Detectors of salient points or regions

Václav Hlaváč

Czech Technical University in Prague

Czech Institute of Informatics, Robotics and Cybernetics

160 00 Prague 6, Jugoslávských partyzánů 1580/3, Czech Republic

http://people.ciirc.cvut.cz/hlavac, vaclav.hlavac@cvut.cz

also Center for Machine Perception, http://cmp.felk.cvut.cz

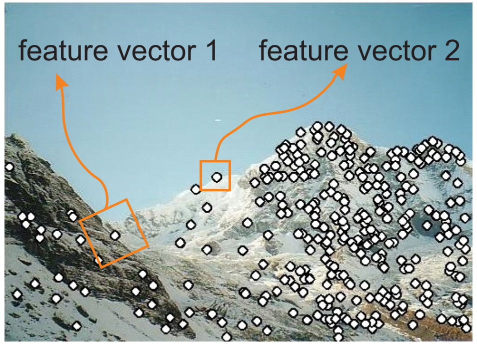
Courtesy: T. Werner.

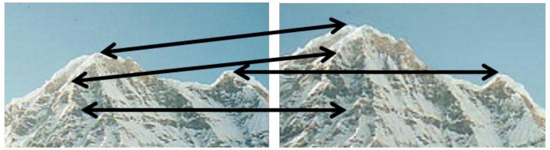
Outline of the talk:

- Global picture: detection, description, matching.
- The aim – create a detector.
- Properties of a good detector.
- Harris corners and MSERs, briefly.
- Harris corners in a more detail.
-

Establishing correspondence across views, a motivation

1. Detection identifies the interest points (also keypoints) or regions.
2. Description calculates descriptors (feature vectors) from the local neighborhood of each interest point or region.
3. Matching compares feature description pairwise across views, ranks them, and selects the most prominent one (or a few most prominent ones to increase robustness).





Some applications of detections/descriptions

- Image alignment
- 3D reconstruction
- Motion tracking
- Indexing and retrieval in the image database
- Object recognition
- Robot navigation

The aim

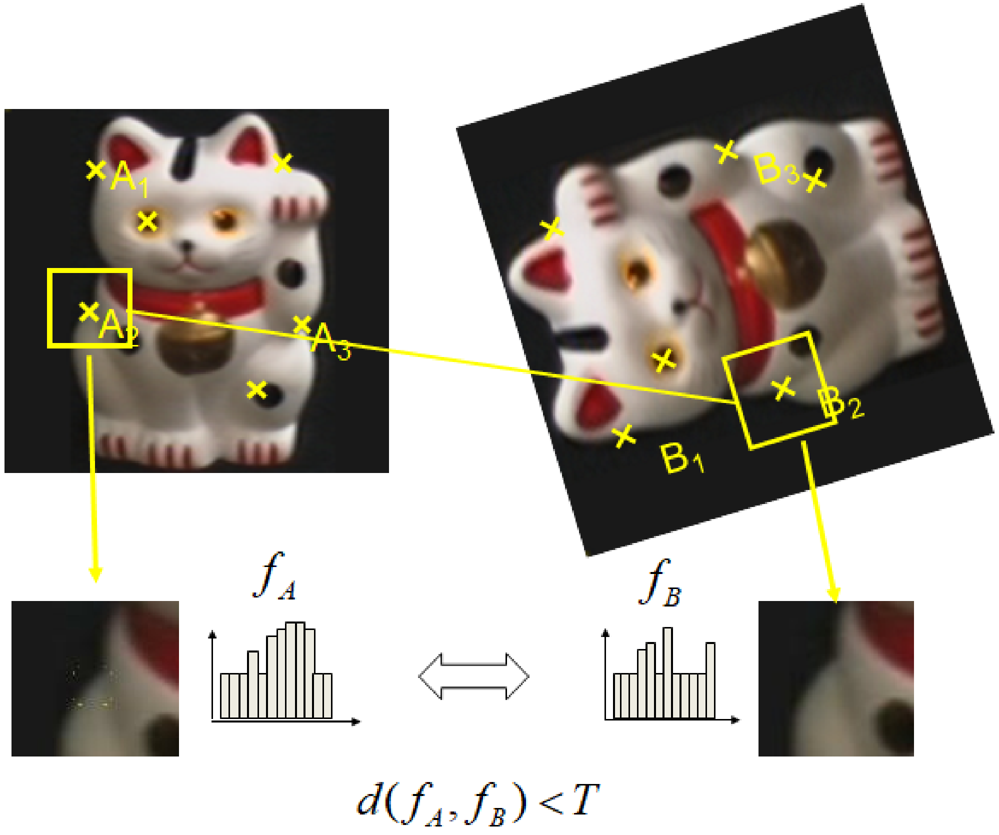
Design a detector/descriptor that finds interest points or interest regions in an image such that:

- There is only a small number of isolated interest points/regions detected as compared to the entire number of pixels in the image.
- Image semantics is not taken into account.
- The interest points/regions and related descriptors are reasonably invariant
  - to small geometric changes as affine transformations, e.g., rotation and scale,
  - to small radiometric variations, e.g., a slight illumination change.
  - different sampling and quantization.

Note: The seminal detector satisfying these requirements is e.g. the Harris corner detector from 1988. However, the idea was around already earlier.

Interest points used in matching, overview, procedure

1. Find a set of distinctive interest points.
2. Define a region around each interest point.
3. Extract and normalize the region content.
4. Compute a local descriptor from the normalized region.
5. Match local descriptors.



Courtesy K. Grauman.

Desired properties of good detections/descriptions

- Repeatability – The same detections with similar descriptors can be found in slightly different images despite geometric and photometric distortions.
- Saliency – Detections/descriptions are distinctive.
- Compactness and efficiency – Significantly less detections/features than image pixels are used.
- Locality – The descriptor is calculated from a small neighborhood of a detection. It brings robustness to clutter and occlusion.
- Invariance to small geometric/photometric changes – Enables the use in diverse applications, cf. next slide.

Points of interest or salient regions are usually stable across viewpoint or illumination changes and provide a good localization (cf. a line vs. a cross).

References

Seminal papers:

- C. Harris, M. Stephens. *A Combined Corner and Edge Detector*. Proceedings of the 4th Alvey Vision Conference, 1988, pages 147–151.
- J. Matas, O. Chum, M. Urban, T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. British Machine Vision Conference, 2002, pages 384–396.

Review papers introducing other detectors:

- C. Schmid, R. Mohr, C. Bauckhage: *Evaluation of Interest Point Detectors*, IJCV 37(2):151-172, 2000.
- K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, L. Van Gool, *A comparison of affine region detectors*, IJCV 65(1/2):43-72, 2005.

ZOO of useful detectors

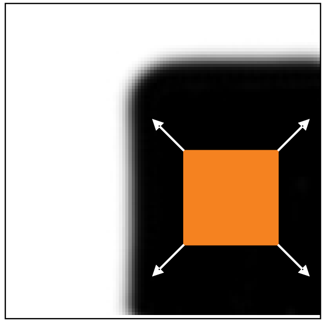
Hessian & Harris	[Beaudet 1978], [Harris 1988]
Laplacian, DoG	[Lindeberg '98], [Lowe 1999]
Harris-/Hessian-Laplace	[Mikolajczyk & Schmid 2001]
Harris-/Hessian-Affine	[Mikolajczyk & Schmid 2004]
EBR and IBR	[Tuytelaars & Van Gool 2004]
MSER	[Matas et al. 2002]
Salient Regions	[Kadir & Brady 2001]
Others ...	

Abbreviations:
DoG – Difference of Gaussians; EBR – edge-based region; IBR – intensity-based region; MSER – maximally stable extremal regions.

