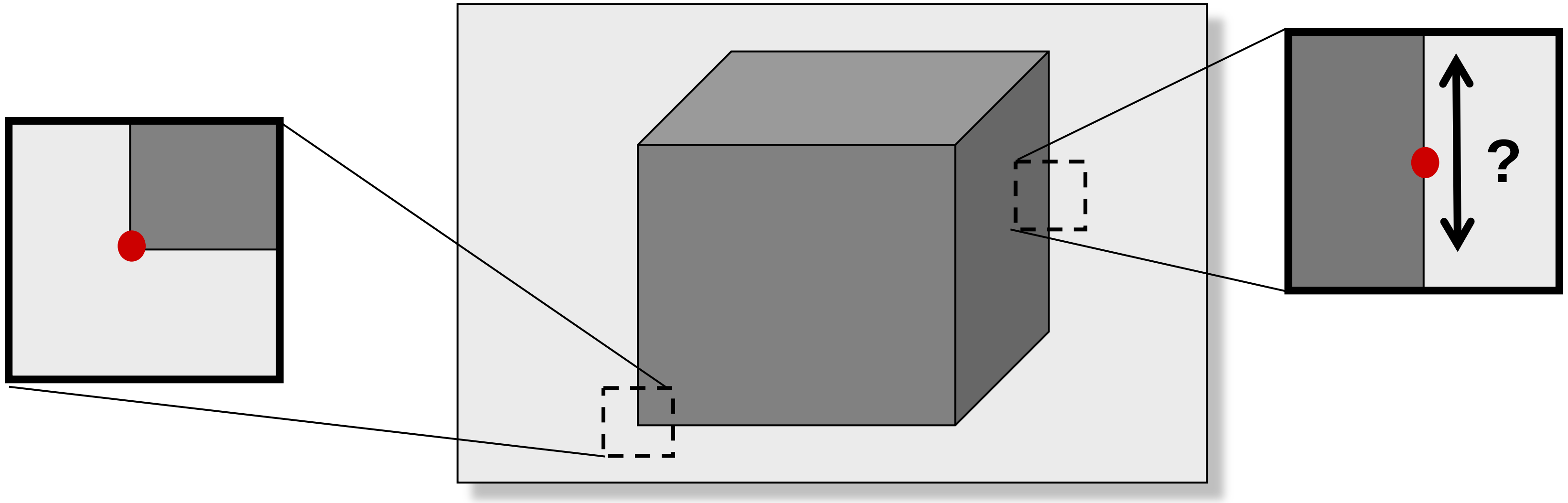
Keypoint detection

- Many applications benefit from features localized in (x,y)
- Edges well localized only in one direction ➔ detect corners?

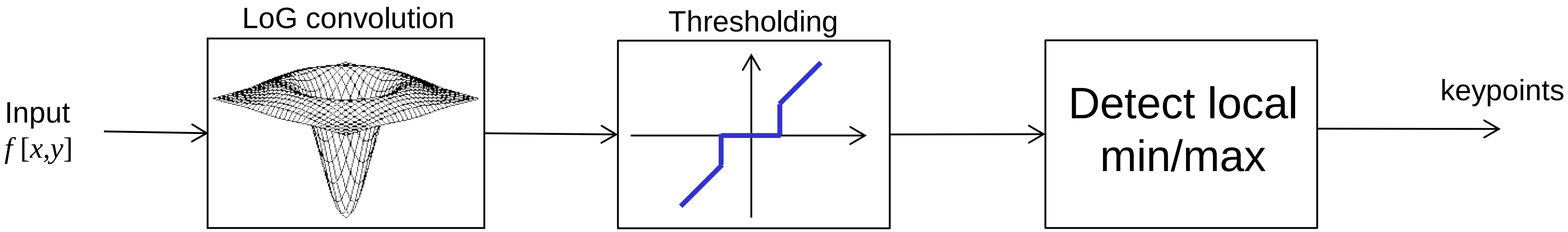


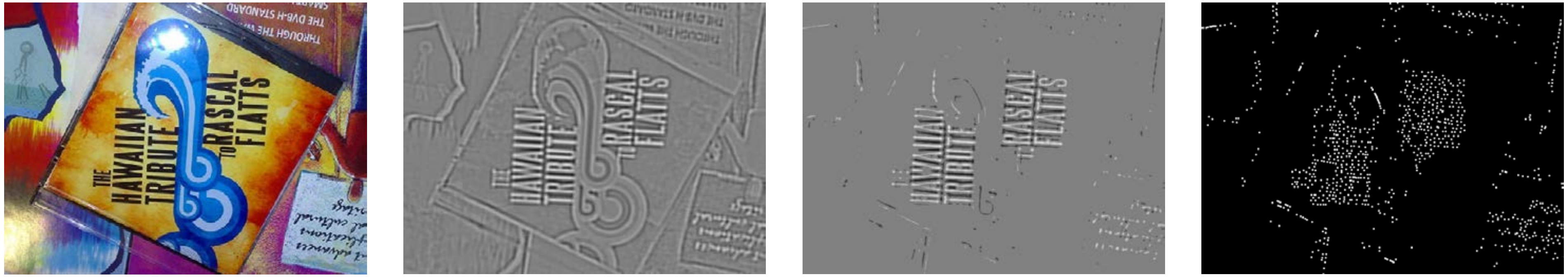
- Desirable properties of keypoint detector
  - Accurate localization
  - Invariance against shift, rotation, scale, brightness change
  - Robustness against noise, high repeatability

Keypoint detection

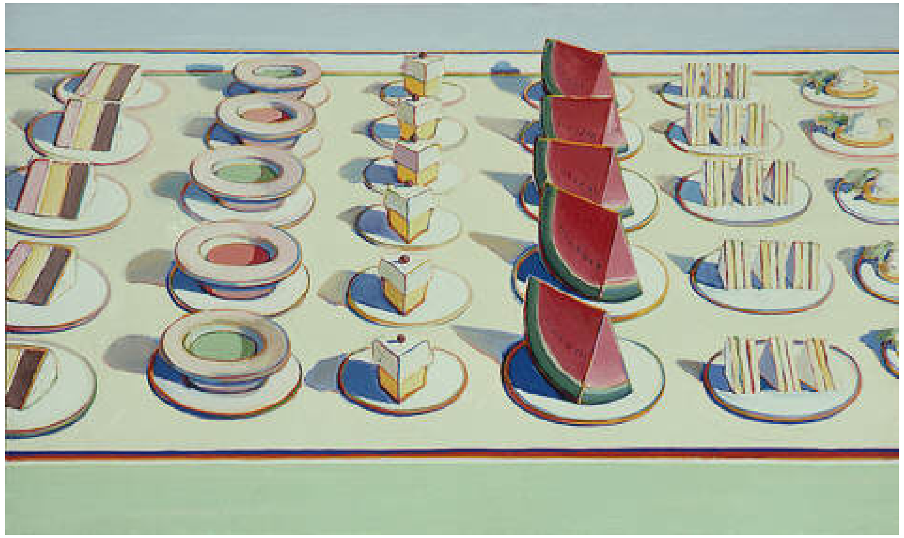
- Laplacian detector
- Determinant of Hessian detector
- Harris detector
- FAST detector

