**ECE 258 – Multirate Digital Signal Processing**

# Reconstruction Quality Analysis of Gaussian, Max Pooling, and Max-Averaging Pyramids


BY-

Dharmesh Mohanraj

1st year M.S in ECE

# Reconstruction Quality Analysis of Gaussian, Max Pooling, and Max-Averaging Pyramids


## 1.Abstract:

This project investigates the reconstruction quality of three image pyramid methods—**Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**—focusing on their base layer (1-level pyramid) to understand their fundamental reconstruction behavior. Using the **Structural Similarity Index (SSIM)** as the primary metric, the project evaluates and compares these methods under various conditions.

The workflow involves tuning the parameters of each pyramid method to their optimal values for fair and legitimate comparisons. For the Gaussian Pyramid, the bandwidth parameter sigma (σ) is analysed using FFT and radial frequency analysis to determine the best suitable value. Across all three pyramids, two processes—**construction** and **reconstruction**—are analyzed. In the reconstruction process, **bicubic interpolation** was identified as the most effective among the three tested techniques (nearest-neighbour, bilinear, and bicubic). Additionally, the analysis of varying **downsampling and upsampling factors** determined that a factor of 2 provides the best balance between reconstruction quality and data reduction.

Once the optimal parameters were established, the computational efficiency (execution time and memory usage) of each method was evaluated. The results show that the **Gaussian Pyramid** is the most computationally efficient, while the **Max-Averaging Pyramid** delivers the highest reconstruction quality, achieving the best SSIM scores across all analyses. Two test images, each containing a mix of texture, high-frequency, and low-frequency features, were used to ensure a comprehensive evaluation.

This study concludes with a comparative analysis of the methods, highlighting their respective strengths, trade-offs, and application suitability for tasks requiring high-quality reconstruction or efficient processing.


## 2.Introduction:

Image pyramids are hierarchical structures that represent images at multiple scales, facilitating efficient processing in various computer vision tasks. The concept was introduced by Burt and Adelson in 1983 [2], emphasizing their utility in image compression and analysis. The Gaussian Pyramid, a widely used technique, involves successive smoothing and subsampling of an image to create progressively lower-resolution representations. This approach is fundamental in applications such as image blending, texture mapping, and multi-scale analysis.

In the realm of deep learning, pooling operations are integral for downsampling feature maps within convolutional neural networks (CNNs). Max Pooling, which selects the maximum value within a specified window, effectively captures prominent features while introducing spatial invariance. This method has been pivotal in enhancing the performance of CNNs across various image recognition tasks [3]. However, standard pooling methods sometimes fail to balance detail retention and global smoothness [4]. Building upon max pooling, the Max-Averaging Pyramid combines max pooling and average pooling operations to retain salient features while reducing artifacts. This hybrid approach provides a more comprehensive representation of the image at different scales.

Evaluating the reconstruction quality of these pyramid methods is crucial, especially when images undergo downsampling and subsequent upsampling. The Structural Similarity Index Measure (SSIM) is a perceptual metric that quantifies image quality degradation, aligning more closely with human visual perception compared to traditional metrics like Peak Signal-to-Noise Ratio (PSNR) [1], [5]. Studies have shown SSIM to be particularly effective in assessing reconstruction quality in the presence of common distortions.

## 2.1 Motivation:

Understanding the reconstruction capabilities of different image pyramid methods is essential for applications like image compression, super-resolution, and multi-scale analysis. While Gaussian Pyramids are well-studied for their smooth representations and anti-aliasing properties [2], the comparative performance of Max Pooling and Max-Averaging Pyramids in reconstruction tasks remains underexplored. Additionally, the trade-offs between reconstruction quality and computational efficiency have not been comprehensively analyzed across these methods. This project aims to fill this gap by systematically comparing their reconstruction abilities and providing insights into their applicability.

## 2.2 Objectives:

The primary objectives of this project are:

**1.Parameter Optimization:** Identify the optimal parameters for each pyramid method to achieve the best reconstruction quality. This includes determining the appropriate bandwidth parameter sigma (σ) for the Gaussian Pyramid and selecting suitable downsampling and upsampling factors for all methods.

**2.Interpolation Method Analysis:** Evaluate the impact of different interpolation methods (nearest-neighbour, bilinear, and bicubic) during the upsampling process on the reconstruction quality of each pyramid method.

**3.Downsampling Factor Evaluation:** Investigate the effects of varying downsampling factors (2, 4, 8) on the reconstruction quality and determine the most effective factor for each pyramid method.

**4.Computational Efficiency Assessment:** Compare the computational time and memory usage of each pyramid method under optimal parameter settings to assess their efficiency.

**5.Comprehensive Comparison:** Conduct a thorough comparison of the three pyramid methods, highlighting their benefits, trade-offs, and suitability for various applications based on reconstruction quality and computational efficiency.

By achieving these objectives, the project seeks to provide a clear understanding of the performance and applicability of Gaussian, Max Pooling, and Max-Averaging Pyramids in reconstruction tasks.

# 3.Methodology:

This section outlines the methodologies employed in constructing and analysing three image pyramid methods: **Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**. The focus is on their construction, reconstruction processes, and the analytical framework used to evaluate their performance.

## 3.1 Gaussian Pyramid:

The Gaussian Pyramid is a hierarchical representation of an image, where each subsequent level is a smoothed and downsampled version of the previous one. This technique is fundamental in multiresolution image processing and is widely used in applications like image compression, progressive transmission, and multi-scale analysis [6].

**Construction of a 1-Level Gaussian Pyramid**

The construction of a Gaussian Pyramid as shown in **Fig 1**. involves the following steps:

1. **Smoothing (Low-Pass Filtering)**:
   - A Gaussian filter is applied to the original image to suppress high-frequency components and reduce noise.
   - The **Gaussian function** in two dimensions is given by:

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

   - Here, sigma(σ) (standard deviation) determines the bandwidth of the filter. A larger sigma(σ) results in stronger smoothing, effectively removing fine details, while a smaller sigma(σ)preserves details but may retain noise [6].
2. **Downsampling (Subsampling)**:
   - After smoothing, the image is reduced in resolution by retaining every alternate pixel in both the horizontal and vertical directions. This effectively decreases the image size by a factor of four.
   - This step captures a coarser representation of the image, removing redundant high-frequency details [6], [7].

The resulting image forms the next level of the Gaussian Pyramid, representing the original image at a lower scale.

**Reconstruction from a 1-Level Gaussian Pyramid**

Reconstruction involves reversing the downsampling process to approximate the original image:

1. **Upsampling**:
   - The low-resolution image is expanded by inserting zeros between pixels in both dimensions, effectively doubling its resolution.

