Efficient Interest Point Detectors & Features

Instructor - Simon Lucey

16-623 - Designing Computer Vision Apps

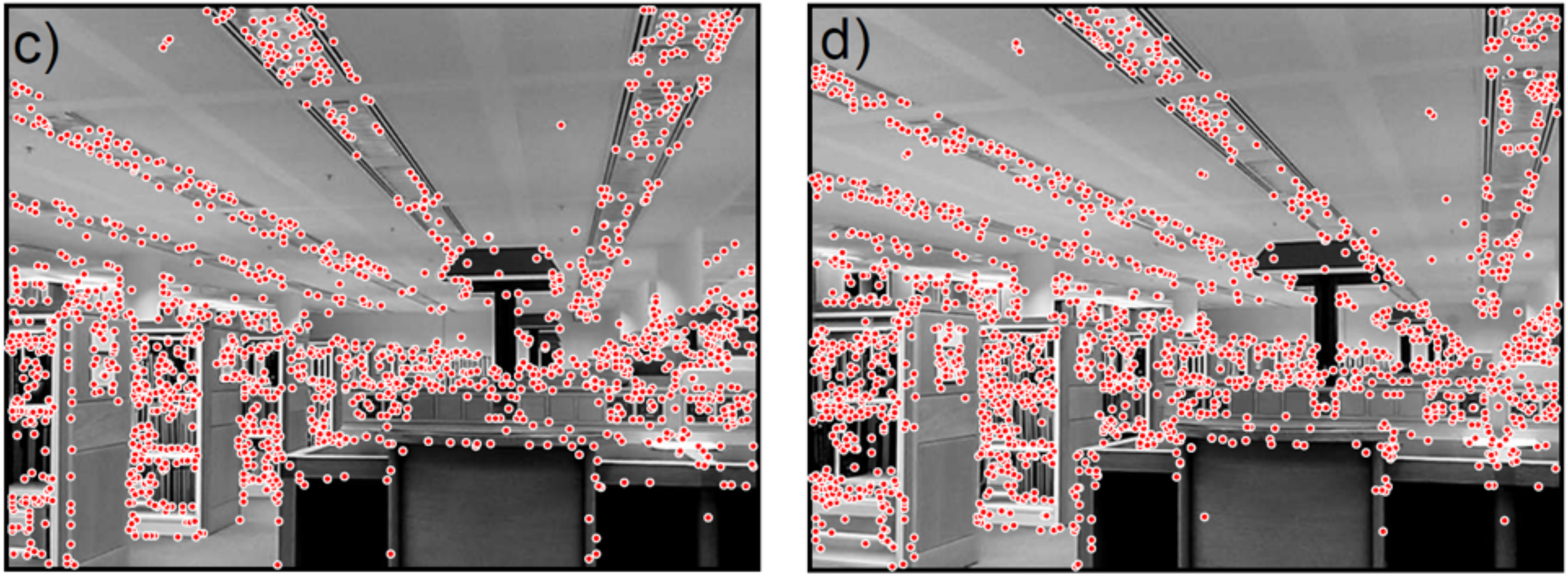
Carnegie Mellon
THE ROBOTICS INSTITUTE

Today

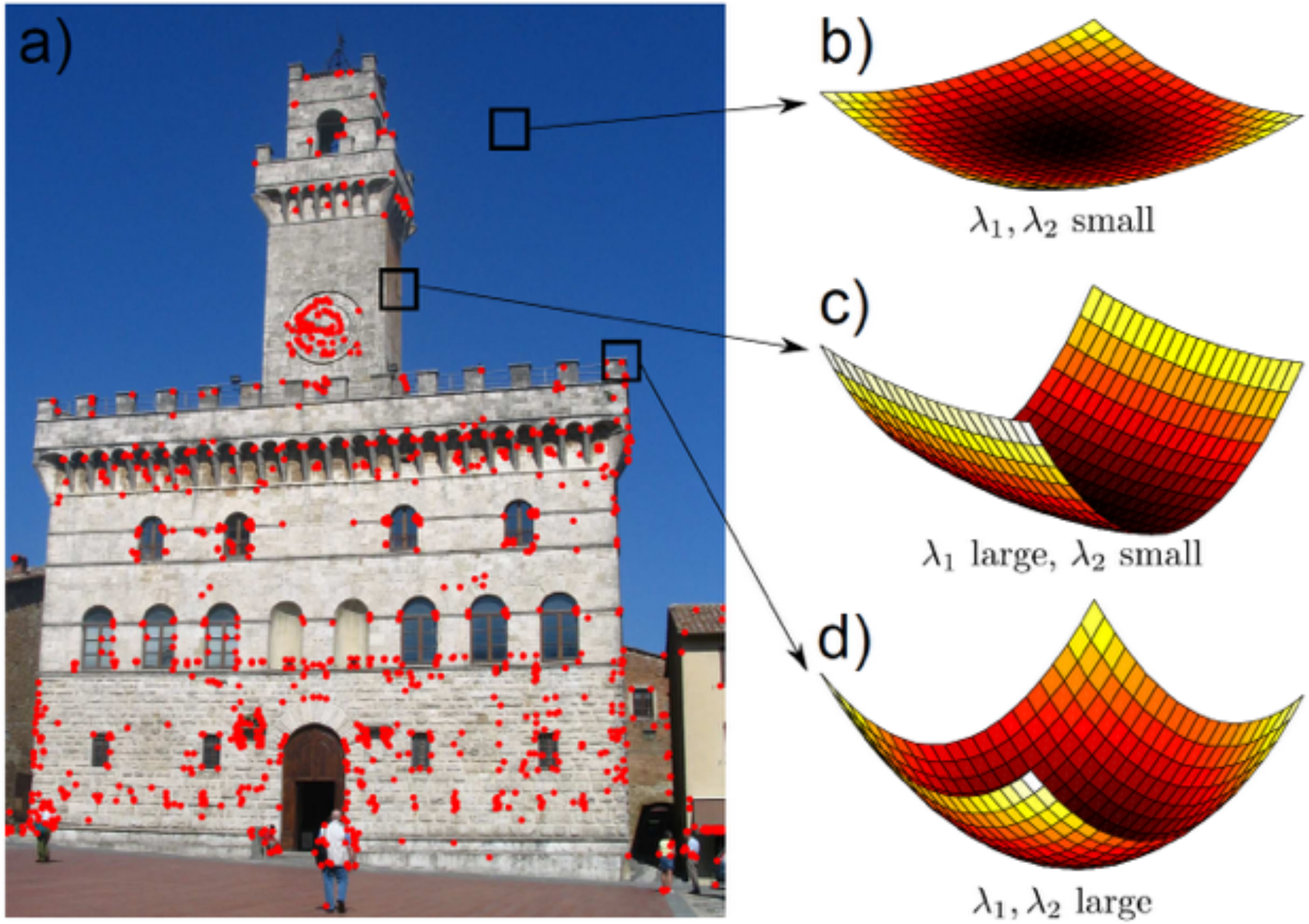
- Review.
- Efficient Interest Point Detectors.
- Efficient Descriptors.

Review

- In classical Structure from Motion (SfM) computer vision pipeline there are four steps,
    1. Locate interest points.



Review - Harris Corner Detector



Make decision based on image structure tensor

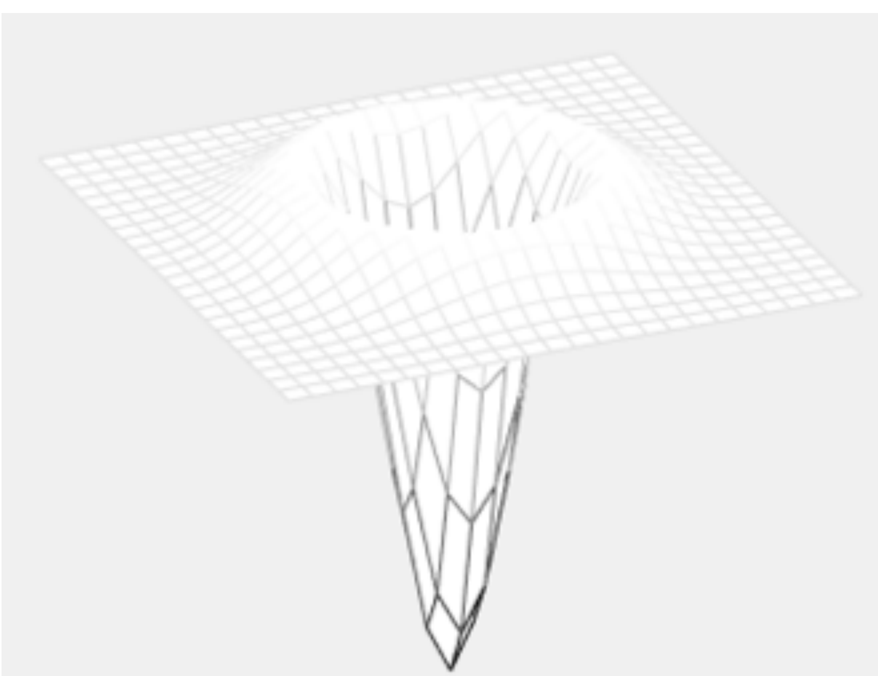
$$\mathbf{H} = \sum_{i \in \mathcal{N}} \frac{\partial I(\mathbf{x}_i)}{\partial \mathbf{x}} \frac{\partial I(\mathbf{x}_i)}{\partial \mathbf{x}}^T$$

Scalar Measures of “Cornerness"

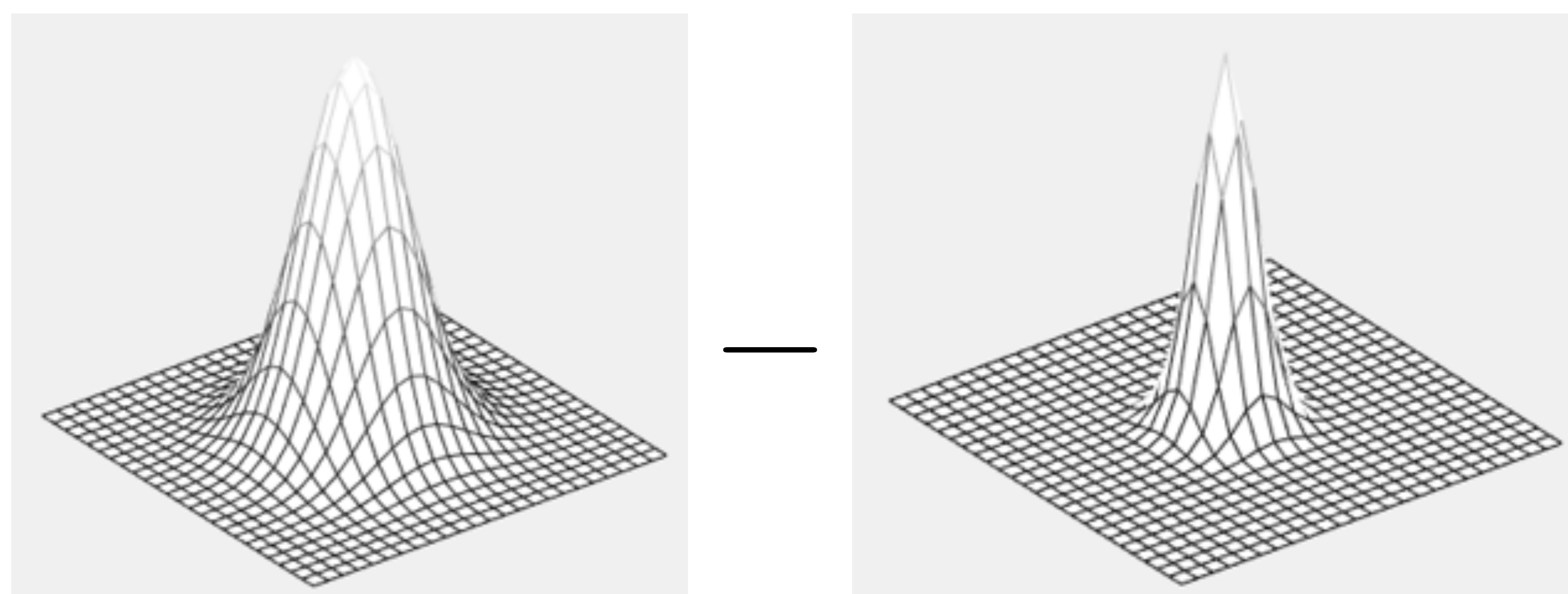
- A popular measure for measuring a corner $\lambda_1 + \lambda_2$,

$$\mathrm{tr}[\mathbf{H}_{(x,y)}] = ||\nabla_x * I(x,y)||_2^2 + ||\nabla_y * I(x,y)||_2^2$$

$$\approx ||L * I(x,y)||_2^2$$



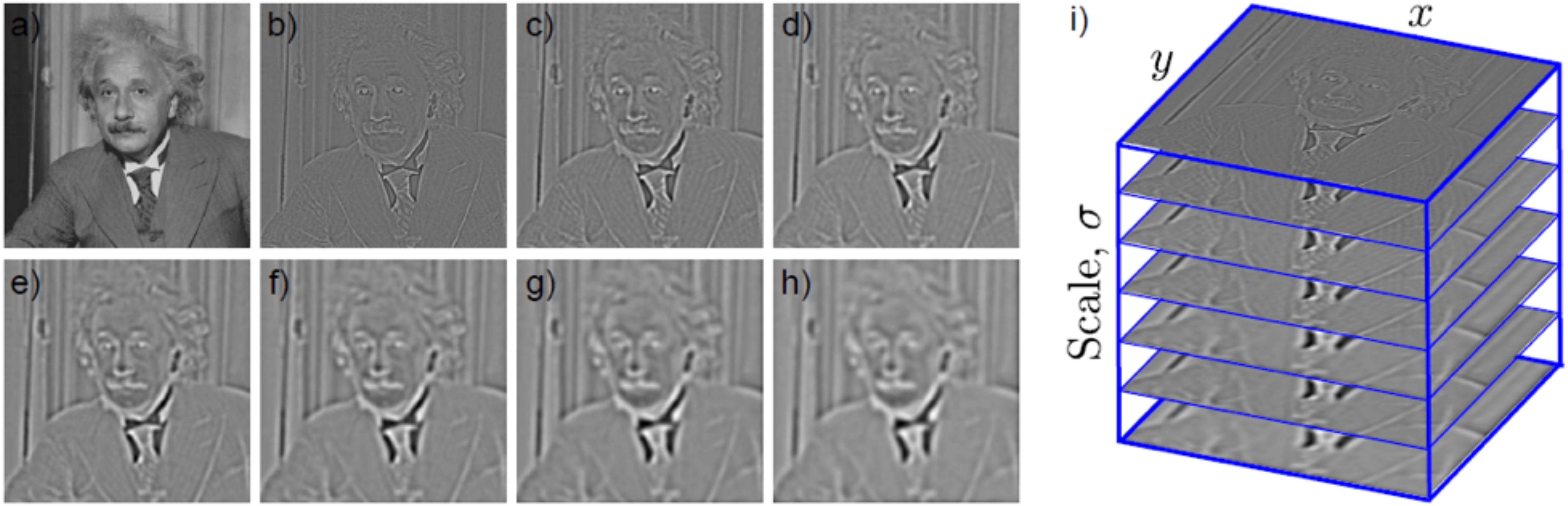
$\approx$



“Laplacian (L)”

