Machine Visions and
Digital Image Analysis

Scale Invariant Feature Transform

developed by David Lowe
University of British Columbia
Initial paper ICCV 1999
Newer journal paper IJCV 2004

LAPPEENRANTA
UNIVERSITY OF TECHNOLOGY

Hani M. A. Fahmi

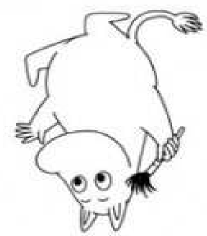
Introduction

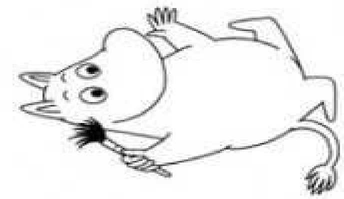
Scale-Invariant Feature Transform (SIFT) [1][2]

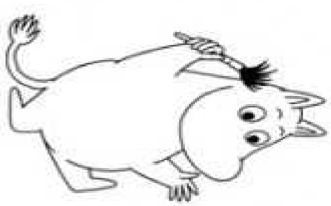
- by David G. Lowe, 1999,2004
- Feature detection algorithm
- Feature usefull for : Image matching, object or scene recognition, solving for 3D structure from multiple images, stereo correspondence and motion tracking,...
- Features properties (invariance) :
  - Scale
  - Rotation
  - Affine transform
  - 3D viewpoint
  - Illumination changes
  - Occlusion

Features invariance

LAPPEENRANTA
UNIVERSITY OF TECHNOLOGY







Images source : wikipedia.org

Sift Method Steps [1]

LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

SIFT algorithm:

1.Scale-space extrema detection

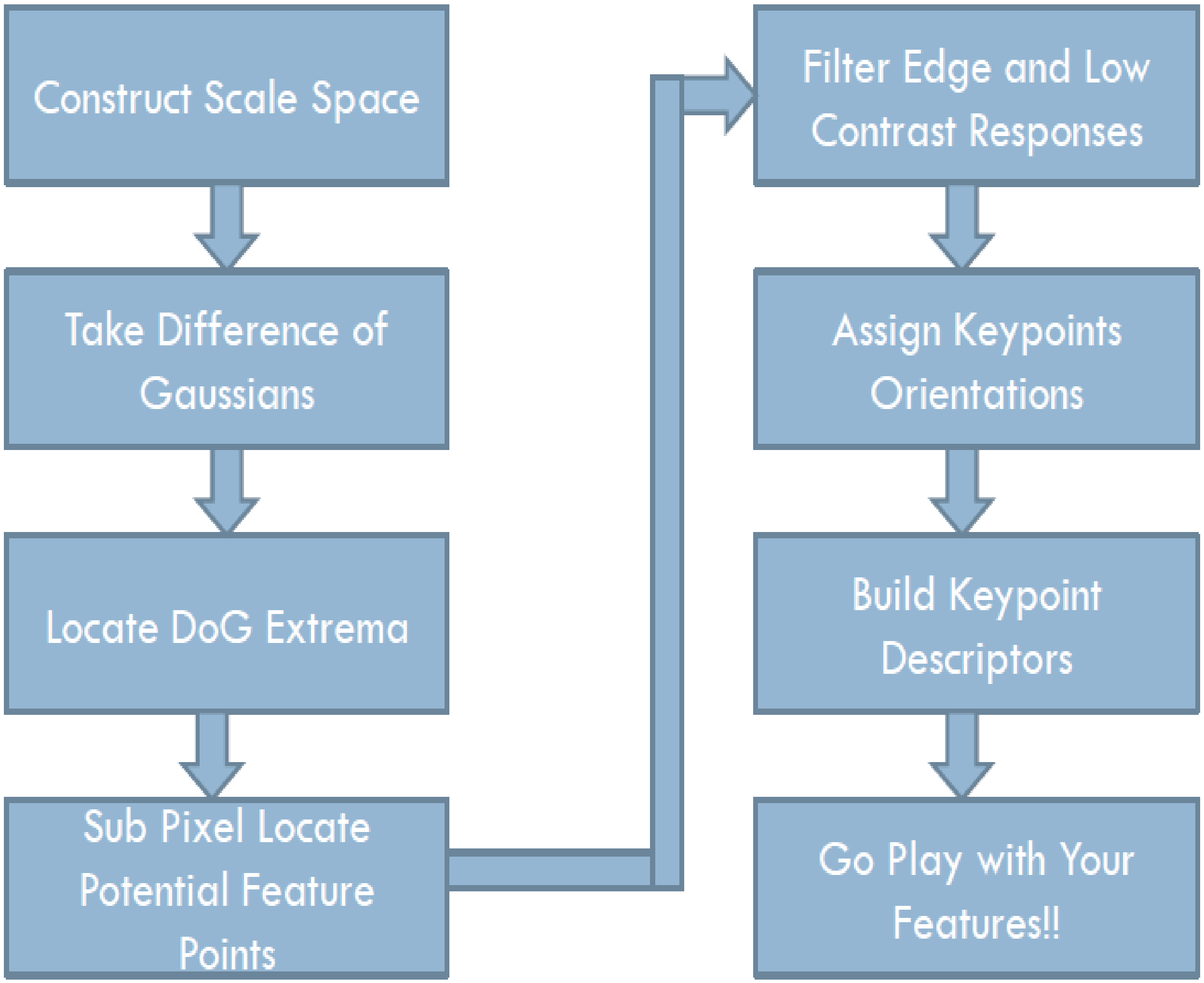
2. Keypoint localization.

