Introduction to Robotics for cognitive science

Dr. Andrej Lúčny

KAI FMFI UK

lucny@fmph.uniba.sk

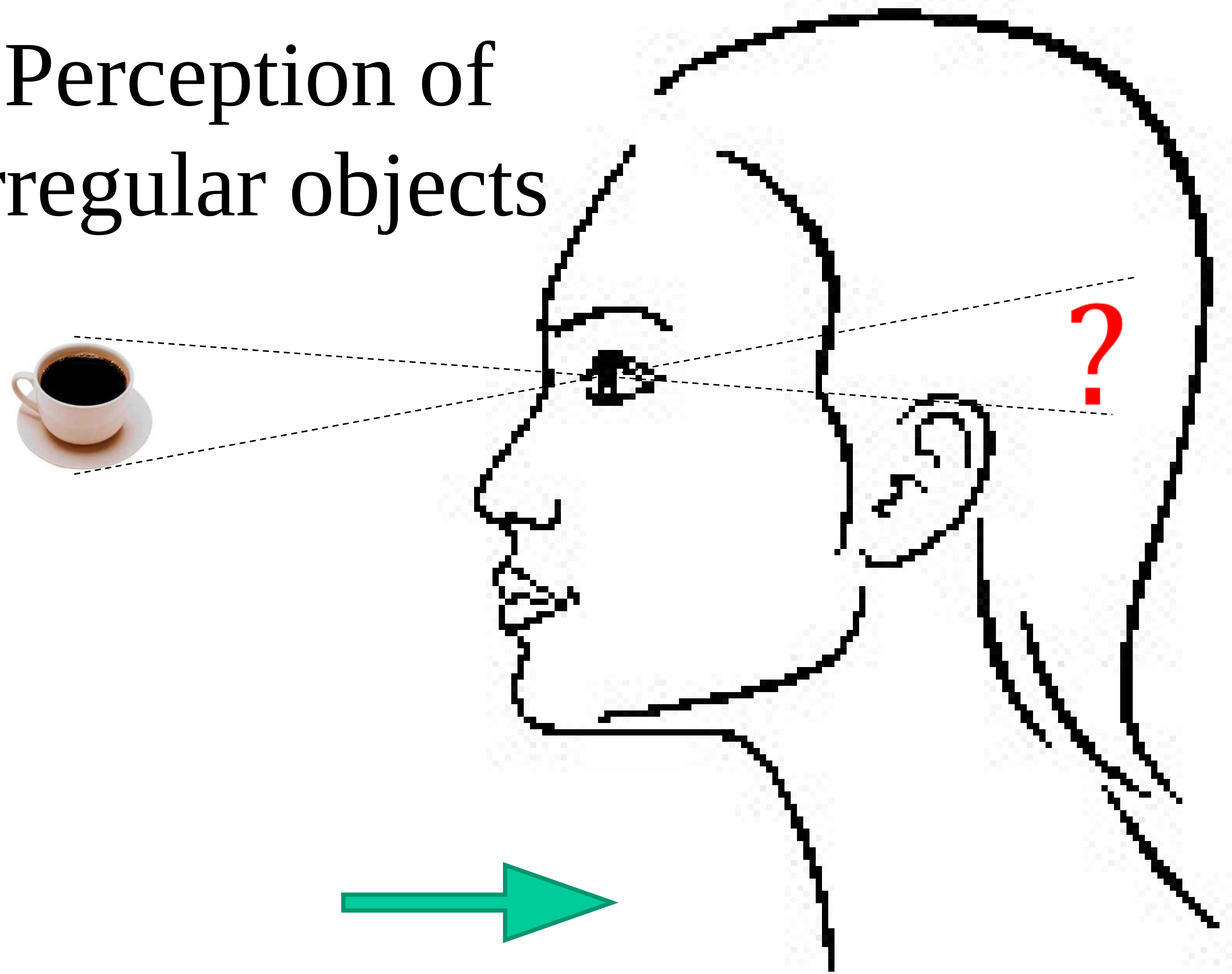
Web page of the subject

www.agentspace.org/kv

Irregular objects

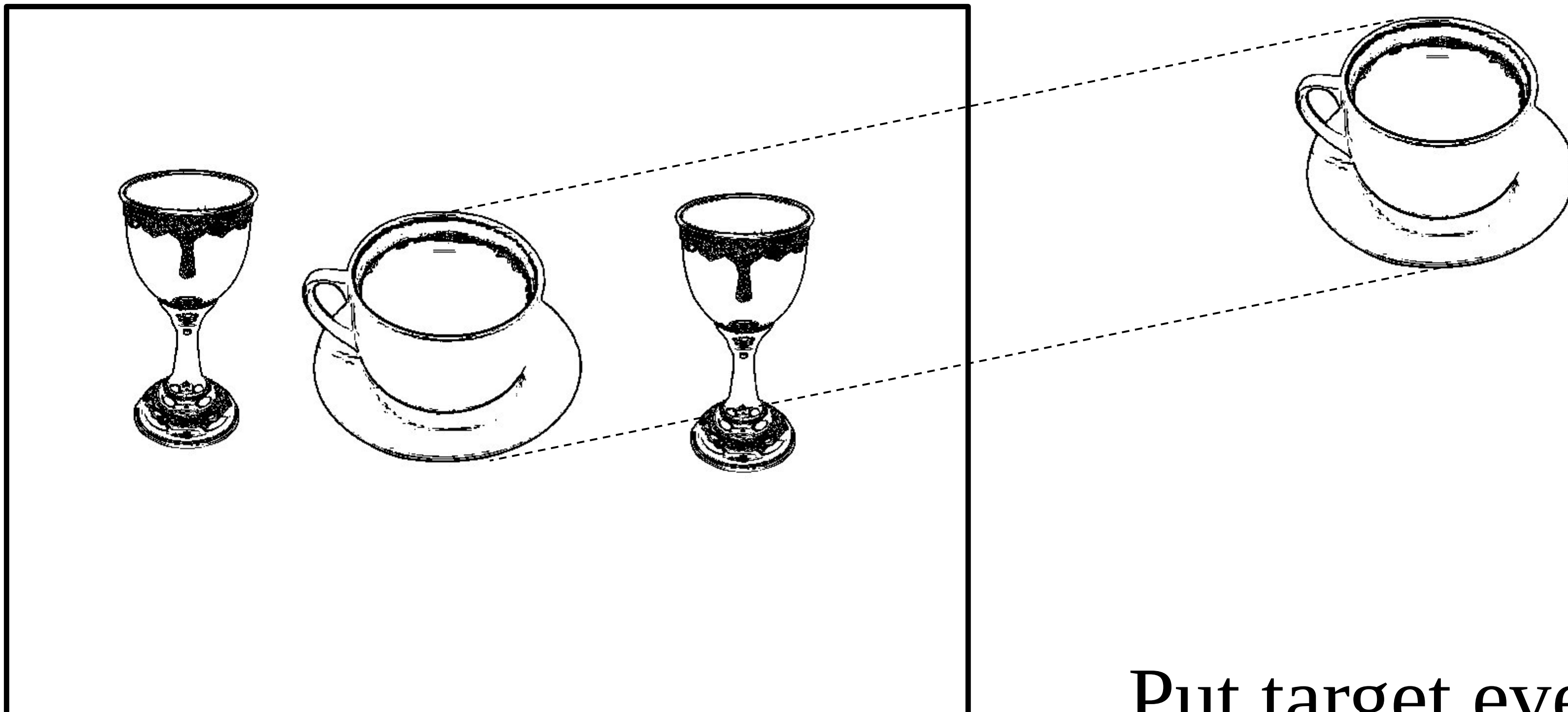
Too many parameters, no rendering algorithm

