an ideal imaging system.
Y	distribution of photons after that has been spread by the PSF of the camera.
$\rho(x, y)$	continuous intensity function.

Multiresolution analysis

$L^2(\mathbb{R})$, $L^2(\mathbb{R}^2)$	spaces of square integrable functions
V_{2^j}	coarse approximation spaces
O_{2^j}	detail spaces
$\phi(x)$	scaling function
$\psi(x)$	wavelet or mother wavelet
$\psi_{(a,b)}(x)$	wavelet at scale a and position b
$\psi^h_{(2^j,k_x,k_y)}(x, y)$	horizontal wavelet at scale 2^j and position (k_x, k_y)
$\psi^v_{(2^j,k_x,k_y)}(x, y)$	vertical wavelet at scale 2^j and position (k_x, k_y)
$\psi^d_{(2^j,k_x,k_y)}(x, y)$	diagonal wavelet at scale 2^j and position (k_x, k_y)
a	scale parameter of the wavelet
b	position parameter of the wavelet
$h_0(n)$	defined by Equation (2.12)
$h_1(n)$	defined by Equation (2.16)
$h(n)$	lowpass analysis filter
$g(n)$	highpass analysis filter
$\tilde{h}(n)$	lowpass synthesis filter
$\tilde{g}(n)$	highpass synthesis filter
$\downarrow 2$	keep one sample out of two
$\uparrow 2$	put one zero between each sample

1D Mallat algorithm

$F_{f(x)}(2^j, k)$	approximation coefficient or scaling coefficient of $f(x)$ at scale 2^j and position k.
$W_{f(x)}(2^j, k)$	wavelet coefficient of $f(x)$ at scale 2^j and position k.
$F_{f(x)}(2^{j+1})$, $F_{F_{f(x)}(2^j)}$	vector of approximation coefficients at scale 2^{j+1}
$W_{f(x)}(2^{j+1})$, $W_{F_{f(x)}(2^j)}$	vector of wavelet coefficients at scale 2^{j+1}
H, G	analysis operators for the 1D Mallat algorithm
$\widetilde{H}$, $\widetilde{G}$	synthesis operators for the 1D Mallat algorithm

2D Mallat algorithm

$F_{f(x)}(2^j, k_x, k_y)$	approximation or scaling coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y).
$W_{h,f(x,y)}(2^j, k_x, k_y)$	horizontal wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y).
$W_{v,f(x,y)}(2^j, k_x, k_y)$	vertical wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y).
$W_{d,f(x,y)}(2^j, k_x, k_y)$	diagonal wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y).
$F_{f(x,y)}(2^{j+1})$, $F_{F_{f(x,y)}(2^j)}$	matrix of approximation coefficient at scale 2^{j+1}
$W_{h,f(x,y)}(2^{j+1})$, $W_{h,F_{f(x,y)}(2^j)}$	matrix of horizontal wavelet coefficients at scale 2^{j+1}
$W_{v,f(x,y)}(2^{j+1})$, $W_{h,F_{f(x,y)}(2^j)}$	matrix of vertical wavelet coefficients at scale 2^{j+1}
$W_{d,f(x,y)}(2^{j+1})$, $W_{h,F_{f(x,y)}(2^j)}$	matrix of diagonal wavelet coefficients at scale 2^{j+1}
H_2, G_h, G_v, G_d	analysis operators for the 2D Mallat algorithm
$\widetilde{H_2}$, $\widetilde{G_h}$, $\widetilde{G_v}$, $\widetilde{G_d}$	synthesis operators for the 2D Mallat algorithm.