3. Orientation assignment.

4. Keypoint descriptor.

Output : feature descriptors

Sift Algorithm overview [1][4]



Images source : Jason Clemons presentation [4]

Scale Space Construction [1][4][15]

- Gaussian kernel used to create scale space
- Only possible scale space kernel (Lindberg , 94)

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y),$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}.$$

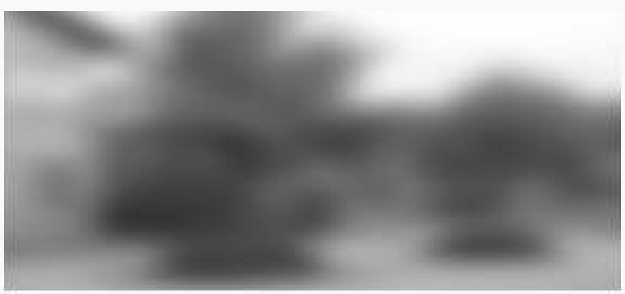
Scale Space representation [1][15]

Scale-space representation $L(x,y;t)$ at scale $t = 0$, corresponding to the original image f

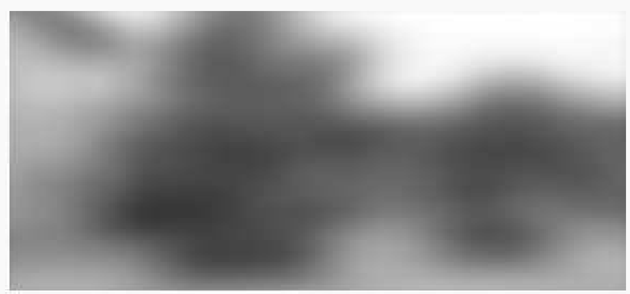
Scale-space representation $L(x,y;t)$ at scale $t = 1$

