CHAPTER 1

INTRODUCTION

Images are ubiquitous in the modern world. We depend on images for news, communication, and entertainment as well as for progress in medicine, science and technology. For better or worse, television images are virtually a sine qua non of modern life. Satellite images provide us with weather and crop information, and they provide our military commanders with timely and accurate information on troop movements. Biomedical research and material science could not proceed without microscopic images of many kinds. The petroleum reserves so essential to our economy are usually found through seismic images, and submarines are located with sonic imaging. These examples and many others that readily come to our mind are ample proof of the importance of imaging systems.

Image processing is a rapidly growing area of communication. Its growth has been fueled by technological advances in digital imaging, computer processors and mass storage devices. Fields which traditionally used analog imaging are now switching to digital systems, for their flexibility and affordability. Important examples are medicine, film and video production, photography, remote sensing, and security monitoring. These and other sources produce huge volumes of digital image data every day, more than could ever be examined manually.

Digital image processing is concerned primarily with extracting useful information from images. Ideally, this is done by computers, with little or no human intervention. Image processing algorithms may be placed at three levels **[Gonzalez 2002]**. At the lowest level are those techniques which deal directly with the raw, possibly noisy pixel values, with denoising and edge detection being good examples. In the middle are algorithms which utilize low level results for further means, such as segmentation and edge linking. At the highest level are those methods which attempt to extract semantic meaning from the information provided by the lower levels, for example, handwriting recognition.

Fig. 1.1 shows different stages of noise contamination in the image acquisition process. The different types of noises are contaminating the original image from image acquisition process to final stage of image application.

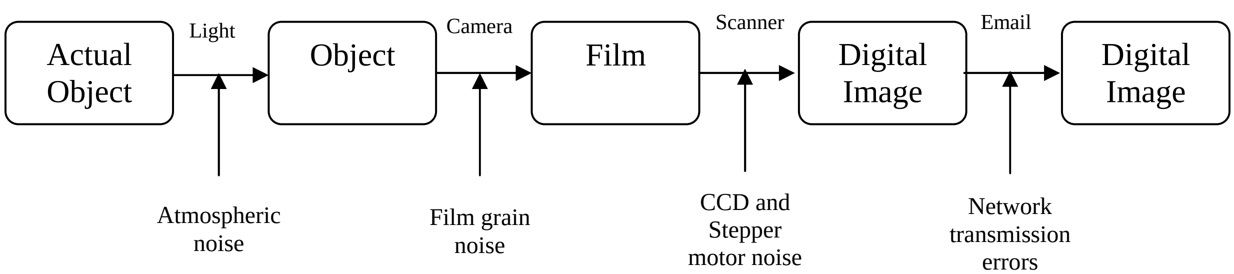

Fig. 1.1 Noise contaminations in the image acquisition process

Image denoising is the process of removing unwanted noise from an image. A denoised image is an approximation to the underlying true image, before it was contaminated. A good image denoising algorithm must simultaneously preserve structure and remove noise. Obviously, to do this the algorithm must be able to identify what structure is present. Local adoption specifically attempts to separate structure from noise on a local scale.

1.1 PRELIMINARIES

Most natural images have smooth intensity variations, with fine details being represented as sharp edges. The smooth variations in intensity are termed as low frequency variations and the sharp variations as high frequency variations. The low frequency components (smooth variations) constitute the base of an image, and the high frequency components (the edges which give the detail) add upon them to refine the image. Hence, the smooth variations demand more importance than details as most of the image energy is in low frequency region **[Mallat 1996]**.

1.1.1 Kinds of Imaging Systems

There are many kinds of objects to be imaged and many mechanism of image formation. Consequently, there are many ways in which imagery system can be classified. One such taxonomy, represented in Table 1.1 **[Wu 2006]**, which classifies systems by kind of radiation or field used to form an image. The most familiar kind of radiation is electromagnetic, including visible light, infrared and ultraviolet radiation. Also under this category we find long-wavelength radiation such as microwaves and radio waves and short-wavelength radiation in the extreme ultraviolet and soft x-ray.