Perception of irregular objects



Brute force

target

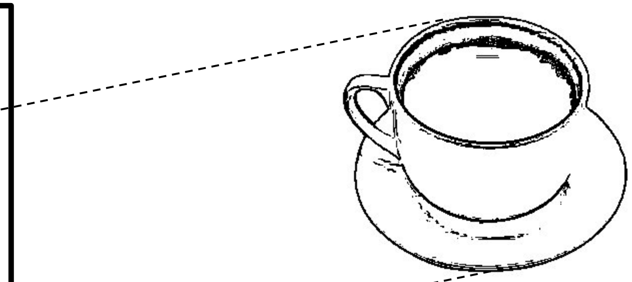


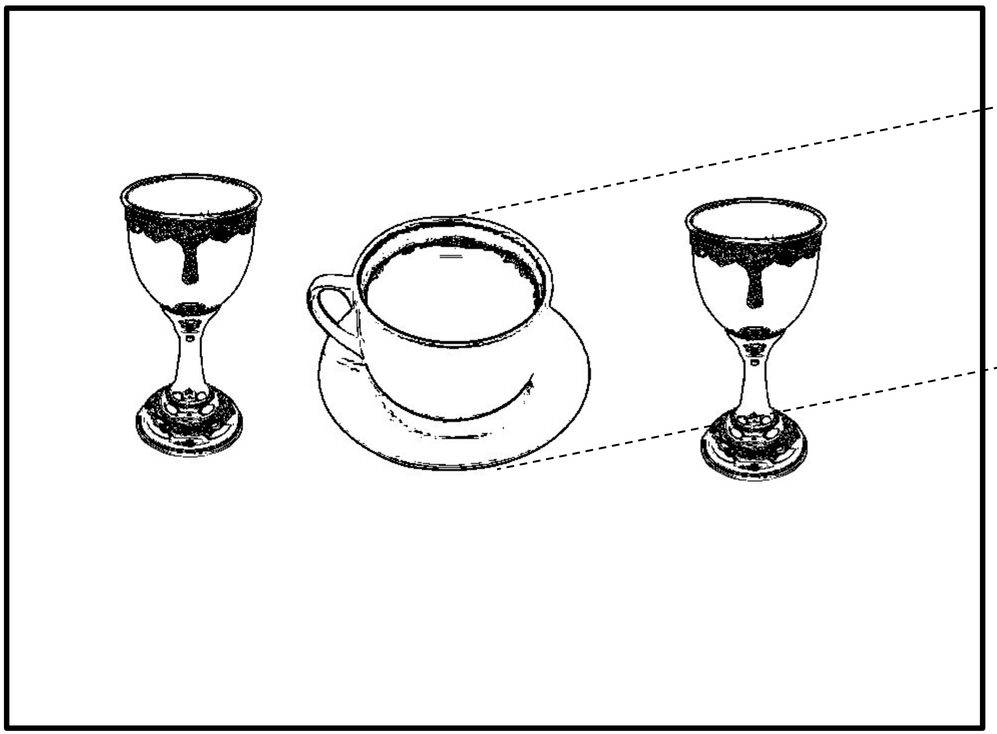
frame

Put target every where, find minimal error

Brute force

reference





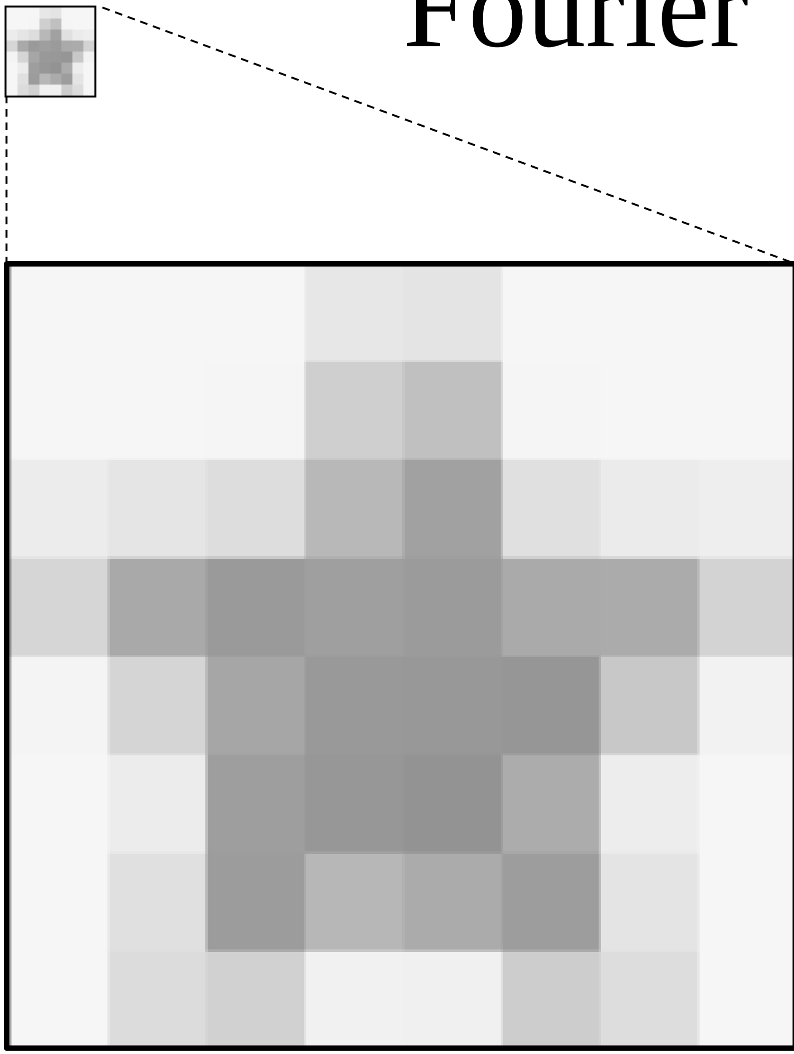
frame

Surprisingly, we have effective algorithm based on Fourier transform

$P \times O(\log P)$

P is number of pixels

Fourier Transform

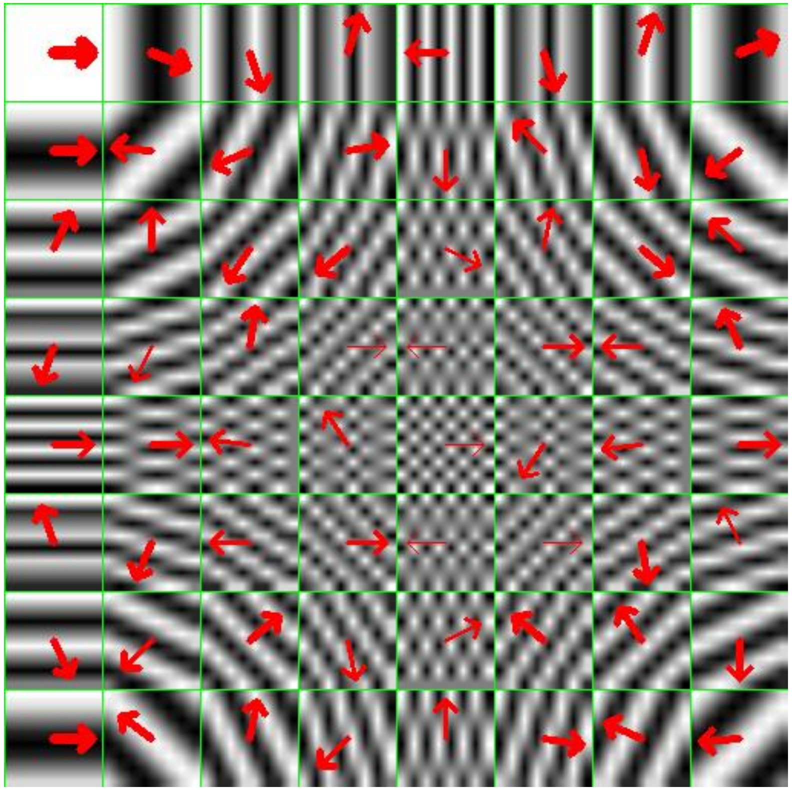


Any image can be effectively expressed as a linear transform of wave images

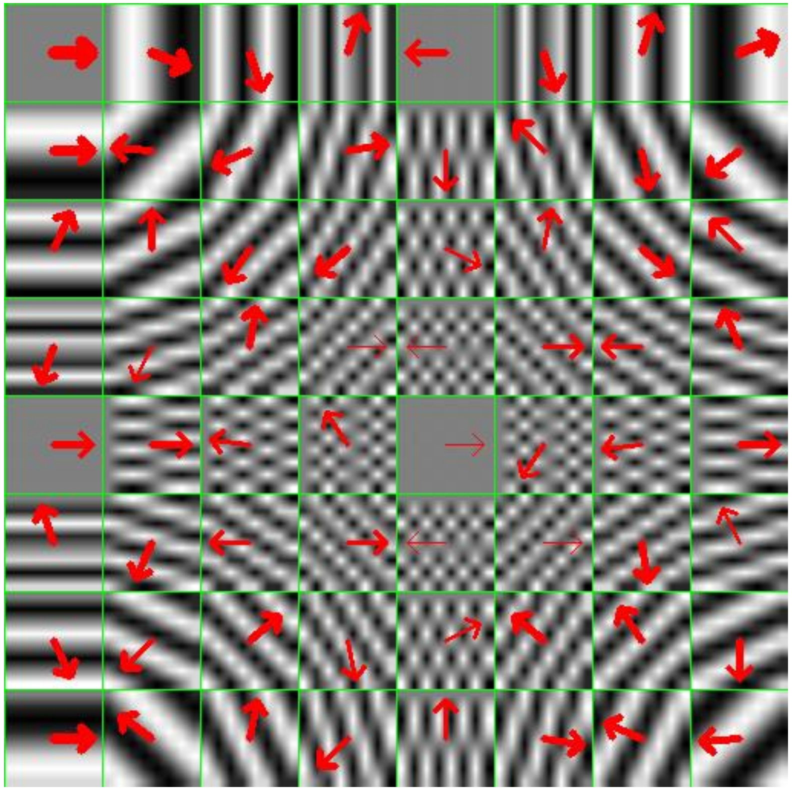
(complex images, complex coefficients)