Scale-space representation $L(x,y;t)$ at scale $t = 4$

Scale-space representation $L(x,y;t)$ at scale $t = 16$



Scale-space representation $L(x,y;t)$ at scale $t = 64$

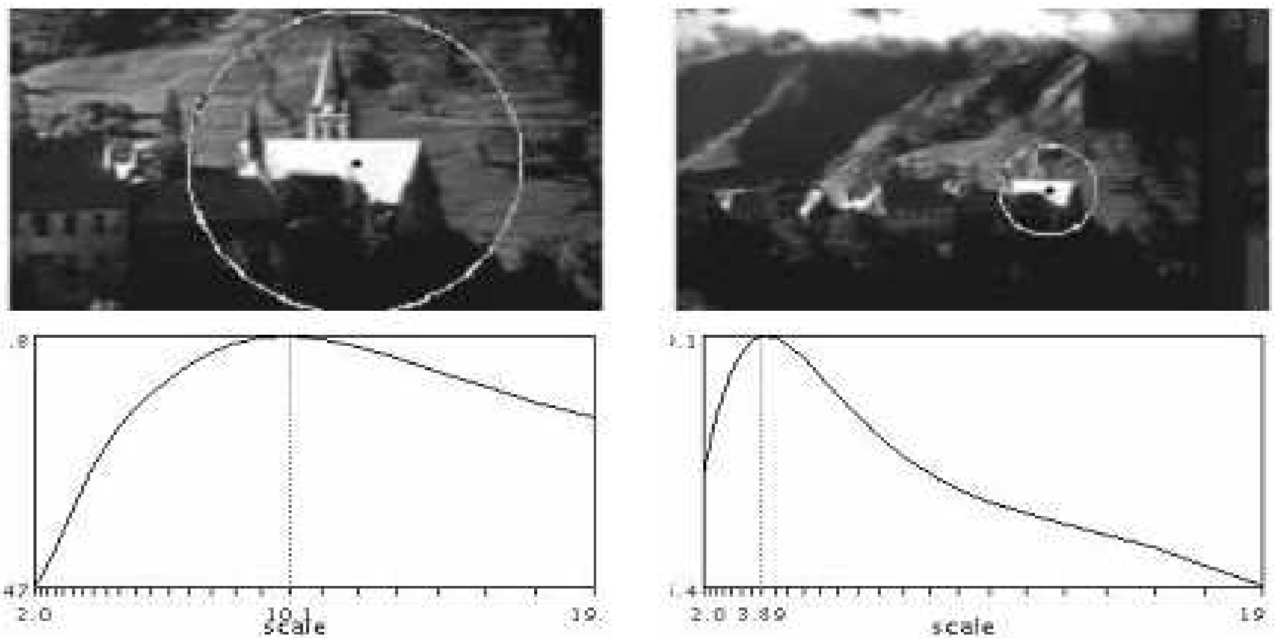


Scale-space representation $L(x,y;t)$ at scale $t = 256$

Images source : wikipedia.org [15]

Scale Space extrema [1][3][12][15]

Mikolajczyk (2002): Experimentally, extrema of LoG gives best notion of scale



Credit and image source : Ilan Shimshoni slides [12]

Scale Space extrema localization

[1][4]

- Approximation of Laplacian of Gaussians

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G.$$

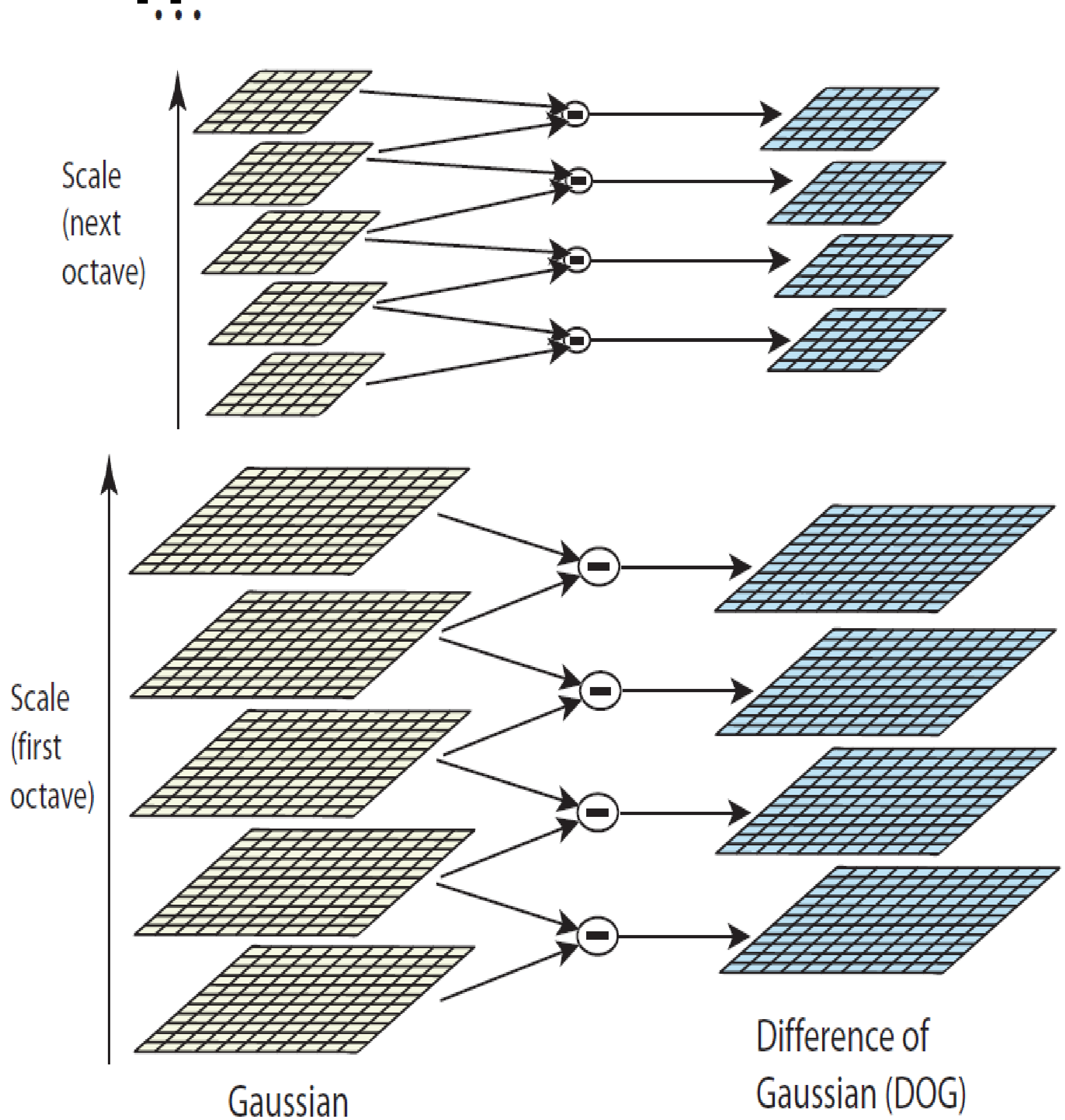
$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2 \nabla^2 G.$$