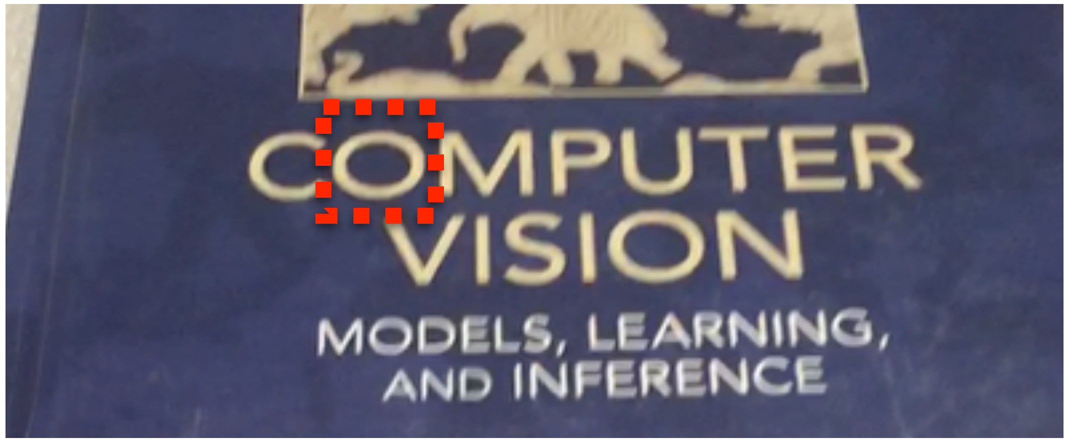
“Difference of Gaussians (DOG)”

Example - DoGs in SIFT



Review

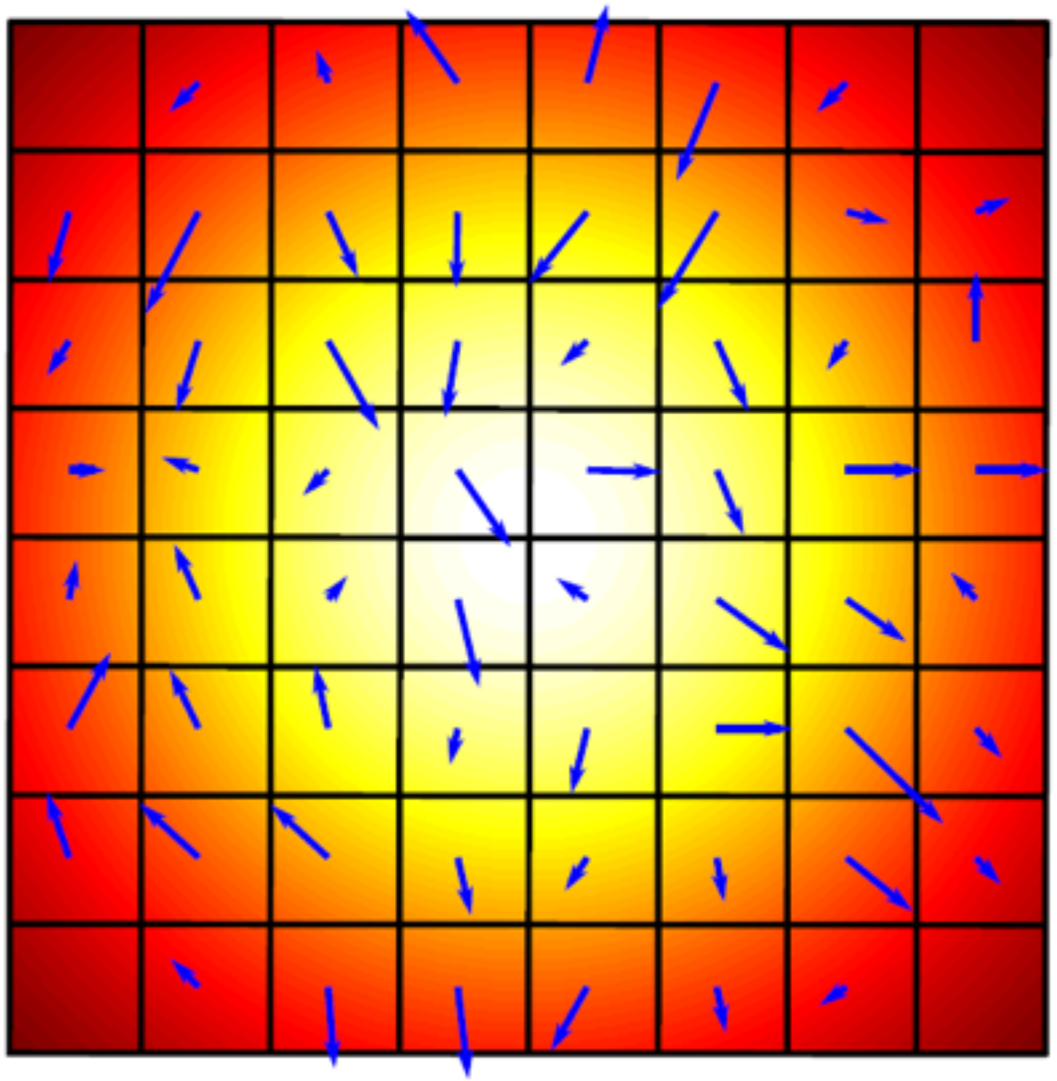
- In classical Structure from Motion (SfM) computer vision pipeline there are four steps,
    1. Locate interest points.
    2. Generate descriptors.

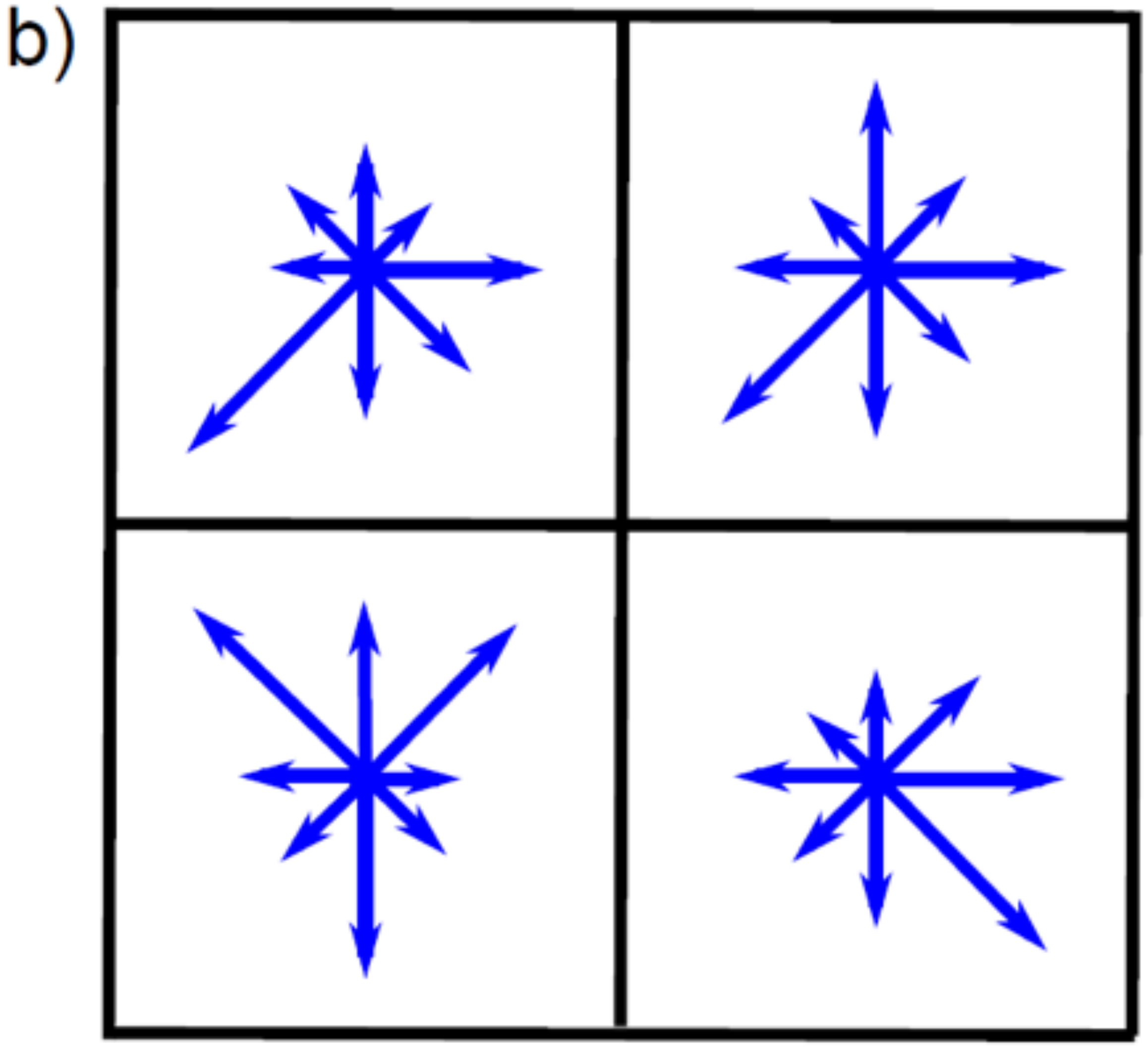


I_1

I_2

Review - SIFT Descriptor

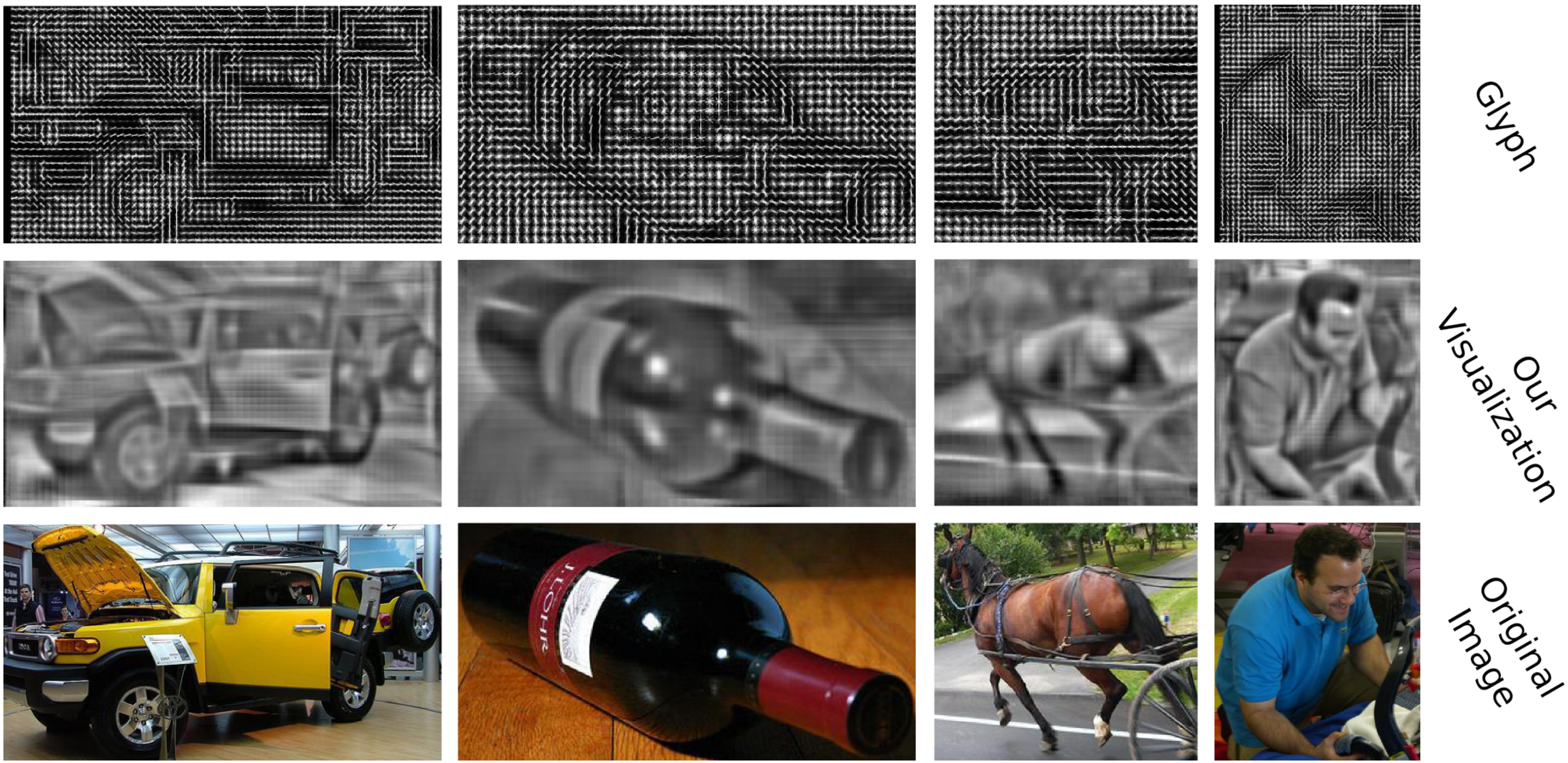




1. Compute image gradients

2. Pool into local histograms
3. Concatenate histograms
4. Normalize histograms

HOGGles

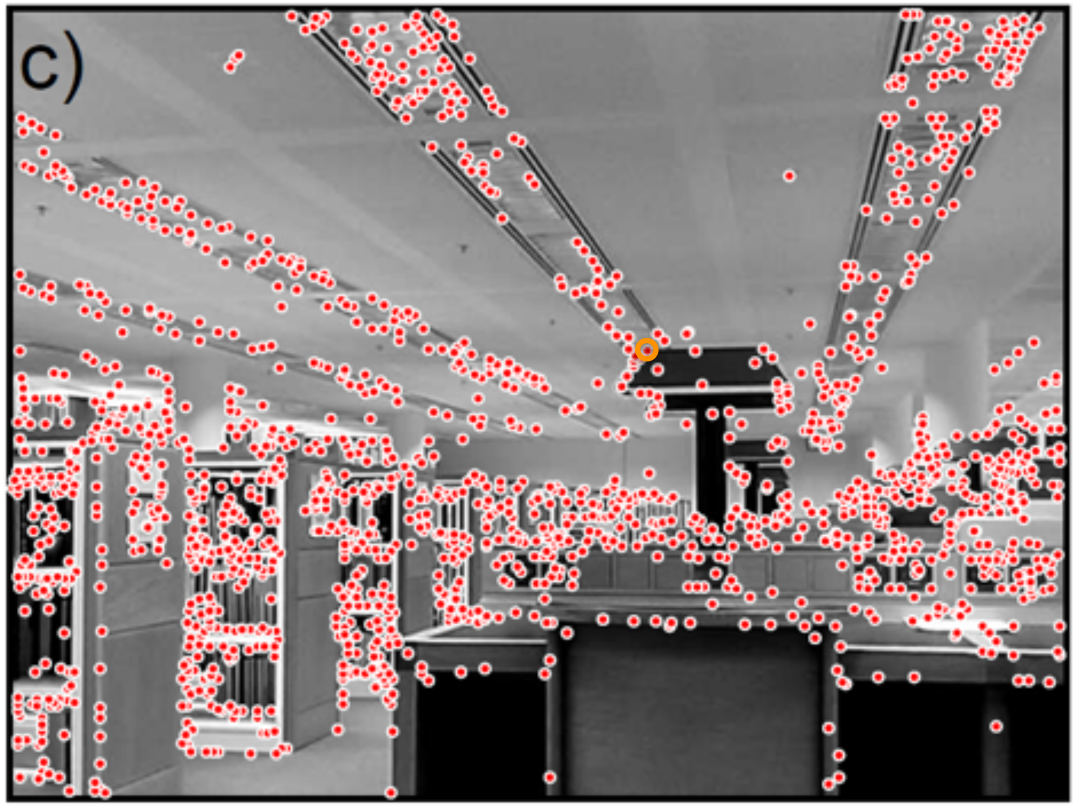


Taken from: Vondrick, C., Khosla, A., Malisiewicz, T., & Torralba, A. (2013). Hoggles: Visualizing object detection features. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1-8).

