Chapter 1

Introduction

There is a Chinese proverb that 'a picture is worth a thousand words'. A digital image is a representation of visual information in a form that can be processed and manipulated by electronic systems, such as computers. It is composed of a grid of picture elements, commonly known as pixels. Each pixel contains information about color and brightness, and when combined, they create the overall visual representation of an image. Digital images can take various forms, including photographs, graphics, illustrations, and more. They can be created using digital cameras, scanners, or generated through computer graphics software. The most common formats for digital images include JPEG, PNG, GIF, and TIFF. Digital images can be edited, resized, and manipulated using image editing software, enabling a wide range of creative possibilities. They are widely used in various fields, such as photography, graphic design, web design, medical imaging, remote sensing images, text images, and scientific research. The ability to store, share, and transmit digital images has greatly facilitated communication and information exchange in the modern digital age.

Image processing encompasses the manipulation and analysis of digital images through various algorithms and techniques. This field involves enhancing, transforming, and extracting information from images to either improve their quality or derive significant insights. Image processing finds applications in numerous areas such as computer vision, medical imaging, remote sensing, and multimedia. This is a multidisciplinary field that combines aspects of computer science, mathematics, and engineering. It plays a crucial role in various applications, from medical diagnostics and satellite imagery analysis to facial recognition and entertainment industry graphics.

Image analysis is a broader term that encompasses the extraction of meaningful information or knowledge from digital images through computational methods. It involves the application of various techniques and algorithms to analyze and interpret visual data. Image analysis can be used in a wide range of fields, including medicine, biology, astronomy, geology, remote sensing, and computer vision. Image analysis often involves a combination of image processing techniques, machine learning, and statistical methods to derive meaningful insights from visual data. It plays a crucial role in automating tasks that would be challenging or time-consuming for humans to perform manually and contributes to advancements in various scientific and technological fields.

Image filtering is a fundamental step in image processing and analysis, often serving as a crucial pre- or post-processing stage. It is essential across all image processing and analysis applications, laying the groundwork for enhancing image quality, reducing noise, and preparing images for further analysis. The effectiveness of image processing and analysis is heavily dependent on the quality of the output image. High-quality images provide a clear foundation for accurate feature detection and interpretation, enabling tasks like object detection, segmentation, and pattern recognition to perform effectively. Clear processing ensures essential features such as edges and textures are preserved, while reducing noise and artifacts minimizes misinterpretation and false results. Additionally, high-quality images improve analysis efficiency by reducing the need for complex corrective algorithms. Overall, well-processed images are critical for achieving accurate, reliable, and efficient outcomes in image analysis.

1.1 Image filtering

Image filtering techniques serve the crucial purpose of enhancing the quality of images. The presence of geometric, radiometric, and sensor-related effects in digital images captured by sensors often introduces noise and distortion. The effectiveness of various image processing and analysis methods significantly relies on the quality of the input and output images. Consequently, image filtering becomes a necessary and integral step in the majority of image processing and analysis applications.

Two primary approaches to image filtering techniques are spatial and frequency domain filtering. Spatial domain filtering is more intuitive as it directly operates on image pixels, while frequency domain filtering works on signals transformed from digital images. Spatial domain image filtering techniques have evolved from simple linear filters to advanced non-linear ones, including edge-preserving filters, adaptive filters, region-based structure preserving filters, and morphological filters. In recent years, developments have shifted from gradient-based edge-preserving filtering to adaptive window-based edge preserving and region statistics-based structure preserving filtering. These structure preserving filtering techniques aim not only to protect individual edges or lines but also to safeguard meaningful structures while eliminating insignificant ones. In contemporary times, semantic-aware structure preserving filtering has emerged as a concept. Defining the semantically meaningful structures within an image remains one of the most challenging aspects of developing such filtering techniques.

The limitations of recent gradient-based filtering methods are that they often rely on directional (horizontal and vertical) image gradients and certain region based statistical measures of them for structure/texture descriptors. This approach frequently fails to effectively preserve the corner portions of objects. Also, the measures of gradients are taken within a fixed-size neighbor window. That is why, to achieve optimal results with these methods, a substantial amount

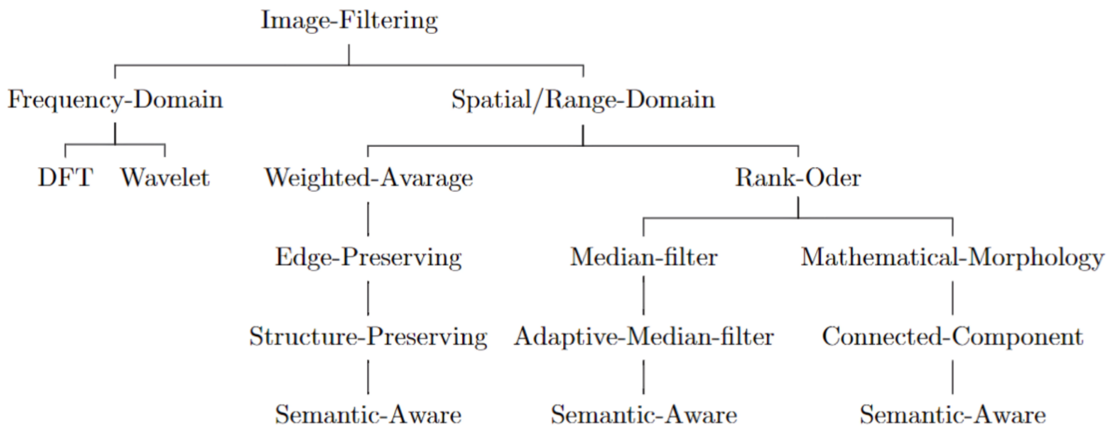

Figure 1-1: Evolution of image filtering

of empirical study is needed to select appropriate neighbor window sizes, region window sizes, or kernel function parameters.

In contrast, morphological filters have the ability to preserve the structure of objects without relying on complex statistical measures. However, their effectiveness is heavily dependent on the size and shape of the structuring element. Connected component filtering, an advanced form of morphological filtering, can also be applied to preserve object structures in images. Nevertheless, it relies on component ordering and threshold values, making it less robust for preserving meaningful image structures.

Figure 1-1 is showing the evolution of all these different conventional spatial and range domain filtering techniques towards more sophisticated and meaningful semantic-aware image filtering approaches.

1.1.1 Weighted average filtering

Image filtering started its journey with simple weighted average filtering. Weighted average filtering [58, 106, 130] in spatial domain mainly takes the information of predefined neighbor pixels to update the centre pixel. As it averages the pixel values, it smooths the images and removes the small noises effectively, but as it operates uniformly throughout the whole image, it can destroy the weak edges and weaken the sharp edges. Weighted average filtering is employed to enhance or denoise images by assigning different weights to neighboring pixels during the filtering process. The basic concept behind weighted average filtering is to replace the intensity value of a pixel with a weighted average of its neighboring pixels. The weights are typically determined by a predefined kernel or filter mask.

A brief step-by-step explanation of the weighted average filtering process is as follows:

- Selection of a kernel/filter mask: A kernel, or filter mask, is a small matrix that defines the weights assigned to neighboring pixels. Commonly used filter masks include the box filter, where all neighboring pixels have equal weights, and the Gaussian filter, which uses a Gaussian distribution over its distance from the central pixel to assign weights.
- Centering on a pixel: The filter mask is centered on each pixel in the image, and the weights are applied to the corresponding neighboring pixels.
- Weighted sum calculation: Multiply the intensity values of each neighboring pixel by its corresponding weight. Sum up the weighted values to get the weighted sum.
- Normalization: Divide the weighted sum by the sum of the weights in the filter mask to normalize the result.
- Replacement of pixel value: Replace the original intensity value of the central pixel with the normalized weighted sum.