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma). \end{aligned}$$

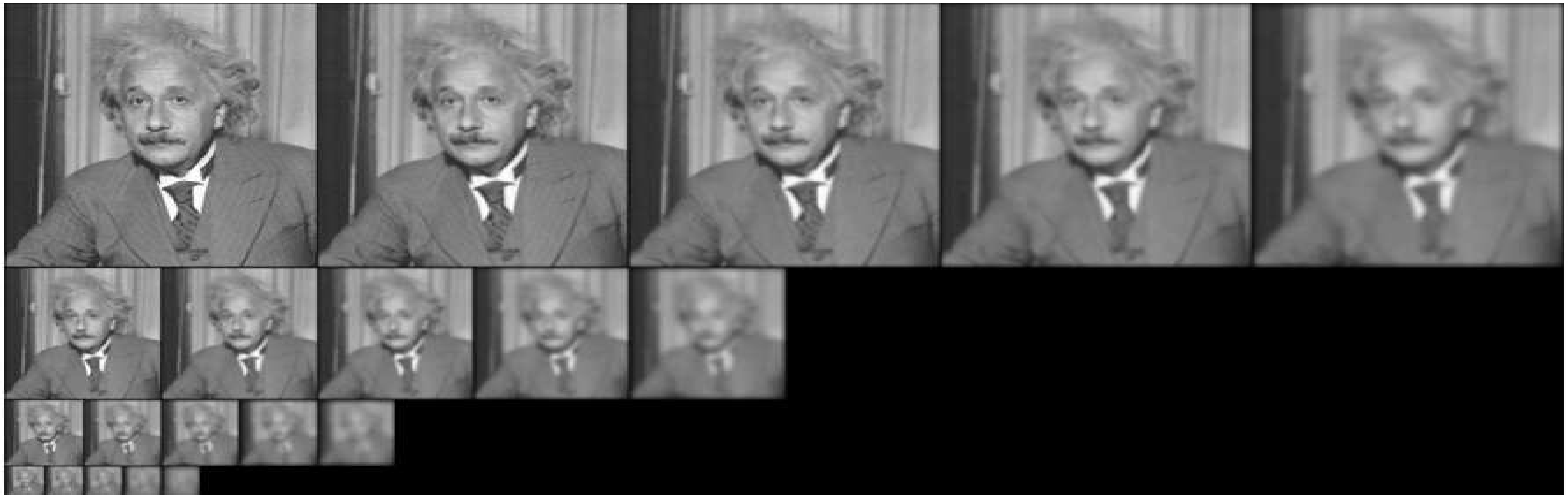
$s = 3$ and $k = 2^{1/s}$.

Scale Space extrema localization (keypoint) overview [1]

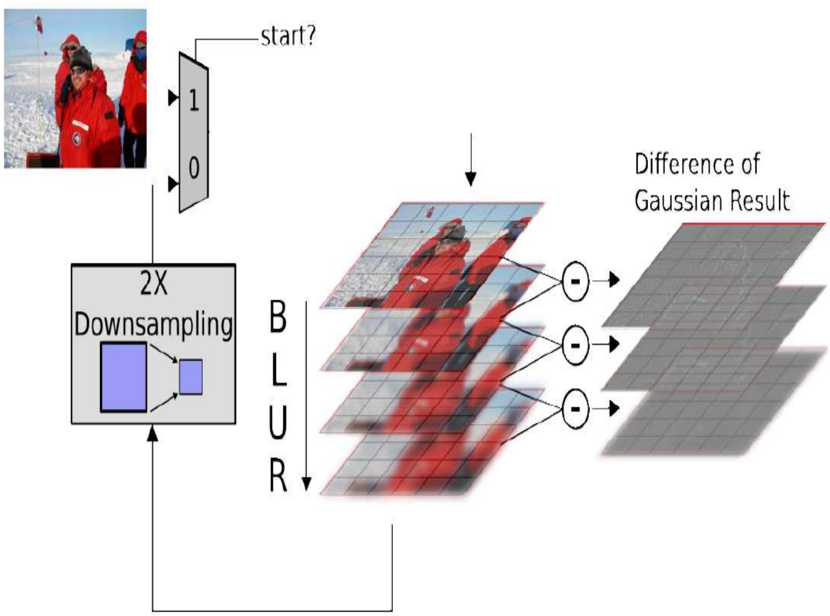


Images source :Lowe 2004 [1]