- This zero-insertion prepares the image for interpolation to fill in the missing pixel values [7].

2. **Interpolation (Smoothing)**:
    - A Gaussian filter is applied to the upsampled image to estimate values for the newly added pixels.
    - This process smooths the transitions and approximates the original resolution. However, due to the loss of high-frequency information during downsampling, the reconstructed image may not perfectly match the original [6], [7].

**Role of σ:**

The parameter σ, often referred to as the bandwidth, plays a crucial role in controlling the extent of smoothing:

- **High σ**:
    - Leads to greater blurring, removing noise and fine details.
    - Useful for applications prioritizing smoothness and noise reduction.
- **Low σ**:
    - Retains more details but may lead to aliasing during downsampling.
    - Suitable for applications requiring high detail preservation.

Determining the optimal sigma(σ)is critical, as it directly affects the pyramid's ability to balance noise suppression and detail retention. In this project, sigma(σ) analysed using FFT and radial frequency methods to select the best value for reconstruction (details in Section 2.2).

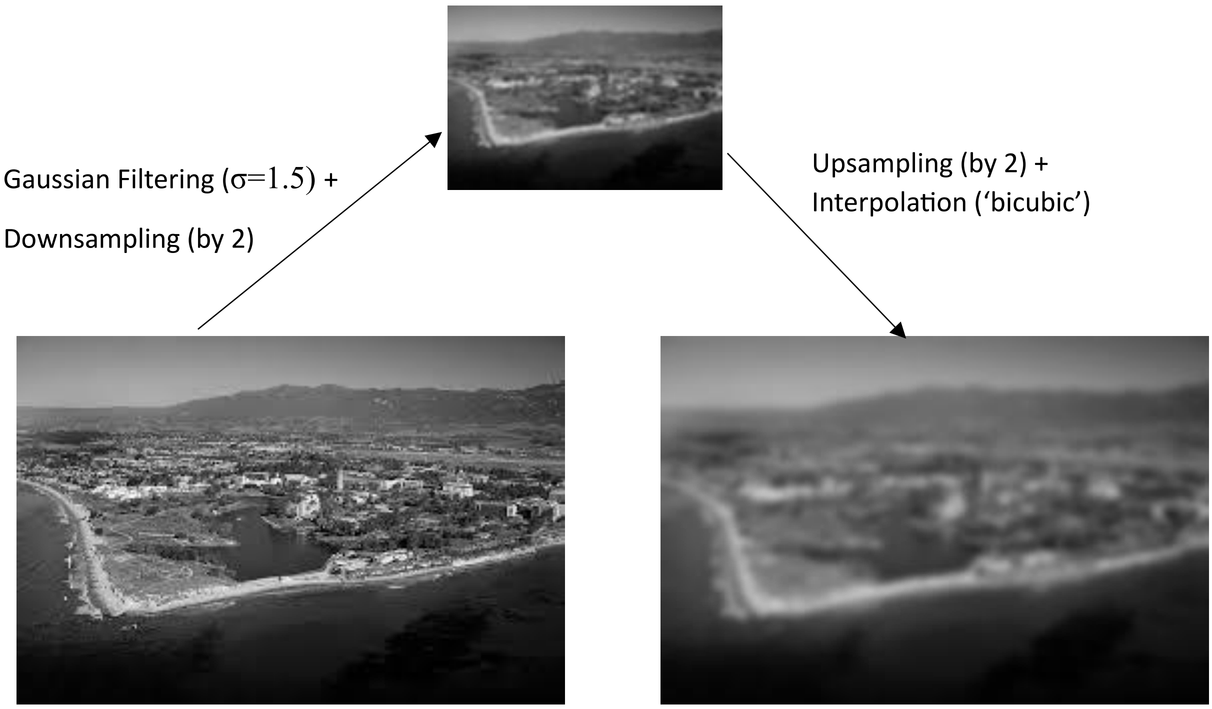


**Fig 1: Gaussian Pyramid (1- Level)**

**Applications**

Gaussian Pyramids are extensively used in tasks such as image blending, texture mapping, and progressive transmission. They are also foundational in many multi-scale vision systems, providing an efficient framework for image representation at different resolutions [6].

## 3.2 Max-Pooling:

The Max Pooling Pyramid [8] is a hierarchical structure that captures the most prominent features of an image by selecting the maximum pixel values within defined regions. This approach effectively reduces the spatial dimensions of the image while preserving essential information, making it valuable in various image processing tasks, including feature extraction and noise reduction. **Fig 2.** depicts the base layer of a Max-Pooling Pyramid.

**Construction of a 1-Level Max Pooling Pyramid:**

1. **Max Pooling (Downsampling):**
    - Divide the input image into non-overlapping blocks of a specified size, commonly 2×2 pixels.
    - For each block, identify and retain the maximum pixel value, discarding the others.
    - This process results in a downsampled image that retains the most significant features from each region.

*Example:*

Consider the following 4×4 pixel image:

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 2 & 4 & 3 \\ 1 & 0 & 6 & 5 \end{bmatrix}$$

Applying 2×2 max pooling yields:

$$\begin{bmatrix} 6 & 8 \\ 9 & 6 \end{bmatrix}$$

Here, each value in the resulting matrix represents the maximum value from each 2×2 block of the original image.

This method effectively reduces the image size while preserving the most prominent features within each region.

**Reconstruction from a 1-Level Max Pooling Pyramid:**

1. **Upsampling:**
   - Expand the downsampled image to the original resolution by replicating each pixel value to form a block corresponding to the original pooling size (e.g., each pixel in the downsampled image becomes a 2×2 block in the upsampled image).

*Example:*

Upsampling the previous 2×2 matrix:

$$\begin{bmatrix} 6 & 8 \\ 9 & 6 \end{bmatrix}$$

Results in:

$$\begin{bmatrix} 6 & 6 & 8 & 8 \\ 6 & 6 & 8 & 8 \\ 9 & 9 & 6 & 6 \\ 9 & 9 & 6 & 6 \end{bmatrix}$$

1. **Interpolation:**
   - Apply interpolation techniques to smooth the upsampled image and approximate the original pixel values. Common methods include:
     - **Nearest-Neighbour Interpolation:** Assigns the value of the nearest pixel to the new pixels.
     - **Bilinear Interpolation:** Calculates the weighted average of the four nearest pixels.
     - **Bicubic Interpolation:** Uses the weighted average of the 16 nearest pixels for smoother results.

The choice of interpolation method affects the quality of the reconstructed image, with more sophisticated methods like bicubic interpolation generally providing better approximations of the original image.