Shift invariant algorithm

$F^T_{f(x)}(2^j, k_x, k_y)$	approximation coefficient or scaling coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y) calculated with the shift invariant algorithm
$W^T_{h,f(x,y)}(2^j, k_x, k_y)$	horizontal wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y) calculated with the shift invariant algorithm
$W^T_{v,f(x,y)}(2^j, k_x, k_y)$	vertical wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y) calculated with the shift invariant algorithm
$W^T_{d,f(x,y)}(2^j, k_x, k_y)$	diagonal wavelet coefficient of $f(x, y)$ at scale 2^j and position (k_x, k_y) calculated with the shift invariant algorithm
$F^T_{f(x,y)}(2^{j+1})$, $F^T_{F^T_{f(x,y)}(2^j)}$	matrix of approximation coefficient at scale 2^{j+1} calculated with the shift invariant algorithm
$W^T_{h,f(x,y)}(2^{j+1})$, $W^T_{h,F^T_{f(x,y)}(2^j)}$	matrix of horizontal wavelet coefficients at scale 2^{j+1} calculated with the shift invariant algorithm.
$W^T_{v,f(x,y)}(2^{j+1})$, $W^T_{h,F^T_{f(x,y)}(2^j)}$	matrix of vertical wavelet coefficients at scale 2^{j+1} calculated with the shift invariant algorithm.
$W^T_{d,f(x,y)}(2^{j+1})$, $W^T_{h,F^T_{f(x,y)}(2^j)}$	matrix of diagonal wavelet coefficients at scale 2^{j+1} calculated with the shift invariant algorithm.
$H_2(2^j)$, $G_h(2^j)$, $G_v(2^j)$, $G_d(2^j)$	analysis operators for the 2D shift invariant algorithm
$\widetilde{H_2}(2^j)$, $\widetilde{G_h}(2^j)$, $\widetilde{G_v}(2^j)$, $\widetilde{G_d}(2^j)$	synthesis operators for the 2D shift invariant algorithm.

Classical approaches to image restoration

$f(k_x, k_y)$	input image or object
$g(k_x, k_y)$	degraded or noisy image
$h(k_x, k_y)$	point spread function of the degradation process

$n(k_x, k_y)$	additive noise term
$F(u, v)$	Fourier transform of the object $f(k_x, k_y)$
$G(u, v)$	Fourier transform of the noisy image $g(k_x, k_y)$
$H(u, v)$	Fourier transform of the PSF $h(k_x, k_y)$
$N(u, v)$	Fourier transform of the additive noise term $n(k_x, k_y)$
$\mathbf{f}$	column vector formed by stacking the rows of $f(k_x, k_y)$
$\mathbf{g}$	column vector formed by stacking the rows of $g(k_x, k_y)$
$\mathbf{H}$	block circulant matrix constructed from $h(k_x, k_y)$
$\mathbf{n}$	column vector formed by stacking the rows of $n(k_x, k_y)$

DeQuant: a flexible multiresolution restoration framework

δ	tolerance parameter that has been introduced to take into account the noise affecting the significant wavelet coefficients.
ϵ	significance level of the hypothesis test
ϵ_0, ϵ_1	convergence parameters
ξ_{L_2}	parameter of the Van Cittert iteration with the Laplacian operator
ξ_M	parameter of the Van Cittert iteration with the operator M
D_x	horizontal gradient operator
D_y	vertical gradient operator
J	scale to which the wavelet transform is calculated
$\mathbf{k}_s$	coordinates (k_x, k_y) of the points where the wavelet coefficients of the image I have been found to be significant.
$\mathbf{k}_{ns}$	coordinates (k_x, k_y) of the points where the wavelet coefficients of the image I have been found to be non significant.
L_2	Laplacian operator
M	operator introduced by Bijaoui and Froeschle to determine image edges

$M^s_{a,I}(2^j, k_x, k_y)$	mask of the significant wavelet coefficients of the image I. a=h, v or d.
$\overline{M}^s_{a,I}(2^j, k_x, k_y)$	mask of the non significant wavelet coefficients of the image I. a=h, v or d.
$t(2^j, k_x, k_y)$	threshold at scale 2^j and position (k_x, k_y)
$W_{a,I}(2^j, k_x, k_y)$	general notation for the calculated wavelet coefficient of I. The parameter a takes the value h for the horizontal, v for the vertical and d for the diagonal wavelet coefficients.
$W^s_{a,I}(2^j, k_x, k_y)$	significant wavelet coefficient. a=h, v or d.
$W^{ns}_{a,I}(2^j, k_x, k_y)$	non significant wavelet coefficient. a=h, v or d.
W^s_I	set of significant wavelet coefficient.
W^{ns}_I	set of non significant wavelet coefficient.
$\mathcal{W}_{a,I}(2^j, k_x, k_y)$	random variable corresponding to the wavelet coefficient of the image I