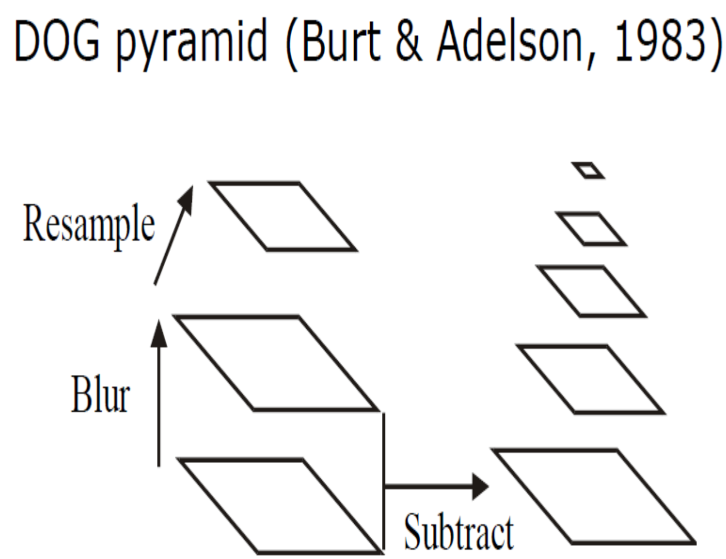
Example of DOG 1/2 [1][8][13]



Images source :Allan Jepson [8]

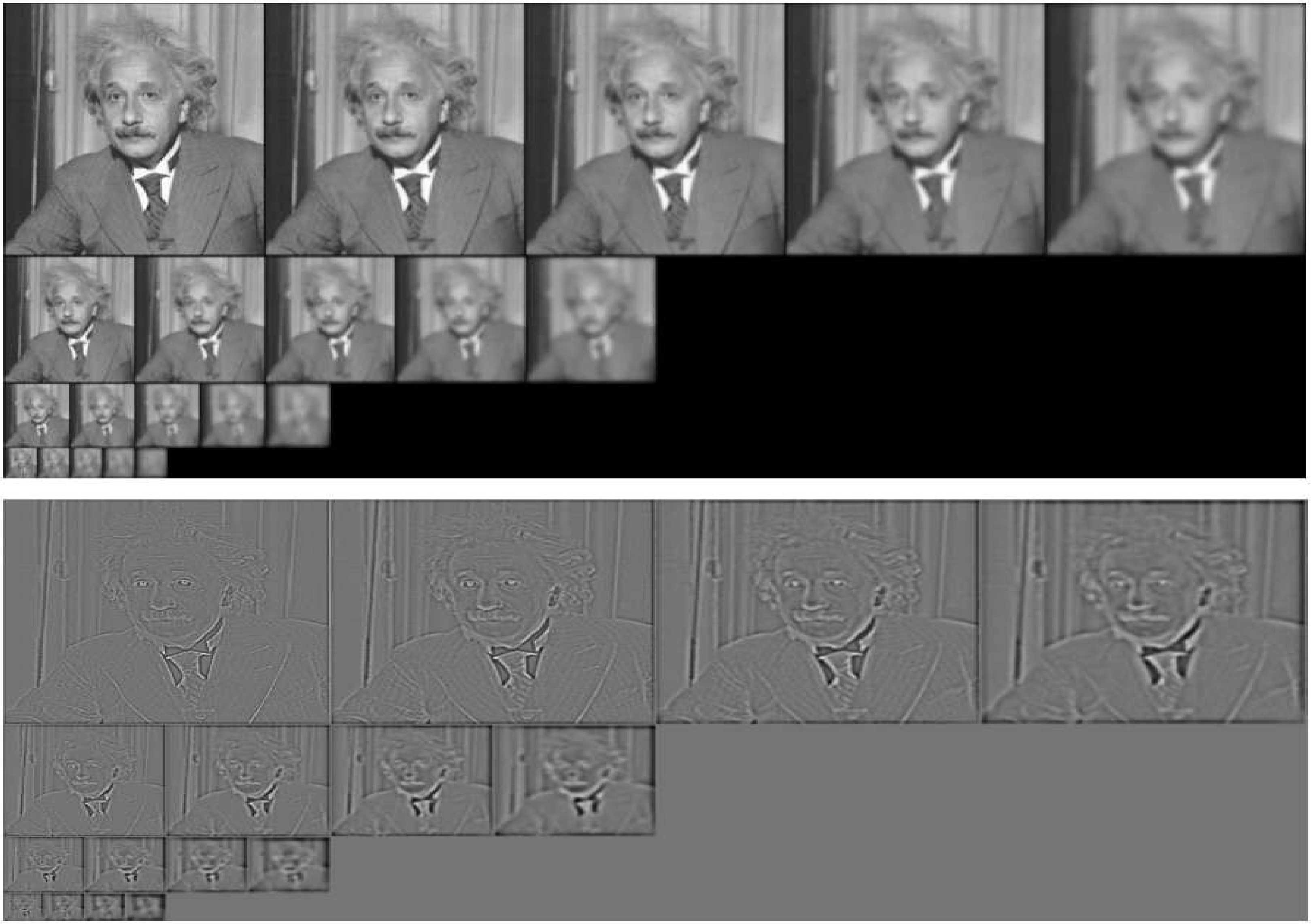


Images source :[13]