The weighted average filtering process effectively reduces noise in the image by smoothing out variations caused by random fluctuations. The choice of the filter mask and its size directly impact the filtering results. Smaller filter sizes preserve finer details but may not eliminate as much noise, while larger filter sizes provide stronger noise reduction but may blur important features. Also, the weight's spatial distribution determines the filtering effects. If the distribution is uniform or flat spatially, then it smooths more, otherwise if the distribution has a sharp peak, then smooth less.

Despite its effectiveness in noise reduction, weighted average filtering has limitations. It can not perform well in scenarios where there are both sharp and weak edges or boundaries in the image, as it treats both edges equally. Therefore, it is crucial to choose the parameters, like the appropriate filter mask and size, based on the characteristics of the image and the desired outcome.

1.1.2 Edge-aware/ Edge preserving filtering

Initial developments in edge preserving image filtering techniques [106, 130] were mainly developed on the assumptions that prominent edges are with comparatively high gradients than noises/textures. They employed the range/intensity difference or gradient information into weight designing function. This made the filtering edge or high gradient aware. Also these methods were applied a uniform operation with fixed parameter (like filtering window/mask size, kernel parameter) through out the whole images. But images having insignificant textures/noises with high gradient values can not be smoothed out with such uniform operation. So most of the gradient based uniform edge preserving image filtering techniques fails for such kind of images. Therefore in later development adaptive window based edge preserving image filtering techniques were proposed in the later developments. But only the gradient based features are not sufficient for developing robust filtering technique for different types of images and for different types of applications also. That is why, instead of several good edge/structure preserving filtering techniques exist in the literature, still there is a scope to develop semantically significant feature extraction to develop semantic filtering techniques. In many applications, such as medical or satellite image analysis, edges are crucial features that must be preserved sharply and without distortion during smoothing or denoising processes.

Initial edge-preserving filters are designed to automatically limit smoothing at structural edges, which are identified by high gradient magnitudes. A weighted average based premier non-linear technique is bilateral filter [130], that is initially designed for edge-preserving and noise reducing smoothing. It defines the weights by combining Gaussian distributions of both spatial and spectral variations such that the weights are inversely proportional to both the spatial distance and spectral difference (i.e the range difference or color difference or depth distance). This key idea take the crucial role in preserving sharp edges. The basic formulation of the bilateral filter is follows:

$$J_p = \frac{1}{W_p} \sum_{q \varepsilon \Omega_p} G_{\sigma_s}(\|q - p\|) G_{\sigma_r}(\|I_q - I_p\|) I_q. \tag{1.1}$$

Where I is the original and J is the corresponding filtered image, G_{σ_s} is the standard Gaussian function over spatial domain with standard deviation σ_s and G_{σ_r} is over range domain of σ_r respectively. It is a non-linear combination of spatial and range filters, where the spatial part smooths the images and the range part used to detect the edges not to be smoothed. In bilateral filtering smoothing and preserving edges is totally depends on the parameter σ_s and σ_r, respectively. Thus, a huge amount of empirical study is needed for different types of images to find the possible values of σ_s and σ_r to get best filtered results.

Anisotropic diffusion is also a premier technique used in image processing and computer vision for smoothing images while preserving and enhancing important structures, such as edges. Unlike isotropic diffusion, which smooths uniformly in all directions, anisotropic diffusion adapts the diffusion process based on local image characteristics. Based on the gradient magnitude of the image the diffusion coefficient is modulated in anisotropic diffusion. The basic idea is to allow diffusion to occur more along regions with weaker intensity gradients (smoother areas

) while limiting diffusion across regions with stronger intensity gradients (edges or boundaries). The preservation of edges and fine details in an image is elevated by this adaptive behavior.

Perona-Malik diffusion is another popular term used for anisotrofic diffusion by its developer names, is become a widely utilised approach for noise removing while preserving significant image contents. It works similarly through the process of scale space creation, by generating a family of smoothed images based on a isotropic diffusion process with a increasing set parameters. Isotropic diffusion basically operates like a convolution of the image with a 2D isotropic Gaussian filter with a increasing width depending on the parameter. This transform the original image by a space-variant linear diffusion, which smooths uniformly in all directions. Anisotropic diffusion adapts the smoothing process based on the local image structure by controlling the diffusion process with a gradient-based function. Consequently, it become a space-variant non-linear transformation of the original image.

The original formulation in [106], is also termed as shape-adapted smoothing or coherence-enhancing diffusion. It simply generalized the usual diffusion equation, in which the constant scalar diffusion coefficient replaced with a function of pixels position. This is designed by a partial differential equations (PDEs) to estimate pixel-wise spatially varying diffusivities. It reduce smoothing at image edges so that preserve important image structures while smoothing out noise or textures. The image is iteratively updated according to the partial differential equation, which governs the diffusion process. The PDE is designed to reduce diffusion in regions with high gradient magnitudes (i.e., at edges) and allow more diffusion in relatively uniform areas. The result is an image with reduced noise and preserved edges, making anisotropic diffusion particularly useful for tasks such as edge detection, image segmentation, and feature extraction. The equation is designed as follows:

$$\frac{\partial I}{\partial t} = \operatorname{div}\left(c(x,y,t)\nabla I\right) = \nabla c \cdot \nabla I + c(x,y,t)\Delta I \tag{1.2}$$

In this context, *div( )* stands for the divergence operator, Δ is the *Laplacian*, ∇ for the *gradient*, and $c(x,y,t)$ represents the diffusion coefficient, which regulates the diffusion rate. Typically, this coefficient is taken as a function of the image gradient to maintain edge integrity in the image. In Perona and Malik [106] two functions are proposed as the diffusion coefficient:

$c\left(\|\nabla I\|\right) = e^{-\left(\|\nabla I\|/K\right)^2}$ and

$c\left(\|\nabla I\|\right) = \dfrac{1}{1+\left(\frac{\|\nabla I\|}{K}\right)^2}$ The constant K determines the sensitivity to edges and is typically selected either through experimentation or depending on the level of the noise in the image.

1.1.3 Mathematical morphology

Mathematical Morphology (MM) [122] proposes a theoretical framework and a set of methodologies for analyzing and manipulating geometric structures. It initially incorporates principles from set theory then lattice theory, topology and random functions. Although MM is primarily used for digital images, it is also applicable to the other spatial structures like surface meshes, solids, and graphs. MM introduces geometrical and topological concepts in both continuous and discrete spaces, including shape, size, geodesic distance, connectivity, and convexity. These concepts are essential to morphological image processing, that uses a series of operators to transform images based on these characteristics.

The basic morphological operators in digital image processing are— dilation, erosion, opening, and closing—play a pivotal role. In contrast to weighted average filtering, morphological operator does not introduce new pixel values during the filtering process. Morphological operators integrate image objects with their background/foreground, determined by structuring elements. Notably, it excels in preserving the fundamental shape of image objects while eliminating spurious or extended edges.

Initially designed for binary images, MM has further evolved to operate on grayscale images or functions. The theoretical foundation of MM is now established by the generalization of it to complete lattices. Morphological filters are notable for preserving object structures without depending on statistical measures, making MM a valuable tool in many applications of image analysis and processing. However, the structuring element's size and shape highly determines the effectiveness of MM.

1.1.3.1 Morphological filter

Morphological filters are image processing techniques that operate on the structure, or morphology, of an image. Computer vision and image processing widely use it to enhance or suppress certain features of an image. These filters are based on mathematical morphology, which deals with the shape and structure of objects. Dilation and erosion are two fundamental operations in morphological filtering. In an image, expansion of the objects boundaries is accomplished through dilation, while erosion involves shrinking them. These operations are performed using a structuring element, which is a small, predefined pattern or shape to determine the neighbor window. The structuring element is moved through each pixel of the image, and the designed interaction determines the result of the morphological operation.

The operational foundation of mathematical morphology relies on the rank-ordering principle. The fundamental operations, dilation and erosion, filter an image by substituting the center pixel of a predefined neighboring region known as the structuring element (SE) with the local maximum and minimum values, respectively. Additional operations like opening and closing are constructed by sequentially applying alternating dilation and erosion. These operations intuitively eliminate impulse noises smaller than the structural element without distorting the fundamental shape of the objects. In essence, mathematical morphology utilizes rank-based comparisons within the structuring element to modify and analyze the structure of objects in an image.