**Importance:**

The Max Pooling Pyramid emphasizes dominant features and reduces spatial dimensions, which is beneficial in feature extraction and noise reduction. This approach is widely used in convolutional neural networks (CNNs) to downsample feature maps, reducing computational complexity while preserving important information. By focusing on the most significant features within each region, max pooling helps in achieving translation invariance and robustness to noise.


**Fig 2: Max-Pooling Pyramid (1-Level)**

For a comprehensive understanding of max pooling operations and their applications in image processing, refer to the article "CNN | Introduction to Pooling Layer" on GeeksforGeeks [8].

## 3.3 Max-Averaging Pyramid:

The **Max-Averaging Pyramid** is a modified version of the traditional Max Pooling Pyramid that combines the principles of **max pooling** and **average pooling**. While there isn't an established name for this method in the literature, it is referred to as **Max-Averaging** in this project to highlight its hybrid nature. This approach aims to balance the strengths of max pooling (preserving prominent features) and average pooling (smoothness) for a more comprehensive image representation. **Fig 3.** Shows the fundamental layer of a Max-Averaging Pyramid.

**Construction of a 1-Level Max-Averaging Pyramid**

1. **Pooling (Downsampling)**:
   - Divide the input image into non-overlapping blocks (e.g., 2×2).
   - For each block, compute:
     - The **maximum value**, which retains the most significant feature.
     - The **average value**, which captures the overall intensity trend.

   Combine the max and average values using a weighted formula:

$$P_{\text{combined}} = \alpha \cdot P_{\max} + (1 - \alpha) \cdot P_{\text{avg}}$$

Here, $\alpha$ ($0 \leq \alpha \leq 1$) controls the balance between max and average pooling.

The resulting value $\mathbf{P_{combined}}$ becomes the representative pixel for the block in the downsampled image.

*Example:*

Consider the following $4 \times 4$ pixel image:

$$\begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 2 & 4 & 3 \\ 1 & 0 & 6 & 5 \end{bmatrix}$$

For a $2 \times 2$ block size and $\alpha = 0.5$:

- Max value of the first block $(1, 3, 5, 6)$: $6$
- Average value of the first block: $3.75$
- Combined value:

$$P_{\text{combined}} = 0.5 \cdot 6 + 0.5 \cdot 3.75 = 4.875$$

After processing all blocks, the downsampled image becomes:

$$\begin{bmatrix} 4.875 & 5.5 \\ 5.5 & 4.5 \end{bmatrix}$$

This process ensures that both edge details and overall intensity trends are preserved.

**Reconstruction from a 1-Level Max-Averaging Pyramid**

1. **Upsampling**:
   - Expand the downsampled image to the original resolution by replicating each pixel value to form a block corresponding to the pooling size (e.g., 2×22 \times 22×2).
2. **Interpolation**:
   - Apply bicubic interpolation (or other methods) to smooth the transitions between repeated values and approximate the original resolution.

*Example:* Upsampling the $2 \times 2$ downsampled image:

$$\begin{bmatrix} 4.875 & 5.5 \\ 5.5 & 4.5 \end{bmatrix}$$

Results in:

$$\begin{bmatrix} 4.875 & 4.875 & 5.5 & 5.5 \\ 4.875 & 4.875 & 5.5 & 5.5 \\ 5.5 & 5.5 & 4.5 & 4.5 \\ 5.5 & 5.5 & 4.5 & 4.5 \end{bmatrix}$$

**Importance of α=0.5**

The choice of α\alphaα significantly impacts the balance between sharpness and smoothness:

- At α=0.5, equal weight is given to max pooling and average pooling, ensuring that:
    - **Prominent features (e.g., edges)** are preserved.
    - **Overall smoothness** is maintained, reducing blocky artifacts.

This mid-point weighting is particularly useful in applications where both edge retention and smooth transitions are crucial. For example:

- In **Image compression**, α=0.5 balances detail preservation with efficient representation.
- In **multi-scale analysis**, it provides a general-purpose pyramid structure suitable for diverse image types.

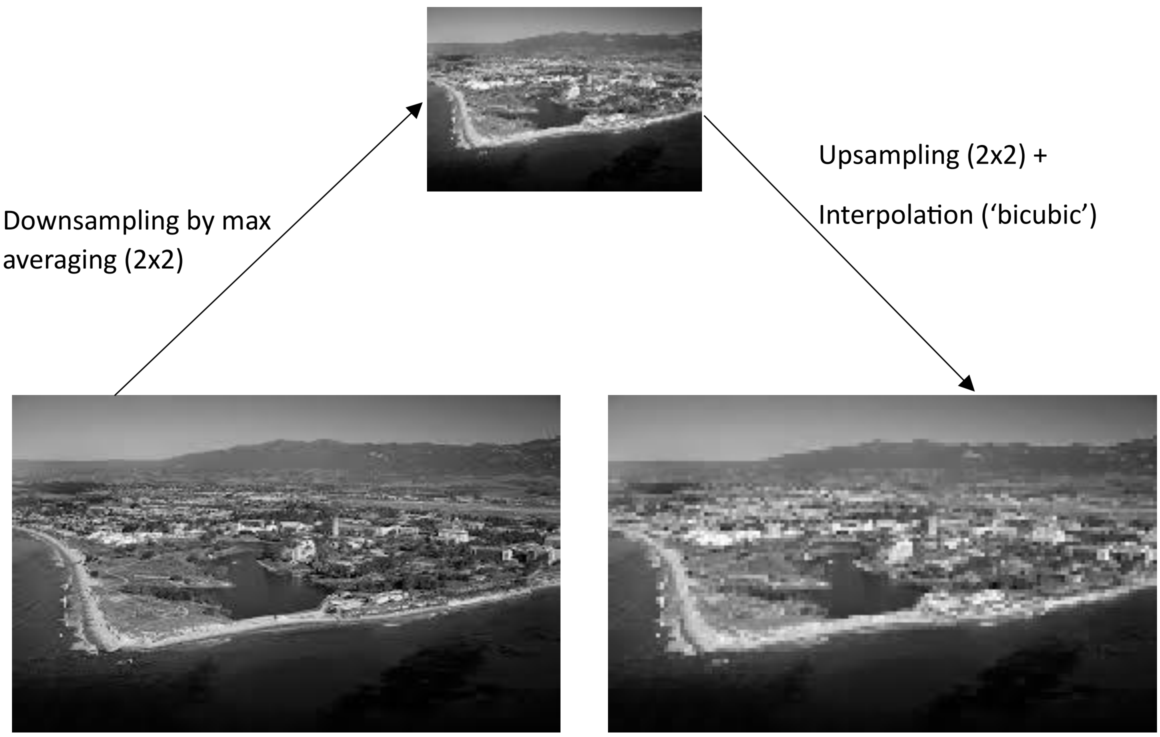


**Fig 3: Max-Averaging Pyramid (1-Level)**

**Applications:**

The Max-Averaging Pyramid is well-suited for scenarios requiring a balance between detail retention and smoothness. By adjusting α, this method can adapt to specific application needs, making it versatile for tasks like feature extraction, compression, and reconstruction.

