Comparing feature detectors

A bias in the repeatability criteria.

I. Rey-Otero M. Delbracio J.M. Morel

CMLA, ENS-Cachan ECE, Duke University

Comparing images...

…an avalanche of applications...

- *Structure from motion*
- *Object recognition*
- *Video stabilisation*
- *Robot self-localisation*
- *Hand gesture recognition*
- *Video query-by-image*

...

… and methods.

- *MOPS*
- *OBR*
- *SURF*
- *SIFER*
- *KAZE*
- *ASIFT*
- *BRIEF*
- *SFOP*
- *PCA-SIFT*
- *Hessian/Harris Laplace*
- *MSER*
- *EBR/IBR*

Motivation

Choosing a detector, a difficult task!

- *Huge number of published detectors*
- *Various types of detectors*
- *Different requirements (types of image transformations)*

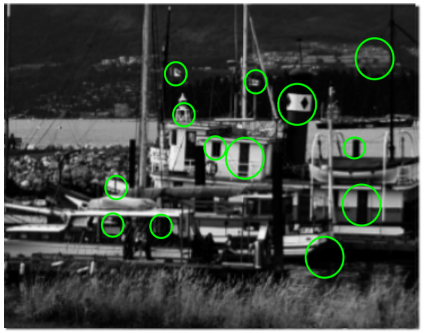
We need a general comparison framework

Repeatability [Mikolajczyk 2005]

$\leftrightsquigarrow$ Homography H_{ab}

u_a u_b

Repeatability [Mikolajczyk 2005]



u_a

$\leftrightsquigarrow$ Homography H_{ab}

u_b

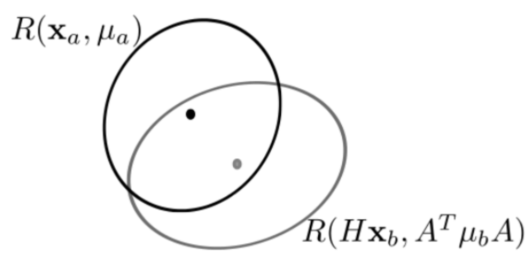
Repeatability *[Mikolajczyk 2005]*

In this framework, a detection is an elliptical region:

$$R(\mathbf{x}, \mu) = \left\{ \mathbf{x}' \in \Omega \mid (\mathbf{x}' - \mathbf{x})^T \mu (\mathbf{x}' - \mathbf{x}) \leq 1 \right\}.$$

A <u>repeated detection</u> is a region that significantly overlaps with a reprojected region:

$$\frac{\left| R(\mathbf{x}_a, \mu_a) \cap R(H\mathbf{x}_b, A^T \mu_b A) \right|}{\left| R(\mathbf{x}_a, \mu_a) \cup R(H\mathbf{x}_b, A^T \mu_b A) \right|} > 60\%$$