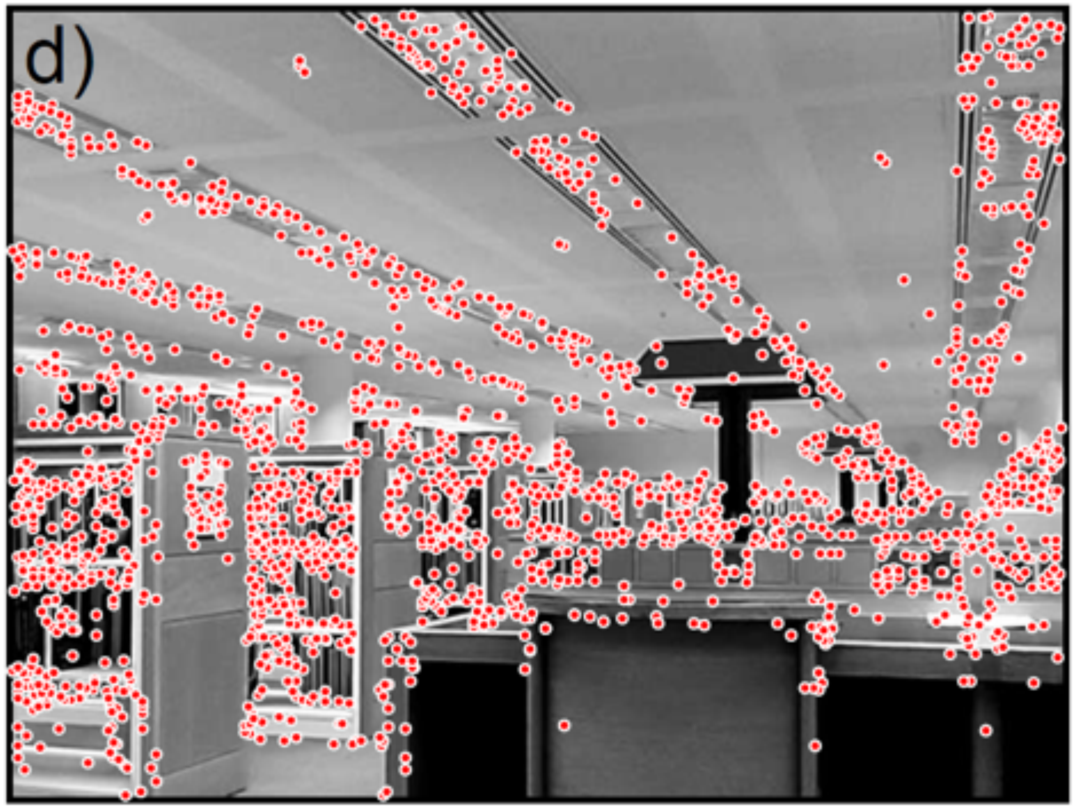
Review

- In classical Structure from Motion (SfM) computer vision pipeline there are four steps,
  1. Locate interest points.
  2. Generate descriptors.
  3. Matching descriptors.

Matching Descriptors

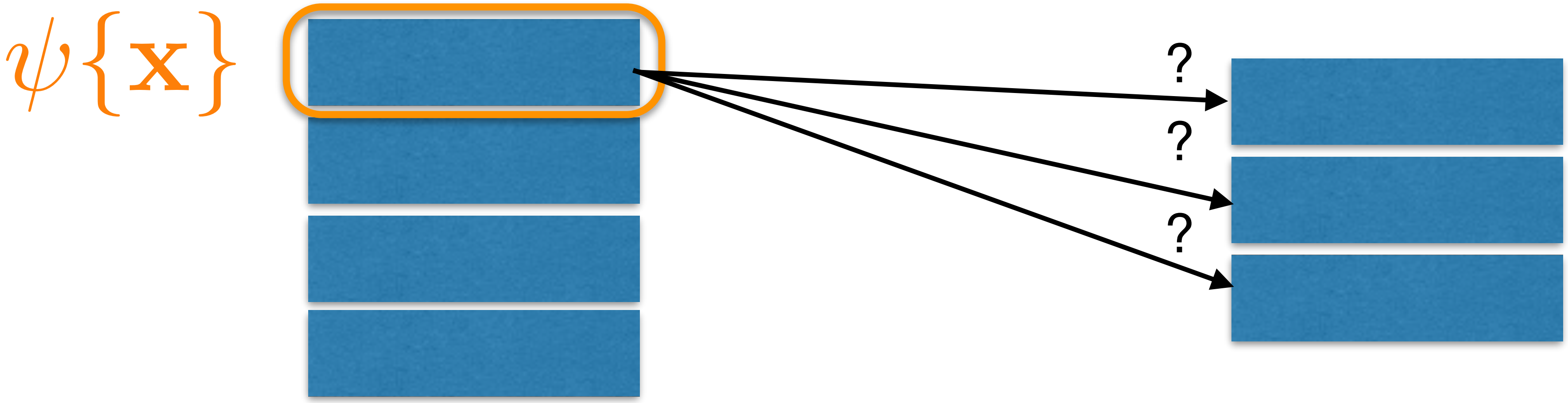


View 1



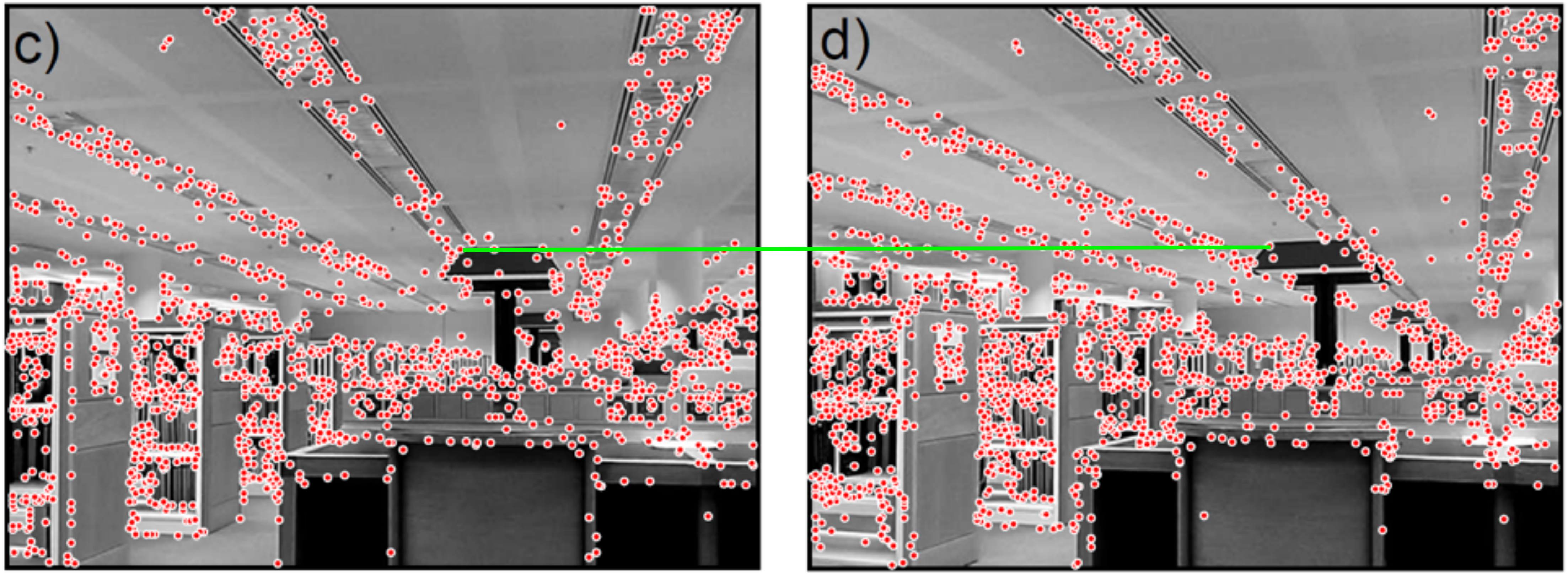
View 2

“Descriptor at position $\mathbf{x}$”



See BFMatcher class in OpenCV!!!

Matching Descriptors



View 1 View 2

$$\zeta(i) = \arg\min_j ||\psi\{\mathbf{x}_i^{(1)}\} - \psi\{\mathbf{x}_j^{(2)}\}||_2^2$$

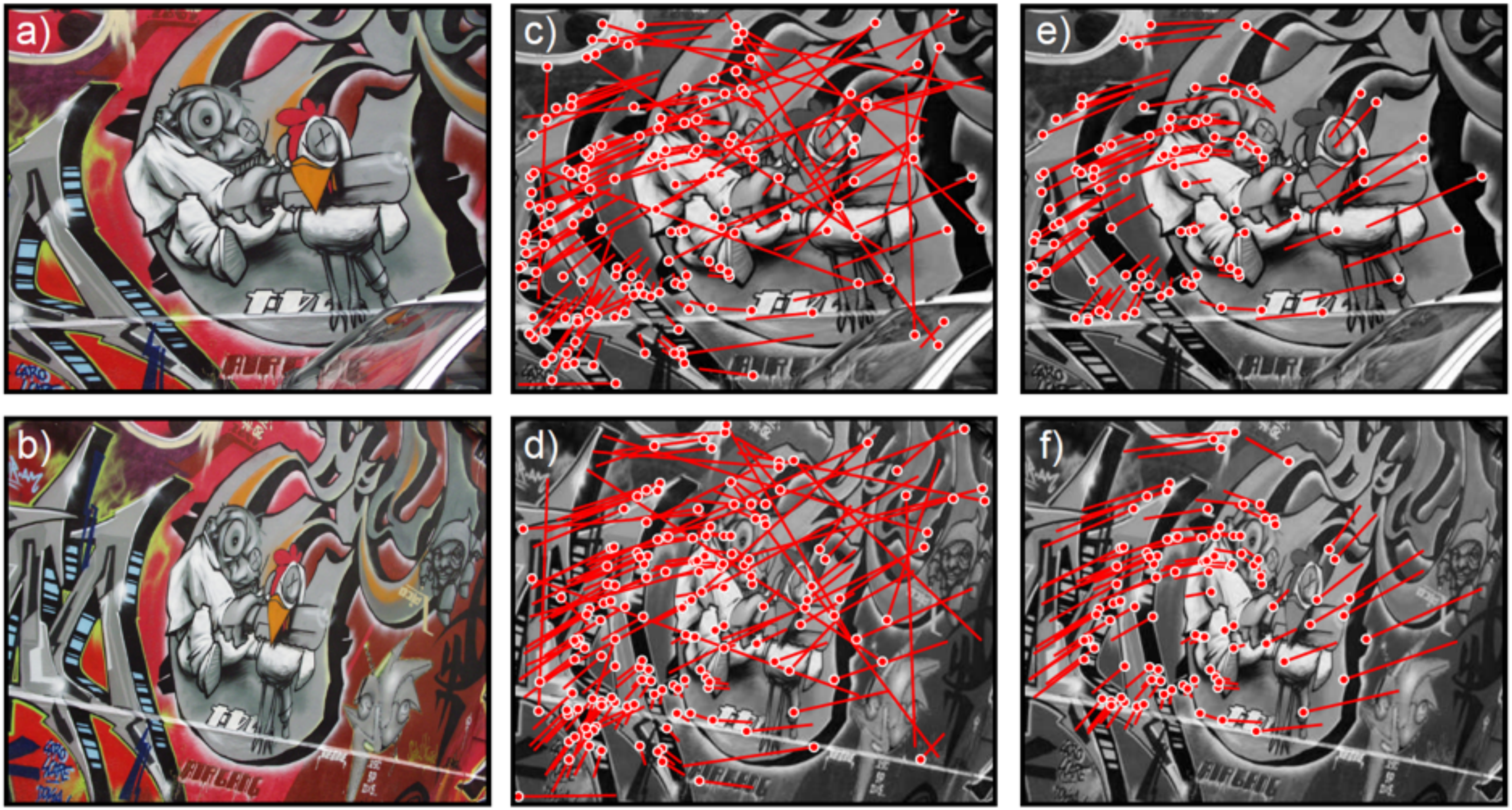
Variants other than nearest neighbor are possible!!!

Review

- In classical Structure from Motion (SfM) computer vision pipeline there are four steps,
    1. Locate interest points.
    2. Generate descriptors.
    3. Matching descriptors.
    4. Robust fit.

$$\arg\min_{\mathbf{\Phi}} \eta\{\mathbf{x}_i^{(1)} - \mathbf{hom}[\mathbf{x}_{\zeta(i)}^{(2)}; \mathbf{\Phi}]\}$$

Review - RANSAC



Original images | Initial matches | Inliers from RANSAC

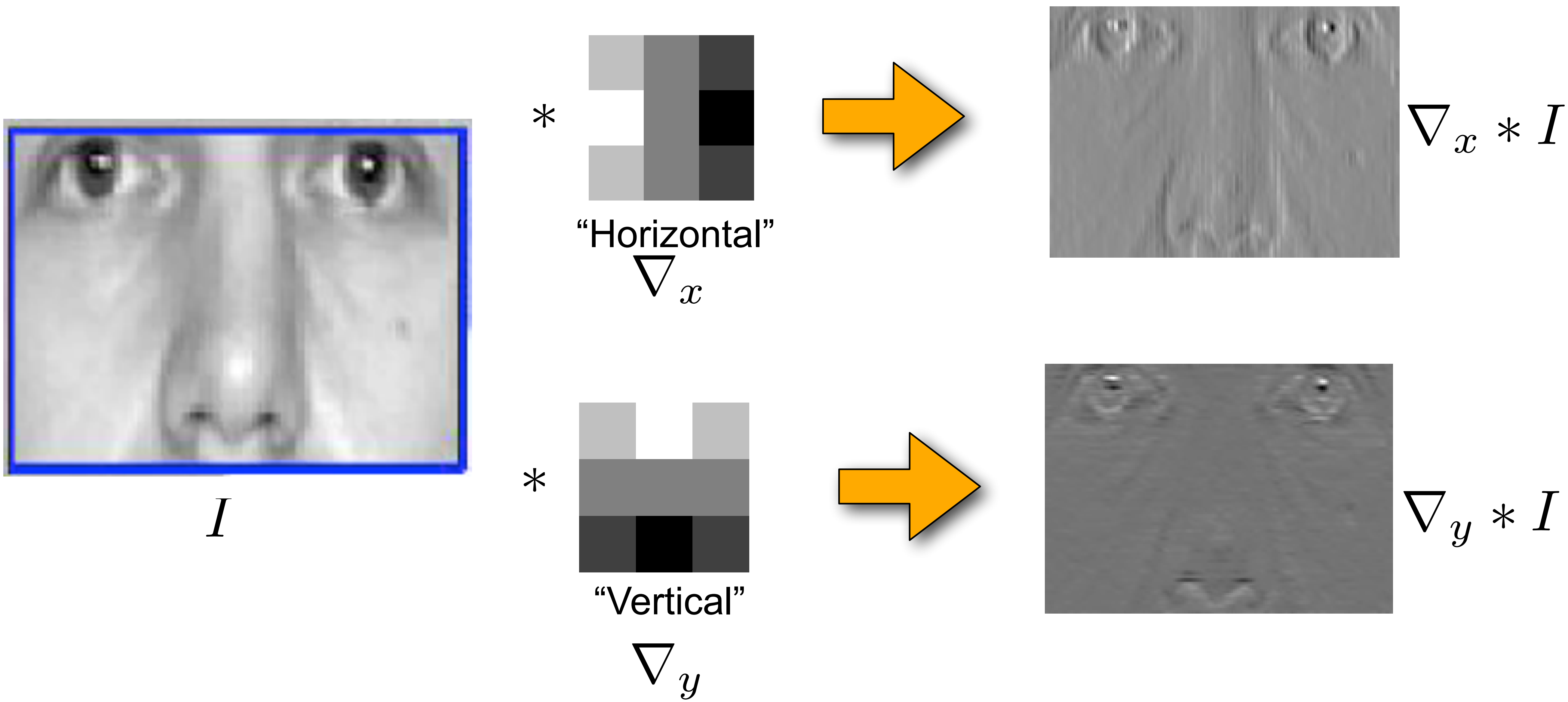
?