Dilation (δ) : Dilation is designed to expand the brighter objects in an image. It is achieved by simply placing the structuring element at each pixel and setting the considered pixel value in the filtered image to the maximum pixel value within the structuring element. This operation is useful for tasks such as joining broken parts of an object, filling in gaps, and making objects more prominent.

Erosion (ε): Erosion, on the other hand, is used to shrink the brighter objects. It involves placing the structuring element at each pixel and setting the considered pixel in the filtered image to the minimum pixel value within the structuring element. Erosion is helpful in tasks like separating objects that are close to each other, removing small details, and smoothing and sharpening object boundaries.

Opening and closing are two other basic compound operations that combine dilation and erosion.

Opening ($\delta(\varepsilon)$): It perform an erosion operation following to a dilation. This helps in removing small objects and fine details.

Closing ($\varepsilon(\delta)$): It involves applying an erosion operation prior to a dilation. Closing is effective in closing small gaps and connecting nearby objects.

Morphological filters are utilized in a variety of image processing tasks, including noise reduction, edge detection, feature extraction and image segmentation. These are particularly useful in scenarios where the object shapes and sizes in the image are critical for analysis.

It's worth noting that the selection of the structuring element is crucial for the effectiveness of morphological filtering. The structuring element's size and shape need to be selected by analysing the attributes of the objects in the input image and the desired outcome of the filtering operation.

Morphological gradient ($\delta - \varepsilon$): Morphological gradient is a morphological operation used in image processing to highlight the boundaries or edges of objects within an image. It is particularly useful for emphasizing transitions between regions of different intensity or texture. The morphological gradient is obtained by computing the difference of the dilation to erosion of an image. It is a powerful tool in image processing, and its application can contribute to various tasks, including image analysis, computer vision, and pattern recognition. It provides a way to enhance and emphasize the edges or transitions in an image, making it valuable for feature extraction and subsequent processing steps.

Morphological texture filtering ($\delta(\varepsilon) + \varepsilon(\delta)$)/2: Morphological texture filtering is a technique used in image processing to analyze and enhance texture information within an image using morphological operations. Morphological operations involve the manipulation of image structures based on the shape and gray scale characteristic of the elements in the image. The opening operator can remove the darker noises/textures and the closing operator can remove the brighter noises and textures. Hence, the combination of these two reduces the sharp oscillation of noises and textures while structural edges remain as they are. The average of opening and closing is thus useful as an initial texture structure discriminator.

In this research work, morphological filters are exploited to improve the discrimination between structural edges and textural regions of the image for further semantic texture structure decomposition.

1.1.4 Median filter

A median filter [7, 60, 64] in signal processing and digital image processing is a non-linear filtering technique applied to reduce noise and remove outliers. Unlike linear filters, such as mean or Gaussian filters, that compute the average or weighted average of pixel values of neighbor pixels, in median filter the median value of the neighbor pixels is computed and replace the centre pixel by it. The median filter operates by sliding a window (also known as a kernel or mask) over the image, and for each position, the pixel value at that position is replaced by the median value of the pixels within the window. The size of the window determines the neighboring pixels considered for computing the median. The key advantage of the median filter is its ability to effectively reduce impulse noise, also known as salt-and-pepper noise, which are the random white and black pixels occur in an image. Also median filter is particularly effective in scenarios where preserving edges and fine details is crucial because it does not compute a weighted average that can blur these features. However, it may not be as effective in reducing certain types of noise, such as Gaussian noise. Unlike weighted average filtering neighbor window size is the only key parameter of it.

It is a popular rank order filter that takes the median value of the defined neighbouring pixels and replaces it to considered centre pixel. As the median value always divides the distribution into equal two parts in the same number of counts, it is naturally not sensitive to the impulse noise/outliers. Unlike the weighted average filters, median filter always takes a pixel value as a substitute, from the existing values within the window if the pixel numbers in the window is odd. If the neighbour pixels follow a positively skewed histogram then the median will be less than the mean. On the other hand, for a negatively skewed histogram the median will be greater than the mean. For a symmetric/near-symmetric histogram (i.e., skewness is near to zero) the median will be equal to or closer to the mean value.

.

In summary, the median filter is an invaluable tool in digital image processing, known for its ability to effectively reduce impulse noise while preserving the integrity of edges and fine details. This makes it indispensable for applications where maintaining image quality through noise reduction is critical. It merges the high intensity textures to its low intensity background if the local distribution is positively skewed and the low intensity texture to its high intensity foreground if the distribution is negatively skewed. But taking traditional fixed size square window can face the similar problem that in average filtering, that it depend on the window size. Hence, an approach of adaptive window (in size or shape or both) consideration can made median filtering best for edge preserving filtering. In our research work, this intuitive property of the adaptive median filter is exploited to develop a robust semantic edge preserving smoothing technique. Where the adaptive (in size and shape both) windows are considered differently for different pixels (edge and non edge pixels) such that edges are preserved semantically while removing noises and textures.

1.1.5 Non-Local approaches in image filtering

Non-local approaches to image filtering, such as Non-Local Means (NLM), patch-based filtering, and dictionary-based methods, have gained significant attention due to their effectiveness in preserving structural and textural image content. These methods rely on the self-similarity property in images, leveraging repeated patterns and textures for more accurate noise suppression. Non-local means (NLM) introduced by Buades, Coll, and Morel [21], exploits the idea that pixels with similar neighborhoods (patches) within an image likely have similar intensities. NLM calculates the weighted average of similar patches found across the entire image, providing an effective noise reduction while preserving fine details. This method is particularly beneficial for natural images with repetitive textures and structures. Variants of NLM include adaptive patch sizes and faster implementations to reduce computational complexity, as presented by Mahmoudi and Sapiro [97] and Kervrann and Boulanger [74]. These enhancements improve performance, making NLM viable for high-resolution images.

Patch based filtering: Patch-based filtering methods, like BM3D (Block-Matching and 3D filtering), introduced by Dabov [34], extend the NLM concept by grouping similar patches and processing them in a 3D transform domain. BM3D's collaborative filtering leverages the self-similarity of patches within an image to achieve state-of-the-art denoising performance, particularly effective for both Gaussian and other types of noise. BM3D has proven especially useful in preserving textures, as it groups similar patches into 3D arrays and applies a joint filter, which averages the noise across the patches. This technique is widely adopted for its robustness in maintaining both fine textures and large-scale structures in the image.

Dictionary based sparse representation methods: Dictionary-based denoising methods aim to represent image patches as sparse linear combinations of basis elements (atoms) in a dictionary, effectively separating noise from signal. Aharon [1] introduced the K-SVD (K-means Singular Value Decomposition) algorithm, which trains dictionaries specific to image patches. By encoding patches sparsely, this method removes noise while retaining essential structures. Dictionary learning has become particularly influential for applications requiring fine detail preservation, such as medical and microscopic imaging [43]. Further advancements include adaptive dictionaries that are tuned for specific types of noise and structures in images, leading to superior denoising outcomes.

Non-local means with deep learning: Recent advances integrate non-local methods with deep learning, where CNNs enhance patch similarity matching. For example, [156] used deep residual learning to improve upon dictionary-based methods by training models on large datasets, achieving noise reduction while retaining non-local structural information. Applications in Complex Noise: These hybrid approaches address limitations in complex noise scenarios where traditional patch-based methods might struggle, making them well-suited for applications like low-light photography and hyperspectral imaging.

Non-local approaches such as NLM, BM3D, and dictionary-based filtering have become cornerstones in image denoising literature due to their capability to preserve textures and structural details. By leveraging self-similarity, these methods offer robust noise reduction while retaining critical image features, proving essential for both natural and medical imaging applications.