Table 1.1 Classification by kind of radiation or field

Electromagnetic Waves	Other waves	Particles	Quasi-static fields
Radio waves Microwaves Infrared Visible light Ultraviolet Soft x-rays	Seismic waves Water waves Ultrasound DeBroglie waves	Neutrons Protons Heavy ions Hard x-rays Gamma rays	Geomagnetic fields Biomagnetic fields Bioelectric fields Electrical Impedance

A second useful taxonomy of imaging system groups them according to the imaging mechanism represented by Table 1.2 **[Wu 2006]**. They are classified as passive and active imaging. In passive imaging, measurements are made without interacting with a source. Familiar examples include ordinary photography of self-illumination sources or of a reflecting source with natural illumination as well as astronomical imaging and medical thermography. By contrast, an active imaging system supplies the radiation being imaged. Systems in this category include flash photography, transmission imaging (x-rays, microcopy, etc.), radar, active SONAR and medical ultrasound.

Table 1.2 Classification by kind of imaging mechanism

Passive systems	Active systems
Fluorescent microscopy Nuclear medicine Lunar imaging with telescope IR thermography Seismology Natural-light photography Biomagnetic imaging	Conventional transmission microscopy Diagnostic radiology Radar ranging of the moon Photoacoustic imaging Geophysical imaging with explosives Magnetic resonance imaging

1.1.2 Degradation of Images

There is various types are degradations in image acquisition and processing. The degradations can be classified as follows **[Sonka 1999]**:

• Spatial degradations, which can be due to the non-zero dimensions of picture elements used in the imaging device, or due to lens errors, defocus, bandwidth limitations in electronics. The effects are usually referred to as blur, unsharpness, ringing, echo, etc.

• Temporal degradations, which are caused by non-zero exposure time of the photosensitive material of the imaging device, by long decay times of light emitting materials, e.g., phosphors in displays. These degradations include motion blur, and temporal flickering artifacts.

• Geometrical degradations, which distorts geometry of the displayed pictures. These degradations can be due to lens aberrations, deflection non-linearity in camera.

• Point-wise degradations, which modifies the grey-level of image elements. These degradations are referred as noise, which is often some kind of random process, e.g., thermal noise, quantum noise, etc; depending on the source of degradation.

In this thesis we focus on point-wise degradations, i.e., we consider noise in image. Images are often distorted by noise during acquisition, recording and transmission. Certain noise sources are located in the camera (acquisition noise) and become amplified under poor lighting conditions, e.g., night vision surveillance cameras. Other noise sources are due to transmission over analogue channels, e.g., satellite or terrestrial broadcasting. Although it is true that digital transmission may avoid introduction of transmission-interference artifacts, there are sound economic reasons to constrain the channel bandwidth to a level where degradations are inevitable, consequently introducing (structured) noise in digital channels. Even in digital camera and output (display) are still most often analogue devices. Hence, noise can be introduced into the analog signal: (i) in image acquiring, before it reaches the digital capturing device or (ii) in image display, after the conversion from the digital. In many image applications the noise can be well approximated by the additive white Gaussian model, which we consider in this thesis.

The main applications for the proposed image denoising schemes, in this thesis, are aimed at visual improvement of natural images. In particular noise suppression in image is gaining importance due to the increasing number of applications where noise is a problem (e.g., infrared images, acquisition by cheap security camera under poor lighting conditions). Although, in these applications noise is not always white Gaussian but rather correlated (in general), in most cases, it can be approximated as the additive white Gaussian (at least locally) and can still be sufficiently well removed. Additionally, the proposed image denoising algorithms can be used as a pre-processing step for further analysis of SAR images.

1.2 NOISE MODELS

Noise in an image may be present due to imperfect capturing device, inadequate or non uniform lighting, or an undesirable view angle. Noise is generally modeled as additive noise. The addition of noise to an image can be represented by following equation:

$$f' = f + n \tag{1.1}$$

where f' is resultant noisy image.
f is original image and
n is the added noise to the image.

Depending on the nature of the image and the environment it was acquired, the distribution of the noise can vary. In most of the cases, the distribution is considered as a Gaussian **[Bosdogianni 1999]**. In this thesis, the additive white Gaussian noise is considered in natural images. More especially the noise is assumed to be an additive wide-sense stationary (WSS) white Gaussian noise (AWGN) process with zero mean and constant variance, which is formed independently of the original noise-free image.