# 4. Analysis Framework:

**4.1 Structural Similarity Index Measure (SSIM):**

**Objective**: To quantitatively assess the similarity between the original and reconstructed images, thereby evaluating the effectiveness of each pyramid method in preserving image quality.

**Rationale for Choosing SSIM**: Traditional metrics like Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) focus on pixel-wise differences, which may not align with human visual perception. SSIM, however, considers structural information, luminance, and contrast, providing a perceptually relevant measure of image quality

**SSIM Formula**: The SSIM index between two image windows $x$ and $y$ is defined as:

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

where:

- $\mu_x$ and $\mu_y$: Mean intensities of $x$ and $y$.
- $\sigma_x^2$ and $\sigma_y^2$: Variances of $x$ and $y$.
- $\sigma_{xy}$: Covariance between $x$ and $y$.
- $C_1$ and $C_2$: Stabilization constants to prevent division by zero.

These components are combined to form the SSIM index, reflecting the perceived quality of the reconstructed image relative to the original.

**Application in This Study**: SSIM is computed between the original and reconstructed images for each pyramid method to quantify reconstruction fidelity. A higher SSIM value indicates better preservation of image quality.

**Standardized Test Images**: The analyses utilize two standardized images commonly employed in image processing research:

- **Lena**: A 512×512-pixel image featuring diverse textures and frequency components.
- **UCSB**: Another 512×512-pixel image with a mix of smooth and detailed regions.

These images are widely used for benchmarking image processing algorithms due to their varied content.

### 4.2 Bandwidth (Variance) Analysis for Gaussian Pyramid

**Objective**: To determine the optimal bandwidth parameter (σ) for the Gaussian Pyramid, which controls the degree of smoothing during downsampling and reconstruction.

**Methodology**:

1. **Frequency Domain Analysis**:
    - Apply Gaussian filters with varying sigma(σ)values to the original image.
    - Compute the Fast Fourier Transform (FFT) of the filtered images to analyse their frequency spectra.
    - Perform radial frequency analysis to evaluate the attenuation of high-frequency components for each σ.
2. **Visual Inspection**:
    - Assess the filtered images for visible aliasing or excessive blurring to identify the sigma(σ)value that balances detail preservation and noise reduction.

**Rationale**: Analysing the frequency response helps in selecting a sigma(σ)that effectively suppresses high-frequency noise while retaining essential image details.

### 4.3 Interpolation Method Analysis

**Objective**: To identify the most effective interpolation technique for upsampling during the reconstruction phase in each pyramid method.

**Methodology**:

1. **Reconstruction with Different Interpolations**:
    - Reconstruct images using three interpolation methods: nearest-neighbour, bilinear, and bicubic.
    - Compute SSIM between the original and reconstructed images for each method.
2. **Comparative Evaluation**:
    - Compare SSIM scores across interpolation methods to determine which provides the highest reconstruction quality for each pyramid technique.

**Rationale**: Interpolation methods vary in complexity and their ability to preserve image structures. Evaluating their performance ensures the selection of an appropriate method that maintains image fidelity during reconstruction.

### 4.4 Downsampling Factor Analysis

**Objective**: To assess the impact of different downsampling factors on the reconstruction quality of each pyramid method.

**Methodology**:

1. **Downsampling and Reconstruction**:

- Downsample the original image by factors of 2, 4, and 8.
- Reconstruct the images to their original size using the predetermined optimal interpolation method.
- Compute SSIM between the original and reconstructed images for each downsampling factor.

2. **Analysis**:
    - Examine the trend of SSIM values as the downsampling factor increases to identify the factor that offers a balance between data reduction and reconstruction quality.

**Rationale**: Larger downsampling factors reduce data size but may lead to information loss. This analysis helps in selecting a downsampling factor that minimizes quality degradation.

**4.5 Computational Efficiency Analysis**

**Objective**: To evaluate the computational performance of each pyramid method in terms of execution time and memory usage.

**Methodology**:

1. **Performance Measurement**:
    - Record the time taken for construction (downsampling) and reconstruction (upsampling) processes.
    - Monitor memory consumption during these processes using profiling tools.
2. **Comparison**:
    - Compare the execution time and memory usage across the three pyramid methods to determine their computational efficiency.

**Rationale**: Understanding the computational requirements is crucial for practical applications, especially in resource-constrained environments.

# 5. Results and Discussion:

This section presents the results of the analyses performed on the three pyramid methods: **Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**. The results are discussed with respect to the reconstruction quality (measured using **SSIM**) and computational efficiency. Each subsection corresponds to an analysis described in the **Methodology** section.

**Preprocessing: Conversion to Grayscale:**

Before analysis, the test images were converted to grayscale as shown in **Fig 4.** to simplify the data and focus on luminance, which is critical for structural similarity. Grayscale images reduce computational complexity by having a single intensity channel, aligning with SSIM's purpose of evaluating luminance and structure. This ensures uniformity and consistency across all analyses.

**Fig 4: The Original image (left) and it grayscale version (right)**

## 5.1 Bandwidth (Sigma) Analysis for Gaussian Pyramid:

This section presents the results of the **sigma analysis** performed to evaluate the effect of varying bandwidth parameters (σ) in the Gaussian Pyramid as shown in **Fig 5**. The analysis includes frequency domain analysis (FFT and radial frequency). The aim is to determine the optimal sigma(σ)that balances noise suppression and detail preservation.