PDF of the wavelet coefficient of a Poisson process

α	mean number of events in the interval a. $\alpha = \rho_0 a$
β	defined as $\frac{1}{\sqrt{\alpha}}$
$\phi_{\mathcal{W}}(\nu)$	characteristic function of $\mathcal{W}$
$\phi_g(\nu)$	characteristic function associated to the Gaussian function $g(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2\alpha I_2}}$
a	interval on which the function $\rho(x)$ has been supposed to be constant and equal to ρ_0.
D	compact support of the wavelet. For simplicity D has been taken equal to 1.
$F(\alpha, w)$	repartition function of the wavelet coefficcient
$g(u)$	Gaussian function $g(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$
$H(\psi)$	normalized histogram of the wavelet ψ

$H_{\psi}^{(*n)}(w)$	denotes n autoconvolutions of $H_{\psi}(w)$
$H_{\psi}(w)(\nu)$	Fourier transform of the normalized histogram of the wavelet ψ
I_n	integrals of the function ψ^n, which represents the wavelet function to the power of n.
$I_n(\alpha)$	modified Bessel function
$p_{\mathcal{W}}(\alpha, w)$, $p_{\mathcal{W}}(\beta, u)$	PDF of $\mathcal{W}_{a,I}(2^j, k_x, k_y)$
$P_n(u)$	Hermite polynomials
u	centered and normalized version of w: $u = \frac{w - E(\mathcal{W})}{\sqrt{Var(\mathcal{W})}}$
w	variable used in the PDF
W	calculated wavelet coefficients
$\mathcal{W}$	random variable associated with the wavelet coefficients

Application of DeQuant to nuclear medicine

(x_G, y_G)	coordinates of the object's center of gravity
j	current scale of the analysis
J_{Max}	equals $2\log_2(\sqrt{N_I})$
N_I	number of pixels of the image I

Publications

Journal papers

- On the distribution of the wavelet coefficient for a Poisson noise.
  A. Bijaoui and G. Jammal.
  accepted by the journal *Signal Processing*, December 2000 .

- DeQuant: a flexible multiresolution restoration framework.
  G. Jammal and A. Bijaoui.
  submitted to the journal *Signal Processing*, September 2000.

Conference proceedings

- Multiscale image restoration for photon imaging systems.
  G. Jammal and A. Bijaoui.
  Proceedings of SPIE Medical Imaging, Volume 3661, Pages 1180-1189, February 1999.

- A multiresolution image restoration method for photon imaging systems.
  G. Jammal and A. Bijaoui.
  Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 6, Pages 3401- 3404, March 1999.

- Wavelet based filtering in nuclear medicine.
  G. Jammal and A. Bijaoui.
  Proceedings of the IEE seventh international conference on Image Processing and its Applications, Volume 2, Pages 851-855, July 1999.

- Regularized image restoration in nuclear medicine.
  G. Jammal and A. Bijaoui.
  Proceedings of SPIE annual meeting, Wavelet Applications in Signal and Image Processing VII, July 1999.

- Restauration et débruitage d'images en médecine nucléaire : l'approche multirésolution.
  G. Jammal and A. Bijaoui.
  Proceedings of 17ème colloque GRETSI on signal and image processing, September 1999.

- Denoising and Deconvolution in nuclear medicine.
  G. Jammal and A. Bijaoui.
  Proceedings of IEEE International Conference on Image Processing, October 1999.

International fairs

Presentation of DeQuant at the Hannover Fair 2000

- DeQuant: New method increases diagnostic confidence in nuclear medicine.
  A. Bijaoui, H. Clausert and G. Jammal.
  Journal *"Hessen- hier ist die Zukunft"*, Hannover Fair, March 2000.

