

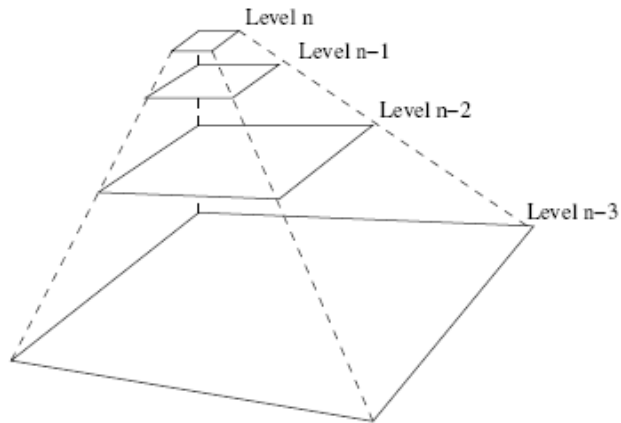
Local descriptors

- Features for local regions in the image
 - Regions obtained by segmentation
 - Regions of interest (RoI) – around interest points (keypoints)
- Interest points: corners, edges and others
- Keypoints: points in images, which are invariant to image translation, scale and rotation, and are minimally affected by noise and small distortions
- Scale-invariant feature transform (SIFT) by David Lowe

SIFT: main steps

1. Scale-space peak selection
 - Using Difference-of-Gaussians (DoG)
2. Keypoint localization
 - Elimination of unstable keypoints
3. Orientation assignment
 - Based on keypoint local image patch
4. Keypoint descriptor
 - Based upon the image gradients in keypoint local neighbourhood

Scale space



Build an image pyramid with resampling between each level



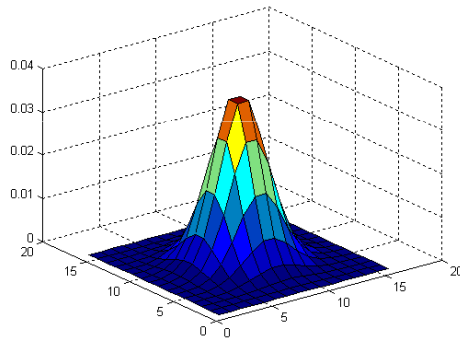
Difference-of-Gaussian

The input image is convolved with Gaussian function:

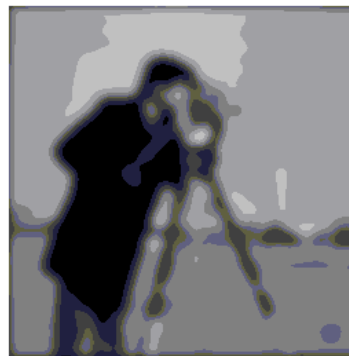
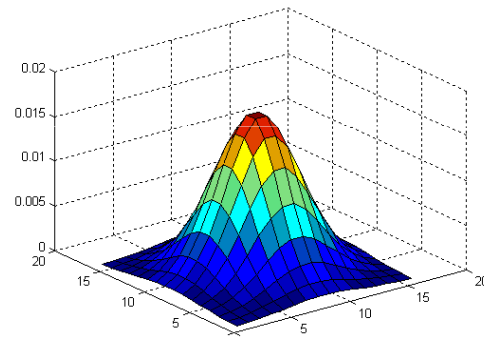
$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}$$