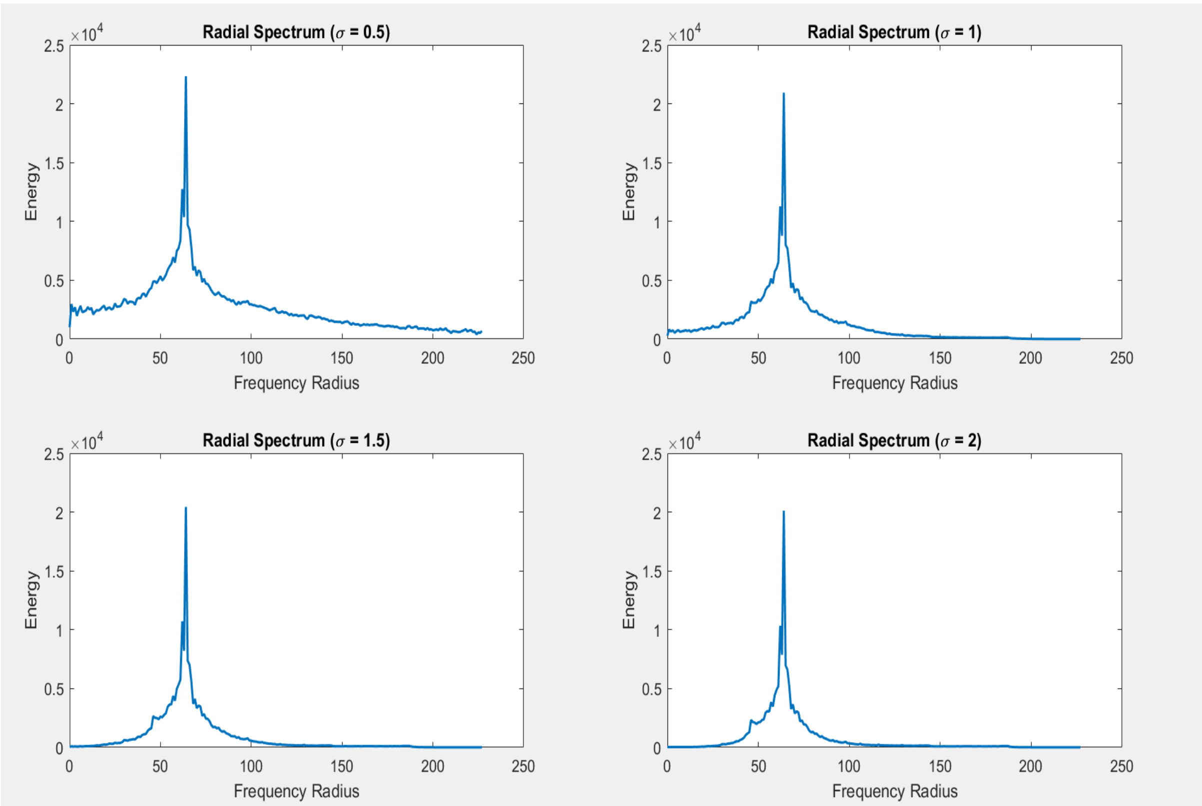


**Fig 5: Radial Frequency analysis for different variances**

**Radial Frequency Analysis (Second Figure)**:

- **Radial Spectrum (σ=0.5)**:
  - High energy is retained at higher frequency radii, indicating inadequate attenuation of high-frequency components.
- **Radial Spectrum (σ=1.0)**:

- A balanced reduction of high-frequency components is observed, with energy concentrated in the lower frequencies.

- **Radial Spectrum (σ=1.5)**:
    - Optimal suppression of high-frequency components, with most energy focused at lower frequency radii.
- **Radial Spectrum (σ=2.0)**:
    - Excessive suppression of frequency components results in significant loss of both high and mid-range frequencies.

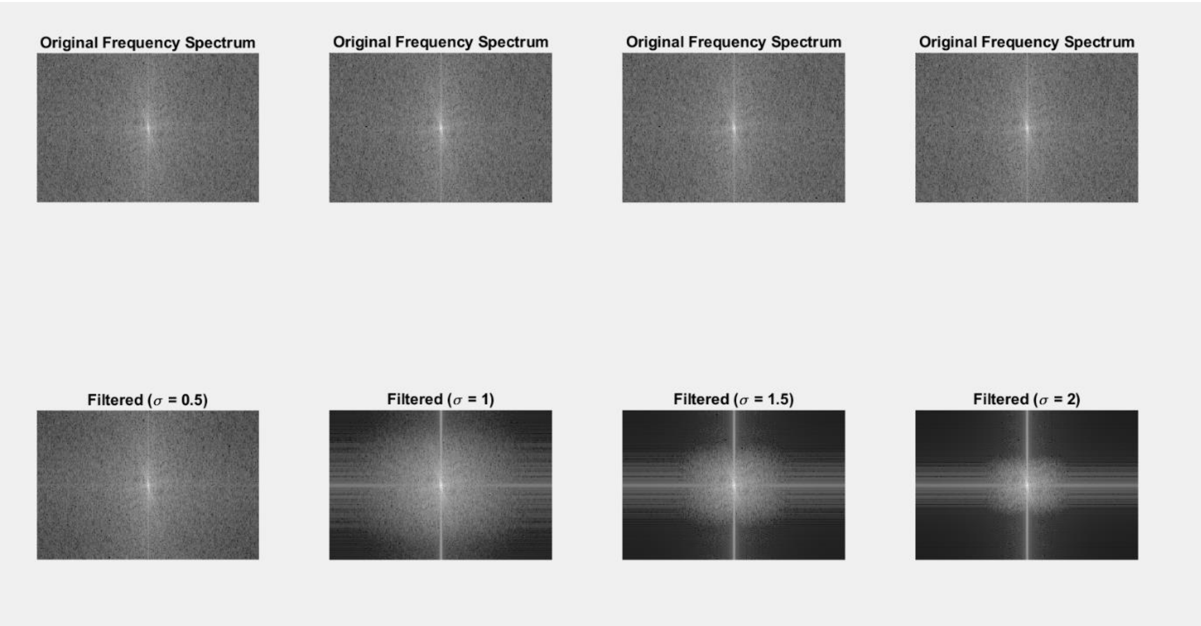


**Fig 6: FFT analysis on the given image with varying variance**

**Results of Fig 6:**

1. **Filtered Images:**
    - **Original Spectrum**:
        - The original FFT spectrum shows a wide range of high-frequency components distributed across the image.
    - **Filtered (σ=0.5)**:
        - Minimal smoothing is applied. High-frequency components remain prominent, indicating aliasing.
    - **Filtered (σ=1.0)**:
        - Moderate smoothing reduces some high-frequency noise while retaining image details.
    - **Filtered (σ=1.5)**:

    - Sufficient smoothing eliminates most high-frequency noise while preserving structural details.
  - **Filtered (σ=2.0)**:
    - Strong smoothing suppresses most high-frequency components, leading to significant detail loss.

**Discussion**

- **Frequency Domain Analysis:** The FFT spectra and radial frequency plots provide insights into the effect of varying the bandwidth parameter σ (standard deviation) in the Gaussian filter. At σ=0.5, the attenuation of high-frequency components is insufficient, leaving noticeable aliasing artifacts in the reconstructed image. As σ increases to σ=1.5, the high-frequency noise is effectively suppressed while preserving structural details, achieving a balanced trade-off between noise reduction and detail retention. At σ=2.0, the filter becomes overly aggressive, leading to excessive suppression of both high- and low-frequency components, which results in visible blurring and loss of important details.
- **Optimal σ:** Based on the FFT and radial frequency analysis, σ=1.5 is identified as the optimal value for this project. It effectively suppresses high-frequency noise while preserving sufficient detail in lower-frequency components, ensuring high reconstruction quality.

