Visibility Recovery on Images Acquired in Attenuating Media.

Application to Underwater, Fog, and Mammographic Imaging

Adrian Galdran

Department of Mathematics
University of the Basque Country (UPV/EHU)

Tecnalia Research and Innovation

Supervised by: David Pardo (UPV/EHU) and Artzai Picón (Tecnalia)

A Dissertation submitted for the degree of

Doctor of Philosophy

Derio, 2015

To Itzi and Luppo

Acknowledgements

These pages are the product of many hours of a hard work. To me, a Ph.D. is like a labyrinth full of tricks and dead-ends; it is hard to find one's way out of it, and I have made great effort until I found it. Nonetheless, it is clear for me that this dissertation does not exist only because of my own sweat. Many people have contributed in different aspects. First of all, Artzai Picón and David Pardo must be acknowledged. Artzai's optimism and enthusiasm is something many of us could learn from. Besides, he is the researcher with the best intuition I have had the pleasure to work with; he grasps deep insights of complex ideas just by glancing over a few equations. On David's side fall the uncountable hours he has invested in the "review-review-review again" process of everything I have written, or listening repeatedly to the same talk essays, which has no price. I must also thank him for teaching me many useful secondary skills that are essential to develop professional research.

I am also glad of having written this dissertation in a research center like Tecnalia. It has been a perfect environment to obtain a balanced mix of the different sides of the R&D world. People at Tecnalia have been really neighbourly all this time, creating a friendly and inspiring place for research. In that sense, I must thank Jone for all the energy she radiates. I am also deeply indebted to all my lab-mates, who have contributed to create an excellent working environment: Aitor, Arantza, Sergio, Estibaliz, Ana, and a large etcetera.

Innovating in such a classical and saturated field as Image Processing is a hard task. I have often found myself thinking of an idea that has come to me at night, or in the train to work, and I could not avoid thinking of how many people would have thought about that very same idea before. In that sense, the short courses given by Vicent Caselles at UPV/EHU on 2012 and Marcelo Bertalmío at BCAM on winter of 2014, together with the visit I paid to his lab on spring of that year, have been key to understand the ideas behind modern Image Processing and the Retinex theory, and how they could be implemented and adapted to different purposes. Chapter 4 and probably 5 have probably come to life thanks to that visit. All his lab at the Pompeu Fabra University gave me a really nice welcome. Special thanks must go to Javier Vazquez-Corral, who enthusiastically contributed to design, test and evaluate the variational image dehazing algorithm I present in chapter 4.

In the personal side of this work, friends and family have also been essential. I must thank my parents Ramon and Carmen, and my brother Ramon, my all-life friends in Alicante and Valencia and my friends and flat-mates at Bilbao. And of course, special thanks go to Itzi, for her continuous support, love, and happiness. Without her, this would not have been possible at all.

Abstract

When acquired in attenuating media, digital images often suffer from a particularly complex degradation that reduces their visual quality, hindering their suitability for further computational applications, or simply decreasing the visual pleasantness for the user. In these cases, mathematical image processing reveals itself as an ideal tool to recover some of the information lost during the degradation process. In this dissertation, we deal with three of such practical scenarios in which this problematic is specially relevant, namely, underwater image enhancement, fog removal and mammographic image processing. In the case of digital mammograms, X-ray beams traverse human tissue, and electronic detectors capture them as they reach the other side. However, the superposition on a bidimensional image of three-dimensional structures produces low-contrasted images in which structures of interest suffer from a diminished visibility, obstructing diagnosis tasks. Regarding fog removal, the loss of contrast is produced by the atmospheric conditions, and white colour takes over the scene uniformly as distance increases, also reducing visibility. For underwater images, there is an added difficulty, since colour is not lost uniformly; instead, red colours decay the fastest, and green and blue colours typically dominate the acquired images. To address all these challenges, in this dissertation we develop new methodologies that rely on: a) physical models of the observed degradation, and b) the calculus of variations. Equipped with this powerful machinery, we design novel theoretical and computational tools, including image-dependent functional energies that capture the particularities of each degradation model. These energies are composed of different integral terms that are simultaneously minimized by means of efficient numerical schemes, producing a clean, visually-pleasant and useful output image, with better contrast and increased visibility. In every considered application, we provide comprehensive qualitative (visual) and quantitative experimental results to validate our methods, confirming that the developed techniques outperform other existing approaches in the literature.

Contents

1 Introduction **1**

2 Background and Literature Review on Image Enhancement and Restoration **5**

2.1 Classical Image Enhancement and Restoration Algorithms 5

2.1.1 Intensity Transformations. Variational Histogram Equalization 6

2.1.2 Spatial Filtering 12

2.1.3 Frequency Domain Methods 13

2.1.3.1 Fourier Methods for Image Processing 14

2.1.3.2 Multiresolution Image Processing via Wavelets . . . 17

2.2 Modern Model-Based Image Enhancement and Restoration Techniques 20

2.2.1 Models of the Visual System 20

2.2.1.1 Automatic Colour Enhancement - ACE 23

2.2.1.2 Improving ACE: Perceptual Colour Correction with Variational Techniques 24

2.2.2 Modelling the Degradation 27

2.2.2.1 Depth-Dependent Degradation and the Dark Channel Method 28

3 Underwater Image Restoration **35**

3.1 Introduction 36

3.2 Red Channel Underwater Image Restoration 37

3.2.1 WaterLight estimation from Red Channel 38

3.2.2 Transmission Estimate from Red Channel 39

3.2.3 Final Inversion 45

3.2.4 Handling Artificial Illumination 46

3.3 Experimental Results 49

3.3.1 Contrast and Visibility Recovery 49

3.3.2 Colour Correction Evaluation 51

3.4 Conclusions on Chapter 3 58

4 A Variational Framework for Image Dehazing **61**

4.1 Introduction 61

4.2 Related work 63

4.3 Enhanced Variational Image Dehazing (EVID) method 65

4.3.1 Variational contrast enhancement . . . 65
4.3.2 Modifying the Gray World assumption . . . 66
4.3.3 Controlling the saturation . . . 68
4.4 Experimental Results . . . 69
4.4.1 Qualitative Evaluation . . . 70
4.4.1.1 Comparisons with contrast enhancement approaches 74
4.4.1.2 Parameter Sensitivity . . . 76
4.4.2 Quantitative Evaluation . . . 80
4.5 Conclusions on Chapter 4 . . . 82

5 Variational Enhancement of Full-Field Digital Mammograms 85
5.1 Introduction . . . 85
5.2 Background and Related Work . . . 87
5.3 A Variational Formulation for Mammographic Image Enhancement . 89
5.3.1 Sparsity Prior for Variational Mammographic Image Enhancement . . . 90
5.4 Algorithm for Mammographic Image Enhancement . . . 92
5.5 Experimental Results . . . 93
5.5.1 Quantitative Performance Measures . . . 95
5.5.1.1 Global Quality Evaluation Metric . . . 95
5.5.1.2 Local Quality Evaluation Metric . . . 95
5.5.1.3 Perceptual Saliency-Based Evaluation . . . 97
5.5.2 Parameter Configuration . . . 98
5.5.3 Numerical Results . . . 102
5.5.3.1 Robustness and Sensitivity to Parameters . . . 103
5.6 Discussions and Conclusion . . . 105

6 Summary and Future Work 107
6.1 Satellite Image Enhancement. The Pansharpening Image Fusion problem . . . 108
6.2 Dehazing Mammograms . . . 110

Appendix A Variational Derivatives 115

Appendix B Main Accomplishments of this Dissertation 119

Bibliography 121

List of Figures

2.1 Simple intensity transformations and their effect 6
2.2 Different gamma correction curves 7
2.3 Result of applying Histogram Equalization to an overexposed image and the correspondent histograms. 8
2.4 Noisy image and application of Gaussian and anisotropic diffusion for denoising. 13
2.5 Result of applying Unsharp Masking with increasing degree of enhancement 14
2.6 Unevenly illuminated image, and the result of applying the Homomorphic filter. Notice the non-uniform illumination field varying from left to right and from top to bottom, that has been corrected by the filtering process. 16
2.7 Result of applying different denoising techniques to an image corrupted with Gaussian noise. 17
2.8 Wavelet denoising of Fig. 2.7a. The values of wavelet coefficients have been inverted (darker means higher) and stretched for better visualization. 19
2.9 Land and McCann experiment setting 21
2.10 Atmospheric Model of Haze, reproduced from [1]. 30
2.11 A haze-free natural scene and its Dark Channel. 31
2.12 A hazy natural scene and its Dark Channel. 31
2.13 Hazy scene with strong depth discontinuities, and illustration of how the unrefined transmission map fails to to capture depth differences. . 33
2.14 (a) Hazy urban scene. (b) Result of applying the Dark Channel corresponding to the image of Fig. 2.14b. 33

3.1 Waterlight detection from Red channel. (a) Original image, the red rectangle marks the region from where waterlight was extracted (b) In red, exact pixel location of waterlight within the rectangle marked in (a) (c) Red Channel of the original image (d) Final restoration result. 39
3.2 Transmission map refinement. 45
3.3 Saturation-based segmentation of artificially illuminated areas. Pixels with $S(x) < 0.4$ are painted on red. 47

3.4 Result of considering or not the Saturation cue. (a) Image with artificial light source on it. (b) Red Channel obtained with Eq. (3.5) (c) Image after processing without saturation prior. (d) Saturation-based segmentation of artificially illuminated areas. (e) Improved Red-Saturation Channel obtained with Eq. (3.18). (f) Image after processing with saturation prior. 48
3.5 The set of images selected for evaluation. 50
3.6 Results on the Diver image. 52
3.7 Visible edge maps of the recovered Diver image. 53
3.8 Results on the *Fishes* image. 53
3.9 Visible edge maps of the recovered *Fishes* image. 54
3.10 Results on the *Open scene* image. 54
3.11 Visible edge maps of the *Open scene* image. 55
3.12 Results on the *Ship* image. 55
3.13 Visible edge maps of the *Ship* image. 56
3.14 Scores for coefficient μ_{diff} in each of the four test images. 57
3.15 Scores for coefficient σ_{diff} in each of the four test images. 58
3.16 Scores for coefficient λ in each of the four test images. 59

4.1 (a) Hazy image, where the degradation increases with distance. A non-uniform illuminant dominates the scene. (b) The result produced by the EVID method. 63
4.2 (a) Hazy open scene. (b) Result of executing the method in Eq. (4.8). (c) Result of executing the method in Eq. (4.13). (d)-(f) Detail of the castle on top of the image (g)-(i) Detail of the buildings on the middle of the image. 70
4.3 Typical benchmark image for dehazing algorithms and result of processing it with several state-of-the-art methods, including EVID. . . . 71
4.4 Hazy image of New York City image and result of processing it with several state-of-the-art methods, including EVID. 72
4.5 Unevenly illuminated hazy image extracted from [2], and result of processing it with several state-of-the-art methods, including EVID. . . . 73
4.6 Unevenly illuminated hazy image of the Thames river, and result of processing it with several state-of-the-art methods, including EVID. . 74
4.7 An image of Paris, with a lack of chromatic information due to haze in the scene, and result of processing it with several state-of-the-art methods, including EVID. 75
4.8 (a) An image of Moscow, with a lack of chromatic information due to haze in the scene, and result of processing it with several state-of-the-art methods, including EVID. 76
4.9 Result of applying a conventional and an advanced contrast enhancement method compared with the result obtained by EVID. (a), (e), (i) Different hazy scenes. (b), (f), (j) Result of applying Histogram Equalization. (c), (g), (k) Result of applying the method in [3]. (d), (h), (l) Our proposed EVID method. 77

4.10 Zoomed details from Fig. (4.9). (a)-(d) Detail of the red wagon in the left of the image for the original image, and the output produced by HE, [3], and EVID. (e) - (f) Detail of the trees in the top part of the scene for the original image, and the output produced by HE, [3], and EVID. (i)-(l)Detail of the crane in the middle of the scene for the original image, and the output produced by HE, [3], and EVID. . . . 78
4.11 A hazy view of the city of Bilbao, and the result of varying contrast parameter γ. (a) Original hazy scene. (b) $\gamma = 0.1$ (c) $\gamma = 0.2$ (d) $\gamma = 0.3$. (e)-(h) Detail of the Guggenheim museum faÃğade. 78
4.12 A hazy view of the city of Bilbao, and the result of varying attachment-to-data parameter β. (a) Original hazy scene. (b) $\beta = 0.25$ (c) $\beta = 0.5$ (d) $\beta = 0.75$. (e)-(h) Detail of the Guggenheim museum faÃğade. . . 79
4.13 A hazy view of the city Bilbao, and the result of varying gray-world parameter β. (a) Original hazy scene. (b) $\alpha = 0.25$ (c) $\alpha = 0.5$ (d) $\alpha = 0.75$. (e)-(h) Detail of the Guggenheim museum faÃğade. 79
4.14 A hazy view of the city of Bilbao, and the result of varying saturation parameter η. (a) Original hazy scene. (b) $\eta = 0$ (c) $\eta = 0.02$ (d) $\eta = 0.04$. (e)-(h) Detail of the park next to the museum. 80
4.15 Scene from the LIVE Color+3D Database [4], and different synthetic heterogeneous fog added on it by following the method in [5]. 83

5.1 Histogram of pixel intensities in the entire INbreast database [6]. . . . 91
5.2 a) Original and b) -f) successive iterations of our method. (g)-(l): a magnification of the area marked in blue in the interior of the breast, containing a mass that progressively becomes more visible, while its contour appears more defined. The ground-truth labeling of the mass is marked in red on (a). 94
5.3 Target and background regions for the computation of the D-coefficient, together with the progressive enhancement produced by our method on a region enclosing a mass (red curve, groundtruth), and the corresponding overlap between the two regions. a) Region on the input mammogram b)-h) successive iterations of our method. 96
5.4 a) Region containing a mass (marked in red), corresponding to Fig. (5.3a) (b) - (f) Succesive enhancenment of the region containing the mass. (g)-(l) Saliency maps corresponding to the regions (a)-(f). Saliency distribution concentrates progressively within the mass. The amount of saliency gain with respect to the initial saliency map is, in each iteration: (h) +0.78% (i) +2.99% (j) + 5.49% (k) +13.41% (l) +16.80%. 98
5.5 a) Contrast fucntion of DWCE b) KNL function of DWCE 100

5.6 Progressive decrease of the amount of visual saliency captured by the mass within a region of increasing size. This saliency loss is partially mitigated by the application of our proposed method. (a) - (f) Region of the input mammogram containing a mass within progressively larger ROIs. (g)-(l) Corresponding saliency maps. (m)-(r) Region of the same mammogram after processing it with our proposed method. (s)-(x) Corresponding saliency maps. 105

6.1 Images showing the difference between original and pansharpened versions of degraded original, as obtained by different methods. Almost every technique leads to a considerable loss of spectral information, with a reduction of color saturation, being the non-local regularization based on Eq. (6.1) the one that best performs in this case. 109

6.2 Digital mammogram containing a cluster of microcalcifications, together with the result of applying the dehazing method in [7], and the CLAHE method. We see how visibility of the cluster increases noticeably. 113

Chapter 1

Introduction

In our modern society, we encounter digital images everywhere, from the internet or mobile devices to hospitals or satellites. The ability to process and enhance them is of paramount importance, and this is reflected by the wide range of advanced techniques for mathematical image processing that lies behind our everyday technology. These include modern digital cinema and television, medical image analysis programs, or consumer software such as Photoshop or Instagram, to name a few. From a more technical point of view, preprocessing input images may represent an essential step in order to improve ulterior computer vision tasks.

Despite being a mature area, mathematical image processing and enhancement is still a remarkably rich and lively field of research. In this sense, there are still strong limitations in the current state-of-the-art to properly process and enhance the quality of images whenever they are acquired in a non-controlled or hostile environment. It is therefore necessary to develop robust and efficient mathematical tools, capable of capturing the physical characteristics of each particular problem.

In this dissertation, we focus on three real-world image enhancement applications. Namely, underwater image enhancement, fog removal, and the enrichment of visual quality of digital mammograms. In the case of fog removal (also known as *image dehazing*), pixels that correspond to regions on the scene that are far away from the observer suffer from a loss of contrast that reduces visibility, as well as a decrease in saturation that leads to poor chromatic information. For underwater images, not only is contrast lost in far away areas of the scene, but also colour is deteriorated in a special manner that has nothing to do with the degradation that the atmosphere produces. Due to the different velocity at which wavelengths travel, colours associated to short wavelengths (*e.g.* red) tend to degrade the fastest, while colours related to longer wavelengths (for instance, green and blue) travel further in the scene and typically dominate the colour of acquired images. Finally, digital mammogram images do not suffer from chromatic degradation, since they are grayscale images. However, there is also low contrast in this class of medical images, and structures of interest lie occluded by *uninteresting* tissue. Therefore, digital mammograms also require specific methods to handle different regions of the image in distinct ways depending on the image content, and conventional methods seem to cope poorly with this necessity.

To address all these challenges, in this dissertation we study traditional as well

as modern techniques to improve image quality, and analyse why they may fail in retrieving adequate visibility in these cases. Our main objective is to develop novel approaches to overcome the described obstacles and design new more effective and efficient numerical methods for image enhancement in the aforementioned applications. To that end, we strongly base our developments on two methodological cornerstones: a) physical models of the propagation of light in attenuating media, and b) the calculus of variations. Regarding the first one, it provides an understanding of the causes of the image degradation; therefore, it comes with valuable information that, if available, no image enhancement approach should neglect. With respect to variational calculus, it equips us with a solid theoretical framework that enables us to design computational models that capture the particularities of each problem. This is achieved by means of the definition of image-dependent functional energies, composed of several terms modelling different parts of the problem. Moreover, these terms are weighted by scalar coefficients that allow the user to easily and intuitively control the visual characteristics present in the final solution.

The main contribution of this dissertation is thus to propose innovative mathematical and computational tools to enhance the quality of images that suffer from a complex degradation produced by an attenuating medium. Specifically, we mathematically represent that degradation through new physical and variational models, and we provide novel efficient numerical methods to solve those models.

The rest of this dissertation is structured as follows:

- In the second chapter of this dissertation we present a broad historical review of the field of mathematical image enhancement. We review some of the most notable methods proposed in the last decades. Two of such methods are of particular relevance for us, and we will analyse them in detail: 1) the variational formulation of histogram equalization proposed by Sapiro and Caselles in [8], and the subsequent works of Bertalmío *et al.* [3,9,10], that extend it, connecting it to the Retinex theory of visual human perception [11], and 2) the simple but powerful dark channel method [7, 12], designed for the task of image dehazing.

- The third chapter of this dissertation is dedicated to the problem of underwater image enhancement. We examine existing approaches to the visual improvement of this kind of images, and notice some of their main drawbacks. We then show how to overcome these drawbacks by extending the dark channel ideas and adapting them to the particularities of the underwater degradation. We provide an exhaustive evaluation of the quality of the images provided by our technique when compared with other relevant methods of the state-of-the-art.

- The fourth chapter of this dissertation deals with the problem of image dehazing. Contrarily to conventional approaches, we adopt here a variational point of view to model the problem. This gives us the ability to split the solution in different sub-tasks, and design an image energy that involves different functional terms defined in order to solve those sub-tasks. The first of these terms allows us to control the amount of deviation from the original input image; the second one provides a modified (lower) average gray value for the output image; the third

term is responsible for locally enhancing contrast, and the last one supplies a better saturation, improving colour quality. The overall action of these terms configures a novel image dehazing method that exhibits a better behaviour than current approaches, typically based on a physical model of haze degradation. This is particularly true when there is an uneven illumination present in the scene we want to enhance, since physical models are unable to consider this additional spatially variant source of degradation.

- In the fifth chapter of this dissertation we deal with a kind of images that are not acquired from real-world scenarios, namely digital mammograms. These images are formed by projecting X-ray beams in the interior of the breast and detecting the amount of electrons that are able to go through it. In other words, we represent the different attenuation capability of the tissues contained in the breast on a digital image. Due to the superimposition on a bi-dimensional image of three-dimensional structures in the interior of the human body, this imaging technique provides low-contrasted images, in which structures of interest suffer from a low visibility. We modify the variational framework proposed in the previous chapter to capture these particularities. After describing our novel mammographic image enhancement method, we provide a comprehensive evaluation on a modern digital mammogram database. Based on existing human visual saliency models, we are also able to assess the ability of our method to increase the amount of computational attention generated by regions containing tumoral masses within the breast.

- We close this dissertation in the sixth chapter, where we draw our main conclusions, and analyse some of the different research directions that arise from the developed work.

Chapter 2

Background and Literature Review on Image Enhancement and Restoration

In this chapter, we first review some of the most well-established methods for image enhancement. These are general methodologies that are usually applied to many image analysis task as a preliminary step, and represent a good benchmark to compare with.

Then, we present a detailed overview of some modern methods that are of interest for the next chapters of this dissertation. In all of our three problems the distortion of the images is spatially varying, due to the distance-dependent degradation that the media lying between the scene and the observer induces. We will introduce two families of enhancement algorithms. The latter relies on a physical model for distance-dependent degradation, that has been intensively used in the literature of image dehazing. We will analyse the different existing approaches to invert this model. The former, rather than attempting to model degradation, is based on how humans perceive colour and contrast, and attempts to implement some perceptual hypothesis of how the Human Visual System works. Quite remarkably, this model can be derived from one of the most simple image enhancement algorithms.

2.1 Classical Image Enhancement and Restoration Algorithms

Among the great number of image enhancement algorithms that have been developed in the last decades, there is a natural classification between methods working on the spatial domain and on the frequency domain. In the sections that follow, we make a short review of the most fundamental methods of both families.

2.1.1 Intensity Transformations. Variational Histogram Equalization

The most basic operations to enhance image contrast are defined pixel-wise. This means that the intensity value of every pixel is changed, not having into account its neighbourhood, or any other part of the image. Simply, to a given intensity value $\mathrm{I}(x, y)$, we associate a new value:

$$\mathrm{J}(x, y) = \mathrm{T}(\mathrm{I}(x, y)), \tag{2.1}$$

where $\mathrm{I}(x, y)$ is the input image and $\mathrm{J}(x, y)$ is the result of enhancing I by means of the transformation T. For now, we consider gray-scale images with domain $\Omega \subset \mathbb{R}^2$ and range being $[0, 256]$ or $[0, 1]$ as convenient. We will follow this notation for input and output images along the rest of the dissertation.

We can represent the transformation process performed by (2.1) by means of the graph of a bijective continuous function. Typical transformations together with their effects on an unevenly illuminated image are displayed in Fig. (2.1).

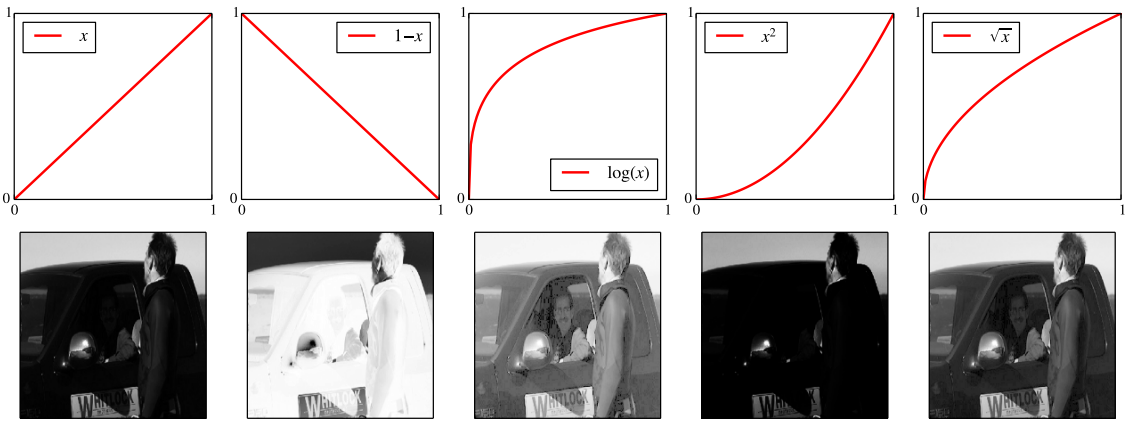


Figure 2.1: Simple intensity transformations and their effect

In the above equation (2.1), T is an operator on I, that acts in this case pixel-wise. If we denote as r and s the intensities of I and J at pixel (x, y) respectively, then $s = \mathrm{T}(r)$ is an intensity transformation function, and $\mathrm{T} : [0, 1] \to [0, 1]$. One of the most popular choices for t is the ***power-law or gamma transformations***, which have the basic form of $s = \alpha r^{\gamma}$, being γ a positive constant. Whenever $\gamma \in [0, 1]$, this transformation will map a narrow range of dark input intensities to a wider range of output values, although it will saturate brighter values, obtaining the opposite effect if $\gamma > 1$, as can be anticipated by looking at Fig. 2.2. Gamma correction is a widely used technique in most modern image capture and display devices [13], although it depends on the a priori knowledge of acquisition parameters that are often not available, which makes this a challenging problem. Blind appropriate gamma correction is still an active field of research [14]. Other recent research directions involve, for example, locally adaptive gamma correction [15, 16], or extensions to other image modalities such as infra-red [17].

As an alternative to gamma correction for contrast increase, one can use a piecewise linear function, which allows for more complex image modifications. Although it is a more flexible technique, it requires the user to provide more information [18].

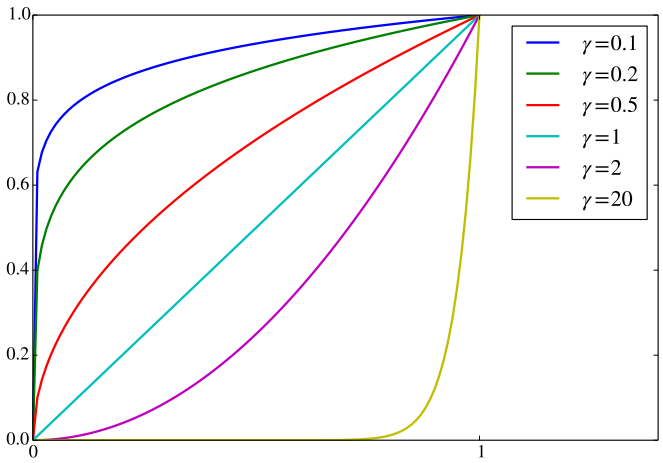


Figure 2.2: Different gamma correction curves

Another prominent family of methods rely on processing the histogram of the image. The histogram of an image with discrete intensity levels in $[0, L]$ is a discrete function h with domain $[0, L]$ that associates to each intensity step r_k the number of pixels n_k that intensity contains, i.e., $h(r_k) = n_k$. Thus, if an image is very bright, high values will occur more often and its histogram will be unbalanced to the right, as we can appreciate in Fig. 2.3.

After dividing each bin count by the number of intensity levels L in the image, we obtain a normalized histogram that can be interpreted as a probability distribution of the intensity levels of the image. The histogram can be used not only to enhance an image, but it also contains useful statistical information of it that can be used for other applications such as segmentation or compression.

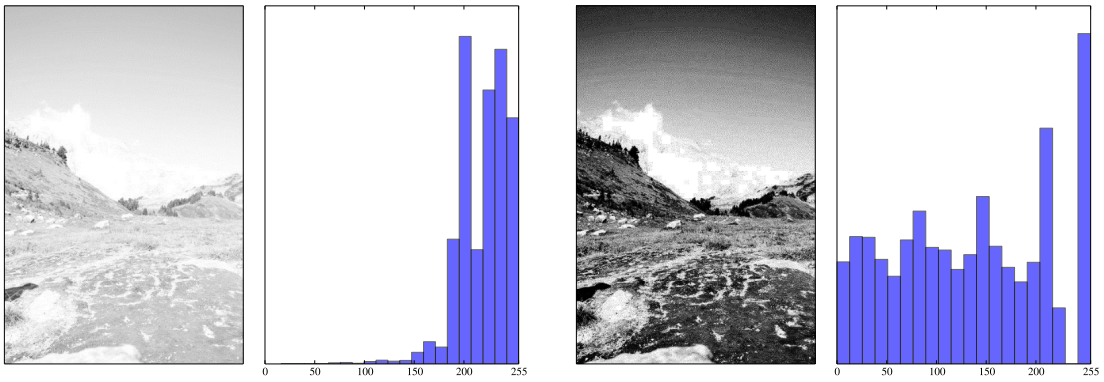
The most popular method that makes use of the histogram for image enhancement purposes is known as ***histogram equalization***. The underlying idea is to transit from an unevenly concentrated distribution such as the one in the left part of Fig. 2.3a to a distribution that is ideally uniform, as in the right part of Fig. 2.3b. The goal is thus to have a similar quantity of pixels lying in the different intensities, aiming at making maximal use of the available dynamic range.

We now switch from the continuous to the discrete representation of images, to allow both counting of occurrences of intensity levels and differentiation/integration. To find the required transform, we impose monotonicity to the transformation. For $\mathrm{I}(x, y) = k$, we will require $\mathrm{J}(x, y) = \mathrm{T}(\mathrm{I}(x, y)) > k$, since we do not want to reverse intensities. Suppose we have the distribution $p_r(r)$, where r is the input intensity. If we want it to be transformed to another distribution $p_s(s)$, where s is the output intensity, by means of a transformation $T(r) = s$ of the input image intensities, then the relationship between both distributions is:

$$p_s(s) = p_r(r)|\frac{dr}{ds}|. \tag{2.2}$$

Now, we assert that a transformation that makes the histogram uniform can be written as follows:

$$T(r) = s = (L-1)\int_0^r p_r(w)dw. \tag{2.3}$$



(a) Overexposed image (b) Result of Histogram Equalization

Figure 2.3: Result of applying Histogram Equalization to an overexposed image and the correspondent histograms.

To prove this, notice that:

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = \frac{d(L-1)\int_0^r p_r(w)dw}{dr} = (L-1)p_r(r), \tag{2.4}$$

so inserting (2.4) in eq. (2.2), we obtain:

$$p_s(s) = p_r(r)|\frac{1}{(L-1)p_r(r)}| = \frac{1}{L-1},$$

since all the above quantities are positive. In the above equation, p_s is a uniform distribution, which means that the transformation in eq. (2.3) achieves the objective of making the output image histogram evenly distributed. We conclude that to equalize the histogram in practice (i.e., in a discrete setting), we can build the cumulative distribution of the histogram. For a pixel of intensity level k we would count the number of occurrences of every pixel intensity from 0 up to k. We then would associate that quantity in the output image to every pixel with intensity k.

Being a relatively simple technique, the study of histogram equalization and its possible variants and extensions still motivates plenty of research works. Often the main goal is to retain the enhancement capabilities of histogram equalization while avoiding typical unpleasant artifacts in the form of noise amplification and over-enhancement. Localized and adaptive histogram equalization algorithms have been proposed, e.g. in [19–22]. In such approaches, usually a sliding window is implemented, which runs across the image pixels, equalizing each of the subimages defined by that window, and finally combining the resulting components to obtain an improved output with less over-enhancement artifacts. Alternatively, a thresholding and weighting approach can be adopted [23], where the histogram of the original image is clipped by an upper and lower threshold, and only the intermediate histogram is transformed, to be combined with the non-transformed part of the original image with proper weights. Optimization-based methods also exist in the literature, *e.g.* in [24] authors propose a linear programming formulation and in [25], where they

employ genetic algorithms in an effort to attain a parameter-free method. There also exist methods that aim to take into account contextual information beyond pixel intensities, such as [26, 27], which use image edge information to weight the histogram. Several other generalizations have been proposed in the literature. In [28], the authors propose to integrate histogram equalization with the white balance technique; in [29], authors extend the 1D conventional to two dimensions in an attempt of capturing the spatial information that is inherently lost in the construction of an image histogram. Histogram equalization has also been extended non-trivially for colour images, for instance in [30] or in [31].

Besides this, histogram equalization is of particular interest for the rest of this dissertation, specially for Chapters 3 and 4, as it forms the basis for the HVS-based methods that we will further adapt to our particular scenarios. The key property of histogram equalization that enables us to adapt it to specific situations is that it can be embedded in a variational formulation that permits its modification to perform joint contrast enhancement and other tasks of interest. This variational formulation was first introduced by Caselles and Sapiro in [8], and we describe in detail it here for the sake of completeness.

Let us assume the input image is defined in a domain $\Omega = [0, W] \times [0, H] \subset \mathbb{R}^2$, and take two pixels in this domain $x = (x_1, x_2), y = (y_1, y_2) \in \Omega$. Consider a smooth function $\psi : [0, 1] \to \mathbb{R}$, $\psi \in \mathcal{C}^{1)}(0, 1)$. We define the following energy functional acting on the space of allowed images:

$$E(\mathrm{I}(x)) = \int_\Omega \psi(\mathrm{I}(x))dx - \frac{1}{WH} \iint_{\Omega^2} |\mathrm{I}(x) - \mathrm{I}(y)| \, dxdy. \tag{2.5}$$

To enable differentiation, we define an approximation of the non-differentiable absolute value in (2.5) by:

$$|z|_\varepsilon = \sqrt{z^2 + \varepsilon^2},$$

the derivative of which is:

$$\frac{d}{dz}|z|_\varepsilon = \frac{z}{\sqrt{z^2 + \varepsilon^2}} = sign_\varepsilon(z),$$

being $sign_\varepsilon$ a function that tends to the sign function as ε tends to 0:

$$sign_0(t) = \begin{cases} 1, & \text{if } t > 0 \\ 0, & \text{if } t = 0 \\ -1, & \text{if } t < 0. \end{cases} \tag{2.6}$$

We define also the positive sign function, given by:

$$sign^+(t) = \begin{cases} 1, & \text{if } t > 0 \\ 1/2, & \text{if } t = 0 \\ 0, & \text{if } t < 0. \end{cases}$$

With this terminology, we can write an approximate version of the energy (2.5) as:

$$E_\varepsilon(\mathrm{I}(x)) = \int_\Omega \psi(\mathrm{I}(x)) \, dx - \frac{1}{WH} \iint_{\Omega^2} |\mathrm{I}(x) - \mathrm{I}(y)|_\varepsilon \, dxdy, \tag{2.7}$$

and we can now compute the variational derivative of (2.7) at the image I, in the direction of the variation δI:

$$\delta E_\varepsilon(\mathrm{I}(x), \delta\mathrm{I}(x)) = \int_\Omega \psi'(\mathrm{I}(x))\delta\mathrm{I}(x)\,dx - \frac{1}{WH}\iint_{\Omega^2} sign_\varepsilon(\mathrm{I}(x)-\mathrm{I}(y))(\delta\mathrm{I}(x)-\delta\mathrm{I}(y))\,dxdy.$$

Using equality $sign_\varepsilon(t) = -sign_\varepsilon(-t)$ and changing variables $x \to y$ in the second term, we can decompose the second integral as:

$$\begin{aligned}&\iint_{\Omega^2} sign_\varepsilon(\mathrm{I}(x) - \mathrm{I}(y))\,\delta\mathrm{I}(x)\,dxdy + \iint_{\Omega^2} sign_\varepsilon(\mathrm{I}(x) - \mathrm{I}(y))\,\delta\mathrm{I}(y)\,dxdy \\ &= 2\iint_{\Omega^2} sign_\varepsilon(\mathrm{I}(x) - \mathrm{I}(y))\,\delta\mathrm{I}(x)\,dxdy,\end{aligned}$$

which allows us to write the above derivative as:

$$\delta E_\varepsilon(\mathrm{I}(x), \delta\mathrm{I}(x)) = \int_\Omega \left(\psi'(\mathrm{I}(x)) - \frac{2}{WH}\int_\Omega sign_\varepsilon(\mathrm{I}(x) - \mathrm{I}(y))\,dy\right)\delta I(x)\,dx.$$

The above equality is the interior product of the variational derivative of E_ε(I) with respect to the direction $\delta I(x)$, which means that we can write:

$$\delta E_\varepsilon(\mathrm{I}(x)) = \psi'(\mathrm{I}(x)) - \frac{2}{WH}\int_\Omega sign_\varepsilon(\mathrm{I}(x) - \mathrm{I}(y))\,dy.$$

When ε tends to 0, we obtain:

$$\delta E(\mathrm{I}(x)) = \psi'(\mathrm{I}(x)) - \frac{2}{WH}\int_\Omega sign_0(\mathrm{I}(x) - \mathrm{I}(y))\,dy,$$

Since $sign_0(t) = 2sign^+(t) - 1$ and $\int_\Omega 1\,dxdy = WH$, we can finally rewrite the variational derivative of the energy (2.5) as:

$$\delta E(\mathrm{I}(x)) = \psi'(I(x)) + 2 - \frac{4}{WH}\int_\Omega sign^+(\mathrm{I}(x) - \mathrm{I}(y))\,dxdy. \tag{2.8}$$

Now, let us write $\mathrm{I}(x) = \lambda$, and define:

$$H_\mathrm{I}(\lambda) = \frac{1}{WH}\int_\Omega sign^+(\lambda - \mathrm{I}(y)\,dy,$$

for $\lambda \in [0, 1]$. Notice that this is a variational version of the cumulative histogram of $\mathrm{I}(x)$. If we had $H_\mathrm{I}(\lambda) = \lambda\ \forall\lambda$, then the histogram of $\mathrm{I}(x)$ would be uniform. In other words, the transformation:

$$\lambda \mapsto \frac{1}{WH}\int_\Omega sign^+(\lambda - \mathrm{I}(y))\,dy$$

performs an equalization of the histogram of the input image I.

Let us observe that the variational derivative in (2.8) can be rewritten in terms of $H_\mathrm{I}(\lambda)$ as:

$$\delta E(\mathrm{I}(x)) = \psi'(\mathrm{I}(x)) + 2 - 4H_\mathrm{I}(\lambda). \tag{2.9}$$

Given that we can find a local minimum of the energy (2.5) by imposing that its variational derivative (2.9) vanishes, a stationary image for this flow will verify:

$$H_{\mathrm{I}}(\lambda) = \frac{\psi'(\lambda) + 2}{4}$$

This means that to equalize the histogram of the input image, we would need to set $\psi' = 4\lambda - 2$, i.e., $\psi(\lambda) = 2\lambda^2 - 2\lambda + K$, with K an arbitrary constant. Writing $K = 1/2$, we obtain $\psi(\lambda) = 2(\lambda - 1/2)^2$. Thus, we conclude that an energy that equalizes the histogram of an input image I when minimized is given by:

$$E(\mathrm{I}(x)) = 2\int_{\Omega}(I(x) - \frac{1}{2})^2\,dx - \frac{1}{WH}\iint_{\Omega^2}|\mathrm{I}(x) - \mathrm{I}(y)|\,dxdy. \tag{2.10}$$

Its derivative would be given by:

$$\delta E(\mathrm{I}(x)) = 4\mathrm{I}(x) - 2 + 2 - \frac{4}{WH}\int_{\Omega} sign^{+}(\mathrm{I}(x) - \mathrm{I}(y))\,dy \tag{2.11}$$

To interpret (2.10), we can define the average contrast of I is:

$$C(I(x)) = \frac{1}{WH}\iint_{\Omega^2}|\mathrm{I}(x) - \mathrm{I}(y)|\,dxdy.$$

With this definition, we see now that minimizing (2.10) amounts to maximizing the average contrast of the image I while penalizing deviation of intensity from a theoretically correct mean of 1/2.

A simple numerical scheme to minimize (2.10) could consist on a descent in the direction of maximum decrease of the energy, i.e., the negative gradient. To that end, we introduce an artificial time parameter t, and we obtain the following algorithm:

- Set $\mathrm{I}(0) = \mathrm{I}$, and set $\dfrac{\partial \mathrm{I}(x)}{\partial t} = -\partial E(\mathrm{I}(x))$.
- Discretize in time, and iterate until steady-state the following expression:

$$\mathrm{I}_{t+1}(x) = \mathrm{I}_t(x) - \Delta t\,\left(\mathrm{I}(x) - \frac{1}{WH}\int_{\Omega} sign + (\mathrm{I}(x) - \mathrm{I}(y))\,dy\right). \tag{2.12}$$

By implementing (2.12), we obtain a sequence of modified versions of the initial image $\mathrm{I}(x)$ that converges to a histogram-equalized version of it. This formulation in itself is of no practical interest, since it is computationally expensive and we are losing the simplicity and efficiency of simple histogram equalization. Nevertheless, it formalizes and throws light to the whole histogram equalization process. The advanced contrast enhancement methods we will analyse and extend in this dissertation build on these ideas. Therefore, it is worth to explain the details of this variational formulation.

2.1.2 Spatial Filtering

One of the main drawbacks of the enhancement techniques presented so far is that they operate in a pixel-wise manner, acting over the intensity distribution of the whole image, and thus remaining "blind" to the intensity of surrounding pixels. This global approach can be useful for overall contrast enhancement, but it is not suitable in the majority of the cases, when one wants to enhance details on specific areas of the image. Histogram processing techniques can be adapted to perform local enhancement by computing at each pixel the histogram of the points within a specified neighbourhood, although this is computationally very expensive. They can also be localized in the intensity range, by not allowing too dramatic intensity changes. This is known as contrast-limited histogram equalization [32]. Another possibility would be to compute the histograms of non-overlapping areas and use them to compute the intensity transform function. However, this often leads to undesired block artifacts in the resulting image [33].

Spatial filtering, on the other hand, works in a neighbourhood of each pixel of the image by performing computations involving the pixel itself and the surrounding pixels, and stores the result in the same location of the output image. Spatial filters receive sometimes the names of masks, kernels or windows. Grossly speaking, one designs a filtering mask, and then convolves the image with this mask with a sliding window approach, padding the image where necessary. In fact, the word filter is inherited from the frequency counterpart that we shall briefly review in the next section. The basic property that relates both points of view is that convolution is equivalent to product in the Fourier transformed space. This allows a reduced computational complexity. However, there is an important difference between spatial and frequency filtering, namely, spatial filtering permits the implementation of non-linear filtering, while this cannot be achieved straightforwardly in the frequency domain, since convolutions are linear operators. An example of this filter is the median filter [34], in which the intensity value of a pixel is substituted by the median of the surrounding intensities, rather than the mean.

The definition of the mask determines the output of the filtering process, that can range from a blurred version of the original image to a sharpened version of it. This choice will depend on the specific applications, such as denoising, removing defocus, or detail enhancement.

A particularly popular filter is the one employing a Gaussian kernel. It is equivalent to blurring the image by convolution with a Gaussian. As long as the noise present in the image matches a Gaussian white noise model, this approach will work. Unfortunately, although white noise is often employed to test the performance of denoising methods, its use is rare in practice. Another drawback of this simple approach is that while denoising the image, a Gaussian filter will also blur informative parts of it.

Gaussian blurring can be seen to be equivalent to applying the heat diffusion equation to the input image. As a consequence, information in the image is diffused in an isotropic manner. This can be improved by designing different filters that depend on the image content. A famous solution was proposed by Perona and Malik

in [35], where the blurring kernel is modified to make it depend on the image gradient, and behave as an impulse function close to edges and other structures that must be preserved in the resultant image. In Fig. (2.4), reproduced from [36], we can appreciate how, while both Gaussian and anisotropic diffusion succeed in removing the noise in the input image, the former preserves better high-frequency content (borders).

(a) Original noisy image (b) Gaussian denoising (c) Anisotropic denoising

Figure 2.4: Noisy image and application of Gaussian and anisotropic diffusion for denoising.

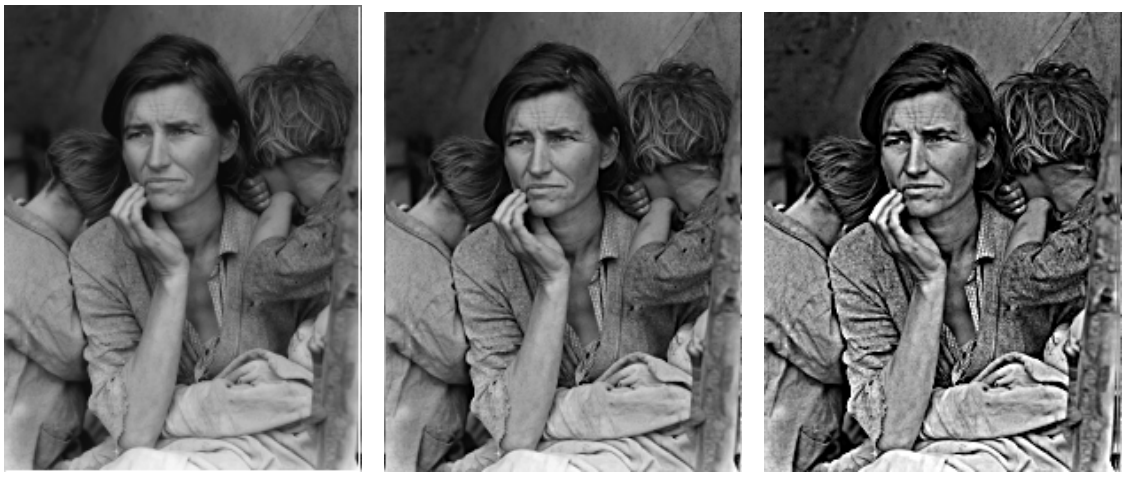
Combinations of different procedures can also be useful to enhance details in an image. That is the case with a popular process called ***unsharp masking*** [37]. It consists on blurring the original image, and building a mask by subtracting the blurred image from the original. Adding that mask to the initial image results in detail enhancement and sharpening. This method can be formulated as:

$$\begin{aligned} \text{mask}(x,y) &= \text{I}(x,y) - (G(x,y) * I(x,y)), \\ \text{J}(x,y) &= \text{I}(x,y) + \alpha \text{mask}(x,y), \end{aligned} \tag{2.13}$$

where $G(x,y)$ is a Gaussian blurring filter and α controls the amount of enhancement. Extensions of this simple idea are still been investigated, as in the recent works of [38] or [39]. The result of applying this basic technique can be appreciated in Fig. (2.5).

2.1.3 Frequency Domain Methods

With the evolution of computers, the frequency domain analysis of images became essential as a tool that must be in the repertory of every practitioner of the field of digital image processing. There are some elements of the human perception that are better understood in terms of frequencies rather than of the spatial distribution of intensities. As a result, a wide range of applications have benefited from the idea of studying the spectrum of frequencies of an image, being possibly the most successful ones image compression and denoising. In the next lines we provide a broad overview of this important approach to image enhancement and restoration.



(a) Original blurry image (b) Unsharp masking, $\alpha = 0.25$ (c) Unsharp masking, $\alpha = 0.5$

Figure 2.5: Result of applying Unsharp Masking with increasing degree of enhancement

2.1.3.1 Fourier Methods for Image Processing

Frequency Domain methods are one of the most powerful tools in the field of signal and image processing, with a large quantity of research devoted to them [40, 41]. When we consider an image I as a continuous function, defined for simplicity in the unit domain $\Omega = (0, 1) \times (0, 1)$, and periodically extended, the information that I encodes can be completely captured into the coefficients that compose its Fourier series representation, that are given by:

$$c_n = c_{n_1,n_2} = \langle \mathrm{I}(x), e^{2\pi i \langle x,n \rangle} \rangle_{L^2(\Omega)},$$

for $n = (n_1, n_2) \in \mathbb{Z}^2$, and being $x = (x_1, x_2) \in \Omega$.

A great computational advantage can be achieved for the Fourier transform when considered in the discrete domain. In that setting, the image domain is a discrete grid. $\Omega = \{0, \ldots, N-1\} \times \{0, \ldots, N-1\}$, and the image can be interpreted as a matrix $\mathrm{I}(j)$, for $j = (j_1, j_2) \in \Omega$. Then, the Fourier coefficients of the image can be computed with the Discrete Fourier Transform (DFT) as:

$$c_n = c_{n_1,n_2} = \sum_{j \in \Omega} \mathrm{I}_j e^{i \frac{2\pi}{N} \langle j,n \rangle},$$

for $n \in \Omega$. The DFT admits a fast implementation known as the Fast Fourier Transform, invented by Cooley and Tukey [42], which is able to improve the performance of a large number of image processing tasks needing to run fast, often in real time [43]. The Fourier transform and its variations (e.g., cosine/sine transform) have been applied in uncountable image processing tasks, apart of image enhancement, such as image compression [44] or medical image reconstruction [45].