Laplacian keypoint detector

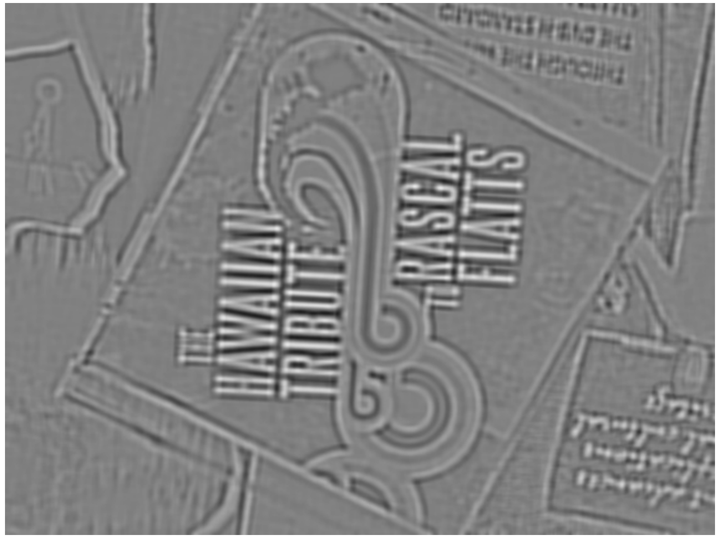


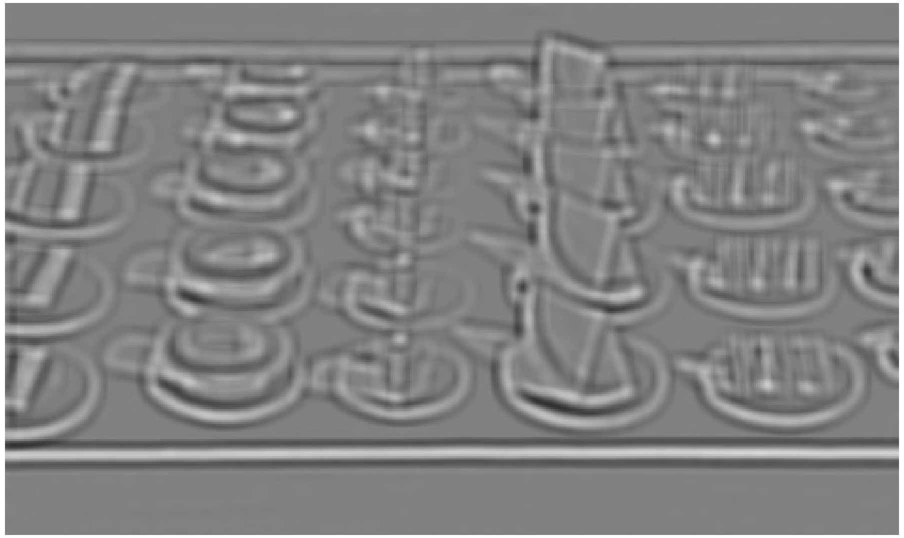


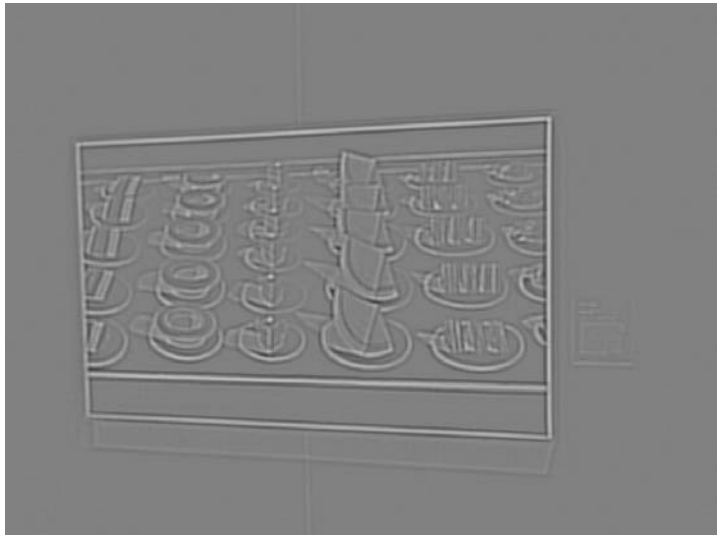
Input images



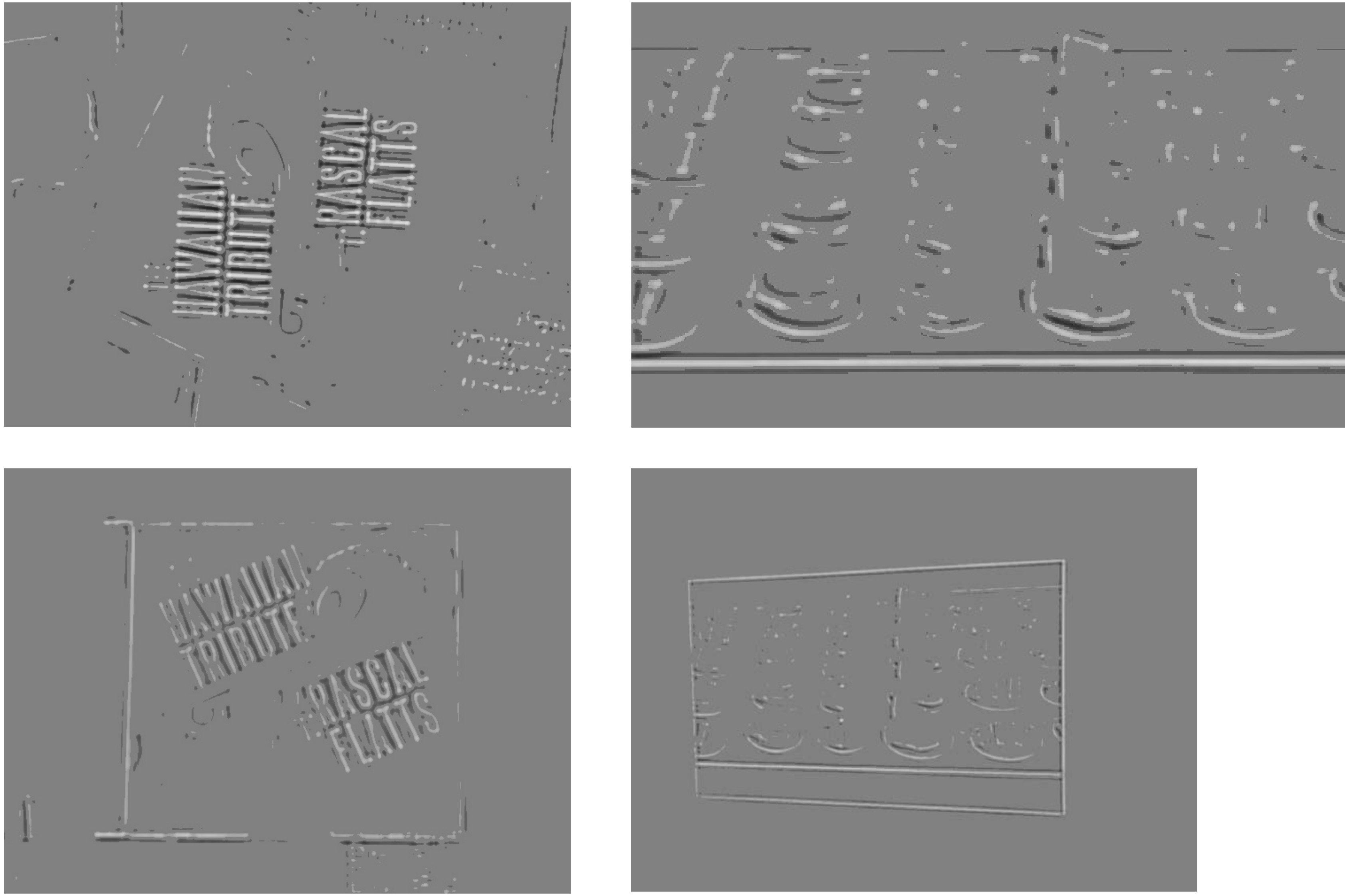
LoG response

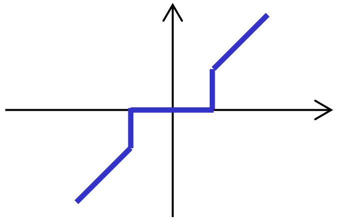




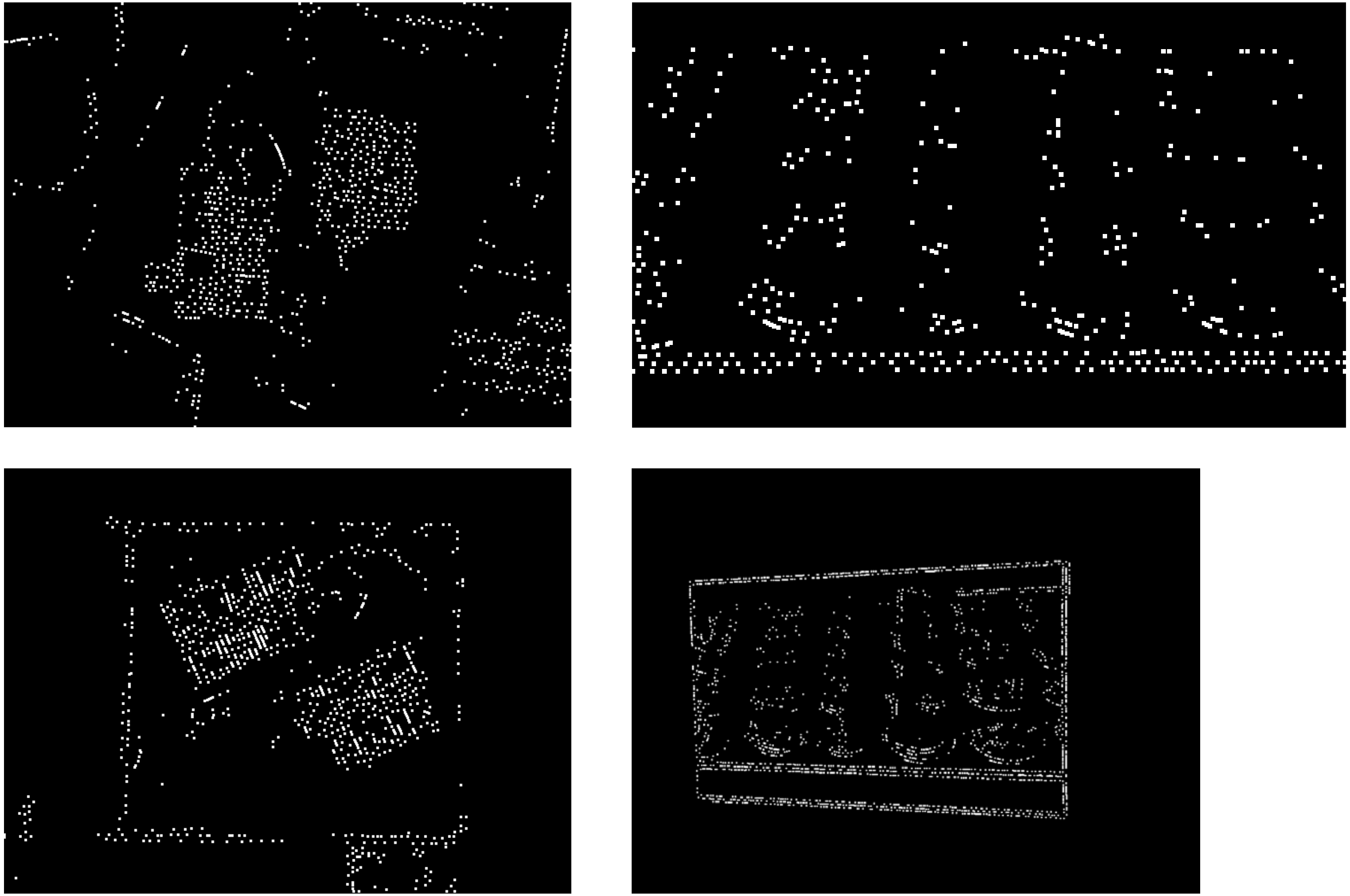


Thresholded LoG response

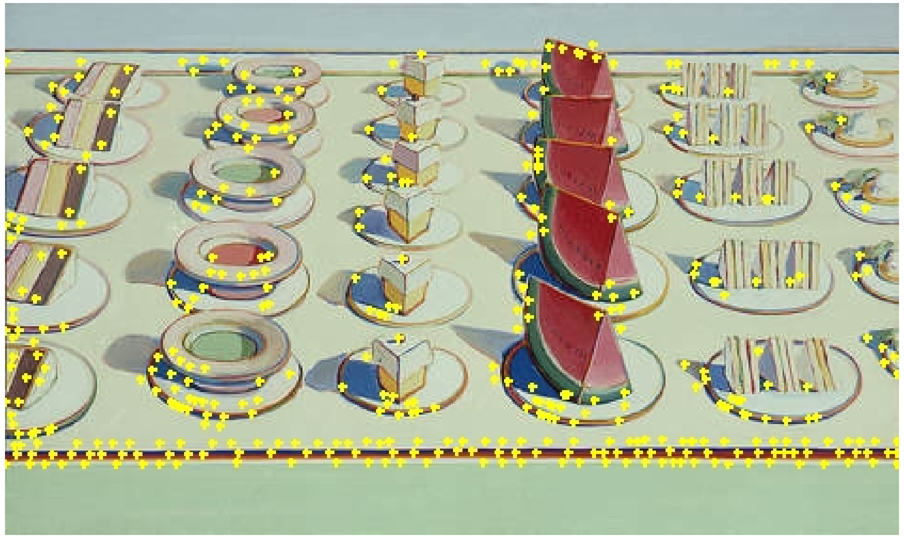




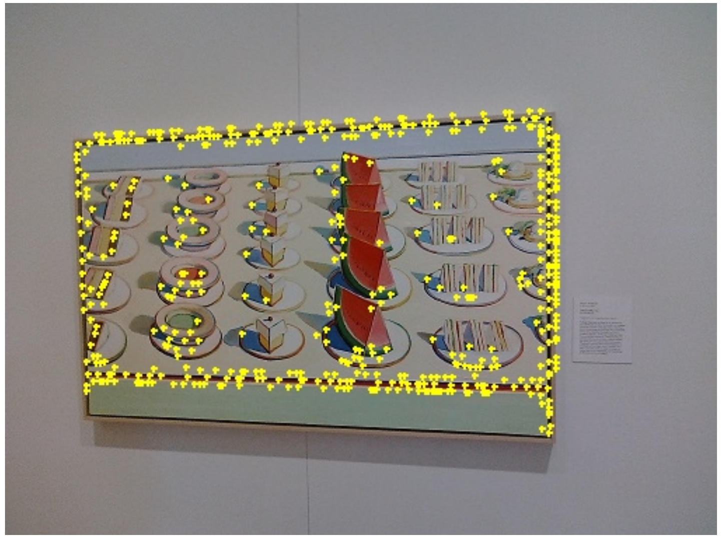
Local extrema of thresholded LoG response



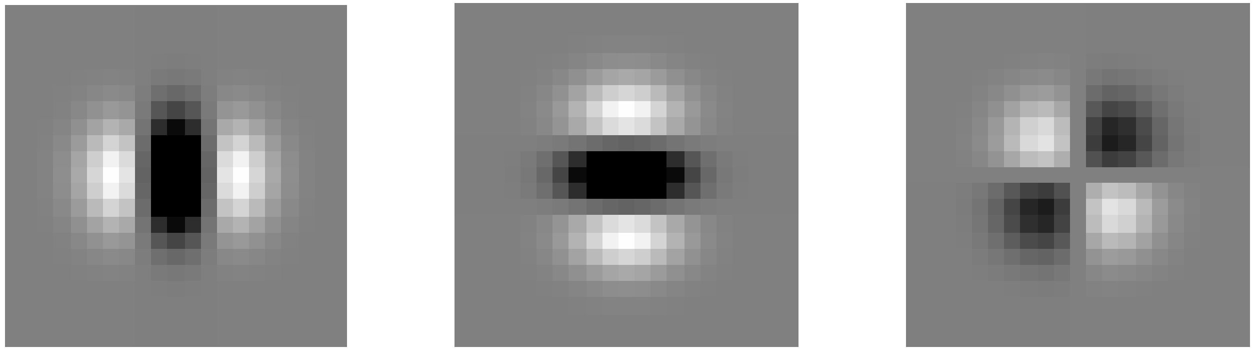