**Inference:**

- Optimal Sigma (σ): **σ = 1.5** is identified as the most suitable value for the Gaussian Pyramid in this application. It strikes a balance between:
  - Effective suppression of high-frequency noise.
  - Preservation of structural details in the image.
  - Avoidance of aliasing or over-smoothing.

This optimal σ will be used in subsequent analyses to ensure consistent and high-quality reconstruction performance.

## 5.2 Interpolation Method Analysis:

This analysis evaluates the impact of different interpolation methods—**nearest-neighbour**, **bilinear**, and **bicubic**—on the reconstruction quality of the three pyramid methods (Gaussian Pyramid, Max Pooling Pyramid, and Max-Averaging Pyramid). The goal is to determine which interpolation method preserves the structural integrity of the original image most effectively.

**Key Evaluation Metric**: The reconstruction quality is measured using the **Structural Similarity Index (SSIM)**, which quantifies the perceptual similarity between the original and reconstructed images.

```
>> Ga_I
SSIM for Gaussian Pyramid with nearest interpolation: 0.6573
SSIM for Gaussian Pyramid with bilinear interpolation: 0.6491
SSIM for Gaussian Pyramid with bicubic interpolation: 0.6586
>> ma_I
SSIM for Max Pooling Pyramid with nearest interpolation: 0.7852
SSIM for Max Pooling Pyramid with bilinear interpolation: 0.7779
SSIM for Max Pooling Pyramid with bicubic interpolation: 0.7946
>> Max_I
SSIM for Max-Averaging Pyramid with nearest interpolation: 0.8162
SSIM for Max-Averaging Pyramid with bilinear interpolation: 0.7942
SSIM for Max-Averaging Pyramid with bicubic interpolation: 0.8190
```

**Fig 7: SSIM values for different interpolation methods**

1. **Gaussian Pyramid**:
    - **Nearest-Neighbour Interpolation**:
      Nearest-neighbour interpolation provides decent reconstruction quality but is prone to blocky artifacts. The Gaussian smoothing applied during the construction phase mitigates these artifacts to some extent, leading to its reasonable SSIM score.
    - **Bilinear Interpolation**:
      Bilinear interpolation smooths transitions between pixels but results in slightly lower SSIM compared to nearest-neighbour and bicubic interpolation. This suggests that the increased smoothness compromises some of the structural details.
    - **Bicubic Interpolation**:
      Bicubic interpolation achieves the highest SSIM for the Gaussian Pyramid, slightly outperforming both nearest-neighbour and bilinear methods. This indicates that it balances smooth transitions and structural detail preservation more effectively.

**Inference**:
For the Gaussian Pyramid, bicubic interpolation is the most effective method, achieving the best balance between smoothness and structural integrity during reconstruction. **Fig 8** shows the implementation of **gaussian pyramid** using the best interpolation method (‘Bicubic’).

**Fig 8: The original image (left) and the reconstructed image (right)**

2. **Max Pooling Pyramid**:

   - **Nearest-Neighbour Interpolation**:
     Nearest-neighbour interpolation preserves sharp edges but introduces noticeable blocky artifacts, which reduce visual smoothness.

   - **Bilinear Interpolation**:
     Bilinear interpolation improves the smoothness of transitions over nearest-neighbour but compromises slightly on the sharpness of edges, leading to a moderate SSIM.

   - **Bicubic Interpolation**:
     Bicubic interpolation achieves the highest SSIM for the Max Pooling Pyramid by effectively mitigating blocky artifacts and preserving dominant features.

**Inference**:
Bicubic interpolation provides the best reconstruction quality for the Max Pooling Pyramid, offering a balanced compromise between edge sharpness and transition smoothness. **Fig 9** shows the implementation of **Max-Pooling pyramid** using the best interpolation method (‘Bicubic’).

**Fig 9: The original image (left) and the reconstructed image (right)**

3. **Max-Averaging Pyramid**:

   - **Nearest-Neighbour Interpolation**:
     Nearest-neighbour interpolation performs well in preserving dominant features but introduces blocky artifacts, which reduce overall perceptual similarity.

   - **Bilinear Interpolation**:
     Bilinear interpolation smooths transitions but slightly blurs finer details, leading to a marginally lower SSIM compared to nearest-neighbour and bicubic methods.

   - **Bicubic Interpolation**:
     Bicubic interpolation achieves the highest SSIM, as it enhances smoothness without significantly compromising the sharpness of prominent features. This aligns well with the hybrid nature of the Max-Averaging Pyramid.

**Inference**:
Bicubic interpolation is the most suitable method for the Max-Averaging Pyramid, providing the best trade-off between structural sharpness and smooth transitions. **Fig 10** shows the implementation of **Max-Averaging pyramid** using the best interpolation method ('Bicubic').

**Fig 10: The original image (left) and the reconstructed image (right)**

**General Inference:**

- **Bicubic interpolation** consistently achieves the **highest SSIM** across all three pyramid methods, demonstrating its superior ability to balance structural detail preservation and smooth transitions during reconstruction.
- Nearest-neighbour interpolation is computationally efficient but prone to blocky artifacts.
- Bilinear interpolation is a middle-ground approach but is outperformed by bicubic interpolation in terms of structural similarity.

For all subsequent analyses, **bicubic interpolation** is selected as the standard interpolation method due to its consistent performance across all pyramid methods.

## 5.3 Downsampling Factor Analysis:

This analysis investigates the impact of varying downsampling factors (**2, 4, and 8**) on the reconstruction quality of the three pyramid methods: **Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**. The goal is to evaluate how the reduction in spatial resolution during construction affects the ability to reconstruct the original image.

**Key Evaluation Metric**:
The reconstruction quality is assessed using the **SSIM** to measure the structural similarity between the original and reconstructed images.

```
>> Ga_S
SSIM for Gaussian Pyramid with downsampling factor=2: 0.6586
SSIM for Gaussian Pyramid with downsampling factor=4: 0.6127
SSIM for Gaussian Pyramid with downsampling factor=8: 0.5302
>> Ma_S
SSIM for Max Pooling Pyramid with downsampling factor=2: 0.7946
SSIM for Max Pooling Pyramid with downsampling factor=4: 0.6200
SSIM for Max Pooling Pyramid with downsampling factor=8: 0.4969
>> Max_S
SSIM for Max-Averaging Pyramid with downsampling factor=2: 0.8190
SSIM for Max-Averaging Pyramid with downsampling factor=4: 0.6545
SSIM for Max-Averaging Pyramid with downsampling factor=8: 0.5397
```

**Fig 11: SSIM values for different Downsampling Factors**

1. **Gaussian Pyramid**:
   - **Downsampling Factor = 2**:
     The SSIM is the highest at this factor, indicating that the Gaussian Pyramid retains sufficient image details while reducing resolution moderately.
   - **Downsampling Factor = 4**:
     The SSIM drops compared to factor 2, as more information is lost during downsampling.
   - **Downsampling Factor = 8**:
     The SSIM further decreases significantly, reflecting a substantial loss of structural details due to aggressive downsampling.