Fourier Transform

Re



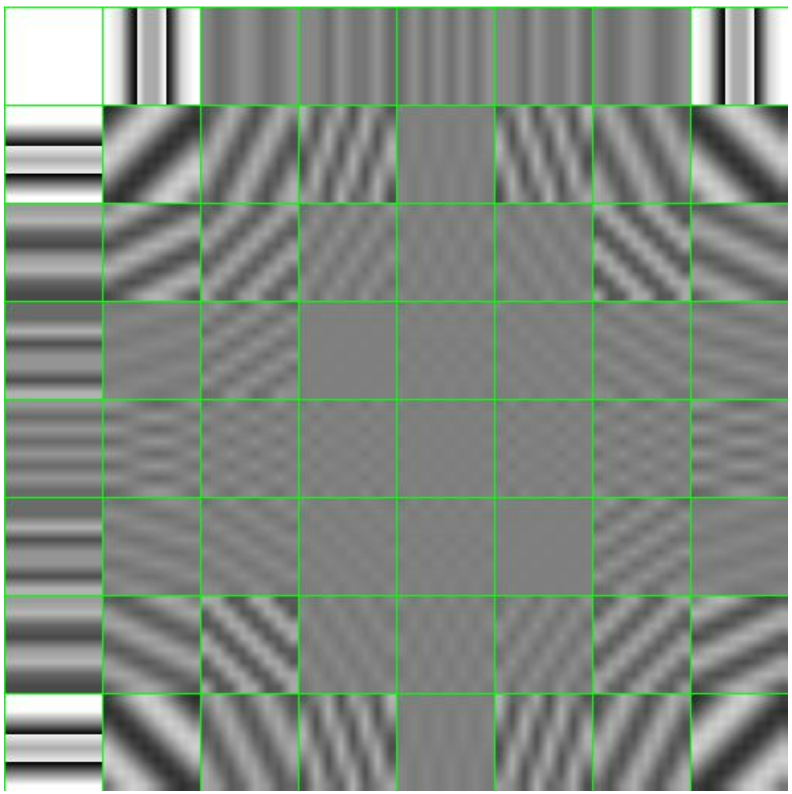
Im



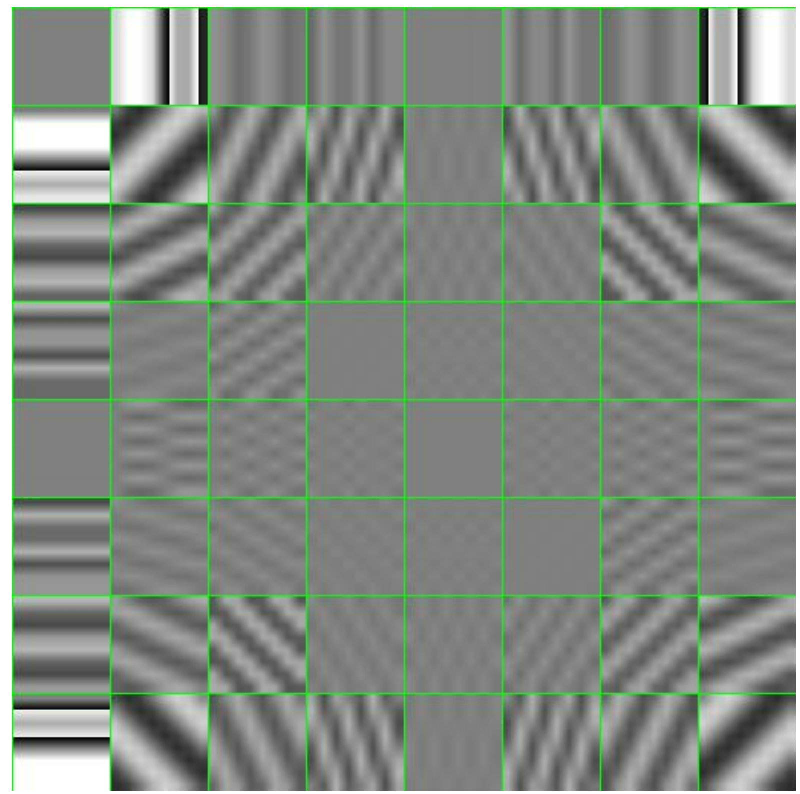
Fourier coefficients are complex numbers, they have amplitude (magnitude) and phase (slope). Phase remains the same when just contrast of image is modified

Fourier Transform

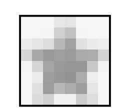
multiplied: Re

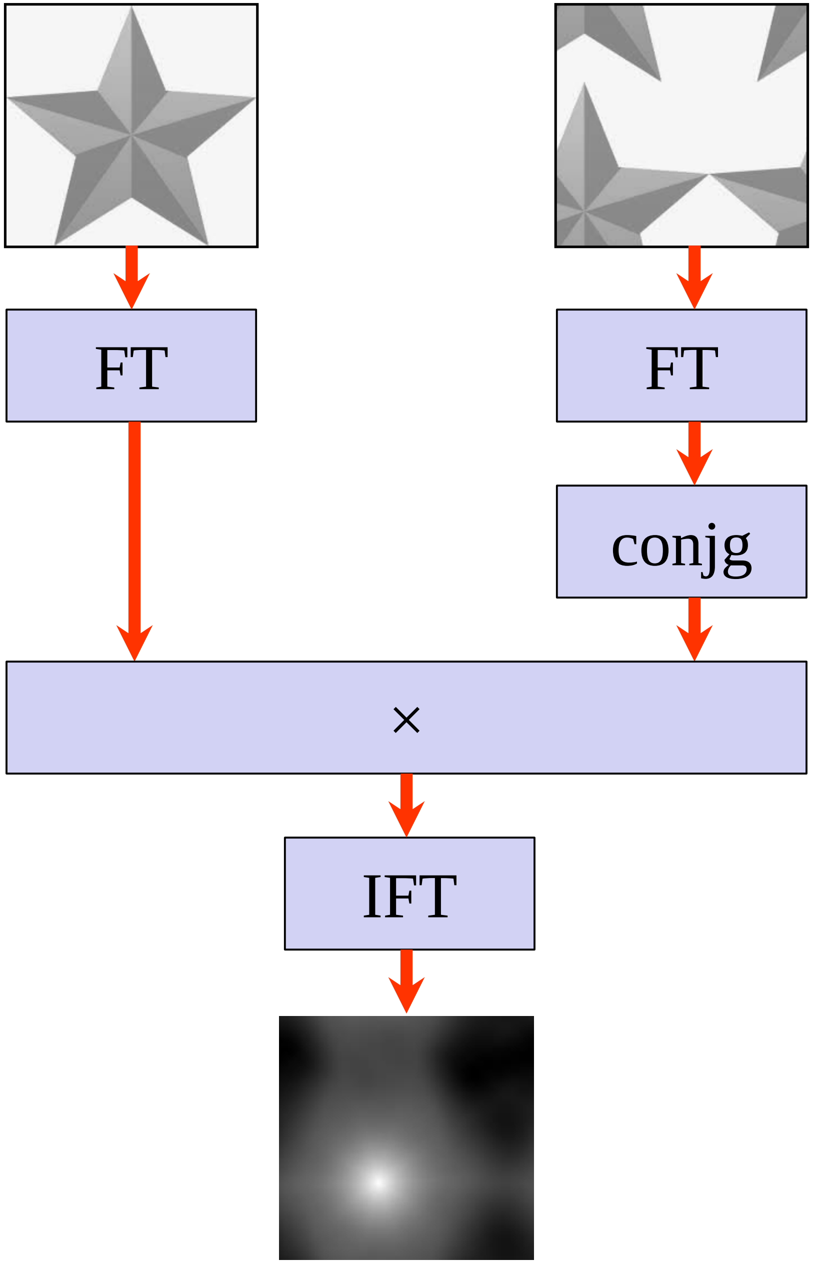


Im



total:

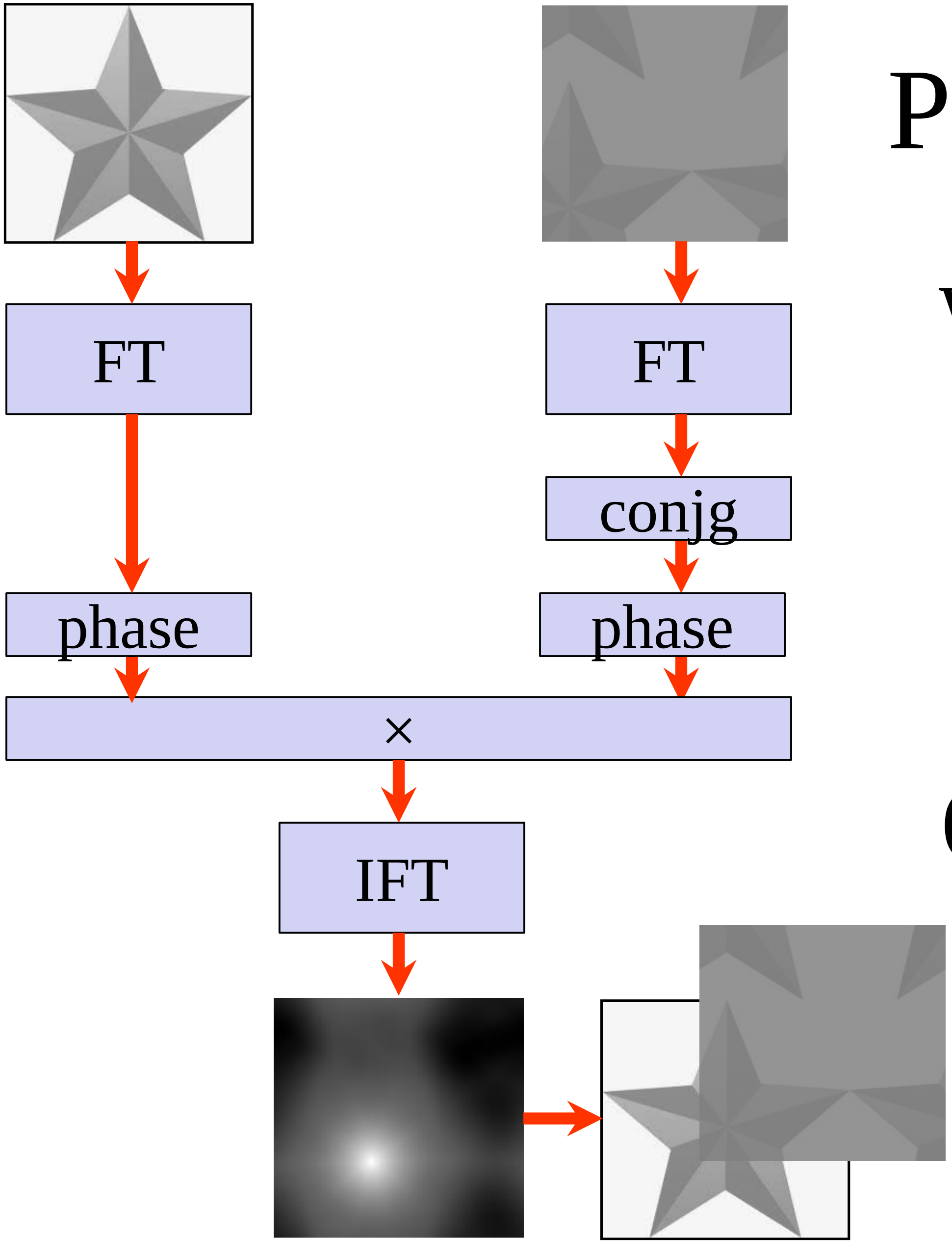




With FT and inverse we can calculate circular convolution of two images.

