8- scale invariant detectors

scale invariant detectors

① harris-laplacian= local maximums of laplacian harris corner detector in scales of the image

② SIFT (scale invariance feature transform) = local maximums of difference of gaussians in space and scale

performance = harris-laplacian > SIFT > harris

interesting point / feature = edges are not interesting, corners also not much

↳ has rich image content (brightness, color variation etc) within the local window

↳ should be invariant to image rotation and scaling

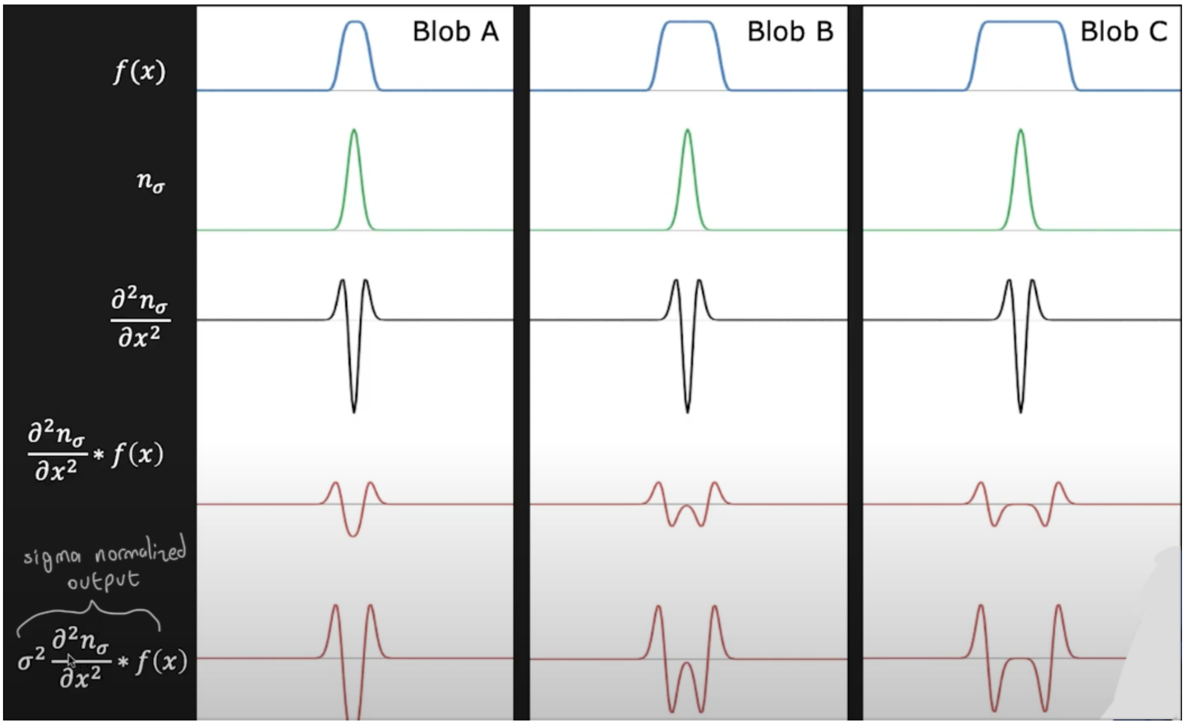
↳ should be insensetive to lighting changes

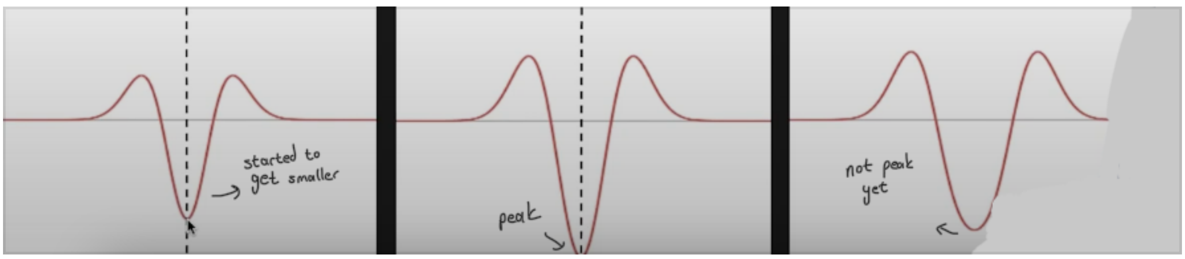
blobs (leke) as interest points = locate → determine size → determine orientation → formulate a discription that is independent of its size and orientation