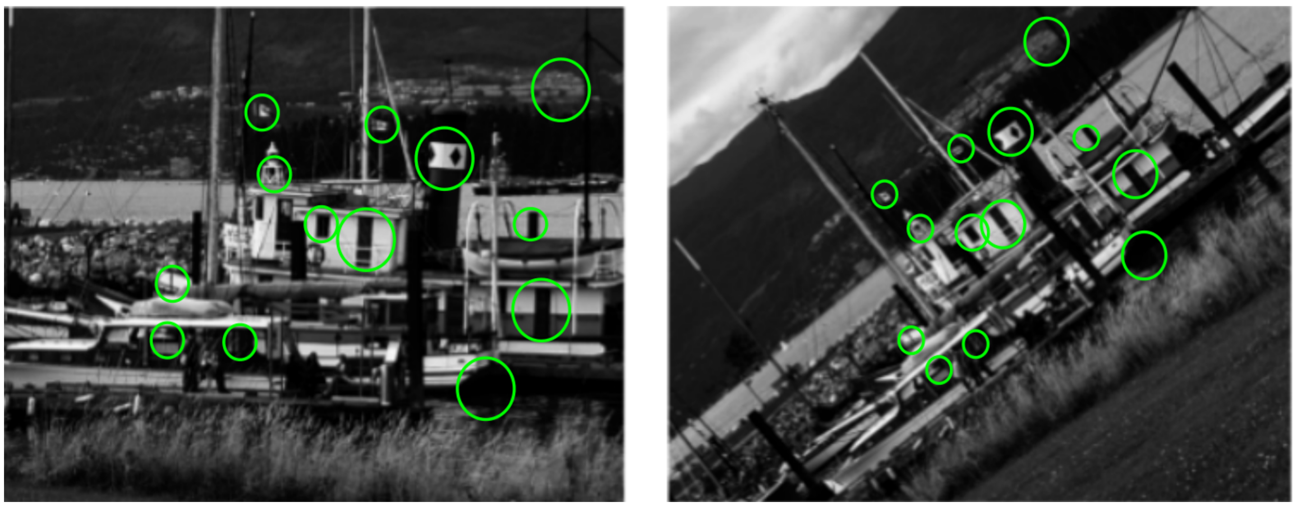
Repeatability *[Mikolajczyk 2005]*

The classic performance metric

$$\text{Repeatability rate} = \frac{\text{Number of repeated detections}}{\text{Total number of detections}}$$

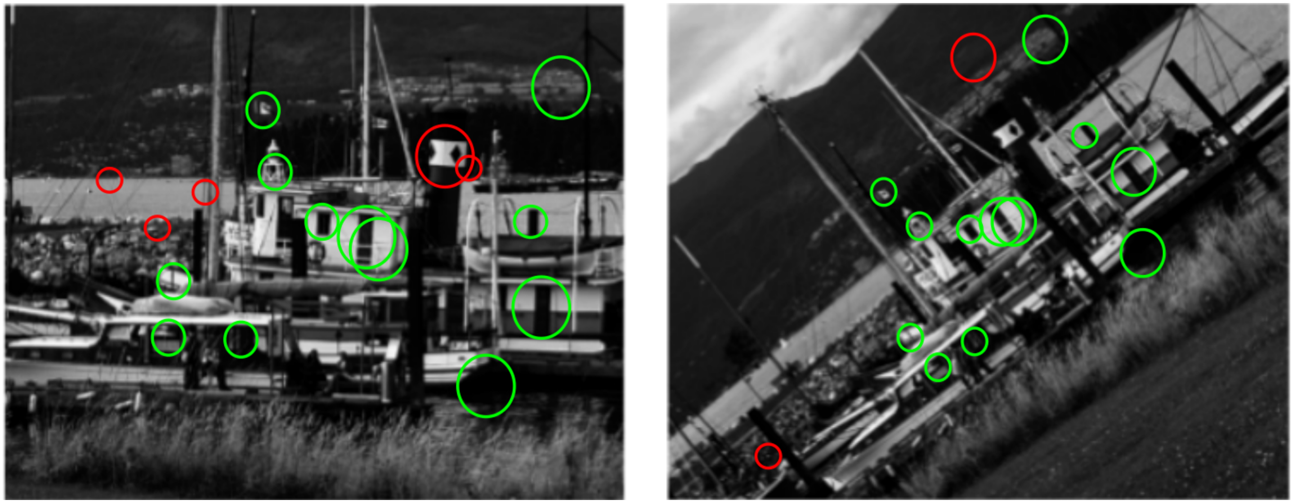
gives an idea of the <u>benefit over cost</u> ratio for a detector