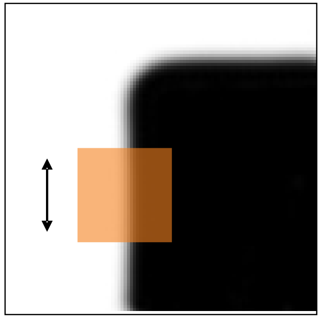
Corner detection, the basic idea pictorially

- The interest point should be detected locally. Consider prospecting the image by a small shifting viewing window.
- A slight shift of this viewing window in any direction should yield a large change in intensity.
- Around a corner, the image gradient has two or more dominant directions.

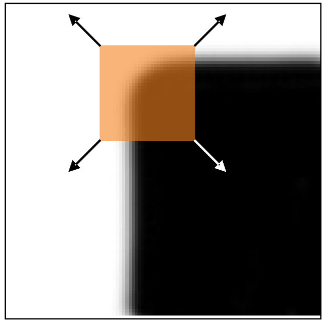
Idea: A. Efros



"flat" region: no change in all directions



"edge": no change along the edge direction

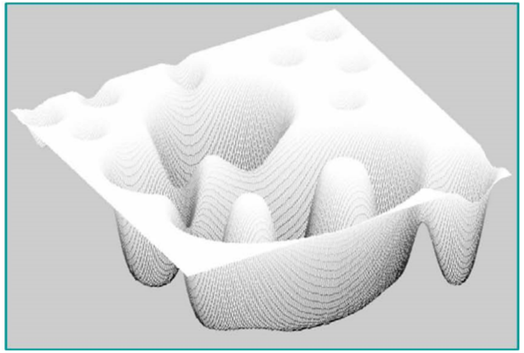


"corner": significant change in all directions

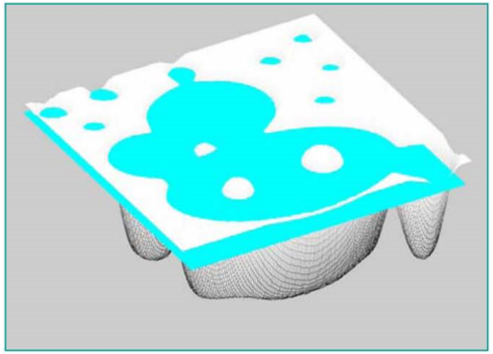
Maximally Stable Extremal Regions (MSERs)
The basic idea

- MSERs (Matas et. al 2002) are regions characterized by an almost uniform intensity. The detected regions are surrounded by a contrasting background.
- MSERs are constructed by trying multiple thresholds while selecting those regions, which maintain the same area over changing thresholds.

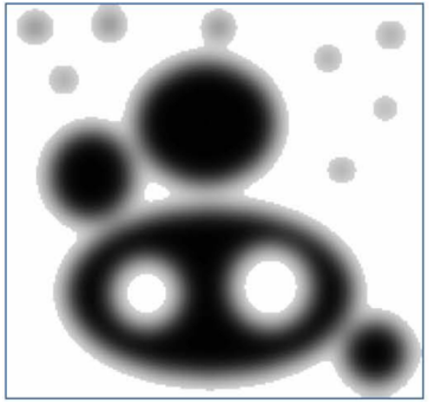
orig. gray scale



landscape



thresholds



MSERs

MSER, idea, 3D animation video

MSER, street image illustration

Harris corners in more detail
Autocorrelation function

- How similar is the image function $I(x, y)$ at point (x, y) to itself, when shifted by $(\Delta x, \Delta y)$?
- This is given by the autocorrelation function

$$c(x, y; \Delta x, \Delta y) = \sum_{(u,v) \in W(x,y)} w(u, v) \big(I(u, v) - I(u + \Delta x, v + \Delta y) \big)^2$$

where

- $W(x, y)$ is a window centered at point (x, y)
- $w(u, v)$ is either constant or (better) Gaussian $\exp \dfrac{-(u - x)^2 - (v - y)^2}{2\sigma^2}$.