One of the most straightforward applications of the Fourier transform for image enhancement is possibly the homomorphic filter [46]. It consists of expressing an

image I as the product of its illumination $j(x)$ and its reflectance $r(x)$. This is known as the illumination-reflectance model of image formation, and it specifies that the intensity of each pixel, *i.e.*, the light reflected by a point in the object and reaching the sensor in the camera, is the product of the illumination of the scene with the reflectance of the object. This can be written as $\mathrm{I}(x) = j(x) \cdot r(x)$, with $j(x)$ being the scene illumination, and $r(x)$ the scene reflectance. In other words, reflectance is produced by the properties of the scene objects themselves, while illumination results from the lighting conditions at the image capture moment.

The goal of homomorphic filtering consists of compensating for a non-uniform illuminant by separating and removing the illumination from the true scene content. The key idea is that illumination fields are generally slowly-varying, while image information mainly lies on the mid-high frequencies. To obtain such a separation, we first take natural logarithms in the previous expression:

$$\ln \mathrm{I}(x) = \ln j(x) + \ln r(x).$$

Notice that the previous equation would lead to non-integrability (and thus the Fourier transform would not be computable) if the input image is dark, so that its logarithm is equal to $-\infty$. This case is not considered here, since for a black image there is no illumination to remove, so we restrict ourselves to the subset of images that take strictly positive values.

Now we move to the spectral domain. Denoting the Fourier Transform of each of the terms of the above equation as $F_{\mathrm{I}} = \mathfrak{F}(\ln \mathrm{I}(x))$, $F_j = \mathfrak{F}(\ln j(x))$, and $F_r = \mathfrak{F}(\ln r(x))$, we can write:

$$F_{\mathrm{I}}(\omega) = F_j(\omega) + F_r(\omega).$$

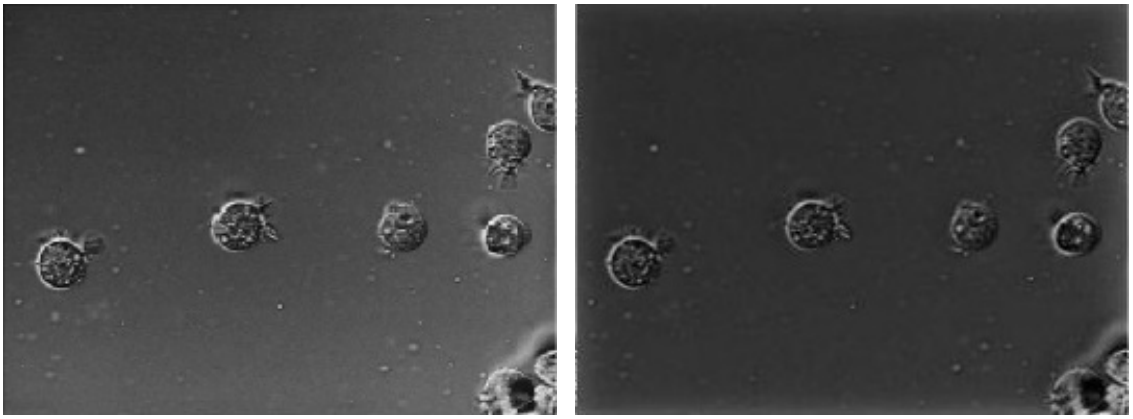
We are indeed interested on recovering an enhanced version of $r(x)$, namely, the original image but free of illumination influence. Having assumed that $j(x)$ lies in the low frequencies (and so does $\ln j(x)$), the idea is to filter the low values on the above equation. That can be achieved by multiplying by high-pass filter $H(\omega)$:

$$H(\omega)F_{\mathrm{I}}(\omega) = \underbrace{H(\omega)F_j(\omega)}_{\approx 0} + H(\omega)F_r(\omega) \approx H(\omega)F_r(\omega)$$

Now we can come back to the spatial domain. Denote the Inverse Fourier Transforms of the filtered images by $\tilde{\mathrm{I}}(x) = \mathfrak{F}^{-1}(H(\omega)F_{\mathrm{I}}(\omega))$ and $\tilde{r}(x) = \mathfrak{F}^{-1}(H(\omega)F_r(\omega))$. The enhanced reflectance of the input image I is given, after undoing the initial logarithmic transformation, by $e^{\tilde{r}(x)}$, which can now be written in terms of the original image:

$$e^{\tilde{r}(x)} = e^{\tilde{\mathrm{I}}(x)} = e^{\mathfrak{F}^{-1}(H(\omega)F_{\mathrm{I}}(\omega))} = e^{\mathfrak{F}^{-1}(H(\omega)\mathfrak{F}(\ln \mathrm{I}(x)))}. \tag{2.14}$$

Eq. (2.14) provides a method to remove illumination effects while enhancing the image content and details. Thus, the homomorphic filters and its extensions have been applied to normalize illumination before segmentation or face recognition tasks [47–49], among others. The effects of applying this procedure can be appreciated on Fig. 2.6:



(a) Original unevenly image (b) Homomorphically filtered image

Figure 2.6: Unevenly illuminated image, and the result of applying the Homomorphic filter. Notice the non-uniform illumination field varying from left to right and from top to bottom, that has been corrected by the filtering process.

Wiener filter [50] is another popular image enhancement technique that relies on the spectral transform of the input image. The fundamental idea is the consideration of images as random variables, as well as their noise. Thus, the goal of the Wiener filter is to derive an estimate $\hat{\mathrm{I}}$ of an ideally clean image I that is optimal in terms of the mean square error between them. The assumptions of the Wiener filter are that the noise and the image are not correlated, one of them has zero mean (they can have different variances) and that intensity values of the estimate of the noise-free image are linearly related to values in its degraded counterpart. Under such conditions, it can be shown [51] that the spectral transform of the minimum of the error function is given by:

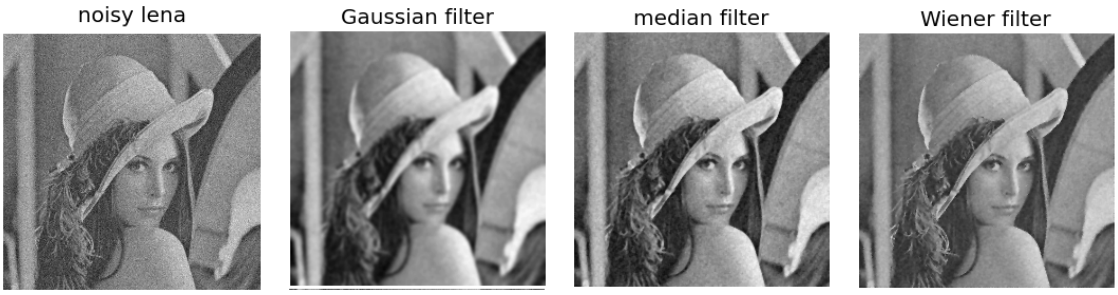
$$\mathfrak{F}\hat{\mathrm{I}} = \frac{1}{H(\omega)} \frac{|H(\omega)|^2}{|H(\omega)|^2 + N(\omega)/S(\omega)}, \tag{2.15}$$

where $H(\omega)$ is a linear function modelling the degradation (in the transformed domain), $N(\omega)$ is the power spectrum of the noise, and $S(\omega)$ is the power spectrum of the undegraded image. Inverse Fourier Transform applied to the output of Eq. (2.15) produces then the restored image. Power spectrum here is defined to be the sum of the squares of the real and the imaginary parts of the noise. Both the spectrum of the noise and the undegraded image must be estimated before Wiener filtering can be applied. A common simplification is to consider noise as white (Gaussian), which leads to a constant spectrum $N(\omega)$. Since noise is of zero mean, only variance needs to be estimated, and it is usually approximated by the local standard deviation on the image. Spectrum of the undegraded image, though, is rarely known, and must be estimated with an auxiliary method. If this is not possible, a frequent simplification is to employ the following formula:

$$\mathfrak{F}\hat{\mathrm{I}} = \frac{1}{H(\omega)} \frac{|H(\omega)|^2}{|H(\omega)|^2 + K}, \tag{2.16}$$

where K is a constant that remains as a parameter of the method. An example of

Wiener denoising, when compared to other conventional techniques, can be found in Fig. (2.7).



(a) Original image with Gaussian noise. (b) Gaussian filter. (c) Median filter. (d) Wiener filter.

Figure 2.7: Result of applying different denoising techniques to an image corrupted with Gaussian noise.

Wiener filter, being itself a fundamental and basic method, still serves these days as a basis for more advanced techniques, with modern applications such as superresolution [52] or dehazing [53]. Wiener filter is also a key component of other recent noise removal techniques, such as BM3D [54], an state-of-the-art denoising method based on the consideration that an image has a locally sparse representation in transform domain. This sparsity is enforced by grouping similar 2D image patches into 3D groups, that are then filtered jointly, and the result combined to obtain a noise-free image.

2.1.3.2 Multiresolution Image Processing via Wavelets

Although Fourier analysis is one of the most popular and widespread tools for image processing and restoration, in the last decades of the 20th century a different frequency-based approach, known as wavelet analysis, has revealed itself as an even more general and powerful approach [55]. Wavelets consists of a localized version of Fourier analysis, capable of operating at different scales.

To better understand the difference between wavelet and Fourier analysis, let us switch momentarily to temporal signals. A signal $f(t)$ defined in the temporal domain and its Fourier transform $\hat{f}(\omega)$, defined in the frequency domain, have the following linking relationships:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{2\pi i \omega t} \, dt,$$

and

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{-2\pi i \omega t} \, d\omega.$$

Thus, Fourier transform characterizes the content of signal f through its frequency components. Given that the support of the basis functions $e^{2\pi i \omega t}$ encompasses the whole time domain (infinite support), $\hat{f}(\omega)$ depends on the values of $f(t)$ for all

times. This makes the Fourier transform a global transform unable to analyse local properties of the original signal.

To be capable of performing a local frequency analysis, a windowed Fourier transform (WFT) was first introduced by means of a window function $w(t)$ acting into the Fourier transform [56]:

$$Sf(\omega, t) = \int_{-\infty}^{\infty} f(\tau) w(t-\tau) e^{-2\pi i \omega \tau} \, d\tau.$$

The energy of these new basis functions $g_{\tau,\xi}(t) = w(t-\tau)e^{-2\pi i \xi t}$ is concentrated in a neighbourhood of time τ over an interval of size σ_t, which corresponds to the standard deviation of $|g|^2$. The Fourier transform of these functions is $\hat{g}_{\tau,\xi} = \hat{w}(\omega - \xi)e^{-2\pi i \tau(\omega-\xi)}$, with energy in frequency domain localized around ξ, over an interval of size σ_ω. The Heisenberg uncertainty principle states that the energy spread of a function and its Fourier transform cannot be simultaneously arbitrarily small, since it must be verified:

$$\sigma_t \sigma_\omega \geq \frac{1}{2}.$$

To analyse local signal structures of various supports and amplitudes in time, it is necessary to use basis functions with different support sizes for different temporal locations. For instance, in the case of high frequency structures, which vary rapidly in time, higher temporal resolution is needed to accurately capture those changes; on the other hand, for a lower frequency, a relatively higher absolute frequency resolution will give a better measurement of frequency values. Wavelet transforms provide a natural representation to accomplish this analysis.

One of the most important features of wavelet transforms is thus their multi-resolution representation. Physiological analogies have suggested that wavelet transforms are similar to low level visual perception. From texture recognition, segmentation to image registration, such multi-resolution analysis provides the possibility of studying a particular problem at various spatial/frequency scales. The main idea of wavelet analysis is to capture that locality by passing from the Fourier orthogonal basis to localized bases, which are designed according to different resolutions, *i*.e., scales of details in the image.

A wavelet function is defined as a function $\psi \in L^2(\mathbb{R})$ with zero average:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0, \tag{2.17}$$

that is normalized so that $\|\psi(t)\| = 1$, and centred in a neighbourhood of $t = 0$. A family of wavelet functions is then derived from ψ by simple translation and scaling of it:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \tag{2.18}$$

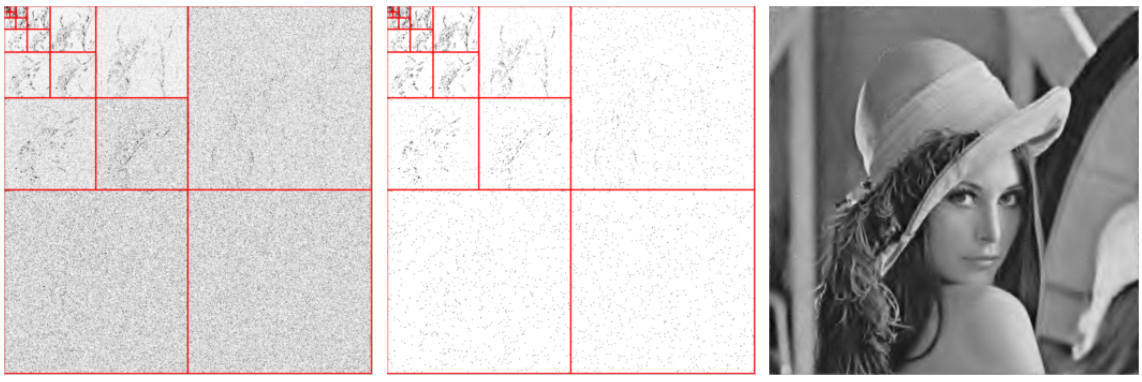
where $a \in \mathbb{R}^+$, $b \in \mathbb{R}$. Thus, a family of wavelet functions originates from a mother wavelet function, that is scaled by the coefficient a and then translated by b. The result is that way localized in both frequency and spatial domain.

Although wavelets have been applied extensively in many image processing fields, such as compression or super-resolution, their most traditional and successful application has been image denoising. An image is often corrupted by noise in its acquisition, transmission, compression, reconstruction, etc. The goal of image denoising is to produce good estimates of the original image from noisy observations.

The general process can be summarized as follows. Consider a noisy image $X = \mathrm{I} + N$, where N is additive noise and I is the noise-free image. Let W and W^{-1} be the forward and inverse wavelet transform operators. Denoting a threshold function of parameter λ as $D(\cdot, \lambda)$, to recover an estimate $\hat{\mathrm{I}}$ of the noise-free image I, the following steps are implemented:

1. To apply wavelet transform on the degraded image: $Y = W(X)$.
2. To threshold wavelet basis coefficients adequately $Z = D(Y, \lambda)$.
3. To reconstruct the image $\hat{\mathrm{I}} = W^{-1}(X)$.

The idea is that small coefficients in the wavelet decomposition are more likely associated to noise, and large coefficients come from important signal features (such as edges). Therefore, in its most basic form, this technique denoises in the orthogonal wavelet domain, where each coefficient is thresholded: if the coefficient is smaller than the threshold, it is set to zero; otherwise it is kept or modified. This is known as hard thresholding. The next natural step is the analysis of different thresholding mechanisms. The most important one was proposed by Donoho *et al.* in [57]. Since then, a large amount of work [58–60] has been devoted to the automatic optimal selection of the wavelet coefficients and their thresholding in order to perform image restoration after it has been corrupted by noise. Wavelet decomposition of the noisy image in Fig. 2.7 can be appreciated in Fig. 2.8, together with the corresponding thresholded wavelet coefficients and the reconstructed noise-free image after inverse wavelet transform of the modified coefficients.



(a) Wavelet coefficients of the noisy image in Fig. 2.7a.

(b) Thresholded coefficients.

(c) Reconstruction of a denoised version of Fig. 2.7a

Figure 2.8: Wavelet denoising of Fig. 2.7a. The values of wavelet coefficients have been inverted (darker means higher) and stretched for better visualization.

2.2 Modern Model-Based Image Enhancement and Restoration Techniques

In this section, we describe two families of specific methods that attempt to enhance images from a substantially different point of view, when compared to the previously outlined techniques.

First, we give a complete description of a series of methods that are based in modelling the way humans perceive the world. The underlying philosophy is that if we understand some of the mechanisms of human perception, we can then use a mathematical model of them to modify a digital image in order to obtain an enhanced version of it, more complying with what a human expects to perceive, hopefully suppressing the present distortions.

Second, we study the modelling of a specific kind of degradation that affects the process of image acquisition when the signal is attenuated by the media. We describe how a simple physical model is able to capture to a reasonable extent the particularities of the degradation caused by the atmosphere, and we review the currently most powerful technique to invert that model and estimate an undegraded image.

2.2.1 Models of the Visual System

In this section, we explain the details of a family of methods that are inspired in the way humans perceive colour and luminance to process an image. These ideas are of great importance for the rest of the dissertation: a considerable part of the methodologies we develop in chapters 4 and 5 builds on the techniques we describe in the following pages.

The underlying mechanisms that govern human perception have inspired the image processing community for many years. One key milestone is the Retinex model, developed by Land and McCann in the seventies [11]. From this work, many other models have been derived for colour image processing. Let us explain qualitatively the ideas underlying it.

When we look at an object from different distances, perspectives, or illuminations, the large changes in the image that is formed in our retina do not affect our perception of the object. This implies that our brain is able to compensate for these modifications and we still perceive its physical properties (such as colour or size) with little variation. In other words, our sensations are much more stable than what we could expect after substantial changes of the retinal image.

When applied to colour invariancy, this phenomenon is known as *colour constancy*. It simply means that our visual system understands the permanency of colours of objects even when it is exposed to large changes in the spectral composition of the light that arrives to our retina. This spectral composition is known to be made of two different factors that combine themselves in a non-linear way. These factors are spectral reflectance of the observed object (the property an object has to reflect certain wavelengths more than others) and spectral composition of the scene's illuminant, i.e., the composition of wavelengths that form the light that falls on the object.

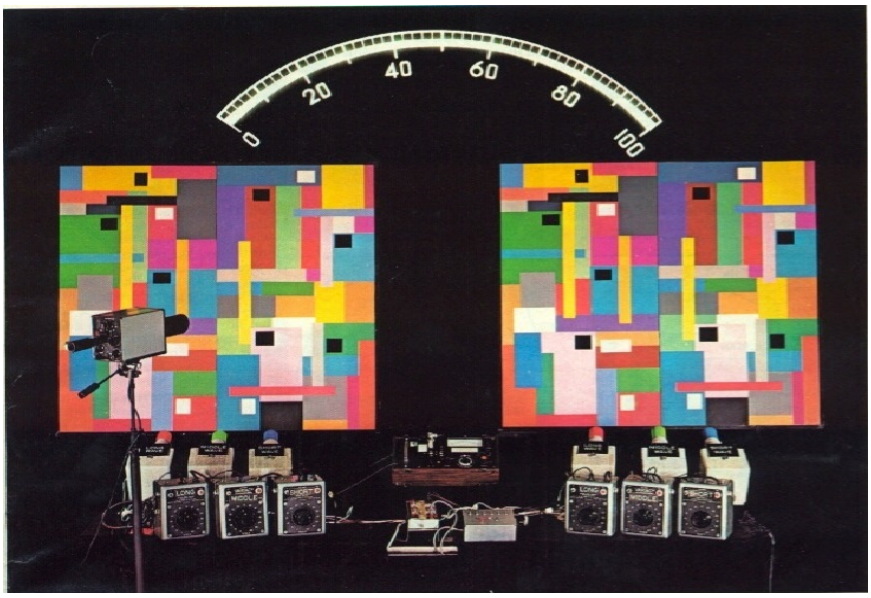


Figure 2.9: Land and McCann experiment setting

Colour constancy explains the fact that when we look at a scene illuminated with, for instance, a tungsten reddish light, and we suddenly turn off the tungsten bulb letting the sunlight illuminate the scene, the perception of the colours of objects on the scene will remain stable, as it just depends on their spectral reflectance. If we were seeing an object as yellow, our visual system will not depart as much from that yellow as we would have expected from such a large change in the illumination. In this same context, an optical camera would fail to compensate for differences in light sources, revealing that it is our visual system, having become robust to such changes through many years of human evolution, who unconsciously performs colour constancy.

Land and McCann conducted a famous series of experiments to prove this property of the Human Visual System (referred to as HVS from now on). They exposed a group of people to a large display of abstract colour patches (this famous pictures are called Mondrians, see Fig. (2.9)), and let them illuminate the patches with a device equipped with colour filters that controlled the brightness and spectral composition of the emitted light. With this setting, they compared the real reflectance of patches, measured with spectrophotometers, and the sensations induced in the observers under changing illuminants. This allowed them to build a lightness perception model that proposes that colour is determined by three lightnesses. Each of them would be computed by performing differential comparisons of intensity information coming from the whole image. To access this far away information, the Land and McCann lightness model states that the HVS calculates ratios of energies at nearby points, and propagates this information by multiplication of ratios to form products that relate different image regions. The mean of many different products of ratios can then be used to predict the lightness and the colour sensation that the observer will report.

Building on the revolutionary work of Land and McCann, many perceptually inspired colour correction algorithms have been designed, with the goal of manipulating an image in a manner that is consistent with the way the HVS actually behaves. A

colour correction algorithm is said to be inspired by the HVS if it maps an input RGB image $\mathrm{I}(x) = (\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x))$ to a processed image $\mathrm{J}(x) = (\mathrm{J}^R(x), \mathrm{J}^G(x), \mathrm{J}^B(x))$ by means of a series of operations derived from some of the HVS internal mechanisms. In general, a HVS-inspired method should implement the following characteristics [61], that are closely related to the observations of Land and McCann:

P.1 It should mimic the *differential nature* of colour perception. The way we perceive an image has been experimentally demonstrated to come from differential comparisons that our brain does between different areas of it.

P.2 It should take into account the *spatially variant nature* of colour perception. Our visual sense is affected by the context in which the object we are looking to is situated.

P.3 It should simulate the *non-linear behaviour* of the HVS. When our visual system is locally or globally stimulated, the response it produces is nonlinear. Consequently, any method that aims at reproducing the way we perceive stimuli should also contain non-linear processes.

P.4 It should try to reproduce the HVS property of *colour constancy*. This means it should be stable to some extent, when the illuminant of the scene varies.

A perceptual algorithm is one that relies on a simplified mathematical model of these (and others) low-level mechanisms of human visual perception. Several image processing and computer vision tasks can be approached from this point of view, such as colour stabilization, tone mapping, feature and object recognition (robust to changes in the illumination, for example), and so on. In the case of interest for us, image restoration and enhancement, the idea is to design a method that simulates the perception of humans by mimicking the above mechanisms. In this way, the objective of perceptually-inspired image enhancement methodologies is to produce an image that is more natural-looking for the observer, since it is processed in a way consistent with respect to human perception.

For colour images, two well established methods that are used to implement lightness and colour constancy are Gray-World and White-Patch [62].

Regarding Gray-World, it is based on the idea that the lightness constancy makes us perceive the objects that reflect the average luminance of a scene as medium gray. To be consistent with this phenomenon, we should enforce that the distribution of intensities in an image has its center of mass around the middle value. Implementing this property independently for each of the chromatic channels can also aid in suppressing some colour cast present in the scene.

The White Patch (WP) hypothesis is based on the fact that the brightest area in each spectral channel reflects all the illuminant energy on that channel. The HVS, to recover constant colours, relies then on the brightest patch on the scene, considering it as its white reference, and normalizing the rest of the reflectances with respect to it. Thus, to be consistent with the perceptual processing of an image, we should normalize its channel values, maximizing towards an hypothetical white reference. In

practice, this translates to the requirement that there is at least one white pixel in the image.

2.2.1.1 Automatic Colour Enhancement - ACE

Let us now analyse a recently developed colour correction method able to implement most of the hypothesis explained above, called ACE, for Automatic Colour Equalization [63]. It is thus a perceptual image enhancement method. Although it was developed as a one-step filter, posterior research revealed deep connections with the variational histogram equalization we presented in section 2.1.1.

Let : $\Omega \times \Omega \to (0, 1]$ be an image, where domain Ω is $[0, 1]$. The ACE filter maps the intensity of a pixel x to a new intensity $\mathrm{ACE}(x)$ in two stages, as follows:

$$\mathrm{I}(x) \mapsto R(x) = \frac{\sum_y \omega(x, y) s_\alpha(\mathrm{I}(x) - \mathrm{I}(y))}{\sum_y \omega(x, y)}, \tag{2.19}$$

$$R(x) \mapsto \mathrm{ACE}(x) = \frac{1}{2} + \frac{R(x)}{2M}, \tag{2.20}$$

where $M = \max_{x \in \Omega}(R(x))$, $\omega(x, y)$ is a weighting function, and $s_\alpha : [-1, 1] \to \mathbb{R}$ is a slope function defined as follows:

$$s_\alpha(t) = \begin{cases} -1, & \text{if } -1 \leq t \leq -\frac{1}{\alpha}, \\ \alpha t, & \text{if } -\frac{1}{\alpha} \leq t \leq \frac{1}{\alpha}, \\ 1, & \text{if } \frac{1}{\alpha} \leq t \leq 1. \end{cases} \tag{2.21}$$

Notice that in the limit of $\alpha \to \infty$, the slope function s_α tends to the sign function (2.6).

The weighting function $\omega : \Omega \times \Omega \to]0, 1[$ is a monotonically decreasing function, capturing the locality of the mutual chromatic influence of pixels, adding the capability to account for property **P.2** of the HVS explained above. Common choices of ω are inverse Euclidean distance function or a Gaussian decaying function centred at pixel x.

Contrarily to Retinex, ACE is implemented through differences of pixel intensities, rather than ratios. This is a more natural choice, if one is to implement the differential nature of spatial comparisons **P.1** jointly with the Gray World hypothesis **P.4**. It is in fact this local weighting and the term 1/2 in Eq. (2.19) who induce a local GW behaviour [61].

Non-linearity of ACE's response **P.3** is induced by the slope function (2.21). This mechanism amplifies small differences of intensity, while saturating large ones. This is because for small intensity differences, falling in the range $-1/\alpha < \mathrm{I}(x) - \mathrm{I}(y) < 1/\alpha$, s_α will increase the contrast as it maps $\mathrm{I}(x) - \mathrm{I}(y)$ to $\alpha(\mathrm{I}(x) - \mathrm{I}(y))$, and α is greater than one. On the other hand, large differences will be mapped to 0 or 1, depending on their sign. This amounts to expanding or compressing the available dynamic range according to the local image content.

Finally, it always exists a pixel x_0 such that $R(x_0) = \max_{x\in\Omega}(R(x)) = M$, and the intensity of that pixel will be mapped to 1 after application of (2.20). This enforces a global White Patch behaviour, further fulfilling **P.4**.

The implementation of all these mechanisms turns ACE into a (very) simplified model of the HVS, as the enhancement it achieves is consistent with human perception.

2.2.1.2 Improving ACE: Perceptual Colour Correction with Variational Techniques

ACE exhibits a remarkable ability to remove colour cast from images, thanks to the GW hypothesis, and is able to locally increase contrast. Also, it is able to deal with over and underexposed images, since it can increase the intensity of some pixels while decreasing intensity of others. However, it suffers from a series of drawbacks. For instance, it is computationally quite intensive. Moreover, uniform areas are mapped to the middle gray, not taking into account the original colour. Besides, it tends to overenhance some image areas, washing out the result.

To circumvent all these issues, in 2007, Bertalmío *et al.* [3] introduced a perceptually-inspired variational framework for contrast enhancement. The authors proposed an energy that, when minimized, leads to a local contrast enhancement of the initial image, while still respecting the GW and WP assumptions. This framework incorporated an *attachment to data* constraint that serves as a regularization mechanism, preventing overenhacement and avoiding chromatic noise introduction.

In its most general form, this energy can be written as:

$$E_{\omega,\tilde{\omega}}(\mathrm{I}) = D_{\tilde{\omega}}(\mathrm{I}) - C_{\omega,J}(\mathrm{I}). \tag{2.22}$$

The first term in the right hand side of Eq. (2.22) is a dispersion measure, meant to control the departure of the original image from an ideal one, and the second term provides local contrast enhancement.

The dispersion term is formulated as a linear combination of two energies, both of the form:

$$D_{\tilde{\omega}}(\mathrm{I}) = \iint_{\Omega^2} \tilde{\omega}(x,y)(\mathrm{I}(x) - \bar{\mathrm{I}}^{\tilde{\omega}}(y)\,dxdy. \tag{2.23}$$

The first of these dispersion term is obtained when we set $\omega(x,y)$ to be a normalized kernel , i.e.,

$$\int_\Omega \omega(x,y)dy = 1 \quad \forall x,$$

and we take $\bar{\mathrm{I}}^{\tilde{\omega}}(y) = \mu$ for every y:

$$D_\mu(\mathrm{I}) = \int_\Omega (\mathrm{I}(y) - \mu)^2\,dy.$$

By further writing $\mu = 1/2$, we obtain a term that enforces the Gray World hypothesis:

$$D_{1/2}(\mathrm{I}) = \int_\Omega (\mathrm{I}(y) - \frac{1}{2})^2\,dy.$$

The second part of the dispersion control is the Attachment to Data term. If $\mathrm{I}_0(x)$ is the original image to be enhanced, and we set:

$$\bar{\mathrm{I}}^{\tilde{\omega}} = \int_\Omega \tilde{\omega}(x,y)\mathrm{I}_0(y)\,dy,$$

we obtain the dispersion term $D_{\tilde{\omega},\mathrm{I}^0}$. For the particular case of the kernel given by $\omega(x,y) = \delta(x-y)$, being δ the Dirac delta, we get:

$$D_{A2D}(\mathrm{I}) = \int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))^2\,dx. \tag{2.24}$$

Regarding the local contrast enhancement measure, given an even functional $\mathrm{J} : \mathbb{R} \to [0,\infty[$, we define:

$$C_{\omega,J}(\mathrm{I}) = \iint_{\Omega^2} \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))\,dx. \tag{2.25}$$

We obtain now a complete energy formulation by writing:

$$E_{\mu,\omega}(\mathrm{I}) = \beta\, D_\mu(\mathrm{I}) + \gamma\, D_{A2D}(\mathrm{I}) - C_{\omega,J}(\mathrm{I}). \tag{2.26}$$

Now, to minimize (2.26) we need to compute first the variational derivatives of every term. These are derived in the following Proposition:

Proposition 2.2.1 *Whenever J is even and differentiable, and ω is symmetric, we obtain:*

1. $\delta C_{\omega,J}(\mathrm{I})(x) = 2\int_\Omega \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))\,dy.$
2. $\delta D_\mu(\mathrm{I})(x) = 2(\mathrm{I}(x) - \mu).$
3. $\delta D_{A2D}(\mathrm{I})(x) = 2(\mathrm{I}(x) - \mathrm{I}_0(x)).$

A proof of these statements can be found in [3], and we develop it also in Appendix A for completeness. Note that in the original work [3], proof is also provided for the more general case of J being an even and convex functional. This includes the absolute value function as a particular case.

Putting the pieces together, we now can write the Euler-Lagrange equation that an image $\mathrm{I}^*(x)$ must satisfy to be a minimizer of the energy functional (2.26):

$$2\beta(\mathrm{I}^*(x) - \mu) + 2\gamma(\mathrm{I}^*(x) - \mathrm{I}_0(x)) - 2\int_\Omega \omega(x,y)J'(\mathrm{I}^*(y) - \mathrm{I}^*(x))\,dxdy = 0. \tag{2.27}$$

Part of the interest of this formulation is that it encompasses both the variational formulation of histogram equalization given in Section 2.1.1 and ACE as particular cases. For the first one, by comparing equations (2.10) and (2.26) when $\mu = 1/2$ and $\gamma = 0$, we directly see that this new energy formulation is a scaled version of the variational histogram equalization functional proposed in [8]. The two key differences are that now we have an attachment to data term that reduces the typical overenhancement that histogram equalization produces, and that the contrast enhancement

term is modulated by a weighting function that incorporates a local behaviour in the energy functional.

To see the connection with ACE, let us write a discrete version of (2.26), again with $\mu = 1/2$ and $\gamma = 0$. This is, consider the following energy:

$$E(\mathrm{I}) = \frac{1}{2}\sum_{x\in\Omega}(\mathrm{I}(x) - \frac{1}{2})^2 + \frac{1}{4M}\sum_{x\in\Omega}\sum_{y\in\Omega}\omega(x,y)S_\alpha(\mathrm{I}(y)\mathrm{I}(x)), \tag{2.28}$$

being S_α such that $S'_\alpha = s_\alpha$. Since s_α was odd, S_α will be an even function. Intuitively, we can think in a regularized version of the absolute value as an example of S_α, and the sign function as its derivative. Euler-Lagrange equations of the energy (2.28) can be computed in the same way as in Proposition (2.2.1), yielding that if $\mathrm{I}(x)$ is a minimum, then:

$$(\mathrm{I}(x) - \frac{1}{2}) - 2\frac{4}{M}\sum_{y\in\Omega}\omega(x,y)s_\alpha(\mathrm{I}(x) - \mathrm{I}(y)) = 0,$$

that can be rewritten as:

$$\mathrm{I}(x) = \frac{1}{2} + \frac{1}{2M}\sum_{y\in\Omega}\omega(x,y)s_\alpha(\mathrm{I}(x) - \mathrm{I}(y)).$$

The above equality is exactly the process described in Equations (2.19) and (2.20). This shows that a minimum of energy (2.28) must be a fixed point of ACE, i.e., $\mathrm{I}(x) = \mathrm{ACE}(\mathrm{I})(x)$.

Regarding the numerical minimization of (2.28), the straightforward approach is to use a descent on the gradient direction, i.e., make $\mathrm{I}(x)$ evolve as $\frac{\partial \mathrm{I}}{\partial t} = -\delta E(\mathrm{I})$. Discretizing explicitly in time, we have:

$$\frac{\mathrm{I}^{k+1}(x) - \mathrm{I}^k(x)}{\Delta t} = (\frac{1}{2} - \mathrm{I}^k(x)) + \frac{1}{2M}R_{\mathrm{I}^k}(x), \tag{2.29}$$

where R_{I^k} is the contrast operator defined in Equation (2.19), applied to the k-th iteration. Eq. (2.29) can in turn be rewritten as:

$$\mathrm{I}^{k+1}(x) = (1 - \Delta t)\mathrm{I}^k(x) + \Delta t(\frac{1}{2} + \frac{1}{2M}R_{\mathrm{I}^k}(x)). \tag{2.30}$$

Selecting $\Delta t = 1$ and $M = \max_{x\in\Omega}\{R_{I_k}(x)\}$ yields the evolution $I^{k+1}(x) = \mathrm{ACE}^k(\mathrm{I}^0)$.

Unfortunately, as ACE does, evolving until steady state (2.30) leads to overenhanced images. It is here where the attachment to data term of Eq. (2.24) provides a regularizing mechanism useful to avoid this problem. Activating the regularizer by setting $\gamma = \lambda/2$ for $\lambda > 0$, and also writing for simplicity $\beta = \mu = 1/2$, we can write an explicit gradient descent for the full energy (2.22) as:

$$\frac{\mathrm{I}^{k+1}(x) - \mathrm{I}^k(x)}{\Delta t} = \frac{1}{2M}\sum_{y\in\Omega}\omega(x,y)s_\alpha(\mathrm{I}(x) - \mathrm{I}(y)) + (\frac{1}{2} - \mathrm{I}^k(x)) + \lambda(\mathrm{I}^k(x) - \mathrm{I}^0(x)),$$

that reduces to the descent given by:

$$\mathrm{I}^{k+1}(x) = (1-(1+\lambda)\Delta t)\mathrm{I}^k(x) + \Delta t(\underbrace{\frac{1}{2} + \frac{R_{\mathrm{I}^k}(x)}{2M}}_{\mathrm{ACE}(\mathrm{I}^k(x))}) + \lambda\Delta t\mathrm{I}^0(x).$$

The above equation reveals that ACE is part of this new framework, that also includes a term penalizing the departure from $\mathrm{I}^0(x)$ by adding together in the kth iteration scaled versions of $\mathrm{I}^k(x)$, $\mathrm{ACE}(\mathrm{I}^k(x))$, and $\mathrm{I}^0(x)$.

This work has been further generalized in several directions. In [10], Palma-Amestoy *et* al. suggested to employ, rather than a quadratic dispersion term given by:

$$D^{quad}_{\beta,\gamma}(\mathrm{I}) = \beta\sum_{x\in\Omega}(\mathrm{I}(x) - \frac{1}{2}) + \gamma\sum_{x\in\Omega}(\mathrm{I}(x) - \mathrm{I}^0(x))^2, \quad \beta,\gamma > 0,$$

an entropic dispersion term. To that end, the authors take the function $f(s) = a\log(a/s) - (a-s)$, with $a > 0, s \in]0,1[$. Then, $f'(s) = 1 - a/s$ and $f''(s) = a/s^2 > 0$, so $s = a$ is a global minimum of f. Defining:

$$D^{entr}_{\beta,\gamma}(\mathrm{I}) = \beta\underbrace{\sum_{x\in\Omega}\frac{1}{2}\log(\frac{1/2}{\mathrm{I}(x)}) - (\frac{1}{2} - \mathrm{I}(x))}_{f(\mathrm{I}(x)),\ \text{with}\ a=1/2} + \gamma\underbrace{\sum_{x\in\Omega}\mathrm{I}^0(x)\log(\frac{\mathrm{I}^0(x)}{\mathrm{I}(x)}) - (\mathrm{I}_0(x) - \mathrm{I}(x))}_{f(\mathrm{I}(x)),\ \text{with}\ a=\mathrm{I}^0(x)},$$

we have a minimum of this regularization term whenever $\mathrm{I}(x) = 1/2$ and $\mathrm{I}(x) = \mathrm{I}^0(x)$. The entropic term represents a compromise of both assumptions, balanced by γ and β.

Another generalization proposed in [10] was to substitute the term in Eq. (2.25) by the minimization of different inverse contrast measures, such as:

$$C^{\log}_{\omega}(\mathrm{I}) = \frac{1}{4}\sum_{x\in\Omega}\sum_{y\in\Omega}\omega(x,y)\log(\frac{\min(\mathrm{I}(x),\mathrm{I}(y))}{\max(\mathrm{I}(x),\mathrm{I}(y))}). \tag{2.31}$$

In fact, the authors claimed that any non-increasing function of the variable $t = \dfrac{\min(\mathrm{I}(x),\mathrm{I}(y))}{\max(\mathrm{I}(x),\mathrm{I}(y))}$ represents a local contrast-increase. Notice that, as long as $\max(\mathrm{I}(x),\mathrm{I}(y))$ increases and/or $\min(\mathrm{I}(x),\mathrm{I}(y))$ decreases, minimizing a term like the one in (2.31) is equivalent to performing a local contrast stretching.

2.2.2 Modelling the Degradation

In this section, we approach the problem of improving an image from a different point of view. The idea now is, if we (at least partially) know the physical phenomenon that produced the degradation of an input image, and we are able to model it mathematically, we could try an invert that model, apply to the corrupted image the resulting method and expect to find a restored image with better visual features. A quite general model for image degradation can be formulated as:

$$\mathrm{I}(x) = D(\mathrm{J}(x)) + \eta(x), \tag{2.32}$$

where D is the degradation function, and η represents additive noise.

In practice, however, solving for J in the above equation is a really challenging problem. Multiple degradation processes can occur that are usually quite difficult to model, tending to have many degrees of freedom that are not easy to fix. Fortunately, in some situations simplified versions of these models can still do a good job in restoring a degraded image. This is the case for example when equation (2.32) is a completely linear process and spatially invariant. Then, it can be shown that the degraded image can be written in the spatial domain as:

$$\mathrm{I}(x) = h(x) * \mathrm{J}(x) + \eta(x), \tag{2.33}$$

where now h is the spatial representation of the degradation function, and the symbol $*$ indicates spatial convolution. Given that convolution in the spatial domain amounts to multiplication in frequency, we can write the model (2.33) in its equivalent frequency domain representation as:

$$G(u) = H(u)F(u) + N(u), \tag{2.34}$$

where G, H, F, N are the Fourier transform of $\mathrm{I}, h, \mathrm{J}, \eta$ respectively. The pair of equations (2.33) and (2.34) are the basis to a large number of image restoration methods.

Particular cases of equation (2.33) instantiate different image restoration problems. For example, when $h(u) = u$, i.e. the identity operator, we face the denoising problem. When η is null, but h is not the identity, the resulting problem is known as deblurring. Usually the size and shape of the blur is assumed to be constant, or at most Gaussian. Dropping this assumption leads to a really challenging problem known as blind image deconvolution, that has historically attracted much interest in the image processing community [64–66].

We now turn to a particular kind of degradation, known as distance-dependent degradation. This process cannot be modelled linearly, as some parts of the image are affected more than others, and thus, it is spatially variant. However, it is of great interest for us, since the kind of degradation we will try to undo in the next chapters is precisely distance-dependent, and it is the media that lies in between the observer and the imaged objects what produces it.

2.2.2.1 Depth-Dependent Degradation and the Dark Channel Method

Of particular relevancy for this dissertation is a kind of degradation that occurs with more intensity as the object in the scene is further away from the observer. This is the case for example under bad weather conditions. The longer the path the signal travels until it reaches the observer is, the more it is attenuated. Unfortunately, the direct approach of restoring that signal by amplifying it in each pixel at the same time would also amplify noise, and close by objects would be over-enhanced in an attempt to improve contrast of far way parts of the scene.

In recent years, the particular case of restoring a degradation that has been produced by the atmosphere has been tackled with reasonable success by the Computer Vision community thanks to the introduction of different techniques able to invert a

simplified model of that degradation. Later in this dissertation we will review some of these methods, but one of them, the Dark Channel Method introduced in [7, 12], is of particular interest to us, since it forms the basis for the underwater restoration technique we introduce in the next chapter.

The Dark Channel technique, as many other related approaches, attempts to invert a physical model of image formation under atmospheric degradation that was suggested by Kochsmieder in [67]:

$$\mathrm{I}(x) = \mathrm{J}(x)\mathrm{t}(x) + \mathrm{A}(x)(1 - \mathrm{t}(x)). \tag{2.35}$$

This equation deserves a detailed analysis. It describes a three dimensional scenario like the one depicted in Fig. (2.10). In (2.35), x is a point located in the scene, and:

- $\mathrm{I}(x)$ is the observed intensity, *i*.e., the input hazy colour image.
- $\mathrm{J}(x)$ is the scene radiance, which corresponds to the non-degraded image.
- $\mathrm{t}(x)$ is a scalar quantity called transmission, that describes the amount of light that is not scattered nor absorbed, and reaches the observer. According to the Beer-Lambert law, transmission is exponentially decreasing with respect to the distance. Hence, given an object in the scene, its transmission depends both in the composition of the portion of the atmosphere that the light beam has traversed to reach it and the distance from the observer. A common simplification is to assume the atmosphere is homogeneous, which allows us to write:

  $$\mathrm{t}(x) = e^{-\beta d(x)}, \tag{2.36}$$

  where $d(x)$ is the scene depth and β is the (constant) scattering coefficient; in atmospheric models, it is usually considered that absorption is negligible compared to scattering effects, i.e., that atmosphere does not absorb light.
- The second term, $\mathrm{A}(x)(1-\mathrm{t}(x))$, represents the degradation due to haze effect. Notice that according to (2.35), the scene radiance first goes through a multiplicative distortion, which result is the direct attenuation $\mathrm{J}(x)\mathrm{t}(x)$, and then through an additive distortion, produced by this second term. Here the airlight A plays the role of the colour of the haze, which is usually considered constant over the scene, and therefore in a channel-wise formulation it is a scalar value.

  One can see that at zero depth, transmission becomes $\mathrm{t}(x) = 1$, and airlight attenuation has no influence in (2.35), that reduces to $\mathrm{I}(x) = \mathrm{J}(x)$. Conversely, as the depth in the scene increases, transmission decreases to zero, making the airlight contribution in (2.35) exponentially more significative, until it dominates the whole scene, properly modelling what happens in the horizon of natural hazy scenes:

  $$\lim_{\mathrm{t}(x)\to 0} \mathrm{I}(x) = \mathrm{A}(x).$$

  Notice also that the hypothesis of the homogeneity of the atmosphere has another simplifying consequence: the airlight can be considered as constant across the scene, i.e., $\mathrm{A}(x) = \mathrm{A}$.

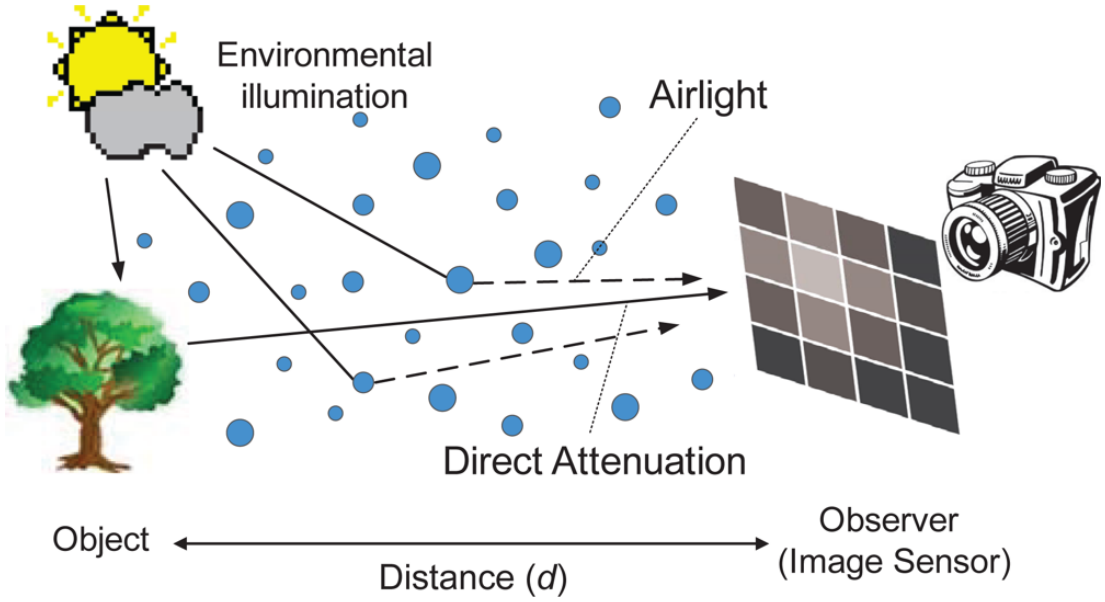


Figure 2.10: Atmospheric Model of Haze, reproduced from [1].

When we deal with two-dimensional RGB images, the goal of a dehazing algorithm is to invert (2.35) and recover J. Computationally, we have an RGB input image I, and the goal is to estimate a haze-free version J, also RGB. But the transmission $\mathrm{t}(x)$, a scalar image, and the airlight, another RGB image, need to be calculated too in (2.35), which leaves us with 3 equations and 7 unknowns for each pixel in the image. This means $3N + N + 3 = 4N + 3$ unknowns and $3N$ constraints in total. This accounts for the ambiguity of not knowing whether the colour of an object in our $2D$ image is far away from the observer, and thus affected by the haze, or it is its true colour, and the object is in fact located near the observer in the real scene.

Scene radiance recovery is theoretically possible from equation (2.36), if we assume airlight and transmission are known, by simply rearranging and writing:

$$\mathrm{J}(x) = \frac{(\mathrm{I}(x) - \mathrm{A})\mathrm{t}(x) + \mathrm{A}}{\mathrm{t}(x)}.$$

However, this is not the general case. As a result, there is a need to introduce some kind of prior information. In the Dark Channel method, this prior is introduced through the observation that a haze-free outdoor image exhibits shadows and textures everywhere, and thus, locally at least one colour channel has some pixel with a low intensity, near the black colour. The authors propose to build an image in which for each pixel they store the minimum value among the R, G and B channels of the corresponding image in the input image, and inspect the degree of "darkness" of it. This new image is called Dark Channel, and it is defined as:

$$\mathrm{J}^{\mathrm{dark}}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{R,G,B\}} \mathrm{J}^c(y) \right),$$

where $\Omega(x)$ is a neighbourhood of pixels around the x location, and the minimum is

first over the three colour channels of the observed image J and then over each local patch. The Dark Channel Prior then simply states that for outdoor haze-free images, for almost every x, it holds that:

$$\mathrm{J}^{\mathrm{dark}}(x) \to 0. \tag{2.37}$$

In Fig. (2.11), we can see a haze-free image together with its dark channel, which does not present bright areas as expected (unless in the sky region), and in Fig. (2.12) we can find a hazy image, with its bright dark channel.

Figure 2.11: A haze-free natural scene and its Dark Channel.

Figure 2.12: A hazy natural scene and its Dark Channel.