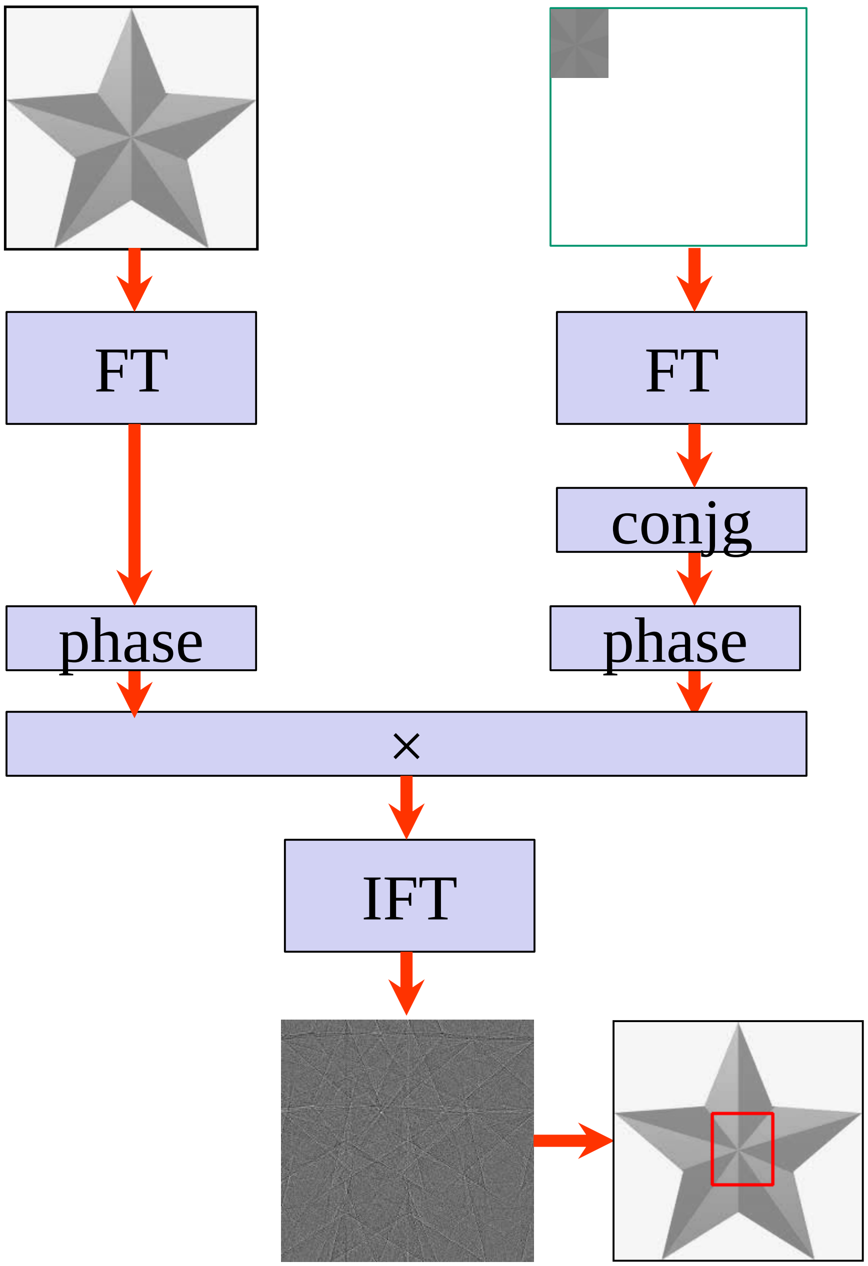
When we reverse one of them and roll by (1,1) we calculated sum of images multiplication for all possible shifts which corresponds to matching error

Phase correlation



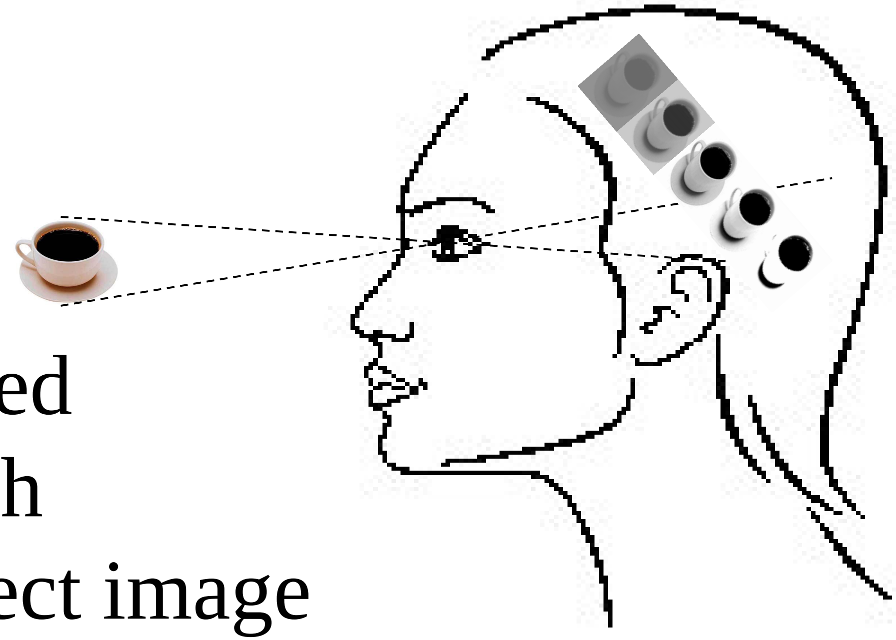
When we pay attention just to the phase, the search becomes independent from lighting condition (contrast modification)

Phase correlation



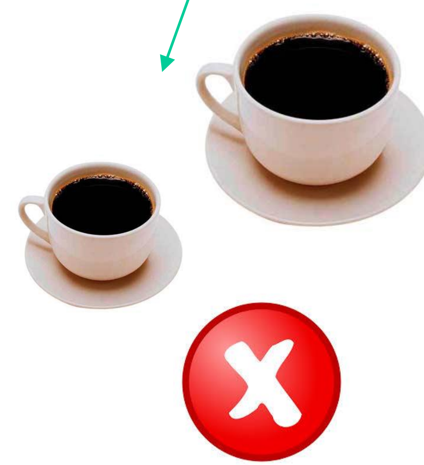
With some risk we can also use padding on part of the image (even with modified contrast) and search for it on the given image

Phase correlation



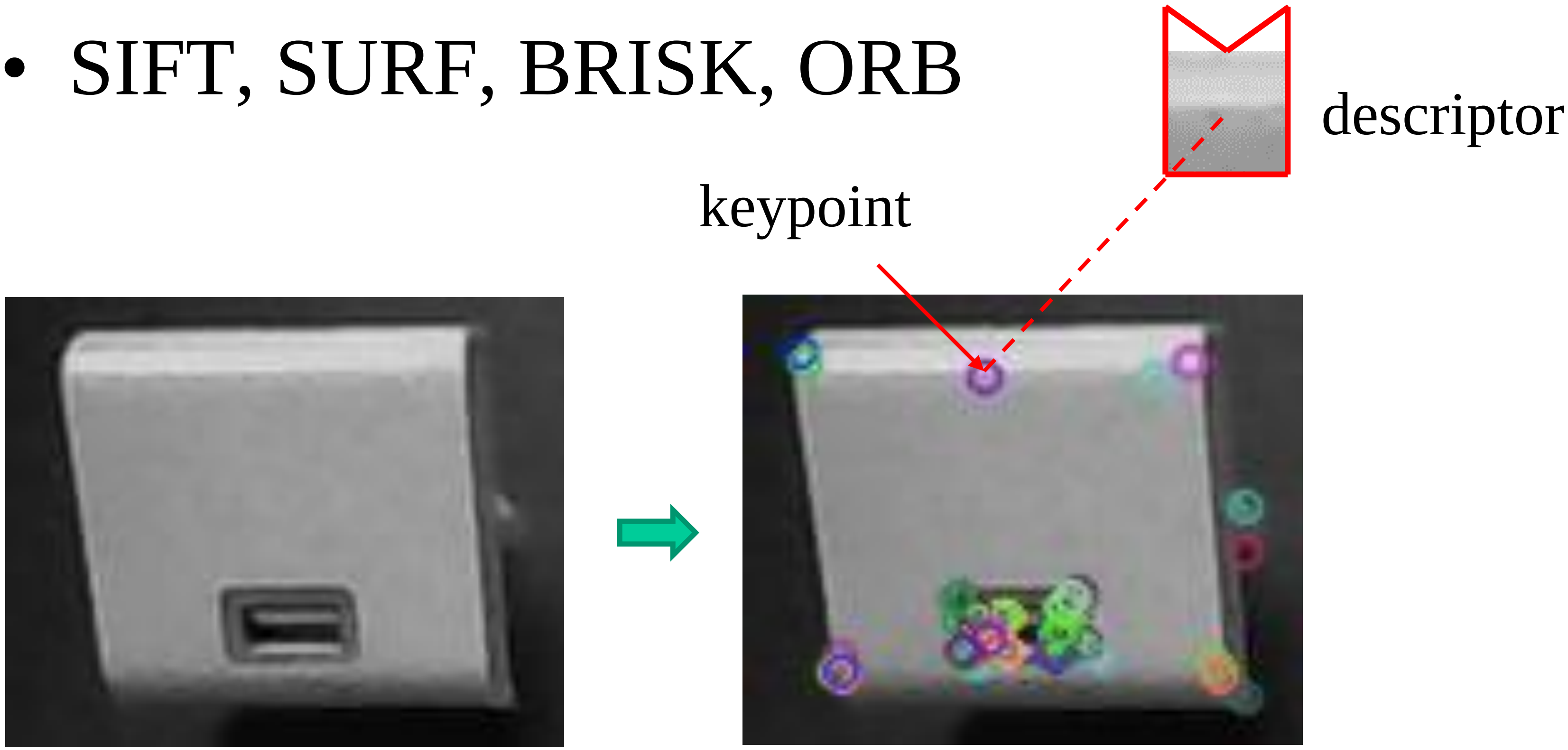
- Object is represented by something which corresponds to object image after application of all possible contrast modifications