Internal reports

- Wavelets, a tutorial.
  F. Bock and G. Jammal.
  Technische Universität Darmstadt, Institut für Netzwerk und Signaltheorie, May 1997.

Internet publication on commercial websites

The Matlab Graphical User Interface that we developed for DeQuant is to be published in April 2001 on the MathWorks (editor of Matlab) website as an example of the use of Matlab in biotechnologies.

Bibliography

[Ans48] F. Anscombe. The transformation of Poisson, binomial and negative binomial data. *Biometrika*, 35:246–254, 1948.

[AS64] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Dover, 1964.

[BB97] Y. Bobichon and A. Bijaoui. A regularized image restoration algorithm for lossy compression in astronomy. *Experimental astronomy*, 7(3):239–255, 1997.

[Bec73] P. Beckmann. *Orthogonal Polynomials for Engineers and Physicists*. The GOLEM Press, 1973.

[BF80] A. Bijaoui and M. Froeschlé. A new algorithm to determine image edges. *Astron. Astrophys.*, 87:250–251, 1980.

[BGG98] C.S. Burrus, R.A. Gopinath, and H. Guo. *Introduction to wavelets and wavelet transforms: a primer*. Prentice-Hall, Inc., 1998.

[BJ00] A. Bijaoui and G. Jammal. On the distribution of the wavelet coefficient for a Poisson noise. *submitted to Signal Processing*, August 2000.

[Bob97] Y. Bobichon. *Restauration d'image en Multirésolution*. PhD thesis, Université de la Nice Sophia Antipolis, 1997.

[Bou00] L. Bourdichon. *Vergleich von zwei Bilrestaurationsverfahren für die Nuklearmedizin*. Final year project, Darmstadt University of Technology, 2000.

[CD95a] R.R. Coifman and D. Donoho. Translation invariant denoising. *Technical Report 475, Dept. of Statistics, Stanford University*, May 1995.

[CD95b] R.R. Coifman and D.L. Donoho. *Translation invariant denoising, In Wavelets and statistics*. A. Antoniadis, ed. Springer Verlag, Heidelberg, 1995.

[Dau92] I. Daubechies. *Ten lectures on wavelets*. SIAM, Philadelphia, Pennsylvania, 1992.

[Dem89] G. Demoment. Image reconstruction and restoration: overview of common estimation structures and problems. *IEEE transactions on acoustics, speech and signal processing*, 37(12):2024–2036, Dec 1989.

[DJ94] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[DJKP95] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: asymptopia? *Journal Royal Statistical Society B.*, 57(2):301–337, 1995.

[Don75] D.L. Donoho. De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627, May 1975.

[Don93a] D.L. Donoho. Nonlinear wavelet methods for recovery of signals, densities and spectra from indirect and noisy data. *In Proc. Sympos. Appl. Math.*, 47, 1993.

[Don93b] D.L. Donoho. Wavelet shinkage w.v.d - a ten minutes tour. *Technical report TR-416, Statistics Department, Stanford University*, January 1993.

[Eca99] O. Ecabert. *Multiscale image restoration for photon imaging systems*. Final year project, Darmstadt University of Technology, 1999.

[GM52] A. Gray and G.B. Mathews. *A treatise on Bessel functions and their applications to physics*. Macmillan and co. London, 1952.

[Gro89] Grossmann. *Reading and understanding continuous wavelet transform, in J.M Combes, A. Grossmann and P. Tchamitchian editors, Wavelets, time frequency methods and phase space*. Springer-Verlag, Berlin, 1989.

[GW93] R.C. Gonzales and R.E. Woods. *Digital image processing*. Addison Wesley publishing company, 1993.

[Had23] J. Hadamard. *Lectures on the Cauchy problem in linear partial differential equations*. Yale University Press, 1923.

[Hel67] C.W. Helstrom. Image restoration by the method of least squares. *J. Opt. Soc. Am.*, 57:297–303, 1967.

[IN67] T.A. Iiuma and T. Nagai. Image restoration in radioisotope imaging systems. *Phys. med. biol.*, 12(4):501–509, 1967.