After assuming the Dark Channel prior, the estimation of airlight and transmission becomes straightforward. Supposing for the moment that the airlight has already been computed and divide Eq. (2.35) by it, we obtain:

$$\frac{\mathrm{I}^c(x)}{\mathrm{A}^c} = \mathrm{t}(x)\frac{\mathrm{J}^c(x)}{\mathrm{A}^c} + 1 - \mathrm{t}(x), \qquad c \in \{R, G, B\}. \tag{2.38}$$

A second assumption needs to be done now. We suppose that transmission (and thus depth) is locally constant in a natural image. Denoting this approximated transmission by $\tilde{\mathrm{t}}(x)$, we now take minima over every channel on (2.38), and again over

each local patch; since $\tilde{\mathrm{t}}(x)$ is constant it can be taken out of the minima operators, yielding:

$$\min_{y\in\Omega(x)}\left(\min_{c}(\frac{\mathrm{I}^c(x)}{\mathrm{A}^c})\right)=\tilde{\mathrm{t}}(x)\min_{y\in\Omega(x)}\left(\frac{\mathrm{J}^c(x)}{\mathrm{A}^c}\right)+1-\tilde{\mathrm{t}}(x),\qquad c\in\{R,G,B\}. \tag{2.39}$$

Now, we use the fact that J is a haze-free image, and its dark channel is almost a close-to-zero image. As the airlight must be a positive quantity, we can write:

$$\min_{y\in\Omega(x)}\left(\frac{\mathrm{J}^c(x)}{\mathrm{A}^c}\right)\approx 0. \tag{2.40}$$

Inserting (2.40) in (2.38) provides the estimate for the transmission we were looking for:

$$\tilde{\mathrm{t}}(x)=1-\min_{y\in\Omega(x)}\left(\min_{c}(\frac{\mathrm{I}^c(x)}{\mathrm{A}^c})\right). \tag{2.41}$$

The fact that we can recover transmission from the Dark Channel of a hazy image makes physical sense, since haze usually becomes denser as distance from the observer increases, and it is that information what the Dark Channel prior captures. Notice that equation (2.36) directly relates transmission and depth of the scene.

As we see in Fig. (2.13), the assumption that the transmission map can be considered patch-wise constant is unrealistic. If the patch has inside a depth jump, this hypothesis fails and leads to artifacts in the recovered scene. To overcome this issue, a refinement step must be performed, in which the rough transmission-map estimate is bilaterally filtered with the original image to recover fine structures in the depth image.

Regarding the airlight, represented by the colour of the horizon, it can be extracted from the Dark Channel of the input image. Since the effect of haze is additive on a haze-free image, one can consider that a bright dark channel contains haze only, and estimate the airlight A from it. The brightest pixel in the dark channel can be considered to be the representative of the haze colour, and its correspondent pixel in the hazy image is selected to be the airlight A.

The final step is simply to invert in (2.35). Usually a lower bound is inserted in the denominator to prevent dividing by small numbers:

$$\mathrm{J}(x)=\frac{\mathrm{I}(x)-\mathrm{A}}{\max(\mathrm{t}(x),\mathrm{t}_0)})+\mathrm{A}, \tag{2.42}$$

where a typical value for t_0 can be 0.1.

Equation (2.42) represents one of the most powerful methods to remove haze from weather-degraded images. Its strength lies on its simplicity. Previous attempts to fulfil the dehazing goal were cumbersome, some of them requiring complex image acquisition techniques or very involved numerical methods. A sample of the impressive visual results produced Dark Channel Method can be appreciated in Fig. (2.14).

(a) Hazy input image

(b) Corresponding Dark Channel.

(c) Corresponding transmission map.

(d) Detail of the lamp-post, containing a depth jumps, and its computed transmission.

Figure 2.13: Hazy scene with strong depth discontinuities, and illustration of how the unrefined transmission map fails to to capture depth differences.

Figure 2.14: (a) Hazy urban scene. (b) Result of applying the Dark Channel corresponding to the image of Fig. 2.14b.

Chapter 3

Underwater Image Restoration

In the last section of the previous chapter, we have discussed the problem of restoring an image that has lost its contrast in far away regions of the scene due to a distance-dependant degradation. The main related application is that of image dehazing, since fog fits nicely with this definition of degradation. On a hazy regime, contrast in the scene is adequate in areas of the image that lie near the observer, while it uniformly degrades as we move away from such observer, exhibiting colours exponentially increasing towards white.

However, it turns out that not only haze, but also other phenomena follow these scheme of degradation. The overall idea of distance-dependant degradation is that the media lying in between the observer and objects in the scene is not transparent, and it gets "thicker" with distance. In this chapter, we analyse the case of images acquired in an underwater environment, and the problem of improving their quality.

Underwater images also suffer from low contrast as a result of the exponential decay that light suffers as it travels. Hence, this problem shares some of the characteristics of image dehazing. However, underwater images also exhibit a particular colour distortion associated to different wavelengths having different attenuation rates, being the red wavelength the one that attenuates the fastest. In other words, while in a hazy scenario every colour degrades at the same rate towards white, in the underwater setting, wavelengths associated typically to the blue and green colours dominate, while the wavelength associated with the red colour tends to vanish rapidly.

To overcome this unbalanced loss of contrast and colour distortion, we will develop a new approach, based on a modified Dark Channel prior. We call this technique the Red Channel method. The *Red Channel method* is designed to restore the lost contrast while recovering colours associated to short wavelengths. After briefly reviewing some previous work related to the specific application of underwater image enhancement, we will explain the details behind this extension of the Dark Channel method to underwater images. In particular, we will explain how to deal with artificially illuminated areas, which are typical of underwater images. A comprehensive set of experimental results demonstrate that the proposed technique is capable of achieving a natural colour correction and superior or equivalent visibility improvement when compared to other state of the art methods.

3.1 Introduction

The field of underwater imaging has drawn much attention over the last years due to the wide range of existing applications, from marine biology and archaeology [68] to ecological research [69]. The improvement of remotely operated vehicles (ROV) navigation capabilities [70] is also a very important field of application. For example, underwater vehicle navigation can be guided through video mosaics [71]. Mosaicking requires a previous registration of different input images [72], and image processing techniques are usually implemented as a preprocessing step to improve the results.

The problem of restoring images degraded by an underwater environment is challenging, in part because light travelling underwater suffers from two combined degradations, known as scattering and absorption. The former consists on a change of direction of light after collision with particles, and the latter explains how particles absorb light. The main consequences of these degradation processes are a decrease in the visibility distance and a colour distortion that depends on the wavelength of each light beam. The shorter wavelengths (green and blue colours), reach further depths in the scene than the rapidly vanishing longest ones (red), transferring to the final image a characteristic bluish-greenish tone. Together with turbidity of the water, and the organic particles suspended on the medium, this yields to a hard restoration problem, since overall algorithms performance becomes highly dependent on environmental conditions. Often, an artificial light source is added to the imaging device to try to increment the visibility range in the scene, but this entails some disadvantages: the amount of power needed to illuminate underwater scenarios can be prohibitively expensive; moreover, illumination is not obtained in a global uniform way and a bright area in the center of the image tends to appear, as backscattering occurs from nearby particles causing visual disturbance.

The restoration of images degraded by an underwater environment with image processing techniques has been studied in various ways. For instance, in [73, 74] authors suggest an algorithm that takes as input two images obtained through a polariser that is rotated to work at different orientations, obtaining in this way extra information about the scene that facilitates the inversion of the image formation process. Also, a supervised algorithm is presented in [75], where parameters of colour correction are learnt over training data. The above methods require additional information, and do not work on single images without extra input. Single-image approaches can be found in [76], where they employ an image fusion strategy, or [77] and [78], where authors are more focused on colour treatment. A more exhaustive review of these and other works can be found in the recent survey [79].

Underwater image restoration can also be considered as a distance-dependent degradation problem. From this point of view, it is interesting to recall the atmospheric degradation model of Kochsmieder, presented in the previous chapter:

$$\overrightarrow{\mathrm{I}}(x) = \overrightarrow{\mathrm{J}}(x)\mathrm{t}(x) + \overrightarrow{\mathrm{A}}(1 - \mathrm{t}(x)) \tag{3.1}$$

In (3.1), $\overrightarrow{\mathrm{I}}$ is the observed *intensity*, the input degraded colour image, t is the *transmission*, which describes the amount of light that is not scattered nor absorbed, and

reaches the observer (inversely related to the depth in the scene), $\overrightarrow{\mathrm{A}}$ is the *airlight*, physically related to the colour of the haze, and $\overrightarrow{\mathrm{J}}$ is the scene *radiance*, which is associated to the non-degraded image.

As outlined in the previous chapter, one of the most powerful methods to invert eq. (3.1) is the recently developed Dark Channel method. Unfortunately, the nature of the degradation induced by a marine environment prevents from directly applying the Dark Channel method. Absorption, which is not present in the atmosphere, since it is considered to be transparent, influences the way in which colours are lost in an underwater scene. Red channel rapidly loses intensity, while the green and blue ones keep their intensity longer. In this situation, the Dark Channel Prior has no sense anymore: being the image degraded or not, there is almost always one colour channel with low intensity (the red one).

Several authors have proposed various modifications of the Dark Channel Method method to solve model (3.1) adapted to the underwater environment. In [80], authors modify the algorithm prior to reflect the difference in attenuation between colours, while in [81] they keep the algorithm essentially untouched and prefer to modify the model by adding some parameters reflecting the disparity of degradation in different wavelengths and depths. Unfortunately, while in [80] the different rates of colour degradations are not taken into account, the method in [81] strongly depends on the choice of the parameters that govern the algorithm, limiting the capability of both approaches.

In the next pages, a Red Channel method is proposed that is suitable for underwater images. It can be interpreted as a variant of the Dark Channel method employed in dehazing of atmospheric images. **The Red Channel method is simple and robust, and it recovers part of the lost visibility range while correcting the colour distortion produced by absorption. Additionally, it has fewer free parameters than previously existing methods, and can cope efficiently with artificial light sources possibly present in the scene.**

3.2 Red Channel Underwater Image Restoration

Examining in detail (3.1), we observe that the scene radiance $\overrightarrow{\mathrm{J}}(x)$ first goes through a multiplicative distortion and then through an additive distortion. According to the Beer-Lambert law, transmission t is exponentially decreasing with respect to the distance. For the case of haze in the atmosphere, the transmission of an object in the scene depends both on the composition of the portion of air that the light beam has traversed to reach it, and the distance from the observer. Assuming that the atmosphere is homogeneous allows us to write:

$$\mathrm{t}(x) = e^{-\beta d(x)}, \tag{3.2}$$

where $d(x)$ is the scene depth and β is the (constant) attenuation coefficient. From Eq. (3.2), we see that at zero depth, the transmission becomes $\mathrm{t}(x) = 1$, and airlight attenuation has no influence in Eq. (3.1), which reduces to $\overrightarrow{\mathrm{I}}(x) = \overrightarrow{\mathrm{J}}(x)$. Conversely, as the depth in the scene increases, transmission decreases to zero, making

the airlight contribution in (3.1) exponentially more significative, until it dominates the whole scene, effectively modelling what happens in the horizon of natural hazy scenes, $\overrightarrow{\mathrm{I}}(x) \approx \overrightarrow{\mathrm{A}}$.

All the above intuition can be translated to the underwater environment with some adaptations. To that end, we rearrange the original model (3.1) and write:

$$\begin{aligned} \mathbf{1} - \mathrm{I}^R &= \mathrm{t}(1 - \mathrm{J}^R) + (\mathbf{1} - \mathrm{t})(\mathbf{1} - \mathrm{A}^R) \\ \mathrm{I}^G &= \mathrm{t}\mathrm{J}^G + (\mathbf{1} - \mathrm{t})\mathrm{A}^G \\ \mathrm{I}^B &= \mathrm{t}\mathrm{J}^B + (1 - \mathrm{t})\mathrm{A}^B, \end{aligned} \tag{3.3}$$

being $\overrightarrow{\mathrm{I}} = (\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B)$ and $\overrightarrow{\mathrm{J}} = (\mathrm{J}^R, \mathrm{J}^G, \mathrm{J}^B)$ the degraded and the original images respectively. Notice that (3.3) is completely equivalent to (3.1). Therefore, it still reflects the fact that the light attenuates with distance, as it actually happens underwater. The main difference we must have into account is that red intensity decays faster as distance increases. To reflect this fact, a possible solution is the following modification of the Dark Channel prior, denoted as Red Channel Prior. It states that:

$$\mathrm{J}^{\mathrm{RED}}(x) := \min\left(\min_{y\in\Omega(x)}(\mathbf{1} - \mathrm{J}^R(y)), \min_{y\in\Omega(x)}(\mathrm{J}^G(y)), \min_{y\in\Omega(x)}(\mathrm{J}^B(y))\right) \approx 0, \tag{3.4}$$

for a non-degraded underwater image. Here, $\Omega(x)$ is a neighbourhood of pixels around the x location. Notice that for a degraded image near the observer, the red channel still keeps some intensity, so its reciprocal $1 - \mathrm{J}^R$ is low, and the prior is still true. However, as distance increases, red intensity rapidly decays, and its weight in the Red Channel image decreases, so the prior begins to be false. This will allow us to estimate the depth in the scene, as well as the colour of the water veil.

Remark 3.2.1 *Notice that we denote the Red Channel of an image* $\overrightarrow{\mathrm{J}} = (\mathrm{J}^R, \mathrm{J}^G, \mathrm{J}^B)$*, given by Eq. (3.4), as* $\mathrm{J}^{\mathrm{RED}}$*, while its red component is denoted by* J^R*.*

3.2.1 WaterLight estimation from Red Channel

The first thing we need to estimate is the colour of the water. One ideally wants to pick a pixel that lies at the maximum depth with respect to the camera. Since we are assuming that the degradation is distance dependent, this location naturally corresponds to the maximum values in the Red Channel of the original image.

Let us denote $\overrightarrow{\mathrm{A}} = \overrightarrow{\mathrm{I}}(x_0) = (1 - \mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B)$. We choose as the waterlight x_0 the pixel in the degraded image that corresponds to the brightest pixel in its Red Channel:

$$\overrightarrow{\mathrm{A}} = (\mathrm{I}^R(x_0), \mathrm{I}^G(x_0), \mathrm{I}^B(x_0)) \text{ such that } \mathrm{I}^{\mathrm{RED}}(x_0) \geq \overrightarrow{\mathrm{I}}^{\mathrm{RED}}(x) \quad \forall x.$$

The analogous method to the one suggested in [7] would be to take the top 10% brightest pixels in the Red Channel image, and among this set of pixels, select the

(a) (b) (c) (d)

Figure 3.1: Waterlight detection from Red channel. (a) Original image, the red rectangle marks the region from where waterlight was extracted (b) In red, exact pixel location of waterlight within the rectangle marked in (a) (c) Red Channel of the original image (d) Final restoration result.

one that is brightest in the degraded image. This pixel does not need to be the brightest pixel in the Red Channel. However, the brightest pixel in the degraded image is an ambiguous definition. In the underwater case, picking among that 10% of pixels the one that has lower red component obtains the best results in the performed experiments, see Fig. 3.1.

3.2.2 Transmission Estimate from Red Channel

After estimating the waterlight, we are now able to produce an estimate for the transmission of the scene. This is achieved with the following result:

Theorem 3.2.2 *Given an estimate* $\overrightarrow{\mathrm{A}} = (\mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B)$ *for the waterlight, assuming that transmission is locally constant, and under the Red Channel hypothesis (3.4),* $\mathrm{t}(x)$ *can be estimated as:*

$$\widetilde{\mathrm{t}}(x) = \mathbf{1} - \min\left(\frac{\min\limits_{y\in\Omega(x)}(\mathbf{1}-\mathrm{I}^R(y))}{\mathbf{1}-\mathrm{A}^R}, \frac{\min\limits_{y\in\Omega(x)}\mathrm{I}^G(y)}{\mathrm{A}^G}, \frac{\min\limits_{y\in\Omega(x)}\mathrm{I}^B(y)}{\mathrm{A}^B}\right). \tag{3.5}$$

Proof: It is a simple adaptation of the steps of the derivation in [7]. Starting from Eqs. (3.3), we divide them by the waterlight:

$$\left(\frac{\mathbf{1}-\mathrm{I}^R}{\mathbf{1}-\mathrm{A}^R}, \frac{\mathrm{I}^G}{\mathrm{A}^G}, \frac{\mathrm{I}^B}{\mathrm{A}^B}\right) = \left(\mathrm{t}\frac{\mathbf{1}-\mathrm{J}^R}{\mathbf{1}-\mathrm{A}^R} + \mathbf{1} - \mathrm{t}, \mathrm{t}\frac{\mathrm{J}^G}{\mathrm{A}^G} + \mathbf{1} - \mathrm{t}, \mathrm{t}\frac{\mathrm{J}^B}{\mathrm{A}^B} + \mathbf{1} - \mathrm{t}\right),$$

and we take minima over local neighbourhoods on every image. Since waterlight are scalar images build out of constant values, they can be extracted from the minima operators. This is also true for the estimate of the transmission map, which we

consider to be a locally constant image, $\min_{\Omega(x)} \widetilde{\mathrm{t}}(x) = \widetilde{\mathrm{t}}(x)$, so we obtain:

$$\left(\frac{\min_{\Omega}(\mathbf{1} - \mathrm{I}^R)}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{\Omega} \mathrm{I}^G}{\mathrm{A}^G}, \frac{\min_{\Omega} \mathrm{I}^B}{\mathrm{A}^B} \right) =$$

$$\left(\widetilde{\mathrm{t}} \frac{\min_{\Omega}((\mathbf{1} - \mathrm{J}^R))}{\mathbf{1} - \mathrm{A}^R} + \mathbf{1} - \widetilde{\mathrm{t}}, \widetilde{\mathrm{t}} \frac{\min_{\Omega} \mathrm{J}^G}{\mathrm{A}^G} + \mathbf{1} - \widetilde{\mathrm{t}}, \widetilde{\mathrm{t}} \frac{\min_{\Omega} \mathrm{J}^B}{\mathrm{A}^B} + \mathbf{1} - \widetilde{\mathrm{t}} \right)$$

After taking minima at both sides of the above equation, we obtain:

$$\min \left(\frac{\min_{\Omega}(\mathbf{1} - \mathrm{I}^R)}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{\Omega} \mathrm{I}^G}{\mathrm{A}^G}, \frac{\min_{\Omega} \mathrm{I}^B}{\mathrm{A}^B} \right) =$$

$$\widetilde{\mathrm{t}} \min \left(\frac{\min_{\Omega}((\mathbf{1} - \mathrm{J}^R))}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{\Omega}(\mathrm{J}^G)}{\mathrm{A}^G}, \frac{\min_{\Omega}(\mathrm{J}^B)}{\mathrm{A}^B} \right) + \mathbf{1} - \widetilde{\mathrm{t}}$$

Direct application now of the Red Channel prior (3.4) cancels out the first term of the right hand side of the above equation. Simple rearrangement leads to the desired conclusion. □

Let us stress that model (3.3) is in fact a convex combination with coefficients t and $1 - \mathrm{t}$, so the estimate $\widetilde{\mathrm{t}}$ of t must also lie in $[0, 1]$. This is a non-trivial fact that relies on how we performed the estimation of the waterlight. We have:

Theorem 3.2.3 *Transmission estimate (3.5) verifies* $\widetilde{\mathrm{t}}(x) \in [0, 1]$ *for all pixel* x *in the image.*

Proof: According to Eq. (3.5), it is enough to prove that:

$$0 \leq \min \left(\frac{\min_{\Omega(x)} \mathbf{1} - \mathrm{I}^R(x)}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{\Omega(x)} \mathrm{I}^G(x)}{\mathrm{A}^G}, \frac{\min_{\Omega(x)} \mathrm{I}^B(x)}{\mathrm{A}^B} \right) \leq 1, \tag{3.6}$$

for every pixel x. Being every quantity in Ineq. (3.6) positive, the first inequality is automatically verified. To see that the inner expression is bounded by 1, let us recall that Airlight $\overrightarrow{\mathrm{A}}$ was defined to be located at the pixel x_0 verifying that:

$$\overrightarrow{\mathrm{A}} = (\mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B) = (\mathrm{I}^R(x_0), \mathrm{I}^G(x_0), \mathrm{I}^B(x_0)) \quad \text{s.t.} \quad \mathrm{I}^{RED}(x_0) \geq \mathrm{I}^{RED}(x) \quad \forall x.$$

Then, we take an arbitrary pixel location x^*, and we apply the definition of the Red Channel given in Eq.(3.4) to the degraded image $\overrightarrow{\mathrm{I}}$:

$$\min \left(\min_{y \in \Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y)), \min_{y \in \Omega(x^*)} (\mathrm{I}^G(y)), \min_{y \in \Omega(x^*)} (\mathrm{I}^B(y)) \right) \leq$$

$$\min \left(\min_{y \in \Omega(x_0)} (\mathbf{1} - \mathrm{I}^R(y)), \min_{y \in \Omega(x_0)} (\mathrm{I}^G(y)), \min_{y \in \Omega(x_0)} (\mathrm{I}^B(y)) \right) \tag{3.7}$$

Since we always have $x_0 \in \Omega(x_0)$, we can write:

$$\begin{aligned}
\min_{y\in\Omega(x_0)} (\mathbf{1} - \mathrm{I}^R(y)) &\leq \mathbf{1} - \mathrm{I}^R(x_0), \\
\min_{y\in\Omega(x_0)} \mathrm{I}^G(y) &\leq \mathrm{I}^G(x_0), \\
\min_{y\in\Omega(x_0)} \mathrm{I}^B(y) &\leq \mathrm{I}^B(x_0).
\end{aligned}$$

With this, we can rewrite Ineq. (3.7) as:

$$\begin{aligned}
\min\left(\min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y)), \min_{y\in\Omega(x^*)} \mathrm{I}^G(y), \min_{y\in\Omega(x^*)} (\mathrm{I}^B(y))\right) \leq \\
\min(\mathbf{1} - \mathrm{I}^R(x_0), \mathrm{I}^G(x_0), \mathrm{I}^B(x_0)) = \min(1 - \mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B). \qquad (3.8)
\end{aligned}$$

Now, for the pixel x^*, there are three possibilities regarding the left-hand-side of Ineq. (3.8), depending on which of the three inner quantities attains the minimum.

First Possibility: Suppose it is the first quantity, i.e.:

$$\min\left(\min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y)), \min_{y\in\Omega(x^*)} (\mathrm{I}^G(y)), \min_{y\in\Omega(x^*)} (\mathrm{I}^B(y))\right) = \min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y))$$

In this case, Eq. (3.8) can be used to write:

$$\min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y)) \leq \min(1 - \mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B) \leq 1 - \mathrm{A}^R$$

and we obtain that:

$$\frac{\min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y))}{\mathbf{1} - \mathrm{A}^R} \leq 1 \qquad (3.9)$$

Finally, using Ineq. (3.9), we arrive to the desired conclusion:

$$\min\left(\frac{\min_{y\in\Omega(x^*)} \mathbf{1} - \mathrm{I}^R(y)}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^G(y)}{\mathrm{A}^G}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^B(y)}{\mathrm{A}^B}\right) \leq \frac{\min_{y\in\Omega(x^*)} (\mathbf{1} - \mathrm{I}^R(y))}{\mathbf{1} - \mathrm{A}^R} \leq 1.$$

Second Possibility: Suppose now the minimum on the left hand side of Ineq. (3.8) is attained by $\min_{y\in\Omega(x)} \mathrm{I}^G(y)$. Just as in the first case, we can write:

$$\min_{y\in\Omega(x^*)} \mathrm{I}^G(y) \leq \min(1 - \mathrm{A}^R, \mathrm{A}^G, \mathrm{A}^B) \leq \mathrm{A}^G,$$

and we deduce now that:

$$\frac{\min_{y\in\Omega(x^*)} \mathrm{I}^G(y)}{\mathrm{A}^G} \leq 1,$$

which leads us to:

$$\min\left(\frac{\min_{y\in\Omega(x^*)} \mathbf{1} - \mathrm{I}^R(y)}{\mathbf{1} - \mathrm{A}^R}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^G(y)}{\mathrm{A}^G}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^B(y)}{\mathrm{A}^B}\right) \leq \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^G(y)}{\mathrm{A}^G} \leq 1.$$

Third Possibility: If the minimum in the left hand side of Ineq. (3.8) is attained by $\min_{y\in\Omega(x^*)} \mathrm{I}^B(y)$, similar derivations as in the previous cases lead us to:

$$\min\left(\frac{\min_{y\in\Omega(x^*)} \mathbf{1}-\mathrm{I}^R(y)}{\mathbf{1}-\mathrm{A}^R}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^G(y)}{\mathrm{A}^G}, \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^B(y)}{\mathrm{A}^B}\right) \leq \frac{\min_{y\in\Omega(x^*)} \mathrm{I}^B(y)}{\mathrm{A}^B} \leq 1.$$

With any of the three possibilities, the result of the theorem is verified. As the choice of pixel x^* was arbitrary, this concludes the proof.

□

Nevertheless, this straightforward extension of the Dark Channel Method to include the Red Channel Prior fails to take into account the velocity of different wavelengths in the light. To incorporate this behaviour in our model, we suggest two possibilities. The first one is performed in the transmission estimate step, and consists of extending the model (3.3) to involve three transmission maps, one for each colour component. The second one occurs during the final inversion step, and consists of weighting the additive contribution of each waterlight component. In this dissertation, we select the former approach.

Incorporating Colour Correction I: Vectorial Transmission

Starting from (3.3) and distinguishing three different transmission maps, we have:

$$\begin{aligned} \mathbf{1}-\mathrm{I}^R &= \mathrm{t}^R(1-\mathrm{J}^R)+(\mathbf{1}-\mathrm{t}^R)(\mathbf{1}-\mathrm{A}^R) \\ \mathrm{I}^G &= \mathrm{t}^G\mathrm{J}^G+(\mathbf{1}-\mathrm{t}^G)\mathrm{A}^G \\ \mathrm{I}^B &= \mathrm{t}^B\mathrm{J}^B+(1-\mathrm{t}^B)\mathrm{A}^B \end{aligned} \tag{3.10}$$

In (3.10), we need to estimate the three components of the waterlight, and the three transmission maps, one for each colour. However, these three terms are not independent: using (3.2), we conclude that we only need to estimate one matrix and two scalar numbers. We have:

$$\begin{aligned} \mathrm{t}^R(x) &= e^{-\beta^R d(x)} \\ \mathrm{t}^G(x) &= e^{-\beta^G d(x)} = \left(e^{-\beta^R d(x)}\right)^{\frac{\beta^G}{\beta^R}} = (\mathrm{t}^R(x))^{\lambda_G} \\ \mathrm{t}^B(x) &= e^{-\beta^B d(x)} = \left(e^{-\beta^R d(x)}\right)^{\frac{\beta^B}{\beta^R}} = (\mathrm{t}^R(x))^{\lambda_B}, \end{aligned} \tag{3.11}$$

where $\lambda_G = \beta^G/\beta^R$ and $\lambda_B = \beta^B/\beta^R$. This leads to a slightly more difficult (but still possible) estimation of the transmission, by modification of the Red Channel prior. We have following property:

Theorem 3.2.4 *Assuming that we have already estimated the waterlight, and the attenuation coefficients for the green and blue transmission maps, λ_G and λ_B, and*

that:

$$\min\Bigg(\min_{y\in\Omega(x)}(\mathbf{1}-\mathrm{J}^R(y))/(\mathbf{1}-\mathrm{A}^R),\mathrm{t}_R^{\lambda_G-1}(x)(\min_{y\in\Omega(x)}\mathrm{J}^G(y))/\mathrm{A}^G+$$
$$\mathbf{1}-\mathrm{t}_R^{\lambda_G-1}(x),\mathrm{t}_R^{\lambda_B-1}(x)(\min_{y\in\Omega(x)}\mathrm{J}^B(y))/\mathrm{A}^B+\mathbf{1}-\mathrm{t}_R^{\lambda_B-1}(x)\Bigg)\equiv 0, \tag{3.12}$$

the red, green, and blue components of the transmission map can be estimated as:

$$\begin{aligned}\widetilde{\mathrm{t}}^R(x)=&\mathbf{1}-\min\left(\frac{\min\limits_{y\in\Omega(x)}\mathbf{1}-\mathrm{I}^R(y)}{\mathbf{1}-\mathrm{A}^R},\frac{\min\limits_{y\in\Omega(x)}\mathrm{I}^G(y)}{\mathrm{A}^G},\frac{\min\limits_{y\in\Omega(x)}\mathrm{I}^B(y)}{\mathrm{A}^B}\right)\\ \widetilde{\mathrm{t}}^G(x)=&(\widetilde{\mathrm{t}}^R(x))^{\lambda_G}\\ \widetilde{\mathrm{t}}^B(x)=&(\widetilde{\mathrm{t}}^R(x))^{\lambda_B}\end{aligned} \tag{3.13}$$

Proof: Starting from Eqs. (3.10), we divide them by the waterlight:

$$\left(\frac{\mathbf{1}-\mathrm{I}^R}{\mathbf{1}-\mathrm{A}^R},\frac{\mathrm{I}^G}{\mathrm{A}^G},\frac{\mathrm{I}^B}{\mathrm{A}^B}\right)=\left(\mathrm{t}^R\frac{\mathbf{1}-\mathrm{J}^R}{\mathbf{1}-\mathrm{A}^R}+\mathbf{1}-\mathrm{t}^R,\mathrm{t}^G\frac{\mathrm{J}^G}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}^G,\mathrm{t}^B\frac{\mathrm{J}^B}{\mathrm{A}^B}+\mathbf{1}-\mathrm{t}^B\right),$$

and we take minima over local neighbourhoods on every image. As in the proof of Theorem (3.2.2), waterlight and transmission map estimates can be extracted from the minima operators, since they are locally constant images, yielding:

$$\begin{aligned}&\left(\frac{\min\limits_{\Omega}(\mathbf{1}-\mathrm{I}^R)}{\mathbf{1}-\mathrm{A}^R},\frac{\min\limits_{\Omega}\mathrm{I}^G}{\mathrm{A}^G},\frac{\min\limits_{\Omega}\mathrm{I}^B}{\mathrm{A}^B}\right)=\\ &\left(\mathrm{t}^R\frac{\min\limits_{\Omega}((\mathbf{1}-\mathrm{J}^R))}{\mathbf{1}-\mathrm{A}^R}+\mathbf{1}-\mathrm{t}^R,\mathrm{t}^G\frac{\min\limits_{\Omega}(\mathrm{J}^G)}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}^G,\mathrm{t}^B\frac{\min\limits_{\Omega}(\mathrm{J}^B)}{\mathrm{A}^B}+\mathbf{1}-\mathrm{t}^B\right)\end{aligned} \tag{3.14}$$

We want to compute the minimum of the three components of both sides of (3.14). Notice that in our case, the transmission maps can only be extracted from this second minimum operator after the following manipulations:

$$\begin{aligned}\mathrm{t}^G\frac{\min\limits_{\Omega}\mathrm{J}^G}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}^G&=\mathrm{t}_R\ \mathrm{t}_R^{\alpha_G-1}\frac{\min\limits_{\Omega}\mathrm{J}^G}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}_R+(\mathrm{t}_R-\mathrm{t}_R^{\alpha_G})\\ &=\mathrm{t}_R\ \mathrm{t}_R^{\alpha_G-1}\frac{\min\limits_{\Omega}\mathrm{J}^G}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}_R+\mathrm{t}_R(\mathbf{1}-\mathrm{t}_R^{\alpha_G-1})\\ &=\mathrm{t}_R\left(\mathrm{t}_R^{\alpha_G-1}\frac{\min\limits_{\Omega}\mathrm{J}^G}{\mathrm{A}^G}+(\mathbf{1}-\mathrm{t}_R^{\alpha_G-1})\right)+\mathbf{1}-\mathrm{t}_R\end{aligned}$$

and analogously,

$$\mathrm{t}^B\frac{\min\limits_{\Omega}\mathrm{J}^B}{\mathrm{A}^B}+\mathbf{1}-\mathrm{t}^B=\mathrm{t}_R\left(\mathrm{t}_R^{\alpha_B-1}\frac{\min\limits_{\Omega}\mathrm{J}^B}{\mathrm{A}^B}(\mathbf{1}-\mathrm{t}_R^{\alpha_B-1})\right)+\mathbf{1}-\mathrm{t}_R$$

where (3.11) has been used.

Inserting the previous result in the above equation, Eq. (3.14) becomes:

$$\left(\frac{\min\limits_{\Omega}(\mathbf{1}-\mathrm{I}^R)}{\mathbf{1}-\mathrm{A}^R},\frac{\min\limits_{\Omega}\mathrm{I}^G}{\mathrm{A}^G},\frac{\min\limits_{\Omega}\mathrm{I}^B}{\mathrm{A}^B}\right)= \tag{3.15}$$

$$\mathrm{t}^R\left(\frac{\min\limits_{\Omega}((\mathbf{1}-\mathrm{J}^R))}{\mathbf{1}-\mathrm{A}^R},\mathrm{t}_G^{\lambda_R-1}\frac{\min\limits_{\Omega}(\mathrm{J}^G)}{\mathrm{A}^G}+\mathbf{1}-\mathrm{t}_R^{\lambda_R-1},\right.$$

$$\left.\mathrm{t}^{\lambda_B-1}\frac{\min\limits_{\Omega}(\mathrm{J}^B)}{\mathrm{A}^B}+\mathbf{1}-\mathrm{t}^{\lambda_R-1}\right)+(\mathbf{1},\mathbf{1},\mathbf{1})-(\mathrm{t}_R,\mathrm{t}_R,\mathrm{t}_R).$$

Now we can take the minimum on both sides of Eq. (3.15):

$$\min\left(\frac{\min\limits_{\Omega}(\mathbf{1}-\mathrm{I}^R)}{\mathbf{1}-\mathrm{A}^R},\frac{\min\limits_{\Omega}\mathrm{I}^G}{\mathrm{A}^G},\frac{\min\limits_{\Omega}\mathrm{I}^B}{\mathrm{A}^B}\right)=$$

$$\mathrm{t}^R\min\left(\frac{\min\limits_{\Omega}((\mathbf{1}-\mathrm{J}^R))}{\mathbf{1}-\mathrm{A}^R},\mathrm{t}_G^{\lambda_R-1}\frac{\min\limits_{\Omega}(\mathrm{J}^G)}{\mathrm{A}^G}+\right.$$

$$\left.\mathbf{1}-\mathrm{t}_R^{\lambda_R-1},\mathrm{t}^{\lambda_B-1}\frac{\min\limits_{\Omega}(\mathrm{J}^B)}{\mathrm{A}^B}+\mathbf{1}-\mathrm{t}^{\lambda_R-1}\right)+\mathbf{1}-\mathrm{t}_R.$$

Here the vector multiplying t_R cancels out due to the hypothesis of the generalized Red Channel prior (3.12), and we obtain:

$$\mathrm{t}^R(x)=\mathbf{1}-\min\left(\frac{\min\limits_{\Omega}\mathbf{1}-\mathrm{I}^R}{\mathbf{1}-\mathrm{A}^R},\frac{\min\limits_{\Omega}\mathrm{I}^G}{\mathrm{A}^G},\frac{\min\limits_{\Omega}\mathrm{I}^B}{\mathrm{A}^B}\right).$$

And the other two transmission maps are immediately deduced from Eqs. (3.11). □

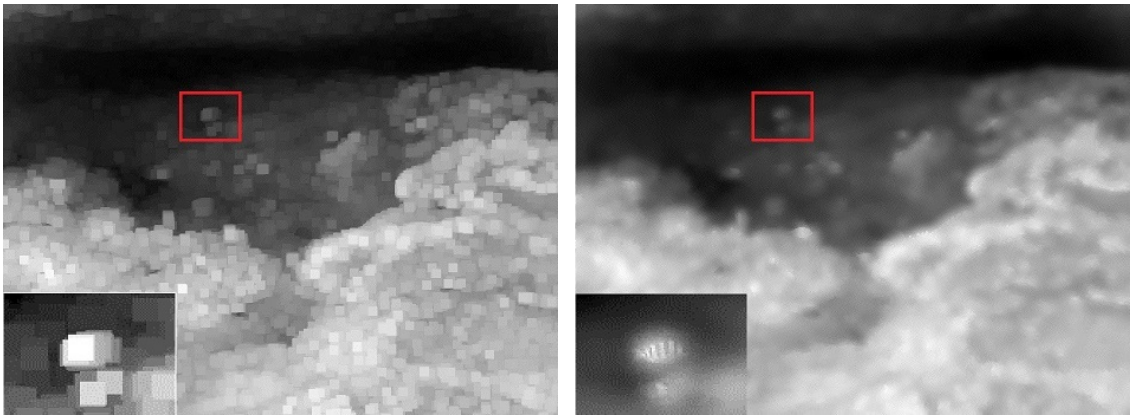
Remark 3.2.5 *Formula (3.12) contains two convex combinations,*

$$\mathrm{t}_R^{\lambda_G-1}\left(\min_{\Omega}\mathrm{J}^G/\mathrm{A}^G\right)+(\mathbf{1}-\mathrm{t}_R^{\lambda_G-1})\mathbf{1} \tag{3.16}$$

$$\mathrm{t}_R^{\lambda_B-1}\left(\min_{\Omega}\mathrm{J}^B/\mathrm{A}^B\right)+(\mathbf{1}-\mathrm{t}_R^{\lambda_B-1})\mathbf{1},$$

where these quantities vary from $\min\limits_{\Omega}\mathrm{J}^G/\mathrm{A}^G$ *and* $\min\limits_{\Omega}\mathrm{J}^B/\mathrm{A}^B$ *respectively, to* $\mathbf{1}$*. Let us analyse the extreme cases of this generalized prior.*

- *When these two quantities are equal to* $\mathbf{1}$*, then* $\mathrm{t}_R^{\lambda_G-1}=\mathrm{t}_R^{\lambda_G}/\mathrm{t}_R=0$*, and* $\mathrm{t}_R^{\lambda_B-1}=\mathrm{t}_R^{\lambda_B}/\mathrm{t}_R=0$*, implying that the transmission is null, so the depth is maximum. In this case, assumption (3.12) reduces to the Red Channel prior (3.4), so we are enforcing our main hypothesis in regions of high depth in the scene.*



(a) Rough transmission estimate (b) Refined transmission estimate.

Figure 3.2: Transmission map refinement.

- *When they are* $\min_{\Omega} \mathrm{J}^G/\mathrm{A}^G$ *and* $\min_{\Omega} \mathrm{J}^B/\mathrm{A}^B$ *respectively, then* $\mathrm{t}_R^{\lambda_G - 1} = \mathrm{t}_R^{\lambda_G}/\mathrm{t}_R = 1$ *and* $\mathrm{t}_R^{\lambda_B - 1} = \mathrm{t}_R^{\lambda_B}/\mathrm{t}_R = 1$, *so* $\mathrm{t}_R = \mathrm{t}_G = \mathrm{t}_B$. *This situation can only happen when we are near the observer, where the degradation is minimum, and more intense red can be found, so* $\left(\min_{y \in \Omega(x)} \mathbf{1} - \mathrm{J}^R(y)\right) / \left(\mathbf{1} - \mathrm{A}^R\right)$ *is near zero, making hypothesis (3.12) valid.*

The fact that the three derived transmission maps still lie in the interval $[0, 1]$ follows directly from Theorem 3.2.3.

Remark 3.2.6 (Transmission Map Refinement) *In all these calculations, we assume that the transmission map can be considered patch-wise constant, but as in the dehazing case, this is not realistic. If the patch contains a depth jump, this hypothesis fails and leads to artifacts in the recovered scene.*

A possible solution to overcome this problem is to refine the calculated rough transmission map using the guided filter [82], a version of the bilateral filter that efficiently captures the fine details of the degraded image, and incorporate them in the estimated transmission map, see Fig. 3.2. This method is faster and has no noticeable loss of detail with respect to the method based on the image matting technique [83], originally used in [7].

3.2.3 Final Inversion

To solve (3.1), we need to invert channelwise Eqs. (3.3). Just as in the Dark Channel method, we prevent divisions by small numbers by introducing a lower bound on the

denominator:

$$\mathrm{J}^R(x) = \frac{\mathrm{I}^R(x) - \mathrm{A}^R}{\max(\mathrm{t}^R(x), \mathrm{t}_0)} + \mathrm{A}^R$$
$$\mathrm{J}^G(x) = \frac{\mathrm{I}^G(x) - \mathrm{A}^G}{\max(\mathrm{t}^G(x), \mathrm{t}_0)} + \mathrm{A}^G$$
$$\mathrm{J}^B(x) = \frac{\mathrm{I}^B(x) - \mathrm{A}^B}{\max(\mathrm{t}^B(x), \mathrm{t}_0)} + \mathrm{A}^B$$

where a typical value for t_0 can be 0.1.

Incorporating Colour Correction II: Weighted WaterLight

Attenuation coefficients λ_G and λ_B in (3.13) are not easy to determine, as they depend on the type of water. They are usually specified heuristically or left as free parameters. In this method, a simplified version of the general technique can be obtained by directly building t as in (3.5) and employing the additive waterlight to weight the contributions of each wavelength. This yields a slightly different inversion formula:

$$\mathrm{J}^R(x) = \frac{\mathrm{I}^R(x) - \mathrm{A}^R}{\max(\mathrm{t}(x), \mathrm{t}_0)} + \rho^R \mathrm{A}^R,$$
$$\mathrm{J}^G(x) = \frac{\mathrm{I}^G(x) - \mathrm{A}^G}{\max(\mathrm{t}(x), \mathrm{t}_0)} + \rho^G \mathrm{A}^G,$$
$$\mathrm{J}^B(x) = \frac{\mathrm{I}^B(x) - \mathrm{A}^B}{\max(\mathrm{t}(x), \mathrm{t}_0)} + \rho^G \mathrm{A}^B,$$

where parameters ρ^R, ρ^G, ρ^B weight the contribution of each component. The multiplicative part involving $\overrightarrow{\mathrm{I}}$ takes care of the restoration of the deep parts of the scene, while the additive part removes the colour cast. Experience shows that there is no need to estimate these coefficients, since the reciprocal of the waterlight coefficients can automatically fulfil this task. Thus, a final version of the algorithm would include the following inversion formula:

$$\mathrm{J}^\alpha(x) = \frac{\mathrm{I}^\alpha(x) - \mathrm{A}^\alpha}{\max(\mathrm{t}(x), \mathrm{t}_0)} + (1 - \mathrm{A}^\alpha)\mathrm{A}^\alpha,$$

for $\alpha \in \{R, G, B\}$. This implementation no longer guarantees that the resultant image lies in $[0, 1]$. However, a simple min-max normalization of the intensity values to carry them to the unit interval easily fixes this issue.

3.2.4 Handling Artificial Illumination

Artificially illuminated underwater images are quite common. Unfortunately, high intensity of the pixels does not necessarily indicate the presence of artificial light. A

(a) Image with artificial light on it (b) Segmented artificial light

Figure 3.3: Saturation-based segmentation of artificially illuminated areas. Pixels with $S(x) < 0.4$ are painted on red.

much more adequate cue to characterize it is the *saturation*. The saturation component represents the purity of the chromaticity of a pixel: when a colour is in the pure spectrum, it is completely saturated, containing no white light. A colour loses saturation when we add white light, which contains power at all wavelengths [84]. The lack of saturation can thus be interpreted as the amount of white light appearing within its colour. Saturation is defined as:

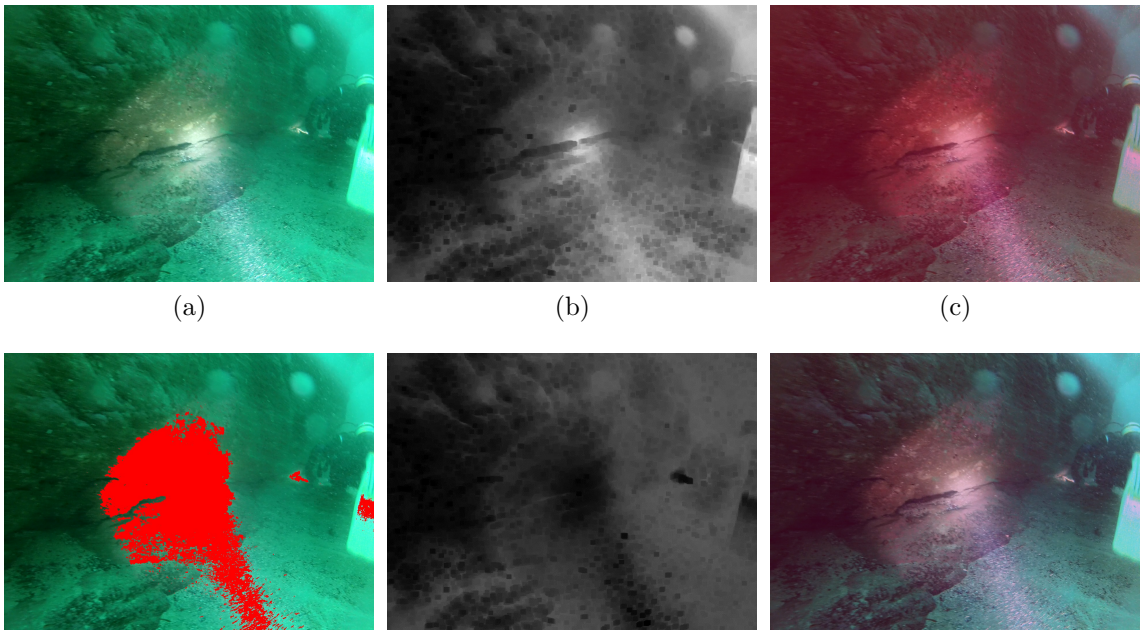
$$\mathrm{Sat}(\overrightarrow{\mathrm{I}}) = \frac{\max(\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B) - \min(\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B)}{\max(\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B)} \tag{3.17}$$

In Fig. 3.3, a saturation-based segmentation effectively separates areas with artificial light from the rest of the image. Additionally, when no light is present in the scene, the saturation is notably different from zero. This reflects the fact that artificial illumination forces a pixel to have similar intensity values rather than more green/blue than red. Then, a low saturation indicates that $\max(\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B) \simeq \min(\mathrm{I}^R, \mathrm{I}^G, \mathrm{I}^B)$, i.e., $\mathrm{I}^R \simeq \mathrm{I}^G \simeq \mathrm{I}^B$.

It is easy to incorporate saturation into the Red Channel prior. We just extend (3.4) to the following *Red-Saturation Prior*:

$$\mathrm{Jj}^{\mathrm{RED-SAT}}(x) = \min\left(\min_{y\in\Omega(x)}(\mathbf{1} - \mathrm{J}^R(y)), \min_{y\in\Omega(x)}(\mathrm{J}^G(y)), \min_{y\in\Omega(x)}(\mathrm{J}^B(y)), \min_{y\in\Omega(x)}\mathrm{Sat}(y)\right) \equiv 0 \tag{3.18}$$

This extension handles illumination gracefully. If a pixel exhibits disparity in the three channels, it is because the Red Channel has already lost intensity, so the object is far away from the observer and needs to be restored. However, if a pixel has a value for the three channels that is far from zero, either it lies in a location near the observer or in an artificially illuminated area. Without adding saturation to (3.4), an artificially illuminated area that is not close to the observer can have intermediate



(a) (b) (c)

(d) (e) (f)

Figure 3.4: Result of considering or not the Saturation cue. (a) Image with artificial light source on it. (b) Red Channel obtained with Eq. (3.5) (c) Image after processing without saturation prior. (d) Saturation-based segmentation of artificially illuminated areas. (e) Improved Red-Saturation Channel obtained with Eq. (3.18). (f) Image after processing with saturation prior.

values in the three channels, *e.g.* $(0.5, 0.5, 0.5)$. The Red Channel Prior considers this pixel to be far away, while (3.18) classifies it as artificially illuminated, and will not attempt to restore it, as desired.

To implement this idea within the Red Channel methodology, we simply reformulate Theorem 3.2.2 as:

Theorem 3.2.7 *Under artificial illumination, we can compute an improved estimate of* $\mathrm{t}(x)$ *as:*

$$\widetilde{\mathrm{t}}(x) = \mathbf{1} - \min\left(\frac{\min\limits_{y\in\Omega(x)}(\mathbf{1} - \overrightarrow{\mathrm{I}}^R(y))}{\mathbf{1} - \overrightarrow{\mathrm{A}}^R}, \frac{\min\limits_{y\in\Omega(x)}\overrightarrow{\mathrm{I}}^G(y)}{\overrightarrow{\mathrm{A}}^G}, \frac{\min\limits_{y\in\Omega(x)}\overrightarrow{\mathrm{I}}^B(y)}{\overrightarrow{\mathrm{A}}^B}, \lambda \min_{y\in\Omega(x)} \mathrm{Sat}(y)\right),$$

where $\lambda \in [0, 1]$ *is a scalar multiplier that can be manually adjusted to suit the amount of artificial light we want to take into account.*

Proof: It is a simple extension of Theorem 3.2.2.