Today

- Review.
- Efficient Interest Point Detectors.
- Efficient Descriptors.

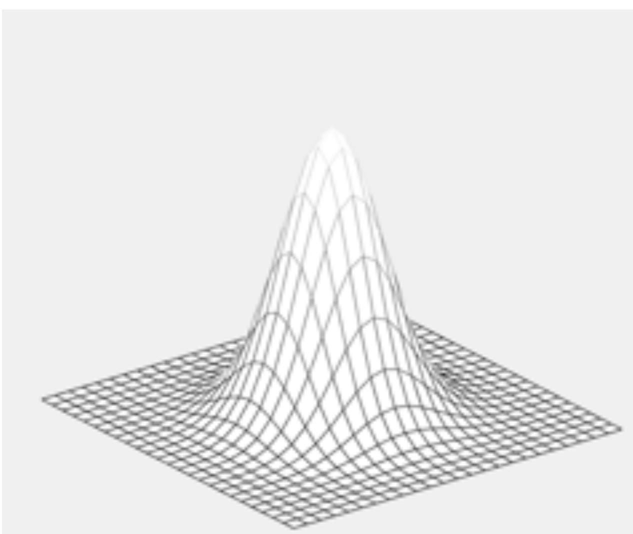
Efficient Interest Point Detection

- Most classical interest point detectors require the employment of oriented edges.



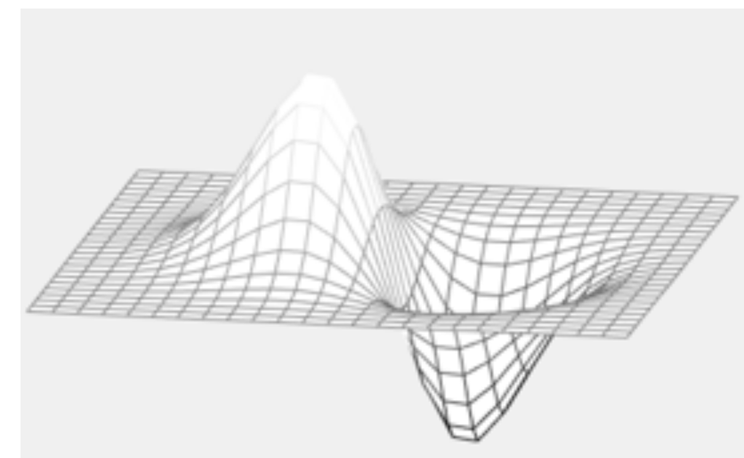
Problem - Gaussian Filtering is Slow

- Naively, filtering with Gaussians is relatively slow on most modern architectures.



$$* \begin{bmatrix} 1, 0, -1 \\ 2, 0, -2 \\ 1, 0, -1 \end{bmatrix}$$

(Sobel)

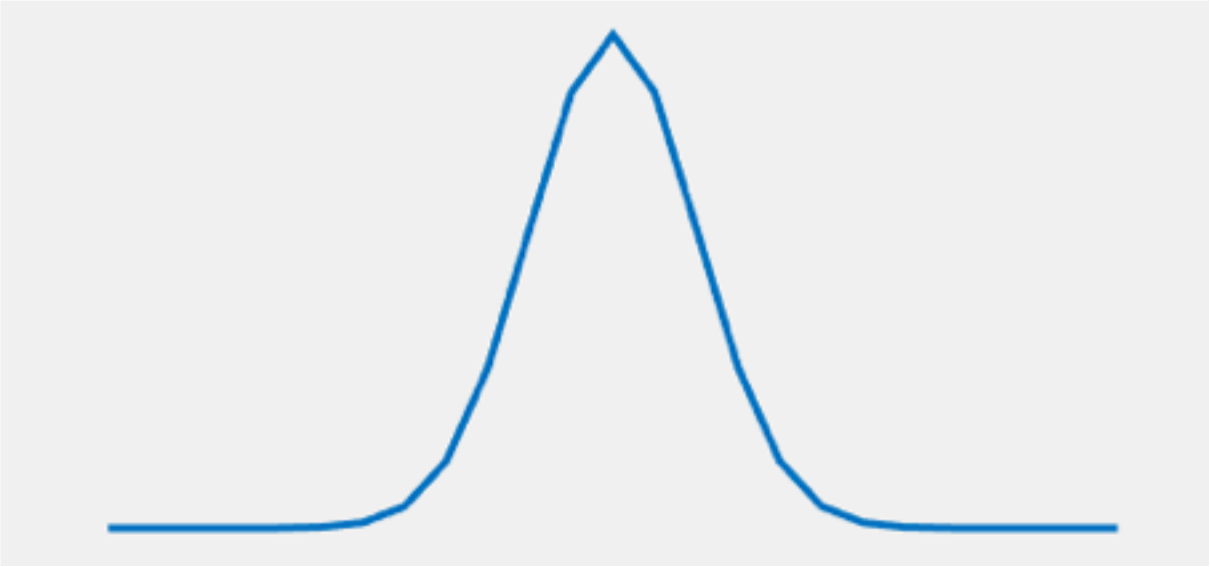


- Does not lend itself well to parallelization as the variance of the Gaussian filter increases (even with FFT).
- Computational cost increases dramatically as a function of the size of the filter.

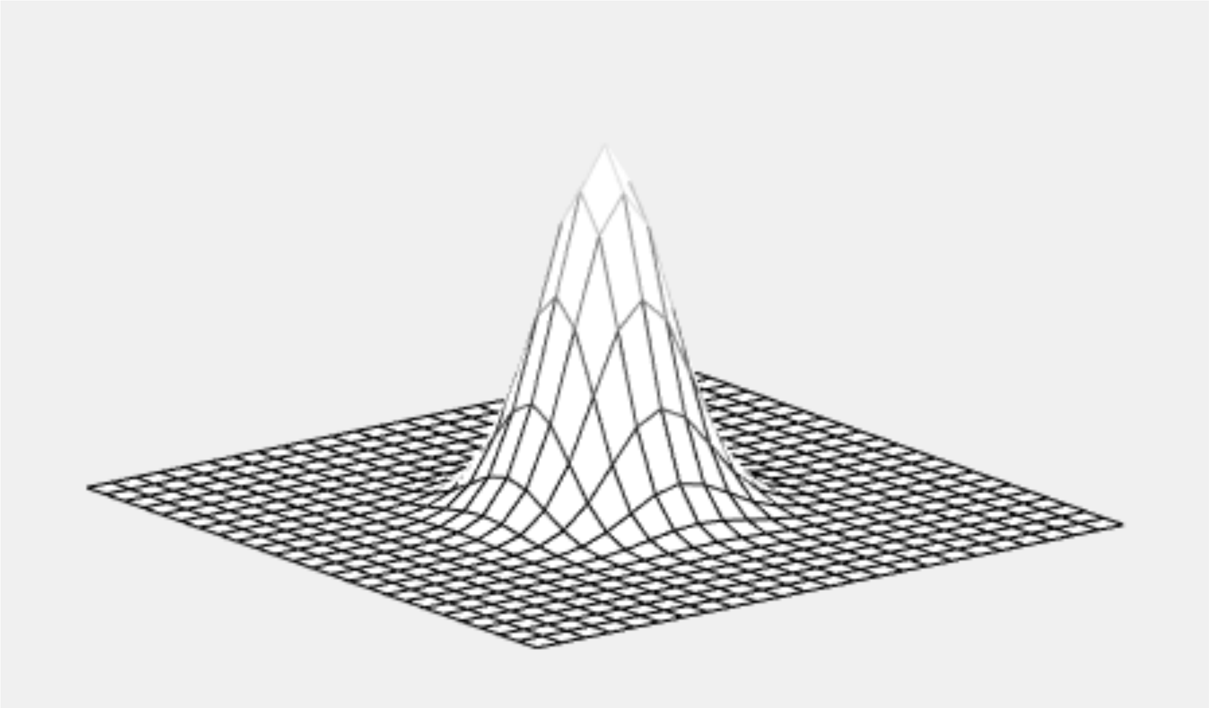
Gaussian Filter is Separable

In MATLAB,

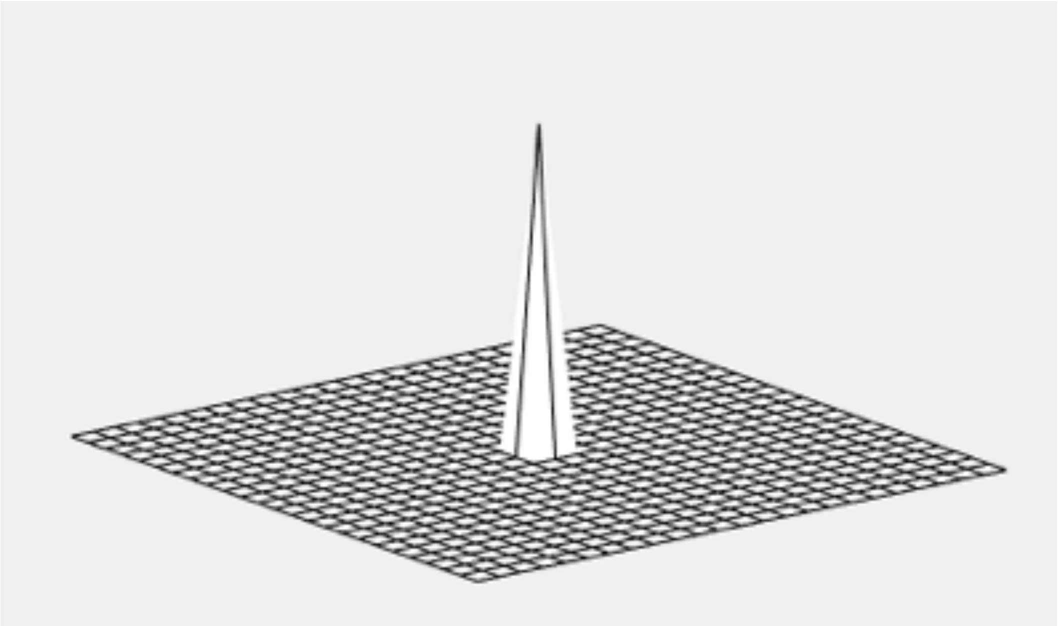
```
>> h1D = fspecial('gaussian',[25,1],3);
```



```
>> h2D = kron(h1D,h1D');
```

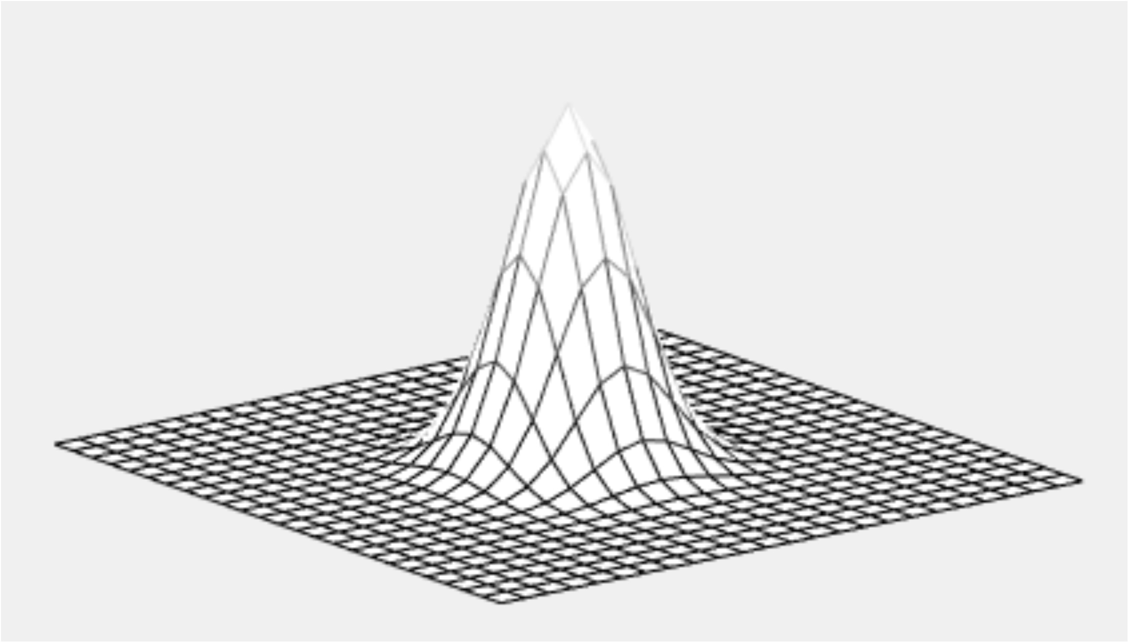


Gaussian Filter is Separable



In MATLAB,

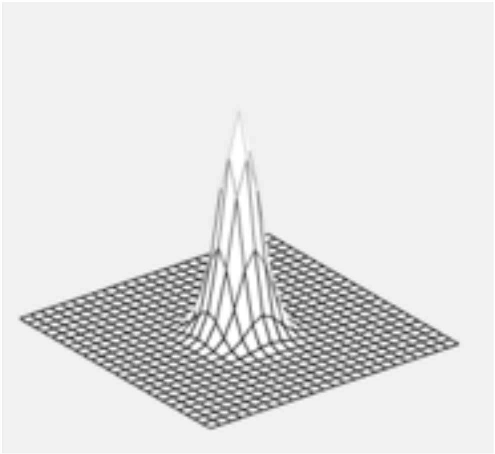
```
>> mesh(i)
```



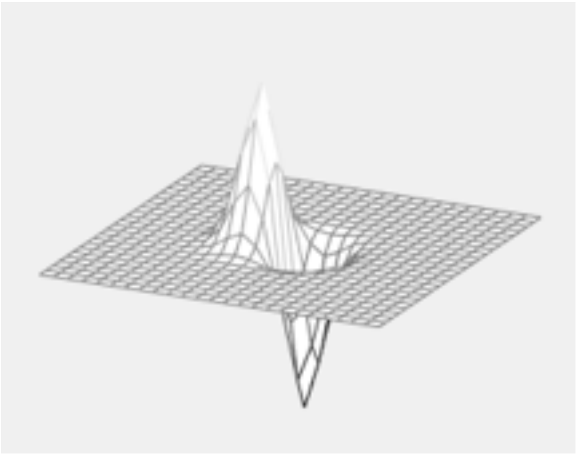
```
>> h2D = imfilter(i, h1D);
>> h2D = imfilter(h2d, h1D');
>> mesh(h2D)
```

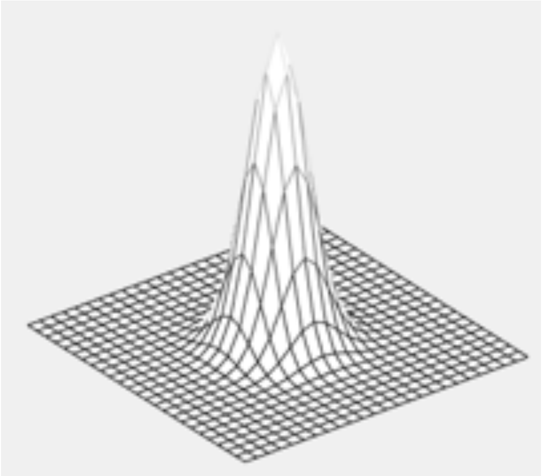
More Problems - Scale

- However, even slower when you have to process things across multiple scales.