Images source :
[13]

Example of DOG 2/2 [1][8]



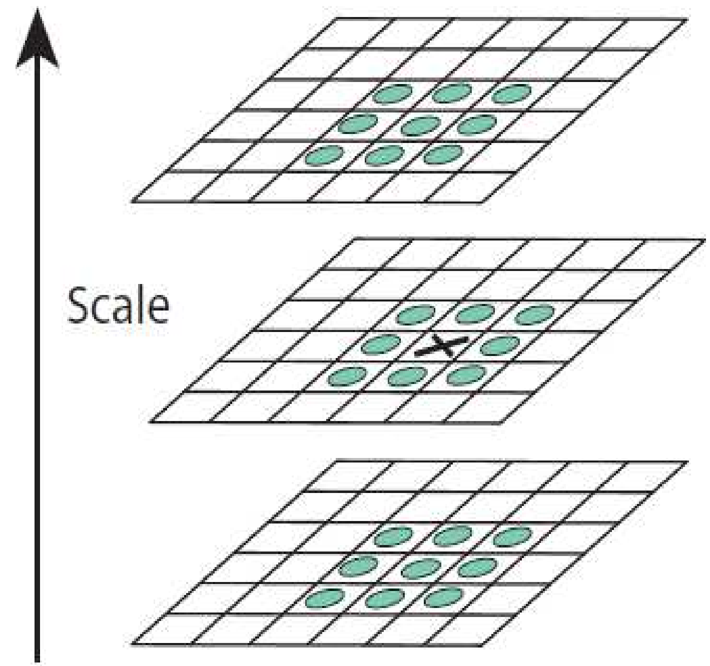
Range: [−0.11, 0.131]
Dims: [959, 2044]

Images source :Allan Jepson [8]

DOG exterma localization [1][4]

- ✓Scan each DOG image
- ✓Look at all neighboring points (including scale)
- ✓Identify Min and Max

26 Comparisons



Slide credit : Jason clemons [4]

Images source :Lowe 2004 [1]

Extra keypoints elimination 1/2 [1] [4]

LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

3D Curve Fitting for localization : Taylor Series Expansion + differentiation

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}}^T \mathbf{x} + \frac{1}{2}\mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x}$$

$$\hat{\mathbf{x}} = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1} \frac{\partial D}{\partial \mathbf{x}}.$$

Low Contrast Points Filter

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2}\frac{\partial D}{\partial \mathbf{x}}^T \hat{\mathbf{x}}.$$

Slide credit : Jason clemons [4]

Extra keypoints elimination 2/2 [1] [4]

Edge Response Elimination :

- ✓Hessian
- ✓Eigenvalues Proportional to principle
- ✓CurvaturesTrace and Determinant

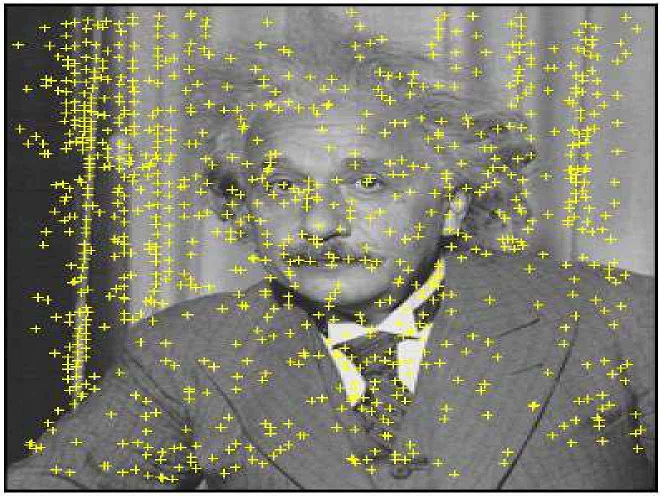
$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