could be extended to comply



Feature detectors

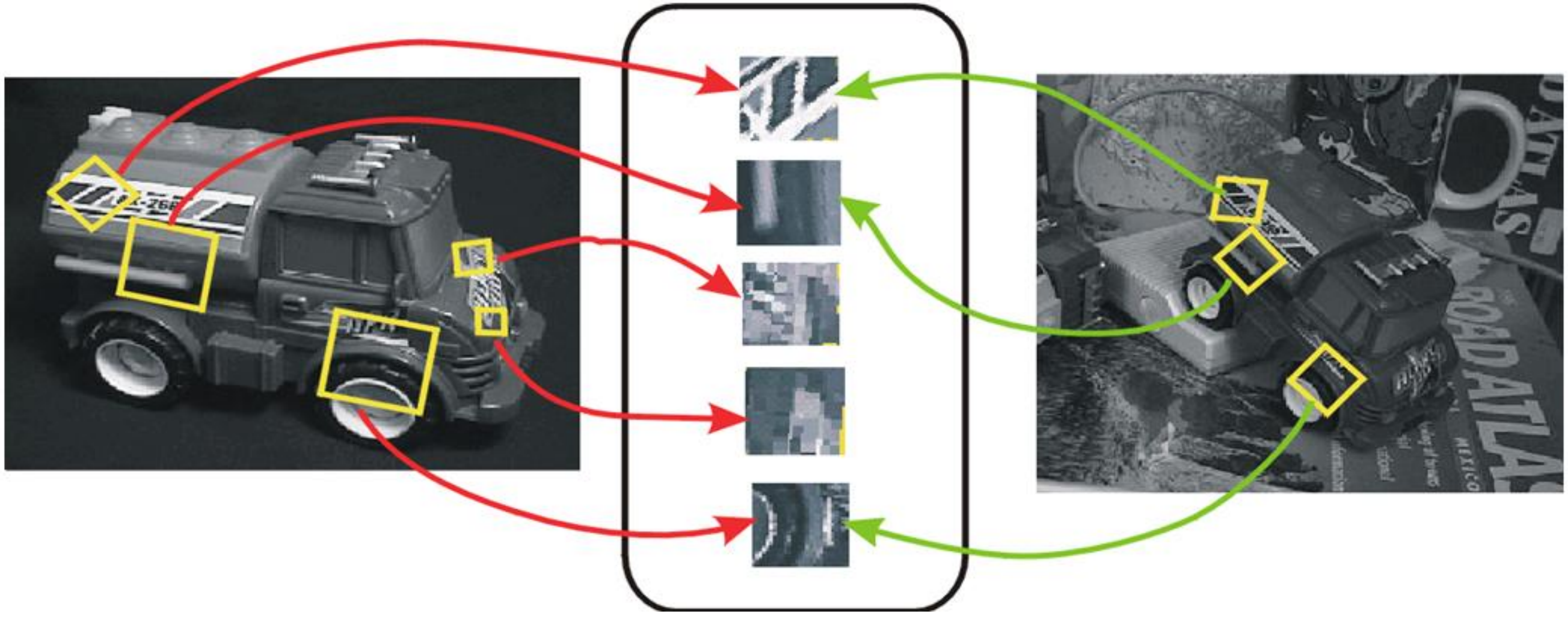
- SIFT, SURF, BRISK, ORB



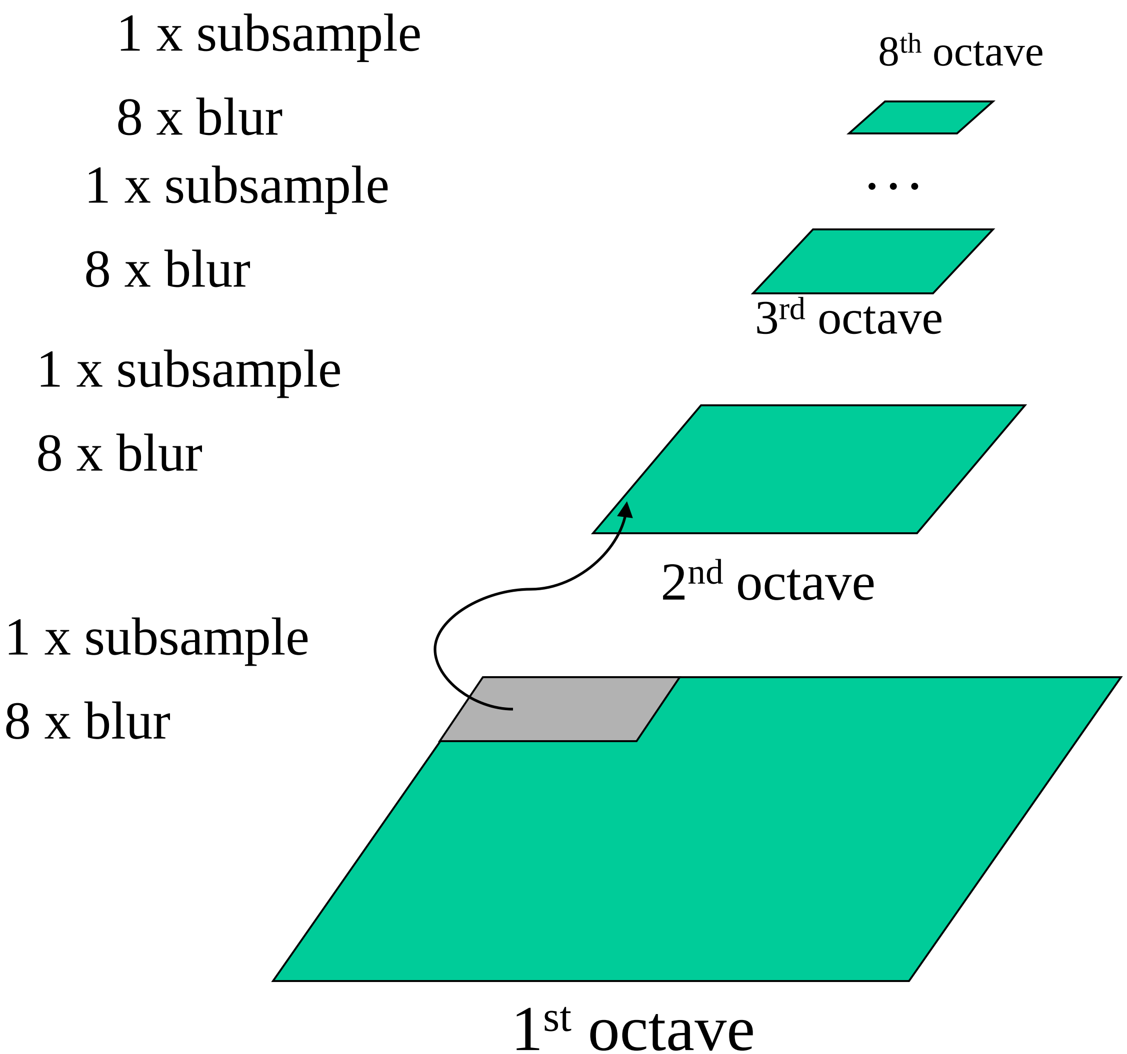
Feature detector provides set of interesting points and describe their vicinity by descriptor in such a way that it is almost invariant to translation, scale and rotation

SIFT (Scale Invariant Feature Transform)