Superimposed LoG keypoints



500 strongest keypoints



Determinant of Hessian keypoint detector



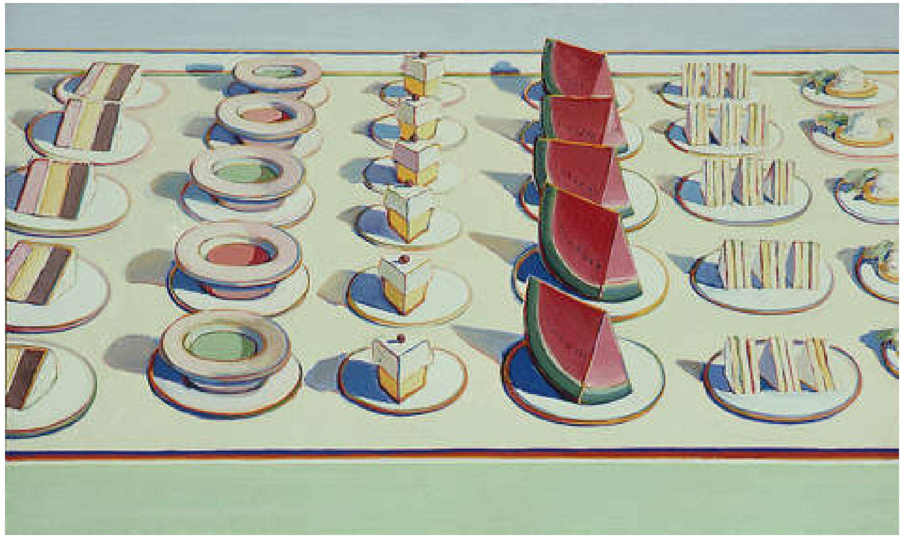
D_{xx} D_{yy} D_{xy}

$$\mathbf{H}[x,y] = \begin{bmatrix} f_{xx}[x,y] & f_{xy}[x,y] \\ f_{xy}[x,y] & f_{yy}[x,y] \end{bmatrix}$$

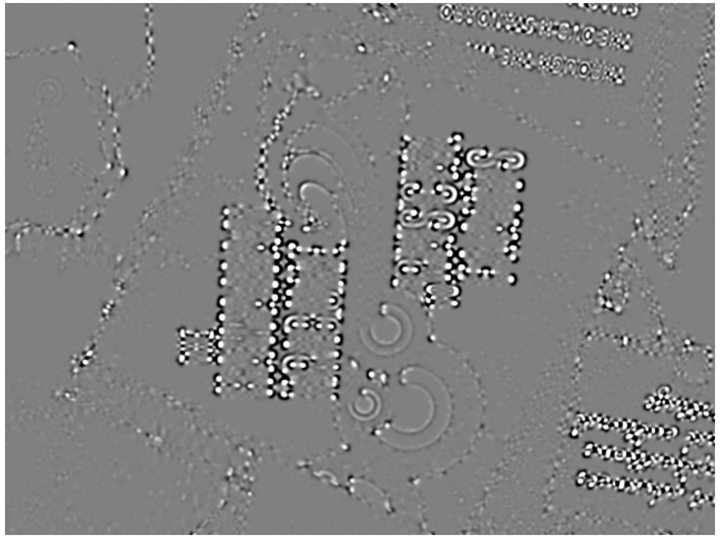
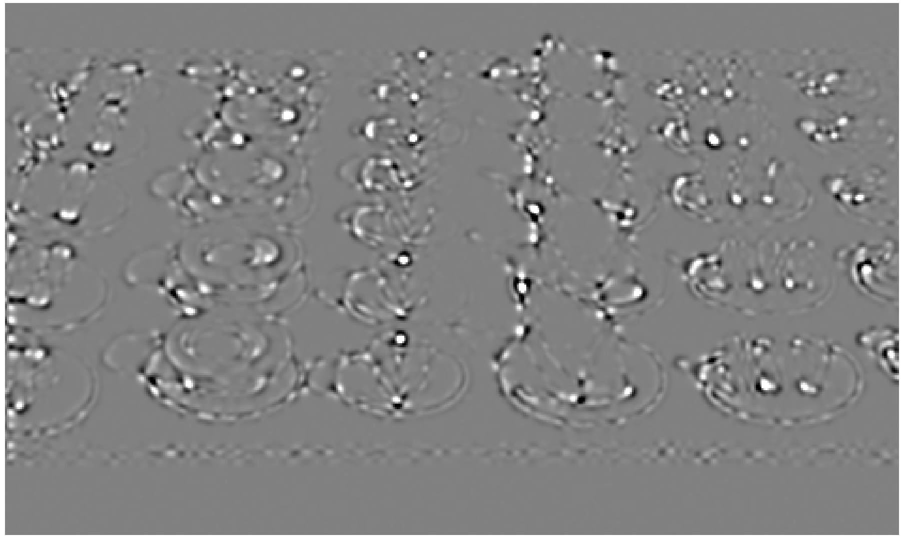
$$= \begin{bmatrix} D_{xx}[x,y] * f[x,y] & D_{xy}[x,y] * f[x,y] \\ D_{xy}[x,y] * f[x,y] & D_{yy}[x,y] * f[x,y] \end{bmatrix}$$