1.1.6 Structure preserving filtering

Over the past two decades, advancements have progressed from gradient-based edge-preserving filtering and swift to adaptive window-based edge-preserving filtering and structure preserving filtering based on region statistics. The objective of structure preserving filtering extends beyond the preservation of individual edges or lines; it aims to preserve meaningful structures while eliminating insignificant textures. Notably, anisotropic diffusion [106] and bilateral filters [130] solely rely on directional (horizontal and vertical) image gradients to depict edges or textures. But in later developments for structure preserving filtering, various patch-based statistical metrics, including region covariance, total variation, relative total variation (RTV), modified RTV, and context features based on local max-min, are employed as texture descriptors, utilizing measures of directional gradients of the first and second order. In equation 1.3, showing that in place of direct gradient in equation 1.1, similarity or dissimilarity measure of region features (like the covariance C_p and C_q) of the predefined region around center pixel p and neighboring pixels q, respectively adopted as the edge or texture descriptor between p and q. However, these methods are also contingent on three parameters: variance σ of kernel function, region patch size, and neighbor filtering window size.

$$J_p = \frac{1}{W_p} \sum_{q \varepsilon \Omega_p} G_\sigma(\|C_q - C_p\|) I_q. \tag{1.3}$$

1.1.7 Applications

Image filtering plays a fundamental role in improving visual quality, enhancing features, and reducing noise, which are essential for various applications across image denoising, enhancement, tone mapping, and hyperspectral image classification. By selecting and applying the appropriate filters, these processes can meet unique requirements and improve accuracy, visual clarity, and interpretability of images.

1.1.7.1 Image denoising

Image denoising is essential for reducing noise in images due to environmental factors or sensor limitations, with applications spanning medical imaging, astrophotography, and remote sensing. Key aspects of image filtering in denoising. Filters are selected based on the type of noise present. Gaussian filters are commonly used for images affected by Gaussian noise, while median filters are preferred for impulse noise, as they preserve edges while removing outliers [67].

Edge preserving filtering is a set of techniques in image processing that smooths or reduces noise in an image while preserving important edges. These filters are widely used in computer vision tasks like denoising, tone mapping, and non-photorealistic rendering. Filters like Bilateral filter, Anisotrofic diffusion and non-local means based filters reduce noise while maintaining structural details, crucial for medical images where clear boundaries are essential [19]. Frequency-based filters are a class of image processing techniques that operate in the frequency domain. These filters analyze and manipulate the frequency components of an image, rather than directly working on the pixel intensities in the spatial domain. The frequency domain is obtained by transforming an image using a Fourier Transform, which decomposes the image into sinusoidal components of varying frequencies. Techniques like wavelet-based denoising filter images based on frequency content, efficiently separate noise from details and make it suitable for spatially varying noise types in satellite imaging [29].

Image denoising is essential for restoring clean images affected by various noise types, such as additive Gaussian noise, multiplicative speckle noise, or signal-dependent Poisson noise. Traditional approaches like mean and Gaussian filters [19], wavelet-based denoising [117], and adaptive filters like Lee and Frost [50] for multiplicative noise have shown effectiveness. For impulse noise, median and adaptive median filters are commonly used [18]. Advanced techniques such as Total Variation (TV) denoising [116], anisotropic diffusion [106], and Fourier domain filtering for periodic noise [58] enhance edge preservation. Signal-dependent noise like Poisson is handled using variance-stabilizing transforms [98] and hybrid methods [98]. Recently, deep learning approaches, including CNNs like DnCNN [157] and [81], have outperformed traditional methods in adaptability and performance, making them suitable for complex noise patterns. Hybrid methods combining spatial and frequency-domain techniques [108] further improve denoising results, reflecting a shift towards more adaptive and learning-based solutions.

1.1.7.2 Image enhancement and tone mapping

Image enhancement improves the perceptibility of features in images, making it easier for observers to analyze details. Applications include digital photography, medical imaging, and surveillance. Image filtering aspects for enhancement include several approaches. Contrast enhancement filters like histogram equalization redistribute intensity values by improving contrast. They are widely used in satellite and medical imaging for enhanced visibility of critical features [58]. Edge enhancing filters like Unsharp masking and high-pass filters enhance edges by amplifying high-frequency details, which is especially useful in facial recognition and surveillance imaging [2]. Adaptive filters modify their parameters based on local image characteristics, providing context-sensitive enhancement and preserving significant details, which is useful in medical imaging to improve visibility of tissue structures [120].

Tone mapping adapts high dynamic range (HDR) images for standard displays, balancing brightness levels while preserving details across shadows and highlights. This is vital in computer graphics, photography, and HDR visualization. Key aspects of filtering in tone mapping include global tone mapping filters those apply a uniform adjustment across the entire image, ideal for scenes with even lighting but often lacking in detailed balance for scenes with wide dynamic range [113]. Local tone mapping filters use local operators adjust tones based on local neighborhood characteristics, preserving spatial details across brightness variations, essential in HDR photography where shadow and highlight details are both important [38]. Gradient domain filters applied in the gradient domain enhance detail consistency and reduce halos around high-contrast edges, providing high-quality visual results in professional photography and CGI [48].

1.1.7.3 Classification of hyperspectral images

Hyperspectral imaging captures extensive spectral data, enabling detailed classification of materials and objects in fields like agriculture, environmental monitoring, and mineral exploration. Filtering aspects important in hyperspectral classification include spatial filters for noise reduction which is a widely used technique for denoising hyperspectral images (HSI) by leveraging spatial information from neighboring pixels to suppress noise while preserving essential details. Since HSI data is rich in spectral and spatial information, effective denoising is crucial to improve the accuracy of subsequent analyses, such as classification, segmentation, and target detection. Spectral filters are also applied in noise reduction in hyperspectral images which often suffer from noise due to atmospheric interference. Filters like spectral smoothing and matched filtering are used to enhance signal quality and extract relevant spectral features [153]. Spatial spectral convolutional filters are recently developed using convolutional neural networks (CNNs) that combine spatial and spectral filtering have enhanced hyperspectral classification accuracy, particularly useful in agricultural mapping and geological surveys [152].

1.1.7.4 Semantic segmentation of natural images

Semantic segmentation [6, 11, 26, 68, 73, 77, 100, 151] is a pivotal computer vision task that assigns class labels to each pixel in an image, enabling applications like autonomous driving, medical imaging, and satellite analysis. One of its key challenges is achieving precise boundary delineation and reducing noise in segmentation maps. Filtering techniques, particularly edge-aware filters such as bilateral and guided filters, play a critical role in addressing these issues. These filters are used both in preprocessing, where they smooth noisy inputs while preserving edges, and in postprocessing, where they refine segmentation outputs to align boundaries with actual object edges. Additionally, advanced methods like Conditional Random Fields (CRFs) and edge-preserving neural architectures integrate filtering principles to enhance segmentation accuracy. By ensuring noise suppression, boundary precision, and feature enhancement, filtering significantly elevates the reliability and quality of semantic segmentation in complex real-world scenarios.

1.2 Challenges

Despite significant advancements towards structure preserving filtering techniques, many of them grapple with the challenge of blurring the original image, leading to information loss and weakening the edges, as they unable to consider semantic information while smoothing.

Blurring and information losses: Image filtering is all about to update the pixel values from the neighbour information to remove the insignificant details or noises. To do so filtering process blur the overall image and causes information losses. So one of the primary challenge to develop a effective filtering technique is to preserve the significant objects properly. The ultimate aim of image filtering is to remove noises while preserving significant objects with minimal information loss. The recent researches focuses on this by incorporating 1) textural features, 2) structural edge features or 3) both textural and structural semantic feature.

Semantically detect significant structures: One of the most challenging tasks in image filtering is to differentiate between significant objects/structures to be preserved and noises/textures to be removed. Most of the existing approaches directly rely on gradient or range difference information for this task. Only gradient information can not differentiate texture and structure. But incorporating semantic information, or semantically defining significant objects is the most crucial and challenging task. In this research some novel semantic features have been proposed to develop semantic-aware structure preserving filtering techniques.