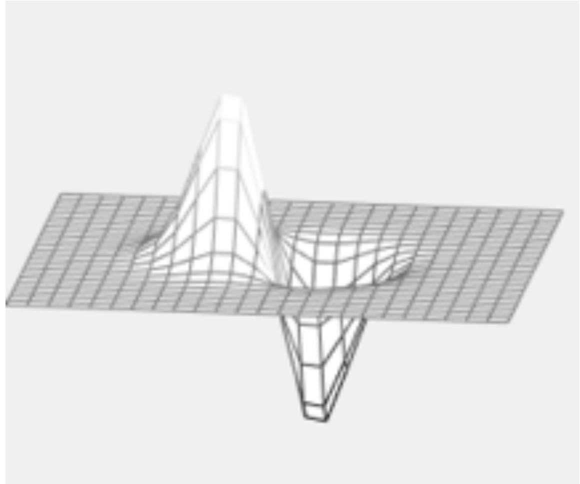


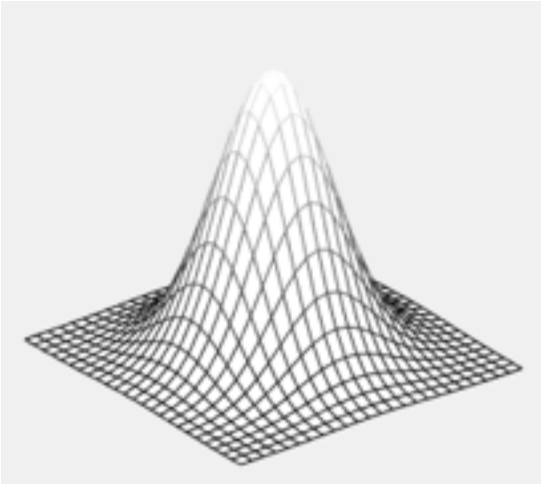
$$* \begin{bmatrix} 1, 0, -1 \\ 2, 0, -2 \\ 1, 0, -1 \end{bmatrix}$$



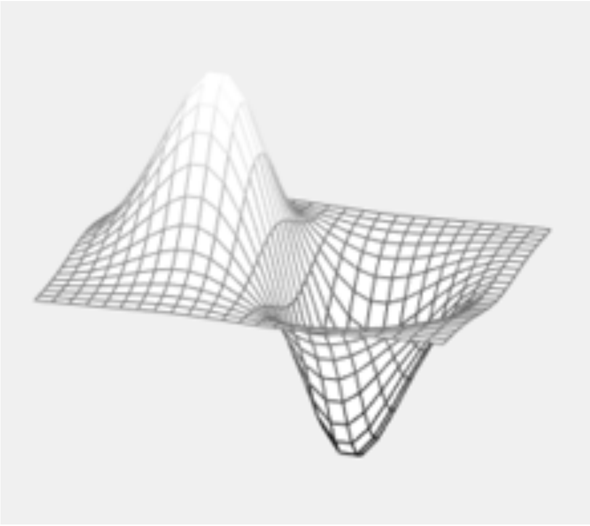


$$* \begin{bmatrix} 1, 0, -1 \\ 2, 0, -2 \\ 1, 0, -1 \end{bmatrix}$$



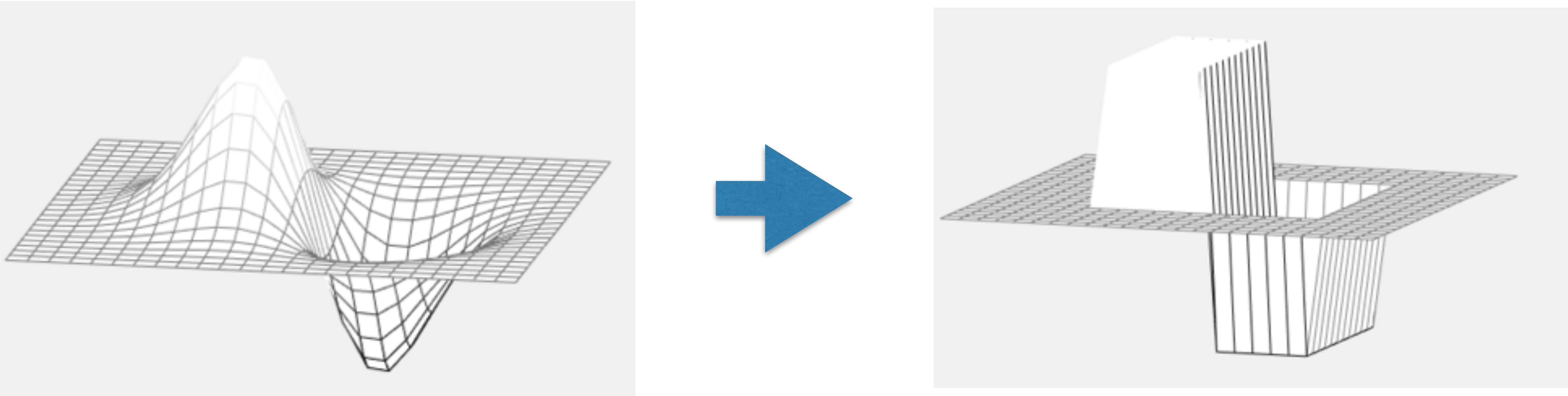


$$* \begin{bmatrix} 1, 0, -1 \\ 2, 0, -2 \\ 1, 0, -1 \end{bmatrix}$$



Solution - Box Filters

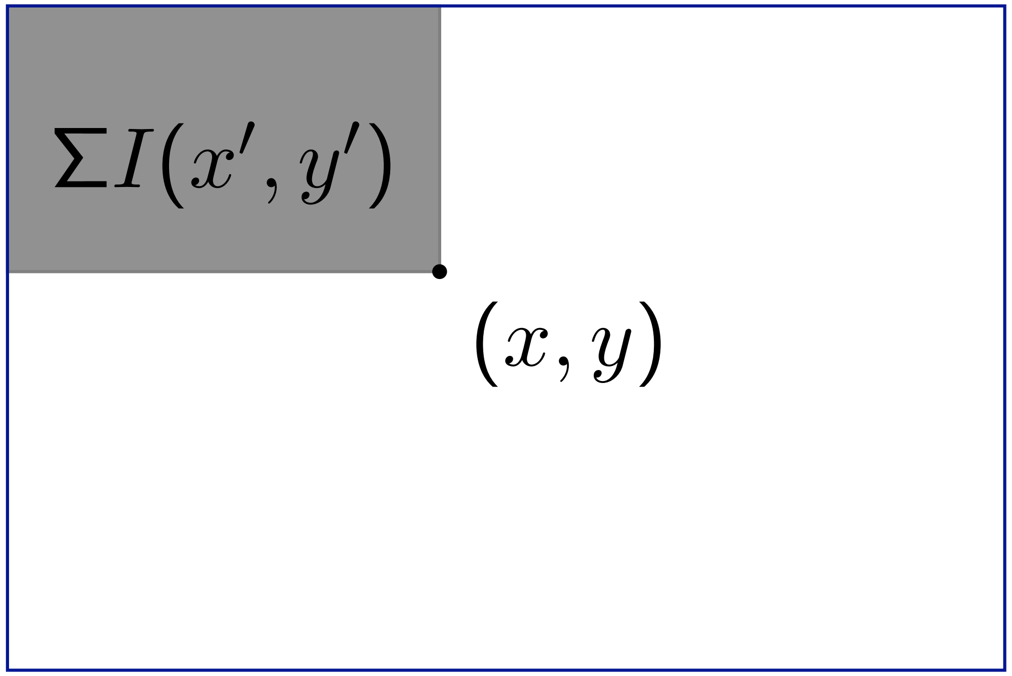
- One strategy has been to approximate oriented filters with box filters.
- Most notably the SURF (Speed Up Robust Feature) descriptor of Bay et al. ECCV 2006.



Integral Image Trick

- We need to compute the box filter values many, many times and we must do it very fast!

$$II(x,y) = \sum_{x' \le x,\, y' \le y} I(x', y')$$

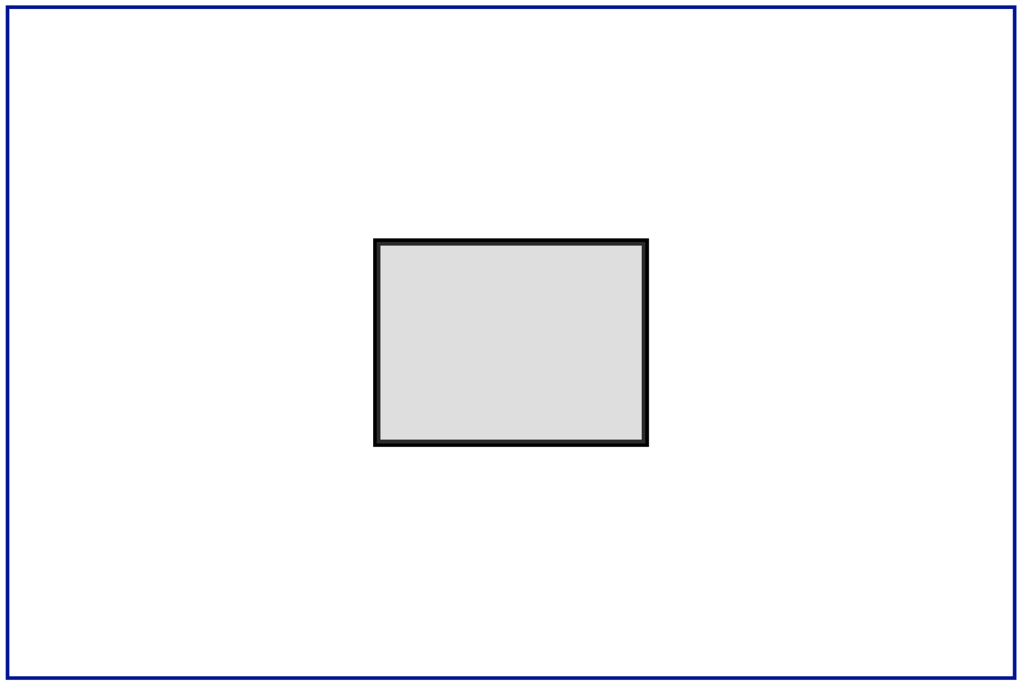


(Black)

Computing Integral Images

- Computing sum of pixels in a rectangular area:

$$f(A) =$$

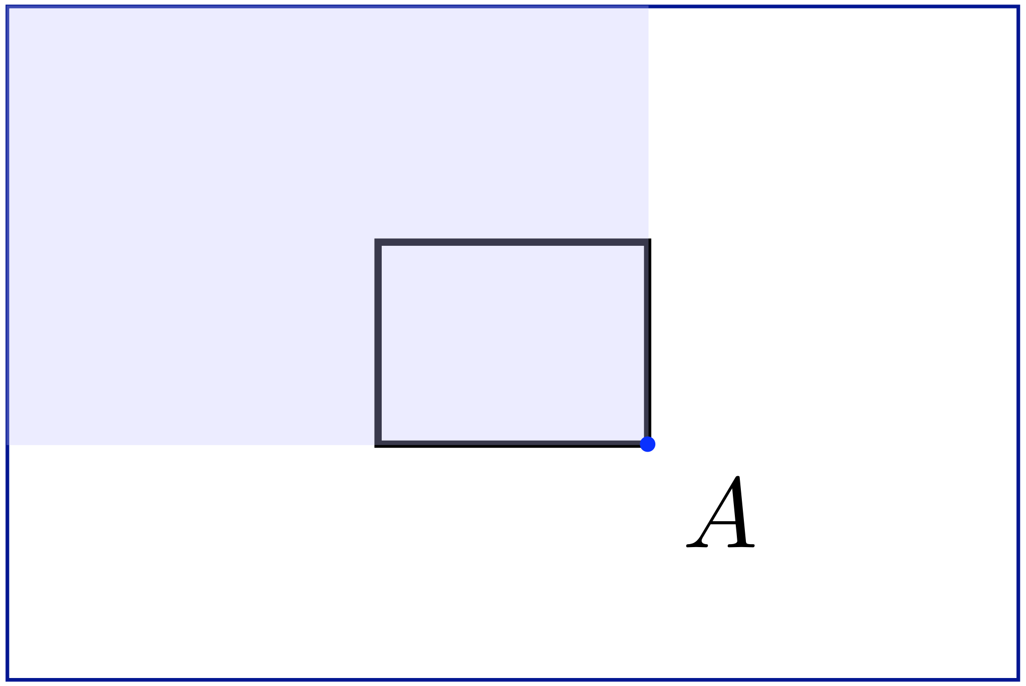


(Black)

Computing Integral Images

- Computing sum of pixels in a rectangular area:

$$f(A) = II(A)$$

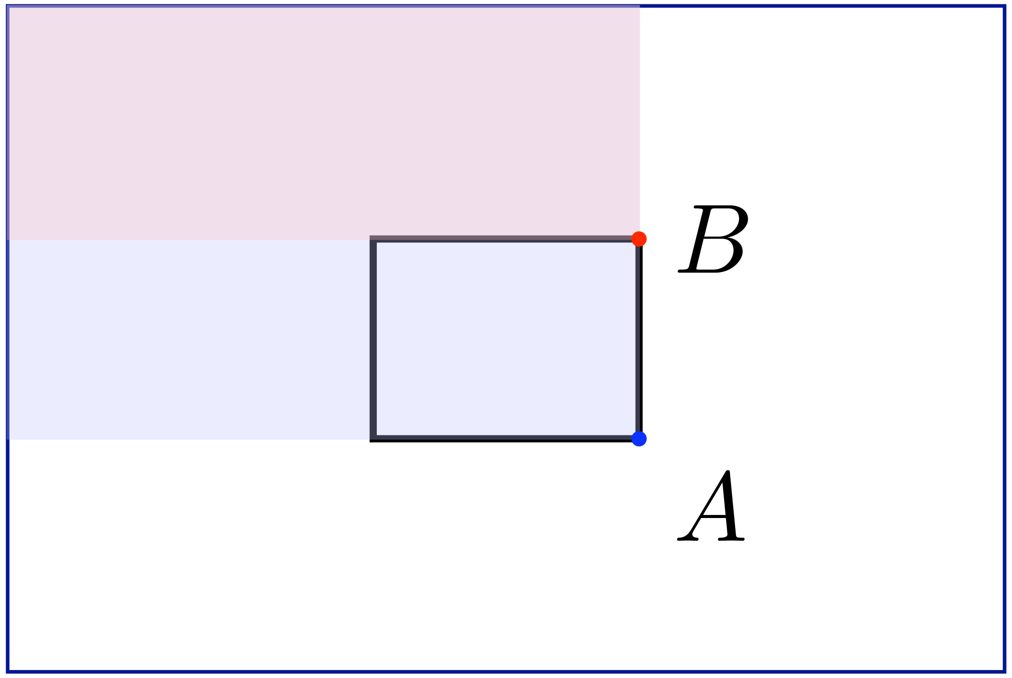


(Black)