**Inference**:
The Gaussian Pyramid performs best at a downsampling factor of 2, as it balances resolution reduction and detail preservation. Larger factors lead to substantial degradation in reconstruction quality. **Fig 12.** Shows the downsampled images for factors 2,4,8 along with the original image going right.

**Fig 12: Depicts the Gaussian pyramid for different dowsampling factors of 2,4,8**

2. **Max Pooling Pyramid**:
   - **Downsampling Factor = 2**:
     The SSIM is highest, highlighting the ability of max pooling to retain dominant features while reducing resolution moderately.

- **Downsampling Factor = 4**:
  A notable drop in SSIM occurs, indicating a loss of finer details as the pooling operation aggregates more pixels.
- **Downsampling Factor = 8**:
  The lowest SSIM reflects the limitations of max pooling at extreme downsampling levels, where critical details are lost.

**Inference**:
The Max Pooling Pyramid achieves the best reconstruction quality at a downsampling factor of 2. Larger factors lead to diminishing returns, as blocky artifacts and detail loss become more pronounced. **Fig 13.** Shows the downsampled images for factors 2,4,8 along with the original image going right.

**Fig 13: Depicts the Max-Pooling pyramid for different dowsampling factors of 2,4,8**

3. **Max-Averaging Pyramid**:
   - **Downsampling Factor = 2**:
     The highest SSIM is observed, indicating that the hybrid pooling method effectively balances edge retention and smoothness at moderate resolution reduction.
   - **Downsampling Factor = 4**:
     SSIM decreases due to increased loss of finer details, as the averaging component dilutes edge information.
   - **Downsampling Factor = 8**:
     The SSIM drops further, reflecting significant degradation in structural details at this extreme downsampling level.

**Inference**:
The Max-Averaging Pyramid performs best at a downsampling factor of 2, where it maintains a balance between sharpness and smoothness. Higher factors lead to excessive smoothing and loss of structural fidelity. **Fig 14.** Shows the downsampled images for factors 2,4,8 along with the original image going right.

**Fig 14: Depicts the Max-Averaging pyramid for different dowsampling factors of 2,4,8**

**General Inference**

- Across all three pyramid methods, a downsampling factor of 2 consistently provides the best reconstruction quality, as measured by SSIM. This factor effectively reduces resolution while preserving critical image details.
- At higher downsampling factors (4 and 8), significant structural degradation occurs across all methods due to excessive loss of information.

For subsequent analyses, **a downsampling factor of 2** will be used as the standard to ensure optimal reconstruction quality and fair comparisons across methods.

## 5.4 Computational Efficiency Analysis:

This analysis evaluates the computational efficiency of the three pyramid methods—**Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**—in their most optimized configurations. By fixing key parameters to their optimal values as determined in previous analyses, we aim to compare the time taken and memory consumed during the construction (downsampling) and reconstruction (upsampling) processes.

```
>> Ga_T
Gaussian Pyramid:
Time for Downsampling: 0.0023 seconds
Time for Reconstruction: 0.0028 seconds
Memory Usage: 11.42 MB
>> ma_T
Max Pooling Pyramid:
Time for Downsampling: 0.0271 seconds
Time for Reconstruction: 0.0009 seconds
Memory Usage: 11.45 MB
>> max_T
Max-Averaging Pyramid:
Time for Downsampling: 0.0628 seconds
Time for Reconstruction: 0.0008 seconds
Memory Usage: 11.64 MB
```

**Fig 15: Results of Computational Analysis for different pyramids**

1. **Gaussian Pyramid**:

- **Downsampling Time**: 0.0023 seconds
  The Gaussian Pyramid is the fastest for downsampling due to the simplicity of applying a Gaussian filter followed by a subsampling operation.
- **Reconstruction Time**: 0.0028 seconds
  The bicubic interpolation, combined with a second Gaussian filter for reconstruction, maintains high computational efficiency.
- **Memory Usage**: 11.42 MB
  The Gaussian Pyramid consumes the least memory due to its straightforward filtering and downsampling operations, making it the most memory-efficient method.

**Inference**:
The Gaussian Pyramid is the most computationally efficient among the three methods in terms of both execution time and memory usage. It is ideal for applications requiring speed and low resource consumption.