Parameter sensitivity: Another important problem of most of the existing filtering techniques is their parameters dependency or sensitivity. The existing techniques have several parameters that need to be wisely set manually within a predefined continuous interval to get the best results for different images. This parameters dependency make them inappropriate for applying in different applications such as classification, segmentation, denoising, etc. This research has focused on this issue and developed some effective filtering techniques which are less parameters sensitive in a way such that only need to select parameters values from few discrete options.

1.2.1 Preservation of significant structures

Due to wide varieties of images used for different applications, preservation of significant structures/edges during image filtering is become very challenging. The term significant structure is vague in sense. For different images used in different type of applications it may differ in sense. Significant object/structure may be in the sense of human visual perception or in the purpose of some particular applications perspective. So, the filtering technique has to define generalized robust features for preservation of significant structures.

1.2.1.1 Semantic features for preserving significant structures

There are several techniques developed for structure/edge preserving image filtering. Most of them either considered intensity difference or image gradients or some statistical measure for structure/edge detection during smoothing process. But there are many cases in which only image gradient or gradient based features can not differentiate significant structures/edges and insignificant textures/noises. Here the term significant can be in some semantical sense of human visual perception or some scientific image analysis and applications perspective. During filtering for different types of images, semantically recognise the significant structures/edges to be preserved and insignificant textures/noises to be removed is a challenging and much needed work. Consideration of appropriate semantic features in filtering process can significantly improve the quality of the filtered image.

1.2.1.2 Semantic-aware filtering

Edge-preserving image filtering aims to smooth an image while maintaining its edges. Recently, several techniques have been developed to differentiate between structures and textures at varying scales. These techniques typically use a userselected scale measurement to control texture smoothing. However, natural images often contain objects of varying sizes, making it difficult to describe them with a single scale measurement. Additionally, edge and contour detection, which is closely related to edge-preserving filtering, has seen significant advancements. Despite this progress, many state-of-the-art filtering techniques rely heavily on image gradients for edge detection and are dependent on multiple parameters. This limitation has led us to propose the use of semantic-aware edge features to better distinguish between textural and structural edges.

In this research work we have developed some efficient semantic-aware structure preserving image filtering techniques and tested those in realm of few applications of computer graphics like image denoising, image enhancement, tone mapping as well as in image classification.

1.3 Evolution and developments of image filtering techniques

Linear filters are the oldest image filtering techniques implemented with the help of a convolution mask. Due to their simplicity and low computational cost, traditional linear filtering techniques like mean filtering, box filtering, Gaussian filtering [58, 67] are still relevant for many image processing applications. One of the primary disadvantages of linear filters is that during filtering they destroyed lines, edges, and other fine details of the image. To reduce this, last three decade a variety of non-linear filters have been developed [27, 32, 42, 55, 61, 65, 82]. The main aim of all these developments is to preserve the image details and local geometries in much better way, while removing the undesired textures from the images.

In early 1990's an important concept anisotropic diffusion is introduced for non-linear filtering [106]. It exploits Partial Differential Equations (PDE) to estimate pixel wise spatially varying diffusivity from image gradients. These diffusivity prevent excessive smoothing at image edges, thereby preserving important image structures while effectively reducing noise. The other popular edge preserving filtering techniques based on PDE that proposed in the literature are non-linear isotropic (spatial) diffusion [137], anisotropic (directional) diffusion [17], level set methods [99], total variation methods ([24, 115]) *etc.*

Although the above PDE based filtering methods show their robustness to preserve edges, they are implemented through an iterative process, which can often be slow and may pose challenges related to stability and efficiency. As an alternative bilateral filter was first introduced by [130] based on the works [8, 124], and later it improved in [42]. Bilateral filter (BF) is a simple nonlinear filtering approach, that combines domain and range considerations for filtering. The spatial domain and the range function controls the smoothing and edge preserving effects, respectively. Because of its simplicity and effectiveness, bilateral filtering has garnered significant attention from researchers and has been successfully applied to various computational photography applications. In the image processing literature, several modifications of it has also been developed to handle the drawbacks of the existing methods.

Although bilateral filters are non-iterative in nature, their brute force (direct) implementation is very slow. The computational complexity of such brute-force implementation is $O(Nr^2)$ (N is pixel numbers in the image), which is eventually become high when the radius r of kernel is large. Therefore, several adaptive techniques [27, 28, 49, 55, 104, 107, 134, 146, 154] have been developed to make computationally faster the bilateral filters using different kernel implementation techniques and data structures. However, achieving fast or real-time implementation remains a challenging problem. To address this challenge, a new filtering approach called guided filtering ([63]) was introduced. This method effectively preserves edges while denoising by taking into account the content of a guidance image. One advantage of the guided filter compared to the bilateral filter is its $O(N)$ time complexity, ensuring efficient implementation without being dependent on the kernel radius r. A filtering method that used self-similarity and non-locality characteristics of images is proposed in [19]. A few extension of this non-local filtering are presented in [33], [96]. Manifold learning based filtering for high dimensional image and video processing are found in [53–55] which make the BF implementation faster.

One of the limitation of the above different approaches of BF techniques is that they directly rely on image gradient for describing edge/structure, which is prone to fail to preserve the different type of edges properly or create false edges. To address this issue, local statistical measure based adaptive and iterative filtering models draw more attention in the last decade. These filtering techniques introduced the concept of texture-structure decomposition. Structure-preserving texture filtering is an area of ongoing development with diverse applications in computational photography and image analysis. This operation decomposes an image into its prominent structures and fine-scale details, facilitating various image manipulations such as tone mapping, detail enhancement, visual abstraction, scene understanding, and more. The large number of filtering methods proposed in the literature are mainly divided into four categories: average based [63, 145], optimization based [22, 135, 138, 139], rank order based [10, 94] and patch based [32, 72, 129]. The filtering methods in the first two are particularly effective for removing noise and fine-scale geometric details. For example, bilateral filter [5, 130], weighted least square filter [47, 91, 135], L0 smoothing [128, 138], and the methods using anisotropic diffusion [103, 106, 143] are successfully used to filter low-contrast noise but failed to smooth high-contrast noises or textures. Rank order based filtering techniques are capable of filtering high-contrast noises or textures using a local histogram where the color/intensity of each pixel is replaced by some order ranked value (like the maximum, minimum, median, etc) of the neighbouring pixels. Thus, they preserve the original contrast in the filtered image. Texture filtering [89, 127, 158] is all about removing the insignificant regular or irregular patterns (textures) from the filtered image while preserving the significant structures or objects of the image. Since the measure of the visual patterns of regular or irregular textures is not considered by the filtering methods of the first three categories, they are not that effective in texture filtering. Whereas, patch based filters are capable of considering the measurements of visual patterns. Thus, more suitable for texture filtering. Generally, visual patterns are measured with axis-aligned patches or by computing the patch based statistical measures. There are patch based techniques presented in [15, 20] performed filtering by applying some non-local measurements, but the textural measurements are not taken into consideration. As a result, these methods are less effective for texture filtering. The spatial relationship, frequency, and symmetry/asymmetry of the textural pattern of regular or near-regular shapes are taken into consideration for texture identification and filtering is presented in [62, 93]. However, for developing a robust texture filtering technique, filtering out arbitrary textures of both regular or irregular shapes/patterns is important. Though, finding out the insignificant irregular textures and removing them in a semantic sense is always a challenging task. In this regard, Rudin et al. [115] have effectively enforced total variation (TV) regularization constraints to preserve large-scale edges. In order to further improve the texture-structure separation based on TV, Xu et al. [139] introduced relative total variation (RTV). Both the methods are developed based on the concept of the aggregation of signed gradient values within a local window/patch often produces a larger absolute value for structural edges than its textural counterpart. As the texture regions usually produce inconsistent gradients within a local window, the gradient measure is used to differentiate textural and structural edges. However, such a measure may misidentify small structural edges as textures. Moreover, the use of larger neighbouring windows overlapping different regions may result in very similar RTV values for neighbouring pixels of different regions. As a result, these methods are prone to produce over-smooth structural edges. In [72], Karacan et al. introduced the region co-variance descriptor as a textural measurement that uses second-order statistics extracted from the fixed size patches. Although this descriptor performs better separation between texture and structure, it fails to avoid the over-blur effect at the structural edges, as the overlap patches near an edge always share almost similar region co-variances.