$$\det \mathbf{H}[x,y] = f_{xx}[x,y] f_{yy}[x,y] - \left(f_{xy}[x,y] \right)^2$$

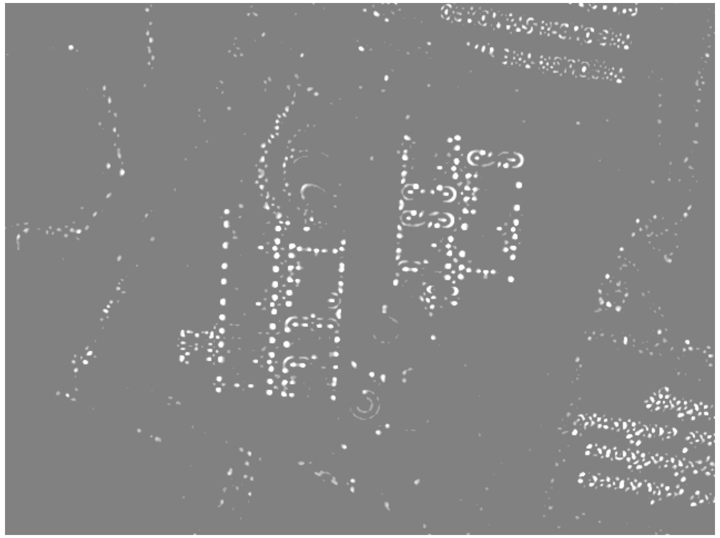
Input images

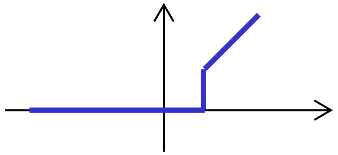


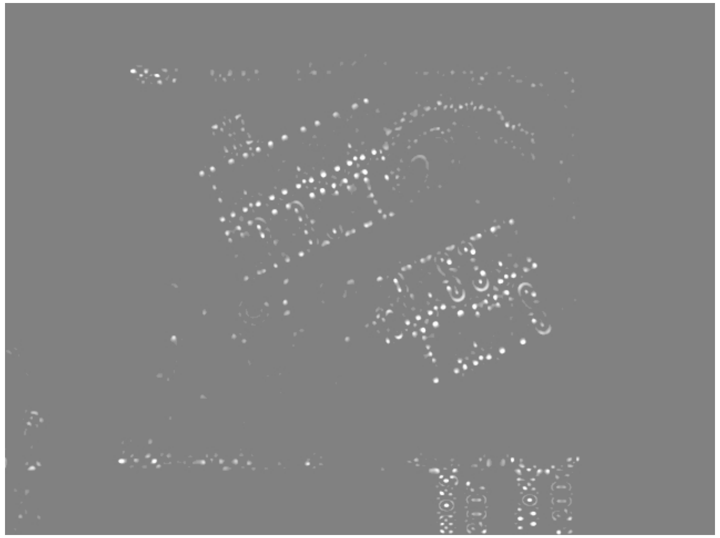
DoH response

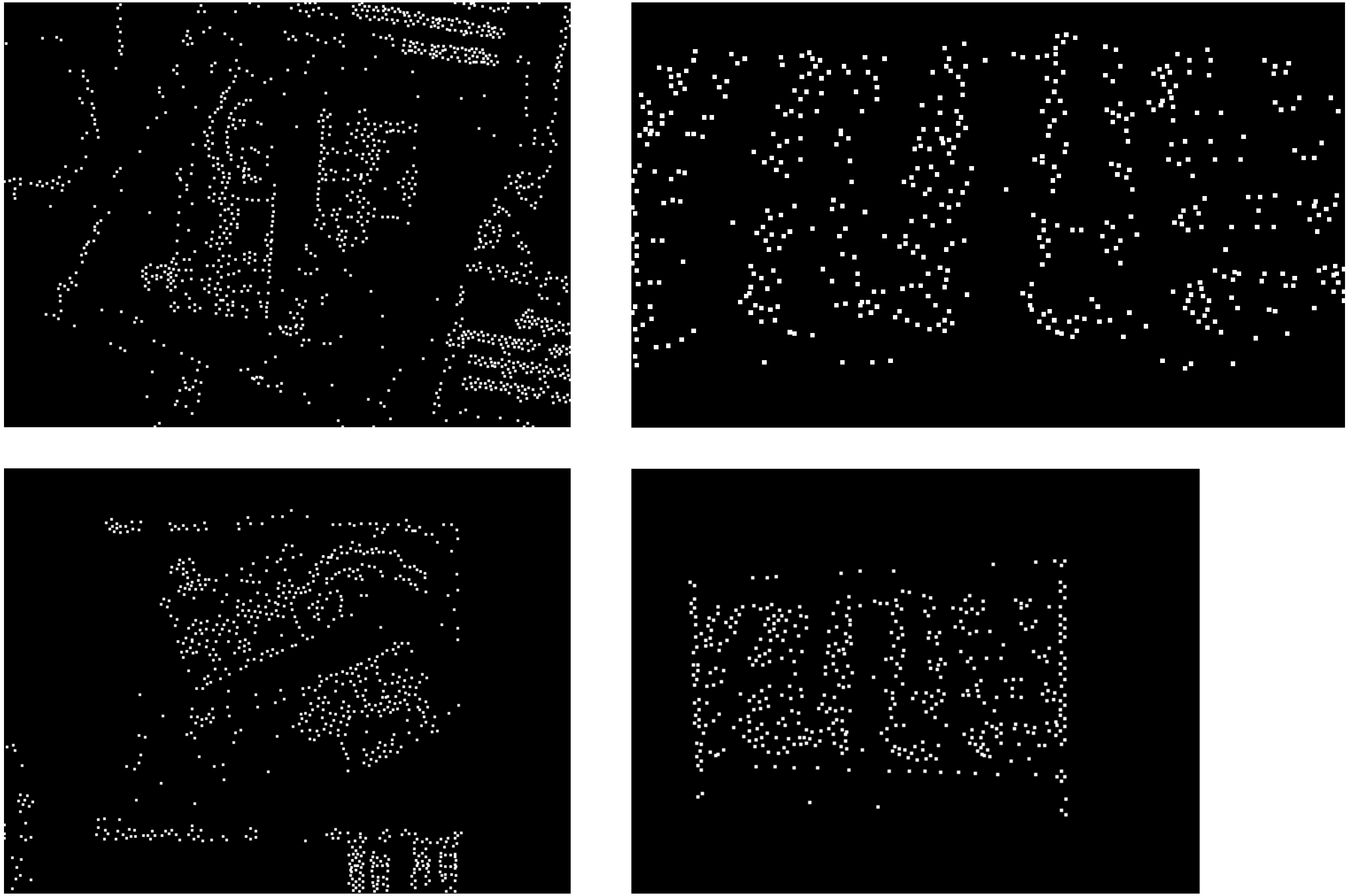
Thresholded DoH response



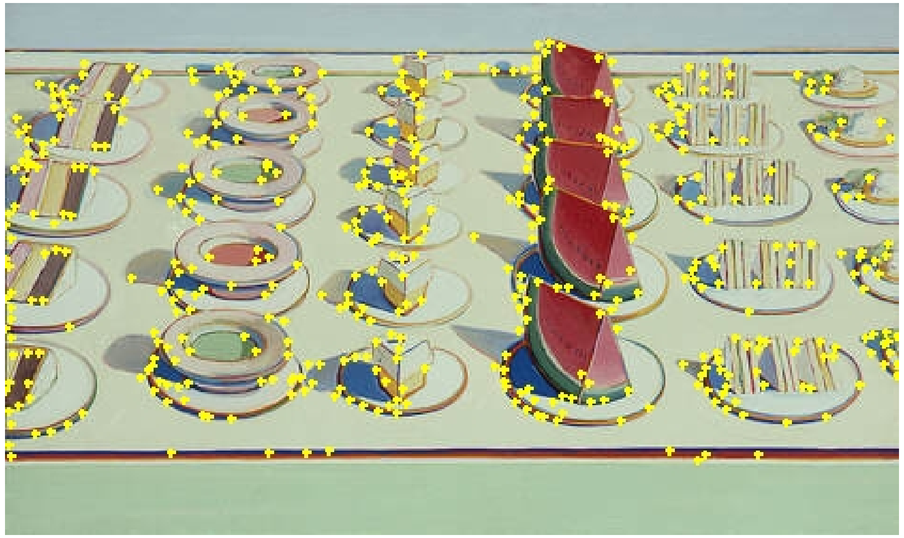




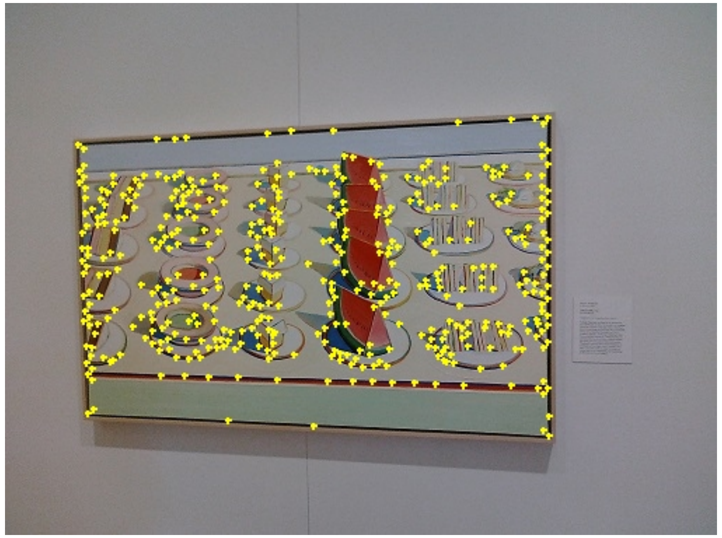
Local maxima of DoH response

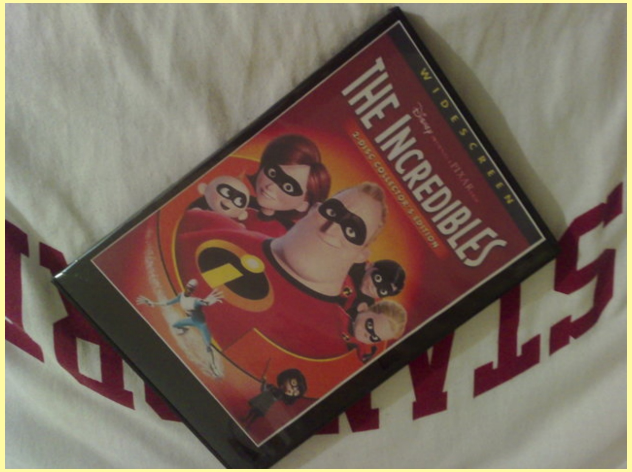


Superimposed DoH keypoints

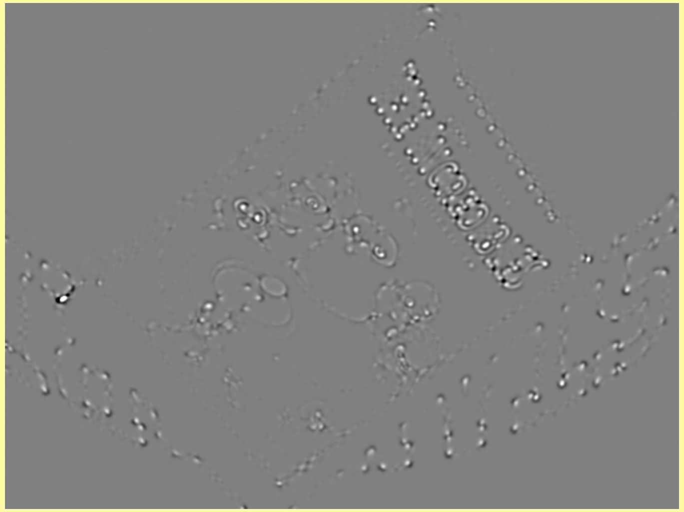


500 strongest keypoints





Which of the following images is the Laplacian of Gaussian and which is the Determinant of Hessian of the above image?



What patterns can be localized most accurately?

- Local displacement sensitivity (assuming continuous $f(x,y)$)

$$S(\Delta x,\Delta y)=\sum_{(x,y)\in window}\left[f(x,y)-f(x+\Delta x,y+\Delta y)\right]^2$$

- Linear approximation for small $\Delta x, \Delta y$

$$f(x+\Delta x,y+\Delta y)\approx f(x,y)+f_x(x,y)\Delta x+f_y(x,y)\Delta y$$

$f_x(x,y)$ – horizontal image gradient