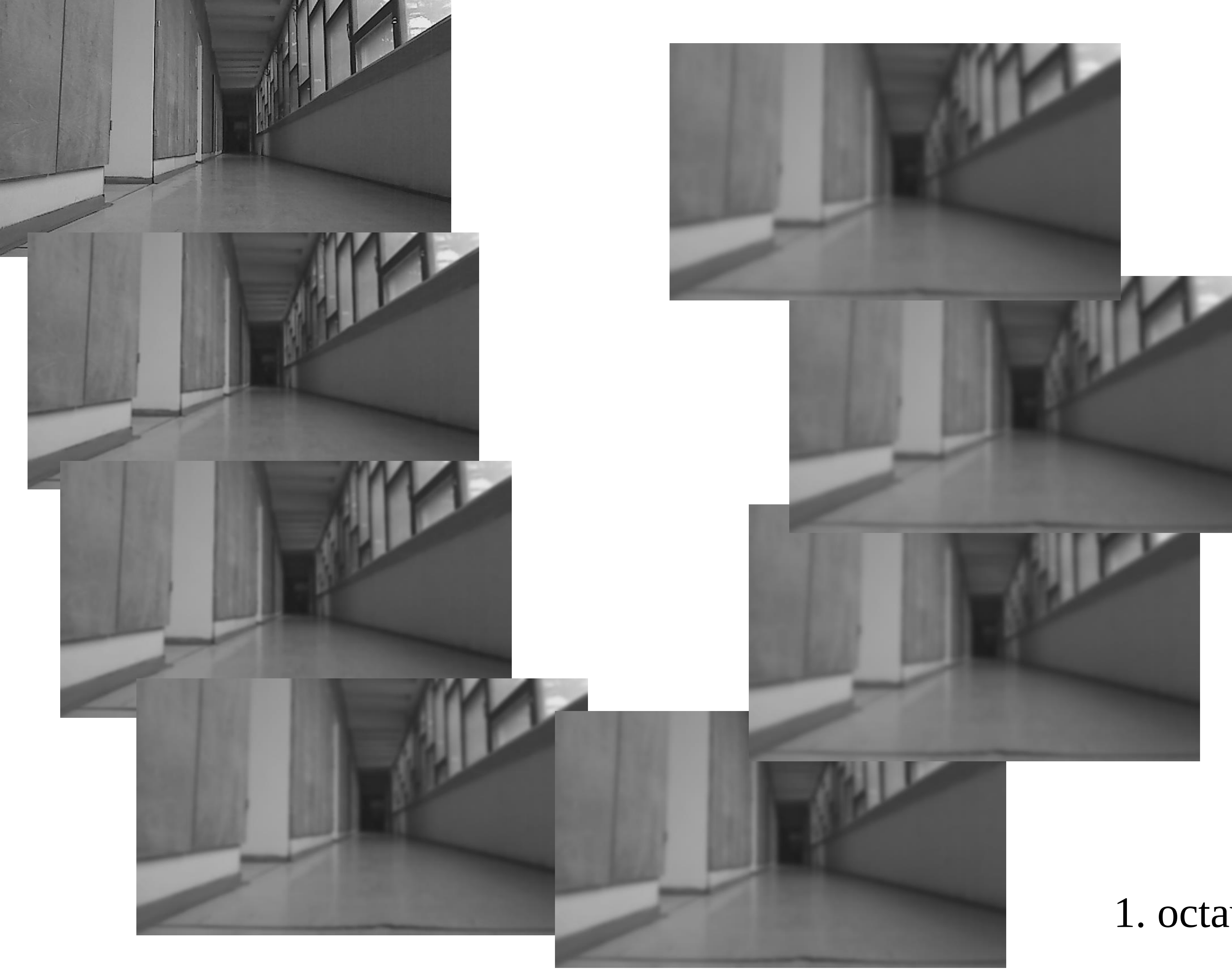
As a result we find points with similar descriptors on two images of the same object and some of them we can pair



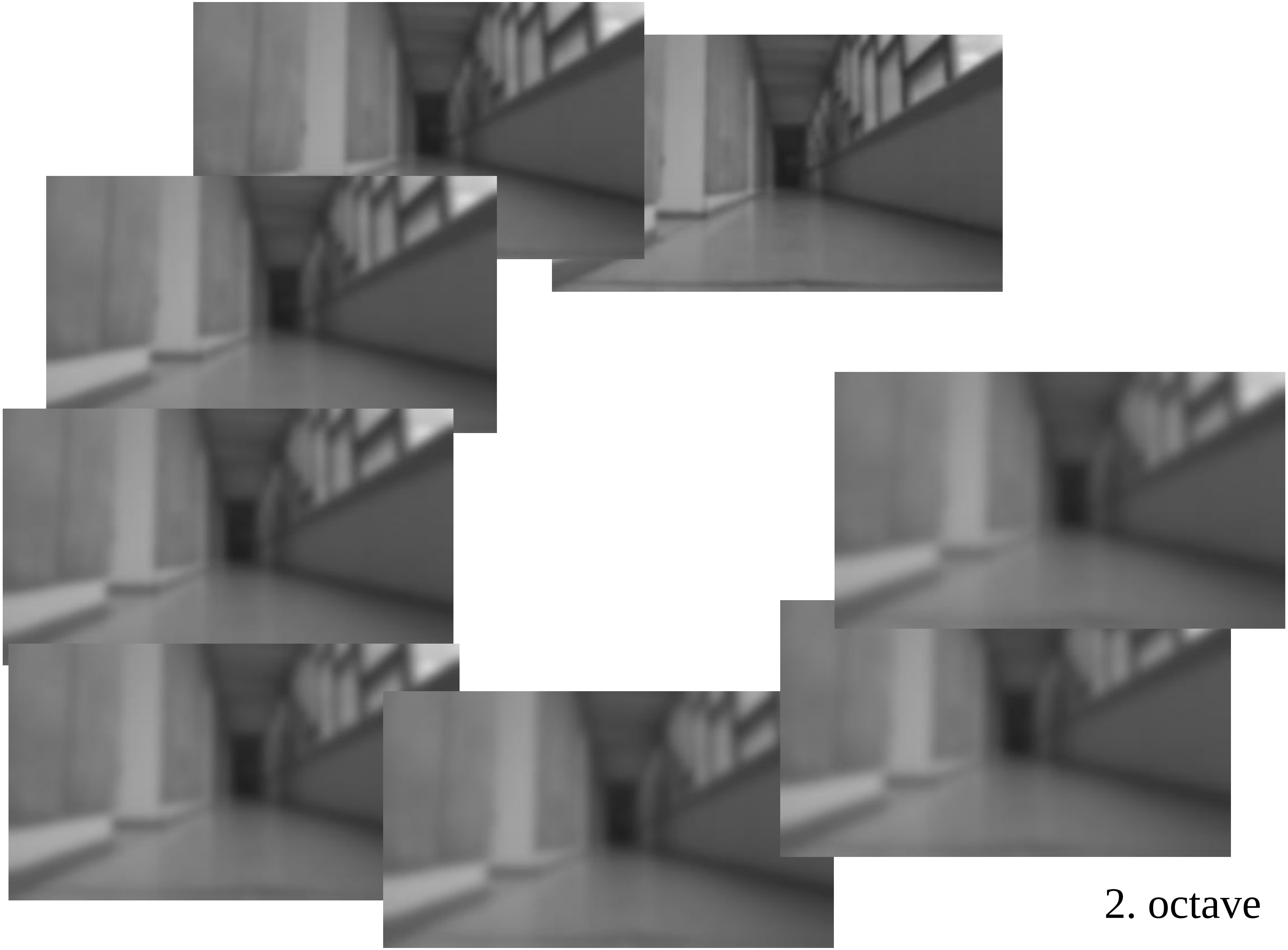
SIFT (Scale Invariant Feature Transform)



Keypoints correspond to significant details lost in process of subsampling of the image