The speckle noise is considered for SAR image which is multiplicative in nature. For denoising SAR images, the image is transformed into logarithmic scale then the noise is considered as additive. The rest of the processing steps are as additive Gaussian noise.

In many image denoising algorithms it is useful or even essential to obtain an estimate of the noise level, for the automatic and the optimal adaptation of the method to varying noise levels present in image. In this thesis we therefore investigate different approaches for noise estimation.

1.3 Denoising Techniques

Image denoising is the process of removing unwanted noise from an image. The noise can take a variety of forms and is introduced in differing amounts at each step during the acquisition of the image. The literature is abundant in denoising techniques, but they may be classified broadly into temporal, spatial, and transform domain approaches:

1.3.1 Temporal Domain

Temporal filtering averages multiple sampled versions of the same scene. If the true image is f_i and N samples of image $f_1, f_2, f_3 .... f_N$ are taken then the temporally filtered image $\hat{f}$ can be calculated as:

$$\hat{f} = \frac{1}{N}\sum_{i=1}^{N} f_i \tag{1.2}$$

If each pixel in f_i is corrupted by the addition of symmetric zero-mean noise of variance σ_n^2, then for any one pixel the expected noise variance, $E[(f_i - \hat{f})^2]$ is σ_n^2. However, the expected noise for the ensemble average reduces to σ_n^2 / N. If the noise is not additive in nature, for example impulse noise, then a better averaging function, such as the median, should be used. Temporal filtering is ideal if multiple version of the same image can be obtained.

1.3.2 Spatial Domain

The temporal filtering uses multiple versions of each pixel value at the same position with different times but the spatial filtering uses pixels at the same time with different spatial coordinates in the image. The pixels are expected to have the same noiseless value at the same position but different times in temporal filtering. When those pixels are not available, the best alternative is pixels near the current pixel. The spatial filtering is the only option for single noisy image. This is the basis of early image denoising algorithms.

The simplest spatial filter is the averaging filter which replaces a pixel with the average of itself and the pixels in the local neighborhood. The most common averaging filter is the box filter, which gives equal weight to all the nine pixels in a 3 x 3 window. This choice of weights maximally reduces the noise variance when the noise is additive. Box filtering is similar to temporal filtering except that samples from adjacent pixel positions are used as approximations to multiple samples of the pixel at the same position.

Although an averaging filter performs well for additive noise in homogeneous regions, it tends to blur edges and other image structure in heterogeneous regions. To combat this deficiency, most research is concerned with structure preserving denoising algorithms.

To preserve image structure while removing noise implies the ability to distinguish between the two. It will be shown in later chapters that neighboring pixels is a good basis to differentiate between structure and noise in an image. Features correspond to homogeneous regions and boundaries between segments are structure that needs to be preserved. Existing spatial denoising algorithms can be examined in terms of the way they use neighboring pixels.

1.3.3 Transform Domain

Image transformation is performed on an image and the resulting image may have properties which make it more suited to a particular purpose than the original.

Transforms are widely used in image denoising, image compression, description, etc; they were actively studied at the end of the 1960s and in the beginning of the 1970s. Image transform theory is well-known area characterized by a precise mathematical background and represent powerful, unified area of image processing. There are many books **[Qian 1996]**, **[Strang 1996]**, **[Gershenfeld 1999]**, **[Milman 1999]**, **[Percival 2000]**, **[Keinert 2004]**, **[Michel 2003]**, **[Sayood 2006]**, and published papers **[Rioul 1991]**, **[Shensa 1992]**, **[Habibi 1995]**, **[Bruce 2001]**, **[Nason 2002]**, **[Fotiou 2003]**, **[Unser 2003]**, **[Addition 2004]** with extensive coverage of image transforms. There are many transform for image processing applications such as discrete cosine transform, wavelet transform, fast fourier transform, etc **[Bosdogianni 1999]**. It is observed that wavelet transform performance is better than other transforms in general **[Balster 2003]**.