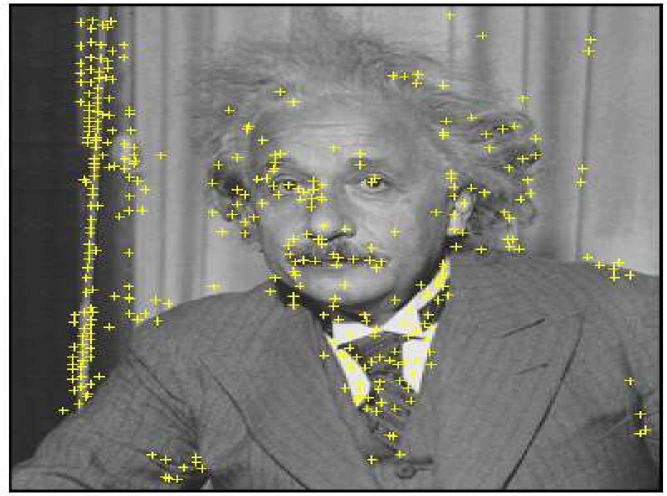
$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$

$$Tr(H) = D_{xx} + D_{yy}$$
$$Det(H) = D_{xx} D_{yy} - (D_{xy})^2$$

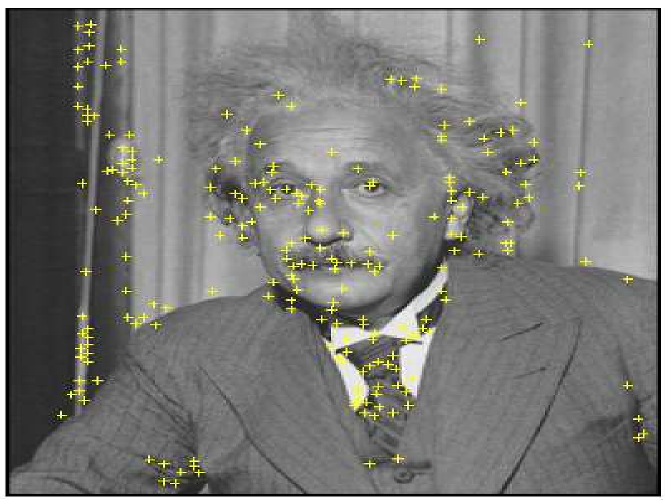
Slide credit : Jason clemons [4]

Extra keypoints elimination example [8]





a) b)



c)

Images source :Allan Jepson [8]

Orientation and feature descriptor [1][4][6]

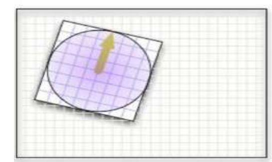
Compute Gradient for each blurred image

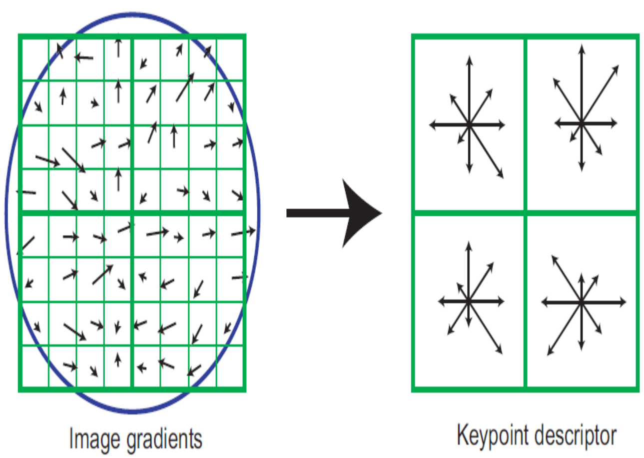
$$L(x,y)=G(x,y,\sigma)*I(x,y)$$

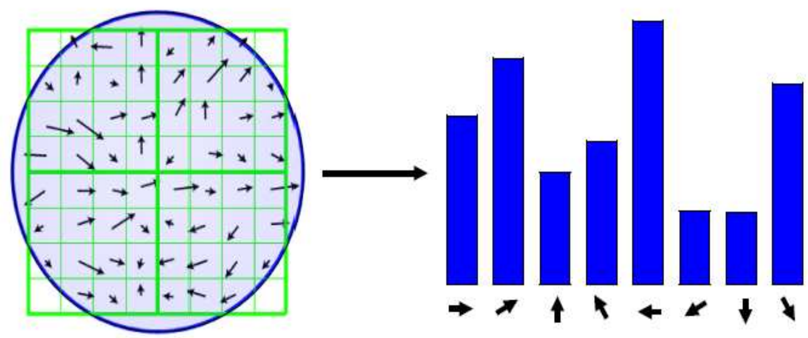
$$m(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2}$$

$$\theta(x,y)=\tan^{-1}\left(\frac{(L(x,y+1)-L(x,y-1))}{(L(x+1,y)-L(x-1,y))}\right)$$

Implementation uses 4x4 descriptors
from 16x16 which leads to a 4x4x8=128 element vector