The effect of including saturation can be clearly appreciated in Fig. 3.4. Therein, Fig. (3.4a) displays an artificially illuminated scene. Fig. (3.4b) shows the Red Channel calculated as in Eq. (3.5), where artificially illuminated pixels are classified as if they were in the background of the scene. In Fig. (3.4d), we see the segmentation

of illuminated areas, while Fig. (3.4e) shows the Red-Saturation Channel, computed using Eq. (3.18). In this figure, misclassified pixels are now darker, implying that they are closer to the observer, and do not need to be restored. Figs. (3.4c) and (3.4f) show the results of restoration with and without the Red-Saturation Prior respectively. We can appreciate a chromatic artifact that renders the restored image in a reddish tone at illuminated areas due to a wrong depth estimation. The integration of Saturation in the transmission estimation considerably reduces this effect, as seen in Fig. (3.4f).

3.3 Experimental Results

It is extremely challenging to assess the performance of an underwater image restoration algorithm, since there is not groundtruth available. Hence, no standard non-referenced image quality metric, such as PSNR or SSIM, can be used in our situation. In this dissertation, we evaluate the restoration output in a twofold manner. First, we will assess the amount of visibility that the presented algorithm, together with other five state-of-the-art methods, is able to recover. Second, we propose to evaluate colour correction based on the examination of three indicators of the quality of the restoration, that will be described in the next section. Each of these indicators reveals the weaknesses and strengths of the methods we compare in a different aspect of colour quality.

For comparison purposes, the four test images displayed in Fig. 3.5 have been selected, since they have been captured under different conditions, with a varied type of waters, representing different scene configurations and with a varying degree of artificial illumination.

3.3.1 Contrast and Visibility Recovery

In this dissertation, to evaluate the amount of contrast recovered we consider a non-referenced metric proposed in [85]. This metric was employed in [86] for purposes of visibility recovery assessment on images corrupted by fog. It is meaningful to use this metric, since it aims at finding whether the algorithm has retrieved some of the edges that were lost by the scattering effect. It works by computing three coefficients that measure the amount of new visible edges that a given contrast restoration method produces, as well as the restoration quality, since retrieving many edges alone is not an indicator of the quality of the algorithm.

Extensive comparisons of the performance of our algorithm with other five state-of-the-art techniques for underwater image restoration and enhancement have been performed, by computing the coefficients e and r, see [85]. The first of these coefficients, e, is calculated by first building a map of visible edges on the restored images, and then, counting the amount of edges on the original image, n_r, and on the restored image, n_o, and setting:

$$e = \frac{n_r - n_o}{n_o}.$$

This formula evaluates the quantity of edges that were not present at I_o but are in I_r.

(a) *Diver* image

(b) *Open scene* image

(c) *Shipwreck* image

(d) *Fishes* image

Figure 3.5: The set of images selected for evaluation.

Complementarily, for each pixel i belonging to a visible edge, it computes the ratio of the gradient in the restored image and in the original image r_i, and geometrically averages it to obtain the coefficient r as:

$$r = \frac{1}{n_r} \sum_i \log(r_i).$$

Finally, it also computes the number n_s of pixels that the algorithm saturates to black or white, normalized by the size of the image, obtaining the third coefficient σ:

$$\sigma = \frac{n_s}{dim_x \times dim_y}.$$

The visual results of restoring these images with the algorithms described in [76, 80, 81, 87, 88] and the Red Channel approach are shown in Figs. (3.6), (3.8), (3.10), (3.12), and the edge maps computed by the evaluation metric we are using are given in (3.7), (3.9), (3.11), (3.13). Tables 3.1 and 3.2 show the e and r coefficients that the different methods scored in these set of images. The σ coefficient is not displayed, since all methods exhibit values very close to zero.

We see that only on the Diver and Fishes images, there is one method, Ancuti's algorithm, that outperforms the Red Channel method's results. However, in Table

Table 3.1: Visibility recovery coefficient e compared on four images

e coefficient	Diver	Open Scene	Shipwreck	Fishes
Ancuti *et al.* [76]	23.7132	0.06377	1.4890	7.5038
Malkasse *et al.* [87]	2.4075	0.07222	0.5599	3.0220
Carlevaris *et al.* [80]	0.1617	-0.2039	1.3534	0.5345
Chiang *et al.* [81]	0.1046	0.1705	0.3123	0.1842
Serikawa *et al.* [88]	0.4414	0.14903	0.0381	0.9163
Red Channel	16.4695	0.4011	1.8982	4.5547

Table 3.2: Restoration quality coefficient r compared on four images

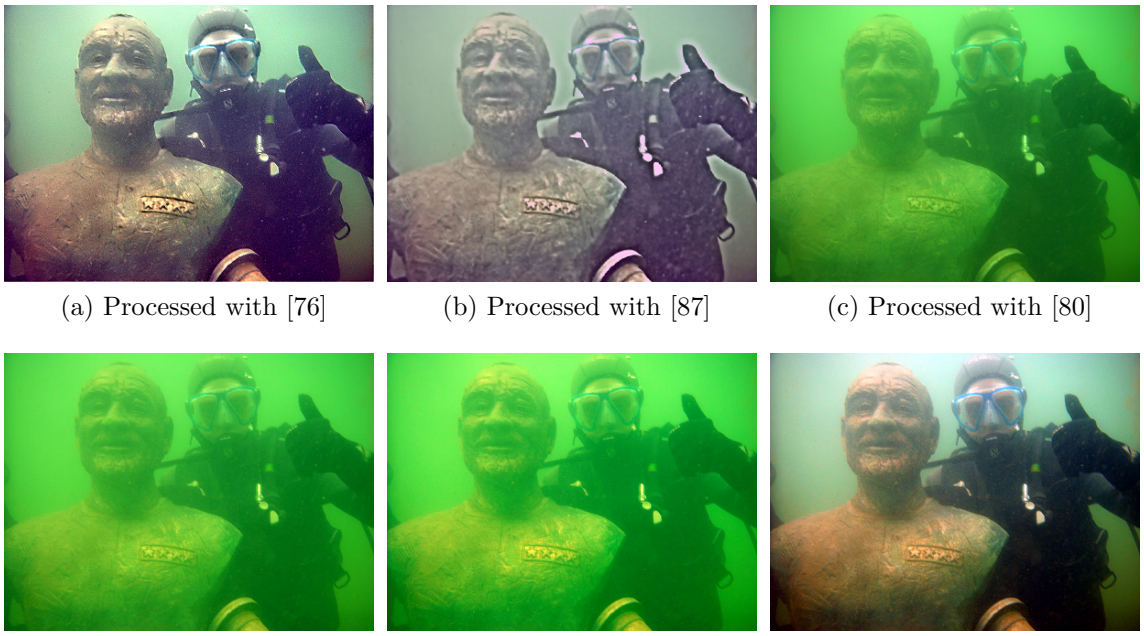
r coefficient	Diver	Open Scene	Shipwreck	Fishes
Ancuti *et al.* [76]	4.7713	1.7827	1.4890	5.3905
Malkasse *et al.* [87]	5.9437	3.6533	4.2686	6.3280
Carlevaris *et al.* [80]	1.2083	1.3625	2.6018	0.9813
Chiang *et al.* [81]	1.2976	1.7812	1.7677	1.4304
Serikawa *et al.* [88]	1.5719	2.1414	2.0557	1.8693
Red Channel	2.1197	1.1235	1.9810	2.3799

3.2 we can see that Ancuti's method is producing a higher r value. That means that this method is recovering *spurious edges*, as confirmed by visual inspection of Figs. (3.7a) and (3.7f), as well as of (3.9a) and (3.9f). We see there how Ancuti's algorithm is recovering much more noise than the Red Channel algorithm on areas of the image that contain no object, only water.

In every other case, the proposed algorithm performs better than the other state-of-the-art methodologies in terms of visibility recovery, analysed through the e coefficient. The values of the r coefficient in the Red Channel method tend to be higher than Carlevaris, Chiang's and Serikawa's, but this occurs because these three methods are not recovering as much visibility as the Red Channel approach does, as reflected by their scores at the e coefficient table. Since they retrieve substantially less edges, they also recover less *spurious edges*.

3.3.2 Colour Correction Evaluation

In terms of colour correction, as there is no groundtruth of how should the colour on the scene really be, we can only evaluate the ability of each method to remove the colour cast and recover a more natural colour distribution. Since *naturalness* of colours is a rather subjective concept, it is very hard to measure the quality of the



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

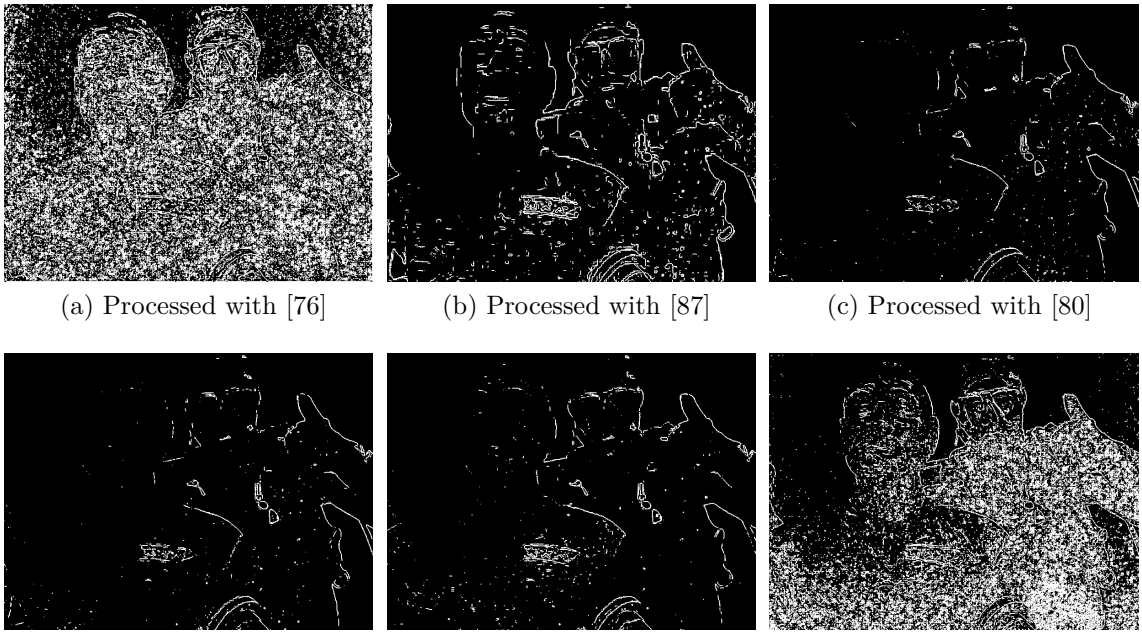
Figure 3.6: Results on the Diver image.

recovered colours.

Other works, see for instance [80, 81, 89], try to assess the ability of a method to recover real colours by taking an image of a colour board before and after immersion in a water body, then processing the water-degraded image and comparing the result to the colours of the image acquired in the atmosphere. This can be a meaningful evaluation approach for a single method performance analysis. Unfortunately, this approach has limited usability for comparison tasks, as it is very particular, and hardly reproducible, since each immersion takes place in different waters and under very different conditions that could favour one or another method. Moreover, it is arguable whether an underwater image processing algorithm should have the goal to convert colours to their equivalent in the surface. The goal here is to obtain an improved image, in terms of contrast and visibility increase, with realistic colours, which do not necessarily need to be the real ones as one would see them in the atmosphere. After all, colour correction must be performed in a way that the resultant image looks natural to the human eye, and this is inevitably a subjective task that does not necessarily match the idea of converting an underwater image to its surface counterpart.

We analyse the behaviour of the different methods regarding three visual components that are affected by colour corruption, namely, colour dominance, colour cast and colour fading which leads to low colour saturation. One metric is proposed for each of these aspects, allowing to quantitatively evaluate the performance of each method in a way that proves to be consistent with subjective perceptual criteria when applied to the test images we handle here.

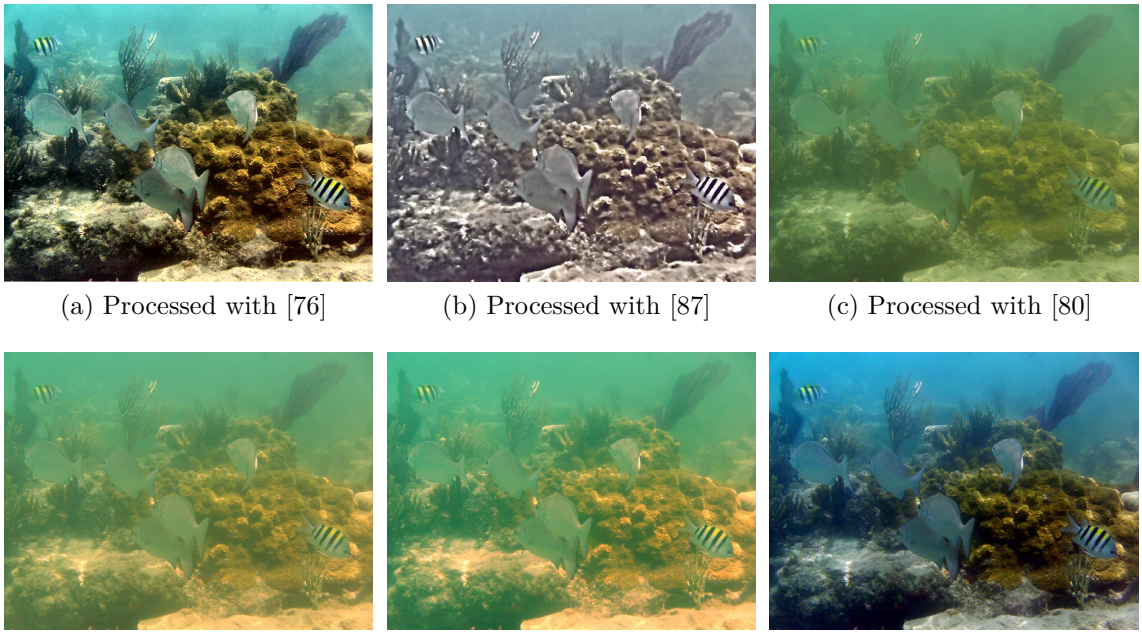
As advocated in [10], we can associate an image with colour cast to the fact that



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

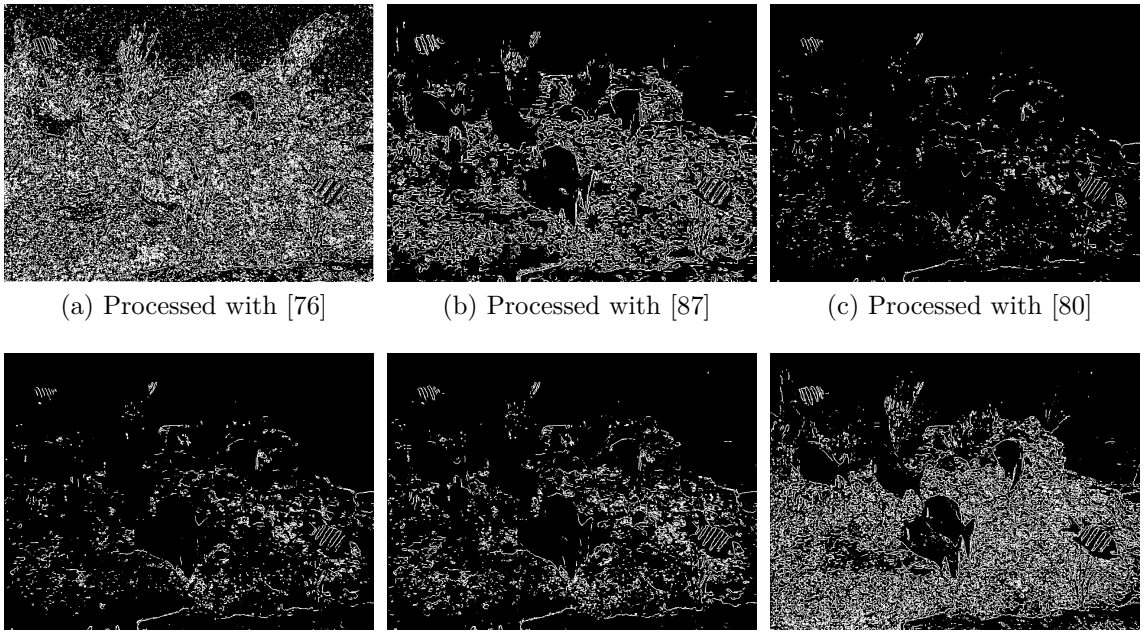
Figure 3.7: Visible edge maps of the recovered Diver image.



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

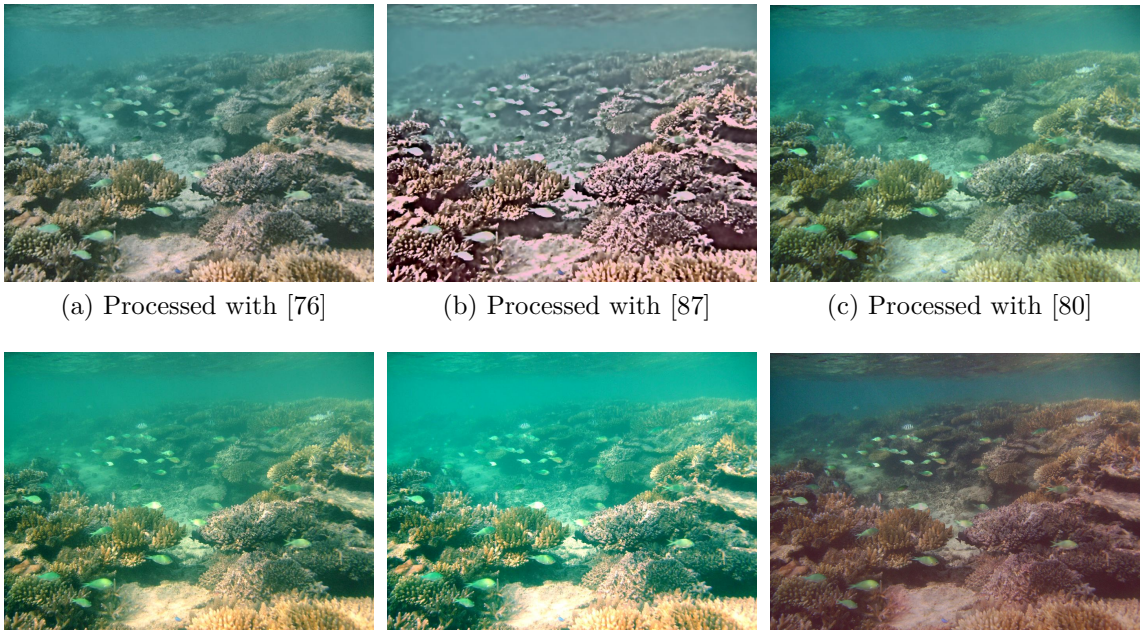
Figure 3.8: Results on the *Fishes* image.



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

Figure 3.9: Visible edge maps of the recovered *Fishes* image.

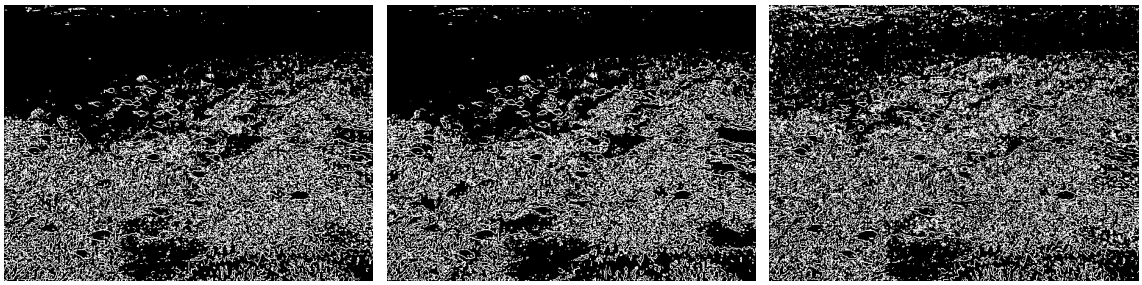


(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

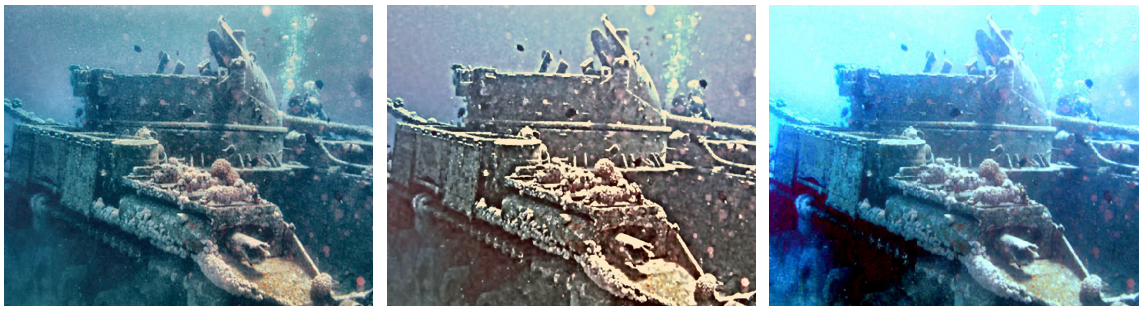
Figure 3.10: Results on the *Open scene* image.

(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]



(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

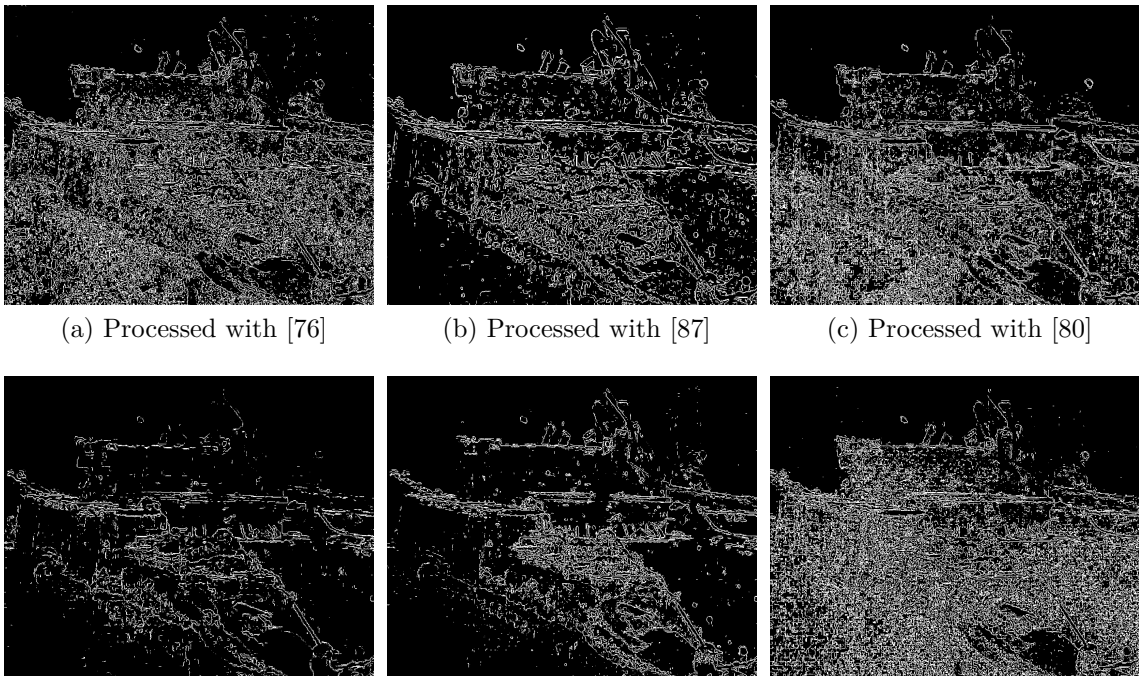
Figure 3.11: Visible edge maps of the *Open scene* image.



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

Figure 3.12: Results on the *Ship* image.



(a) Processed with [76] (b) Processed with [87] (c) Processed with [80]

(d) Processed with [81] (e) Processed with [88] (f) Processed with Red Channel.

Figure 3.13: Visible edge maps of the *Ship* image.

one chromatic channel exhibits a particularly different standard deviation with respect to others. Moreover, colour cast should not be confused with colour dominance. In the latter case, it is the average value of a channel that predominates. This interpretation further motivates the idea that contrast enhancement in colour corrupted images can help to decrease colour cast, as it spreads the intensity values of all colour channels, decreasing thus the dispersion of the correspondent standard deviations.

We can thus set as a meaningful goal the reduction of colour cast without increasing colour dominance on the three chromatic components of an underwater image. To measure the achievement of these goals, we can look at the relative dispersion of the three RGB channels, μ_R, μ_G, μ_B, and of its standard deviations, σ_R, σ_G, σ_B. To quantify this dispersion, we simply pick the maximum of their mutual distances, i.e.:

$$\mu_{\text{diff}} = \max(|\mu_R - \mu_G|, |\mu_R - \mu_B|, |\mu_B - \mu_G|) \tag{3.19}$$

and

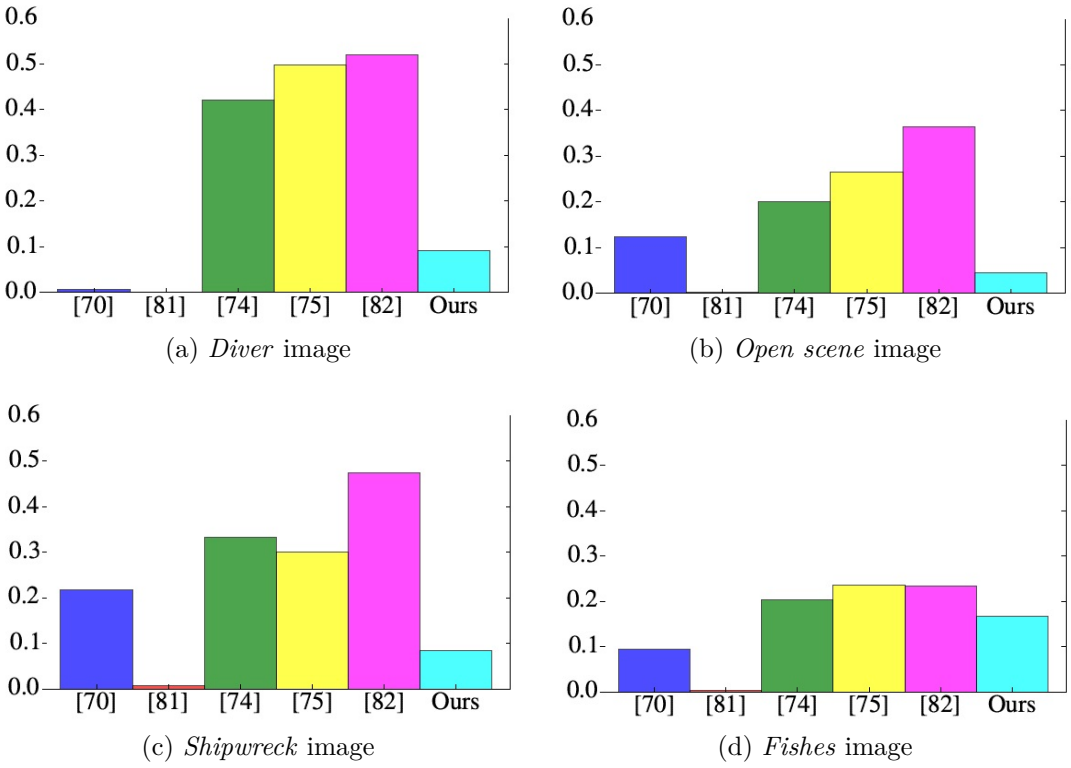
$$\sigma_{\text{diff}} = \max(|\sigma_R - \sigma_G|, |\sigma_R - \sigma_B|, |\sigma_B - \sigma_G|) \tag{3.20}$$

By enforcing that these coefficients remain relatively low, we are requiring that colour cast and dominance are reduced. Notice that a completely gray image would achieve the lowest score in both μ_{diff} and σ_{diff}. To incorporate the requirement of recovering a variety of colours, we can simply consider a measure of saturation. To keep the consistency of requiring low value to reflect optimal behaviour, we introduce a third coefficient, given by:

$$\lambda = 1 - \text{mean}(\text{Sat}(\text{I})), \tag{3.21}$$

where Sat(I) is defined as in Eq. (3.17). Each of these metrics measures the ability to reduce colour dominance, remove colour cast and retrieve rich colours, respectively.

Regarding colour dominance, Fig. (3.14) reproduces the scores that each method achieved in the metric given by Eq. (3.19), for each image in our test set.



(a) *Diver* image (b) *Open scene* image

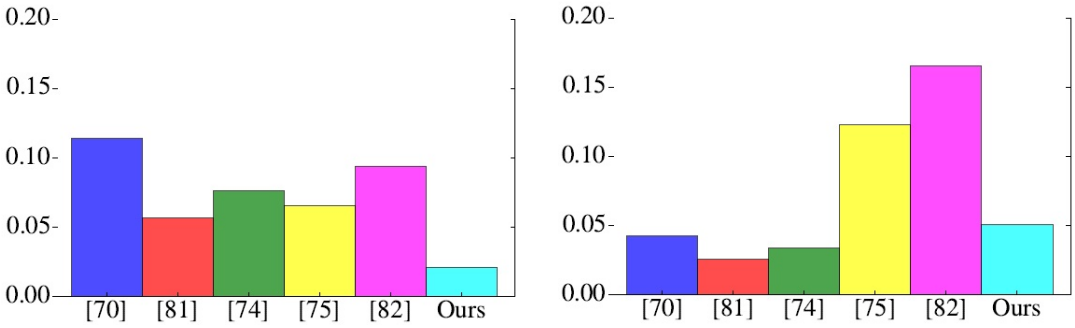
(c) *Shipwreck* image (d) *Fishes* image

Figure 3.14: Scores for coefficient μ_{diff} in each of the four test images.

We observe that the method in [87] is able to carry the chromatic means to a common value, obtaining the best score in colour dominance reduction. Ancuti's algorithm [76] and the Red Channel approach are ranked in second/third position here, while the rest of the methods fail to remove colour dominance.

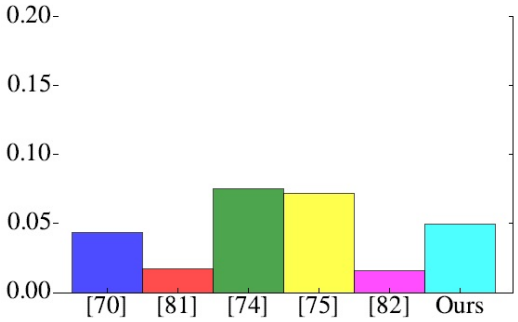
Regarding colour cast reduction, again the method in [87] scores in the best positions, only beaten by the Red Channel method on Image 1 and obtaining similar results as [88] on Image 3.

However, these results should be judged cautiously. The algorithm in [87] removes colour dominance and cast by sacrificing the recovery of intense colours, i.e., it is obtaining rather grayish images, see for example Fig. 3.8b. This is confirmed by inspection of Fig. 3.16. Therein, we realize that [87] is the method that recovers the least vivid colours. On the other hand, the methods in [80], [81] and [88] are obtaining very saturated colours, but clearly sacrificing the reduction of colour dominance, and to some extent, also colour cast. We thus see that, among the three methods that are able to reduce colour dominance and colour cast, Ancuti's method [76] and the Red Channel technique are again ranked as the first/second with respect to the ability to retrieve intense colours. This quantitative evaluation seems to soundly match

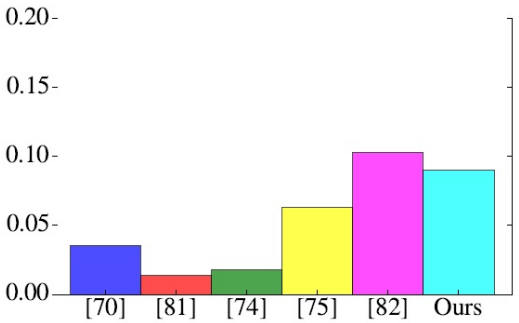


(a) *Diver* image

(b) *Open scene* image



(c) *Shipwreck* image



(d) *Fishes* image

Figure 3.15: Scores for coefficient σ_{diff} in each of the four test images.

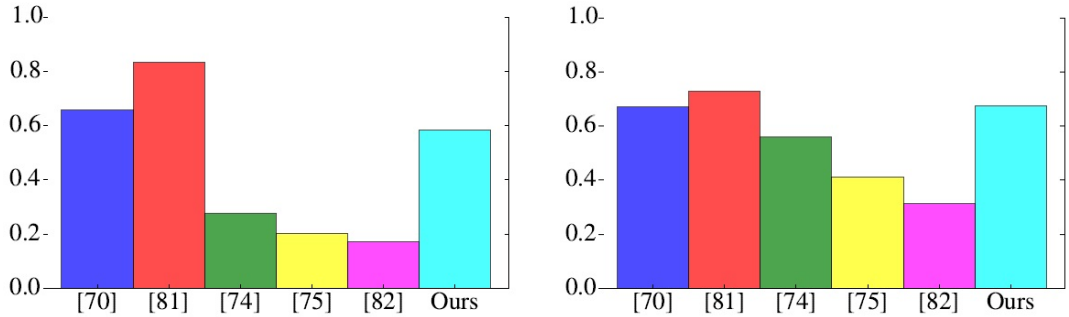
perceptual criteria, as visual inspection of the experimental results in the previous section reveals.

3.4 Conclusions on Chapter 3

In this chapter, a new method for the restoration of underwater images has been developed. It tackles both visibility loss and colour corruption. This technique extends the Dark Channel Method, adapting to the way these images are degraded. A general methodology to locate artificially illuminated areas within an underwater scene has also been studied, together with a method to handle these regions properly, avoiding colour artifacts that can appear as a result of depth miscalculations..

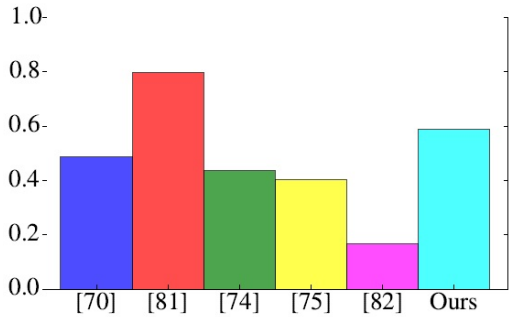
Regarding the experimental results, a twofold evaluation has been carried out. First, to measure the improvement of image contrast, a metric proposed in [85] has been employed to quantify visibility enhancement on images degraded by fog. Second, to evaluate the quality of the recovered colours, three basic indicators that adapt nicely to the problematic of assessing the quality of underwater colours after image restoration have been specifically developed.

Images restored with the Red Channel method have been compared with five different state-of-the-art algorithms. Results show that our approach obtains good-quality images, with a visibility enhancement comparable a better than other recent

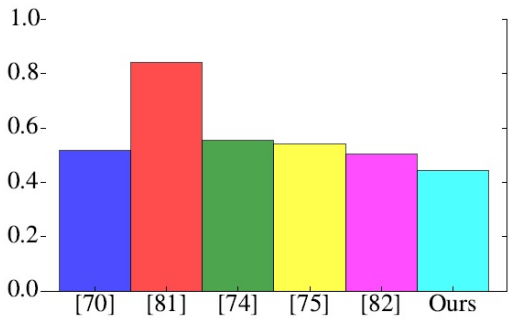


(a) *Diver* image

(b) *Open scene* image



(c) *Shipwreck* image



(d) *Fishes* image

Figure 3.16: Scores for coefficient λ in each of the four test images.

methods. As for colour recovery, the Red Channel technique retrieves natural colours, consistently ranking between the best positions among different images, regardless of the different water conditions.

The experiments have been performed on four different test images that attempt to capture the wide variability of underwater images. Although this test set encompasses rather different images, it is obvious that it is impossible to reflect the whole range of possibilities regarding the different water qualities, depth from the atmosphere, illuminations and a large number of parameters that affect underwater image acquisition. To improve potential comparison and evaluation, an online repository has been prepared[1] where the reader can find all the images that appear in this chapter at full resolution, as well as restoration results in a wider set of test images.

[1] `https://github.com/agaldran/UnderWater`

Chapter 4

A Variational Framework for Image Dehazing

Images obtained under adverse weather conditions, such as haze or fog, typically exhibit low contrast and faded colours, which may severely limit the visibility within the scene. This parallels in some sense the process suffered by underwater images. In the same way, unveiling the image structure under the haze layer and recovering vivid colours out of a single image remains a challenging task, since the degradation is again depth-dependent and conventional methods are unable to overcome this problem. In this chapter, we extend the perception-inspired variational framework reviewed in the chapter 3 of this dissertation to make it account for such a particular kind of degradation. Two main improvements are proposed. First, we replace the value used by the framework for the grey-world hypothesis by an estimation of the mean of the clean image. Second, we add a set of new terms to the energy functional for maximizing the inter-channel contrast. We present experimental results demonstrating that the developed Enhanced Variational Image Dehazing method (from now on EVID) outperforms other state-of-the-art methods both qualitatively and quantitatively. In particular, when the illuminant is uneven, the EVID method is the only one that recovers realistic colours, avoiding the appearance of strong chromatic artifacts.

4.1 Introduction

Under good visibility conditions, the human eye can resolve two different objects only within a range of a few kilometres. Indeed, when looking at mid and long range open natural scenes, colours of far away parts of the scene tend to mix with the colour of the atmosphere, gradually fading away and vanishing against the horizon. However, distance is not the only factor that can deteriorate perceptibility, but also bad weather conditions can accentuate visibility decrease. The presence of tiny suspended particles in the atmosphere lying between the observer and the objects in the scene can cause the deflection of light beams that travel from the latter to the former. This is produced by a wide range of factors, including dust, pollution, fog, haze, rain, or snow.

Recently, the problem of restoring and properly enhancing images affected by bad

atmospheric conditions has attracted much attention in the image processing and computer vision community. From a practical point of view, many vision systems operating in real-world outdoor scenarios assume that the input is the unaltered scene radiance. These techniques designed for clear weather images may suffer under bad weather conditions where, even for the human eye, discerning image content can represent a serious challenge. Therefore, robustly recovering visual information in bad weather conditions is essential for several machine vision tasks, such as autonomous robot/vehicle navigation [90], remote sensing [91] or video surveillance systems [92, 93]. Aerial and remotely sensed images, related to applications such as land cover classification [94, 95], and even underwater image restoration [81] can also benefit from efficient dehazing techniques.

From a physical point of view, models of the propagation of light through the atmosphere can be traced back to the work of Kochsmieder [67], who formulated a simple linear model relating the luminance reflected by an object to the luminance reaching the observer. This model identified the distance between the observer and the source, as well as the particle composition of the atmosphere, as the key parameters that govern the decay of the luminance. From then on, studies of the interaction of light with the atmosphere continued growing as a rich field of research in the area of applied optics [96, 97].

Based on these physical models, many image dehazing methods have been proposed. They can be roughly classified into multiple image ones [98–103], if they use more than one input image to estimate depth information, or single image ones [7, 86, 104], that work only with an initial degraded image. We can also find works that directly compute the albedo. In this case a depth map can be obtained as a by-product of these methods. Examples of this approach include the work of Tan [105], where the albedo is estimated by imposing a local maximization of contrast, or Fattal [106], where it is assumed that depth and surface shading are locally uncorrelated. Unfortunately, both of these approaches rely on the assumption that depth is locally constant, and as a consequence, the obtained images may sometimes suffer from artifacts and tend to over-enhance the results.

The work presented in the following lines is a combination of the advantages of the two approaches for image enhancement presented in section 2.2, namely, the modelling of the Human Visual System (HVS) and the modelling of a distance-dependent degradation. In this sense, two main improvements over the HVS developed in section 2.2.1.2 are proposed in the following lines. First, we modify the Gray World hypothesis to make it comply with the physical model of haze degradation presented in section 2.2.2.1 . Second, we extend the original functional, adding extra energy terms that allow us to control the degree of image saturation on the output. We also conduct an exhaustive quantitative evaluation versus a large set of recent methods. As the experiments demonstrate, the minimization of this new model provides high-quality haze-free images, even in challenging scenarios with uneven illumination or remarkably poor chromatic information. These seem to be situations where the majority of the state-of-the-art methods tend to encounter difficulties in restoring the degraded image without introducing unnatural colour artifacts.

The rest of the chapter is structured as follows. In the next section we review

(a) (b)

Figure 4.1: (a) Hazy image, where the degradation increases with distance. A non-uniform illuminant dominates the scene. (b) The result produced by the EVID method.

recent methods for image dehazing. Next, we formulate the image dehazing problem in a variational setting, and we develop an Enhanced Variational Image Dehazing (EVID) method. Section 4.4 is devoted to experimental results and comparison to other state-of-the-art methodologies. We end up in section 4.5 by summarizing the entire approach.

4.2 Related work

Most previous works on image dehazing are based on solving the image formation model presented by Koschmieder [67] that can be computed channel-wise as follows

$$\mathrm{I}(x) = \mathrm{t}(x)\mathrm{J}(x) + (1 - \mathrm{t}(x))\mathrm{A}, \tag{4.1}$$

where x is a pixel location, $\mathrm{I}(x)$ is the observed intensity, $\mathrm{J}(x)$ is the scene radiance, corresponding to the non-degraded image, transmission $\mathrm{t}(x)$ is a scalar quantity that is inversely related to the scene's depth, while A, known as airlight, plays the role of the haze colour, which is usually considered constant over the scene, and therefore in a channel-wise formulation it is a scalar value.

Solving Eq. (4.1) is an under-constrained problem, i.e. there exist a large number of valid solutions. To constrain this indeterminacy, extra information in different forms has been introduced in the past. A large part of these works can be grouped under the category of methods that work with multiple input images or other data, which must be fused to take profit of this extra-information. For example, in [102], several instances of the same scene are acquired under different weather conditions. Then, differences in pixel intensities allow to estimate depth discontinuities, computing thus a scene structure map that enables the recovery of a clear weather scene. Schaul *et al.* acquire visible and near-infra-red (NIR) images of the same scene [107], taking profit of the fact that NIR images are less sensitive to haze. Both images are then fused by means of a multi-resolution scheme to obtain a dehazed image. The

work in [108] retrieves depth information from geo-referenced digital urban and terrain models. An interactive registration of the photograph and its model is performed, after which depth information is available, making possible to remove haze effects in the input image. In [100], multiple images acquired through a polarizer at different orientations are used. This approach relies on the fact that the light scattered by atmospheric particles is partially polarized. Thus, it is possible to invert a physical model of image formation, enhancing scene contrast and correcting colour. Unfortunately, all these methods depend on the acquisition of extra information, which is often unavailable, and this hinders the practical use of these techniques.

Dehazing is particularly challenging when only a single input image is available. In this case, the majority of existing methods are also focused on solving Eq. (4.1) by inferring depth information. The method in [109] assumes that hazy images are usually acquired outdoors, allowing to conclude certain characteristics about the geometry of hazy scenarios. For example, the top part of the image is assumed to contain the most hazy regions in the image. Tarel *et al.* propose in [86] a fast algorithm that estimates the atmospheric veil (equivalent to the depth map) through an optimization procedure in which they impose piecewise smoothness. The main advantage of this method is its low complexity, that they show to be a linear function of the number of pixels. More recently, several fusion-based dehazing strategies have also been proposed. In [110], a multi-scale fusion strategy is adopted, in which the authors compute several weighting maps (based on luminance, chromaticity and saliency) that they use to fuse a white-balanced and a contrast enhanced version of the original degraded image. In [2], the depth map is obtained by a fusion procedure in a probabilistic framework.

When inverting models like the one in Eq. (4.1), there is a risk of enhancing both the underlying signal and the noise. Some works have also studied this issue. For example, in [111], a variational formulation is proposed, in which a distance-dependent and edge-preserving regularization term is introduced, resulting in a numerical scheme that is able to dehaze the input image while denoising far away regions. The variational point of view has also been recently exploited in [112], where Fang *et al.* proposed a variational formulation intended to dehaze a degraded image while denoising it. After estimating the depth information using a variation of the dark channel prior, the authors formulated an image energy, that they minimized using the primal-dual Chambolle-Pock algorithm [113]. Existence of a minimizer and convergence of the numerical schemes were also discussed.

A special mention deserves the Dark Channel Method [7], probably the most successful technique to date, due to its simplicity and effectiveness. This method is based on the statistical observation that haze-free images are colourful and contain textures and shadows, therefore lacking locally the presence of at least one of the three colour components. On the contrary, hazy images present less contrast and saturation. As depth increases and the haze takes over the image, the contrast and saturation further decrease, providing an estimate of the depth information based on which it becomes possible to invert Eq. (4.1), obtaining high-quality results. This method has been explained in detail on section 2.2.2.1.

Several methods that are independent of an initial estimation of the scene depth

have also been devised. Tan [105] observes that a haze-free image must have more contrast than a hazy one, and also that the amount of haze varies smoothly. Considering that nearby pixels tend to lie in the same depth layer of the scene, and thus are equally affected by haziness, he formulates a cost function attempting to maximize contrast while preserving depth smoothness locally. This cost function is optimized within a Markov random field framework, yielding a haze-free image. In [106], Fattal estimates scene depth by extending the image formation model to take into account both transmission and surface shading. He then looks for a haze-free solution of the model in which the resulting shading and transmission functions are locally uncorrelated in an statistical sense. In a subsequent work [104], the same author builds on the concept of colour-lines, establishing that for natural images, pixels exhibit a local one-dimensional distribution in the RGB space. A local image formation model is devised to adapt the colour-lines concept to hazy images, and transmission in the scene is estimated by means of a Markov random field that also allows to control noise. A substantially different approach is adopted by Nishino *et al.* in [114]. Rather than attempting to estimate depth in an initial stage, the authors resort to a Bayesian probabilistic framework, in which they are able to jointly estimate the albedo and the depth, by considering them as two statistically independent latent layers.

4.3 Enhanced Variational Image Dehazing (EVID) method

The majority of current dehazing algorithms are based on an estimation of the image depth (or transmission). Therefore, these methods are susceptible to fail when the physical assumptions underlying Eq. (4.1) are violated, as for example, when there is a source of light hidden by the haze, and in virtually-generated images that add different types of fog. Methods that do not estimate the model depth do not suffer from this problem, but they usually result in over-enhanced images due to the special characteristics of the degradation associated with haze. More conventional contrast enhancement algorithms, such as histogram equalization, are not suitable either. Fortunately, recent spatially-variant contrast enhancement techniques can be adapted to perform well for image dehazing tasks. In the following, we develop a variational framework for image dehazing that enforces contrast enhancement on hazy regions of the image throughout an iterative procedure that allows us to control the degree of restoration of the visibility in the scene.

4.3.1 Variational contrast enhancement

In 2007, Bertalmío *et al.* [3] presented a perceptually-inspired variational framework for contrast enhancement. This technique has been analysed in detail in section 2.2.1.2, and we briefly review it here, in its discrete version, for the convenience of the reader. The method is based on the minimization of the following functional for

each image channel I:

$$E(\mathrm{I}) = \frac{\alpha}{2}\sum_x \left(\mathrm{I}(x) - \frac{1}{2}\right) + \frac{\beta}{2}\sum_x (\mathrm{I}(x) - \mathrm{I}_0(x))^2 - \frac{\gamma}{2}\sum_{x,y}\omega(x,y)|\mathrm{I}(x) - \mathrm{I}(y)|, \quad (4.2)$$

where I is a colour channel (red, green or blue) with values in $[0,1]$, I_0 is the original image, x, y are pixel coordinates, α, β, γ are positive parameters, and $\omega(x,y)$ is a positive distance function with its value decreasing as the distance between x and y increases. This method extends the idea of variational contrast enhancement presented by Sapiro and Caselles [8] and it also shows a close connection to the ACE method [63]. Bertalmío and co-authors later revealed connections between this functional and the human visual system: they generalized it to better cope with perception results [10], and they established a very strong link with the Retinex theory of colour [9]. All these features of the above energy have been studied in depth in the first chapter of this dissertation.