Computing Integral Images

- Computing sum of pixels in a rectangular area:

$$f(A) = II(A) - II(B)$$

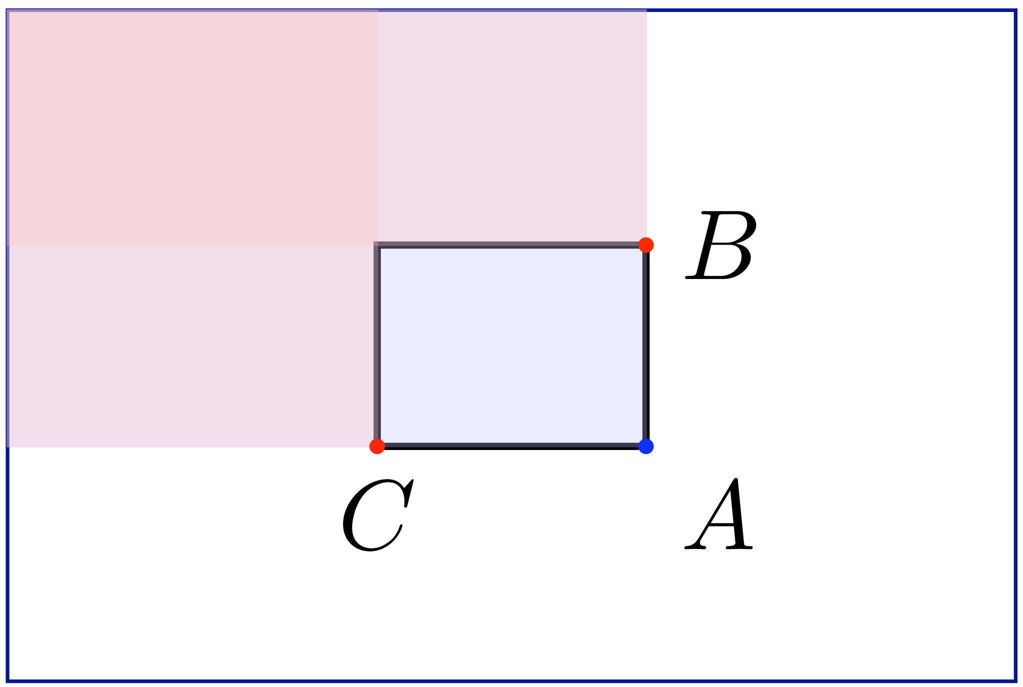


(Black)

Computing Integral Images

- Computing sum of pixels in a rectangular area:

$$
\begin{aligned}
f(A) &= II(A) - II(B) \\
&\quad - II(C)
\end{aligned}
$$

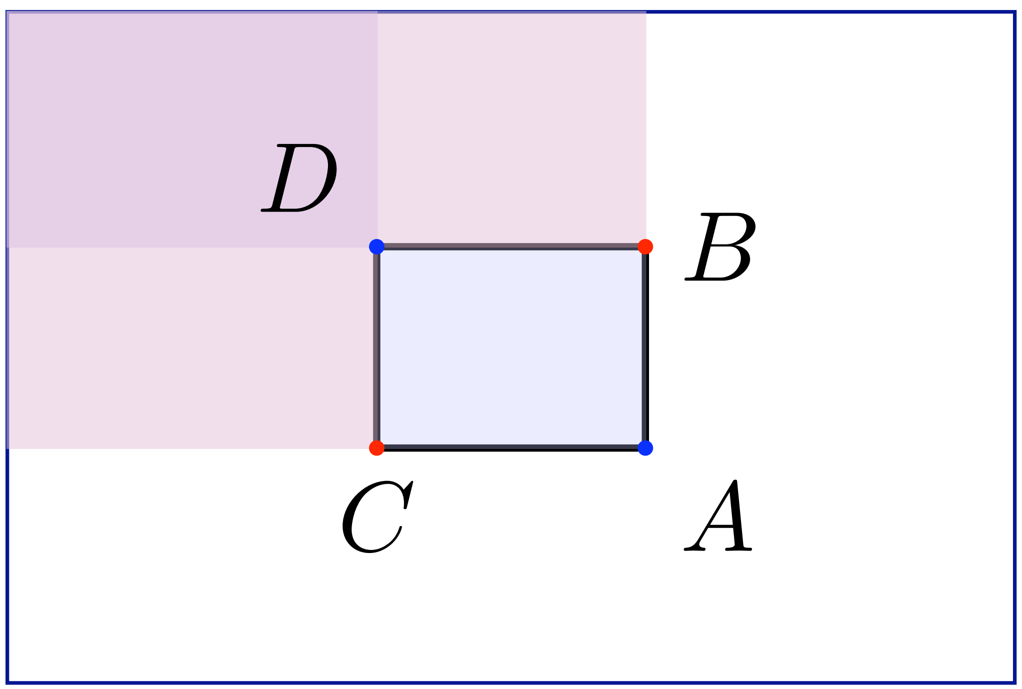


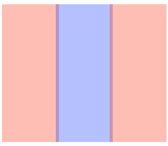
(Black)

Computing Integral Images

- Computing sum of pixels in a rectangular area:

$$\begin{aligned} f(A) &= II(A) - II(B) \\ &\quad - II(C) + II(D) \end{aligned}$$

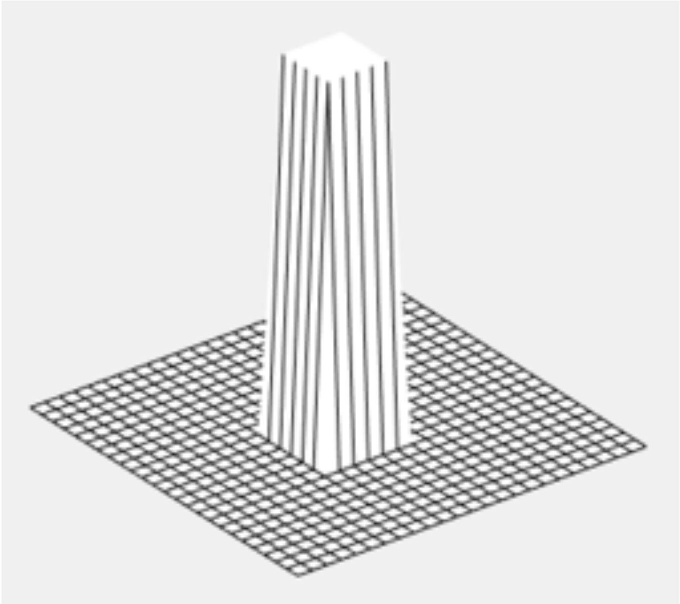


- A 3 box filter array  takes only 8 lookups.

(Black)

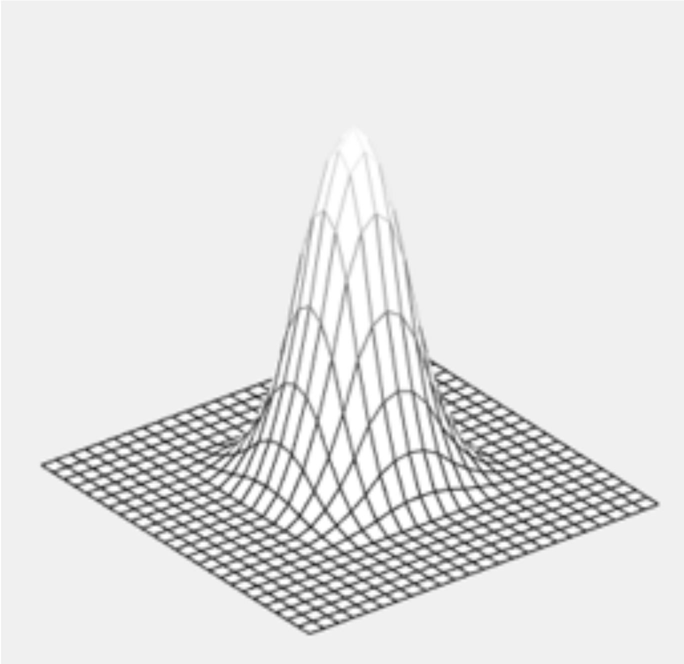
Fast Gaussian Filtering

- Iterative box filters can also be applied to obtain extremely efficient Gaussian filtering,



In MATLAB,

```
>> mesh(b)
```



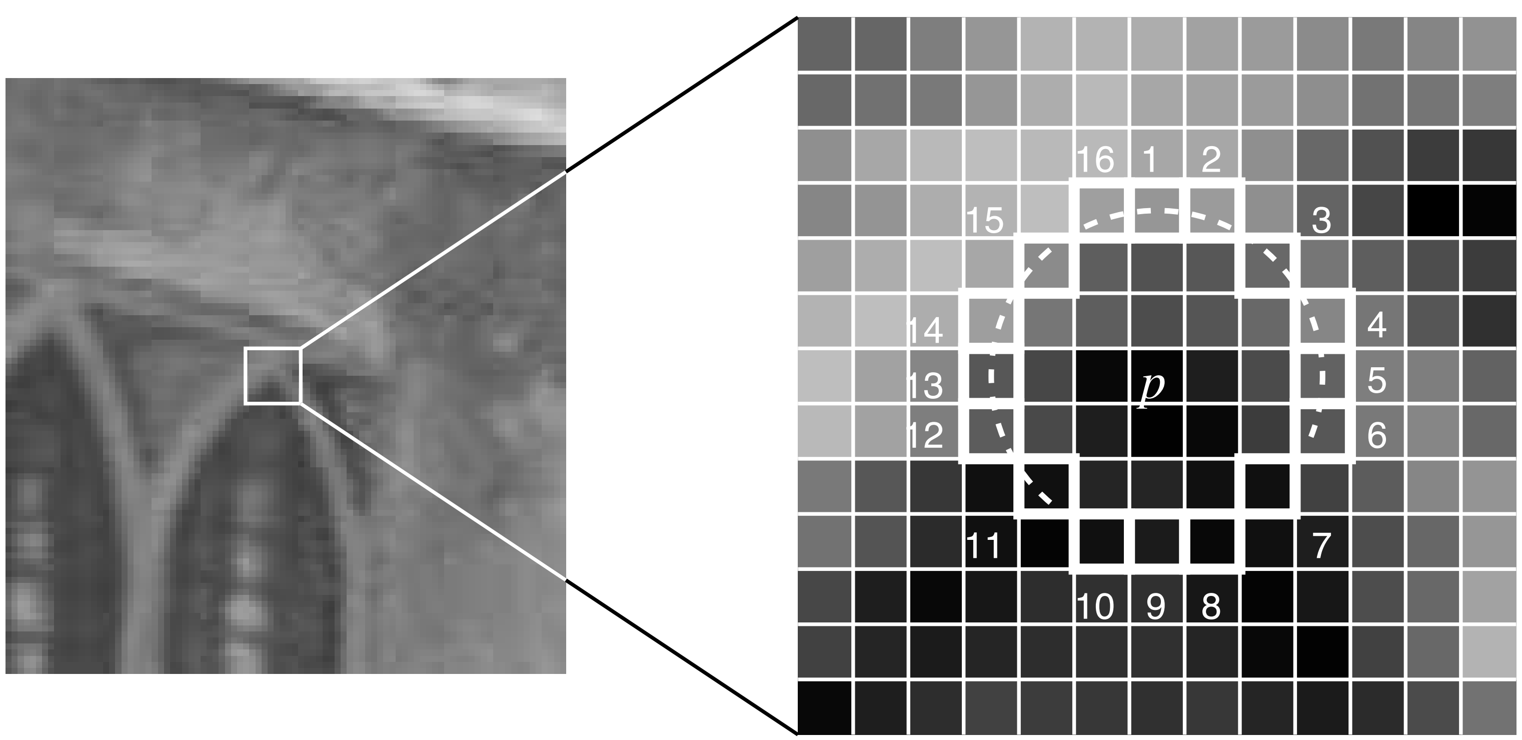
```
>> mesh(imfilter(imfilter(b,b),b))
```

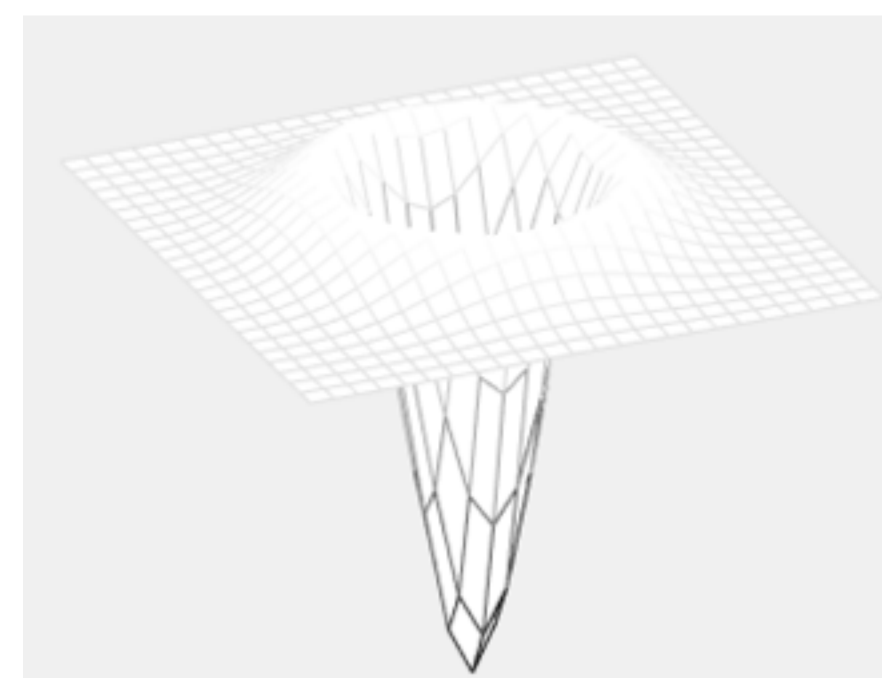
SURF - Efficient Computation

- Positives:-
  - Filters can be efficiently applied irrespective of size.
  - Integral images well suited in particular to SIMD.
  - Can take advantage of fixed integer arithmetic.
- Negatives:-
  - Due to recursive nature cannot be easily parallelized.
  - All pixels in local region need to be touched.
  - Outputs floating/integer point metric of interest.

FAST

- Features from Accelerated Segment Test (FAST) - basis for most modern day computationally efficient interest point detectors.
- Proposed by Rosten et al. PAMI 2010.
- Operates by considering a circle of sixteen pixels around the corner candidate.





Why is it FAST?

- FAST relies heavily upon simple binary test,

$$I(\mathbf{x}_p) - t > I(\mathbf{x}_n)$$

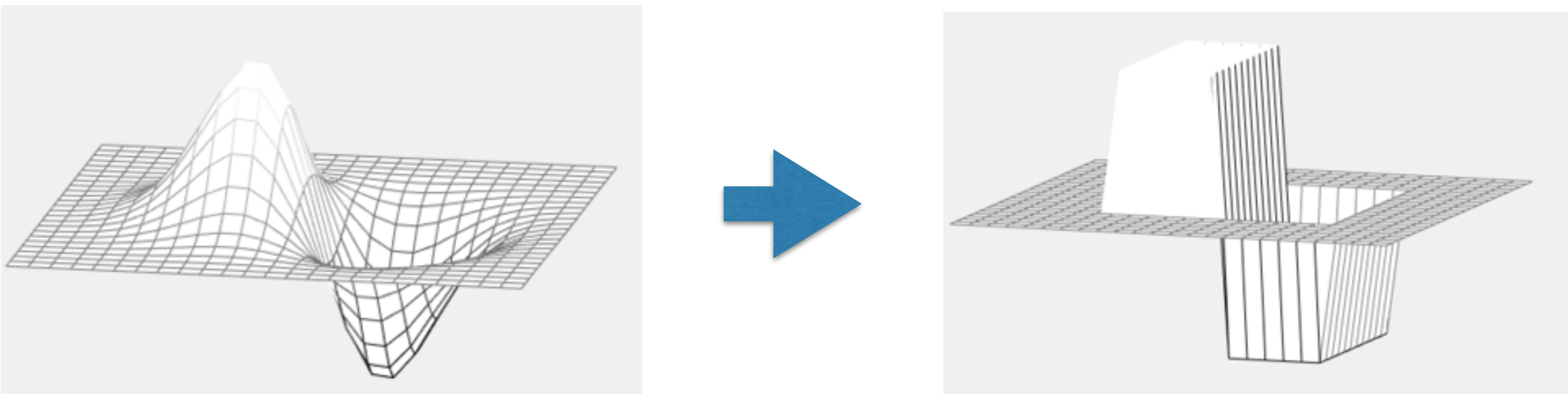
- Does not have to touch all pixels before making a decision.
- Again, lends itself strongly to the employment of SIMD.
- Does not rely on integral images.
- Very good at finding possible corner candidates.
- Can still fire on edges, (can use Harris to remove false positives).
- Is NOT multi-scale.