$f_y(x,y)$ – vertical image gradient

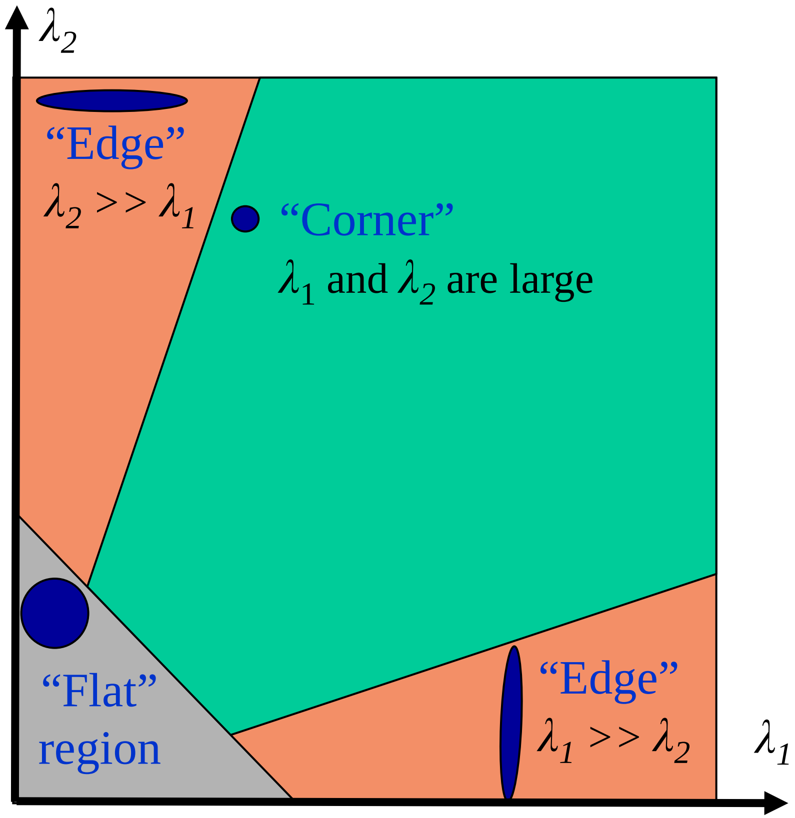
$$S(\Delta x,\Delta y)\approx\sum_{(x,y)\in window}\left[\begin{pmatrix} f_x(x,y) & f_y(x,y)\end{pmatrix}\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix}\right]^2$$

$$=\begin{pmatrix}\Delta x & \Delta y\end{pmatrix}\left(\sum_{(x,y)\in window}\begin{bmatrix} f_x^2(x,y) & f_x(x,y)f_y(x,y)\\ f_x(x,y)f_y(x,y) & f_y^2(x,y)\end{bmatrix}\right)\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix}$$

$$=\begin{pmatrix}\Delta x & \Delta y\end{pmatrix}\mathbf{M}\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix}$$

- Iso-sensitivity curves are ellipses

Harris detector



Based on eigenvalues λ_1, λ_2 of "structure matrix" (aka "normal matrix" aka "second-moment matrix")

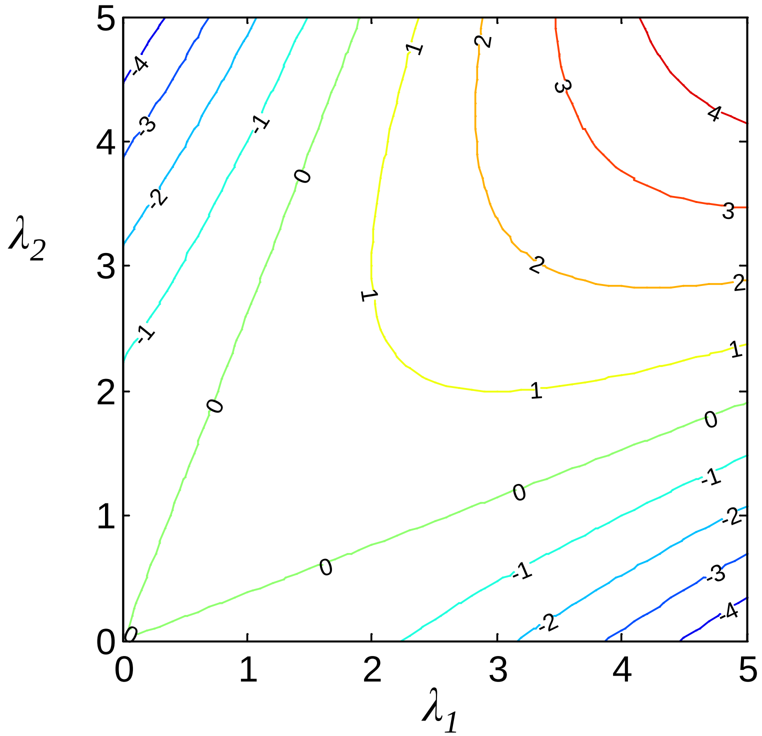
$$\mathbf{M} = \begin{bmatrix} \sum\limits_{[x,y]\in window} f_x^2[x,y] & \sum\limits_{[x,y]\in window} f_x[x,y] f_y[x,y] \\ \sum\limits_{[x,y]\in window} f_x[x,y] f_y[x,y] & \sum\limits_{[x,y]\in window} f_y^2[x,y] \end{bmatrix}$$