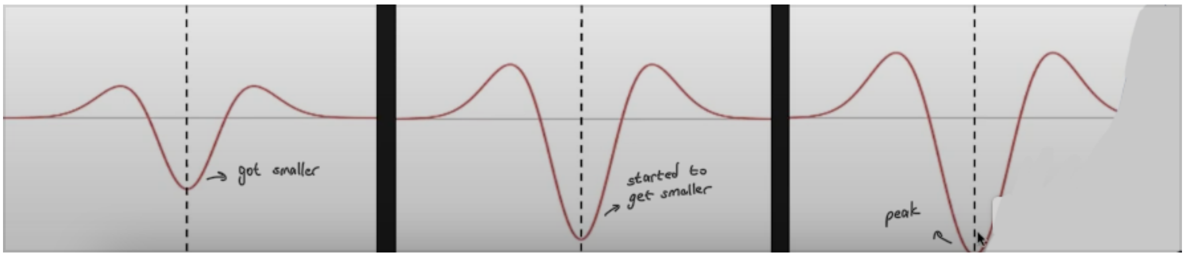
edge detection = $f=$ gaussian(n_σ) = $n_\sigma * f$ = remove noise by smoothing

↳ ① → 1^{st} derivative of gaussian $\nabla(n_\sigma)$ = $\nabla(n_\sigma)*f$ = extremum denotes an edge

↳ ② → 2^{nd} derivative of gaussian $\nabla^2(n_\sigma)$ = $\nabla^2(n\sigma)*f$ = zero crossing denotes an edge







- as sigma σ gets wider peak value begins to fall ⟶ response of the operator reduces
- change sigma to find different scaled blobs, with finding max points among them

↳ local extrema in (x, σ) - space represent blobs

locations of blobs ↲ ↳ characteristic scale for that max point

in the example above: $\underbrace{\sigma_A^* = \sigma_1, \quad \sigma_B^* = 2\sigma, \quad \sigma_C^* = 3\sigma_3}$

↳ characteristic scale ∝ size of blob (they are proportional)
(size of the blob)

2D blob detection = normalized laplacian of gaussian (NLoG) is used

↳ laplacian = $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ ↳ $\sigma^2 \cdot \nabla^2 n_\sigma$ → find extreme locals among many scales

* increasing sigma lowers the resolution

Scale space = stack created by filtering an image with gaussians of different sigma (σ)

↳ $S(x,y,\sigma) = \underbrace{n(x,y,\sigma)}_{NLoG} * Image(x,y)$

↳ $(x^*, y^*, \sigma^*) = \arg\max_{(x,y,\sigma)} |\sigma^2 \nabla^2 n_\sigma * I(x,y)|$
→ (x^*, y^*) = position of the blob
→ σ^* = size of the blob

fast NLoG approximation = difference of gaussian (DoG) = $n_{s\sigma} - n_\sigma \approx (s-1)\underbrace{\sigma^2 \nabla^2 n_\sigma}_{NLoG}$ normalized laplacian of gaussian

↳ instead of computing normalized laplacian of gaussian, simply take the difference between stack layers in gaussian scale-space

to make it scale invariant = $\frac{\sigma_1^*}{\sigma_2^*}$: ratio of blob size, match accordingly

to make it rotation invariant = compute image gradient directions in each pixel (in blobs region), choose the most prominent/repeated gradient direction, macth these in both images
↳ using histograms

Summary = given same two images with different scales ⟶ goal: finding interest points independently

↳ solution: search for max of functions in scale and in space

↳ methods: harris-laplacian = max laplacian over scale, harris measure of corner response over the img
↳ SIFT = max difference of gaussians over scale and space

matching SIFT descriptors =

① create histograms of gradient directions over spatial regions

↳ normalized histogram = invarient to rotation, scale, and brightness

↳ very large image gradients are usually from illumination effects

↳ to reduce this, clamp all values in vector to be ≤0.2, then normalize again

(best is 8 orientation bins and 4x4 histograms)
16 bins
8x16 = 128 numbers

② compare them using
→ L2 distance: zero perfect match
↳ normalized correlation: 1 is perfect match
↳ intersection: larger, better the match (overlap)

applications of local invariant features = motion tracking, panoramas, 3D reconstruction ...