[Jai89] A.K. Jain. *Fundamentals of digital image processing.* Prentice Hall, 1989.

[Jay57] E.T. Jaynes. *Physical Review*, (106):620–630, 1957.

[JB99a] G. Jammal and A. Bijaoui. Denoising and deconvolution in nuclear medicine. *Proceedings of the IEEE International Conference on Image Processing*, October 1999.

[JB99b] G. Jammal and A. Bijaoui. A multiresolution image restoration method for photon imaging systems. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 4:3401–3404, Mars 1999.

[KE90] A.K. Katsaggelos and S.N. Estradiadis. A class of iterative signal restoration algorithms. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(5):778–786, May 1990.

[Kol97] E.D. Kolaczyk. Nonparametric estimation of gamma-ray burst intensities using haar wavelets. *The astrophysical journal*, 483:340–349, July 1997.

[Kol99] E.D. Kolaczyk. Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds. *Statistica Sinica*, 9:119–135, 1999.

[Lan51] L. Landweber. An iteration formula for fredholm integral equations of the first kind. *American journal of mathematics*, 73:615–624, 1951.

[LLR83] M.A. Leadbetter, G. Lindgrea, and H. Rootzén. *Extremes and related properties of random sequences and processes.* Springer Verlag, New York, 1983.

[LMR97] A.K. Louis, P. Maaß, and A. Rieder. *Wavelets, theory and applications.* John Wiley and Sons, 1997.

[Luc74] L.B Lucy. An iteration technique for the rectification of observed distributions. *Astronomical journal*, 79:745–754, 1974.

[Mal89a] S. Mallat. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. on acoustics, speech and signal processing*, 37(12):2091–2110, Dec. 1989.

[Mal89b] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on pattern analysis and machine intelligence*, 11(7):674–693, July 1989.

[Mal97] S. Mallat. *A wavelet tour of signal processing*. Academic Press, 1997.

[Mar96] P. Marchand. *Graphics and GUIs with Matlab*. CRC Press, 1996.

[Mey86] Y. Meyer. Principe d'incertitude, bases hilbertiennes et algèbres d'opérateurs. *Borbaki seminar*, (662), 1985–1986.

[MHT89] J. Morlet M. Holschneider, R. Kroland-Martinet and P. Tchamitchian. A real time algorithm for signal analysis with the help of the wavelet transform. *Wavelets, Time-Frequency methods and Phase Space*, Springer-Verlag, Berlin, 1989.

[Moo91] A.S. Moore. *Statistics concepts and controversies*. W.H. Freeman and Company New York, 1991.

[NB97] R.D. Nowak and R.G. Baraniuk. Wavelet-domain filtering for photon imaging systems. *Proc. SPIE, Wavelet Applications in Signal and Image Processing V*, 3169:55–66, August 1997.

[Now96] R.D. Nowak. Optimal signal estimation using cross-validation. *IEEE Signal Processing Letters*, 3(12):23–25, 1996.

[OS89] A.V. Oppenheim and R.W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[Pap84] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill, Inc., 1984.

[PTK85] T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 317(26):314–319, September 1985.

[RB98] R.M. Rao and A.S. Bopardikar. *Wavelet transforms: Introduction to theory and applications*. Addison-Wesley, 1998.

[Sap90] G. Saporta. *Probabilités, Analyse des données et Statistiques*. Editions Technip, 1990.

[Sha48] C.E. Shannon. *Bell Systems Technical Journal*, (27):379 and 623, 1948.

[SHW93] D.L. Snyder, A.M. Hammoud, and R.L. White. Image recovery from data acquired with a charge coupled device camera. *J. Opt. Soc. Am.*, 10(5):1014–1023, 1993.

[SM99] J.L. Starck and F. Murtagh. Multiscale entropy filtering. *Signal Processing*, (76):147–165, 1999.

[SMB95] J.L. Starck, F. Murtagh, and A. Bijaoui. Multiresolution support applied to image filtering and restoration. *Graphical Models and Image Processing*, 57(5):420–431, 1995.