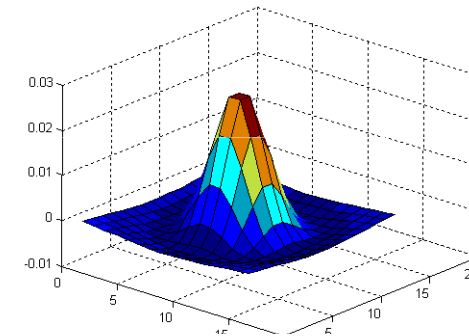
$\sigma_1 = 2$



$\sigma_2 = 4$



$\sigma_1 - \sigma_2$



Difference-of-Gaussian

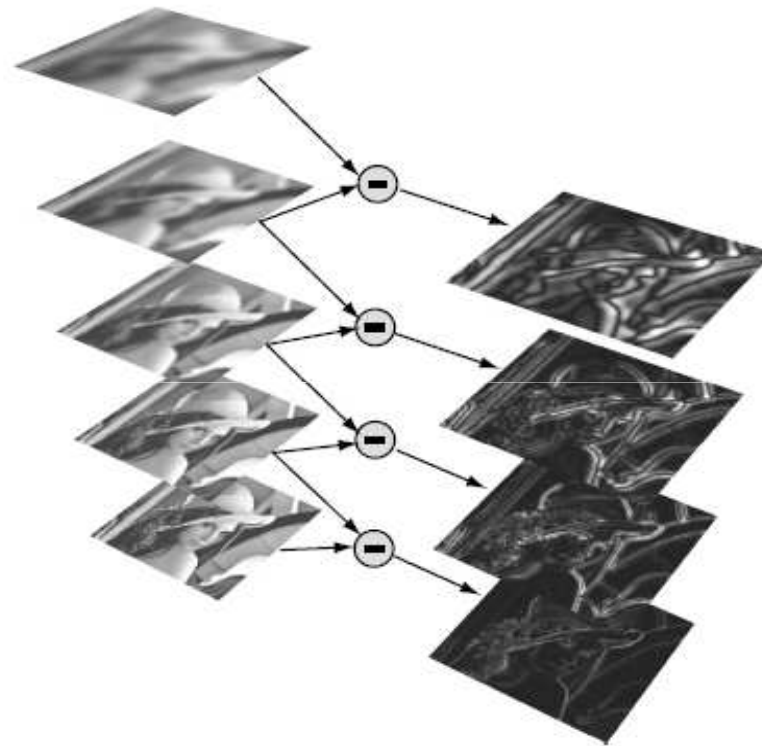
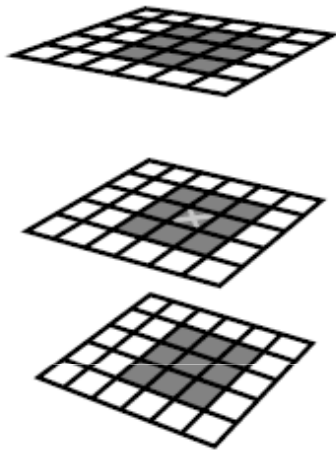


Figure 9.1: A Difference-of-Gaussian octave. The five images in the left stack are incrementally smoothed versions of the input image. The right stack shows the resulting DoG.