To mitigate this problem, some later techniques have tried to optimize the position of patches [32] or the size of patches [150]. For computing the textural measure, Cho et al. [32] assumed that a pixel may not be always well represented by the patch centered on it. Finding out the nearby patch which is least likely to contain structural edges can resolve the patch overlapping issue. Thus, they have proposed a patch shifting technique to avoid the neighboring window overlapping problem, which has improved the effectiveness of the textural measures. The main challenge in this technique is to choose an effective patch size and its position. To address this issue [140] introduced a concept of edge aware dynamic shape and size based filtering approach. In [69] a scale aware structure preserving filtering mechanism is proposed by using directional Relative Total Variation (dRTV) measure. Some graph based minimum spanning tree/shortest path filtering concept was presented in [10, 37, 101]. There is another kind of filtering techniques exist in the literature based on mathematical morphology. Some developments of adaptive structure preserving morphological filtering was proposed by several researchers [59, 80, 82, 83]. They introduced the concept of adaptive morphological filtering and morphological amoebas, which dynamically define structuring elements size and shape for each pixel. [4] proposed a morphological bilateral filtering approach by defining a spatially-variant bilateral structuring function, where the adaptive structuring elements are obtained by taking thresholds on the bilateral functions. Connected component filtering, an advanced version of morphological filtering found in [16, 23, 35, 39, 119, 142] applied for hyper-spectral image denoising. The semantic edge-aware structure preserving filtering concept was introduced in [84, 145, 149]. In [145] a learning-based edge-preserving filtering technique is presented. In [87], to avoid over-blurring smaller structural edges, a smaller patch (shrink version of the original patch) is considered for the pixels located at the nearby structural edges, while a larger patch (original patch) is used for the pixels representing the texture regions. Since the technique is only able to shrink the original patches and is unable to expand them, it prevents the original patches from being as large as needed. Thus, filtering large-scale textures is not efficiently done by this method. For multiple scaled texture smoothing, a scale-aware method is presented in [69]. Depending on the patch-based statistics, this method finds an optimal per-pixel adaptive smoothing scale for texture filtering. It used RTV based statistics for the textural measure. Because of the over-blurring issue with RTV, this method may fail to estimate the scales of textures properly. As a result, large-scale texture filtering may not be efficiently performed by this method. As the scale of the textures is related to the gradients, Lee et al. [79] proposed an adaptive gradient smoothing method by introducing the interval gradient measure. An adaptive relative total variation filter (ARTVF) is proposed in [149]. Since the estimation of the scales of textures depends on the statistics of gradients, the method is effective when the scales of the textures are similar. The textures with large, varying texels are not properly handled by this method, as the statistics of gradients also change drastically. So, the main challenge for developing a robust structure preserving filtering technique lies in handling varying scales of textural patterns.

To address this, some recent developments have proposed dynamic shape adaptive window generation by taking into consideration the linear characteristics of the structural edges [88, 94, 140]. They approximate the linearity of the structural edges by local gradient directions. As the structural edges are not always axis-aligned, approximating the linearity of the edges locally is not an easy task. There are some studies directed towards the development of edge-aware filters [25, 54, 145, 149, 155, 160]. Some of these methods define an edge-aware adaptive window for filtering depending on a prior edge-map generated by introducing modified edge features or using existing edge detection techniques. As the generation of a proper edge-map is a challenging and developing task, semantic-aware filtering has undergone more exploration in recent research.

Thus more recent studies are focusing on the development of semantic edge-aware filtering techniques [25, 76, 90, 95, 109, 111, 145, 147, 149, 159, 160]. These techniques proposed structural-edge feature or edge-aware textural features. [54] proposed edge-aware filtering by reducing the dimension of the RGB image. Edge detection techniques are used for developing learning-based edge-aware filtering in [145, 149]. An artificial neural network (ANN) is trained for edge-aware smoothing in [95, 159] and [75, 76] used a convolution neural network (CNN) based framework to exploit deep variational prior for the same task. An edge-aware filtering by combining bilateral filter and weighted least square is presented in [90]. [25] proposed an optimization framework for filtering by combining edge detection and L_0 gradient minimization recursively. [147] proposed a novel soft clustering algorithm using a restricted Gaussian mixture model for the same. Although, these edge-aware filtering techniques provided satisfactory results even for varying scale irregular textures, they have multiple parameters and the filtering result is very much dependent on the values of these parameters set manually. In this research, we proposed a novel edge-aware filtering method that provides better filtering results for a wide variety of images with minimum fine-tune of its parameters. The aim of the present work is to develop some semantic edge-aware structure preserving filtering techniques which explore semantic information that has not been exploited earlier.

1.4 Methodologies/approaches

Recent developments in bilateral filtering techniques have successfully used regional statistical features to describe regular or irregular textures in the image. Although the statistical features are robust to distinguish different structures present on the image, they cannot exploit the geometric features of the structures. As a