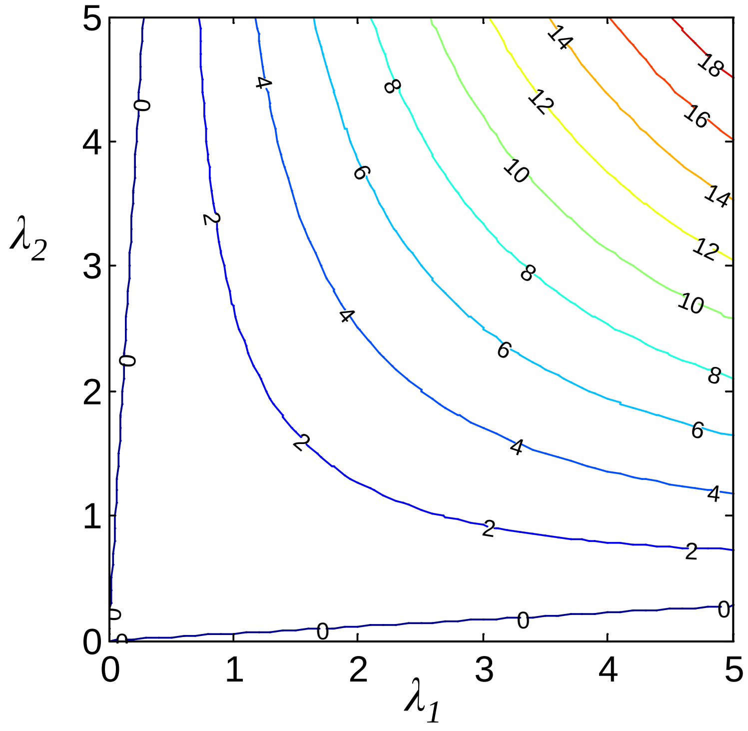
$f_x[x,y]$ – horizontal image gradient
$f_y[x,y]$ – vertical image gradient

Harris cornerness

$$C = \det(\mathbf{M}) - k \cdot \left(trace\left(\mathbf{M}\right) \right)^2 = \lambda_1 \lambda_2 - k \cdot \left(\lambda_1 + \lambda_2 \right)^2$$