The minimization of the image energy in Eq. (4.2) represents a competition between two positive terms and a negative one. The two positive terms prevent the solution from departing too much from the original image (second term) and preserve the gray-world hypothesis (first term). The negative competing term attempts to maximize the contrast. By focusing on this negative term of Eq. (4.2), we can observe a very useful relation with dehazing methods. It can be written as:

$$\sum_{x,y}\omega(x,y)|\mathrm{I}(x) - \mathrm{I}(y)| = \sum_{x,y}\omega(x,y)\left(\max(\mathrm{I}(x),\mathrm{I}(y)) - \min(\mathrm{I}(x),\mathrm{I}(y))\right). \quad (4.3)$$

We see from the above equation that the contrast term is maximized whenever the minimum decreases or the maximum increases, corresponding to a contrast stretching. Notice that minimization of local intensity values is one of the premises of a haze-free image, according to the Dark Channel prior, see Eqs. (2.2.2.1) and (2.37). The second premise of this prior is the presence, locally, of a low intensity of Red, Green or Blue. We will extend our method to handle also this hypothesis is subsection 4.3.3.

4.3.2 Modifying the Gray World assumption

In the image dehazing context, the Gray World hypothesis implemented in Eq. (4.2) is not adequate, since we want to respect the colours of the haze-free image, not to correct the illuminant of the scene. Different modifications of this hypothesis have already been proposed for several problems [115, 116]. In this chapter, to approximately predict which should be the mean value of a dehazed scene, we rely on the model of Eq. (4.1), that is written channel-wise as:

$$\mathrm{I}^j = \mathrm{J}^j\mathrm{t} + (1-\mathrm{t})A^j, \quad (4.4)$$

where $j \in \{R, G, B\}$. By rearranging and taking the average of each term, we can write:

$$\mathrm{mean}(\mathrm{J}^j \cdot \mathrm{t}) = \mathrm{mean}(\mathrm{I}^j) - \mathrm{mean}((1-\mathrm{t})A^j).$$

Now, we make the following two assumptions:

1. On a hazy image, colour of a pixel is depth-dependent. On a haze-free image, however, we can expect colours to be independent of where the object is located in the scene. Thus, we can assume that J^j and t are uncorrelated, allowing us to write:
$$\mathrm{mean}(\mathrm{J}^j \cdot \mathrm{t}) = \mathrm{mean}(\mathrm{J}^j) \cdot \mathrm{mean}(\mathrm{t})$$
2. We assume also that t has a uniform distribution across the image, i.e., depth values are equally distributed. This can be expressed simply as $\mathrm{mean}(\mathrm{t}) = 1/2$.

Using the above assumptions, we can estimate $\mathrm{mean}(\mathrm{J}^j)$ as:

$$\mathrm{mean}(\mathrm{J}^j)/2 \approx \mathrm{mean}(\mathrm{I}^j) - (1/2)\ \mathrm{mean}(\mathrm{A}^j).$$

The airlight A takes a constant value for each channel that can be roughly approximated by the maximum intensity value on each component, since haze regions have usually higher intensity. Thus, a reasonable approximation for the mean value of the haze-free scene, and in consequence a new gray-world value, is given by:

$$\mu^j = \mathrm{mean}(\mathrm{J}^j) \approx 2\ \mathrm{mean}(\mathrm{I}^j) - \mathrm{A}^j. \tag{4.5}$$

We can now rewrite the energy functional as:

$$E(\mathrm{I}^j) = \frac{\alpha}{2}\sum_x (\mathrm{I}^j(x) - \mu^j)^2 + \frac{\beta}{2}\sum_x (\mathrm{I}^j(x) - \mathrm{I}_0^j(x))^2 - \frac{\gamma}{2}\sum_{x,y}\omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^j(y)|. \tag{4.6}$$

To minimize the above energy, we first need to compute its Euler-Lagrange derivative. We have already formulated the condition a minimizer of (4.6) must satisfy. This was:

$$\nabla E(\mathrm{I}^j) = \alpha(\mathrm{I}^j(x) - \mu_j) + \beta(\mathrm{I}^j(x) - \mathrm{I}_0^j(x)) - \gamma R(\mathrm{I}^j)(x) = 0, \quad j \in \{R, G, B\}$$

where the function $R(\mathrm{I})$ is a contrast enhancement operator:

$$R(\mathrm{I})(x) = \frac{\sum_y \omega(x,y)s(\mathrm{I}(x) - \mathrm{I}(y))}{\sum_y \omega(x,y)}, \tag{4.7}$$

and s is a smooth approximation of the sign function, that accounts for the first derivative of the absolute value. In the next subsection, we formulate an extended version of the energy given in (4.6). We give in the Appendix a proof of the computation of its variational derivative.

We can now apply a gradient descent strategy. To this end, we solve $\dfrac{\delta \mathrm{I}}{\delta t} = -\nabla E(\mathrm{I})$, being t the evolution parameter. For the case of the energy given by Eq. (4.6), with the modified gray world assumption, after an explicit discretization in time, we have:

$$\mathrm{I}_{k+1}^j = \mathrm{I}_k^j(1 - \Delta t(\alpha + \beta)) + \Delta t(\alpha\mu^j + \beta\mathrm{I}_0) + \Delta t\gamma R(\mathrm{I}_k^j), \quad j \in \{R, G, B\} \tag{4.8}$$

The initial condition for this descent is $\mathrm{I}_{k=0}^j = \mathrm{I}^j(x)$. The computation of operator R is reformulated in terms of convolutions by means of Fast Fourier Transforms. This brings a significant computational improvement to the method, since the effort to compute expression (4.7) falls down from $\mathcal{O}(N^2)$ to $\mathcal{O}(N\log(N))$. Details of this argument for complexity reduction can be found in [3, 117].

4.3.3 Controlling the saturation

The variational framework has several advantages over traditional closed-form solutions. First, it provides an iterative procedure that allows the user to stop the iterations as soon as the image meets a desired degree of quality. Also, the effect of the modification of the parameters governing the evolution of Eq. (4.8) is simple to understand and has a clear meaning. Furthermore, the energy-based formulation eases the combination of different image processing objectives. For example, we can combine Eq. (4.6) with a depth map coming from any of the algorithms that are able to estimate a 3D structure in the scene, to enforce denoising or deblurring tasks in far away areas of the scene. As an additional feature of the proposed method, in this section we illustrate how we can add an extra term to Eq. (4.6) in order to recover more vivid colours when an output image seems too washed-out, or if the user considers there is too much saturation in the recovered scene.

Hazy images are not only characterized by a loss of contrast, but they also exhibit a low saturation that makes colours look faint. The degree of saturation of a pixel x in a RGB image is defined by the following formula:

$$\begin{aligned} \mathrm{S}(x) &= \frac{\max(\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x)) - \min(\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x))}{\max(\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x))} \\ &= 1 - \frac{\min(\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x))}{\max(\mathrm{I}^R(x), \mathrm{I}^G(x), \mathrm{I}^B(x))} \end{aligned} \tag{4.9}$$

This implies that saturation of a pixel is augmented as long as the maximum increases or the minimum decreases in the above formula. This idea complies in part with the Dark Channel prior, formulated in Eqs. 2.2.2.1 and 2.37, which states that in a local neighbourhood around a haze-free pixel, we should find a low value of intensity in some of the RGB channels. We have already included the spatial component of this assertion in Eq. (4.3). The maximization of the expression there enforces locally the presence of low intensity pixels. To incorporate the chromatic component, we simply modify the contrast increase term given by Eq. (4.3) to account for the value of the pixel in the different channels:

$$\sum_{x,y} \omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^l(y)| = \sum_{x,y} \omega(x,y) \left(\max(\mathrm{I}^j(x), \mathrm{I}^l(y)) - \min(\mathrm{I}^j(x), \mathrm{I}^l(y))\right),$$

where $j \neq l$, $j, l \in \{R, G, B\}$. This amounts to extending the energy in Eq. (4.6) by introducing an inter-channel stretching term:

$$\begin{aligned} E(\mathrm{I}^j) =& \frac{\alpha}{2} \sum_x (\mathrm{I}^j(x) - \mu^j)^2 + \frac{\beta}{2} \sum_x (\mathrm{I}^j(x) - \mathrm{I}_0^j(x))^2 - \frac{\gamma}{2} \sum_{x,y} \omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^j(y)| \\ & - \left(\frac{\eta}{2} \sum_{x,y} \omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^{j+1}(y)| + \frac{\eta}{2} \sum_{x,y} \omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^{j+2}(y)| \right), \end{aligned} \tag{4.10}$$

where $j \in \mathbb{Z}_3$. In here, $\{R, G, B\}$ is identified with the space of integers modulo 3, $\mathbb{Z}_3$, meaning for instance that $\mathrm{I}^2 = \mathrm{I}^B$, and $\mathrm{I}^3 = \mathrm{I}^R$. The derivative of this extended

energy is given by the following expression:

$$\begin{aligned}\nabla E(\mathrm{I}^j) =&\alpha(\mathrm{I}^j(x) - \mu_j) + \beta(\mathrm{I}^j(x) - \mathrm{I}_0^j(x)) \\ &- \gamma R(\mathrm{I}^j, \mathrm{I}^j)(x) - \eta[R(\mathrm{I}^j, \mathrm{I}^{j+1}) + R(\mathrm{I}^j, \mathrm{I}^{j+2})], \quad j \in \mathbb{Z}_3,\end{aligned} \tag{4.11}$$

where we have introduced an extension of the operator (4.7) by considering the operator $R(\mathrm{I}^1, \mathrm{I}^2)$ defined as follows:

$$R(\mathrm{I}^1, \mathrm{I}^2)(x) = \frac{\sum_y \omega(x,y) s(\mathrm{I}^1(x) - \mathrm{I}^2(y))}{\sum_y \omega(x,y)}. \tag{4.12}$$

Notice that when $\mathrm{I}^1 = \mathrm{I}^2$, this reduces to the mono-channel contrast enhancement operator defined in (4.7). Thus, the energy in Eq. (4.10) includes the one in Eq. (4.6) as a particular case, and so does its derivative. We include a proof of equation (4.11) in Appendix A.

A complete numerical gradient descent that increases local contrast and saturation can be written as:

$$\begin{aligned}\mathrm{I}_{k+1}^j =&\mathrm{I}_k^j(1 - \Delta t(\alpha + \beta)) + \Delta t(\alpha\mu^j + \beta \mathrm{I}_0) \\ &+ \Delta t(\gamma R(\mathrm{I}_k^j, \mathrm{I}_k^j) + \eta[R(\mathrm{I}_k^j, \mathrm{I}_k^{j+1}) + R(\mathrm{I}_k^j, \mathrm{I}_k^{j+2})]),\end{aligned} \tag{4.13}$$

where $j \in \mathbb{Z}_3$. Parameter η controls the degree of saturation increase. If necessary, a change of sign in the last term of (4.13) would accomplish a local saturation decrease.

The iteration of Eq. (4.13) until steady-state configures the proposed Enhanced Variational Image Dehazing (EVID) method. EVID provides the user with a mechanism to control the amount of saturation on the recovered colours. As an example, in Fig. (4.2a) we display a hazy open scene. The enhanced versions of the input image appearing in Fig. (4.2a) have been obtained by using Eqs. (4.8) and the EVID method, respectively. The convenience of applying EVID over the method of Eq. (4.8) is demonstrated in the close-up details appearing in Figs. (4.2d) to (4.2i). We can appreciate how the addition of the inter-channel saturation term produces more intense colours both in far away regions of the scene and in middle-range areas.

4.4 Experimental Results

In this section, the dehazing capability of our EVID method is tested qualitatively and quantitatively by comparing it with several existing techniques. To qualitatively evaluate the method, we have executed it over various hazy images that are quite popular in the literature of image dehazing for evaluation purposes, as well as in other realistic scenarios, in which the illumination affecting the scene is non-uniform and the physical model in Eq. (4.1) becomes invalid. Likewise, for the quantitative evaluation of the experimental results, we make use of real depth data provided by [4] to produce both homogeneous and heterogeneous haze layers over real-world scenarios. We then compare the ability of our method to remove that haze layer

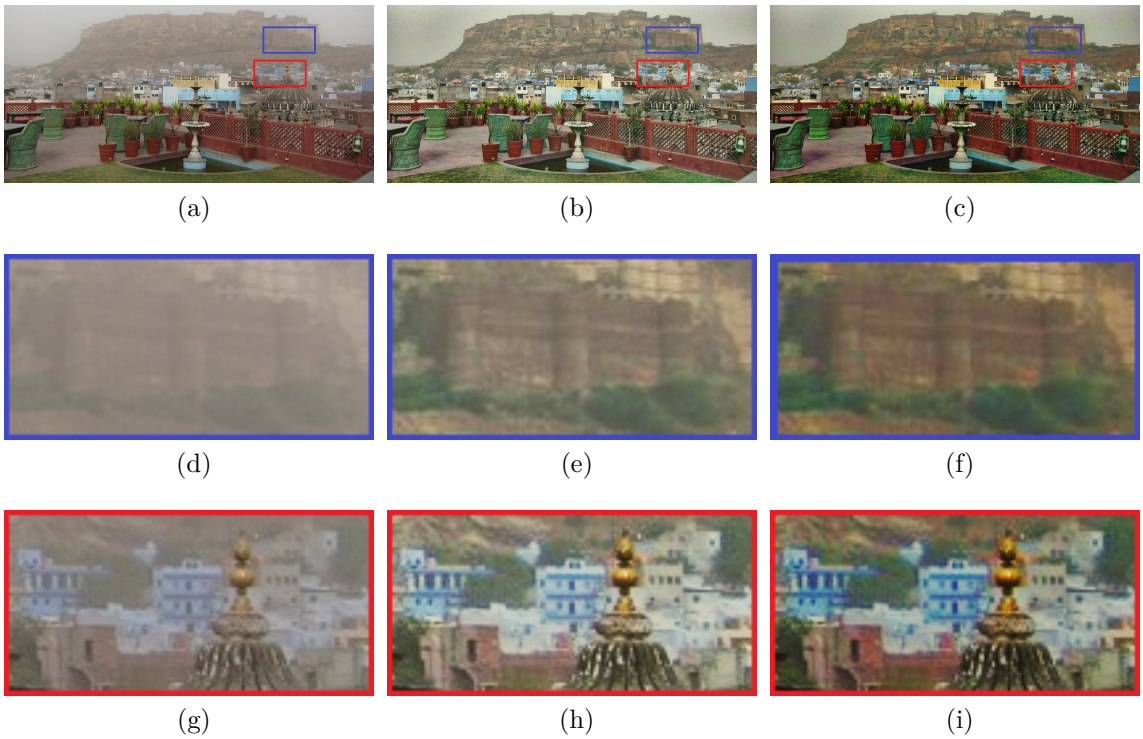


Figure 4.2: (a) Hazy open scene. (b) Result of executing the method in Eq. (4.8). (c) Result of executing the method in Eq. (4.13). (d)-(f) Detail of the castle on top of the image (g)-(i) Detail of the buildings on the middle of the image.

by means of several standard full-reference metrics, comparing also against various state-of-the-art approaches.

Regarding the parameter configuration, we have considered Eq. (4.13) with the following baseline values: $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 0.2$, and $\eta = 0.02$. Both distance functions were defined as Gaussians with kernel of standard deviation equal to 50 pixels. The time step was set to $\Delta t = 0.15$, and we considered that a steady-state of the gradient descent was achieved when the difference between the images of two consecutive iterations was below 0.020. A sensitivity study with respect to the parameters of our methods is provided in Section 4.4.1.2.

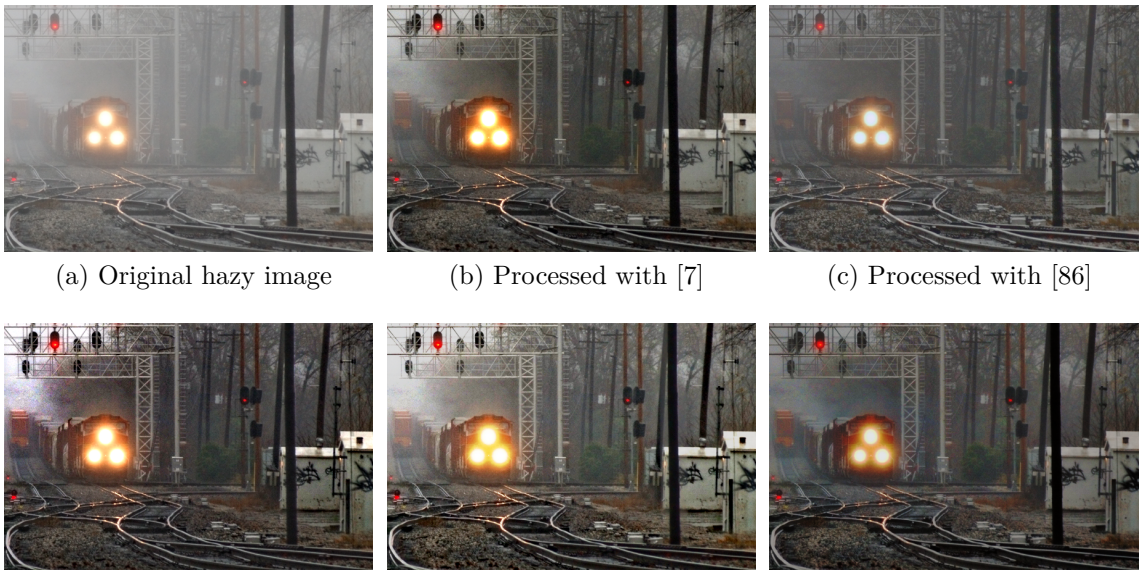
4.4.1 Qualitative Evaluation

To evaluate the performance of a dehazing technique, the most popular criteria are often the amount of retrieved visibility for far away details, together with the plausibility of the colours that appear in the recovered scene. In this section, we consider also these subjective measurements to evaluate the performance of our method when compared with other existing approaches.

Fig. 4.3 displays a hazy scene typically used as a benchmark for testing a dehazing method. For this example, all methods effectively recover visibility of hidden details at the bottom of the scene, being the technique of Meng *et* al. [118] the one that

retrieves slightly more structure than the rest. Unfortunately, this seems to happen at the price of introducing some colour artifacts at the leftmost part of the image. The rest of the methods are free from these artifacts.

On the other hand, the image in Fig. (4.3a) is a particularly interesting test case, since the headlights of the train are useful to evaluate the way each method handles other light sources different from the ambient light. We can appreciate how only our EVID method and the one of Tarel *et* al. [86] are capable of correctly preserving the shape of the three lights. For this example our method recovers more vivid colours than [86], as can be seen by observation of the red wagon at the left of Figs. (4.3c) and (4.3f).



(a) Original hazy image (b) Processed with [7] (c) Processed with [86]

(d) Processed with [118] (e) Processed with [53] (f) Processed with EVID

Figure 4.3: Typical benchmark image for dehazing algorithms and result of processing it with several state-of-the-art methods, including EVID.

Figure 4.4 displays an image of New York City, also often considered for benchmarking of dehazing algorithms. Therein, our EVID method is compared against the works in [7, 86, 104–106, 108]. Visibility of distant objects is again recovered by every method up to a reasonable degree. The recent method of Fattal [104] and the Dark Channel method [7] are possibly the ones recovering most realistic colours, although a close analysis reveals that they fail to recover the information underlying in the horizon (see upper right corner in Figs. (4.3e) and (4.3f)). Additionally, the method by Tan [105] suffers of noticeable over-saturation artifacts.

The previous two examples demonstrate how for typical images, existing algorithms (including ours) can handle haze effectively, recovering visibility up to some extent. On the other hand, each of the methodologies restores chromatic information in a different way, although the majority of the available techniques produce rather plausible colours. Differences in the performance are subtle and only little details reveal whether a method is performing better than another in particular regions of the image.

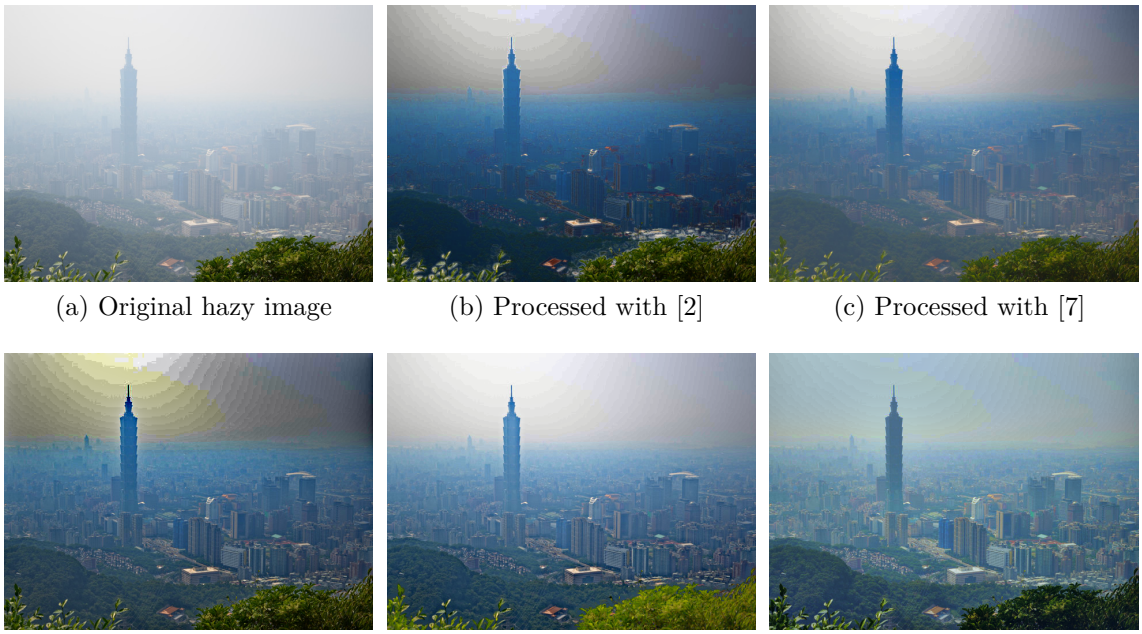
(a) Original hazy image (b) Processed with [106] (c) Processed with [105] (d) Processed with [108]

(e) Processed with [104] (f) Processed with [7] (g) Processed with [86] (h) Processed with EVID

Figure 4.4: Hazy image of New York City image and result of processing it with several state-of-the-art methods, including EVID.

Unfortunately, little research has addressed the problem of image dehazing in a more challenging and also realistic scenario, such as the one depicted in Fig. 4.5 or Fig. 4.6. Most of the state-of-the-art methodologies rely on the previous computation of a depth map of the scene. Thus, they usually resort to a physical model of the image formation under haze and bad weather conditions, such as Eq. (4.1). This model assumes constant illumination in the scene. When this assumption is violated, the airlight cannot be considered to be constant. The result is a transmission underestimate or overestimate in unevenly illuminated areas, and colour distortions characterized by dark blue regions appear in the restored images.

The method proposed in this chapter benefits from the advantage of depending only mildly on physical considerations. The EVID method does not need to compute any depth information prior to restoration. Thanks to this feature, unevenly illuminated regions are handled properly, and scene structure can be recovered without introducing excessive colour distortion. Fig. 4.5 clearly demonstrates this point. The original image in Fig. (4.5a) was extracted from [2]. The sun appears behind the haze at the upper left part of the image, producing a non-uniform illuminant: the leftmost part of the sky appears brighter than the rightmost side. The methods that rely on an accurate estimate of the airlight and the depth are fooled by this non-uniformity,



(a) Original hazy image (b) Processed with [2] (c) Processed with [7]

(d) Processed with [118] (e) Processed with [53] (f) Processed with EVID

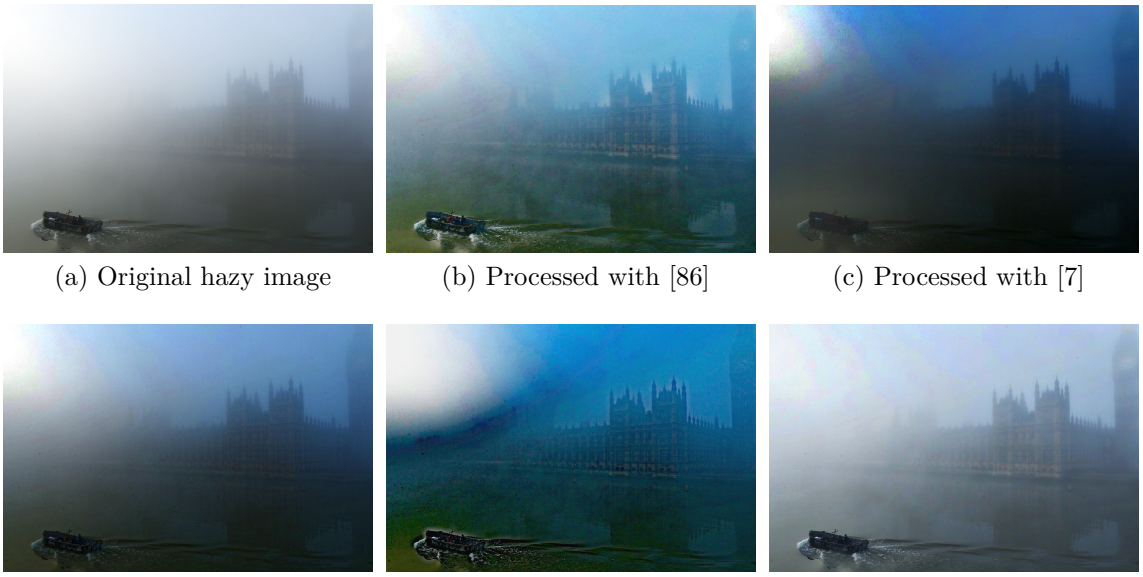
Figure 4.5: Unevenly illuminated hazy image extracted from [2], and result of processing it with several state-of-the-art methods, including EVID.

and generate strong colour artifacts. For this example, only our method is capable of retrieving some of the buildings structure without producing a chromatic degradation.

The same challenging scenario is faced in Fig. 4.6. Again, the compared methods cannot avoid the creation of colour distortions in the effort to increase contrast in the upper rightmost part of the original hazy image. In this case, the EVID method was executed with a single parameter modification: γ was set to 1, forcing the solution to stay closer to the original image. This represents another advantageous feature of the variational approach, since the attachment to data term provides the user with an intuitive way of controlling whether the contrast increase is excessive or insufficient.

The example in Fig. 4.7 shows another useful characteristic of our approach. Most works rely on the presence of enough colour information in the scene so as to recover the depth structure. When this chromatic information is weak or missing, the result is often also an image with unpleasant colour artifacts. The EVID method operates in a channel-wise manner, handling thus more robustly the lack of colour cues in the input image, as can be observed in the results obtained by the other physically-based techniques when compared to our variational approach, see Fig. (4.7f). As in the previous image, a single parameter variation was considered for this image: γ was set to 1.5, forcing the solution to stay closer to the original image.

Fig. 4.8 presents another image suffering from poor chromatic information. As in the previous case, our method seems to be the only one capable of restoring content in far away areas of the scene without introducing a remarkable colour degradation.



(a) Original hazy image (b) Processed with [86] (c) Processed with [7]

(d) Processed with [53] (e) Processed with [114] (f) Processed with EVID

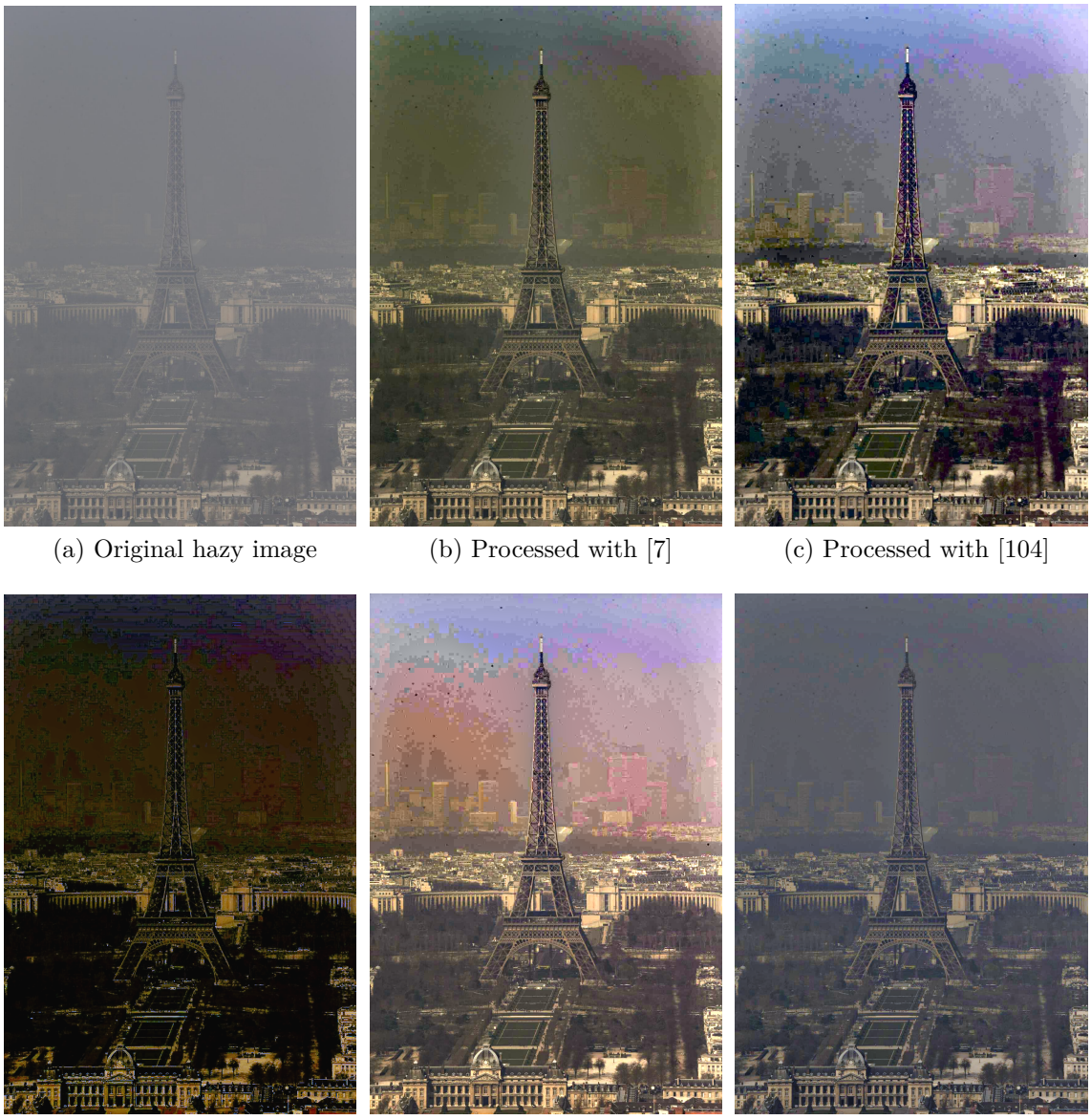
Figure 4.6: Unevenly illuminated hazy image of the Thames river, and result of processing it with several state-of-the-art methods, including EVID.

4.4.1.1 Comparisons with contrast enhancement approaches

The results described in the previous section clearly illustrate the advantages of the EVID method over other recent dehazing techniques. Most of these advantages come from the fact that EVID does not compute a transmission map of the hazy scene. On the other hand, we do incorporate in the model a mild physical constraint, since our modified Gray-World hypothesis relies on an estimate of the mean of the haze-free scene, and this is obtained from the physical model given by Eq. (4.1). This also explains the improvement of EVID with respect to other contrast enhancement methods that do not consider any physical cue. To illustrate this point, we show in Fig. (4.9) the result of processing three hazy scenes with EVID, as well as with a conventional contrast enhancement method (Histogram Equalization, HE) and the advanced variational contrast enhancement model on which the EVID method builds, introduced in [3] and described by Eq. (5.3). Notice that for a fair comparison, this model is run here with a parameter configuration analogous to the one of EVID.

In the three examples shown, we can clearly appreciate how EVID produces more vivid colours while increasing contrast. In Fig. (4.9a) we see the same hazy image of a train as in Fig. (4.3a). HE saturates large parts of the image, in an attempt to obtain a global uniform histogram. On the other hand, the method in [3] does not saturate bright areas, but it seems unable to retrieve plausible colours in far-away regions of the scene, such as the red wagon or the green bush next to the train.

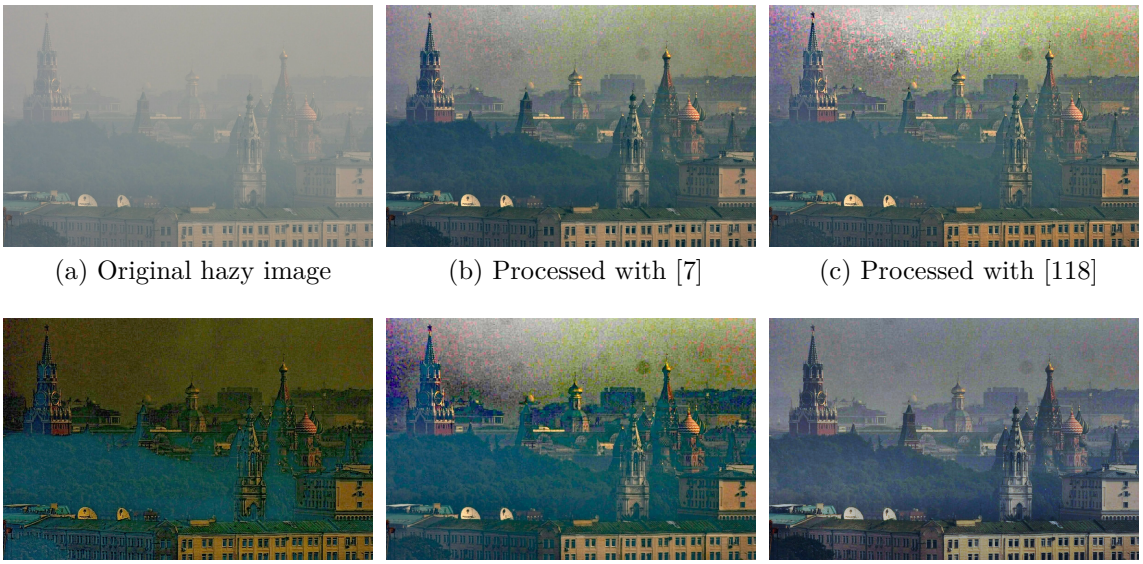
In Fig. (4.9e), we have a natural landscape with a forest on it. Again, HE burns dark regions while saturating to white on brighter areas. However, in this case, the method in [3] is capable of retrieving a similar contrast when compared to EVID. Unfortunately, due to the Gray World hypothesis implemented by [3], the model



(a) Original hazy image (b) Processed with [7] (c) Processed with [104]

(d) Processed with [114] (e) Processed with [118] (f) Processed with EVID

Figure 4.7: An image of Paris, with a lack of chromatic information due to haze in the scene, and result of processing it with several state-of-the-art methods, including EVID.



(a) Original hazy image (b) Processed with [7] (c) Processed with [118]

(d) Processed with [114] (e) Processed with [104] (f) Processed with EVID

Figure 4.8: (a) An image of Moscow, with a lack of chromatic information due to haze in the scene, and result of processing it with several state-of-the-art methods, including EVID.

lacks the ability to produce intense colours.

Finally, in Fig. (4.9i) we include a hazy view of the city of Taipei. We can appreciate similar drawbacks for HE, which forces dark areas to become black and handles incorrectly the sky areas, saturating to white the top part of the image. Also, the method in [3] recovers some contrast, but again it produces grayish colours due to the Gray World assumption it incorporates, further supporting the idea that a smart contrast enhancement is not enough to approach the problem of haze removal. In this case too, the EVID method is able to recover contrast while producing richer colours. All these observations can be better appreciated in Fig. 4.10, where we have included close-up details of the image areas marked by blue, red, and yellow rectangles on Figs. (4.9a), (4.9e), and (4.9i).

4.4.1.2 Parameter Sensitivity

In this section, we study the influence of the different parameters presented in the model (degree of contrast, attachment to initial input, and saturation enhancement).

Let us first analyse the influence of the parameter γ, which controls the degree of contrast enhancement provided by the method. Fig. (4.11a) shows a hazy scene of the city of Bilbao. The restored version displayed in Fig. (4.11c) has been obtained with the configuration of parameters recommended above, namely $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 0.2$, and $\eta = 0.02$. The images appearing in Figs. (4.11b) and (4.11d) have been produced by varying parameter γ. Specifically, we have set $\gamma = 0.1$ and $\gamma = 0.3$. We can appreciate how this modification produces a smooth variation of contrast in the scene, with an insufficient increase in contrast for $\gamma = 0.1$, and a slightly over-enhanced output for $\gamma = 0.3$. This can better appreciated in the close-up details of

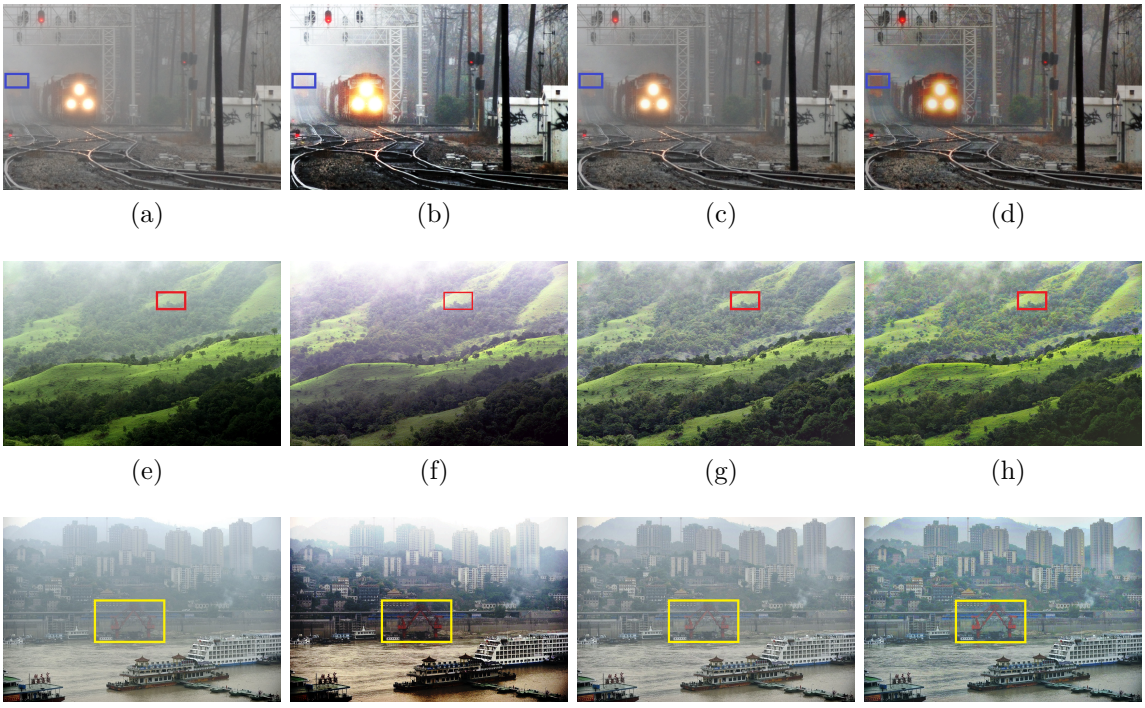


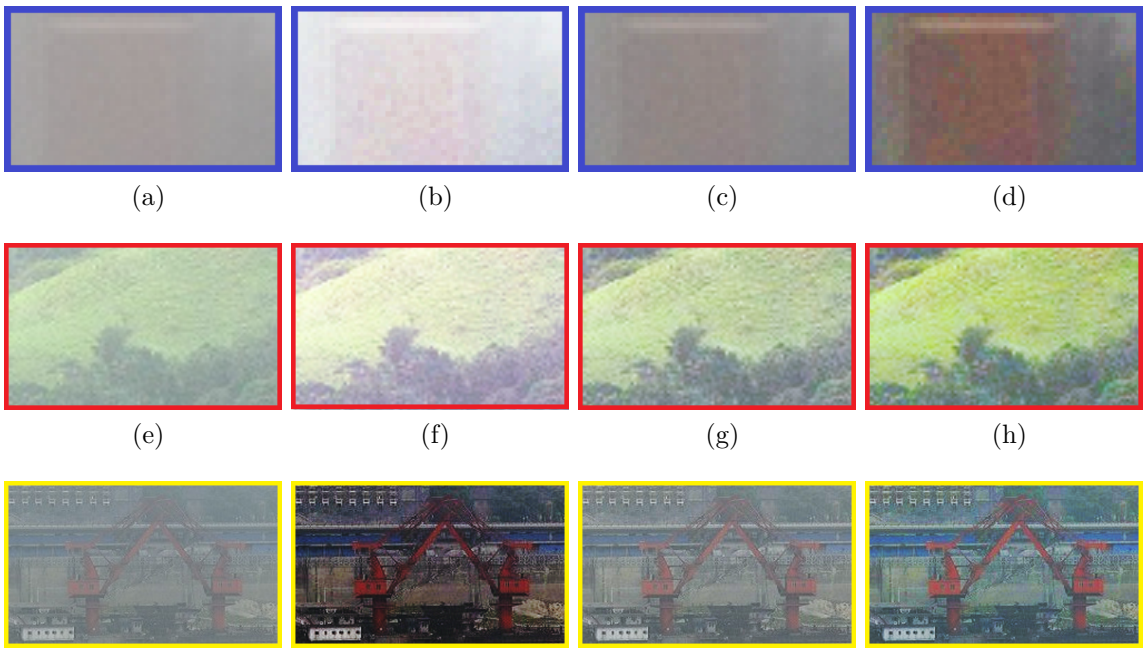
Figure 4.9: Result of applying a conventional and an advanced contrast enhancement method compared with the result obtained by EVID. (a), (e), (i) Different hazy scenes. (b), (f), (j) Result of applying Histogram Equalization. (c), (g), (k) Result of applying the method in [3]. (d), (h), (l) Our proposed EVID method.

the museum's faÃğade displayed on Figs. (4.11e) to (4.11h).

To assess the effect of a varying attachment-to-data parameter, we have executed again EVID on the hazy image (4.11a), varying β in the set $\{0.25, 0.5, 0.75\}$ and keeping fixed the other parameters on its baseline configuration, *i*.e. $\gamma = 0.2$, $\beta = 0.5$ and $\eta = 0.02$. The effect produced by increasing parameter β is opposite to the one produced by increasing the contrast parameter γ, as expected. We can observe the result of this on Fig. 4.12. There, we can appreciate how the variation of the attachment-to-data parameter influences the degree of separation of the output image from the original one, allowing the user to control the degree of enhancement provided by the method.

The gray-world parameter captures the degree of confidence we have on the estimate of the mean of the haze-free image in accordance to Kochsmieder model. The effect of varying this parameter can be appreciated on Fig. (4.13), where we have executed the EVID method with the parameter reference configuration, and with $\alpha \in \{0.25, 0.5, 0.75\}$. A slight colour shift can be noted in the different outputs. This is a global effect that can be better appreciated with a close look into the high resolution versions of the processed images.

Finally, the effect of modifying parameter η affects to the amount of saturation that colours in the scene experiment. This is illustrated on Fig. (4.14), where he have

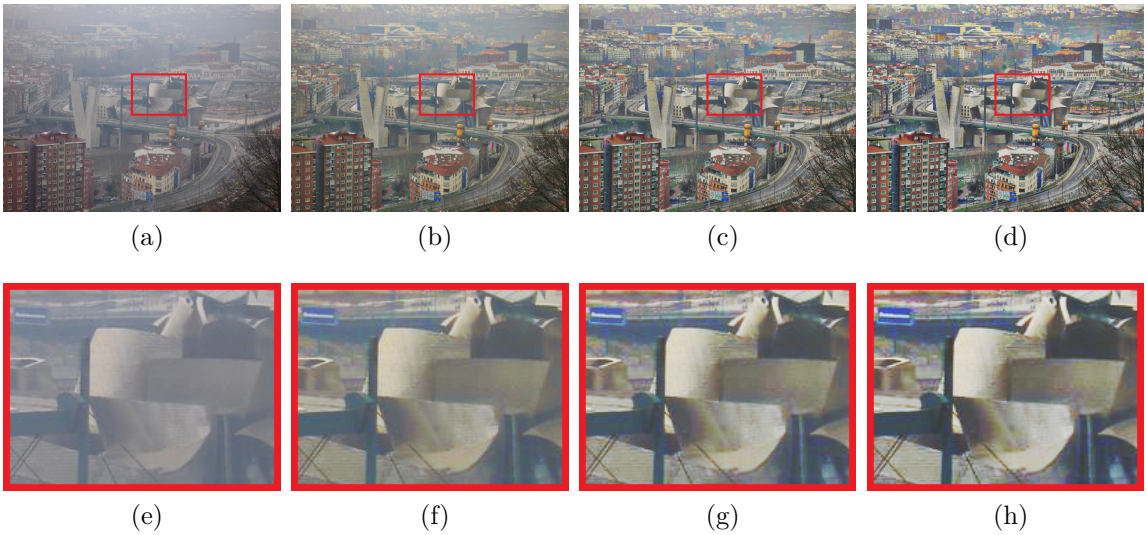


(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k) (l)

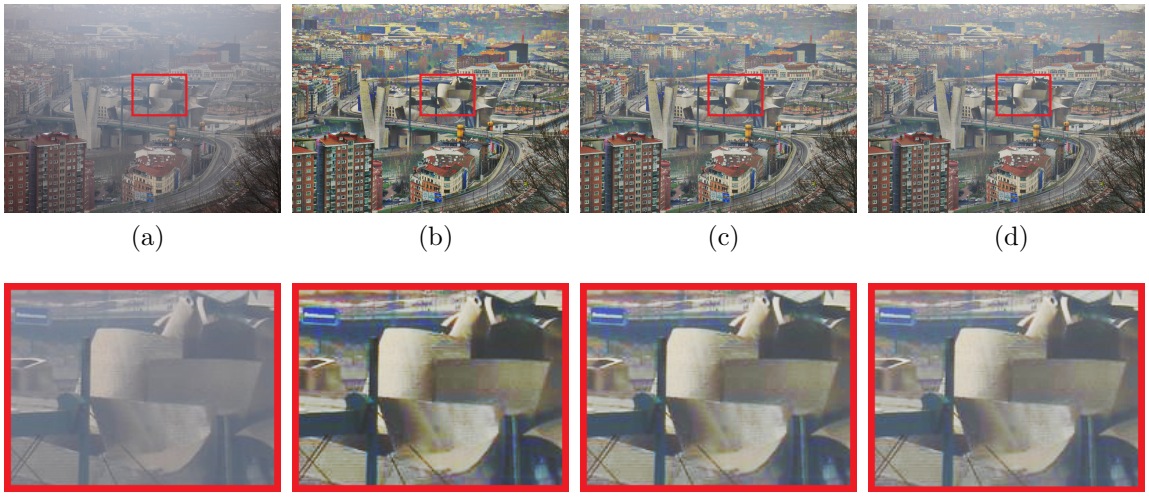
Figure 4.10: Zoomed details from Fig. (4.9). (a)-(d) Detail of the red wagon in the left of the image for the original image, and the output produced by HE, [3], and EVID. (e) - (f) Detail of the trees in the top part of the scene for the original image, and the output produced by HE, [3], and EVID. (i)-(l)Detail of the crane in the middle of the scene for the original image, and the output produced by HE, [3], and EVID.



(a) (b) (c) (d)

(e) (f) (g) (h)

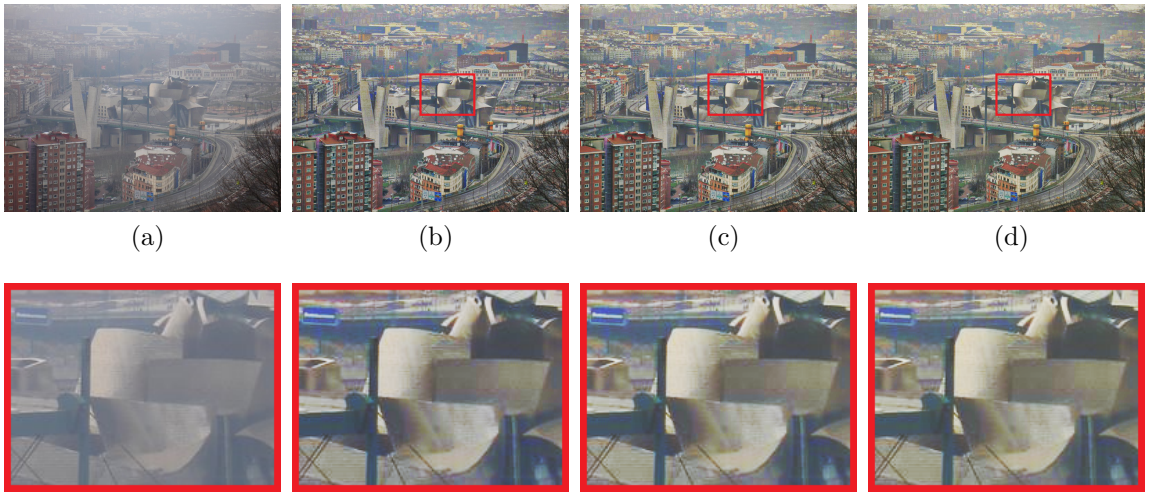
Figure 4.11: A hazy view of the city of Bilbao, and the result of varying contrast parameter γ. (a) Original hazy scene. (b) $\gamma = 0.1$ (c) $\gamma = 0.2$ (d) $\gamma = 0.3$. (e)-(h) Detail of the Guggenheim museum faÃğade.