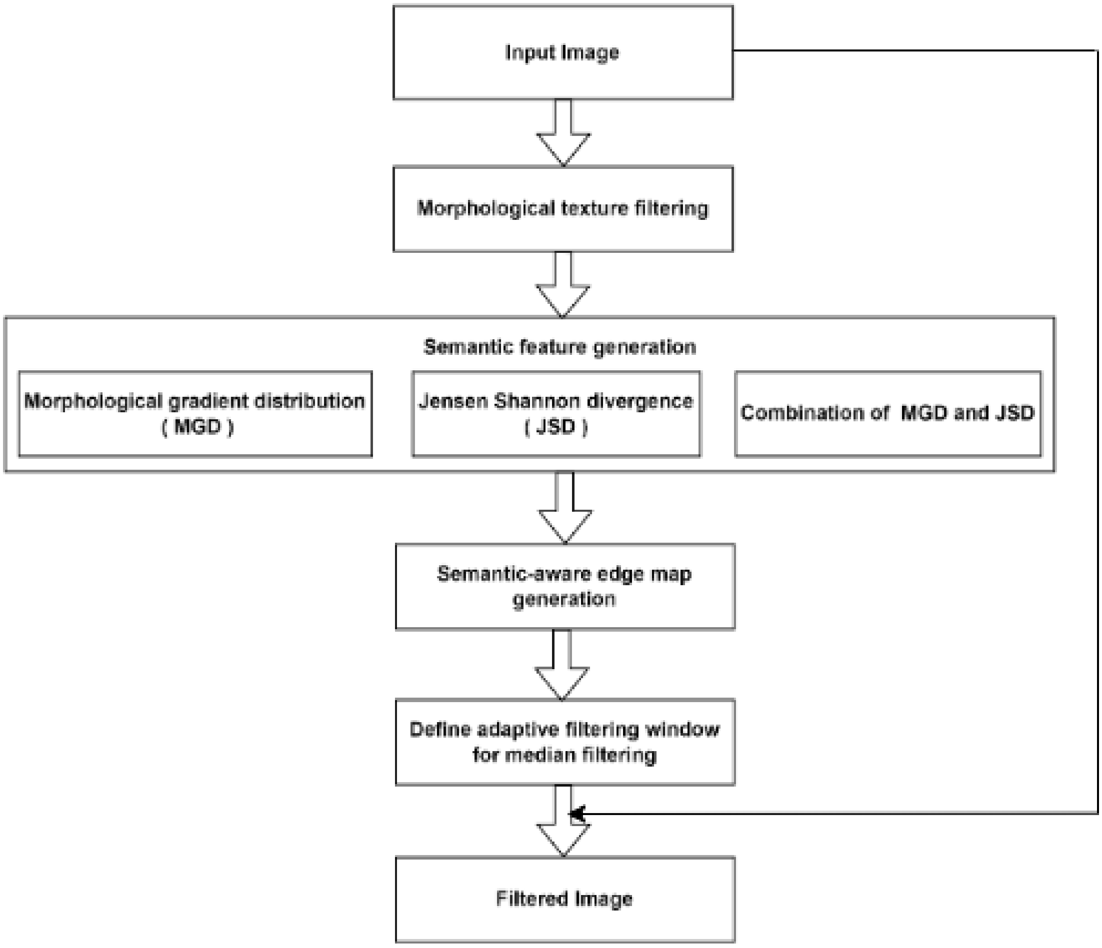

Figure 1-2: Proposed framework

result these features are very much parameter sensitive. On the other hand, morphological filters intuitively preserve the different shapes of the structures (objects). But the filtering results are dependent on the considered structure element or threshold value. Moreover, in the literature it is found that median filter provides excellent results when the size of the dynamic window of the median filter are chosen properly. Thus, developing filtering technique by exploiting both, patch-based region statistics and morphological filters to define adaptive window of the median filter could solve some critical issues present in filtering techniques. Unfortunately, such a filtering technique is seldom presented in the image processing literature. In this thesis, our main objective is to develop few robust structure preserving filtering techniques that exploit semantic information of the image by using region-based statistics and/or morphological filters for defining the adaptive dynamic window of the median filter with minimal fine tune of the parameters. The major steps of this approach are as follows:

Propose semantic features for texture-structure decomposition- In recent research, there has been an introduction of a concept termed semantic-aware structure preserving filtering. While the definition of semantic-aware structure preserving filtering remains somewhat ambiguous due to the vagueness of the term ”semantic-aware”, its essence lies in the identification of significant objects. Unlike traditional methods that primarily detect individual edges based on sharp intensity differences, semantic-aware filtering focuses on recognizing structures in a meaningful manner, whether to preserve or remove them. The challenge lies in defining what constitutes semantically meaningful structures within an image. In our work, we have proposed few novel semantic features for addressing this challenge. First, we have employed morphological gradient distribution to define semantic feature. Second, we have utilized the Jensen-Shannon divergence (JSD) for the same. Third both, morphological gradient distribution and JSD are exploited to define robust semantic features for texture-structure decomposition. Moreover, to generate such semantic features our aim is to define less number of parameters which required minimal fine tune.

Semantic-aware edge-map generation- Once the semantic features of the pixels are obtained, a semantic-aware edge-map of the input image is generated by using these features. Depending on the semantic features different algorithms are proposed to generate the edge-map of the input image. The main focus of these algorithm is to preserve the edges of the significant objects and remove the edges of the insignificant objects or textures.

Adaptive dynamic window for median filter- After generation of semantic edge-map of the input image, a novel technique needs to be proposed to define the dynamic window of the median filter with help of the generated semantic edge-map. To develop such technique care should be taken so that the proposed algorithm either define its parameters automatically or less sensitive to its parameters. Figure 1-2 shows the framework of our approaches.

1.5 Experimental validation

To evaluate the effectiveness of our developed filtering techniques the common approach is to compare the filtered image and its zoomed parts of different structural and textural areas by human visual perception. As well as the quality of the filtered image can be assessed by some standard image quality measurement index (IQA). For image classification problem, the classification results is used to validate the effectiveness of our filtering technique.

Visual perception:- The common way of assessing filtered image quality subjectively is human visual perception based. Specially for structure preserving image filtering techniques need to check how well the filtered images preserving significant objects' structure while removing textural fine scale details. As the term significant is depends on manual observation or some specific applications aspect, visual perception is a useful way to validate it.

IQA:- IQA stands for Image Quality Assessment, which is a field of study and a set of techniques used to measure and evaluate the perceived quality of images. The goal of IQA is to develop metrics that can objectively quantify the degree of distortion or degradation in an image compared to a reference or original image. This is particularly essential in numerous applications such as image compression, transmission, and processing, where maintaining high visual quality is crucial. Common metrics used in IQA include Signal to Noise Ratio (SNR), Peak Signal to Noise Ratio (PSNR), Mean Squared Error (MSE), and Structural Similarity Index (SSIM), among others are the reference based measure, i.e it measure the filtered image quality with respect the the original input image. IQA techniques play a significant role in optimizing and validating image processing algorithms and systems. There are other IQAs like perceptual based image quality evaluator (PIQE) or naturalness image quality evaluator (NIQE) which are non-reference based.

HSI classification:- Hyperspectral Image (HSI) classification is the process of categorizing different materials or land cover types within an image captured by a hyperspectral sensor. Unlike traditional images that have three bands (RGB), hyperspectral images have numerous bands, each capturing information across a specific wavelength range. HSI classification entails assigning each pixel in the image to a specific class or category based on its spectral and spatial characteristics. The classifiers such as Neural network, K-nearest neighbours, support vector machine (SVM) and random forest (RF) are widely used for HSI classification. Evaluations of the performance of the classification was done by using validation data that was not used during training. Common metrics include overall accuracy (OA), average accuracy (AA), kappa coefficient (κ) obtained from confusion matrix are used to measure the classification performance. In this thesis the effectiveness of the proposed filtering techniques for incorporating better spatial information are assessed by using such metrics.

1.5.1 Performance measure

To assess the effectiveness of our developed filtering techniques in this thesis, the quality of the filtered image is assessed by visual perception as well as with the help of some standard image quality measurement index. The conventional Image Quality Assessment (IQA) measures such as Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR) are not useful for assessing the quality of the filter images produced by the structure preserving filtering techniques [140]. We used three subjective reference IQA metrics: Structural Similarity Index (SSIM) [136], Multi-scale SSIM (MSSIM), and Mutual Information (MI) [131] and one subjective no-reference metric: Perception Image Quality Evaluator (PIQE)

[132] for quantitative comparisons.

SSIM:- Structural Similarity Index (SSIM), is a metric used to quantify the similarity of two images. It is commonly employed in the field of image processing and computer vision to assess the perceived quality of an image. The SSIM index is designed to capture changes in structural information, luminance, and contrast that are perceptible to the human visual system. The SSIM value of two images x and y is computed as follow:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{1.4}$$

where μ_x, σ_x^2 and σ_{xy} represent the mean, variance of x and covariance of x, y, respectively. C_1 and C_2 are constants added to prevent instability when the denominators approach to zero. The SSIM index ranges from -1 to 1, with 1 signifying perfect similarity between the two images. A higher SSIM value implies a closer resemblance.

MSSIM:- Multi-Scale Structural Similarity Index (MSSIM) is an extension of the Structural Similarity Index (SSIM) that takes into account information at multiple scales. It addresses limitations in SSIM, particularly its sensitivity to changes in scale and orientation. The MSSIM metric is designed to provide a more comprehensive evaluation of image quality across various resolutions. The formula for MSSIM involves averaging the SSIM values obtained at different scales as follow:

$$MSSIM(x,y) = \frac{1}{L}\sum_{i=1}^{L} SSIM_i(x,y) \tag{1.5}$$

MSSIM provides a more robust assessment of image quality because it considers structural similarities across a range of spatial frequencies. This is particularly useful when evaluating the impact of distortions or compression artifacts that may affect different scales differently. MSSIM is widely used in image quality assessment tasks, and it has become a standard metric in evaluating the performance of image and video compression algorithms. It offers a more nuanced and informative evaluation compared to traditional single-scale similarity metrics.