1.4 ASSESSMENT OF IMAGE QUALITY

The original image and image with additive white Gaussian noise are shown in Fig. 1.2. There are two different situations that may occur when measuring the quality of a denoised image: i) when the original image is available, and ii) when it is unknown. The first case usually occurs in an experimental situation where a known image is artificially corrupted with noise. The original image is ground truth to which any denoised images may be compared directly. The second case is more realistic situation whereby a noisy image has been sourced, say remotely from SAR system, and one wish to denoise it before further processing. Here there is no ground truth with which to compare. In either case, there are subjective and objective techniques for assessing the quality of denoised images. In most of the experiments performed in this thesis, an original image f has been contaminated by controlled amounts of noise n it to produce a noisy version f'. A denois-

(a) (b)
Fig. 1.2 (a) Original Image of Lena (b) Noisy Image

-ing algorithm produces an estimate $\hat{f}$ of the original image. In denoising process it is desired that only noise, and not image information, be removed. Thus the aim is to produce $\hat{f}$ as close as possible to f . An example is given in Fig. 1.3. The subjective and objective measures are explained in next section.

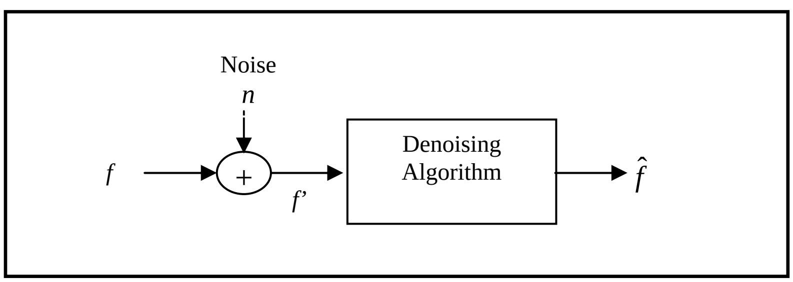

Fig. 1.3 Noise Contamination Model

1.4.1 Subjective Measures

Visual inspection is a subjective process whereby a human viewer assesses the quality of an image. In the case where ground truth is available, one or more assessors may perform a side by side comparison of the denoised and original images. The comparison is usually rated using predefined quality classes, such as “excellent”, “fine”, “passable”, “marginal”, “inferior” and “unusable” **[Wang 2002]**. This type of experiment requires that the viewing conditions be strictly controlled. This includes factors like ambient lighting levels, the display hardware, and the position of the assessor relative to the displayed images.

1.4.2 Objective Measures

An objective measure for the quality of a denoised image would be very useful. One traditional measure for the closeness of two data sets is the mean squared error (*MSE*) **[Wu 2006]**. When ground truth is available, the two data sets could be the denoised image $\hat{f}$, and the original image f . The *MSE* is calculated as:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\hat{f}(i,j) - f(i,j)\right]^2 \qquad (1.3)$$

where MN is size of image.

The *MSE* is proportional to the disparity between two images. In the case of two equivalent images, it is zero. For the case of additive zero-mean Gaussian noise, the *MSE* between the noisy and original images is exactly equal to the noise standard deviation.

The peak signal to noise ratio (*PSNR*), is derived from *MSE*, and is measured in decibels (dB). This logarithmic measure is computed as:

$$PSNR(dB) = 10\log_{10}\left(\frac{R^2}{MSE}\right) \qquad (1.4)$$

where R is the maximum possible pixel intensity. When the pixel intensity is an r -bit discrete variable, the signal energy can be replaced by the maximum input symbol energy $(2^r\text{-}1)^2$. For the common case of 8 bits per pixel (bpp) elements of input image, $R = 2^8 - 1 = 255$.

MSE and *PSNR* use the square of the pixel difference. These measures penalize large errors heavily. For example, a single pixel in error by 100 will have the same contribution as 10,000 pixels in error by 1. An alternative measure, which aims to alleviate this potential problem, is the mean absolute error (*MAE*), calculated using equation (1.5). The *MAE* penalizes errors by their magnitude, and is less likely, compared to *MSE*, to be biased by occasional large errors. The *MAE* is calculated as:

$$MAE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left| f(i,j) - \hat{f}(i,j) \right| \tag{1.5}$$

The *MSE* and *MAE* are useful in that they provide a number which can be compared objectively. Their drawback is that they do not take into account the spatial distribution of the pixel differences. Many small differences may be more tolerable than fewer larger errors, especially if those errors occur at “busy” locations in the image. In fact, this perceptual masking effect is exploited by lossy image and audio compression algorithms. Large errors clumped together, or near the image margins, may be preferred to the same errors spread around the image. There is evidence to suggest that MSE may be well correlated to human observer’s subjective opinions. This fact, combined with its simple formula, has allowed PSNR to be used widely throughout the literature.