(a) (b) (c) (d)

(e) (f) (g) (h)

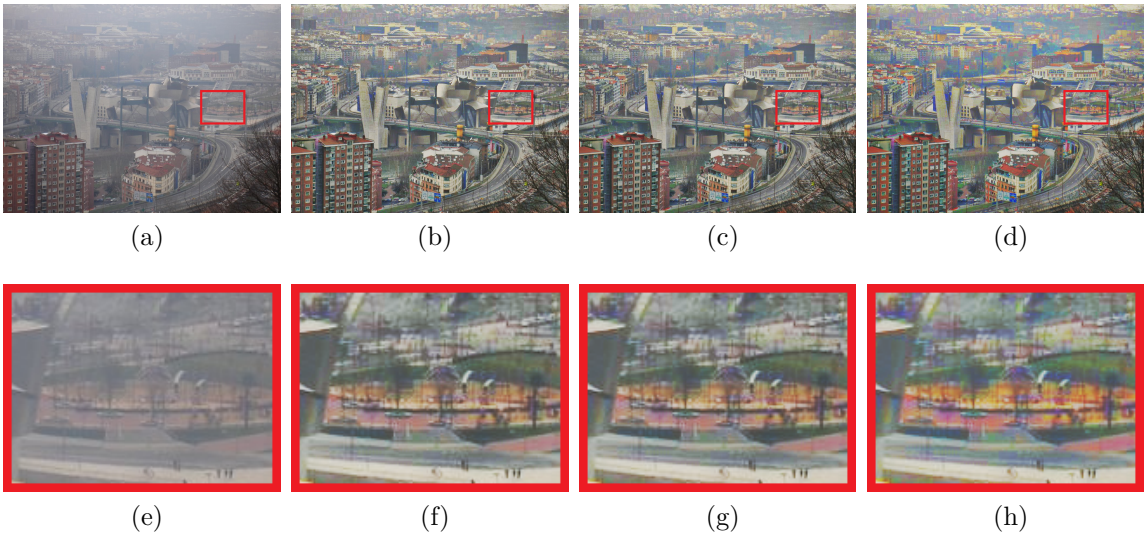
Figure 4.12: A hazy view of the city of Bilbao, and the result of varying attachment-to-data parameter β. (a) Original hazy scene. (b) $\beta = 0.25$ (c) $\beta = 0.5$ (d) $\beta = 0.75$. (e)-(h) Detail of the Guggenheim museum faÃğade.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 4.13: A hazy view of the city Bilbao, and the result of varying gray-world parameter β. (a) Original hazy scene. (b) $\alpha = 0.25$ (c) $\alpha = 0.5$ (d) $\alpha = 0.75$. (e)-(h) Detail of the Guggenheim museum faÃğade.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 4.14: A hazy view of the city of Bilbao, and the result of varying saturation parameter η. (a) Original hazy scene. (b) $\eta = 0$ (c) $\eta = 0.02$ (d) $\eta = 0.04$. (e)-(h) Detail of the park next to the museum.

executed the EVID method with the baseline parameter configuration, and setting $\eta \in \{0.01, 0.02, 0.03\}$. Colour saturation gradually grows as η increases. We can also appreciate in the far away buildings and trees the effect produced by a large value of η, leading to the appearance of colour artifacts on large uniform areas.

4.4.2 Quantitative Evaluation

Evaluating the performance of a dehazing method is always a hard task due to the inherent lack of ground-truth. To overcome this difficulty, some previous works have resorted to the generation of a synthetic fog layer [5, 119, 120]. However, to be able to generate a realistic haze, we first need depth data on the scene. To this end, we employ in this work the LIVE Color+3D Database provided in [4]. This database includes twelve different scenes together with real depth data. Data was acquired using an advanced range scanner with a digital camera mounted on top of it.

To incorporate realistic fog layers on top of these images, we employed the approach of Tarel *et al.* [5]. The authors provide software to generate up to four different kinds of fog. This is achieved by slightly modifying Eq. (4.1) to introduce some variability, adding Perlin's noise to both transmission t and airlight A. The only parameter required by the method was the visibility distance, which was set to 60 meters for all the experiments. An example of an original image in the LIVE Color+3D Database, together with its ground-truth depth data, and the different hazy versions of the input image can be seen in Fig. (4.15).

We have therefore a total of 48 images (12 scenes under 4 different fogs). We have compared EVID versus current state-of-the-art methods, namely the ones of He *et al.* [7], Meng *et al.* [118], Nishino *et al.* [114], Tarel and Hautière [86], and Gibson and Nguyen [53]. Our method was computed with the same fixed parameters for all

scenes, as pointed out at the beginning of section 4.4.

We have used five different measures, three mainly based on the Mean Square Error (MSE), and two based on the correlation. The first considered MSE based measure computes the MSE between the images for each particular channel, and then combines the three channels using by the l_2 norm:

$$MSE_{split} = \sqrt{MSE(I_r^{GT}, I_r^m)^2 + MSE(I_g^{GT}, I_g^m)^2 + MSE(I_b^{GT}, I_b^m)^2}, \tag{4.14}$$

where I^{GT} stands for the ground-truth image, I^m stands for the method used, the subindexes represent each colour channel and

$$MSE(I^1, I^2) = \frac{\sum_{i=1}^{M}(I^1(i) - I^2(i))^2}{M}, \tag{4.15}$$

being M the number of pixels.

The second error measure is the MSE between the luminance of the two images. Defining $I_{lum}^{GT} = 0.2126I_r^{GT} + 0.7152I_g^{GT} + 0.0722I_b^{GT}$ and $I_{lum}^m = 0.2126I_r^m + 0.7152I_g^m + 0.0722I_b^m$, the error measure becomes:

$$MSE_{lum} = MSE(I_{lum}^{GT}, I_{lum}^m). \tag{4.16}$$

We have also defined a new error measure in order to consider also the relation between the three colour channels. We denote this error measure as $l_2 - colour$ and we define it as:

$$l_2 - colour = \frac{\sum_{i=1}^{M}\sqrt{(I_r^{GT}(i) - I_r^m(i))^2 + (I_g^{GT}(i) - I_g^m(i))^2 + (I_b^{GT}(i) - I_b^m(i))^2}}{M}. \tag{4.17}$$

Measures based on the correlation are similar to the first two defined measures, except for the fact that the MSE measure has been replaced by the correlation between the channels. Mathematically,

$$Corr_{split} = \sqrt{Correlation(I_r^{GT}, I_r^m)^2 + Correlation(I_g^{GT}, I_g^m)^2 + Correlation(I_b^{GT}, I_b^m)^2}, \tag{4.18}$$

and

$$Corr_{lum} = Correlation(I_{lum}^{GT}, I_{lum}^m), \tag{4.19}$$

where $Correlation$ gives us the correlation coefficient between the two images.

Results are shown in Table 4.1. Our EVID method outperforms all the others in the five different measures when the full set of images is considered (table rows labelled 'All'). When dividing the images depending on the fog model, EVID also outperforms the remaining ones for the homogeneous, heterogeneous attenuation and heterogeneous airlight fogs in most of the measures, being very close to the best method (Tarel and Hautiere method [86]) for the remaining fog types.

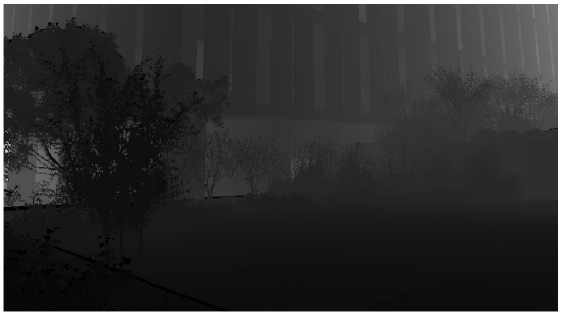
Error measure	Fog model	[53]	[86]	[114]	[118]	[7]	EVID
$l_2 - colour$	het. air.	**4.8639**	5.2898	6.824	6.3354	5.5232	4.8699
	het. air. and att.	4.0625	**3.7592**	6.3964	5.9873	4.193	4.0129
	het. att.	4.9991	**4.4006**	6.1187	6.2086	5.0711	4.4687
	homo.	6.7278	7.2462	6.8091	6.9326	6.6519	**5.857**
	All	5.1633	5.174	6.5371	6.366	5.3598	**4.8021**
MSE_{lum}	het. air.	0.0249	0.0271	0.0484	0.0513	0.0355	**0.023**
	het. air. and att.	0.0168	**0.0135**	0.0434	0.0494	0.0202	0.0162
	het. att.	0.0269	**0.0202**	0.0394	0.0513	0.0305	0.0209
	homo.	0.0519	0.0573	0.0481	0.0566	0.0532	**0.0365**
	All	0.0301	0.0295	0.0448	0.0522	0.0349	**0.0242**
MSE_{split}	het. air.	0.2691	0.281	0.3608	0.3328	0.3117	**0.2582**
	het. air. and att.	0.221	**0.1956**	0.3391	0.3123	0.2379	0.2113
	het. att.	0.2785	0.2393	0.3249	0.3276	0.2929	**0.2377**
	homo.	0.3825	0.4042	0.3611	0.3701	0.3851	**0.3191**
	All	0.2878	0.28	0.3465	0.3357	0.3069	**0.2566**
$Corr_{split}$	het. air.	**1.2215**	1.1914	0.738	0.9422	1.0925	1.2199
	het. air. att.	**1.442**	1.3639	0.8806	1.0961	1.3315	1.4163
	het. att.	1.2529	1.2573	0.9301	1.0125	1.2379	**1.2884**
	homo.	1.028	1.045	0.8417	0.8959	1.0498	**1.0765**
	All	1.2361	1.2144	0.8476	0.9867	1.1779	**1.2503**
$Corr_{lum}$	het. air.	0.6958	0.6826	0.3937	0.5203	0.6144	**0.6991**
	het. air. att.	**0.8297**	0.7839	0.4779	0.6299	0.7627	0.8180
	het. att.	0.7147	0.7198	0.5036	0.5811	0.7062	**0.7395**
	homo.	0.5598	0.5825	0.4547	0.4848	0.5742	**0.5970**
	All	0.7	0.6922	0.4575	0.554	0.6644	**0.7134**

Table 4.1: Quantitative results under the five different measures using the four different types of fog.

4.5 Conclusions on Chapter 4

In this chapter, we have proposed two substantial enhancements for the well-known perceptual colour correction framework of [3] in order to perform image dehazing and defogging. First, we have adapted the grey-world hypothesis to deal with hazy scenes. Second, we have included a set of terms to maximize inter-channels contrast, leading to the formulation of an Enhanced Variational Image Dehazing (EVID) method. The numerical experiments demonstrate the excellent behaviour of the EVID method, which is qualitatively comparable versus state-of-the-art methods for normally degraded images while outperforming them in more challenging scenarios. We have also shown how the proposed method quantitative outperforms other state-of-the-art methods in various fog scenarios.

(a) Original haze-free scene.



(b) Corresponding depth range data.

(c) Fog with heterogeneous airlight added

(d) Fog with heterogeneous attenuation added

(e) Fog with heterogeneous airlight and attenuation added

(f) Homogeneous fog added

Figure 4.15: Scene from the LIVE Color+3D Database [4], and different synthetic heterogeneous fog added on it by following the method in [5].

Chapter 5

Variational Enhancement of Full-Field Digital Mammograms

In this chapter, we deal with digital mammograms, a class of images that are not acquired from real-life scenes. Image processing and computer vision algorithms are often employed in mammographic image analysis, typically to: (a) improve the visibility of abnormalities within the breast, and (b) serve as a pre-processing step facilitating ulterior tasks such as segmentation, detection, and/or classification. In the first chapter of this dissertation, we analysed the substantial development experimented in the design of advanced and smart contrast enhancement techniques in the natural image processing area in recent years. However, this progress has not yet been translated to the medical imaging setting, where relatively simple and conventional methods are still employed for image enhancement. In this chapter, we propose to reformulate and adapt the modern mathematical model for perceptual colour correction and contrast enhancement described in Section 2.2.1.2 to the context of Full-Field Digital Mammograms (FFDM). The proposed model considers both spatial and intensity information, and also incorporates the empirical observation that sparsity should be enforced to better reveal the underlying structures of interest within a mammogram. All this information is integrated in a variational framework that allows to devise an efficient iterative numerical scheme for FFDM quality enhancement. A comprehensive twofold evaluation is proposed: we first perform a quantitative analysis of the gain of global and local contrast produced by our method, and second, we assess the improvement of visual saliency achieved on regions of interest of the mammogram that contain tumoral masses. Experimental results demonstrate that the proposed technique consistently increases the quality of the input mammogram both in terms of contrast and perceptual quality.

5.1 Introduction

Breast cancer has been ranked in the last decades as the second most prevalent cancer worldwide, and the most common in women [121]. Despite huge research efforts, prevention remains the most effective long-term strategy for the control of

breast cancer. One of the main tools to increase early cancer detection are public screening programs, a process in which women over a certain age are periodically invited to have a mammographic test. There, radiologists acquire and systematically examine mammograms to search for possible early signs of abnormalities and track their evolution over time. The main benefit of mammographic image analysis is that a mammogram may reveal subtle abnormalities which can be omitted by manual examination, and which may constitute early evidences of disease. Yet, statistics show that screening programs imply considerable risks of overdiagnosis and overtreatment, as well as psychological distress due to false positive findings [122]. Thus, any tool that has the capability of improving the accuracy of mammogram analysis is of high interest and may have a strong impact in patient welfare and also in the health care expenditure.

A mammogram is a low-dose x-ray image of the breast obtained with dedicated imaging equipment. In modern full-field digital mammography, images are recorded digitally and displayed or stored immediately on the computer system. Additionally, more and more digital data is rapidly becoming available to medical practitioners. This has led to a progressive introduction of image processing and computer vision algorithms in the field of mammographic image analysis.

Digital image processing algorithms have the potential to allow the radiologist to examine more accurately certain areas of interest within the image, and it has been demonstrated that they can have a highly favourable impact in diagnosis performance [123]. Moreover, they also serve as a necessary preprocessing step for further application of Computer Aided Detection and Diagnosis (CAD) techniques. In a typical CAD system, there are four steps involved [124], the first of which is usually image preprocessing. Then, a segmentation of suspicious regions that may contain an abnormality follows. In the next stage, a pre-trained statistical classifier is usually fed with these candidate regions, in order to classify them according to whether they contain an abnormality or not, and finally, a diagnosis of the lesion as malignant or benign is given as output.

Unfortunately, due to the lack of specialized techniques capable of properly handling the particularities of mammographic images, most of the research that focus on abnormality detection, segmentation or classification, either entirely omits the potentially useful preprocessing step, or employs conventional generic image preprocessing methodologies that ignore the specific challenges associated to mammogram enhancement. On the clinical side, specialized enhancement techniques are inevitably complex and parameter-depending. Radiologists are often unaware of most of the underlying mathematical details of these algorithms. Thus, they may end up being exposed to a cumbersome manipulation of the main parameters involved in such algorithms in order to adjust their behaviour.

In this chapter, we propose a novel mathematical model specifically designed for mammographic image enhancement. The proposed approach relies on a variational formulation of histogram equalization (HE), but adapted to the specific features of mammograms. Our approach operates globally on the mammogram, while taking into account spatial/intensity contextual information, and enforcing sparsity within the breast area to remove fat and other irrelevant tissue. The resulting iterative algorithm

is efficiently implemented thanks to the Fast Fourier Transform, and provides a set of intermediate images where the visibility of abnormalities is gradually increased. Moreover, it depends on few parameters that have a greatly intuitive interpretation.

The rest of this chapter is organized as follows. In the next section, we describe some of the most popular mammographic image enhancement methods available in the literature. Then, we review the basis of a well-known variational formulation for image contrast enhancement and colour correction. The variational point of view allows us to understand the reasons of the failure of some traditional techniques when applied to mammograms, and serves as the main setting for the derivation of our method. In section 5.4, we describe our proposed variational enhancement method. Section 5.5 is dedicated to the performance evaluation of our method when compared to other existing approaches. We provide a thorough analysis of the visibility improvement capability of our method on a set of breast masses in terms of two popular global and local performance metrics, and a novel evaluation approach based on the obtained gain in perceptual visual saliency. The paper is concluded with a discussion on the benefits and limitations of our approach as well as possible future research directions.

5.2 Background and Related Work

Mammograms are low-contrast complex images. In addition, signs of early breast cancer are usually subtle, with a great variety in shape, size and location. As a consequence, the design of efficient ad-hoc mammographic image enhancement techniques represents a challenging problem. In the past, a number of contributions have appeared trying to improve the performance of well-known general-purpose image enhancement methods.

A considerable number of recent works related to mammographic image interpretation and understanding still rely on classical or relatively simple image preprocessing methodologies. The reduction in size of the image and/or the pre-segmentation of the breast to save computational time, together with some denoising, frequently by means of simple Gaussian [125] or median [126] filters, as well as anisotropic diffusion [127] or morphological operations [128], are the most common preprocessing steps encountered in various works, e.g. [129, 130].

Likewise, basic manipulations of the dynamic range of the mammogram are often employed. For instance, direct nonlinear rescaling is implemented in [131]. In [132], a simple logarithmic transformation is performed prior to feature extraction, while in [133], scaled gamma correction is used. In [134, 135], direct histogram equalization is applied, while in [136], a very popular variant, known as Adaptive Contrast Limited Histogram Equalization (CLAHE) [19], is used as the only preprocessing step. This basic technique was demonstrated to be beneficial for abnormality detection in mammograms in [137], and it is still the state of the art in visual mammogram enhancement. In [138], a combination of CLAHE, Gaussian filtering and gamma correction is applied. In [139], CLAHE is combined with morphological operators and anisotropic diffusion, and in [140] a fuzzy version of CLAHE is employed.

Sophisticated enhancement techniques that are better adapted to the mammographic image modality have a clear potential to help improving posterior mammographic image analysis tasks. Developing such specialized methods has been a subject of research for several decades. Early work on mammographic image enhancement goes back to the eighties, and relied on the application of fundamental contrast enhancement techniques, such as simple local contrast amplification [141], edge enhancement [142], unsharp masking [143], or their combination [144].

Starting with the work [145], multi-resolution approaches soon became very popular for mammogram enhancement, with the goal of accounting for the intrinsic multi-scale nature of the different abnormalities. The main idea was to decompose the mammogram into multiple scales and modify the coefficients of the decomposition in order to reconstruct an enhanced image. This idea persists in all multi-resolution based works, and it is the choice of the decomposition and the coefficient correction techniques what varies. Wavelets are a common approach to compute suitable multi-scale decompositions. For instance, the discrete dyadic wavelet representation [146], integrated wavelets [147], contourlets [148] or redundant discrete wavelet transforms [149] have been suggested. Also, different techniques for modifying and enhancing or denoising the decomposition coefficients have been tested, such as nonlinear steerable filters [150], adaptive unsharp masking [151], wavelet shrinkage function [152], sigmoid-type transfer mapping [153] or even direct contrast modification [154, 155]. A thorough review of the details involved in the wavelet-based approach is included in [156].

Apart from the frequency analysis provided by wavelets, other ways to account for the intrinsic multiscale nature of the different abnormalities possibly present in a mammogram, such as Laplacian pyramids, have been proposed in [157, 158]. On the other hand, some authors have remained septic regarding the benefits of adopting a multiscale representation of a mammogram [159]. Moreover, some multiscale-based algorithms may require the a priori knowledge of the range of sizes of the image structures they expect to enhance, demanding thus a fine tuning of the parameters governing the algorithm, which may become critical in order to achieve good performance.

Another research direction that has drawn some attention is the application of fuzzy logic approaches in an attempt to handle the uncertainty associated with the vague definition of image features such as edges, shape, or contrast in the context of digital mammograms. In [160], the authors employ the maximum fuzzy entropy principle to map the input image into a fuzzy domain. There, the fuzzy entropy is used to locally enhance contrast, and the enhanced mammogram is mapped back into the spatial domain by *defuzzification*. The posterior work in [161] extended this approach by combining it with a structure tensor operator.

Specialized filtering techniques are also very popular for mammogram processing. In [162], authors build density maps of the breast, which are used to weight the amount of contrast enhancement for each pixel. After denoising, adaptive Wiener filtering is performed in [163]. In [164], different polynomial filters designed for edge and contrast enhancement are linearly combined to obtain an improved version of the input mammogram. The iris filter, introduced in [165], has also been reported

to be able to improve the quality of mammographic images [166]. It consists of an adaptive filter capable to enhance low-contrast rounded structures locally using maps of gradient orientations at each point.

Advanced denoising techniques, specifically adapted for mammographic image processing, have also been designed in [167–169], as well as normalization methods based on modelling the physical formation process of a mammogram [170].

Other common approaches are based on extensions of traditional well-known methods. Starting from simple histogram equalization, more elaborated techniques have been developed, such as the one encountered in [171]. The classical unsharp masking technique has also been the basis for recent methods, see [172]. However, this extension requires optimization of several parameters. They are selected by defining contrast measures, and maximizing the output contrast according to this measure. Since this is a non-differentiable metric, the optimization of the parameters is performed in a greedy manner, which is a computationally intensive approach. Also, the choice of the initial parameters may severely affect the result. Additional details on these and other related methods can be found in [173] and [174].

5.3 A Variational Formulation for Mammographic Image Enhancement

In this section, we explore a variational framework for colour correction and contrast enhancement that, when adapted to the specificities of mammograms, allows for the design of an iterative numerical scheme that gradually improves their contrast and visibility. We start from a variational formulation for histogram equalization that was proposed in [8], consisting on minimizing the following energy:

$$\mathrm{E(I)} = 2\int_{\Omega}(\mathrm{I}(x) - \frac{1}{2})^2\, dx - \frac{1}{\mathrm{WH}}\iint_{\Omega\times\Omega}|\mathrm{I}(x) - \mathrm{I}(y)|\, dxdy,$$

where x, y are spatial pixel coordinates belonging to the domain Ω of I, the size of this domain is $\mathrm{W} \times \mathrm{H}$, and the range of available intensities is normalized to $[0, 1]$.

This formulation provides some insight on the histogram equalization process: the above energy formulation can be interpreted as the combination of two qualitatively different terms: the first one assesses the separation from the mean value $1/2$, and the second one is a global measure of the contrast of the image, obtained after computing the sum of the absolute values of all pixel intensities differences.

In spite of its popularity, histogram equalization presents several drawbacks when applied to mammographic contrast improvement. First, the contrast measure in the previous equation is global, implying that the contrast associated to a given pixel is computed using information coming from every region within the image domain. Shot noise is habitual in mammographic images, and it appears as isolated white dots in uniformly dark areas. Histogram equalization overlooks neighbourhood information around these pixels, and it is thus prone to noise overenhancing. Furthermore, mammograms can be considered as relatively sparse images, in the sense that without any prior segmentation of the breast area, almost every background pixel has an intensity

close to 0. This decreases the global mean value, that tends to be low. Equalizing the histogram of a mammogram, thus, inevitably leads to an artificial increase of pixel intensity values.

There exist several extensions of histogram equalization that attempt to overcome the described weaknesses. One popular approach is Contrast Limited Adaptive Histogram Equalization (CLAHE). This technique restricts the amount of enhancement a pixel can experiment by limiting it, and tries to localize its action adaptively, computing separate histograms in different subimages. Unfortunately, the constraint in the amount of enhancement is not dependent on the image content, but rather on the contrast limiting parameter selected by the user. As a result, CLAHE performs better than simple Histogram Equalization, but its contrast enhancing capabilities remain limited.

Recently, a variational framework that extends histogram equalization by adding locality and a mechanism of attachment to data has been introduced by Bertalmío et al. in [3]. This method builds on top of the variational histogram formulation proposed in [8]. To introduce locality and penalize excessive departure from the original image, the following energy is proposed in [3]:

$$\mathrm{E(I)} = \frac{\alpha}{2}\int_{\Omega}(\mathrm{I}(x)-\frac{1}{2})^2\,dx + \frac{\beta}{2}\int_{\Omega}(\mathrm{I}(x)-\mathrm{I_0}(x))^2\,dx - \frac{\eta}{2}\iint_{\Omega\times\Omega}\omega(x,y)|\mathrm{I}(x)-\mathrm{I}(y)|\,dxdy, \tag{5.1}$$

where $\mathrm{I_0}$ is the original image, α, β, η are positive parameters and $\omega(x, y)$ is a positive weight function (properly normalized to account for image dimensions, see [3] for the details) with its value decreasing as the distance between x and y increases.

This methodology was originally developed for colour correction. The above energy is minimized in each *RGB* channel of the input image, and it keeps strong links with some perceptual mechanisms of the Human Visual System, mainly locality, colour constancy and white patch [3]. It has been later extended in [10], and further connections with the Retinex theory of colour constancy [11] were established in [9]. The main idea is that by minimizing Eq. (5.1), one is still penalizing departure from a mean gray value (*Gray World* hypothesis), but also now from the original image trough the second term, while promoting a local increase in contrast through the third term, that can in fact be written as:

$$\iint_{\Omega\times\Omega}\omega(x,y)\left(\max(\mathrm{I}(x),\mathrm{I}(y))-\min(\mathrm{I}(x),\mathrm{I}(y))\right)\,dxdy.$$

This expression is minimized whenever the minimum decreases and/or the maximum increases, corresponding to a local contrast stretching.

5.3.1 Sparsity Prior for Variational Mammographic Image Enhancement

To statistically confirm that mammograms are a relatively sparse class of images, we have conducted a simple experiment. We have considered the INbreast database ([6]),

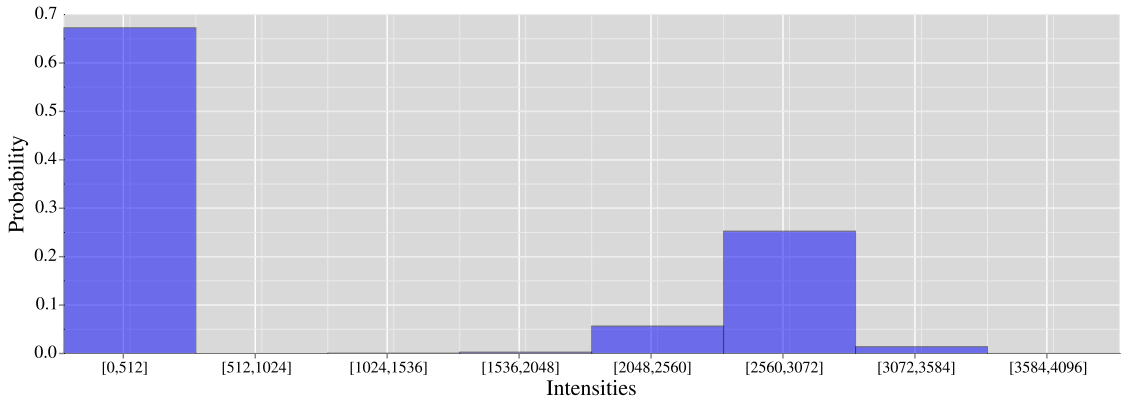


Figure 5.1: Histogram of pixel intensities in the entire INbreast database [6].

a modern FFDM database containing 410 high resolution images with a bitdepth of 2^{12}, and we have computed the mean of the intensities of every pixel in this set of images. The result is illustrated in Fig. 5.1, where pixel intensities have been grouped in sub-intervals of 512 gray levels for better visualization. We see there how the majority of the pixel intensities lies in the lower half of the available range, supporting this way our argument.

Fig. 5.1 also provides insight on some interesting features of mammograms. Namely, from the intensity distribution we can appreciate that a mammogram has qualitatively different regions. Accordingly, an enhancement algorithm should proceed differently in each of these regions. In particular, it should not enhance the background, and it should operate only within the breast. Therein, the image is not sparse anymore, but rather, medium to high values predominate. Abnormalities are also contained in that range of values, but are masked by fat and other dense tissues, which occlude their visibility.

Ideally, a completely visible abnormality would be represented by high values corresponding to abnormalities, surrounded by darker regions, i.e., lesions would exhibit better contrast with respect to the background. In other words, our goal should be to smoothly separate the central range of pixels in Fig. 5.1 into two classes in order to unveil the abnormalities within the breast. Therefore, an ideal processing of the mammogram would render an even sparser image.

The process modelled by minimizing the energy in Eq. (5.1) approximates to some extent a local histogram equalization [175], and for this reason, it shares the same drawbacks as CLAHE. To amplify contrast in the input mammogram while promoting sparsity, we modify the energy in Eq. (5.1) by removing the hypothesis that the mean value of the enhanced mammogram is 1/2, and adding a mechanism to enforce more sparsity in the resultant image. We propose to minimize the following energy:

$$\mathrm{E(I)} = \frac{\beta}{2}\int_{\Omega}(\mathrm{I}(x)-\mathrm{I}_0(x))^2\,dx + \frac{\gamma}{2}\int_{\Omega}\mathrm{I}(x)\,dx - \frac{\eta}{2}\iint_{\Omega\times\Omega}\omega(x,y)|\mathrm{I}(x)-\mathrm{I}(y)|\,dxdy. \tag{5.2}$$

This new formulation produces a local increase in contrast around abnormalities, while pushing down the intensity values of uninteresting pixels, and keeping the attachment to data term to manage excessively aggressive or too low enhancement. Moreover, although the model operates on the entire input mammogram, spatial/intensity contextual information is present in Eq. (5.2) through $\omega(x, y)$, which weights the influence of the contrast term by contextual local information around each pixel, integrating the Gaussian-weighted intensity profile of its neighbours.

5.4 Algorithm for Mammographic Image Enhancement

Formulating the problem in terms of an energy minimization naturally leads to an iterative algorithm that generates a set of images approaching a minimum of Eq. (5.2) progressively. Next, we give the details of this method.

In order to minimize the above energy, we first compute its derivative. This is detailed in the following Theorem.

Theorem 5.4.1 *The variational derivative of the energy in Eq. (5.2) is given by:*

$$\nabla \mathrm{E}(\mathrm{I})(x) = \beta(\mathrm{I}(x) - \mathrm{I}_0(x)) + \frac{\gamma}{2} - \eta \int_\Omega \omega(x, y)\mathrm{sign}(\mathrm{I}(x) - \mathrm{I}(y))\ dy. \tag{5.3}$$

Proof: Let us first introduce some notation. We denote by $\mathrm{E}_{\mathrm{A2D}}(\mathrm{I})$ to the part of the energy of Eq. (5.2) that accounts for the attachment to data; $\mathrm{E}_{\mathrm{C}}(\mathrm{I})$ is the term responsible of the local contrast increase, and $\mathrm{E}_{\mathrm{sparse}}(\mathrm{I})$ the term accounting for the sparsity prior, i.e.:

$$\begin{aligned}
\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}) &= \int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))^2\ dx, \\
\mathrm{E}_{\mathrm{sparse}}(\mathrm{I}) &= \int_\Omega \mathrm{I}(x)\ dx, \\
\mathrm{E}_{\mathrm{C}}(\mathrm{I}) &= \iint_{\Omega\times\Omega} \omega(x, y)|\mathrm{I}(x) - \mathrm{I}(y)|\ dxdy.
\end{aligned}$$

With this notation, the energy we want to differentiate can be written as:

$$\mathrm{E}(\mathrm{I}) = \frac{\beta}{2}\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}) + \frac{\gamma}{2}\mathrm{E}_{\mathrm{sparse}}(\mathrm{I}) - \frac{\eta}{2}\mathrm{E}_{\mathrm{C}}(\mathrm{I}).$$

The variation of $\mathrm{E}_{\mathrm{A2D}}(\mathrm{I})$ and $\mathrm{E}_{\mathrm{C}}(\mathrm{I})$ in the direction of a perturbation $\delta\mathrm{I}$ of the input image I has been already computed in Proposition (2.2.1) in the more general case of an even convex functional. If we consider here the sign function, we can write:

$$\begin{aligned}
\delta\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}, \delta\mathrm{I}) =& 2\int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))\ \delta\mathrm{I}(x)\ dx \\
\delta\mathrm{E}_{\mathrm{C}}(\mathrm{I}, \delta\mathrm{I}) =& \int_\Omega \left(\int_\Omega \omega(x, y)\mathrm{sign}(\mathrm{I}(x) - \mathrm{I}(y))\ dy \right) \delta\mathrm{I}(x)dx.
\end{aligned}$$

The variation of the sparsity term is a trivial application of the chain rule, and gives:

$$\delta \mathrm{E}_{\mathrm{sparse}}(\mathrm{I}, \delta \mathrm{I}) = \int_{\Omega} 1 \ \delta \mathrm{I}(x) \ dx.$$

By assembling the variations of the three terms, we obtain the following expression for the joint variation of the energy:

$$\mathrm{E}(\mathrm{I}, \delta \mathrm{I}) = \int_{\Omega} \left(\beta(\mathrm{I}(x) - \mathrm{I}_0(x)) + \frac{\gamma}{2} - \eta \int_{\Omega} \omega(x, y) \mathrm{sign}(\mathrm{I}(x) - \mathrm{I}(y)) \ dy \right) \delta \mathrm{I}(x) dx.$$

Thus, the variational derivative of the full energy is indeed given by Eq. (5.3). □

Theorem 5.4.1 enables to design a numerical algorithm to minimize the energy. First, we discretise in space the energy and its derivatives, i.e., integrals in the continuous domain Ω become sums across a discretised version of it. Let us also denote the point-wise contrast enhancement operator appearing in Eq. (5.3), in its discrete version, as:

$$\mathrm{R}_I(x) = \sum_y \omega(x, y) \mathrm{sign}(\mathrm{I}(x) - \mathrm{I}(y)). \tag{5.4}$$

There exist multiple strategies to advance towards a minimum of the energy. We adopt here a simple gradient descent, that is, we make I evolve in the direction opposite to the gradient of E,

$$\frac{\partial \mathrm{I}}{\partial t} = -\nabla \mathrm{E}(\mathrm{I}).$$

Then, by using a simple forward explicit numerical scheme, we obtain:

$$\frac{\mathrm{I}^{t+1}(x) - \mathrm{I}^t(x)}{\Delta t} = \beta(\mathrm{I}_0^t(x) - \mathrm{I}^t(x)) - \frac{\gamma}{2} + \eta \mathrm{R}_{\mathrm{I}^t}(x).$$

After some manipulations we obtain an iterative scheme, given by the following equation:

$$\mathrm{I}^{t+1}(x) = \mathrm{I}^t(x)(1 - \beta \Delta t) + \Delta t(\beta \mathrm{I}_0^t(x) + \eta \mathrm{R}_{\mathrm{I}^t}(x) - \frac{\gamma}{2} \tag{5.5}$$

Now, starting from $\mathrm{I}^{t=0} = \mathrm{I}_0$ (input image), we iterate the above equation until we arrive to a steady state, which is the final result of our algorithm. We consider steady-state has been reached as soon as the MSE (mean square error) between one iteration and the next falls below a small fixed stopping criterion. This procedure is outlined below.

Fig. (5.2) shows the result of applying the described procedure. It can be appreciated how the iterations generated by Eq. (5.5) progressively make the occluding tissue disappear, while the structures of the underlying abnormalities are gradually revealed. In the next section, we provide an exhaustive quantitative evaluation of our method.

5.5 Experimental Results

A range of different types of abnormalities can be found in a mammogram, the most interesting of which, regarding the diagnosis value, are micro-calcification clusters

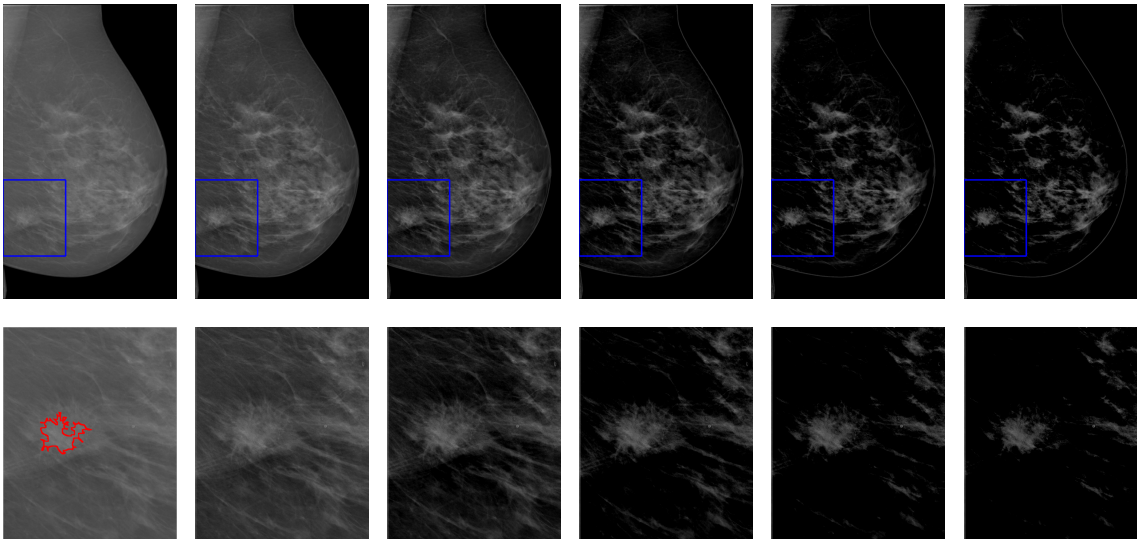


Figure 5.2: a) Original and b) -f) successive iterations of our method. (g)-(l): a magnification of the area marked in blue in the interior of the breast, containing a mass that progressively becomes more visible, while its contour appears more defined. The groundtruth labeling of the mass is marked in red on (a).

and masses. Micro-calcifications are tiny calcium deposits ranging from 20 to a few hundred microns [124], while breast masses are space-occupying lesions that are visible on two different mammogram projections. In our case, the evaluation has been performed on a subset of the INbreast database [6] consisting of mammograms containing a mass. These are Full-Field Digital Mammograms with spatial resolution of 3328×4084 and a bitdepth of 2^{12}. Since the groundtruth supplied with the images is a fine delineation of the outline of the mass, we can easily quantify the visibility gain around it with well-defined quantitative measures. The proposed approach has been tested on 107 mammograms, containing a total of 116 masses.

Algorithm 1 Iterative Algorithm for minimizing Eq. (5.2)

Initialize contribution of different terms γ, β, η
Define Δt, stop
Set $\mathrm{I}^{t=0} = \mathrm{I}$
while MSE$(\mathrm{I}^{t+1}, \mathrm{I}^{t})$ <stop: **do**
    for *pixel* in image **do**
        Compute $\mathrm{R}_{I^t}(pixel)$
        Set $\mathrm{I}^{t+1}(pixel) = \mathrm{I}^{t}(pixel)\ (1 - \beta\Delta t) + \Delta t \beta \mathrm{I}_0^t(pixel) + \eta \mathrm{R}_{\mathrm{I}^t}(pixel) - \frac{\gamma}{2}$
        Store iteration I^t
    end for
end while
Return I^{t+1}

5.5.1 Quantitative Performance Measures

Measuring the improvement in contrast and quality of a mammogram after processing is a challenging task. A considerable amount of research has been devoted to this issue [33,176–178]. For a comprehensive evaluation, in this work we include a measure to asses the global increase in contrast for mammograms proposed in [178], as well as a local measure of visibility designed specifically for regions containing masses, developed in [176]. Furthermore, we propose a new perceptually-based evaluation metric, based on an established model of visual saliency. All these performance measures are described in the following sections.

5.5.1.1 Global Quality Evaluation Metric

In [33], a measure called EME was proposed to quantify the enhancement capabilities of an image processing technique. The basic idea is to divide the image into blocks of size $k_1 \times k_2$, and compute at each block a measure of the contrast ratio given by the quotient of the maximum and the minimum intensity values within the block. Accumulating these measurement samples throughout the image, the EME measure is defined as:

$$\mathrm{EME}^{k_1,k_2}(\mathrm{I}) = \frac{1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} 20 \ln \left(\frac{\mathrm{I}^{k,l}_{\max}}{\mathrm{I}^{k,l}_{\min}} \right), \tag{5.6}$$

where $\mathrm{I}^{k,l}_{\max}$ and $\mathrm{I}^{k,l}_{\min}$ denote the maximum and minimum intensities inside the (k,l) block, that has size $k_1 \times k_2$[1]. Starting from Eq. (5.6), many variants have been proposed. From the set of possible extensions of EME, it is advocated in [178] that the most convenient to measure general contrast improvement for the mammographic image enhancement scenario is the AMEE:

$$\mathrm{AMEE}^{k_1,k_2}_{\alpha}(\mathrm{I}) = \frac{-1}{k_1 k_2} \sum_{l=1}^{k_1} \sum_{k=1}^{k_2} \alpha \left(\frac{\mathrm{I}^{k,l}_{\max}}{2} \right)^{\alpha} \ln \left(\frac{\mathrm{I}^{k,l}_{\min}}{2} \right). \tag{5.7}$$

AMEE is a non-reference image quality metric designed to measure the overall global contrast of the entire image rather than the local measurement given by Eq. (5.6). A lower AMEE measure implies a better enhancement of the image contrast.

5.5.1.2 Local Quality Evaluation Metric

Despite of the sometimes ill-defined border of masses, the groundtruth of the INbreast database for this type of lesions is particularly fine. This allows us to determine a sub-region of the mammogram containing the corresponding mass, as well as a surrounding sub-region. These are easily obtained by simple morphological operations, i.e., dilating the binary mask containing the mass, and then computing the difference between the dilated mask and the original one. An example of this procedure is shown in Fig. 5.3. The border of the inner mask is displayed in red, while the

[1] In this work, we select k_1 and k_2 as one tenth of the width and height of the mammogram, respectively.

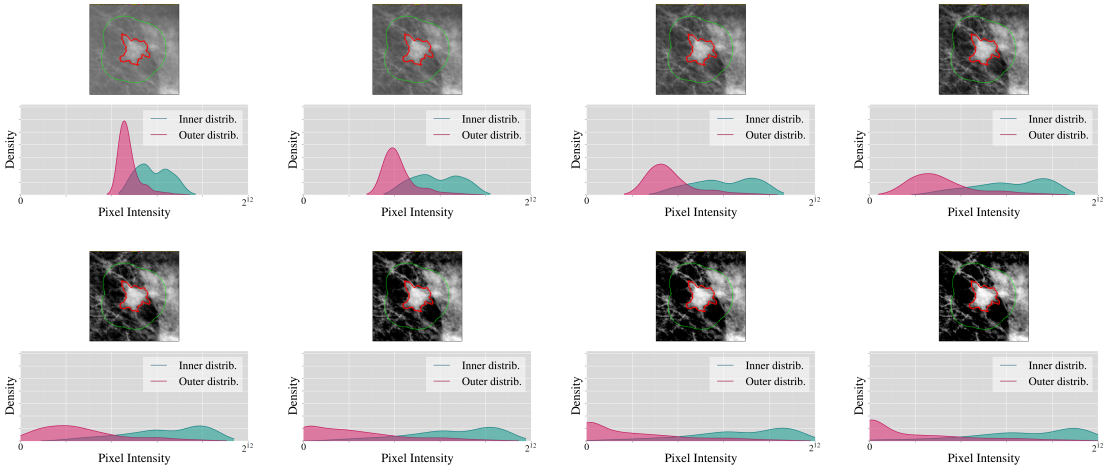


Figure 5.3: Target and background regions for the computation of the D-coefficient, together with the progressive enhancement produced by our method on a region enclosing a mass (red curve, groundtruth), and the corresponding overlap between the two regions. a) Region on the input mammogram b)-h) successive iterations of our method.

border of the surrounding outer mask is shown in green. The inner mask is thus the region enclosed by the red curve, while the outer mask is given by the region enclosed in between the red and the green curves. We include also in Fig. 5.3 the successive iterations produced by minimizing Eq. (5.2). We can observe there the ability of our proposed method to separate the intensities of each region in their histograms along these iterations, resulting in a better visibility of the mass. In what follows, as suggested in [176], we refer to the inner mask as target region, and to the outer mask as the background region.

The work in [176] introduced several quality measures to quantify the degree of enhancement obtained after processing a given mammogram. The following three error measures were defined:

$$\begin{aligned} \text{DSM} &= (|\mu_T^E - \mu_B^E|) - (|\mu_T^B - \mu_B^B|) \\ \text{TBC}_\text{s} &= \frac{(\mu_T^E/\mu_B^E) - (\mu_T^O/\mu_B^O)}{\sigma_T^E/\sigma_T^O} \\ \text{TBC}_\varepsilon &= \frac{(\mu_T^E/\mu_B^E) - (\mu_T^O/\mu_B^O)}{\varepsilon_T^E/\varepsilon_T^O}, \end{aligned}$$

where $\mu_T^O, \mu_B^O, \sigma_T^O, \sigma_B^O$ refer to the mean and standard deviation, for the original image, of the target and the background region respectively, and μ_T^E, μ_B^E, σ_T^E, σ_B^E are the corresponding counterparts for the enhanced image. Also, ε_T^E and ε_T^O are the entropy of the target region in the original and enhanced images, see [176] for further details.

Each of the above coefficients measure contrast enhancement in a different way. To merge them in a unique figure of image quality, the authors of [176] propose to compute the following joint measure:

$$\text{D} = \sqrt{(1 - \text{DSM})^2 + (1 - \text{TBC}_\text{s})^2 + (1 - \text{TBC}_\varepsilon)^2} \tag{5.8}$$

Here, the best enhancement method to process a mammogram is marked by the lowest D score.

5.5.1.3 Perceptual Saliency-Based Evaluation

Computational models of visual saliency have been heavily studied in recent years by the computer vision community. They represent an attempt to capture the ability of the Human Visual System to rapidly localize the most important information on an image. These models have been extensively validated in several consistent ways, including the study of their response to psychological patterns or the examination of their ability to predict human eye fixational movements in psychophysical experiments.

Most of the existing works on computational visual saliency rely on computing colour and contrast cues of the original image, in order to process them and obtain a saliency map that assigns higher values to salient objects. After normalization into the $[0, 1]$ interval, these saliency maps can be understood as probability distributions, where the most visually relevant areas in the scene hold higher probability values than the less relevant ones. Here we follow the works in [179] and [180], which consider saliency detection on a frequency domain approach.

The main advantage of the approach in [179] for us is that it can process grayscale images. The saliency operator proposed in that work is defined as:

$$m = g_k * (\bar{x} \circ \bar{x}), \tag{5.9}$$

where g_k is a Gaussian kernel of standard deviation k, $*$ stands for the convolution operator, $\circ$ is the Hadamard pixelwise product operator, and $\bar{x}$ represents the inverse Fourier Discrete Cosine Transform (DCT) of the signature of the image x, which is defined as $\hat{x} = \text{sign}(\text{DCT}(x))$, being x the input image.

However, the output of the above formula depends on the choice of the Gaussian kernel g_k, and it is unable to properly handle the multi-scale nature of the lesions inside the breast. To overcome this drawback, we combine Eq. (5.9) with the technique proposed in [180] to model visual saliency by means of a scale-space analysis. In that work, the authors compute a saliency map S_k at several scales k, and then the optimal scale k_p is selected to be the one that minimizes the following criterion:

$$k_p = \arg\min_k \mathcal{H}(S_k), \tag{5.10}$$

where $\mathcal{H}(x) = -\sum_{i=1}^{n} p_i \log p_i$ is the entropy operator, with p_i being the probability of occurrence in the image x of the intensity level i, which can be calculated from its histogram, and n is the number of gray levels in the dynamic range of x. The entropy of an image can be understood as a statistical measure of the randomness present in the image. More details about these two computational visual saliency models can be found in [179, 180].