MI:- Mutual Information (MI) can serves as a metric for evaluating the quality of filtered images, especially in the context of computer vision and image processing. The basic idea is to measure the amount of information shared between the original and filtered images, where higher mutual information indicates better quality or similarity. In the case of filtered image quality assessment, the MI computed between the original image (X) and the filtered image (Y) can be computed as follows:

$$I(X,Y) = \sum_{x \in X} \sum_{y \in Y} P(x,y) \log\left(\frac{P(x,y)}{P(x)P(y)}\right) \tag{1.6}$$

Where the joint probability mass function of intensity values in the original and filtered images is P(x,y). The marginal probability mass functions of intensity values in the original and filtered images, respectively are P(x) and P(y). The above formula essentially quantifies how well the intensity values in the original and filtered images are related. Higher mutual information suggests that the filtered image retains more information from the original image, indicating better quality. The computation of mutual information involves discretizing the intensity values, especially when dealing with continuous images. Additionally, the joint and marginal probabilities need to be estimated from the pixel intensity distributions. In practical terms, mutual information for filtered image with original image can serve the purpose of evaluating the effectiveness of different image filtering techniques, such as denoising filters or edge-preserving filters. It provides a quantitative measure of how well the filtered image preserves information with respect to the original image.

PIQE:- PIQE stands for Perceptual Image Quality Evaluator. It is a metric used to assess the perceptual quality of images, aiming to mimic human perception of image quality. PIQE is designed to provide a numerical score that correlates with the perceived quality of an image. This metric is commonly used in the field of image and video processing for quality assessment. The PIQE algorithm is based on modeling the human visual system's sensitivity to various image distortions. It takes into account factors such as luminance, contrast, color, and spatial details. The goal is to evaluate the overall visual quality by considering multiple aspects that contribute to human perception. The output of PIQE is typically a single score that represents the quality of the image. A higher PIQE score generally corresponds to better perceptual quality. Conversely, lower scores indicate lower perceived quality or the presence of image distortions. The application of PIQE is relevant in scenarios where it's crucial to objectively measure the quality of processed or compressed images. This can include assessing the impact of compression algorithms, evaluating image enhancement techniques, or monitoring the quality of images in various imaging systems. It's worth noting that there are different versions or implementations of perceptual quality metrics, and the specific details of the PIQE metric may vary depending on the version being used. We have used the Matlab 21 inbuilt function.

1.5.2 Data sets

In this thesis we have proposed few semantic-aware structure preserving image filtering techniques. In order to establish the effectiveness of our developed filtering techniques several images are used which are made publicly available in different publications cited in the thesis and few internet sources like http://www.cse.cuhk.edu.hk/~leojia/projects/texturesep/index.html, http://cg.postech.ac.kr/publications and https://eod-grss-ieee.com/dataset.

1.6 Objectives

Developing an effective structure-preserving image filtering technique requires addressing three primary challenges: (i) minimizing image blurring and information loss, (ii) incorporating suitable semantic information for accurate texture-structure decomposition, and (iii) reducing dependency on adjustable parameters. Existing techniques in the literature rarely meet all three challenges simultaneously. This research aims to design robust structure-preserving image filtering techniques that overcome these challenges. To achieve this goal, four key objectives have been accomplished, marking significant advancements in the field of image filtering.

1. *To develop a semantic edge-aware median morpho-filtering technique:* According to the recent developments, the most challenging task of texture filtering is to filter out the irregular and varying scale texture while preserving the structural contents with minimum distortion. For better handling varying scale irregular textural patterns, in this objective we propose a novel semantic-aware structure preserving filtering technique. The technique, first, generates a semantic edge-map of the input image by analysing the skewness of global and local histograms of the morphological gradient image obtained from the input image. Then, with help of the generated semantic edge-map, a novel method is proposed to define edge-aware adaptive dynamic window of the median filter for filtering each pixel of the input image by excluding its neighbor pixels belonging to different textural or structural regions. Finally, the proposed median filter and a morphological filter iteratively applied to generate the filter image.

   Although, the filtering technique proposed in the first objective produced satisfactory results for a wide varieties of input images, it has one critical parameter. To incorporate semantic information of a pixel for determining whether it is an edge or non-edge pixel, the skewness of the local histogram of gradient image constructed by considering a fixed size window is computed.

So, the semantic edge-map generated by the proposed technique is highly sensitive to the size of this window, which is varied from image to image. In the next two objectives our aim is to develop more parameter efficient semantic-aware filtering techniques.

2. *To develop a reduced parameter sensitive semantic edge-aware image filtering technique:* The effectiveness of edge-aware structure preserving filtering techniques is dependent on their ability to identify right structural edges. Since in most of the images structural and textural edges are confused with each other, only tonal and sharpness dissimilarity are inadequate to discriminate them. In this objective by exploiting Jensen-Shannon (JS) divergence, a set of novel semantic features are proposed to depict the pixels of the input image. These features are defined in such a way that incorporate semantic information of the pixel for determining whether it is a structural edge pixel or not. Projecting all the pixels into the feature space, the technique applied k-means clustering to generate a semantic-aware edge-map of the input image. Once the edge-map is obtained, an edge-aware adaptive median filter which is proposed in first objective is used to generate the filter image. The developed method outperforms for a wide variety of images with the minimal fine tune of its parameters.

3. *To develop another parameter efficient semantic edge-aware image filtering technique:* The aim of this objective to propose another parameter efficient semantic image filtering technique by exploiting both the semantic features proposed in objectives one and two. It generates an edge-map of the input image by exploiting semantic information in two phases. In the first phase, two novel features: i) the semantic gradient, and ii) the semantic skewness gradient, are utilised to generate the semantic gradient image (SGI) of considered input image. In the second phase, with the help of the generated SGI, the window size associated to each pixel of the input image is automatically defined to determine whether it is a structural edge or non-edge pixel

by exploiting JS divergence. Once the semantic edge-map is generated, an edge-aware adaptive median filter which is proposed in first objective is used to generate the filter image. Although the proposed technique required to define multiple windows, the sizes and shapes of most of these windows are either automatically defined or kept fixed, irrespective of the consider input images.

4. *To incorporate spatial information by exploiting proposed filtering techniques for hyperspectral image classification and semantic segmentation of natural imaages:* Structure preserving image filtering techniques differentiate structures and textures in the image by identifying textural and structural edges. The fundamental goal of such techniques is to eliminate noise and smooth the textures while causing the least amount of distortion to the structures of the objects on the image. The development of semantic-aware structure preserving filtering techniques, which leverage semantic information in the image to differentiate structures from textures, has been the subject of recent research. Even if the filter images produced by such sophisticated approaches provide richer spatial in formation, they are seldom used for HSI classification. The last objective of this thesis is to exploits our proposed filtering techniques to incorporate spatial information for spectral-spatial HSI classification and semantic segmentation of natural images.

1.7 Organization of the thesis

The thesis presents few novel semantic-aware structure preserving image filtering techniques and their applications. The structure of this thesis is as follows:

Chapter 1: It gives brief introduction of the thesis. Several image filtering techniques and their usefulness are discussed. The evolution of image filtering techniques with the challenges and setbacks are also presented here. The purpose of thesis and the methodologies followed to developed different semantic-aware filtering techniques are presented. Finally, it lays down the objectives of the thesis.

Chapter 2: Presents a novel semantic-aware structure preserving median morpho-filtering technique for better handling varying scale irregular textural patterns. The technique, first, generates a semantic edge-map of the input image by analysing the skewness of global and local histograms of the morphological gradient image obtained from the input image. Then, with help of the generated semantic edge-map, a novel method is proposed to define edge-aware adaptive dynamic window of the median filter for filtering each pixel of the input image.

Chapter 3: Presents a parameter efficient novel semantic-aware structure preserving filtering technique. It exploits JS divergence to represent the pixels of the input image by a set of semantic features. Projecting all the pixels into the feature space, the technique applied k-means clustering to generate a semantic-aware edge-map of the input image.

Chapter 4: Presents another parameter efficient novel semantic-aware structure preserving filtering technique by exploiting both the semantic features proposed in Chapter 2 and Chapter 3.

Chapter 5: Describes a method that incorporate spatial information by exploiting one of the semantic-aware structure preserving filtering technique proposed in the above chapters for hyperspectral image classification and semantic segmentation of natural images.

Chapter 6: Drawn the concluding remarks of the thesis by summarizing the works done and also point out the lists of the possible further research in this area.