SIFT keypoints

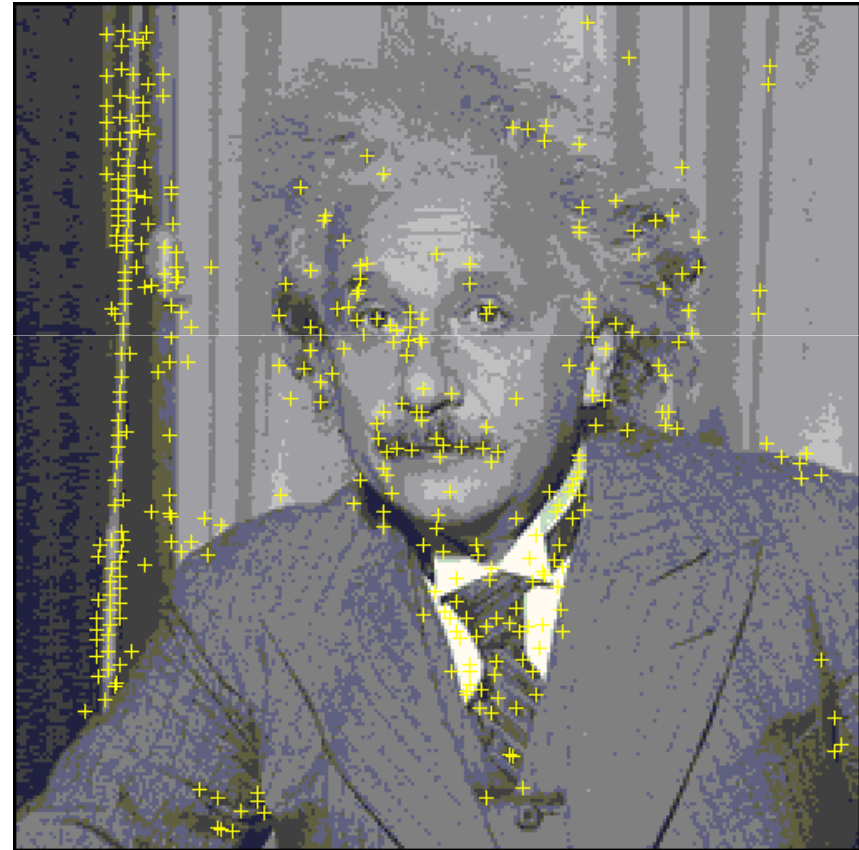
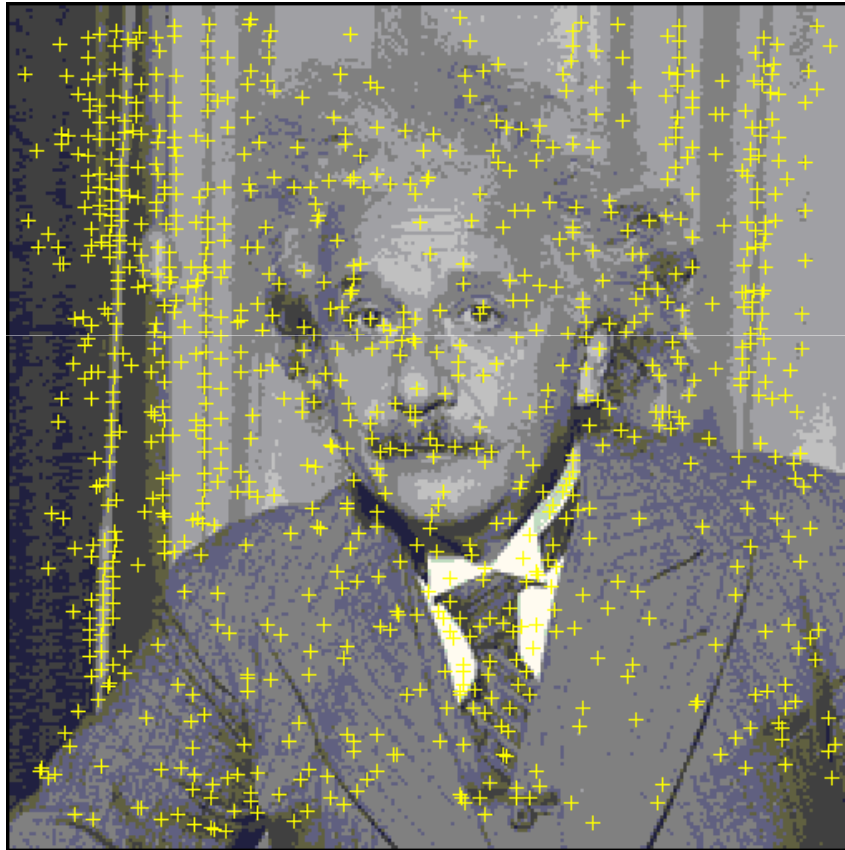


Maxima and minima of DoG applied in scale-space:

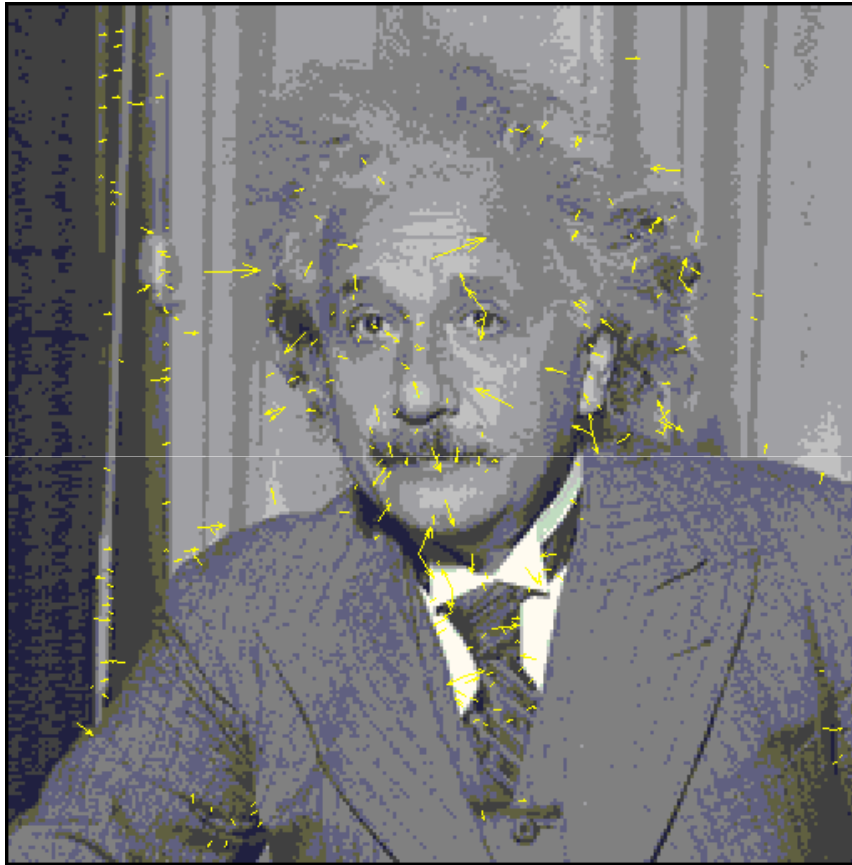
- 1) Extrema detection for the same scale
- 2) Check if it is stable for different scales