The worst case absolute error (*WCAE*) is the magnitude of the single largest difference between two images. It provides a measure of a denoising algorithm's worst case performance. The calculation of the *WCAE* is given as:

$$WCAE = \arg\max_{(i,j)}\left| f(i,j) - \hat{f}(i,j) \right| \tag{1.6}$$

In an experimental situation, artificial noise is added to an original image to create a noisy version thereof. Difficulties can occur if the original image already contains some noise. A good denoising algorithm is likely to remove both the original and artificial noise components, producing a denoised image which is less affected by noise than the original image itself. In this situation one can not expect the denoised image to be the same as the already noisy original. The original noise level is usually low compared to the artificial noise being added, and can be safely ignored.

Equivalent Number of Looks (*ENL*) **[Chan 2003]** is used to measure speckle level in a SAR image over a uniform image region. A large value of *ENL* reflects the better quantitative performance of filter. The value of *ENL* depends on the size of the tested region. The equivalent number of looks is given as:

$$ENL = \frac{\mu^2}{\sigma^2} \qquad (1.7)$$

where μ and σ^2 are mean value, and variance of test region respectively.

1.4.3 Complexity

Complexity is measured both by the arithmetic processing required by an algorithm (typically measured in mips (million instructions per second) and by its memory requirements (measured typically in kilobytes of ROM or RAM). The use of millions instruction per second (mips)or millions flip-flops (mflops) as complexity measure is particularly appropriate for implementations on general purpose DSPs or computers. In application-specific integrated circuits (ASICs), other metrics of complexity (such as number of transistors or gates) can be useful. Complexity is an important parameter of performance mainly for two reasons - the need to minimize cost and the requirement to minimize power dissipation. In this thesis, it is calculated on the basis of the number of basic operations. In applications such as real time processing, it is particularly important to minimize complexity. In off line processing, complexity is a relatively less important issue.

1.5 Objectives and Approach

This thesis addresses denoising of images. In general, image denoising imposes a compromise between noise reduction and preserving significant image details. In our methods we assume white Gaussian noise, which is considered as a good noise model in many applications. We propose some emerging trends in wavelet domain denoising methods. To achieve a good performance in this respect, a denoising algorithm should adapt to image features. The features are detected in form of neighboring pixel's energy and variance. The wavelet representation naturally facilitates the construction of such spatially adaptive algorithms. It compresses the essential information in a signal into relatively few, large coefficients, which represent image details at different resolution scales. In practical applications, we often simplify the theory using heuristics, when this leads to algorithms with lower complexity or higher performance. Then the algorithm is optimized by differential evaluation technique to improve performance. We do not propose a universal denoising method, but rather proposed algorithms suitable for diverse applications.

The remote sensing images are considered for geographical feature extraction. In remote sensing image applications the aim is to enhance specific features. It is affected by speckle noise. The noise is multiplicative in nature. The pre-processing step is required to convert multiplicative noise to additive noise then the process remains same as the additive noise.

1.6 ORGANIZATION

A brief introduction to imaging systems, classification, noise models, objective of this research, and definitions of various terms are given in this introductory chapter.

In chapter 2, detailed literature survey has been presented on the denoising techniques that apply various approaches such as spatial domain, frequency domain and wavelet domain to investigate denoising problem.

Chapter 3 contains a detailed study of noise level estimation and gives a thorough review of the existing approaches for noise estimation in images. Subsequently, the spatial gradient-based noise estimation method has been described. The results of noise estimation methods are evaluated in terms of the noise estimate accuracy varying noise levels.

Chapter 4 includes a detailed description and implementation of various wavelet based image denoising methods. These wavelet image denoising methods has been applied for the comparison purpose in later chapters.

Chapters 5 and 6, discuss the proposed methods, the detailed investigation of existing wavelet methods for the purpose of image denoising has been presented and compared with proposed methods.

In chapter 7, image enhancement techniques have been discussed and one algorithm has been implemented based on regularity parameter.

Finally, this thesis is concluded in chapter 8, where a summary of the undertaken work and findings are presented. The scope for future work and investigations are presented in chapter 9. Following the chapter 9, Appendix A and the list of references is appended.