To evaluate the potential of different mammogram processing methods to improve saliency of masses, we first extract a Region of Interest (ROI) of size $m_x \times m_y$ around a lesion from the unprocessed and the processed mammogram. Then, we compute

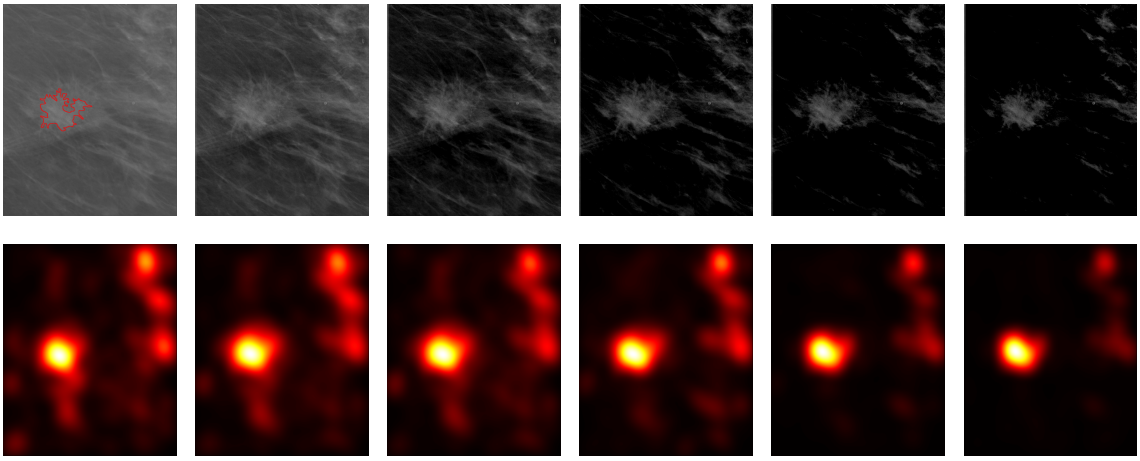


Figure 5.4: a) Region containing a mass (marked in red), corresponding to Fig. (5.3a) (b) - (f) Succesive enhancenment of the region containing the mass. (g)-(l) Saliency maps corresponding to the regions (a)-(f). Saliency distribution concentrates progressively within the mass. The amount of saliency gain with respect to the initial saliency map is, in each iteration: (h) +0.78% (i) +2.99% (j) + 5.49% (k) +13.41% (l) +16.80%.

the saliency map following Eqs. (5.9) and (5.10) for each of the ROIs, and normalize them to the interval $[0, 1]$. By simply comparing the percentage of visual saliency that the interior of the mass accumulates on both images, we can numerically quantify the goodness of a method. If we define the amount of saliency inside the mass enclosed on a ROI denoted by R_{m_x,m_y} as $\mathrm{S_{in}}(R_{m_x,m_y})$, then the *Local Saliency Gain* on that region is defined by:

$$\mathrm{LSG}(m) = \mathrm{S_{in}}(\tilde{R}_{m_x,m_y}) - \mathrm{S_{in}}(R_{m_x,m_y}), \tag{5.11}$$

where $\tilde{R}_{m_x,m_y}$ is the ROI extracted from the mammogram after it has been processed by the corresponding enhancement method m. A visual example of this saliency analysis for the mass appearing in the mammogram of Fig. (5.2) when processed with our method is shown on Fig. (5.4).

5.5.2 Parameter Configuration

Parameters β, γ, and η of Eq. (5.2) govern the overall behaviour of our method, weighting the contributions of each of each of its terms. One of the advantages of the presented method is the intuitive influence these parameters have on its performance: increasing the attachment-to-data parameter leads to a more conservative behaviour of the method, while an increase in γ produces more contrasted images and β controls the degree of *sparsification* of the output. Regarding the weighting function $\omega(x, y)$, it is responsible of the extent of local information that the method considers when increasing contrast. This parameter should be tuned in roughly approximate correspondence to the size of the objects to be enhanced. Given that the average area of a mass in the INbreast database is of 479 mm^2, we have experimented with Gaussian kernels with a standard deviation of $\sqrt{479} = 285$ mm.

The time step was always set to $\Delta t = 1/5$, and we considered that a steady-state of the gradient descent was achieved when the Mean Square Error between I^t and I^{t+1} fell below 0.07. In all the experiments, we observed convergence within $4 - 11$ iterations.

To reduce the dimensionality of the space parameter, we implemented a simple normalization strategy consisting of dividing Eq. (5.2) by $\frac{\beta}{2}$. We thus have the following model for mammogram enhancement:

$$\tilde{E}(\mathrm{I}) = \int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))^2 \, dx + \tilde{\gamma} \int_\Omega \mathrm{I}(x) \, dx - \tilde{\eta} \iint_{\Omega\times\Omega} \omega(x,y) |\mathrm{I}(x) - \mathrm{I}(y)| \, dxdy, \tag{5.12}$$

where $\tilde{\gamma} = 2\gamma/\beta$ and $\tilde{\eta} = 2\eta/\beta$. These two parameters were optimized on a random training set of 85 images and then fixed on the remaining 21 test mammograms. This process was repeated five times in a stratified cross-validation approach, selecting the parameter configuration with the best average performance on the resulting test sets. We varied parameters $\tilde{\gamma}$ and $\tilde{\eta}$ across the sets $\{0.5, 0.75, 1, 1.25, 1.5\}$ and $\{0.25, 0.37, 0.5, 0.67, 0.75\}$ respectively. The optimality criterion was the D measure detailed in the previous section. Once an optimal pair $(\tilde{\gamma}, \tilde{\eta})$ was found, it was fixed for all the experiments in this section at average values of $\tilde{\gamma} = 1.3$ and $\tilde{\eta} = 0.396$.

We have tested the performance of our method in comparison to several other popular and recent methods in the state-of-the-art. These include Contrast-Limited Adaptive Histogram Equalization (CLAHE, [137]), Density Weighted Contrast Enhancement (DWCE, [162]), Robust Polynomial Filtering (RPF, [164]), Fuzzy CLAHE (FCLAHE, [140]), and Histogram-Modified Local Contrast Enhancement (HMLCE, [171]). For a fair comparison, the described cross-validation procedure was carried out for all these methods in order to achieve a similar optimal parameter specification. Next, we provide a brief description of all these techniques and the parameters that were optimized for our experiments.

CLAHE

Contrast Limited Adaptive Histogram Equalization [137] is a classical method that still represents the state of the art in most commercial clinical systems to improve the quality of visualization on medical images. Differently from conventional global histogram equalization, CLAHE first divides the input image into small non-overlapping blocks or tiles, and each of these tiles is subjected to histogram equalization. In such a local neighbourhood, the presence of noise can influence the results severely. To control noise overenhancing, if a histogram bin has a higher value than the specified contrast limit, the associated pixels have their intensity clipped, and this amount is distributed uniformly over the other bins previous to equalizing the histogram of the tile. This clipping is performed according to a parameter κ, being a typical default parameter $\kappa = 0.01$. A higher κ results in a stronger contrast enhancement in the image, but also an increase of saturation and noise levels. To find a good trade-off, we made κ vary in the set $\{0.01, 0.015, 0.2, 0.25\}$, and carried out a cross-validation

process similar to the one described above, which led to an optimal average value of $\kappa = 0.024$.

DWCE

The DWCE technique was designed to improve mass detection on digitized mammograms [162]. It relies on a decomposition of the mammogram onto a density and a contrast image, F_D and F_C respectively. After normalizing the input mammogram on the interval $[0, 1]$, the density image is obtained applying a low-pass filter and the contrast mammogram is produced by applying a high-pass filter to the input mammogram. Next, each pixel in the density image is weighted by a multiplication factor defined by a non-linear filtering K_M of the contrast image, obtaining a density-weighted intermediate image given by $F_{KC}(x, y) = K_M(F_D(x, y)) \times F_C(x, y)$. With this, the final output of the DWCE technique is a non-linear filter of F_{KC}, given by $F_E(x, y) = K_{NL}(F_{KC}(x, y)) \times F_{KC}(x, y)$. Fig. 5.5 displays the two non-linear filters K_M and K_{NL}, the shape of which is controlled by the length of the intervals $[0, a], [a, b]$ and $[b, 1]$, as well as $[0, c], [c, d]$ and $[d, 1]$. In the original work [162], $a = 1/8, b = 2/8, c = 3/8, d = 7/8$ were employed. In our analysis, we have perturbed these values by $\pm 1/20$, so that these four coefficients vary as:

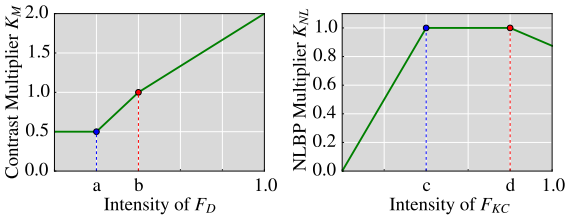


Figure 5.5: a) Contrast fucntion of DWCE b) KNL function of DWCE

$$a \in \{\frac{1}{8} - \frac{1}{20}, \frac{1}{8}, \frac{1}{8} + \frac{1}{20}\},\ b \in \{\frac{2}{8} - \frac{1}{20}, \frac{2}{8}, \frac{2}{8} + \frac{1}{20}\},$$
$$c \in \{\frac{3}{8} - \frac{1}{20}, \frac{3}{8}, \frac{3}{8} + \frac{1}{20}\}, d \in \{\frac{7}{8} - \frac{1}{20}, \frac{7}{8}, \frac{7}{8} + \frac{1}{20}\}.$$

After cross-validation, an optimal configuration for DWCE was found to be $a = 3/40, b = 3/10, c = 13/40, d = 33/40$.

Robust Polynomial Filtering

The Robust Polynomial Filtering framework (RPF) proposed in [164] consists of a set of one linear filter and two polynomial (quadratic) filters that are combined to yield an enhanced mammogram. These filters are complex and require a parameter specification coming from a large parametric space. Part of these coefficients can be fixed thanks to symmetry and isotropy properties imposed on the filters, leading to

the following defining equations:

$$
\begin{aligned}
y_{\text{linear}} &= \theta_0 x_5^{2a} + \theta_1(x_1^{2b} + x_3^{2b} + x_7^{2b} + x_9^{2b}) + \theta_2(x_2^{2c} + x_4^{2c} + x_6^{2c} + x_8^{2c}) \\
y^0_{\text{quadratic}} &= \phi_3(x_1^b x_2^c + x_1^b x_4^c + x_2^c x_3^b + x_3^b x_6^c + x_4^c x_7^b + x_6^c x_9^b + x_7^b x_8^c + x_8^c x_9^b) \\
&\quad + \phi_4(x_1^b x_5^a + x_3^b x_5^a + x_5^a x_7^b + x_5^a x_9^b) \\
&\quad + \phi_5(x_2^c x_5^a + x_4^c x_5^a + x_5^a x_6^c + x_5^a x_7^b) + \phi_6(x_2^c x_4^c + x_2^c x_6^c + x_4^c x_8^c + x_6^c x_8^c) \\
y^2_{\text{linear}} &= \phi_7(x_1^b x_3^b + x_1^b x_7^b + x_3^b x_9^b + x_7^b x_9^b) + \phi_8(x_1^b x_9^b + x_3^b x_7^b) + \phi_9(x_2^c x_8^c + x_4^c x_6^c) \\
&\quad + \phi_{10}(x_1^b x_6^c + x_1^b x_8^c + x_2^c x_7^b + x_2^c x_9^b + x_3^b x_4^c + x_3^b x_8^c + x_4^c x_9^b + x_6^c x_7^b).
\end{aligned}
$$

This three filters are finally linearly combined to produce the final output, see [164] for the details. To deal with the remaining free parameters, the authors experimentally find the following values: $\theta_0 = 0.2, \theta_1 = \theta_2 = 0.1, \phi_0 = 8\epsilon, \phi_1 = \phi_2 = \phi_6 = -\epsilon, \phi_3 = -0.5\epsilon, \phi_4 = \phi_5 = \phi_1 0 = \epsilon, \phi_7 = -2\epsilon, \phi_8 = -4\epsilon, \phi_9 = 4\epsilon$, and $a = 8\mu, b = c = \mu$. This configuration makes the filters dependent of only the two parameters ϵ and μ, that the authors propose to set as $\epsilon = 0.15$ and $\mu \in [0.5, 0.7]$. Consequently, in our experiments we set $\epsilon \in \{0.10, 0.15, 0.20\}$ and $\mu \in \{0.5, 0.6, 0.7\}$ to perform cross-validation, learning an optimal pair of parameters of $\epsilon = 0.12$ and $\mu = 0.7$.

FCLAHE

The FCLAHE method [140] is an extension of CLAHE that adds a non-linear fuzzy filtering stage to provide an extra enhancement of the gray levels of the image. The FCLAHE method does not operate on the entire mammogram, but rather on a Region of Interest, located by specifying an initial seed point on the center of a mass within the breast. The fuzzy function employed in the original work was $F(p) = \frac{1}{1+\beta|d|}$, where d is the intensity difference between p and the seed point. Parameter β controls the degree of opening of F, and thus the amount of enhancement applied within the neighbourhood of the seed point. With the help of F, FCLAHE processes input gray level intensities by means of the following formula:

$$
\mathrm{I}_{out}(x, y) = A_r(\mathrm{I}_{in}(X, Y)) + [\mathrm{I}_{in}(x, y) - A_r(\mathrm{I}_{in}(X, Y))] \cdot F(|\mathrm{I}_{in}(x, y) - A_r(\mathrm{I}_{in}(X, Y))|),
$$

where coordinates X, Y are related to the location of the seed point, *i.e.*, $X = x - x_{seed}$, $Y = y - y_{seed}$, and $A_r(\mathrm{I}_{in}(X, Y))$ computes the average of the gray levels within a distance $r = \sqrt{X^2 + Y^2}$ from the seed point. In the original work [140], an optimal β was found to be $\beta = 0.12$. Here, we have performed the same cross-validation process as for all the other methods, varying beta across the set $\{0.008, 0.010, 0.012, 0.014, 0.016\}$. We found that an optimal β for this database was $\beta = 0.0156$.

HMLC

The Histogram-Modified Local Contrast Enhancement technique was proposed in [171] as an alternative to the conventional Histogram Equalization method, in an attempt to smartly control its enhancement level. It is based on a two-step procedure.

The first step consists of a modification of the input histogram, given by the solution of a bi-criteria optimization problem, expressed as $\min \|\bar{g} - g_i\| + \psi\|\bar{g} - u\|$, where g_i is the histogram of the input mammogram, $\tilde{g}$ is the histogram of the enhanced mammogram, u is a uniform histogram and ψ is the enhancement parameter. An analytical solution to that problem is given by $\bar{g} = \dfrac{g_i + \psi u}{1 + \psi} = \dfrac{1}{1 + \psi} g_i + \dfrac{\psi}{1 + \psi} u$. In this way, the authors obtain an enhanced mammogram with a modified histogram that enters into the second step of the method. This consists of a local contrast enhancement control on $\bar{g}$, given by the following rule:

$$g(x, y) = \begin{cases} E\bar{g}(x, y) & \text{if } m_{S_{xy}} \leq K_0 M \text{ and } K_1 D \leq \sigma_{S_{xy}} \leq K_2 D \\ \bar{g}(x, y) & \text{otherwise,} \end{cases}$$

being M the global mean of the input mammogram, and D its local standard deviation, while $m_{S_{xy}}$ and $\sigma_{S_{xy}}$ are images built by storing at each pixel local mean and standard deviation values. The HMLC technique requires thus the specification of parameters ψ, E, K_0, K_1, and K_2. The authors suggest the values of $\psi = 0.8$, $E = 3.0$, $K_0 = 0.5$, $K_1 = 0.03$ and $K_2 = 0.5$. In here we fixed parameters K_0, K_1, K_2 as proposed in the original work, and optimized parameters ψ and E following the same procedure as in the previous methods, letting ψ, E vary on the sets $\{0.6, 0.7, 0.8, 0.9, 1\}$ and $\{1, 2, 3, 4, 5\}$. An optimal pair of parameters was found to be $\psi = 0.62, E = 1.4$.

5.5.3 Numerical Results

We examine now the numerical results obtained after applying the described techniques to our entire dataset, in terms of the AMEE, D and LSG coefficients.

We present first the numerical results corresponding to the global quality of the processed mammograms according to the AMEE index. Table 5.1 shows the average AMEE coefficient and standard deviation achieved by each considered approach. Let us recall that for AMEE, the lowest score represents the better enhancement quality. We see how our method supplies a substantial global improvement of the input mammograms, outperforming every other method, although HMLC exhibits a similar behaviour in terms of global enhancement. Notice that we do not report results of the FCLAHE method for the AMEE metric, since it is a local technique, meaning that in does not return a full processed mammogram, but rather a sub-region of it.

With respect to the D-score measurement, our method also achieves a greater improvement when compared to other approaches. Only the FCLAHE method provides a better D-coefficient than our approach. However, this methods enjoys the advantage of operating on a ROI containing the mass, previously defined by the radiologist. Unfortunately this requirement blocks the possibility of applying FCLAHE in a fully automated manner.

The D measures discussed above indicates a good local enhancement capability of our method, that is further confirmed by the evaluation with the Local Saliency Gain provided in Table 5.3. In this case, no method matches the increase in saliency achieved by our variational approach, not even the FCLAHE technique, although

Table 5.1: Numerical Evaluation - AMEE metric

	Ours	CLAHE [137]	DWCE [162]
Mean AMEE	**0.1487**	0.2634	0.4252
Std AMEE	0.1379	0.1834	0.2086
	RPF [164]	FCLAHE [140]	HMLC [171]
Mean AMEE	0.4384	–	0.1761
Std AMEE	0.1946	–	**0.1102**

Table 5.2: Numerical Evaluation - D-coefficient

	Ours	CLAHE [137]	DWCE [162]
Mean D-coefficient	1.2132	1.3715	1.6281
Std D	**0.3821**	0.4401	0.7724
	RPF [164]	FCLAHE [140]	HMLC [171]
Mean D-coefficient	1.5365	**1.1929**	1.2937
Std D	0.4511	0.4645	0.6626

it still performs substantially better than the rest of the considered methods, and obtains a smaller standard deviation than ours, meaning that its behaviour is slightly more robust. This agrees with the advantage that FCLAHE enjoys of enhancing only the region of interest around the masses.

Table 5.3: Numerical Evaluation - Local Saliency Gain

	Ours	CLAHE [137]	DWCE [162]
Local Saliency Gain (LSG)	**17.04**%	2.52%	−0.04%
Std LSG	5.47%	1.62%	2.02%
	RPF [164]	FCLAHE [140]	HMLC [171]
Local Saliency Gain (LSG)	2.83%	8.74%	
Std LSG	**1.70**%	3.39%	2.12%

5.5.3.1 Robustness and Sensitivity to Parameters

The parameters that define our proposed method have an intuitive meaning, and are thus useful to provide some flexibility to the user. However, to ensure the usability of the method, it is important to study the effect of varying the different parameters appearing on Eq. (5.12). We show in Table 5.4 the AMEE, D, and

LSG coefficients corresponding to executing our proposed method on our database after perturbing the sparse and contrast parameter, decreasing/increasing them by an amount of $2,5\%, 5\%, 7,5\%$ and 10%. We can see how, although the results vary with the different parameter configurations, they do it up to a relatively moderate amount: for a modification of $\pm 10\%$ of the magnitude of both parameters, the variation of the outputs is of a) 5.38% for AMEE b) 9.32% for the D error c) 4.45% for the LSG. Note also that the best configuration regarding the D coefficient does not necessarily coincide with the one optimizing the AMEE score or the Local Saliency Gain analysis, although they seem to be relatively close one to each other.

Table 5.4: Numerical Evaluation - Robustness against parameter perturbation

AMEE	$\tilde{\gamma} - 10\%$	$\tilde{\gamma} - 5\%$	$\tilde{\gamma}$	$\tilde{\gamma} + 5\%$	$\tilde{\gamma} + 10\%$
$\tilde{\eta} - 10\%$	0.1526	0.1531	0.1508	0.1486	0.1482
$\tilde{\eta} - 5\%$	0.1517	0.1493	0.1492	0.1480	0.1450
$\tilde{\eta}$	0.1505	0.1494	0.1487	0.1426	**0.1407**
$\tilde{\eta} + 5\%$	0.1514	0.1499	0.1481	0.1499	0.1452
$\tilde{\eta} + 10\%$	0.1528	0.1533	0.1517	0.1501	0.1479
D error	$\tilde{\gamma} - 10\%$	$\tilde{\gamma} - 5\%$	$\tilde{\gamma}$	$\tilde{\gamma} + 5\%$	$\tilde{\gamma} + 10\%$
$\tilde{\eta} - 10\%$	1.2710	1.2598	1.2285	1.2301	1.2414
$\tilde{\eta} - 5\%$	1.2849	1.2625	1.2279	1.2276	1.2385
$\tilde{\eta}$	1.3012	1.2755	**1.2132**	1.2419	1.2542
$\tilde{\eta} + 5\%$	1.3177	1.2849	1.2386	1.2432	1.2771
$\tilde{\eta} + 10\%$	1.3263	1.3005	1.2964	1.2689	1.2811
LSG	$\tilde{\gamma} - 10\%$	$\tilde{\gamma} - 5\%$	$\tilde{\gamma}$	$\tilde{\gamma} + 5\%$	$\tilde{\gamma} + 10\%$
$\tilde{\eta} - 10\%$	15.82	15.48%	15.29%	15.12%	14.95%
$\tilde{\eta} - 5\%$	16.64%	16.40%	16.13%	15.94%	15.53%
$\tilde{\eta}$	17.65%	17.30%	17.04%	16.80%	16.58%
$\tilde{\eta} + 5\%$	18.90%	18.64%	18.04%	17.70%	17.42%
$\tilde{\eta} + 10\%$	**21.49**%	19.95%	19.50%	19.03%	18.55%

Another interesting aspect that can be studied to asses the robustness of the local behaviour of our method is to perform the evaluation on increasingly larger ROIs containing the lesions. As can be expected, the largest the ROI enclosing the mass, the less saliency can be captured by its interior. An example of this analysis can be observed in Fig. 5.6, where we show a mass contained in regions of progressively larger sizes, both in the original mammogram and also in the one processed by our method. We see how the Local Saliency Gain decreases with the size of the region, since enlarging the ROI size amounts in practice to introducing more tissue and other visually distracting elements on the analysed region. However, our method seems to have a remarkable ability to preserve the saliency captured by the mass, even when those elements enter the visual field, while in the original mammogram the saliency contained in the mass is gradually displaced to the tissue entering the region from below, see Figs. 5.6l and 5.6x.

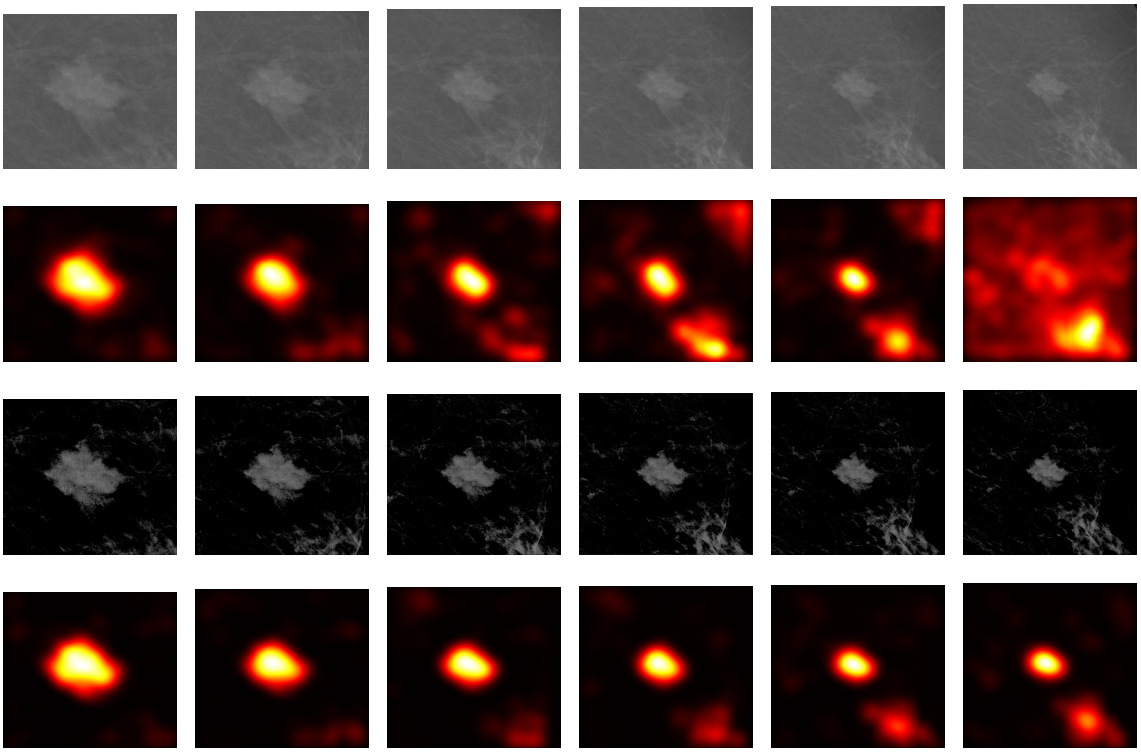


Figure 5.6: Progressive decrease of the amount of visual saliency captured by the mass within a region of increasing size. This saliency loss is partially mitigated by the application of our proposed method. (a) - (f) Region of the input mammogram containing a mass within progressively larger ROIs. (g)-(l) Corresponding saliency maps. (m)-(r) Region of the same mammogram after processing it with our proposed method. (s)-(x) Corresponding saliency maps.

5.6 Discussions and Conclusion

The numerical evaluation detailed in this work demonstrates the advantages of our proposed method with respect to other state-of-the-art methods, as well as traditional ones, for FFDM quality enhancement. Only the FCLAHE method is able to compete with our approach, but it must be taken into account that our model needs no prior localization of the regions of interest, as FCLAHE does. Additionally, our method produces a sequence of images that allow the radiologist to better correlate the original image with the postprocessed one in order to deliver a more accurate diagnostic.

Although the parameter optimization process was carried out focused on the D coefficient, it could be preferred in some situations another kind of behavior in which a stronger or weaker contrast enhancement, or less sparsity, could be desirable. Our method provides enough flexibility to control that behavior, thanks to the attachment to data parameter, which encapsulates in a single parameter the capability to fix degree of separation of the output from the input. Moreover, we have also shown how slight perturbations of the parameters defining our method do not produce unexpected results. These constitute essential features for any medical image enhancement

method, and contribute to ease the potential adoption of the technique by clinical personnel.

In terms of computational efficiency, the result of applying the operator R in Eq. (5.4) can be estimated approximating the sign function with a sigmoid mapping. In this way, it can be reformulated in terms of convolutions and computed using Fast Fourier Transforms. This adds a computational improvement to the method, since the effort to compute expression (5.4) drops from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log(N))$. Details of this argument for complexity reduction can be found in [3]. An unoptimized implementation in C++, running on an Intel Core i7 at 2.20 GHz, takes a few seconds in producing each iteration when processing a mammogram of the spatial resolutions considered here. Moreover, the method operates on the entire image, including dark background pixels that contain no information. A previous coarse segmentation of the breast could further increase the speed of our method.

Chapter 6

Summary and Future Work

Throughout this dissertation, we have extensively studied three qualitatively different image processing applications, namely underwater image restoration, fog removal, and mammographic image enhancement. These three problems have in common that structures of interest lie in areas where there is an undesired lack of contrast that must be recovered„ while other regions of the image do not need such enhancement; this demanded a smart application-driven enhancement approach, rather than the limited general existing techniques.

We have employed different (although related) techniques to solve each of these problems. For the underwater image restoration case, we proposed an extension of the fog removal method known as Dark Channel Prior [7] adapted to the particular characteristics of images acquired in an underwater environment. This implied considerations regarding the particular color cast and loss of contrast typical of these images, in which colors associated to short wavelengths tend to decay remarkably faster. For the image dehazing problem, we noticed the weaknesses of current available models based on physical principles, that do not account, for example, for a possibly uneven illumination. To overcome this deficiency, we developed a variational framework based on existing perceptual color correction techniques [3] that we extended to cope with haze degradation problems. This involved the design of appropriate energy terms that allowed us to increase saturation and contrast without the introduction of noticeable color artifacts, typical of traditional image dehazing methods. In our third application, we extended again the variational framework in [3] in order to adapt it to digital mammograms. A comprehensive evaluation based on perceptual saliency models allowed us to demonstrate the benefits of employing this novel image enhancement technique to increase visibility of tumoral masses.

On the other hand, the employed techniques are variations and extensions of general, flexible and powerful ideas, and there is still room for further research and applications. We describe below two possible future research direction in this context.

6.1 Satellite Image Enhancement. The Pansharpening Image Fusion problem

As an example of a possible further application of the framework developed in Chapter 4 of this dissertation, let us consider the problem of image pansharpening. This problem arises from the acquisition of satellital images. It consists of producing a high-resolution multispectral (being RGB a special case) image, starting with a corresponding low-resolution multispectral and a high-resolution panchromatic image, which are what satellites typically acquire. The variational approach has been recently very successful in solving this problem, since the introduction of the P+XS formulation, see [181]. As usual, this variational formulation involves a data fidelity component and a regularizing term. The data component is composed of two energy terms, responsible of keeping the consistency of the output pansharpened image with the input panchromatic and low-resolution input data, and the regularizing term aims at transferring the geometry of the panchromatic data and the diverse components of the multispectral input.

To formulate variationally this problem, we consider the unknown high-resolution multichannel image $\mathrm{I} = (\mathrm{I}_1, \ldots, \mathrm{I}_m)$, which we want to estimate. The input data captured by the satellite consists of a panchromatic image P, and a series of low-resolution spectral components $\mathrm{I}^S = (\mathrm{I}_1^S, \ldots, \mathrm{I}_m^S)$. While the high-resolution images I and P are defined on an open bounded domain $\Omega \subset \mathbb{R}^n$, the low-resolution components give us data on a sampled grid $S \subset \Omega$. The goal is thus to infer I from the given data P and I^S.

To solve this underdetermined problem, several constraints are imposed in [181], arising from considerations on the image acquisition process:

- *Panchromatic Constraint:* The unknown panchromatic image should be expressed as a linear combination of the different spectral components, with weights $\alpha_1, \ldots, \alpha_m$:
$$P(x) = \sum_{i=1}^{m} \alpha_i \mathrm{I}_i(x) \quad \forall\, x \in \Omega,$$
where $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$.

- *Spectral Constraint:* This is a data fidelity term in which we assume that every low-resolution image has been formed as a low-pass filtering of the high-resolution one, followed by a subsampling:
$$k_i * \mathrm{I}_i(p) = \mathrm{I}_i^S(p) \quad \forall p \in S, \quad 1 \leq i \leq m,$$
where k_i is the impulse response for the spectral component I_i. To ensure the ability to evaluate $k_i * \mathrm{I}_i$ at every point of the subsampling grid S, a mapping operator Π_S and an extrapolation I_i^Ω of I_i^S to the global domain Ω are defined, see [181] for details. The final data fidelity term is defined as:
$$\int_\Omega \Pi_S \cdot (k_i * \mathrm{I}_i(x) - \mathrm{I}_i^\Omega(x))^2 \, dx \quad \forall 1 \leq i \leq m.$$

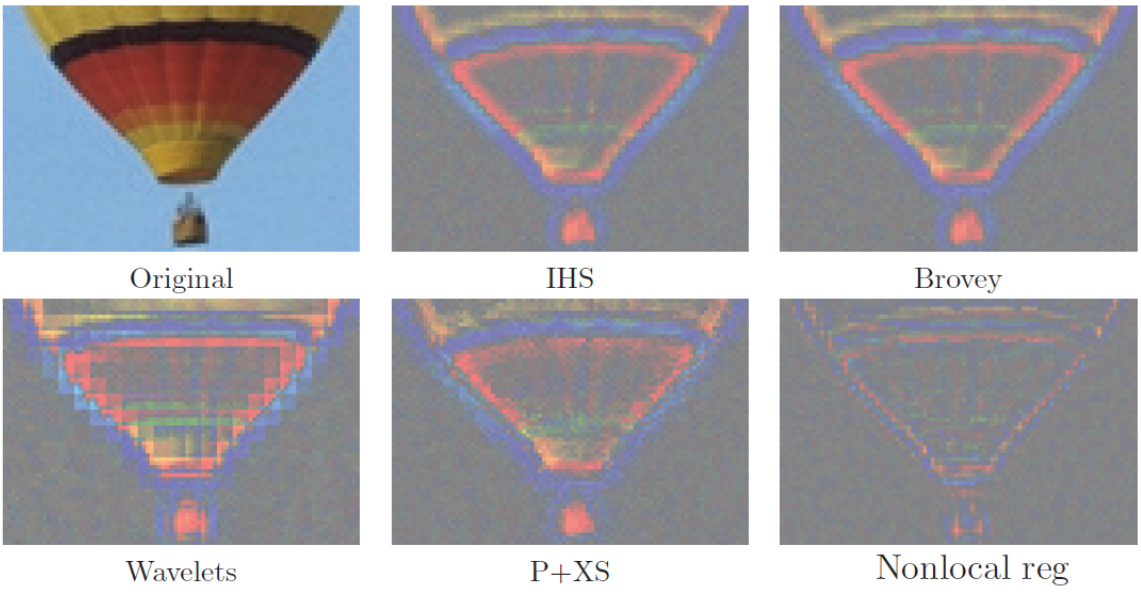


Figure 6.1: Images showing the difference between original and pansharpened versions of degraded original, as obtained by different methods. Almost every technique leads to a considerable loss of spectral information, with a reduction of color saturation, being the non-local regularization based on Eq. (6.1) the one that best performs in this case.

In [182], a non-local regularizing term is added to this formulation, aiming at transferring the geometry of the panchromatic image to the fused output. We denote it here by NL(I), and refer to the original work for the specific details.

By considering the energy terms involving the above quantities, and the non-local regularizing term, we can write the full variational framework proposed in [182]:

$$\mathrm{J(I)} = \mathrm{NL(I)} + \frac{\lambda}{2}\int_{\Omega} P(x)^2 \, dx + \frac{\mu}{2}\sum_{i=1}^{m}\int_{\Omega} \Pi_S(k_i * \mathrm{I}_i(x) - \mathrm{I}_i^{\Omega}(x))^2 \, dx, \tag{6.1}$$

where λ and μ are positive parameters weighting the contributions of each term.

Although Eq. (6.1) represents a sophisticated solution that outperforms every other recent pansharpening method, it still does not seem to fully recover the vivid colors present in the spectral components, see Fig 6.1. This resembles the same drawback we found in the dehazing or the underwater cases. Therefore, we could employ a strategy to increase saturation similar to the one we implemented in chapter 4 of this dissertation. Namely, to add the inter-channel stretching term of Eq. (4.10). With that in mind, we write the following saturation control term:

$$\mathrm{SC(I)} = \sum_{i=1}^{m} \sum_{\substack{x,y \\ j \neq i}} \omega(x,y) |\mathrm{I}_i(x,y) - \mathrm{I}_j(x,y)|,$$

where $\omega(x,y)$ is a symmetric distance function, as defined in Eq. 4.2. This energy term provides a local increase in the saturation of the output pansharpened image.

This strategy targets directly to the main weakness of the method given by the minimization of Eq. (6.1), namely that the spectral components are all treated independently, *i*.e., they all contribute equally to a sum across the available spectral

input. Another interesting possibility could be to bound the amount of stretching by the saturation observed in the input data. This could be achieved by simply writing:

$$\mathrm{SC}'(\mathrm{I}) = \sum_{i=1}^{m} \sum_{\substack{x,y \\ j \neq i}} \omega(x,y) \left(|\mathrm{I}_i - \mathrm{I}_j| - |\mathrm{I}_i^{\Omega} - \mathrm{I}_j^{\Omega}| \right)^2 .$$

The minimization of the two energies defined above, namely $\tilde{\mathrm{J}}(\mathrm{I}) = \mathrm{J}(\mathrm{I}) + \mu\ \mathrm{SC}(\mathrm{I})$ and $\tilde{\mathrm{J}}(\mathrm{I}) = \mathrm{J}(\mathrm{I}) + \mu\ \mathrm{SC}'(\mathrm{I})$, aims to represent a new approach to retain the ability of Eq. (6.1) to build a consistent pansharpened image while increasing the amount of saturation present in the fused image, making it coherent with the saturation observed in the input spectral components in the second case. Further research in this direction may lead to interesting results, although a deeper analysis is needed to validate these ideas.

6.2 Dehazing Mammograms

In Chapter 5 of this dissertation, we have explicitly disregarded the physical information about the image formation process, adopting a pure image processing approach. Our goal was to demonstrate that there is a lack of translation of potentially useful modern advanced image enhancement techniques from the natural scene field to the medical imaging area, where conventional approaches are still the dominant paradigm. However, one can easily conceive that these physical considerations can be incorporated to some extent to this scenario in order to design even more powerful image enhancement methods. Indeed, this has already been explored in the past with various objectives [183–185]. An open question is whether the simple haze/fog formation model represented by Eq. (4.1) can be useful for mammographic image enhancement, *i*.e., whether it makes sense to dehaze mammograms.

To better understand the problem, let us briefly outline the image formation model associated to X-ray imaging. X-rays are produced in the context of medical images by accelerating electrons and making them collide with heavy atoms that lie on a rotating metal disk, known as anode. These collisions take place on a X-ray tube, and they make the electrons lose part of their kinetic energy, producing a radiation. When the colliding electrons have enough energy, they produce a detachment of electrons belonging to the target metal atoms. In that case, electrons from higher states move in order to fill this vacuum, and this process emits X-ray photons at a certain frequency range. From the anode, these photons take off in all directions. A diaphragm acts as a collimator, forming the beams of X-ray photons that enter the patient body. Within it, there are many different tissues, each one with different attenuation properties. Three phenomena can occur:

1. Photons do not interact and go through the tissue meeting no atom in their way (*transmission*).

2. Photons interact with tissue atoms and transfer them some energy (*absorption*).

3. After interacting, photons are diverted into some other direction (*scattering*).

In the case of interaction with tissues, the amount of absorbed energy depends on the atomic mass of the corresponding tissue. For instance, bone and mineral have a large atomic mass, so they absorb a large amount of radiation. Other softer tissues such as fat, muscles, etc. are not so dense, and photons have a higher probability of passing through them. The attenuated radiation reaches a photon detecting film that lies behind the patient, producing a 2D image. In older days, these detectors consisted of a cassette that contained a film and an intensifying phosphor screen behind. X-ray photons went through the screen, and they were captured by the screen, which absorbed the X-rays and emitted fluorescent light that exposed the film. Nowadays, these screens have been replaced by electronic devices, giving rise to digital X-ray imaging techniques.

In any case, when an X-ray beam traversing the body meets, for instance, bone, a low quantity of photons reach the detector, leaving the film whiter. When this number of photons is larger, the film becomes darker. This happens when the beam goes through soft tissue or through air. The attenuation experimented by an X-ray beam of intensity I depends on the intensity of the source I_0, the attenuation coefficient μ of the tissue and the length of the path through that tissue d, and follows an exponential decay law:

$$\mathrm{I} = \mathrm{I}_0 \cdot e^{-\mu \cdot d}.$$

The above equation is a very simplified version of the attenuation process that takes place in the interior of the body. A more realistic (although still rather simplistic) model is the following:

$$\mathrm{I}(r) = \mathrm{I}_0 \cdot \int_{E=0}^{\infty} p(E) e^{-\int_{z=0}^{h(r)} \mu(r,z;E) dz} dE$$

where $p(E)$ is the photon energy spectrum, which depends (and is characterized) on the anode material, the tube voltage and other acquisition parameters; $\mu(r, z; E)$ is the linear attenuation coefficient for a tissue located at the 3D location (r, z), as well as the photon energy E; the inner integral is along the line of length $h(r)$. This model clearly shows complexity of modelling the physical process of X-ray attenuation within the human body.

On the other hand, this and related formulations have been successfully used to implement different quantitative mammographic image analysis tasks, such as the normalized mammographic representation in [184], leading to several interesting applications including image restoration, realistic simulations, or advanced visualizations. In [185] a model-based breast density estimation is proposed. These and related works have in common that the tissues contained in the breast are modelled in terms of two classes: fatty tissue (with a much lower attenuation) and dense, or "interesting" tissue. The latter includes glandular, fibrous or cancerous tissues. However, the price to pay for using such an advanced image acquisition model is at least twofold: first, it requires the availability of several acquisition parameters, and second, it explicitly ignores the presence of calcium.

Calcium is present sometimes in the breast in the form of calcifications, in a relatively very small quantity. Nevertheless, the presence of tiny calcification clusters (microcalcification) is one of the most useful early indicators of the presence of breast cancer. Many research has been devoted to the detection and classification of such clusters, see a comprehensive review in [186]. And it is precisely in this particular application where the simple haze formation model of (Eq. 4.2) can be useful. Let us recall it here for convenience:

$$\mathrm{I}(x) = \mathrm{t}(x)\mathrm{J}(x) + (1 - \mathrm{t}(x))\mathrm{A}. \tag{6.2}$$

Here, x is a pixel location on the foggy scene, $\mathrm{I}(x)$ is the observed intensity, $\mathrm{J}(x)$ is the scene radiance corresponding to the haze-free image, transmission $\mathrm{t}(x)$ is a scalar quantity that is inversely related to the scene's depth, while A, the airlight, plays the role of the haze color, usually considered constant over the scene.

Let us now reinterpret the above model in terms of digital mammograms and microcalcifications. To that end, we interpret that the brighter pixel in the mammogram corresponds to the most radio-opaque element within the breast. This will correspond to a calcification in case it is present. If $\mathrm{t}(x)$ is a scalar function that reaches its maximum for completely non-absorbent tissues (air), and its minimum in completely radio-opaque tissues, and $\mathrm{t}(x) \in [0, 1]$, then we have the following observations:

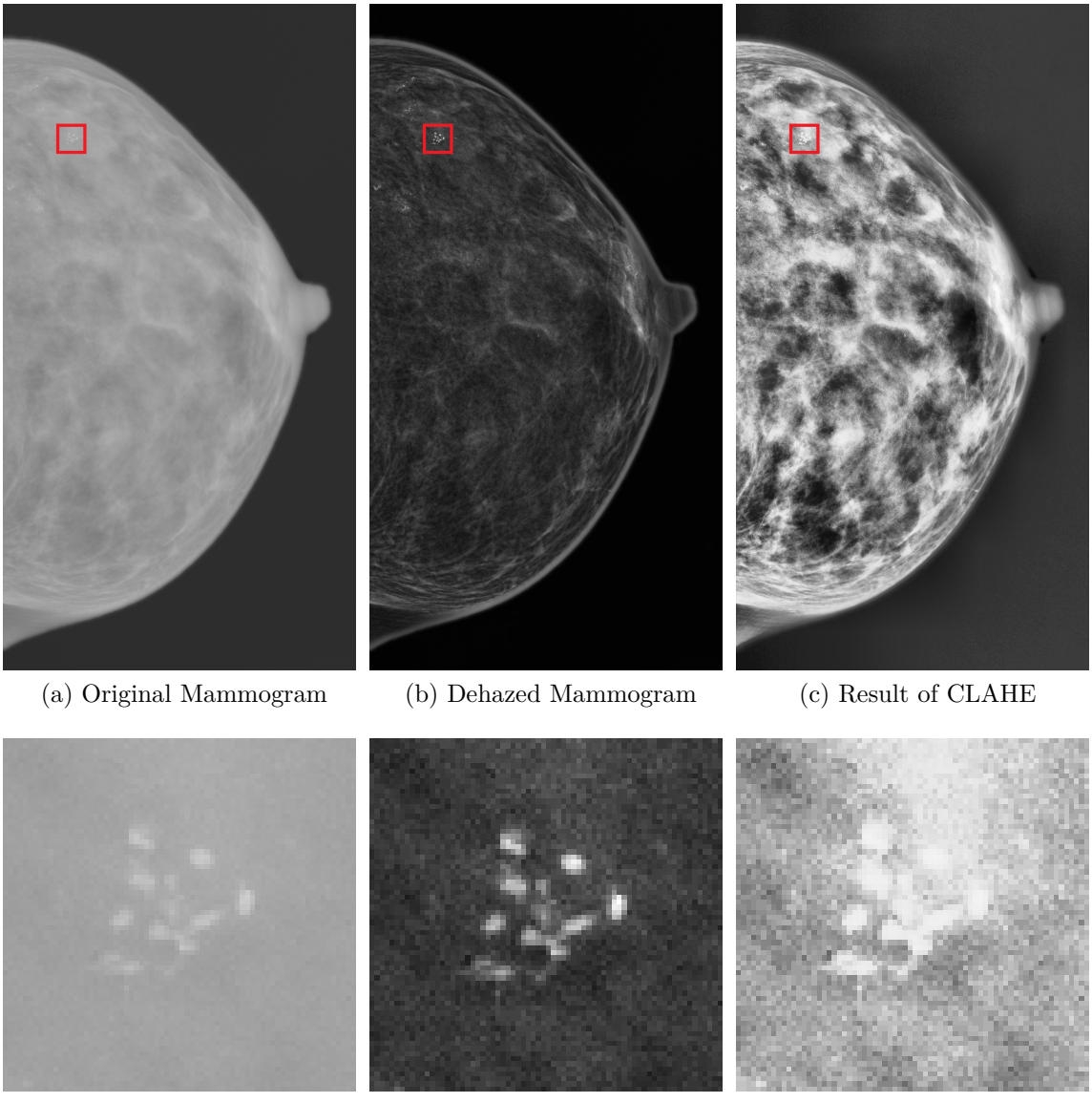
- For $\mathrm{t}(x) = 1$, the model simplifies to $\mathrm{I}(x) = \mathrm{J}(x)$. This implies that for low or non-absorbent tissues, the enhanced mammogram will remain closely similar to the input mammogram.
- For $\mathrm{t}(x) = 0$, the model simplifies to $\mathrm{I}(x) = \mathrm{A}$, the intensity of the most radio-opaque tissue, *i*.e. calcifications.
- For the rest of the mammogram, tissues with similar intermediate intensities will be regarded as equal in terms of absorption, and inversion of the above formula will lead to a similar enhancement regarding pixels associated to them.

From these remarks, we can draw the conclusion that inversion of Eq. (6.2) leads to a selective increase in contrast: non-absorbent tissues will remain unaltered in the output mammogram, and a progressive contrast enhancement will take place in the rest of the image. This contrast increase will reach its maximum power in radio-opaque elements, i.e., in the neighbourhoods of micro-calcification.

Having justified the usefulness of the haze model for the microcalcification scenario, we now turn to the issue of how to solve for $\mathrm{J}(x)$ in Eq. (6.2) to obtain a *dehazed mammogram*. This is the key advantage of adopting the approach described above: as we have reviewed in chapter 4, there are plenty of strategies available from the image dehazing literature that can be seamlessly exported to the field of mammographic image enhancement. The only requisite is that the corresponding dehazing technique is able to work on grayscale images. Unfortunately, this is a non-trivial requirement. A typical image dehazing technique aims to restore a spatially-variant

loss of contrast and saturation. Therefore, they tend to rely on luminance as well as chromatic information present in the hazy image.

Luckily, some of the most successful approaches to image dehazing can be easily particularized to grayscale images. For instance, [86] was directly proposed for both color and grayscale images. Also, a specialization of the method in [7] to the grayscale case can be straightforwardly derived by simply ignoring the chromatic component. Moreover, the method proposed in chapter 4 of this dissertation can also work in grayscale images, at least its basic version given by Eq. (4.8). In Fig. 6.2, we



(a) Original Mammogram (b) Dehazed Mammogram (c) Result of CLAHE

(d) Detail of the cluster (e) Detail of the cluster (f) Detail of the cluster

Figure 6.2: Digital mammogram containing a cluster of microcalcifications, together with the result of applying the dehazing method in [7], and the CLAHE method. We see how visibility of the cluster increases noticeably.

present a visual preliminary result of applying the method in [7] to enhance a digital mammogram containing a microcalcification cluster, as well as a comparison with the popular CLAHE algorithm [187], to illustrate the potential benefits of the use of image dehazing techniques in the context of enhancing the visibility of microcalcifications within the breast.