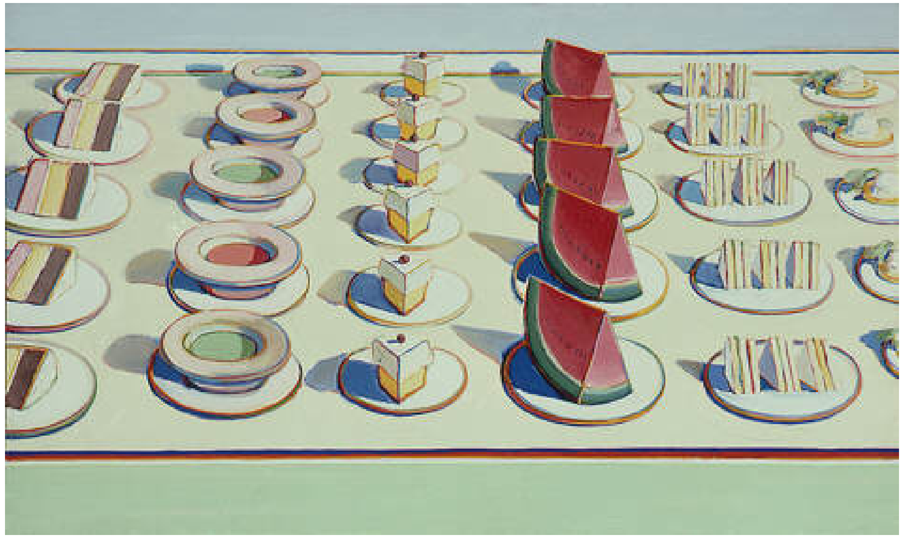
k = 0.2



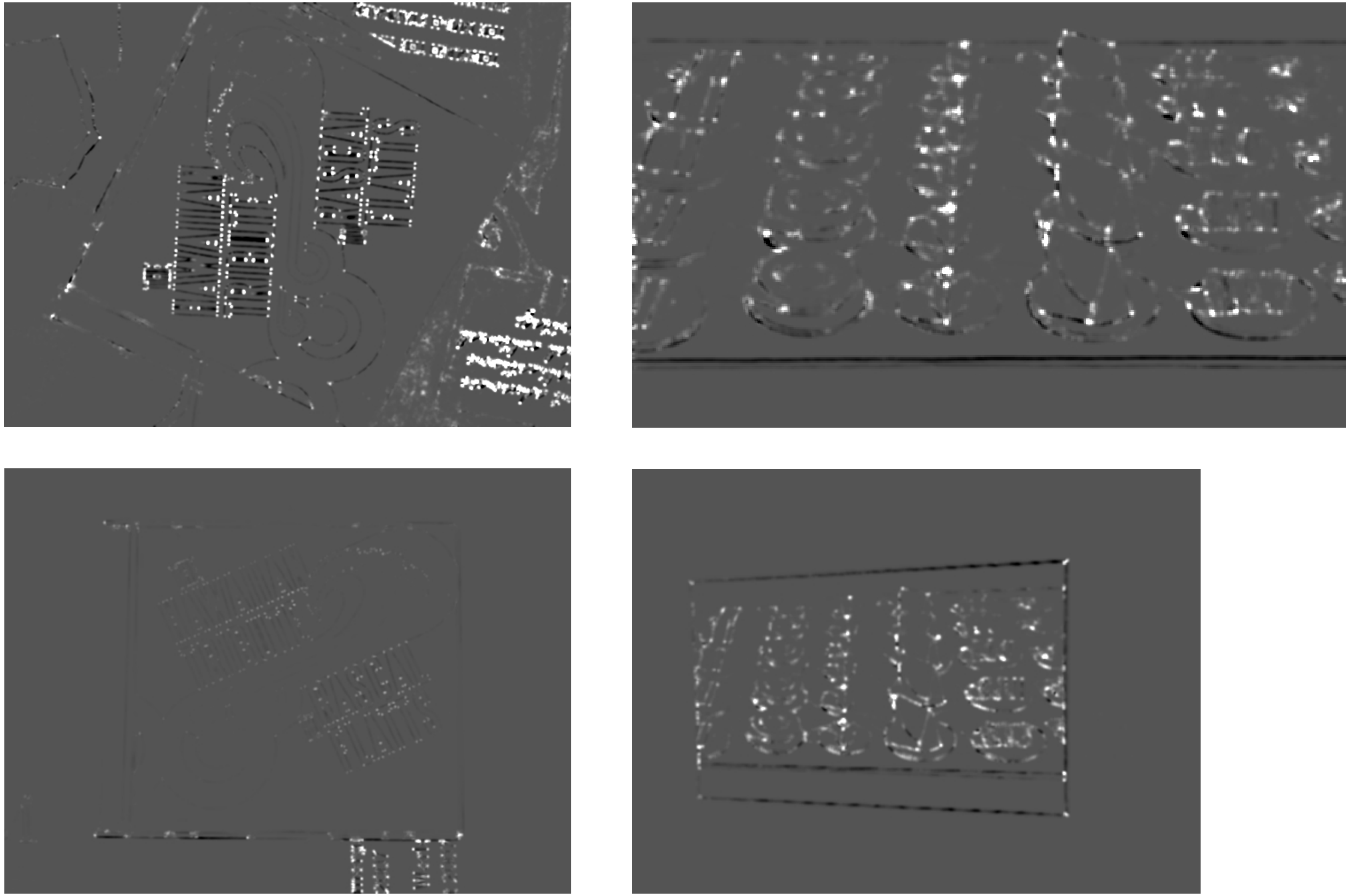


k = 0.05

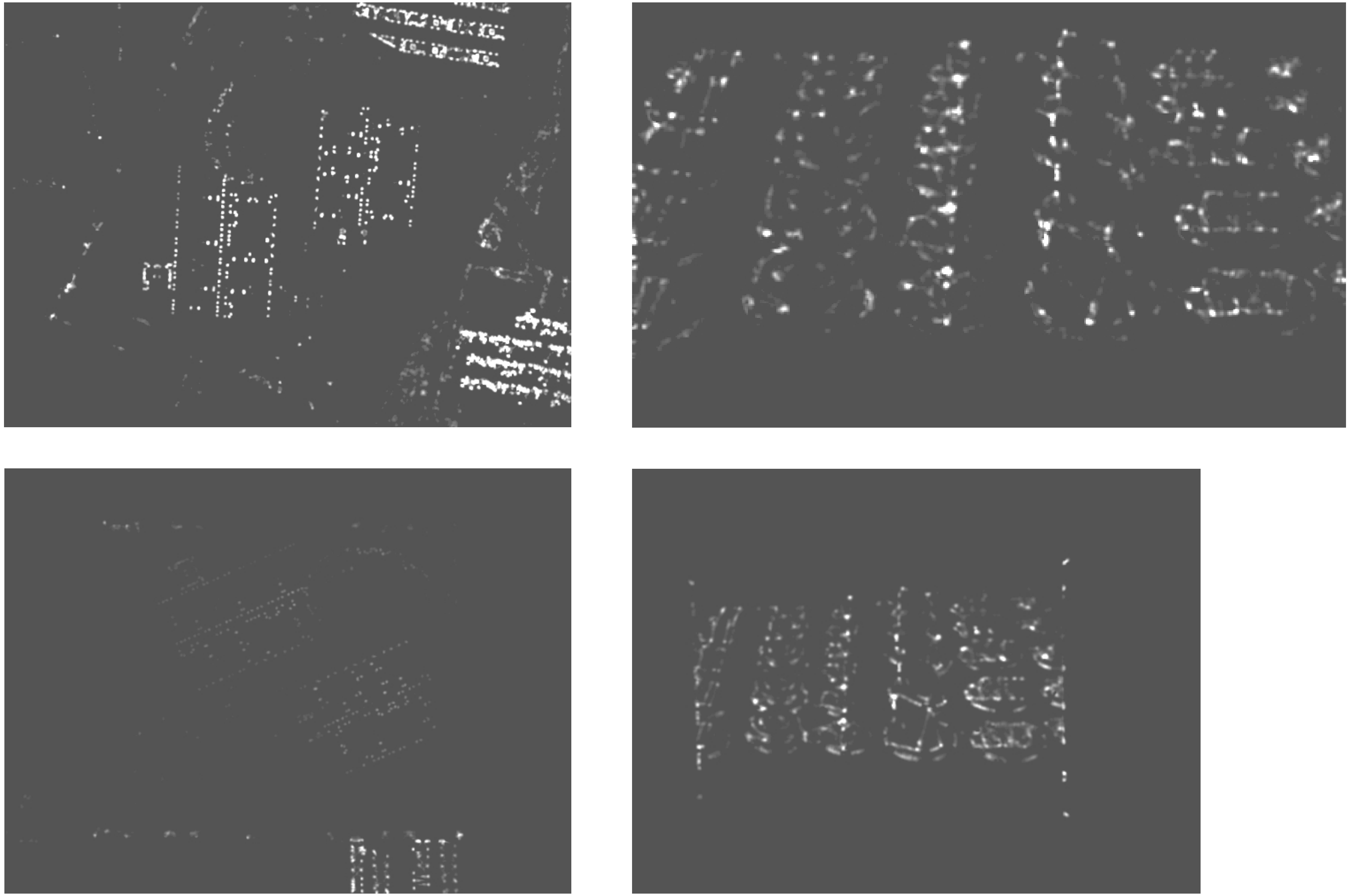
Input images

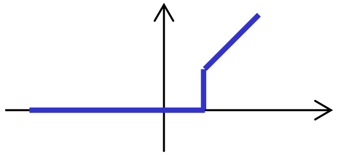


Harris cornerness

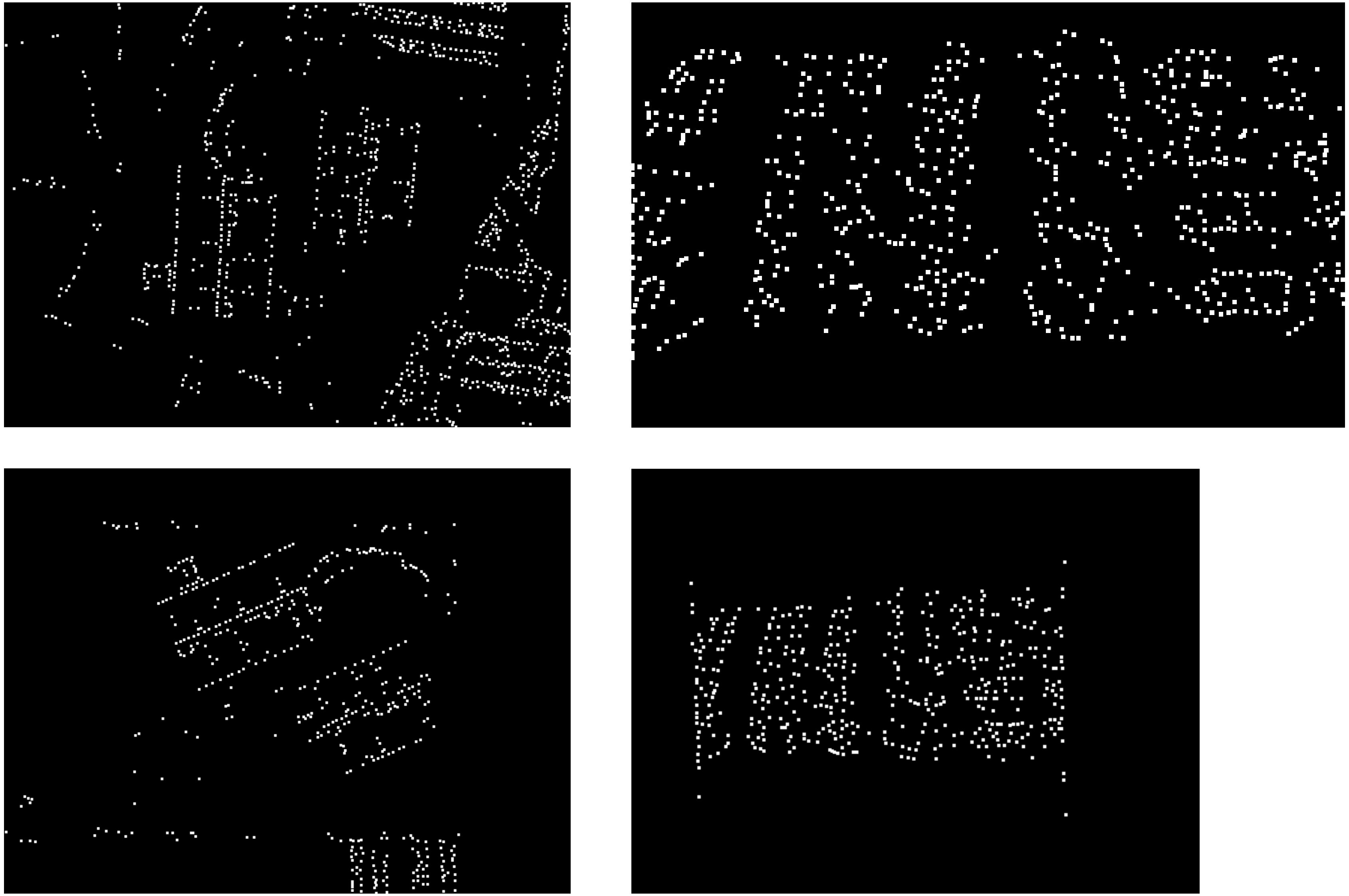


Thresholded cornerness





Local maxima of cornerness

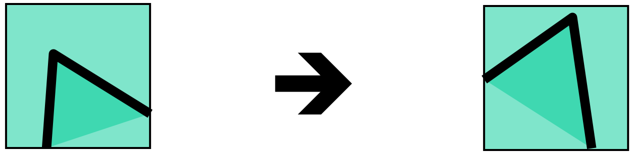


Superimposed Harris keypoints

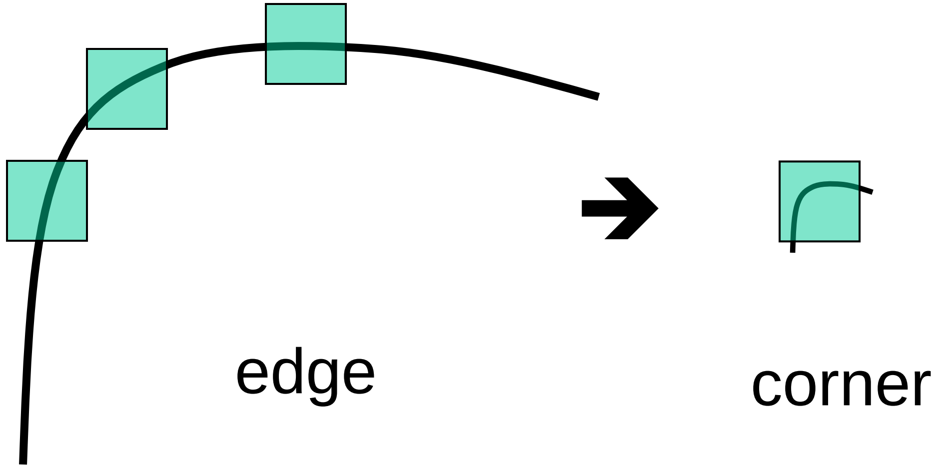
500 strongest keypoints

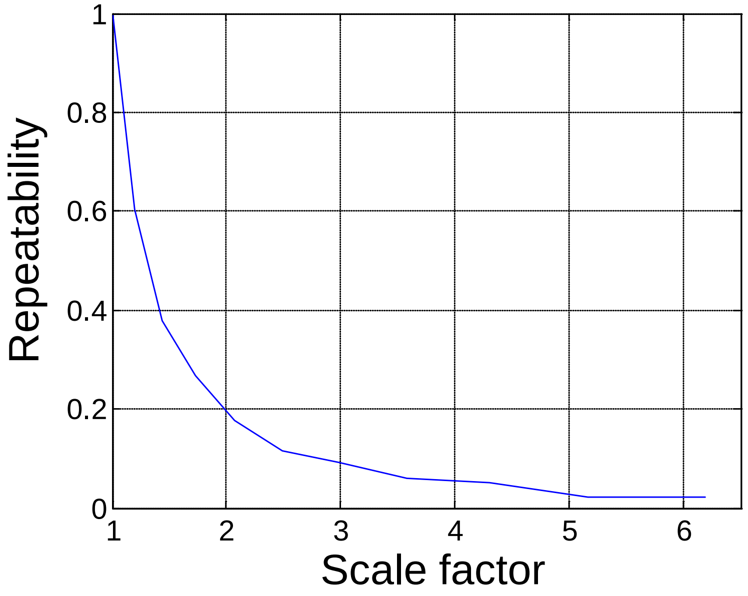
Robustness of Harris detector

- Invariant to brightness offset: $f[x,y] \rightarrow f[x,y] + c$
- Invariant to shift and rotation