Repeatability criterion favors redundancy

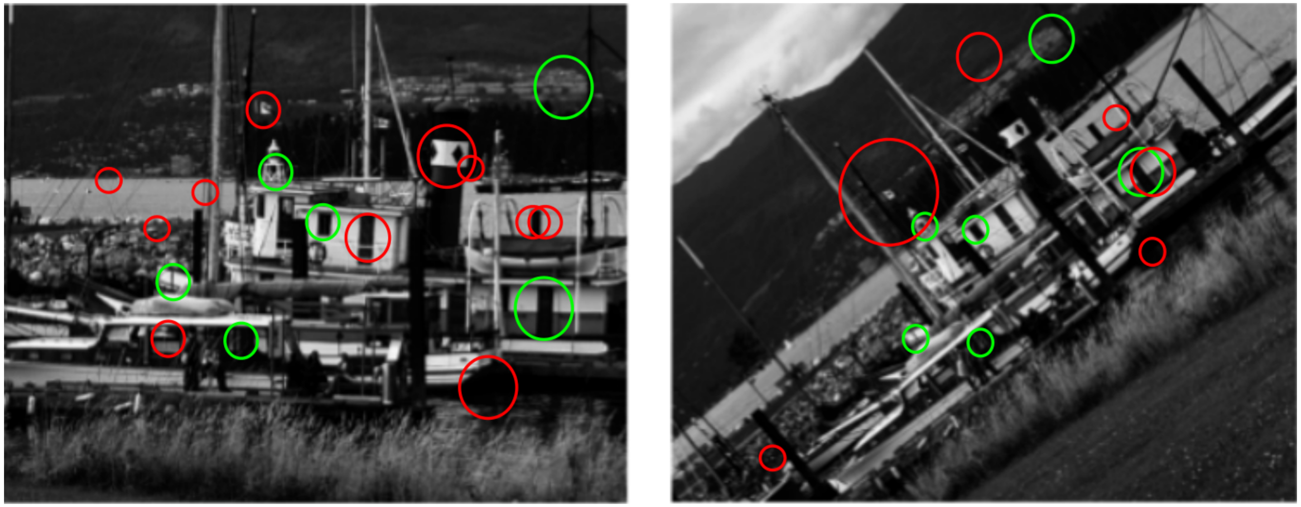


A perfect detector

Repeatability criterion favors redundancy



Repeatability criterion favors redundancy



Repeatability criterion favors redundancy

A useless detector...

...with good repeatability

Some popular methods are redundant

SIFT (17)

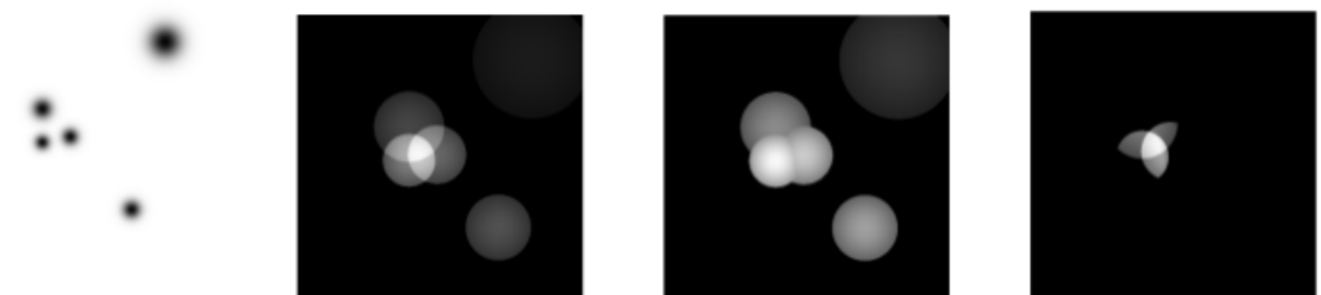
Harris-Laplace (242)

Taking into account descriptors overlap

Assign a mask function to each detection k

$$f_k(\mathbf{x}) = K e^{-\frac{1}{2\zeta^2}(\mathbf{x}-\mathbf{x}_k)^T \Sigma_k^{-1} (\mathbf{x}-\mathbf{x}_k)}$$

$f_k(\mathbf{x})$ denotes the contribution of the pixel **x** to the detection **k**



Input detections $k \in \mathcal{K}$

$\sum_{k\in\mathcal{K}} f_k(\mathbf{x})$

$\max_{k\in\mathcal{K}} f_k(\mathbf{x})$

$\sum_{k\in\mathcal{K}} f_k(\mathbf{x}) - \max_{k\in\mathcal{K}} f_k(\mathbf{x})$

Number of detections

$$K = \int_\Omega \sum_{k\in\mathcal{K}} f_k(\mathbf{x}) d\mathbf{x}$$

Num of *non-redundant* detections

$$K_{\text{nr}} = \int_\Omega \max_{k\in\mathcal{K}} f_k(\mathbf{x}) d\mathbf{x}$$