1. octave



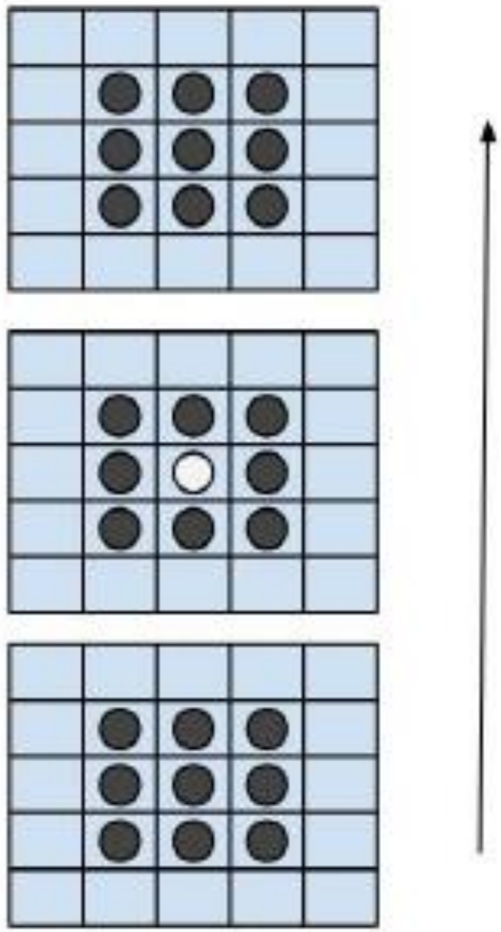
2. octave

3. octave

4. octave

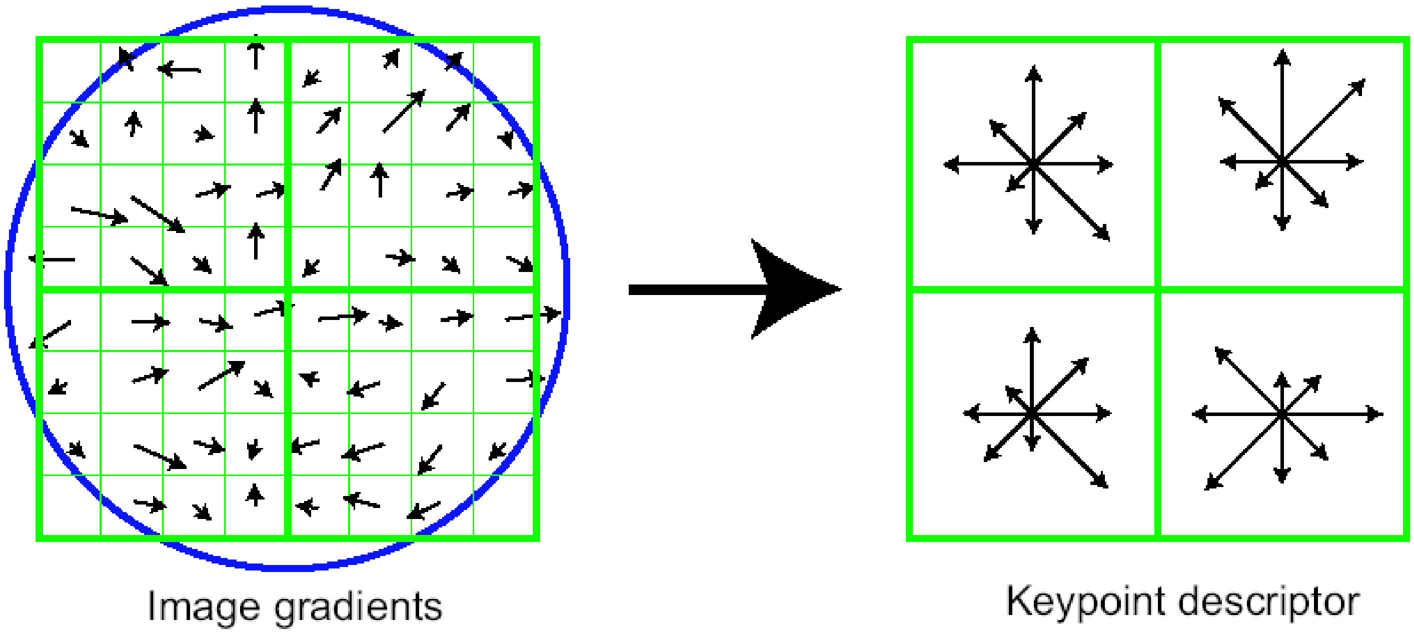
…

SIFT (Scale Invariant Feature Transform)



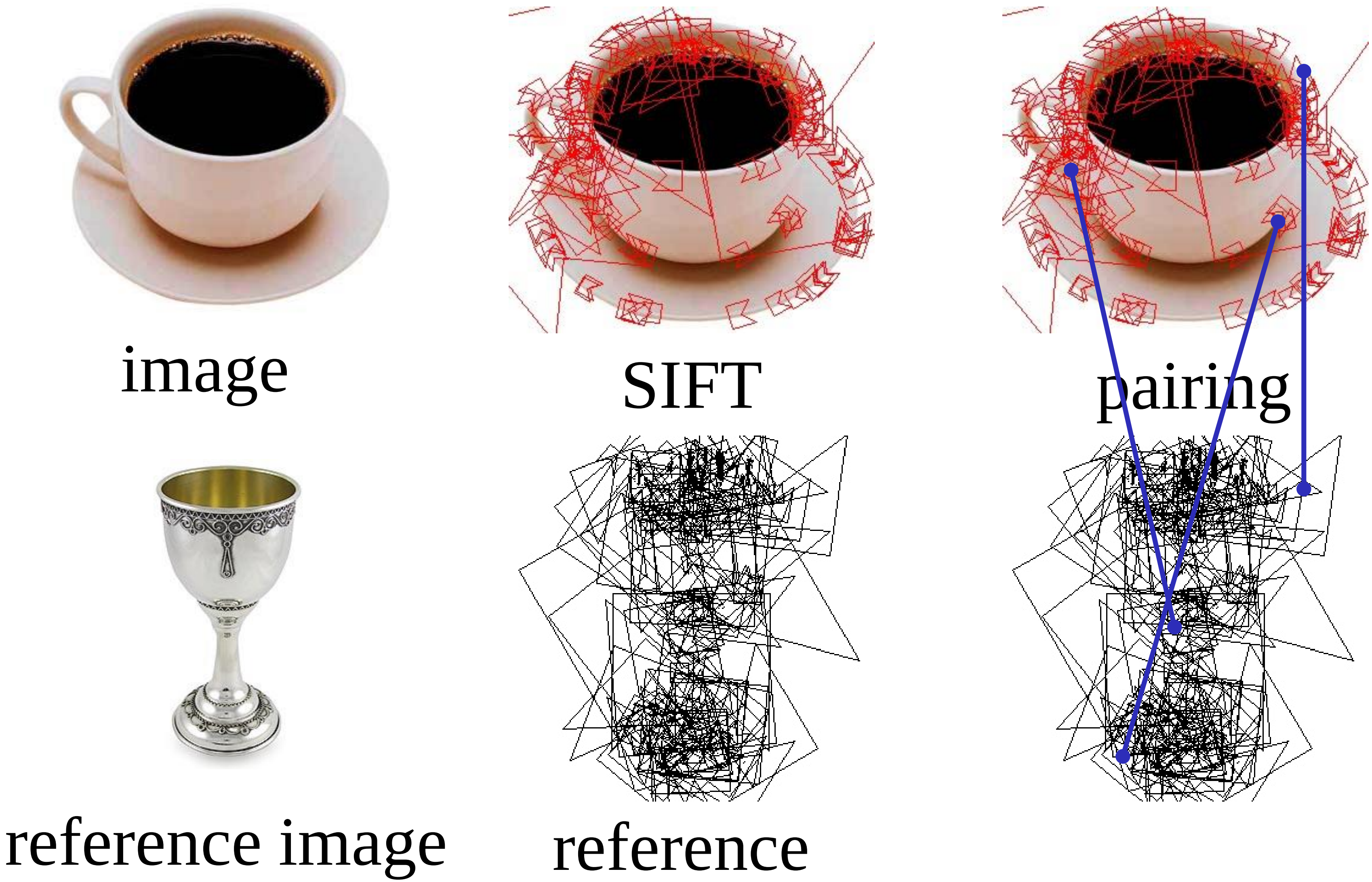
Keypoint has extremal intensity in 3D pyramid compounded from octaves

SIFT (Scale Invariant Feature Transform)



Each keypoint is associated with descriptor of its vicinity: distribution of edge gradients in quadrants

Object detection can be made by statistical pairing of descriptors



$$\begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Scaling matrix

$$\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Rotation matrix

$$\begin{bmatrix} s_x & s_{xy} \\ s_{yx} & s_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Linear transform

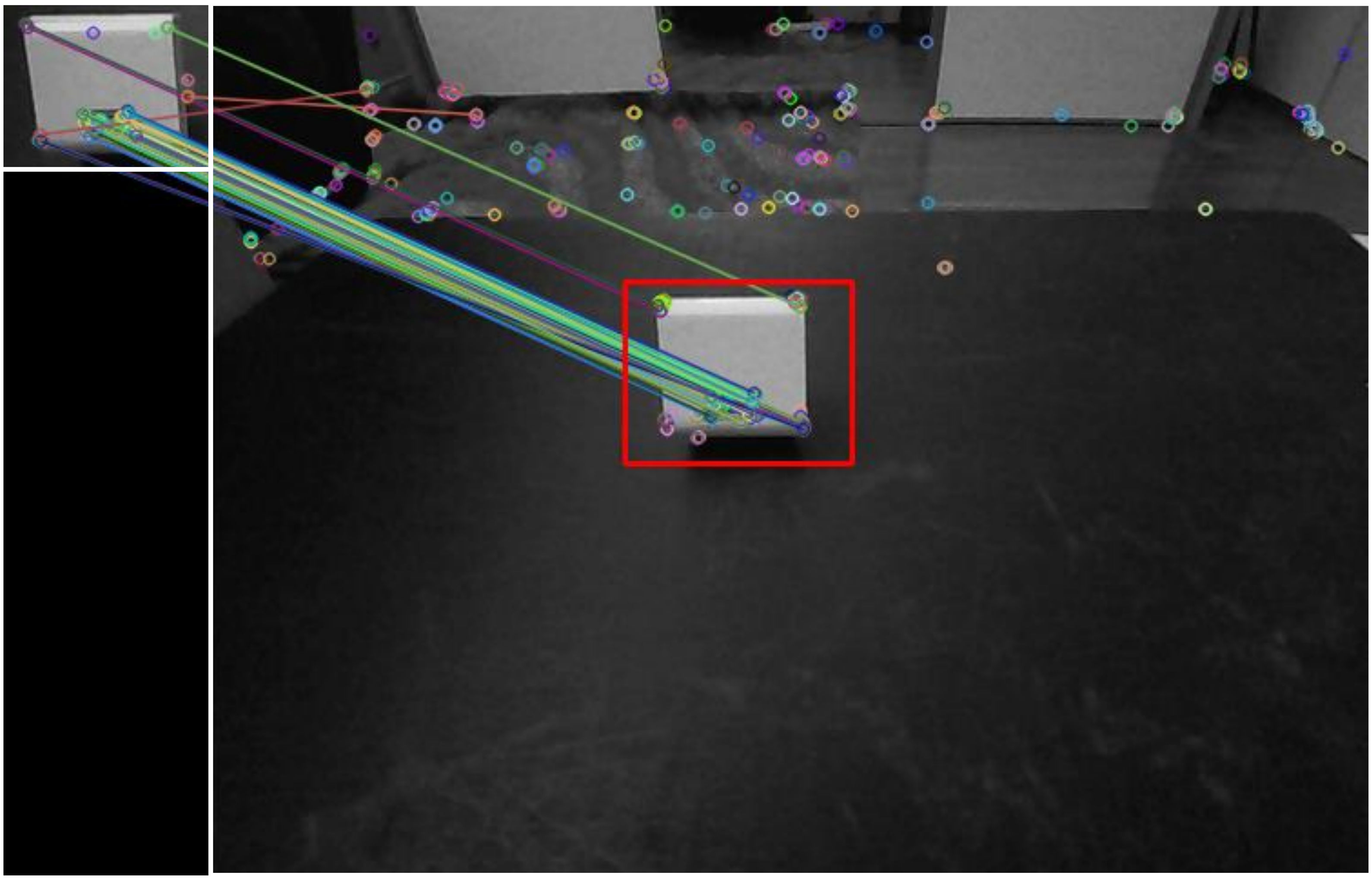
$$\begin{bmatrix} s_x & s_{xy} & t_x \\ s_{yx} & s_y & t_y \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Affine transform

$$\begin{bmatrix} s_x & s_{xy} & t_x \\ s_{yx} & s_y & t_y \\ t_{yx} & t_{xy} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