Today

- Review.
- Efficient Interest Point Detectors.
- Efficient Descriptors.

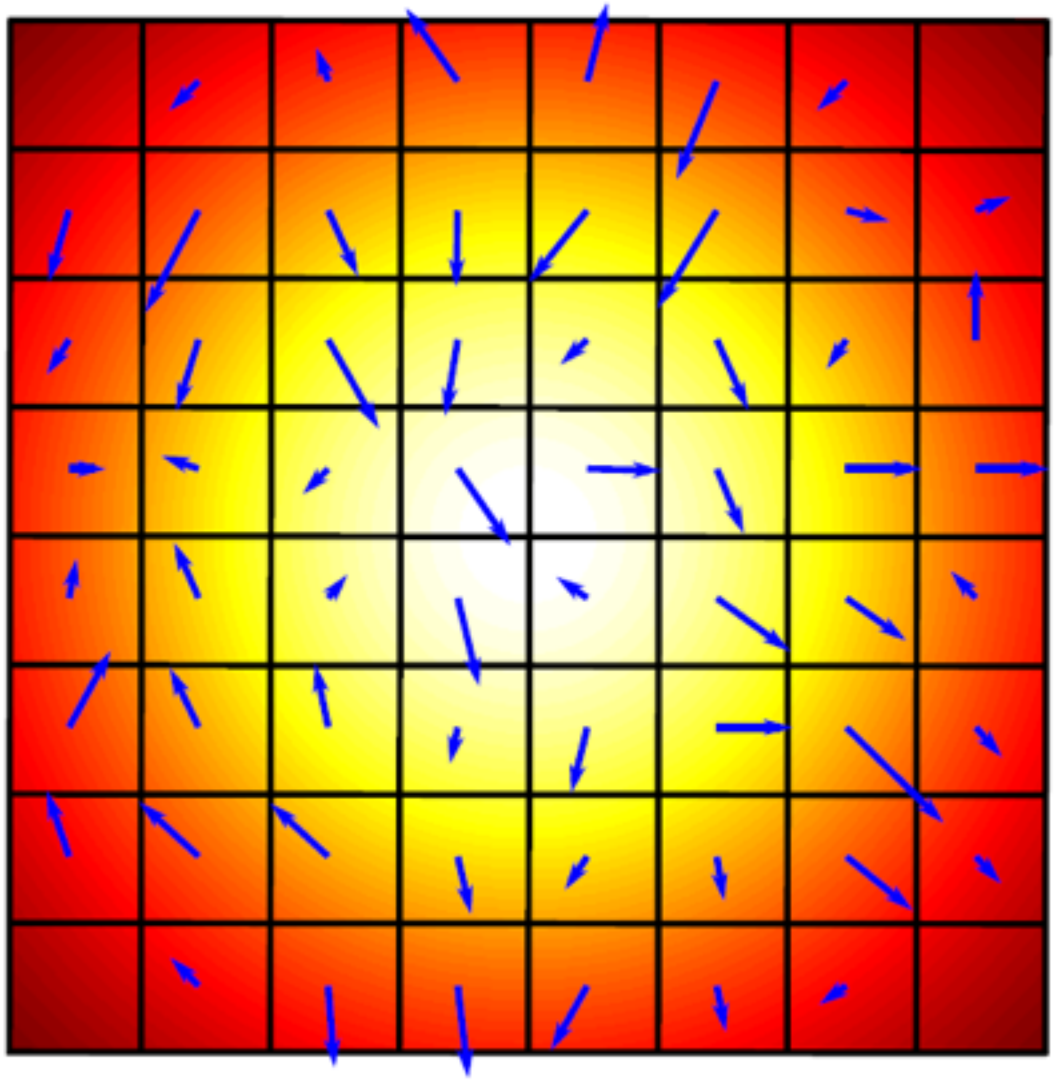
SURF Descriptor

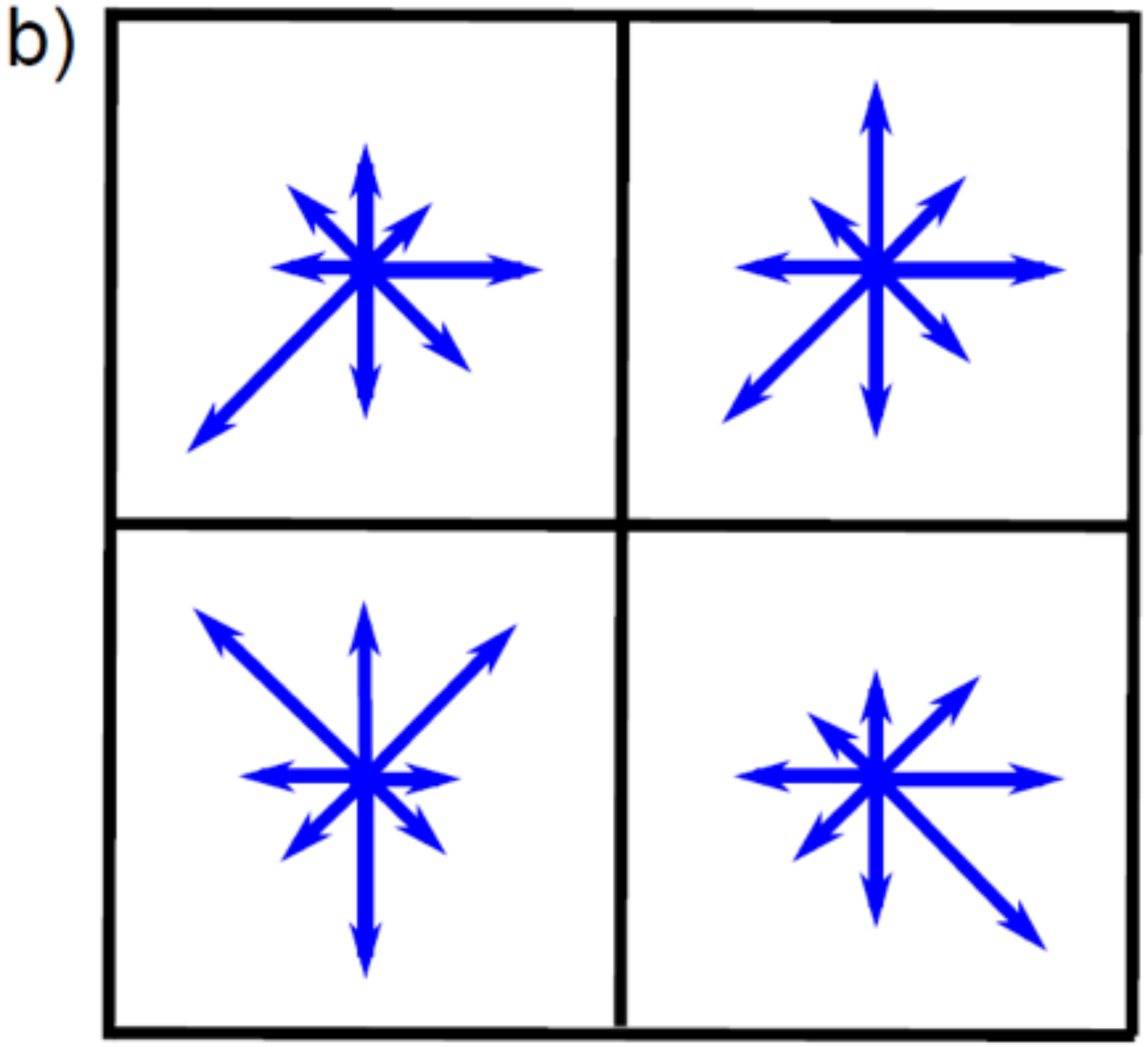
- SURF also proposed a more efficient descriptor extraction strategy using box filters,



- Rest of the descriptor quite similar to SIFT.

Reminder - SIFT Descriptor





1. Compute image gradients

2. Pool into local histograms
3. Concatenate histograms
4. Normalize histograms

BRIEF Descriptor

- Proposed by Calonder et al. ECCV 2010.
- Borrows idea that binary comparison is very fast on modern chipset architectures,

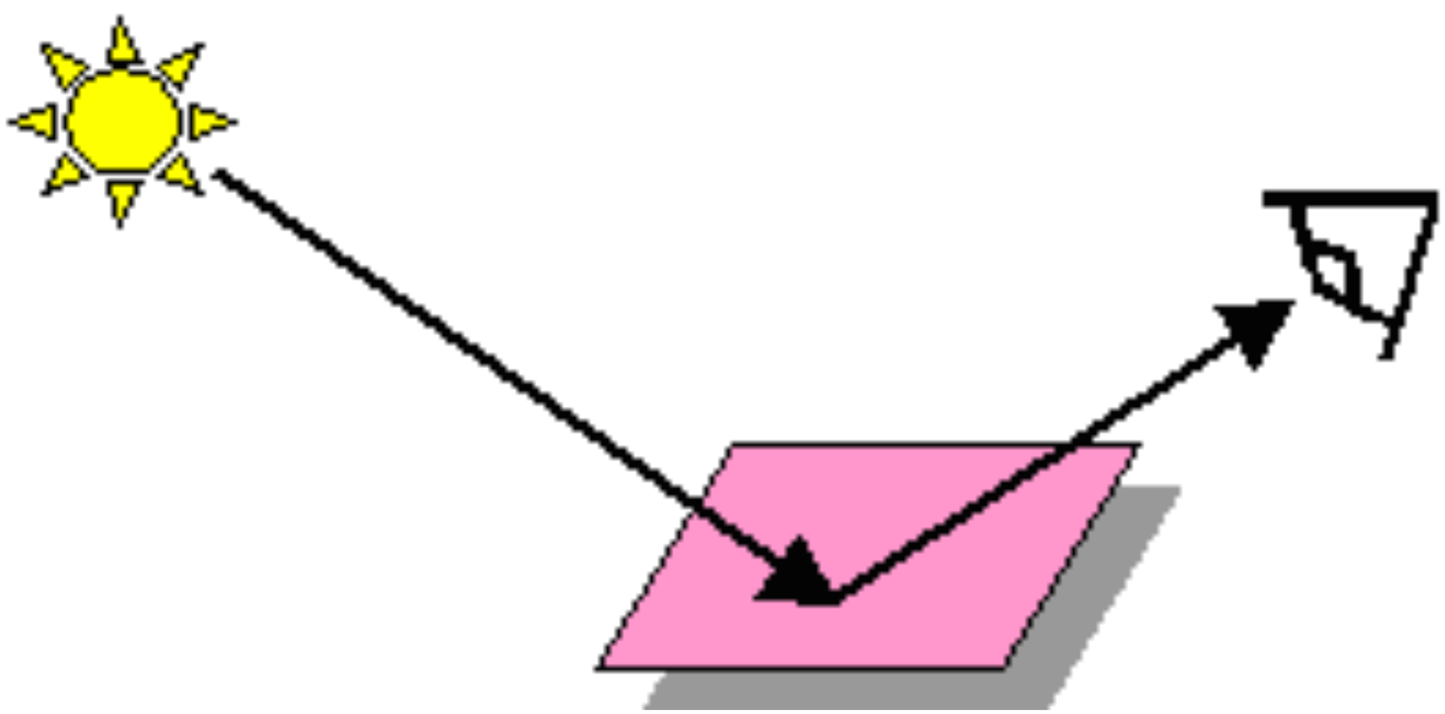
$$\psi_I(\mathbf{x}, \Delta_1, \Delta_2) = \begin{cases} 1 : I(\mathbf{x} + \Delta_1) > I(\mathbf{x} + \Delta_2) \\ 0 : \text{otherwise} \end{cases}$$

- Combine features together compactly,

$$\psi_I(\mathbf{x}) = \sum_i 2^{i-1} \psi_I(\mathbf{x}, \Delta_1^{(i)}, \Delta_2^{(i)})$$

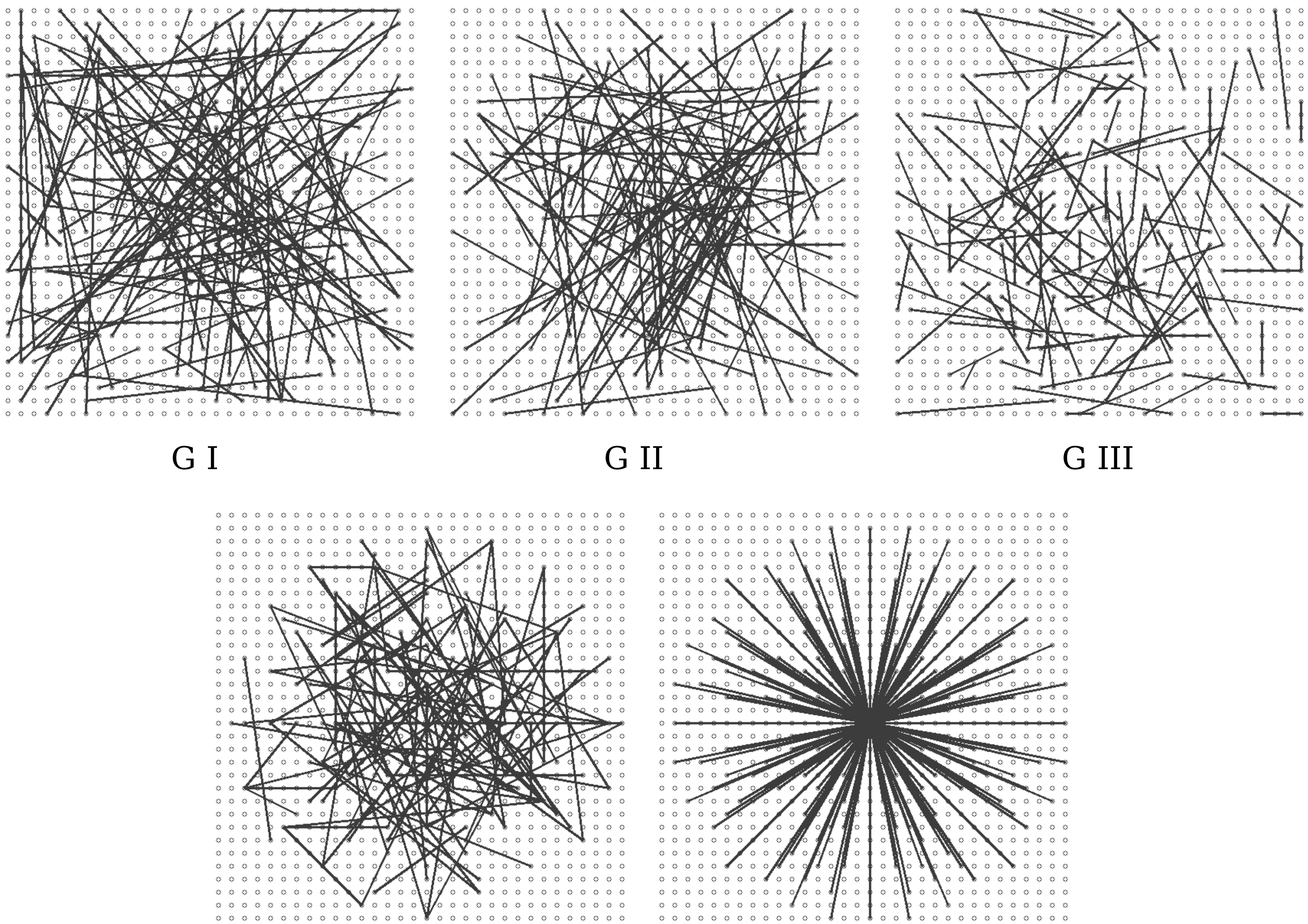
Why do Binary Features Work?