Taking into account descriptors overlap

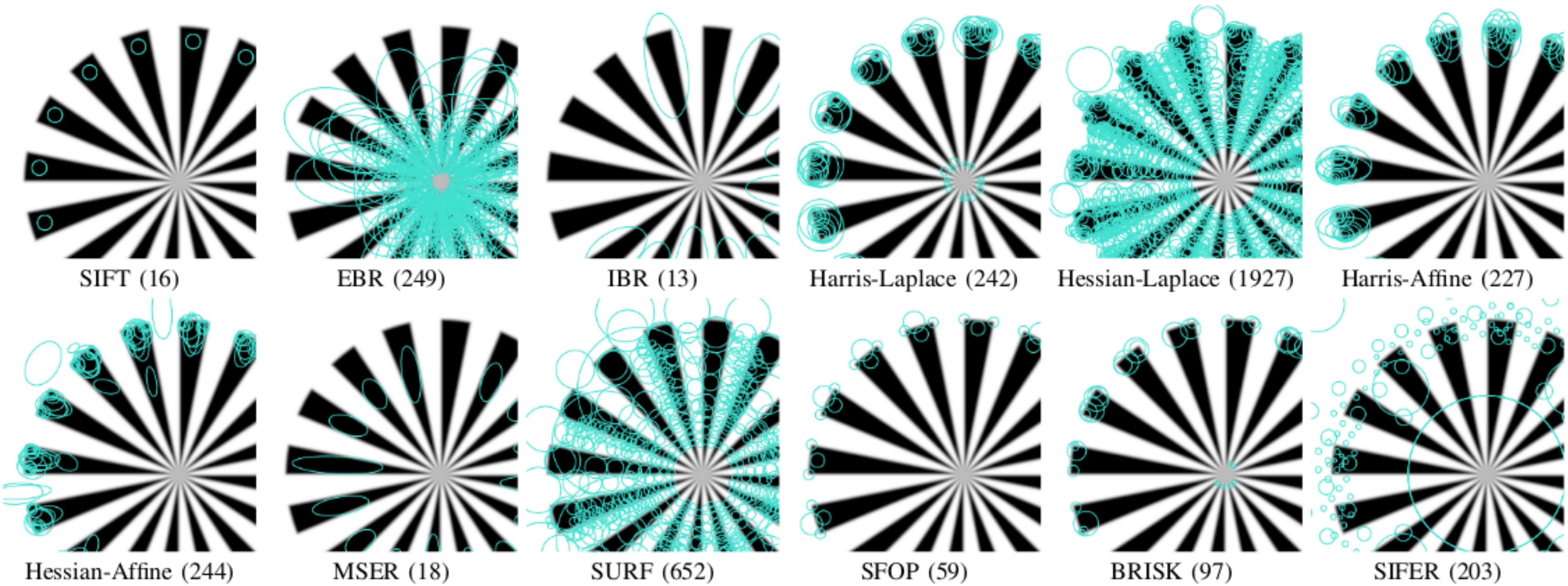
The classic metric does not take the descriptor's overlap into account

$$\cancel{\text{Repeatability rate} = \frac{\text{Number of repeated detections}}{\text{Total number of detections}}}$$

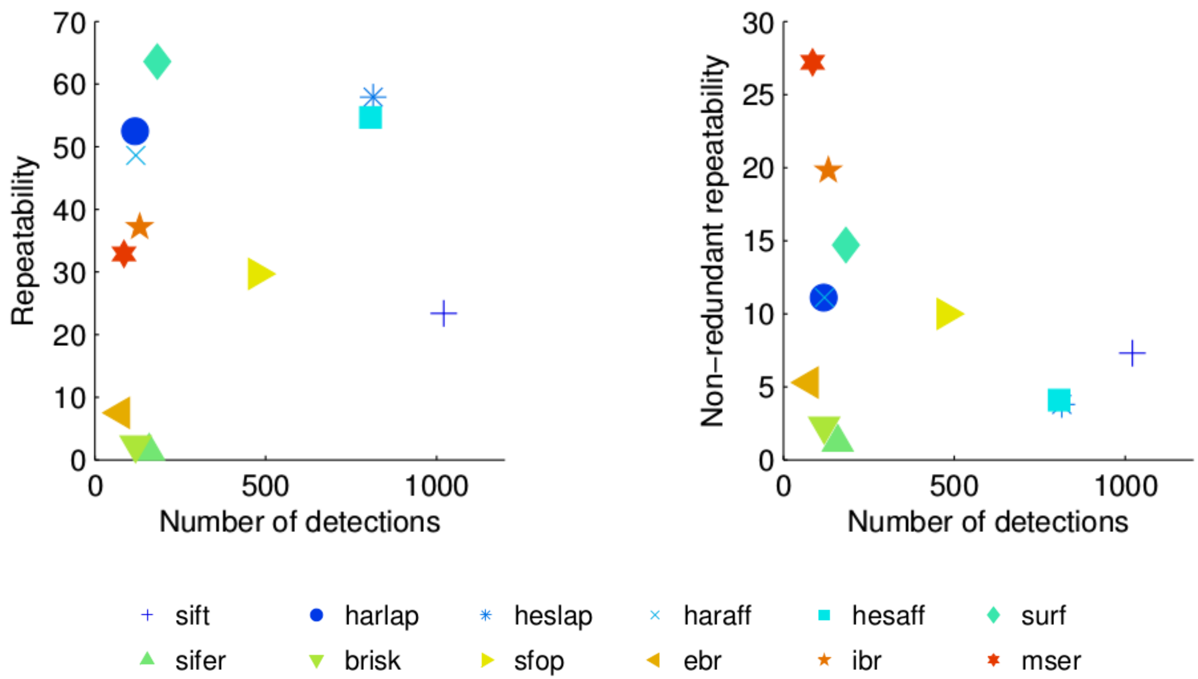
Replace it with a better measure of the <u>expected benefit:</u>