- Not invariant to scaling

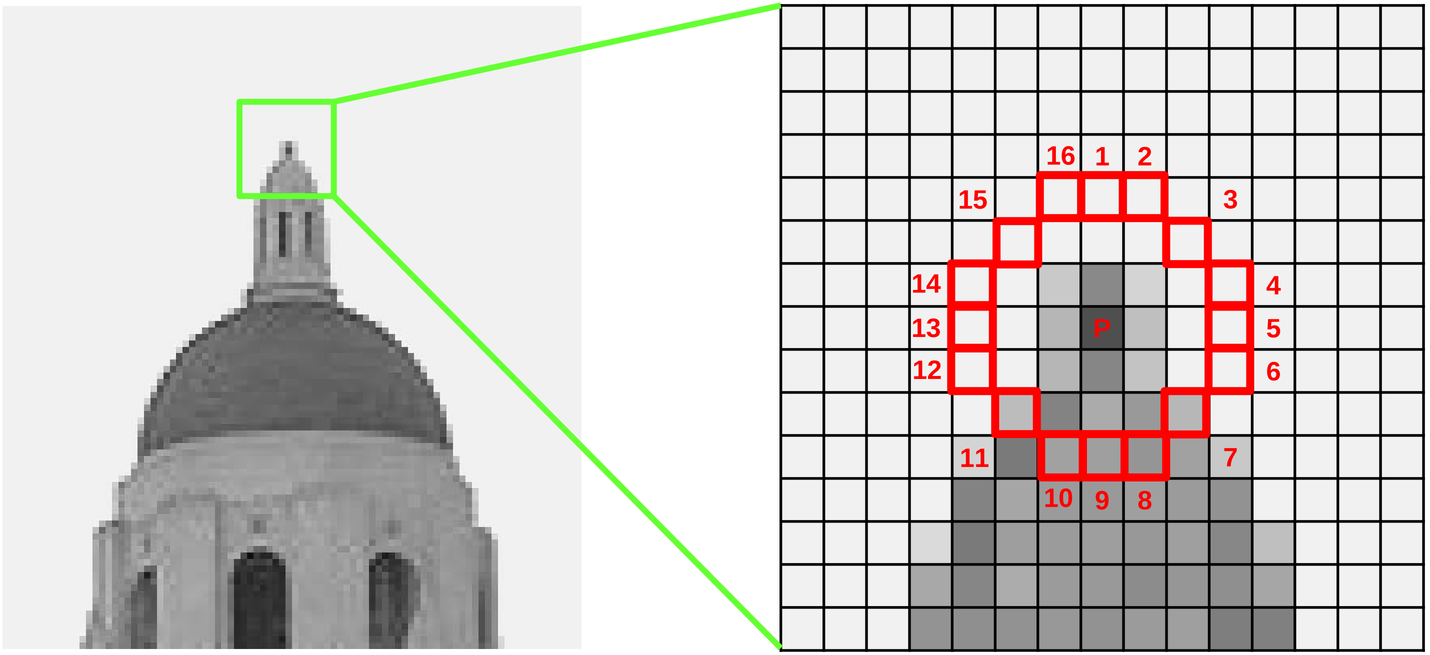




The following structure matrices are examples of the three different cases “Flat”, “Edge” and “Corner” of the Harris detector. Which is which?

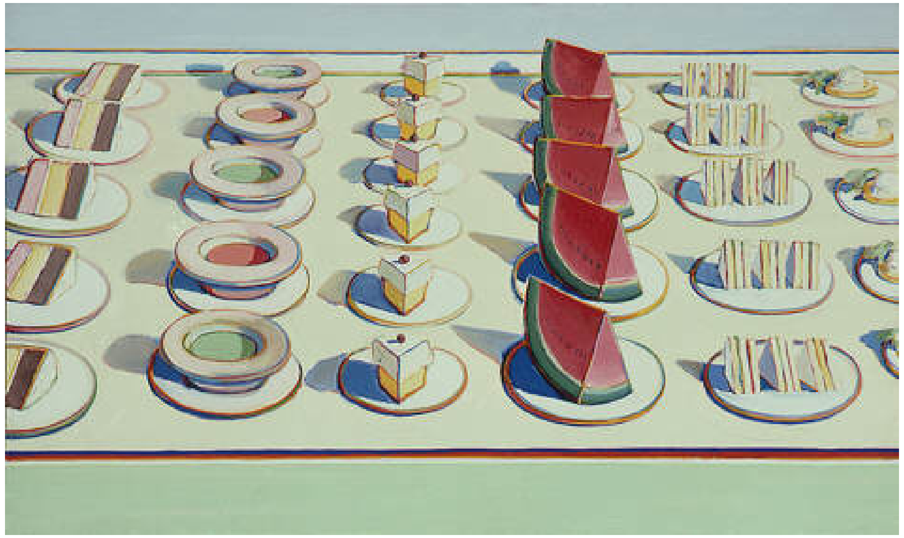
$$M = \begin{bmatrix} 5 & 1 \\ 1 & 4 \end{bmatrix} \qquad M = \begin{bmatrix} 12 & 0.1 \\ 0.1 & .003 \end{bmatrix} \qquad M = \begin{bmatrix} .05 & -0.01 \\ -0.01 & 0.02 \end{bmatrix}$$

Features from Accelerated Segment Test (FAST)

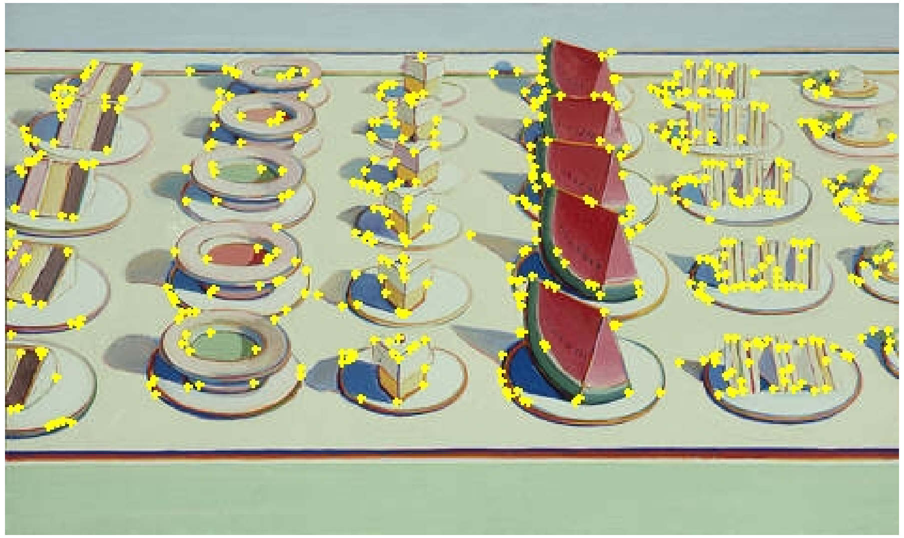


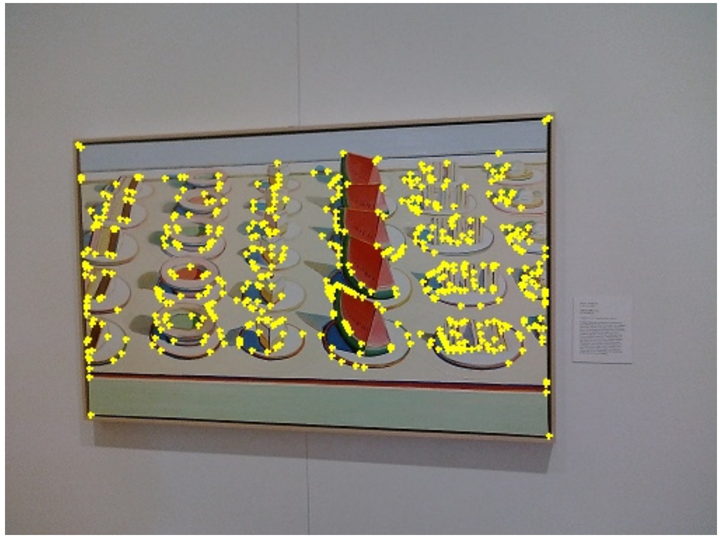
- Compare “nucleus” p to circle of sixteen pixels
- Nucleus is feature point, iff at least $n=9$ contiguous circle pixels are either all brighter, or all darker, by θ
- Optimize pixel comparisons to reject non-corners early

Input images



FAST corners superimposed





FAST corner detection on smartphone

FAST keypoint tracking on smartphone