- Success of binary features says something about perception itself.
- Absolute values of pixels do not matter.
- Makes sense as variable lighting source will effect the gain and bias of pixels, but the local ordering should remain relatively constant.



BRIEF Descriptor

- Do not need to “touch” all pixels, can choose $\{\Delta_1, \Delta_2\}$ pairs randomly and sparsely,



G I G II G III

G IV G V

BRIEF Descriptor

- Measuring distance between descriptors,

$$d(\psi\{\mathbf{x}_i\}, \psi\{\mathbf{x}_j\}) = \text{Hamming distance}$$

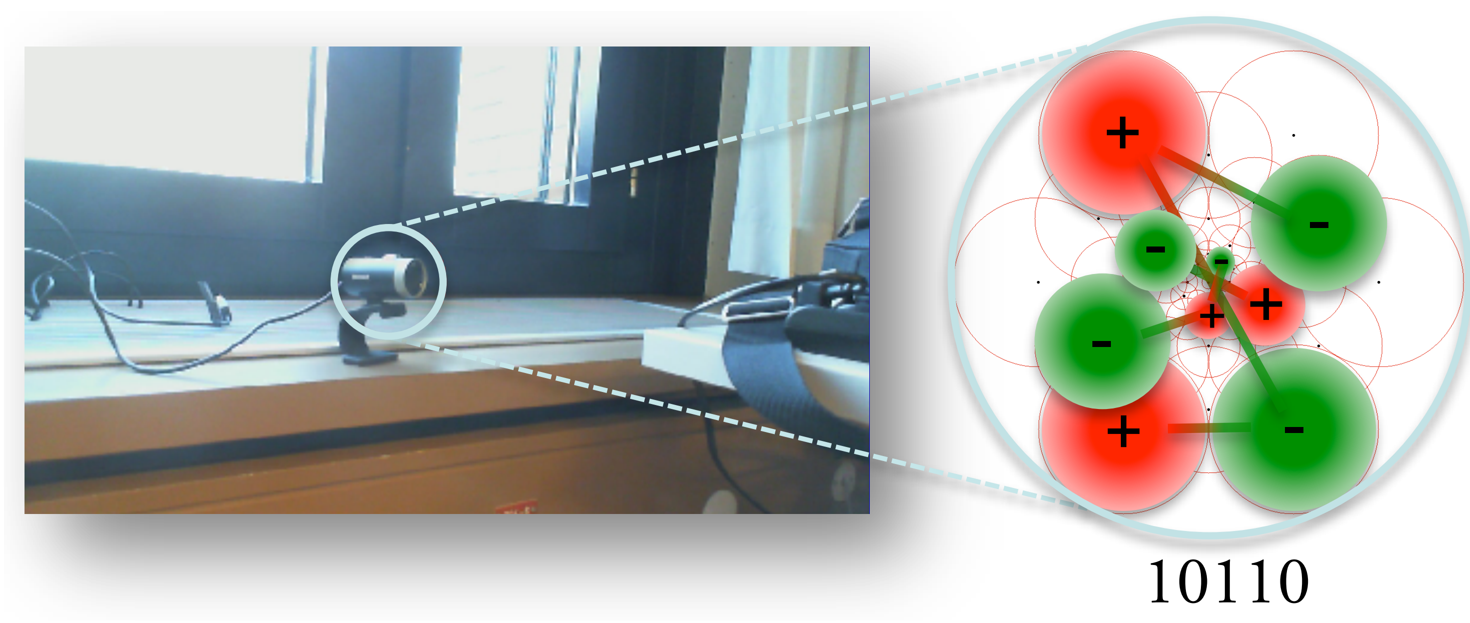
- e.g.,

$$d\left(\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}\right) = 1$$

Why not use Euclidean distance?

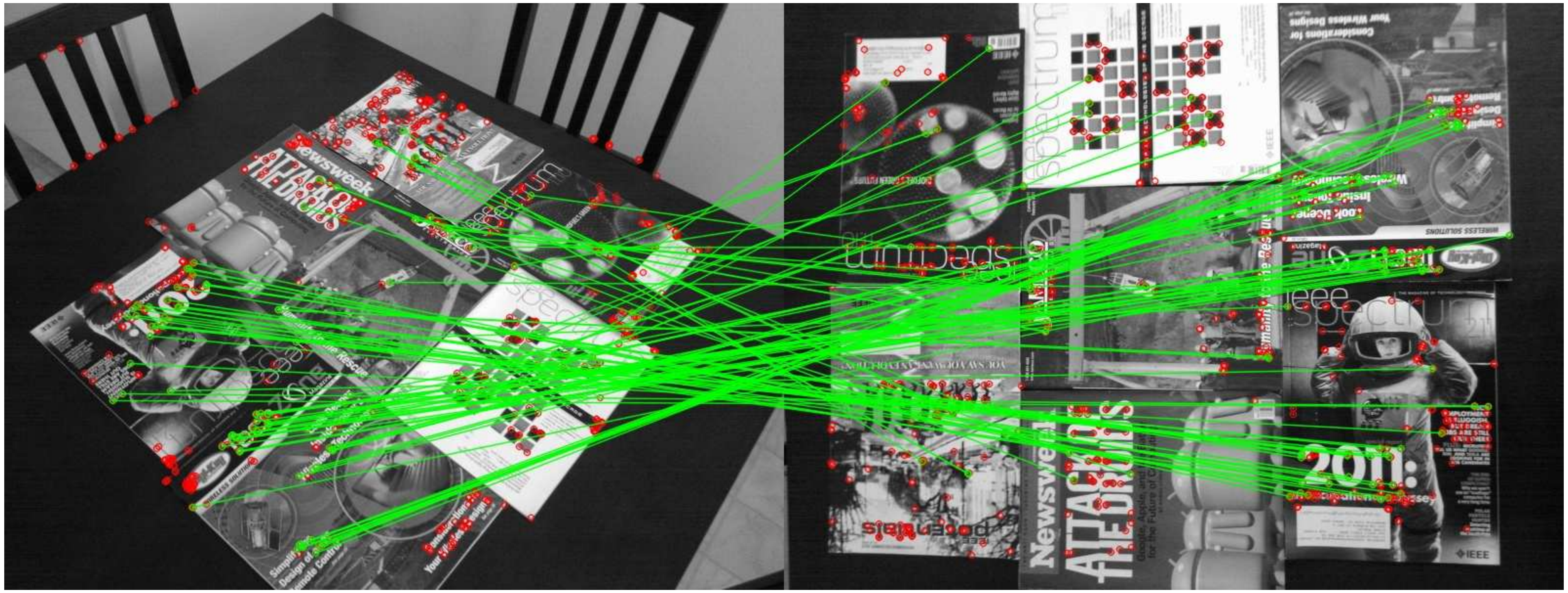
Binary Descriptor Variants

- At the same time that BRIEF was proposed similar variants were also proposed and have gained some popularity.
  - BRISK - **B**inary **R**obust **I**nvariant **S**calable **K**eypoints - Leutenegger 2011.
  - FREAK - **F**ast **RE**tin**A** **K**eypoints - Alahi et al. 2012.
- Both use AGAST for keypoint detection an extension of FAST proposed by Mair et al. 2010.



ORB

- Rublee et al. ICCV 2011 proposed Oriented FAST and Rotated BRIEF (ORB).
- Essentially combines FAST with BRIEF.
- Demonstrated that ORB is 2 **orders** of magnitude faster than SIFT.
- Very useful for mobile devices.



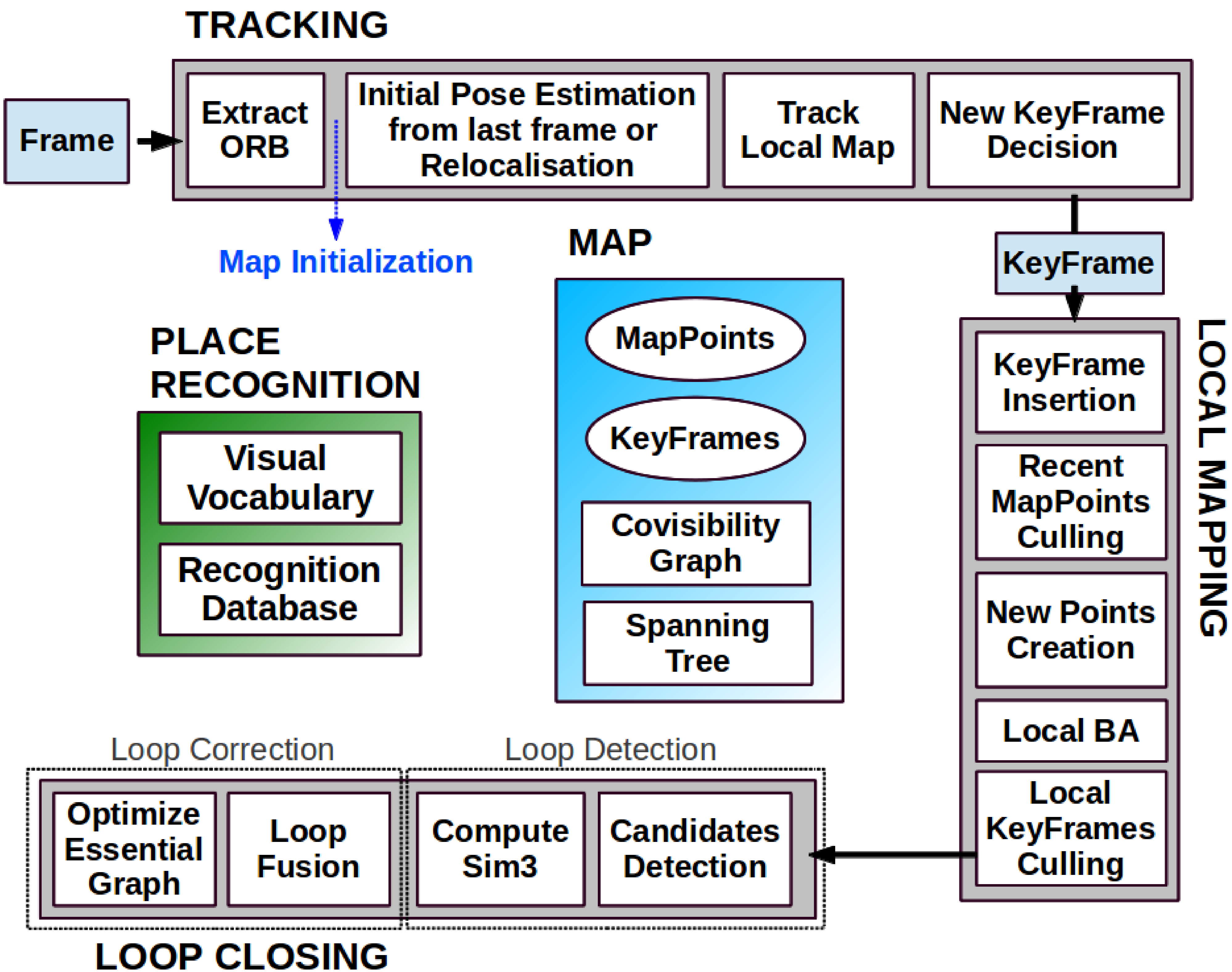
ORB

- ORB is patent free and available in OpenCV,

Detector	ORB	SURF	SIFT
Time per frame (ms)	**15.3**	217.3	5228.7

These times were averaged over 24 640x480 images from the Pascal dataset [9]. ORB is an order of magnitude faster than SURF, and over two orders faster than SIFT.

ORB SLAM



Taken from: Mur-Artal, Raul, J. M. M. Montiel, and Juan D. Tardós. "Orb-slam: a versatile and accurate monocular slam system." IEEE Transactions on Robotics 31.5 (2015): 1147-1163.

ORB SLAM

ORB-SLAM

Raúl Mur-Artal, J. M. M. Montiel and Juan D. Tardós

{raulmur, josemari, tardos} @unizar.es

Instituto Universitario de Investigación
en Ingeniería de Aragón
Universidad Zaragoza

Universidad
Zaragoza
1542

Taken from: Mur-Artal, Raul, J. M. M. Montiel, and Juan D. Tardós. "Orb-slam: a versatile and accurate monocular slam system." IEEE Transactions on Robotics 31.5 (2015): 1147-1163.

More to read…

BRIEF: Binary Robust Independent Elementary Features*

Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua

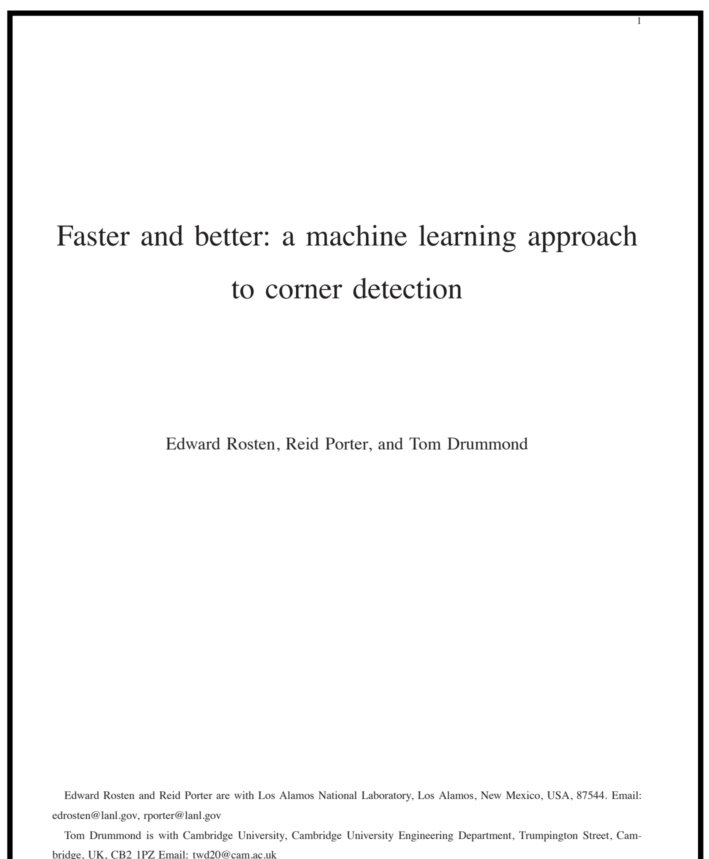
CVLab, EPFL, Lausanne, Switzerland

- E. Rublee et al. "ORB: an efficient alternative to SIFT or SURF", ICCV 2011.
- Calonder et al. "BRIEF: Binary Robust Independent Elementary Features", ECCV 2010.
- E. Rosten et al. "Faster and better: A machine learning approach to corner detection" IEEE Trans. PAMI 2010.
- P. Kovesi "Fast Almost Gaussian Filtering" DICTA 2010.

ORB: an efficient alternative to SIFT or SURF

Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski

Willow Garage, Menlo Park, California



Faster and better: a machine learning approach to corner detection

Edward Rosten, Reid Porter, and Tom Drummond

Edward Rosten and Reid Porter are with Los Alamos National Laboratory, Los Alamos, New Mexico, USA, 87544. Email: edrosten@lanl.gov, rporter@lanl.gov

Tom Drummond is with Cambridge University, Cambridge University Engineering Department, Trumpington Street, Cambridge, UK, CB2 1PZ Email: twd20@cam.ac.uk

2010 Digital Image Computing: Techniques and Applications

Fast Almost-Gaussian Filtering

Peter Kovesi