$$\boxed{\text{Non-redundant repeatability rate} = \frac{\boxed{\int_\Omega \max_{\text{repeated keys}} f_k(x,y)\,dx\,dy}}{\text{Total number of detections}}}$$

Revisiting a popular benchmark [Mikolajczyk 2005]



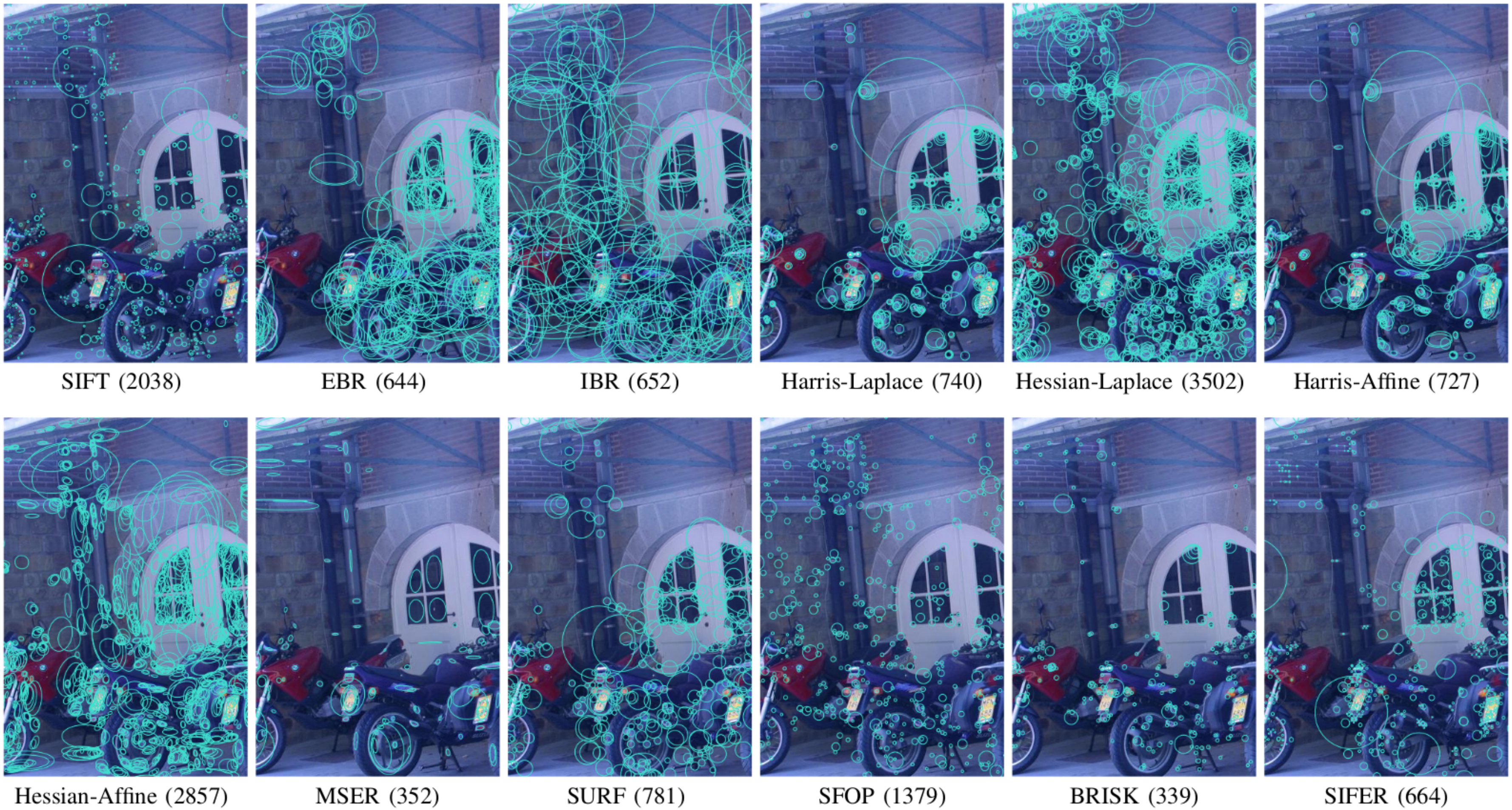
Revisiting a popular benchmark [Mikolajczyk 2005]