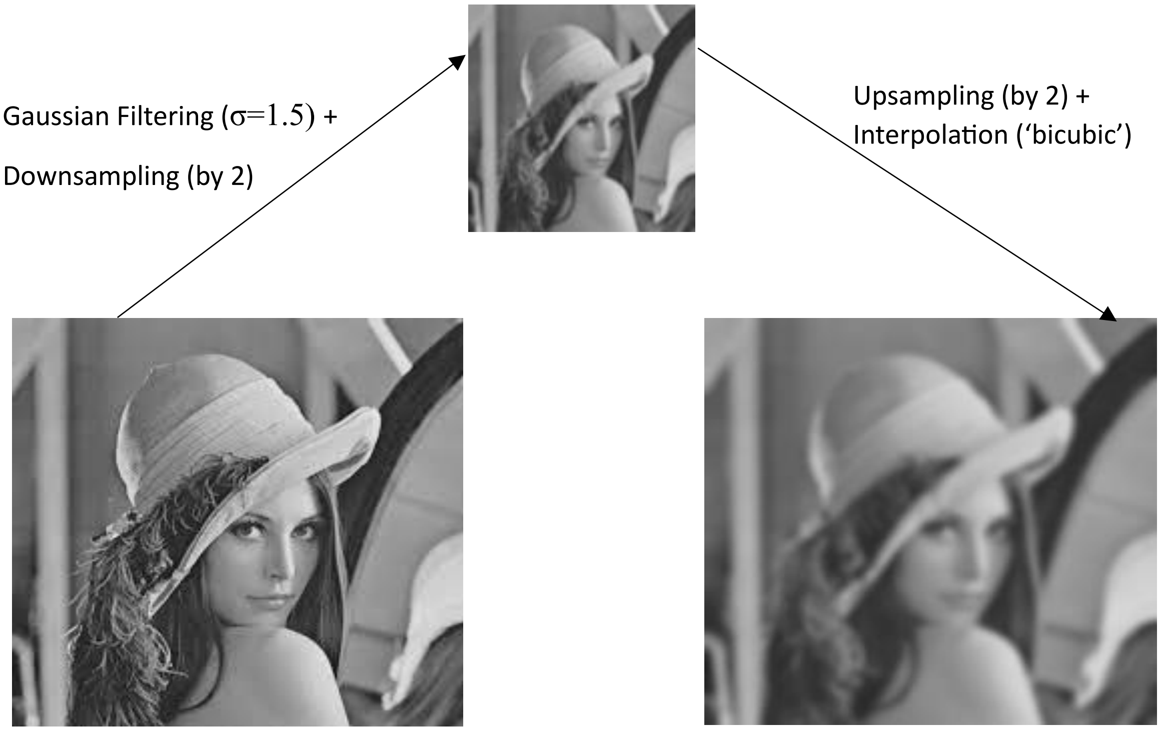


**Fig 16: The Gaussian Pyramid construction and reconstruction (1-level)**

2. **Max Pooling Pyramid**:
   - **Downsampling Time**: 0.0271 seconds
     Max pooling involves finding the maximum value within blocks, which is computationally more intensive than Gaussian filtering, leading to a higher downsampling time.

- **Reconstruction Time**: 0.0009 seconds
  Reconstruction is faster due to the simplicity of bicubic interpolation without any additional filtering.
- **Memory Usage**: 11.45 MB
  Slightly higher memory usage compared to the Gaussian Pyramid, attributed to the intermediate storage of block-wise maximum values.

**Inference**:
The Max Pooling Pyramid offers moderately high computational efficiency during reconstruction but is slower during downsampling. Its ability to retain edge details comes at the cost of slightly higher memory usage and longer downsampling time.


**Fig 17: The Max-Pooling Pyramid construction and reconstruction (1-level)**

3. **Max-Averaging Pyramid**:
   - **Downsampling Time**: 0.0628 seconds
     The hybrid nature of max-averaging pooling, involving both maximum and average calculations, makes it the most computationally intensive method during downsampling.
   - **Reconstruction Time**: 0.0008 seconds
     Like the Max Pooling Pyramid, reconstruction is fast due to the use of bicubic interpolation without additional filtering.

- **Memory Usage**: 11.64 MB
  The highest memory usage among the three methods, attributed to the additional storage required for both max and average values during pooling.

**Inference**:
The Max-Averaging Pyramid incurs the highest computational cost in terms of time and memory for downsampling, but it offers a better balance of detail retention and smoothness. Reconstruction remains efficient due to the use of bicubic interpolation.


**Fig 18: The Max-Averaging Pyramid construction and reconstruction (1-level)**

**General Inference and Trade-offs**

1. **Gaussian Pyramid**:
   - **Strengths**: Fastest and most memory-efficient, suitable for applications where computational resources are constrained.
   - **Trade-off**: While efficient, it may not retain sharp edges and fine details as effectively as the other methods.
2. **Max Pooling Pyramid**:
   - **Strengths**: Good edge retention and moderately efficient reconstruction.
   - **Trade-off**: Higher downsampling time and memory usage compared to the Gaussian Pyramid.
3. **Max-Averaging Pyramid**:

- **Strengths**: Best reconstruction quality due to the hybrid pooling approach.
- **Trade-off**: Highest computational cost, making it less suitable for time-critical or resource-constrained applications.

# 6. Conclusion:

This project systematically analysed the reconstruction quality of three image pyramid methods—**Gaussian Pyramid**, **Max Pooling Pyramid**, and **Max-Averaging Pyramid**—using their 1-level (base layer) implementations. By focusing on reconstruction as the core metric, the study offered insights into the strengths, limitations, and trade-offs associated with each method, along with their suitability for specific applications.

**Summary of Key Findings**

1. **Gaussian Pyramid**:
    - Exhibited the fastest processing times and lowest memory usage, making it the most computationally efficient method.
    - Reconstruction quality improved significantly with bicubic interpolation, achieving the highest SSIM among interpolation methods.
    - While efficient, the Gaussian Pyramid struggled to preserve fine details, particularly at higher downsampling factors.
    - Ideal for real-time or resource-constrained applications requiring moderate reconstruction quality.
2. **Max Pooling Pyramid**:
    - Demonstrated a strong ability to retain dominant image features, such as edges, while being moderately efficient in terms of computation.
    - Bicubic interpolation mitigated blocky artifacts, yielding the best reconstruction quality among interpolation methods for this pyramid.
    - Downsampling introduced blocky artifacts at higher factors, reducing SSIM and structural fidelity.
    - Suitable for feature extraction and general-purpose image processing where edge retention is prioritized.
3. **Max-Averaging Pyramid**:
    - Achieved the best reconstruction quality across all analyses due to its hybrid approach, balancing edge sharpness and smoothness.
    - Consistently attained the highest SSIM scores, particularly when paired with bicubic interpolation.
    - The method was computationally intensive, requiring the most time and memory among the three pyramids.

- Best suited for high-quality reconstruction tasks in applications like image compression, restoration, and medical imaging, where computational resources are less of a constraint.

**Overall Trade-offs and Recommendations**

The choice of pyramid method depends on the specific application requirements:

- **Gaussian Pyramid**: Recommended for applications where speed and low memory consumption are critical, such as real-time image processing.
- **Max Pooling Pyramid**: Offers a balance between computational cost and reconstruction quality, making it suitable for edge-based feature extraction.
- **Max-Averaging Pyramid**: The preferred method for applications requiring high reconstruction quality, where computational efficiency is secondary, such as medical imaging or high-resolution visualizations.

This study highlights the importance of parameter tuning and methodical analysis in multi-rate image processing. The evaluation of bandwidth (σ), interpolation techniques, downsampling factors, and computational efficiency ensures a comprehensive comparison and enables informed decisions regarding the most suitable pyramid method for a given task.

# 7. References:


[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004.

[2] P. Burt and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 532-540, April 1983.

[3] M. D. Zeiler and R. Fergus, "Stochastic Pooling for Regularization of Deep Convolutional Neural Networks," in *Proceedings of the International Conference on Learning Representations*, 2013.

[4] D. Yu, H. Wang, P. Chen, and Z. Wei, "Local Importance-Based Pooling," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 3354-3363.

[5] A. Hore and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in *Proceedings of the 20th International Conference on Pattern Recognition*, 2010, pp. 2366-2369.

[6] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989.

[7] H. Schilling, M. Venugopal, and D. Doermann, "Comparison of Downsampling Techniques for Document Image Analysis," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 876-888, 2014.

[8] *CNN | Introduction to Pooling Layer*, GeeksforGeeks, Accessed [online].