(Further on, we will replace $\sum_{(u,v) \in W(x,y)} w(u, v)$ with $\sum_{W}$ for simplicity)

Quadratic approximation of the autocorrelation function (1)

Approximate the shifted function by the first-order Taylor expansion:

$$
\begin{aligned}
I(u + \Delta x, v + \Delta y) &\approx I(u, v) + I_x(u, v)\Delta x + I_y(u, v)\Delta y \\
&= I(u, v) + [\, I_x(u, v), I_y(u, v) \,] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} ,
\end{aligned}
$$

where I_x, I_y are partial derivatives of $I(x, y)$.

Quadratic approximation of the autocorrelation function (2)

Autocorrelation function $c(x, y; \Delta x, \Delta y)$:

$$
\begin{aligned}
c(x, y; \Delta x, \Delta y) &= \sum_W \Big(I(u, v) - I(u + \Delta x, v + \Delta y) \Big)^2 \\
&\approx \sum_W \left([\, I_x(u, v), I_y(u, v) \,] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\
&= [\, \Delta x, \Delta y \,]\, Q(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} ,
\end{aligned}
$$

$$
\begin{aligned}
\text{where } Q(x, y) &= \sum_W \begin{bmatrix} I_x(x, y)^2 & I_x(x, y) I_y(x, y) \\ I_x(x, y) I_y(x, y) & I_y(x, y)^2 \end{bmatrix} \\
&= \begin{bmatrix} \sum_W I_x(x, y)^2 & \sum_W I_x(x, y) I_y(x, y) \\ \sum_W I_x(x, y) I_y(x, y) & \sum_W I_y(x, y)^2 \end{bmatrix}
\end{aligned}
$$

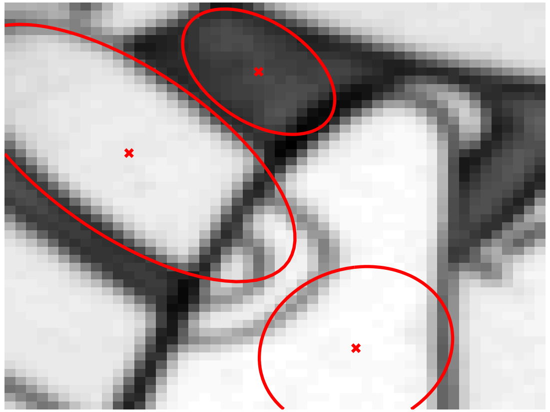
Quadratic approximation of the autocorrelation function (3)

- The autocorrelation function has been approximated by a quadratic function

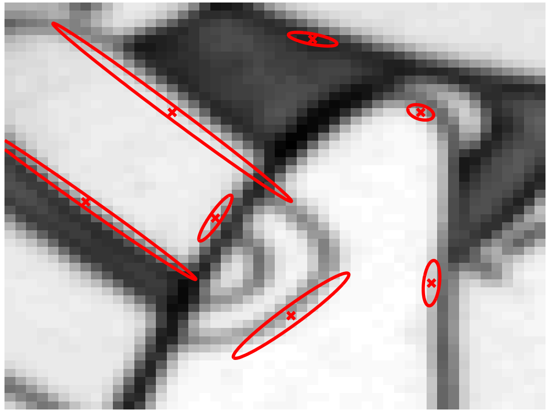
$$c(x,y;\Delta x,\Delta y) \approx [\,\Delta x, \Delta y\,]\, Q(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = [\,\Delta x, \Delta y\,] \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

- Elongation and size of the uncertainty ellipse is given by eigenvalues λ_1, λ_2 of $Q(x,y)$
- The rotation angle of the ellipse is given by eigenvectors of $Q(x,y)$. We do not need the rotation information here.
- Uncertainty ellipses have the equation $[\,\Delta x, \Delta y\,]\, Q(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = 1$:

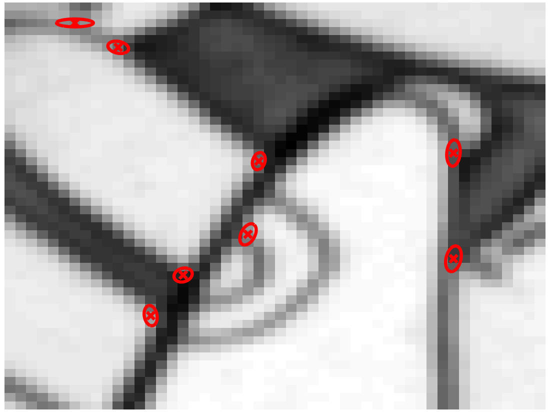
Uncertainty ellippses, illustration



flat region
both eigenvalues small



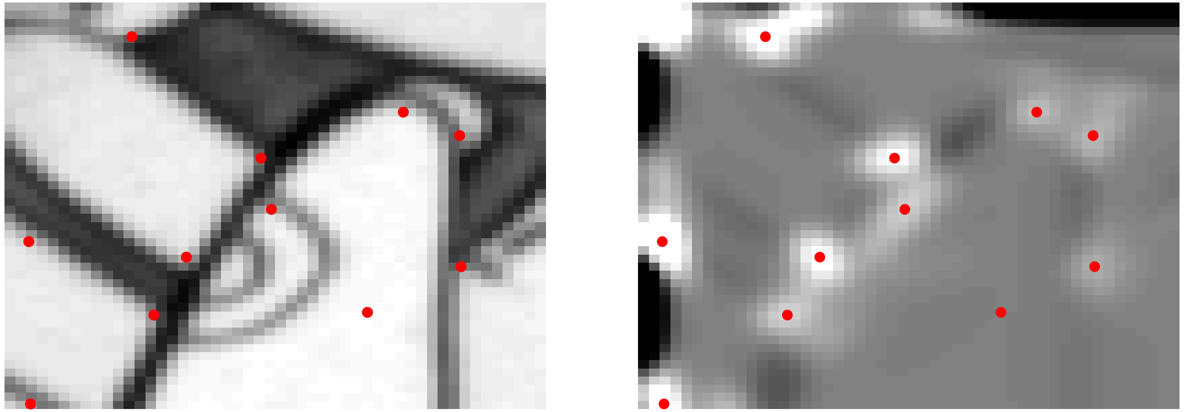
edge
one small, one large



corner
both eigenvalues large

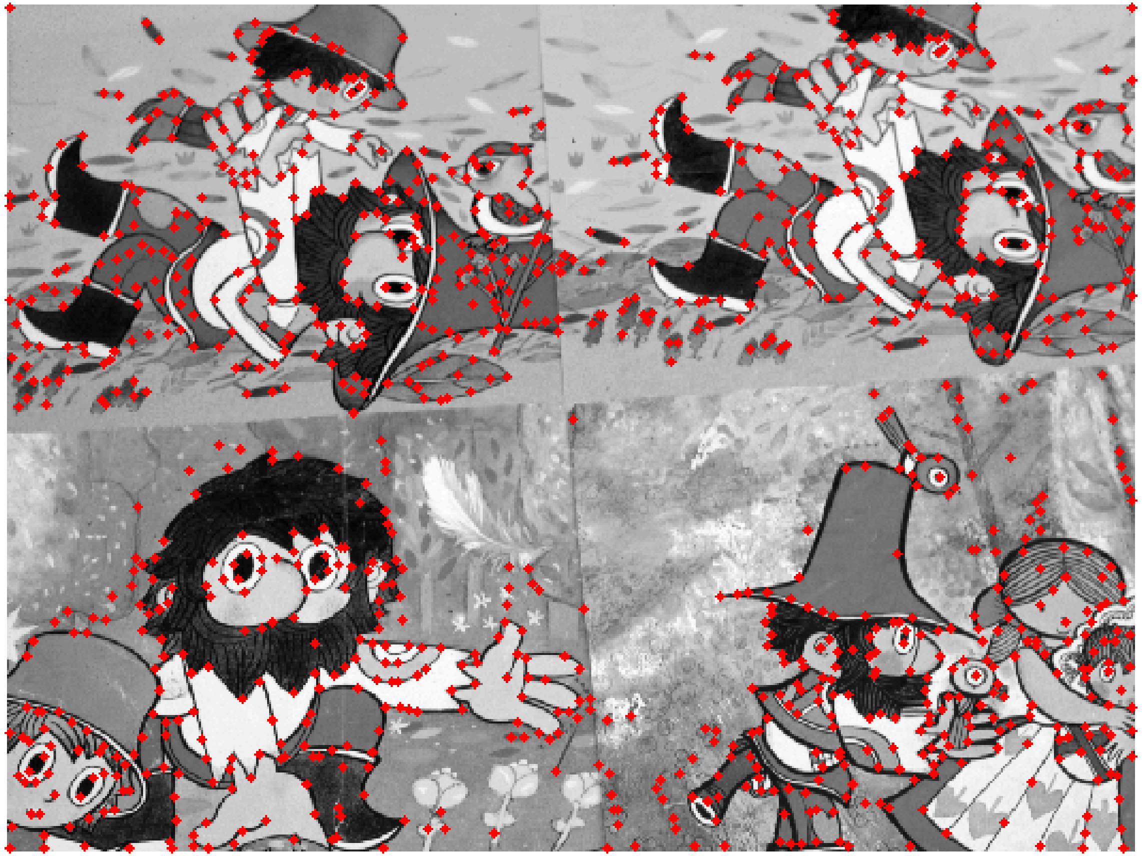
How to find isolated feature points?

- Characterize 'cornerness' $H(x,y)$ by eigenvalues of $Q(x,y)$:
  - $Q(x,y)$ is symmetric and positive definite $\quad \Rightarrow \quad \lambda_1, \lambda_2 > 0$
  - $\lambda_1 \lambda_2 = \det Q(x,y) = AC - B^2, \quad \lambda_1 + \lambda_2 = \operatorname{trace} Q(x,y) = A + C$
  - Harris suggested: Cornerness $H = \lambda_1 \lambda_2 - 0.04(\lambda_1 + \lambda_2)^2$
  - Image $I(x,y)$ and its cornerness $H(x,y)$:



- Find corner points as local maxima of cornerness $H(x,y)$:
  - Local maximum in the image is defined as a point greater than its neighbors (in 3×3 or even 5×5 neighborhood)

Harris corners, typical result (on a larger image)



Harris corners, the algorithm summary

- Compute partial derivatives $I_x(x,y)$, $I_y(x,y)$ by finite differences:

$$I_x(x,y) \approx I(x+1,y) - I(x,y), \quad I_y(x,y) \approx I(x,y+1) - I(x,y)$$

  Before this, it is good (but not necessary) to smooth image with Gaussian with $\sigma \sim 1$, to eliminate noise.

- Compute images

$$A(x,y) = \sum_W I_x(x,y)^2, \quad B(x,y) = \sum_W I_x(x,y)\, I_y(x,y), \quad C(x,y) = \sum_W I_y(x,y)^2$$

  E.g., image $A(x,y)$ is just the convolution of image $I_x(x,y)^2$ with the Gaussian. Use MATLAB function conv2.

- Compute cornerness $H(x,y)$

- Find local maxima in $H(x,y)$. This can be parallelized in MATLAB by shifting the whole image $H(x,y)$ by one pixel left/right/up/down.