- What are the motivations for computing interest points? What are they typically used for?
    - Robustness to Scale and Rotation
    - Distinctiveness
    - Reducing Data
    - Localization
    - Efficiency

Used in:

- Objection Recognition
- Image stitching and panorama creation
- Image registration
- Feature tracking
- 3D Reconstruction
- Augmented Reality
- Camera calibration

- Describe three common interest point detectors including their mathematical definitions.

- Why is scale selection an important operation?

- **Scale invariance:** Objects in the real world can appear at different sizes or scales in images due to variations in distance, viewpoint and imaging conditions. Scale selection ensures that the same object or feature can be detected and described consistently across different scales, making computer vision algorithms more robust and invariant to scale changes
- **Multiscale analys:** Many computer vision trasks, such as objection recognition, object tracking and image matching require analyzing images at multiple scales to account for different object sizes and distances. Scale selection enables the efficient exploration of various scales to identify relevant features or objects
- **Localization accuracy:** Accurate localization of keypoints of features depend on selecting an appropriate scale. By choosing the right scale, you can pinpoint the center or extrent of a feature more precisely, improving the reliability of subsequent processing steps.
- **Efficient processing:** Without scale selection, a computer vision algorithm would need to search for features at all possible scales, which is computationally expensive and unnecessary. Scale selection helps narrow down the search space and focus computation sources on the most relevant scales
- **Feature matching:** In matching tasks, such as image-stitching or object recognition, features extracted from different images must be matched. Scale-invariant features extracted at corresponding scales are essential for accurate and reliable matching.

- Describe how scale selection can be performed in practice.

Scale-space extrema detection algorithm:

1. Given a scale-range $[t_{min}, t_{max}]$, distribute a set of scale levels uniformly in terms of effective scale $\tau = \log t$.
2. Convolve image by a Gaussian kernel to each scale level.
3. For every image point, compute approximations of necessary derivatives and combine these into the desired measures, e.g. $\nabla^2_{norm} L$.
4. Detect local extrema over scale and space in $3 \times 3 \times 3$ neighbourhoods and by thresholding on the magnitude of the response.
5. Optionally, sort the interest points in decreasing order with respect to their scale-normalized magnitude values.

- What is the motivation for using image pyramids in computer vision?
- The motivation for using image pyramids in computer vision is primarily to address the challenges posed by scale variations in images. Image pyramids are multi-scale representations of an image created by resampling or resizing the original image at multiple scales. Here are the key motivations for using image pyramids in computer visions.
    - Scale invations
    - Objust object detections
    - Feature extraction
    - Image alignment
    - Efficient processing
    - Improved recognition and matching
    - Handling multiple resolutions
    - Texture and Pattern analysis

- How are image pyramids computed from image data?
- Start with original image
- Apply Gaussian Smoothing
- Scale-space representation
- Blob or keypoint detection
- Keypoint localization
- Scale selection Criteria
- Non-maximum suppression
    -
- Resulting keypoints

- Describe a basic trade-off issue that arises in hybrid pyramids.
- **Detail preservation:** To maintain fine details in the image, especially in the high frequency components, the hybrid should retain as much high-frequency information as possible. This allows for accurate reconstruction of the original image from the pyramids levels, preserving fine textures, edges and other intricate details.

- **Storage efficiency:** To save storage space or transmission bandwidth, it may be desirable to minize the amount of high frequency information store. High frequency details often require more storage or data bits, and in some cases, it may be impractical to store or transmit the full high-frequnecy content.

- What is the purpose of computing image descriptors at interest points?
- The purpose of computing image descriptors at interests points is to characterize and represent the local visual information surrounding those points. Image description provide a compact and informative representation of the local image content at specific locations in an image.

- **How is the SIFT descriptor defined from image data?**
- The scale-invariant feature transform description is defined from image data through a series of steps that capture local image information around a keypoint or interest points.
    - **Keypoint detection**
        - Start by detecting keypoints
    - **Local region extraction**
        - For each keypoint, define a local imag eregion around the keypoints location. This region is called the “keypoint neighbourhood” or local region”
    - **Subregion division**
        - Divide the local region into smaller subregions or bins. The local region is often divided into a grid of subregions typically 4x4 or 8x8 subregions
    - **Orientation Assignment:**
        - Compute the dominant orientation or orientation histogream for the keypoint neighborhood. This is done by examining the gradient orientations of pixels within the local region. The dominant orientation is the peak in the histogram
    - **Rotation Compensation:**
        - To achieve rotation invariance, rotate the local region and gradient orientations by the negative of the dominant orientation. This step aligns the keypoint neighborhood with a canonical orientation.
    - **Gradient Calculation:**
        - Calculate the gradient magnitude and orientation for each pixel within each subregion. The gradient represents the rate of change of pixel intensities in the x and y directions.

    - **Histogram of Oriented Gradients (HOG):**
        - Create histograms of gradient orientations for each subregion. Each subregion contributes to a histogram of gradient orientations. Typically, the histogram bins cover a 360-degree range and may be quantized into, for example, 8 or 16 bins.

- **Descriptor Vector Formation:**
    - Concatenate the histograms of gradient orientations from all subregions to form a single, high-dimensional descriptor vector. The resulting descriptor is a representation of the local image information surrounding the keypoint.

- **Normalization:**
    - Normalize the descriptor vector to make it robust to changes in illumination and contrast. Common methods include L2 normalization or clipping of large gradient values.

- **Descriptor Clipping (Optional):**
    - Threshold or clip the values in the descriptor to ensure it is not sensitive to extreme gradient variations.

- **Final Descriptor:**
    - The resulting normalized and possibly clipped descriptor vector is the SIFT descriptor for the selected keypoint. It is typically a high-dimensional feature vector that captures the local image characteristics around the keypoint.

- How is the SURF descriptor defined from image data?

- Outline the basic steps in an algorithm that matches interest points with associated image descriptors between two images of the same scene