[SMB98] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis, The multiscale approach*. Cambridge University Press, 1998.

[SO] A. Stuart and J. Keith Ord. *Kendall's advanced Theory of Statistics*. Volume 1.

[SP87] J.A. Sorenson and M.E. Phelps. *Physics in nuclear medicine*. W.B. Saunders Company, 1987.

[SP96] J.L. Starck and E. Pantin. Multiscale maximum entropy images restoration. *Vistas in Astronomy*, 40(4):563–569, 1996.

[SST92] K.K. Shung, M.B. Smith, and B. Tsui. *Principles of medical imaging*. Academic Press, 1992.

[Swe96] W. Sweldens. Wavelets: What next? *Proceedings of the IEEE*, 84(4):680–685, April 1996.

[Tik63] A.N. Tikhonov. Regularization of incorrectly posed problems. *Sov. Math. Dokl.*, 4:1624–1627, 1963.

[TN99] K. Timmermann and R. Nowak. Multiscale modeling and estimation of poisson processes with application to photon-limited imaging. *IEEE transaction on Information Theory*, 45(3):846–862, April 1999.

[VC31] P.H. Van-Cittert. Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien II. *Zeitschrift für Physik*, 69:298–308, 1931.

[VK95] Martin Vetterli and Jelena Kovacevic. *Wavelets and Subband Coding*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1995.

[Wic94] M.V. Wickerhauser. *Adapted wavelet analysis from theory to software*. A K Peter, Wellesley, Massachussets, 1994.

[ZBH75] H.J. Zweig, E.B. Barrett, and P.C. Hu. Noise-cheating image enhancement. *J. Opt. Soc. Am.*, 65(11):1347–1353, November 1975.

Lebenslauf

Zur Person

Name	Ghada JAMMAL
Geboren am	26. Juli 1971 in Beirut
Staatsangehörigkeit	libanesisch

Ausbildung

1986-1989	Gymnasium: Lycée Franco-Libanais, Beirut, Libanon.
09.1989	Abschluß: Abitur mit Schwerpunkt Naturwissenschaften.
1989-1992	Vordiplom (Math. sup. und Math. spé.) an dem Lycée Saint-Louis, Paris, Frankreich. Studienschwerpunkt: Mathematik und Physik.
1992-1995	Diplom an der Ecole Nationale Supérieure de l'Electronique et de ses Applications (ENSEA - Technische Hochschule), Cergy-Pontoise, Frankreich. Studienschwerpunkt: Elektrotechnik - Signalverarbeitung. Gesamtnote: sehr gut.
09.1995	Abschluß: Diplom-Ingenieur (TH), Elektrotechnik.
1996-2000	Promotion an der Technischen Universität Darmstadt.

Sprachkenntnise

Deutsch	sehr gut in Wort und Schrift. 5 Jahre Aufenthalt in Deutschland.
Englisch	sehr gut in Wort und Schrift. 'Certificate of Advanced English' (Cambridge University) Mehrere Aufenthalte in Nordamerika und in England.
Französisch	Muttersprache.
Arabisch	Muttersprache.

Berufstätigkeiten

07.1993 - 08.1993	Zwei Monate Praktikum bei der Fa. THOMSON-CSF in Versailles, Frankreich: Erstellung von Schulungsunterlagen über Radargleichungen.
07.1994	Ein Monat Praktikum bei der Fa. THOMSON-CSF in Versailles, Frankreich: Erstellung von Schulungsunterlagen über den Simulator einer Radarstation.
05.1995 - 07.1995	Drei Monate Diplomarbeit am Imperial College in London, England: New non-gaussianity tests based on Hermite Polynomials. Gesamtnote: sehr gut.
1996-2000	Wissenschaftlische Mitarbeiterin am Institut für Nachrichtentechnik der Technischen Universität Darmstadt. Aufgaben: Unterstützung des Fachgebietsleiters bei der Vorbereitung und Durchführung von Vorlesungen, Übungen und Prüfungen. Betreuung von Studien-und Diplomarbeiten. Beratung der Studierenden im Lernzentrum.