Very different conclusions:

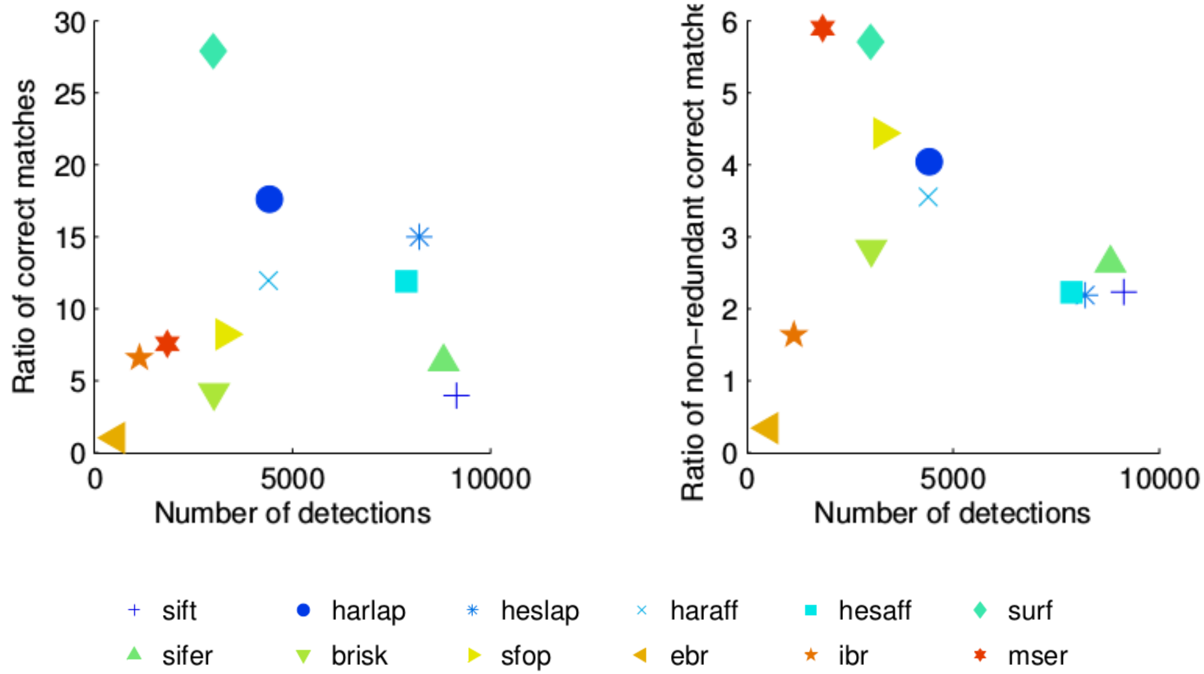
- *Most methods are highly redundant*
- *SIFT performs best for large number of keypoints*

Revisiting a popular benchmark [Mikolajczyk 2005]



Performance in a matching scenario

..using SIFT descriptor and SIFT matching algorithm



- *State-of-the-art turned upside-down*

Overview

1) *Classic repeatability criterion:*

   A bias towards redundant algorithms

2) *An amended criterion:*

   Take spatial redundancy into account

3) *A revisited benchmark:*

   Hierarchy turned upside-down