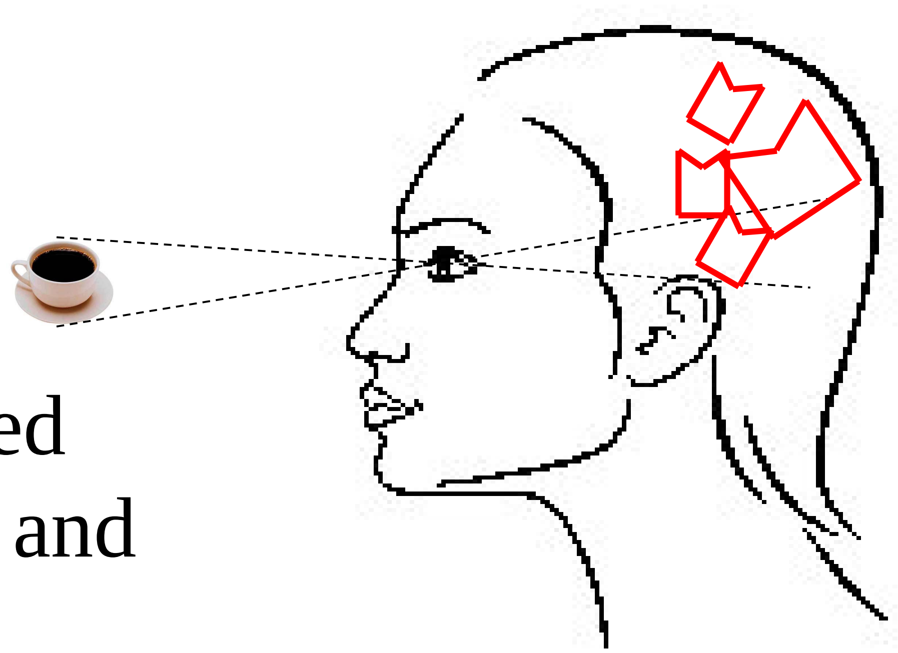
Homography

When found pairs corresponding to the same homography, object is found



Feature detector

- Object is represented by set of keypoints and their descriptors

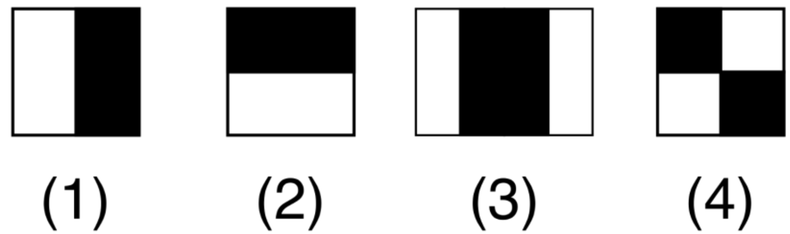


Trackers

- Detector process each image separately
- Tracker works with video and can employ information from previous images
- Simplest tracker: detector + outlier filtering (e.g. by Kalman filter)
- Boosting, MIL, KFC, TLD, MedianFlow

Multiply instance learning (MIL)

- Haar features



we put on ROI set of 2, 3 or 4 rectangles
and calculate difference of white and black
areas totals

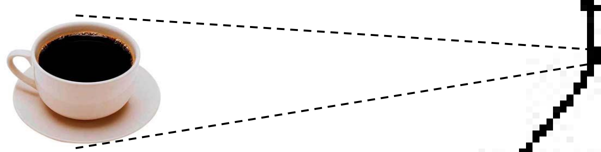
- Initial Region of Interest (ROI)

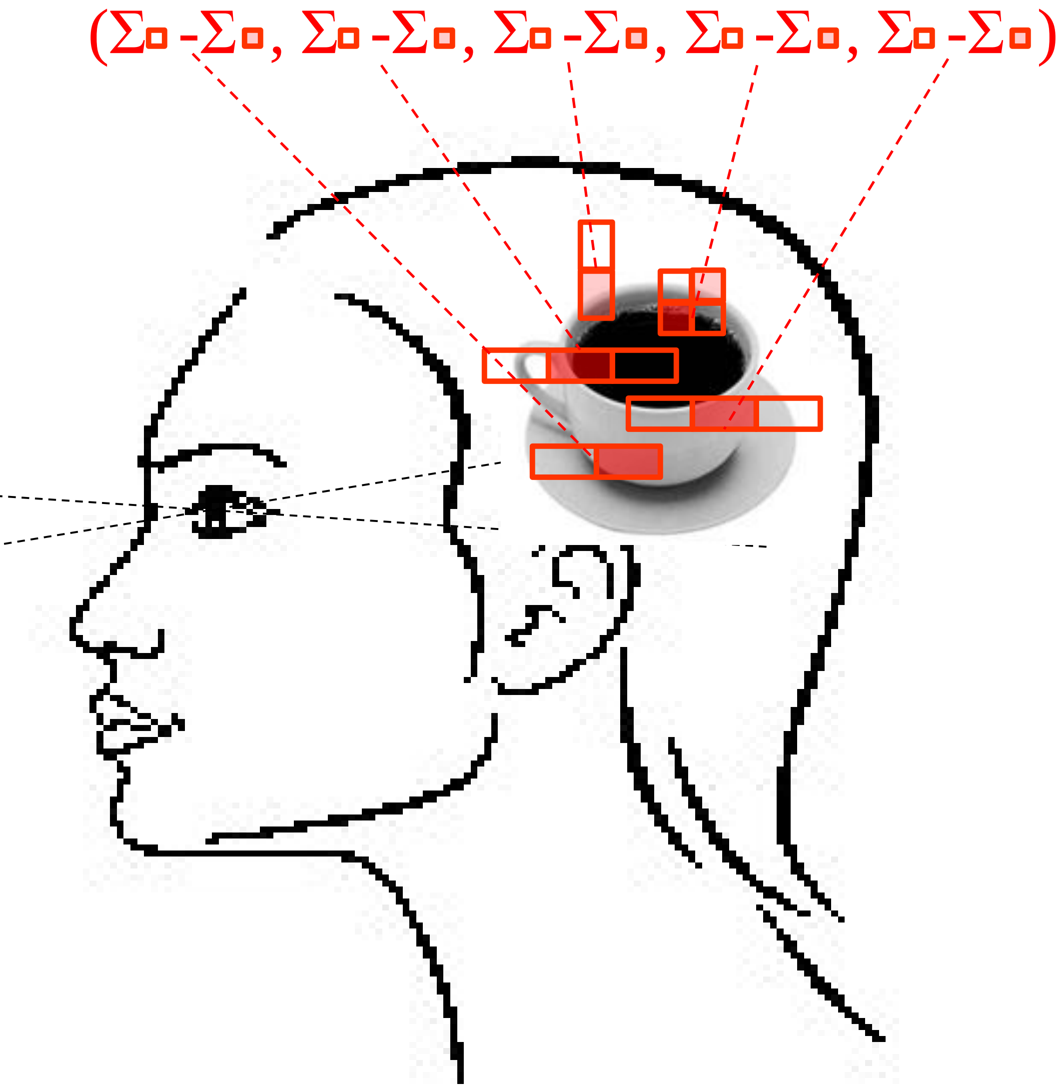
Then we teach Bayes classifier of bags of positive (ROIs close to the current ROI) and negative (far ROIs) examples

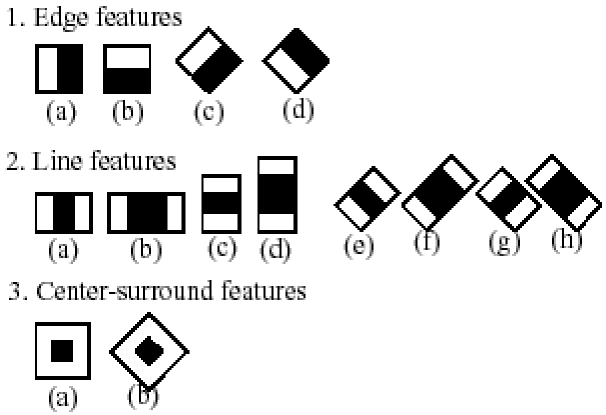
- Motion model: neighborough ROI

we move to neighborough ROI which has the highest ranging provided by the classifier

MIL tracker






1. Edge features
(a) (b) (c) (d)
2. Line features
(a) (b) (c) (d) (e) (f) (g) (h)
3. Center-surround features
(a) (b)