Scale-space extrema detection



Keypoints orientation and scale

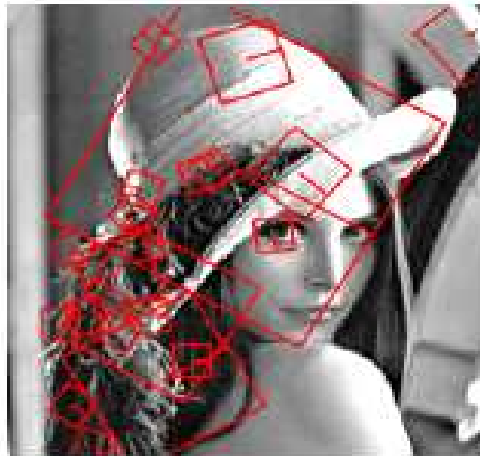


- Extract image gradients and orientations at each pixel
- Each key location is assigned a canonical orientation
- The orientation is determined by the peak in a histogram of local image gradient orientations

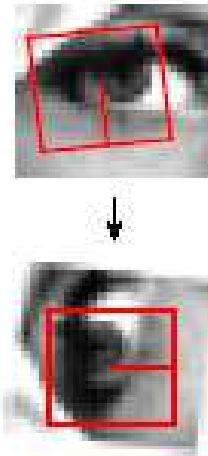
$$M_{ij} = \sqrt{(A_{ij} - A_{i+1,j})^2 + (A_{ij} - A_{i,j+1})^2}$$

$$R_{ij} = \text{atan2}(A_{ij} - A_{i+1,j}, A_{i,j+1} - A_{ij})$$

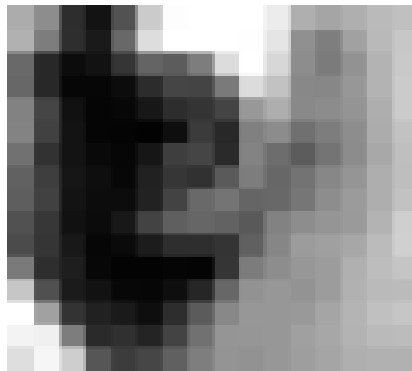
Example



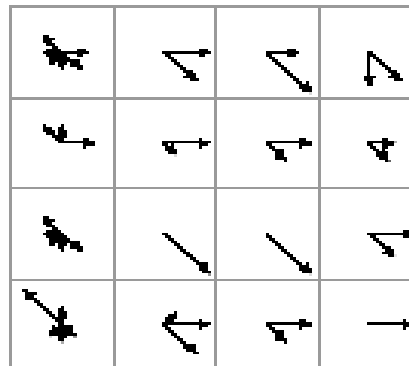
(a) A subset of the extracted interest points, and the associated regions used to create descriptors.



(b) An interest point region covering Lena's eye before and after rotation in respect to the reference orientation of the point of interest.



Interest point region



Interest point descriptor

(c) Computation of a descriptor by determining a 4×4 gradient orientation histogram array from a 16×16 pixels region around the interest point location.