There are two key parameters that are typically common to every dehazing strategy, and which we must carefully take into account, since they become critical for the enhancement of microcalcification clusters: the scale and the enhancement strength.

The scale is usually employed in image dehazing to avoid confusing a white object near the observer with haze. It is often defined so as to determine a spatial neighbourhood that completely encloses such an object. In our case, this scale parameter should be defined as the mean size of a microcalcification, which may vary between 5 and 15 pixels, depending on the scanner resolution.

As for the enhancement strength parameter, dehazing methods normally set it to approximately 0.95, meaning that 5% of the haze present in the scene will be kept. It is usually argued that a complete removal of all the haze in the scene would render the output image unnatural to the human eye. From our point of view, we keep this as a free parameter of the method, which must be tuned, and should vary in a neighbourhood of 1 from below: it controls the amount of enhancement that the mammogram will experiment.

Apart of the encouraging visually pleasing results displayed in Fig. (6.2), it is important to validate the proposed method when applied as a preprocessing step for some Computer Aided Diagnosis task, such as detection or classification of microcalcification clusters. In that sense, we have produced interesting preliminary results considering as the evaluation method the increase in performance of a state-of-the-art automatic detector called casCADe, described in [188]. Numerical experiments point to a surprisingly strong added detection capability, with a decrease in the False Positive rate of the detector. Further experiments are being carried out at the moment in order to support the claim that dehazing techniques represent a powerful approach to this problem, and to better determine the potential of these ideas to impact the field of CAD based on digital mammograms.

Appendix A

Variational Derivatives

Proof of Proposition (2.2.1)

Given a differentiable and even functional J, and a symmetric function $\omega(x, y)$, we must prove:

1. $\delta C_{\omega,J}(\mathrm{I})(x) = 2\int_\Omega \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))\,dy$.
2. $\delta D_\mu(\mathrm{I})(x) = 2(\mathrm{I}(x) - \mu)$.
3. $\delta D_{\mathrm{A2D}}(\mathrm{I})(x) = 2(\mathrm{I}(x) - \mathrm{I}_0(x))$.

Recall that given a scalar function φ and an arbitrary energy:

$$\mathrm{E}(\mathrm{I}) = \int \varphi(\mathrm{I}(x))\ dx,$$

the chain rule states that its derivative in the $\delta\mathrm{I}$ direction is:

$$\delta\mathrm{E}(\mathrm{I}, \delta\mathrm{I}) = \int_\Omega \left.\frac{\partial\varphi}{\partial I}\right|_{\mathrm{I}(x)} \cdot \delta\mathrm{I}(x)\ dx, \tag{A.1}$$

which when applied to $D_\mu(\mathrm{I}) = \int_\Omega (\mathrm{I}(x) - \mu)^2\ dx$ and $D_{\mathrm{A2D}}(\mathrm{I}) = \int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))^2\ dx$ gives:

$$\begin{aligned}\delta D_\mu(\mathrm{I}, \delta\mathrm{I}) =& 2\int_\Omega (\mathrm{I}(x) - \mu(x))\ \delta\mathrm{I}(x),\\ \delta D_{\mathrm{A2D}}(\mathrm{I}, \delta\mathrm{I}) =& 2\int_\Omega (\mathrm{I}(x) - \mathrm{I}_0(x))\ \delta\mathrm{I}(x),\end{aligned}$$

which directly leads to the variations given above. Thus we only need to prove the first of the above equations, involving the contrast term $C_{\omega,J}(\mathrm{I})$. The variation of this functional is given by:

$$\begin{aligned}\delta C_{\omega,J}(\mathrm{I}, \delta\mathrm{I}^j) &= \iint_{\Omega\times\Omega} \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))(\delta\mathrm{I}(x) - \delta\mathrm{I}(y))\\ &= \iint_{\Omega\times\Omega} \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))\delta\mathrm{I}(x) - \iint_{\Omega\times\Omega} \omega(x,y)J'(\mathrm{I}(x) - \mathrm{I}(y))\delta\mathrm{I}(y).\end{aligned}$$

Since J is even, we can use the formula $J'(z) = -J'(-z)$ to rewrite the second term in the right hand side of the last equality:

$$\begin{aligned}\delta C_{\omega,J}(\mathrm{I}, \delta\mathrm{I}) &= \iint_{\Omega\times\Omega} \omega(x,y) J'(\mathrm{I}(x) - \mathrm{I}(y))\delta\mathrm{I}(x) + \iint_{\Omega\times\Omega} \omega(x,y) J'(\mathrm{I}(y) - \mathrm{I}(x))\delta\mathrm{I}(y) \\ &= \iint_{\Omega\times\Omega} \omega(x,y) J'(\mathrm{I}(x) - \mathrm{I}(y))\delta\mathrm{I}(x) + \iint_{\Omega\times\Omega} \omega(y,x) J'(\mathrm{I}(x) - \mathrm{I}(y))\delta\mathrm{I}(y).\end{aligned}$$

In the above equality, we have interchanged the notation $x \leftrightarrow y$ in the last term. Since $\omega(x, y)$ is a symmetric function, and applying Fubini's Theorem, we can interchange the order of integration in the second term of the sum, and gather both terms arriving at:

$$\delta C_{\omega,J}(\mathrm{I}, \delta\mathrm{I}) = 2\int_{\Omega} \left(2int_{\Omega}\omega(x,y) J'(\mathrm{I}(x) - \mathrm{I}(y))\ dy\right) \delta\mathrm{I}(x)\ dx. \tag{A.2}$$

This leads to the formula for the variation of $C_{\omega,J}$ given above.

Proof of Eq. (4.11)

Let us denote each of the different terms appearing in the global energy in Eq. (4.10) as follows:

$$\mathrm{E}_{mv}(\mathrm{I}^j, \mu^j) = \sum_x (\mathrm{I}^j(x) - \mu^j)^2 \tag{A.3}$$

$$\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}^j) = \sum_x (\mathrm{I}^j(x) - \mathrm{I}_0^j(x))^2 \tag{A.4}$$

$$\mathrm{E}_{\mathrm{C}}(\mathrm{I}^j, \mathrm{I}^k) = \sum_{x,y} \omega(x,y)|\mathrm{I}^j(x) - \mathrm{I}^k(y)|. \tag{A.5}$$

When $j = k$, Eq. (A.5) represents the mono-channel contrast enhancement term, and when $j \neq k$, the inter-channel stretching term. With this notation, the full energy in Eq. (4.10) can be rewritten as:

$$E(\mathrm{I}^j) = \frac{\alpha}{2}\mathrm{E}_{mv}(\mathrm{I}^j, \mu^j) + \frac{\beta}{2}\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}^j) - \frac{\gamma}{2}\mathrm{E}_{\mathrm{C}}(\mathrm{I}^j, \mathrm{I}^j) - \frac{\eta}{2}\left(\mathrm{E}_{\mathrm{C}}(\mathrm{I}^j, \mathrm{I}^{j+1}) + \mathrm{E}_{\mathrm{C}}(\mathrm{I}^j, \mathrm{I}^{j+2})\right).$$

For the cases of $\mathrm{E}_{mv}(\mathrm{I}^j, \mu^j)$ and $\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}^j)$, as well as the mono-channel contrast enhancement $\mathrm{E}_{\mathrm{C}}(\mathrm{I}^j, \mathrm{I}^j)$, these are only discrete versions of the results provided in Proposition (2.2.1), when $J(x) = |x|_\epsilon$, *i.*e., a smooth (even) approximation to the absolute value function. Thus, their discrete variational derivatives are given by :

$$\begin{aligned}\delta\mathrm{E}_{mv}((\mathrm{I}^j, \mu^j), \delta\mathrm{I}^j) &= 2\sum_x (\mathrm{I}^j(x) - \mu^j(x))\ \delta\mathrm{I}^j(x) \\ \delta\mathrm{E}_{\mathrm{A2D}}(\mathrm{I}^j, \delta\mathrm{I}^j) &= 2\sum_x (\mathrm{I}^j(x) - \mathrm{I}_0^j(x))\ \delta\mathrm{I}^j(x). \\ \delta\mathrm{E}_{\mathrm{C}}^{\epsilon}(\mathrm{I}^j, \delta\mathrm{I}^j) &= 2\sum_x \left(\sum_y \omega(x,y)\mathrm{sign}_\epsilon(\mathrm{I}^j(x) - \mathrm{I}^j(y))\ dy\right) \delta\mathrm{I}^j(x),\end{aligned}$$

where sign_ϵ represents the derivative of $|\cdot|_\epsilon$, tending to the sign function as ϵ approximates 0. For the case of the inter-channel stretching term, *i*.e., the variation of $\text{E}_\text{C}(\text{I}^j, \text{I}^k)$ in the direction of I^j when $j \neq k$, it is a simplification of the derivations in Proposition (2.2.1) . Since in this case I^k can be considered a constant, we can directly apply a discrete version of the chain rule, obtaining:

$$\delta \text{E}_\text{C}^\epsilon((\text{I}^j, \text{I}^k), \delta \text{I}^j) = 2 \sum_x \left(\sum_y \omega(x,y) \text{sign}_\epsilon(\text{I}^j(x) - \text{I}^k(y)) \right) \delta \text{I}^j(x). \tag{A.6}$$

By letting ϵ tend to zero in the Eq. (A.6) and gathering together all variations, we can write the variation of the full energy in the direction of δI^j as:

$$\begin{aligned}\delta E(\text{I}^j) = \sum_x \Bigg[& \alpha(\text{I}^j(x) - \mu^j) + \beta(\text{I}^j(x) - \text{I}_0^j(x)) - \gamma \sum_y \omega(x,y) \text{sign}(\text{I}^j(x) - \text{I}^j(y)) \\ & - \eta \left(\sum_y \omega(x,y) \text{sign}(\text{I}^j(x) - \text{I}^{j+1}(y)) + \sum_y \omega(x,y) \text{sign}(\text{I}^j(x) - \text{I}^{j+2}(y)) \right) \Bigg] \ \delta \text{I}^j(x).\end{aligned}$$

This shows that the gradient of the full energy amounts to the expression shown in Eq. (4.11).

Appendix B

Main Accomplishments of this Dissertation

During this Ph.D., in addition to develop the work described in this document, we have published the following scientific papers in peer-reviewed international journals and conferences:

- A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, *A Variational Framework for Single Image Dehazing*, in Proc. of the 4th Color and Photometry in Computer Vision Workshop (ECCV 2014), Zurich, Sept. 2014.
- A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, *Automatic Red-Channel Underwater Image Restoration*, Journal of Visual Communication and Image Representation, 26 (2015), pp. 132-145.
- A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, *Enhanced Variational Image Dehazing, SIAM Journal on Imaging Sciences*, accepted.
- A. Galdran, A. Picón, E. Garrote, A. Bereciartua, D. Pardo, *Sparsity-Based Context-Aware Variational Enhancement of Full-Field Digital Mammograms*, submitted.

The following talks have also been presented in different institutions:

- A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, *A Variational Framework for Single Image Dehazing*, in the 4th Color and Photometry in Computer Vision Workshop (ECCV 2014), Zurich, Sept. 12$^{\text{th}}$, 2014.
- A. Galdran, J. Vazquez-Corral, D. Pardo, and M. Bertalmío, *A Variational Framework for Single Image Dehazing*, in the Third BCAM Workshop on Computational Mathematics, Bilbao, July 18$^{\text{th}}$, 2014.
- A. Galdran, Artzai Picón, and D. Pardo, *Image Restoration Techniques on Hazy and Underwater Environments*, in the Mathematics and Applications Days (V-MAD 4), Valparaiso, Chile, January 8$^{\text{th}}$, 2014.

- A. Galdran, Artzai Picón, and D. Pardo, *Distance Dependent Image Degradation. Restoration Algorithms and Applications to Atmospheric and Underwater image enhancement*, in the Third BCAM Workshop on Computational Mathematics, Bilbao, October 11$^{\text{th}}$, 2013.

- A. Galdran, *Mathematical Image Segmentation. Variational and Level Set Techniques*, in Minisymposium: Projects in technology centres of the Basque Country. Santiago de Compostela, July 12$^{\text{th}}$, 2012.

References

[1] J. Yang, B. Zhang, and Y. Shi, "Scattering Removal for Finger-Vein Image Restoration," *Sensors*, vol. 12, pp. 3627–3640, Mar. 2012.

[2] Y. Wang and C. Fan, "Single Image Defogging by Multiscale Depth Fusion," *IEEE Transactions on Image Processing*, vol. 23, pp. 4826–4837, Nov. 2014.

[3] M. Bertalmío, V. Caselles, E. Provenzi, and A. Rizzi, "Perceptual color correction through variational techniques," *IEEE Transactions on Image Processing*, vol. 16, pp. 1058–1072, Apr. 2007.

[4] C.-C. Su, L. Cormack, and A. Bovik, "Color and Depth Priors in Natural Images," *IEEE Transactions on Image Processing*, vol. 22, pp. 2259–2274, June 2013.

[5] J.-P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision Enhancement in Homogeneous and Heterogeneous Fog," *IEEE Intelligent Transportation Systems Magazine*, vol. 4, pp. 6–20, Apr. 2012.

[6] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "INbreast: toward a full-field digital mammographic database," *Academic Radiology*, vol. 19, pp. 236–248, Feb. 2012.

[7] K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2341–2353, Dec. 2011.

[8] G. Sapiro and V. Caselles, "Histogram Modification via Differential Equations," *Journal of Differential Equations*, vol. 135, pp. 238–268, Apr. 1997.

[9] M. Bertalmío, V. Caselles, and E. Provenzi, "Issues About Retinex Theory and Contrast Enhancement," *International Journal of Computer Vision*, vol. 83, pp. 101–119, June 2009.

[10] R. Palma-Amestoy, E. Provenzi, M. Bertalmío, and V. Caselles, "A Perceptually Inspired Variational Framework for Color Enhancement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 458–474, Mar. 2009.

[11] E. H. Land and J. J. McCann, "Lightness and Retinex Theory," *Journal of the Optical Society of America*, vol. 61, pp. 1–11, Jan. 1971.

[12] K. He, J. Sun, and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," in *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009*, pp. 1956–1963, June 2009.

[13] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (The Morgan Kaufmann Series in Computer Graphics)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2005.

[14] H. Farid, "Blind inverse gamma correction," *IEEE Transactions on Image Processing*, vol. 10, pp. 1428–1433, Oct. 2001.

[15] Y. Shi, J. Yang, and R. Wu, "Reducing Illumination Based on Nonlinear Gamma Correction," in *IEEE International Conference on Image Processing, 2007. ICIP 2007*, vol. 1, pp. I – 529–I – 532, Sept. 2007.

[16] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient Contrast Enhancement Using Adaptive Gamma Correction With Weighting Distribution," *IEEE Transactions on Image Processing*, vol. 22, pp. 1032–1041, Mar. 2013.

[17] B. Liu, X. Wang, W. Jin, Y. Chen, C. Liu, and X. Liu, "Infrared image detail enhancement based on local adaptive gamma correction," *Chinese Optics Letters*, vol. 10, pp. 21–28, Feb. 2012.

[18] J. Lisani, A. Petro, and C. Sbert, "Color and Contrast Enhancement by Controlled Piecewise Affine Histogram Equalization," *Image Processing On Line (IPOL)*, vol. 2, pp. 243–265, Oct. 2012.

[19] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer Vision, Graphics, and Image Processing*, vol. 39, pp. 355–368, Sept. 1987.

[20] R. Dale-Jones and T. Tjahjadi, "A study and modification of the local histogram equalization algorithm," *Pattern Recognition*, vol. 26, pp. 1373–1381, Sept. 1993.

[21] H. Zhu, F. H. Y. Chan, and F. K. Lam, "Image Contrast Enhancement by Constrained Local Histogram Equalization," *Computer Vision and Image Understanding*, vol. 73, pp. 281–290, Feb. 1999.

[22] H. D. Cheng and X. J. Shi, "A simple and effective histogram equalization approach to image enhancement," *Digital Signal Processing*, vol. 14, pp. 158–170, Mar. 2004.

[23] Q. Wang and R. Ward, "Fast Image/Video Contrast Enhancement Based on Weighted Thresholded Histogram Equalization," *IEEE Transactions on Consumer Electronics*, vol. 53, pp. 757–764, May 2007.

[24] X. Wu, "A Linear Programming Approach for Optimal Contrast-Tone Mapping," *IEEE Transactions on Image Processing*, vol. 20, pp. 1262–1272, May 2011.

[25] S. Hashemi, S. Kiani, N. Noroozi, and M. E. Moghaddam, "An image contrast enhancement method based on genetic algorithm," *Pattern Recognition Letters*, vol. 31, pp. 1816–1824, Oct. 2010.

[26] T. Arici, S. Dikbas, and Y. Altunbasak, "A Histogram Modification Framework and Its Application for Image Contrast Enhancement," *IEEE Transactions on Image Processing*, vol. 18, pp. 1921–1935, Sept. 2009.

[27] T. Celik and T. Tjahjadi, "Contextual and Variational Contrast Enhancement," *IEEE Transactions on Image Processing*, vol. 20, pp. 3431–3441, Dec. 2011.

[28] H. Xu, G. Zhai, X. Wu, and X. Yang, "Generalized Equalization Model for Image Enhancement," *IEEE Transactions on Multimedia*, vol. 16, pp. 68–82, Jan. 2014.

[29] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognition*, vol. 45, pp. 3810–3824, Oct. 2012.

[30] N. Bassiou and C. Kotropoulos, "Color image histogram equalization by absolute discounting back-off," *Computer Vision and Image Understanding*, vol. 107, pp. 108–122, July 2007.

[31] J.-H. Han, S. Yang, and B.-U. Lee, "A novel 3-D color histogram equalization method with uniform 1-D gray scale histogram," *IEEE Transactions on Image Processing*, vol. 20, pp. 506–512, Feb. 2011.

[32] K. Zuiderveld, "Graphics Gems IV," pp. 474–485, San Diego, CA, USA: Academic Press Professional, Inc., 1994.

[33] S. S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Transactions on Image Processing*, vol. 16, pp. 741–758, Mar. 2007.

[34] T. Huang, G. Yang, and G. Tang, "A fast two-dimensional median filtering algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, pp. 13–18, Feb. 1979.

[35] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, July 1990.

[36] A. Buades, B. Coll, and J.-M. Morel, "Self-similarity-based Image Denoising," *Communications of the ACM*, vol. 54, pp. 109–117, May 2011.

[37] A. Polesel, G. Ramponi, and V. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Transactions on Image Processing*, vol. 9, pp. 505–510, Mar. 2000.

[38] Y.-H. Kim and Y. Cho, "Feature and noise adaptive unsharp masking based on statistical hypotheses test," *IEEE Transactions on Consumer Electronics*, vol. 54, pp. 823–830, May 2008.

[39] G. Deng, "A Generalized Unsharp Masking Algorithm," *IEEE Transactions on Image Processing*, vol. 20, pp. 1249–1261, May 2011.

[40] S. A. Broughton, *Discrete Fourier Analysis and Wavelets: Applications to Signal and Image Processing*. Hoboken, N.J: Wiley, edition: 1 ed., Nov. 2008.

[41] R. L. Easton, *Fourier Methods in Imaging*. John Wiley & Sons, Nov. 2010.

[42] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[43] I. S. Uzun and A. Amira, "FPGA implementations of fast Fourier transforms for real-time signal and image processing," *IET Proceedings of the Vision, Image and Signal Processing Conference*, pp. 283 – 296, June 2005.

[44] G. Mandyam, N. Ahmed, and N. Magotra, "Lossless Image Compression Using the Discrete Cosine Transform," *Journal of Visual Communication and Image Representation*, vol. 8, pp. 21–26, Mar. 1997.

[45] S. Matej, J. Fessler, and I. Kazantsev, "Iterative tomographic image reconstruction using Fourier-based forward and back-projectors," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 401–412, Apr. 2004.

[46] R. S. A. Oppenheim, "Nonlinear filtering of multiplied and convolved signals," *IEEE Transactions on Audio and Electroacoustics*, vol. 16, pp. 437– 466, Sept. 1968.

[47] C.-N. Fan and F.-Y. Zhang, "Homomorphic filtering based illumination normalization method for face recognition," *Pattern Recognition Letters*, vol. 32, pp. 1468–1479, July 2011.

[48] N. Dhinagar and M. Celenk, "Ultrasound medical image enhancement and segmentation using adaptive homomorphic filtering and histogram thresholding," in *2012 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pp. 349–353, Dec. 2012.

[49] K. Hassani, K. Bajelani, M. Navidbakhsh, D. Doyle, and F. Taherian, "Heart sound segmentation based on homomorphic filtering," *Perfusion*, vol. 29, pp. 351–359, Feb. 2014.

[50] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications.* Stationary time series, Cambridge: Technology Press of the Massachusetts Institute of Technology, 1950.

[51] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction.* John Wiley & Sons, 2006.

[52] R. Hardie, "A Fast Image Super-Resolution Algorithm Using an Adaptive Wiener Filter," *IEEE Transactions on Image Processing*, vol. 16, pp. 2953–2964, Dec. 2007.

[53] K. Gibson and T. Nguyen, "Fast single image fog removal using the adaptive Wiener filter," in *2013 20th IEEE International Conference on Image Processing (ICIP)*, pp. 714–718, Sept. 2013.

[54] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Transactions on Image Processing*, vol. 16, pp. 2080–2095, Aug. 2007.

[55] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way.* Amsterdam; Boston: Academic Press, 3 edition ed., Dec. 2008.

[56] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, July 1989.

[57] D. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, pp. 613–627, May 1995.

[58] S. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, pp. 1532–1546, Sept. 2000.

[59] S. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing*, vol. 9, pp. 1522–1531, Sept. 2000.

[60] D. Cho, T. D. Bui, and G. Chen, "Image denoising based on wavelet shrinkage using neighbor and level dependency," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 07, pp. 299–311, May 2009.

[61] E. Provenzi, C. Gatta, M. Fierro, and A. Rizzi, "A Spatially Variant White-Patch and Gray-World Method for Color Image Enhancement Driven by Local Contrast," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1757–1770, Oct. 2008.

[62] A. Rizzi, C. Gatta, and D. Marini, "Color correction between gray world and white patch," in *in IS&T/SPIE Electronic Imaging 2002. The human Vision and Electronic Imaging VII Conference*, pp. 367–375, May 2002.

[63] A. Rizzi, C. Gatta, and D. Marini, "A New Algorithm for Unsupervised Global and Local Color Correction," *Pattern Recognition Letters*, vol. 24, pp. 1663–1677, July 2003.

[64] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine*, vol. 13, pp. 43–64, May 1996.

[65] A. Levin, Y. Weiss, F. Durand, and W. Freeman, "Understanding Blind Deconvolution Algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2354–2367, Dec. 2011.

[66] L. Sun, S. Cho, J. Wang, and J. Hays, "Good Image Priors for Non-blind Deconvolution," in *Computer Vision - ECCV 2014* (D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, eds.), no. 8692 in Lecture Notes in Computer Science, pp. 231–246, Springer International Publishing, Sept. 2014.

[67] H. Koschmieder, *Theorie der horizontalen Sichtweite: Kontrast und Sichtweite*. Keim & Nemnich, 1925.

[68] M. Ludvigsen, B. Sortland, G. Johnsen, and H. Singh, "Applications of Geo-Referenced Underwater Photo Mosaics in Marine Biology and Archaeology," *Oceanography*, vol. 20, pp. 140–149, Dec. 2007.

[69] N. Strachan, "Recognition of fish species by colour and shape," *Image and Vision Computing*, vol. 11, pp. 2–10, Jan. 1993.

[70] M. Costa, P. Goncalves, A. Martins, and E. Silva, "Vision-based assisted teleoperation for inspection tasks with a small ROV," in *Oceans, 2012*, pp. 1–8, Oct. 2012.

[71] N. Gracias and J. Santos-Victor, "Underwater Video Mosaics as Visual Navigation Maps," *Computer Vision and Image Understanding*, vol. 79, pp. 66–91, July 2000.

[72] N. Gracias, M. Mahoor, S. Negahdaripour, and A. Gleason, "Fast image blending using watersheds and graph cuts," *Image and Vision Computing*, vol. 27, pp. 597–607, Apr. 2009.

[73] Y. Schechner and N. Karpel, "Clear underwater vision," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*, vol. 1, pp. I–536–I–543 Vol.1, July 2004.

[74] Y. Schechner and N. Karpel, "Recovery of underwater visibility and structure by polarization analysis," *IEEE Journal of Oceanic Engineering*, vol. 30, pp. 570–587, July 2005.

[75] L. A. Torres-Mendez and G. Dudek, "Color correction of underwater images for aquatic robot inspection," in *Proceedings of the 5th international conference on Energy Minimization Methods in Computer Vision and Pattern Recognition*, EMMCVPR'05, (Berlin, Heidelberg), pp. 60–73, Springer-Verlag, Nov. 2005.

[76] C. Ancuti, C. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 81–88, June 2012.

[77] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi, "Underwater color constancy: enhancement of automatic live fish recognition," in *Proc. SPIE*, vol. 5293, pp. 157–168, Dec. 2003.

[78] K. Iqbal, R. Abdul Salam, O. Mohd, and T. Abdullah Zawawi, "Underwater Image Enhancement Using An Integrated Colour Model," *International Journal of Computer Science*, vol. 32, pp. 239–244, Nov. 2007.

[79] R. Schettini and S. Corchs, "Underwater Image Processing: State of the Art of Restoration and Image Enhancement Methods," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 14, Apr. 2010.

[80] N. Carlevaris-Bianco, A. Mohan, and R. Eustice, "Initial results in underwater single image dehazing," in *Oceans 2010*, pp. 1–8, Sept. 2010.

[81] J. Chiang and Y.-C. Chen, "Underwater Image Enhancement by Wavelength Compensation and Dehazing," *IEEE Transactions on Image Processing*, vol. 21, pp. 1756–1769, Apr. 2012.

[82] K. He, J. Sun, and X. Tang, "Guided Image Filtering," in *Proceedings of the 11th European Conference on Computer Vision: Part I*, ECCV'10, (Berlin, Heidelberg), pp. 1–14, Springer-Verlag, Sept. 2010.

[83] A. Levin, D. Lischinski, and Y. Weiss, "A Closed-Form Solution to Natural Image Matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 228–242, Feb. 2008.

[84] K. N. Plataniotis and A. N. Venetsanopoulos, *Color Image Processing and Applications*. Springer, May 2000.

[85] N. Hautiere, J.-P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Analysis & Stereology*, vol. 27, pp. 87–95, May 2011.

[86] J. P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *2009 IEEE 12th International Conference on Computer Vision*, pp. 2201–2208, Sept. 2009.

[87] J. Malkasse, L. Jaulin, I. Quidu, and S. Bazeille, "Automatic underwater image pre-preprocessing," *Sea Tech Week Caracterisation du Milieu Marin*, Oct. 2006.

[88] S. Serikawa and H. Lu, "Underwater Image Dehazing Using Joint Trilateral Filter," *Computers & Electrical Engineering*, vol. 40, pp. 41–50, Jan. 2014.

[89] B. Henke, M. Vahl, and Z. Zhou, "Removing color cast of underwater images through non-constant color constancy hypothesis," in *2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 20–24, Sept. 2013.

[90] N. Hautiere, J. P. Tarel, and D. Aubert, "Towards Fog-Free In-Vehicle Vision Systems through Contrast Restoration," in *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07*, pp. 1–8, June 2007.

[91] J. Long, Z. Shi, W. Tang, and C. Zhang, "Single Remote Sensing Image Dehazing," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, pp. 59–63, Jan. 2014.

[92] I. Yoon, S. Kim, D. Kim, M. Hayes, and J. Paik, "Adaptive defogging with color correction in the HSV color space for consumer surveillance system," *IEEE Transactions on Consumer Electronics*, vol. 58, pp. 111–116, Feb. 2012.

[93] Y. Shiau, H. Yang, P. Chen, and Y. Chuang, "Hardware Implementation of a Fast and Efficient Haze Removal Method," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, pp. 1369–1374, Aug. 2013.

[94] Y. Zhang and B. Guindon, "Quantitative assessment of a haze suppression methodology for satellite imagery: effect on land cover classification performance," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1082–1089, May 2003.

[95] A. Makarau, R. Richter, R. Muller, and P. Reinartz, "Haze Detection and Removal in Remotely Sensed Multispectral Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, pp. 5895–5905, Sept. 2014.

[96] W. E. K. Middleton, *Vision through the atmosphere*. University of Toronto Press, 1952.

[97] E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles*. Wiley, Jan. 1976.

[98] S. Nayar and S. Narasimhan, "Vision in bad weather," in *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*, vol. 2, pp. 820–827, 1999.

[99] S. Narasimhan and S. Nayar, "Chromatic framework for vision in bad weather," in *IEEE Conference on Computer Vision and Pattern Recognition, 2000. Proceedings*, vol. 1, pp. 598–605 vol.1, June 2000.

[100] Y. Schechner, S. Narasimhan, and S. Nayar, "Instant dehazing of images using polarization," vol. 1, pp. I–325–I–332 vol.1, June 2001.

[101] S. Narasimhan and S. Nayar, "Vision and the Atmosphere," *International Journal of Computer Vision*, vol. 48, pp. 233–254, July 2002.

[102] S. Narasimhan and S. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 713–724, June 2003.

[103] S. Narasimhan and S. Nayar, "Interactive (de) weathering of an image using physical models," *IEEE Workshop on Color and Photometric Methods in Computer Vision*, vol. 6, p. 1, Sept. 2003.

[104] R. Fattal, "Dehazing using Color-Lines," *ACM Transaction on Graphics*, Dec. 2014.

[105] R. Tan, "Visibility in bad weather from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition,sCVPR 2008*, pp. 1–8, June 2008.

[106] R. Fattal, "Single Image Dehazing," in *ACM SIGGRAPH Papers*, SIGGRAPH '08, (New York, NY, USA), pp. 72:1–72:9, ACM, Aug. 2008.

[107] L. Schaul, C. Fredembach, and S. Susstrunk, "Color image dehazing using the near-infrared," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, pp. 1629–1632, Nov. 2009.

[108] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep Photo: Model-based Photograph Enhancement and Viewing," in *ACM SIGGRAPH Asia 2008 Papers*, (New York, NY, USA), pp. 116:1–116:10, ACM, Dec. 2008.

[109] P. Carr and R. Hartley, "Improved Single Image Dehazing Using Geometry," in *Digital Image Computing: Techniques and Applications, 2009. DICTA '09.*, pp. 103–110, Dec. 2009.

[110] C. Ancuti and C. Ancuti, "Single Image Dehazing by Multi-Scale Fusion," *IEEE Transactions on Image Processing*, vol. 22, pp. 3271–3282, Aug. 2013.

[111] R. Kaftory, Y. Schechner, and Y. Zeevi, "Variational Distance-Dependent Image Restoration," in *IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR '07*, pp. 1–8, 2007.

[112] F. Fang, F. Li, and T. Zeng, "Single Image Dehazing and Denoising: A Fast Variational Approach," *SIAM Journal on Imaging Sciences*, vol. 7, pp. 969–996, Jan. 2014.

[113] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, pp. 120–145, May 2011.

[114] K. Nishino, L. Kratz, and S. Lombardi, "Bayesian Defogging," *International Journal of Computer Vision*, vol. 98, pp. 263–278, July 2012.

[115] S. Ferradans, M. Bertalmío, E. Provenzi, and V. Caselles, "An Analysis of Visual Adaptation and Contrast Perception for Tone Mapping," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 2002–2012, Oct. 2011.

[116] S. Zamir, J. Vazquez-Corral, and M. Bertalmío, "Gamut Mapping in Cinematography Through Perceptually-Based Contrast Modification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 490–503, June 2014.

[117] P. Getreuer, "Automatic Color Enhancement (ACE) and its Fast Implementation," *Image Processing On Line*, vol. 2, pp. 266–277, Nov. 2012.

[118] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient Image Dehazing with Boundary Constraint and Contextual Regularization," in *2013 IEEE International Conference on Computer Vision (ICCV)*, pp. 617–624, Dec. 2013.

[119] J.-P. Tarel, N. Hautiere, A. Cord, D. Gruyer, and H. Halmaoui, "Improved visibility of road scene images under heterogeneous fog," in *2010 IEEE Intelligent Vehicles Symposium (IV)*, pp. 478–485, June 2010.

[120] Y. Lai, Y. Chen, C. Chiou, and C. Hsu, "Single-Image Dehazing via Optimal Transmission Map Under Scene Priors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, pp. 1–14, Jan. 2015.

[121] F. Bray, J.-S. Ren, E. Masuyer, and J. Ferlay, "Global estimates of cancer prevalence for 27 sites in the adult population in 2008," *International Journal of Cancer*, vol. 132, pp. 1133–1145, Mar. 2013.

[122] P. C. Gotzsche and M. Nielsen, "Screening for breast cancer with mammography," *The Cochrane Database of Systematic Reviews*, p. 1877, June 2011.

[123] L. M. Warren, R. M. Given-Wilson, M. G. Wallis, J. Cooke, M. D. Halling-Brown, A. Mackenzie, D. P. Chakraborty, H. Bosmans, D. R. Dance, and K. C. Young, "The effect of image processing on the detection of cancers in digital mammography," *AJR. American journal of roentgenology*, vol. 203, pp. 387–393, Aug. 2014.

[124] A. Oliver, J. Freixenet, J. Marti, E. Perez, J. Pont, E. R. E. Denton, and R. Zwiggelaar, "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, pp. 87–110, Apr. 2010.

[125] N. R. Mudigonda, R. M. Rangayyan, and J. E. Desautels, "Detection of breast masses in mammograms by density slicing and texture flow-field analysis," *IEEE Transactions on Medical Imaging*, vol. 20, pp. 1215–1227, Dec. 2001.

[126] T. Berber, A. Alpkocak, P. Balci, and O. Dicle, "Breast mass contour segmentation algorithm in digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 110, pp. 150–159, May 2013.

[127] M. G. Linguraru, K. Marias, R. English, and M. Brady, "A biologically inspired algorithm for microcalcification cluster detection," *Medical Image Analysis*, vol. 10, pp. 850–862, Dec. 2006.

[128] J. C. Fu, S. K. Lee, S. T. C. Wong, J. Y. Yeh, A. H. Wang, and H. K. Wu, "Image segmentation feature selection and pattern classification for mammographic microcalcifications," *Computerized Medical Imaging and Graphics*, vol. 29, pp. 419–429, Sept. 2005.

[129] S. Yu and L. Guan, "A CAD system for the automatic detection of clustered microcalcifications in digitized mammogram films," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 115–126, Feb. 2000.

[130] J. M. Lesniak, R. Hupse, R. Blanc, N. Karssemeijer, and G. Szekely, "Comparative evaluation of support vector machine classification for computer aided detection of breast masses in mammography," *Physics in Medicine and Biology*, vol. 57, pp. 5295–5307, Aug. 2012.

[131] G. Kom, A. Tiedeu, and M. Kom, "Automated detection of masses in mammograms by local adaptive thresholding," *Computers in Biology and Medicine*, vol. 37, pp. 37–48, Jan. 2007.

[132] G. Braz Junior, S. V. da Rocha, M. Gattass, A. C. Silva, and A. C. d. Paiva, "A mass classification using spatial diversity approaches in mammography images for false positive reduction," *Expert Systems with Applications*, vol. 40, pp. 7534–7543, Dec. 2013.

[133] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono, and E. Colini, "Breast Mass Segmentation in Mammographic Images by an Effective Region Growing Algorithm," No. 5259 in Lecture Notes in Computer Science, pp. 948–957, Springer Berlin Heidelberg, Jan. 2008.

[134] W. B. Sampaio, E. M. Diniz, A. C. Silva, A. C. de Paiva, and M. Gattass, "Detection of masses in mammogram images using CNN, geostatistic functions and SVM," *Computers in Biology and Medicine*, vol. 41, pp. 653–664, Aug. 2011.

[135] P. B. Ribeiro, R. A. F. Romero, P. R. Oliveira, H. Schiabel, and L. B. Vercosa, "Automatic segmentation of breast masses using enhanced ICA mixture model," *Neurocomputing*, vol. 120, pp. 61–71, Nov. 2013.

[136] P. Agrawal, M. Vatsa, and R. Singh, "Saliency based mass detection from screening mammograms," *Signal Processing*, vol. 99, pp. 29–47, June 2014.

[137] E. D. Pisano, S. Zong, B. M. Hemminger, M. DeLuca, R. E. Johnston, K. Muller, M. P. Braeuning, and S. M. Pizer, "Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms," *Journal of Digital Imaging*, vol. 11, pp. 193–200, Nov. 1998.

[138] Q. Abbas, M. E. Celebi, and I. F. Garcia, "Breast mass segmentation using region-based and edge-based methods in a 4-stage multiscale system," *Biomedical Signal Processing and Control*, vol. 8, pp. 204–214, Mar. 2013.

[139] B. Kurt, V. V. Nabiyev, and K. Turhan, "A novel automatic suspicious mass regions identification using Havrda & Charvat entropy and Otsu's N thresholding," *Computer Methods and Programs in Biomedicine*, vol. 114, pp. 349–360, May 2014.

[140] P. Rahmati, A. Adler, and G. Hamarneh, "Mammography segmentation with maximum likelihood active contours," *Medical Image Analysis*, vol. 16, pp. 1167–1186, Aug. 2012.

[141] A. P. Dhawan, G. Buelloni, and R. Gordon, "Enhancement of mammographic features by optimal adaptive neighborhood image processing," *IEEE Transactions on Medical Imaging*, vol. 5, pp. 8–15, Mar. 1986.

[142] R. L. Smathers, E. Bush, J. Drace, M. Stevens, F. G. Sommer, B. W. Brown, and B. Karras, "Mammographic microcalcifications: detection with xerography, screen-film, and digitized film display," *Radiology*, vol. 159, pp. 673–677, June 1986.

[143] H. P. Chan, C. J. Vyborny, H. MacMahon, C. E. Metz, K. Doi, and E. A. Sickles, "Digital mammography. ROC studies of the effects of pixel size and unsharp-mask filtering on the detection of subtle microcalcifications," *Investigative Radiology*, vol. 22, pp. 581–589, July 1987.

[144] M. McSweeney, P. Sprawls, and R. Egan, "Enhanced image mammography," *American Journal of Roentgenology*, vol. 140, pp. 9–14, Jan. 1983.

[145] A. Laine, S. Schuler, J. Fan, and W. Huda, "Mammographic feature enhancement by multiscale analysis," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 725–740, Dec. 1994.

[146] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, "Mammographic Images Enhancement and Denoising for Breast Cancer Detection Using Dyadic Wavelet Processing," *IEEE Transactions on Instrumentation and Measurement*, vol. 57, pp. 1422–1430, July 2008.

[147] P. Heinlein, J. Drexl, and W. Schneider, "Integrated wavelets for enhancement of microcalcifications in digital mammography," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 402–413, Mar. 2003.

[148] J. M. M. Munoz, H. d. J. O. Dominguez, O. O. V. Villegas, V. G. C. Sanchez, and L. O. Maynez, "The Nonsubsampled Contourlet Transform for Enhancement of Microcalcifications in Digital Mammograms," in *MICCAI 2009: Advances in Artificial Intelligence* (A. H. Aguirre, R. M. Borja, and C. A. R. Garcia, eds.), no. 5845 in Lecture Notes in Computer Science, pp. 292–302, Springer Berlin Heidelberg, Jan. 2009.

[149] S. Skiadopoulos, A. Karahaliou, F. Sakellaropoulos, G. Panayiotakis, and L. Costaridou, "Breast Component Adaptive Wavelet Enhancement for Soft-Copy Display of Mammograms," in *Digital Mammography* (S. M. Astley, M. Brady, C. Rose, and R. Zwiggelaar, eds.), no. 4046 in Lecture Notes in Computer Science, pp. 549–556, Springer Berlin Heidelberg, Jan. 2006.

[150] C. M. Chang and A. Laine, "Coherence of multiscale features for enhancement of digital mammograms," *IEEE Transactions on Information Technology in Biomedicine*, vol. 3, pp. 32–46, Mar. 1999.

[151] F. Y. M. Lure, P. W. Jones, and R. S. Gaborski, "Multiresolution unsharp masking technique for mammogram image enhancement," in *SPIE*, vol. 2710, pp. 830–839, Apr. 1996.

[152] J. Scharcanski and C. R. Jung, "Denoising and enhancing digital mammographic images for visual screening," *Computerized Medical Imaging and Graphics*, vol. 30, pp. 243–254, June 2006.

[153] D.-Y. Tsai, E. Matsuyama, and H.-M. Chen, "Improving Image Quality in Medical Images Using a Combined Method of Undecimated Wavelet Transform and Wavelet Coefficient Mapping," *International Journal of Biomedical Imaging*, vol. 2013, p. 11, Dec. 2013.

[154] P. Sakellaropoulos, L. Costaridou, and G. Panayiotakis, "A wavelet-based spatially adaptive method for mammographic contrast enhancement," *Physics in Medicine and Biology*, vol. 48, p. 787, Mar. 2003.

[155] J. Tang, X. Liu, and Q. Sun, "A Direct Image Contrast Enhancement Algorithm in the Wavelet Domain for Screening Mammograms," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, pp. 74–80, Feb. 2009.

[156] K. Ganesan, U. Acharya, C. Chua, L. Min, K. Abraham, and K. Ng, "Computer-Aided Breast Cancer Detection Using Mammograms: A Review," *IEEE Reviews in Biomedical Engineering,*, vol. 6, pp. 77–98, 2013.

[157] S. Dippel, M. Stahl, R. Wiemker, and T. Blaffert, "Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform," *IEEE Transactions on Medical Imaging*, vol. 21, pp. 343–353, Apr. 2002.

[158] A. Rojas Dominguez and A. K. Nandi, "Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection," *Computerized Medical Imaging and Graphics*, vol. 32, pp. 304–315, June 2008.

[159] G. te Brake and N. Karssemeijer, "Single and multiscale detection of masses in digital mammograms," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 628–639, July 1999.

[160] H. D. Cheng and H. Xu, "A novel fuzzy logic approach to mammogram contrast enhancement," *Information Sciences*, vol. 148, pp. 167–184, Dec. 2002.

[161] J. Jiang, B. Yao, and A. M. Wason, "Integration of fuzzy logic and structure tensor towards mammogram contrast enhancement," *Computerized Medical Imaging and Graphics*, vol. 29, pp. 83–90, Jan. 2005.

[162] N. Petrick, H. P. Chan, B. Sahiner, and D. Wei, "An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection," *IEEE Transactions on Medical Imaging*, vol. 15, pp. 59–67, Feb. 1996.

[163] L. C. S. Romualdo, M. A. C. Vieira, H. Schiabel, N. D. A. Mascarenhas, and L. R. Borges, "Mammographic image denoising and enhancement using the Anscombe transformation, adaptive wiener filtering, and the modulation transfer function," *Journal of Digital Imaging*, vol. 26, pp. 183–197, Apr. 2013.

[164] V. Bhateja and M. Misra, "A Robust Polynomial Filtering Framework for Mammographic Image Enhancement From Biomedical Sensors," *IEEESensors Journal,*, vol. 13, pp. 4147–4156, Nov. 2013.

[165] H. Kobatake and S. Hashimoto, "Convergence index filter for vector fields," *IEEE Transactions on Image Processing*, vol. 8, no. 8, pp. 1029–1038, 1999.

[166] H. Kobatake, M. Murakami, H. Takeo, and S. Nawano, "Computerized detection of malignant tumors on digital mammograms," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 369–378, May 1999.

[167] W. J. H. Veldkamp and N. Karssemeijer, "Normalization of local contrast in mammograms," *IEEE Transactions on Medical Imaging*, vol. 19, pp. 731–738, July 2000.

[168] K. J. McLoughlin, P. J. Bones, and N. Karssemeijer, "Noise equalization for detection of microcalcification clusters in direct digital mammogram images," *IEEE Transactions on Medical Imaging*, vol. 23, pp. 313–320, Mar. 2004.

[169] M. Adel, D. Zuwala, M. Rasigni, and S. Bourennane, "Filtering noise on mammographic phantom images using local contrast modification functions," *Image and Vision Computing*, vol. 26, pp. 1219–1229, Sept. 2008.

[170] R. Highnam and M. Brady, *Mammographic Image Analysis*. Springer Science & Business Media, Jan. 1999.

[171] M. Sundaram, K. Ramar, N. Arumugam, and G. Prabin, "Histogram Modified Local Contrast Enhancement for mammogram images," *Applied Soft Computing*, vol. 11, pp. 5809–5816, Dec. 2011.

[172] K. Panetta, Y. Zhou, S. Agaian, and H. Jia, "Nonlinear unsharp masking for mammogram enhancement," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, pp. 918–928, Nov. 2011.

[173] R. M. Rangayyan, F. J. Ayres, and J. E. Leo Desautels, "A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs," *Journal of the Franklin Institute*, vol. 344, pp. 312–348, May 2007.

[174] N. Karssemeijer and P. R. Snoeren, "Image Processing," in *Digital Mammography* (U. Bick and F. Diekmann, eds.), Medical Radiology, pp. 69–83, Springer Berlin Heidelberg, Jan. 2010.

[175] M. Bertalmío, "From image processing to computational neuroscience: a neural model based on histogram equalization," *Frontiers in Computational Neuroscience*, vol. 8, July 2014.

[176] S. Singh and K. Bovis, "An evaluation of contrast enhancement techniques for mammographic breast masses," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, pp. 109–119, Mar. 2005.

[177] F. Zanca, J. Jacobs, C. Van Ongeval, F. Claus, V. Celis, C. Geniets, V. Provost, H. Pauwels, G. Marchal, and H. Bosmans, "Evaluation of clinical image processing algorithms used in digital mammography," *Medical Physics*, vol. 36, pp. 765–775, Mar. 2009.

[178] K. Panetta, A. Samani, and S. Agaian, "Choosing the Optimal Spatial Domain Measure of Enhancement for Mammogram Images," *International Journal of Biomedical Imaging*, vol. 2014, p. 8, Aug. 2014.

[179] X. Hou, J. Harel, and C. Koch, "Image Signature: Highlighting Sparse Salient Regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, pp. 194–201, Jan. 2012.

[180] J. Li, M. Levine, X. An, X. Xu, and H. He, "Visual Saliency Based on Scale-Space Analysis in the Frequency Domain," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 996–1010, Apr. 2013.

[181] C. Ballester, V. Caselles, L. Igual, J. Verdera, and B. Rouge, "A Variational Model for P+XS Image Fusion," *International Journal of Computer Vision*, vol. 69, pp. 43–58, Apr. 2006.

[182] J. Duran, A. Buades, B. Coll, and C. Sbert, "A Nonlocal Variational Model for Pansharpening Image Fusion," *SIAM Journal on Imaging Sciences*, vol. 7, pp. 761–796, Jan. 2014.

[183] R. Highnam, J. Brady, and B. Shepstone, "Computing the scatter component of mammographic images," *IEEE Transactions on Medical Imaging*, vol. 13, pp. 301–313, June 1994.

[184] R. Highnam, M. Brady, and B. Shepstone, "A representation for mammographic image processing," *Medical Image Analysis*, vol. 1, pp. 1–18, Mar. 1996.

[185] S. van Engeland, P. Snoeren, H. Huisman, C. Boetes, and N. Karssemeijer, "Volumetric breast density estimation from full-field digital mammograms," *IEEE Transactions on Medical Imaging*, vol. 25, pp. 273–282, Mar. 2006.

[186] A. Oliver, A. Torrent, X. Llado, M. Tortajada, L. Tortajada, M. Sentis, J. Freixenet, and R. Zwiggelaar, "Automatic microcalcification and cluster detection for digital and digitised mammograms," *Knowledge-Based Systems*, vol. 28, pp. 68–75, Apr. 2012.

[187] E. D. Pisano, E. B. Cole, B. M. Hemminger, M. J. Yaffe, S. R. Aylward, A. D. A. Maidment, R. E. Johnston, M. B. Williams, L. T. Niklason, E. F. Conant, L. L. Fajardo, D. B. Kopans, M. E. Brown, and S. M. Pizer, "Image Processing Algorithms for Digital Mammography: A Pictorial Essay," *RadioGraphics*, vol. 20, pp. 1479–1491, Sept. 2000.

[188] A. Bria, N. Karssemeijer, and F. Tortorella, "Learning from unbalanced data: A cascade-based approach for detecting clustered microcalcifications," *Medical Image Analysis*, vol. 18, pp. 241–252, Feb. 2014.