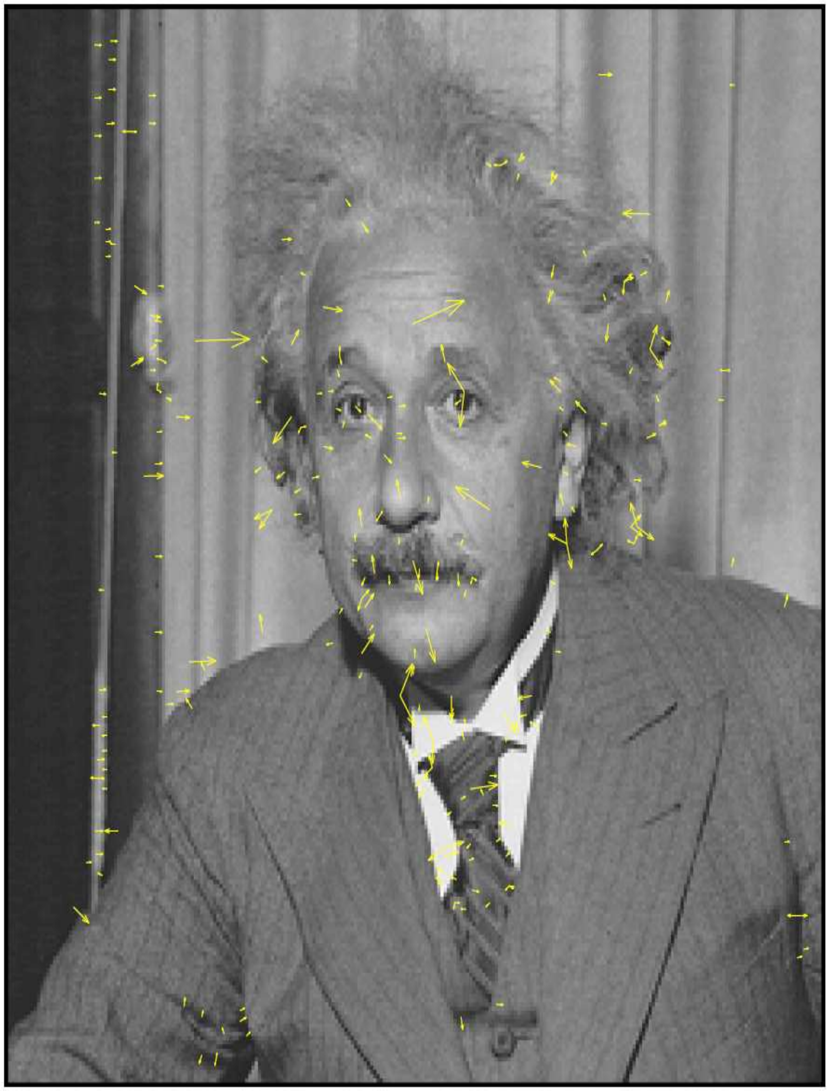


Images source :Lowe 2004 [1]

Slide credit : Jason clemons [4]

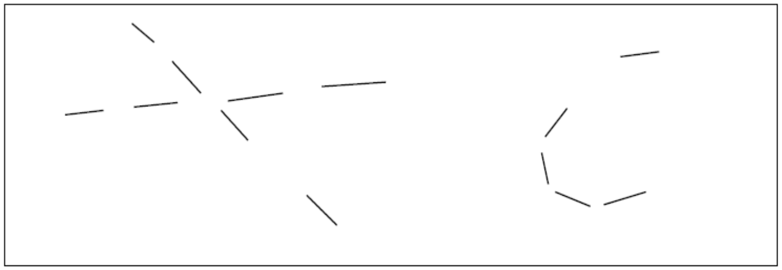
Images source : Ofir Pele [6]

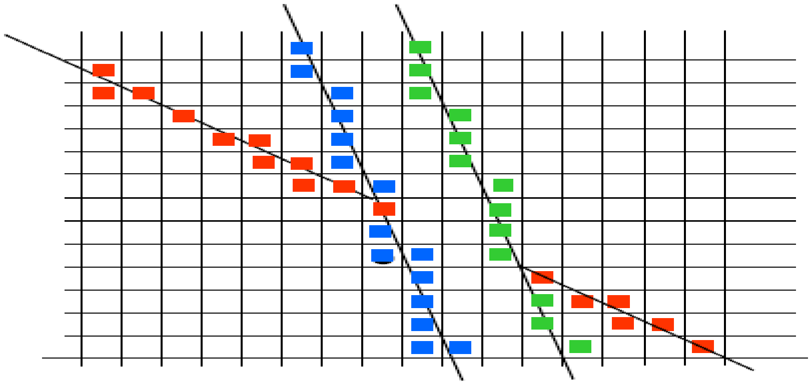
Orientation example [8]

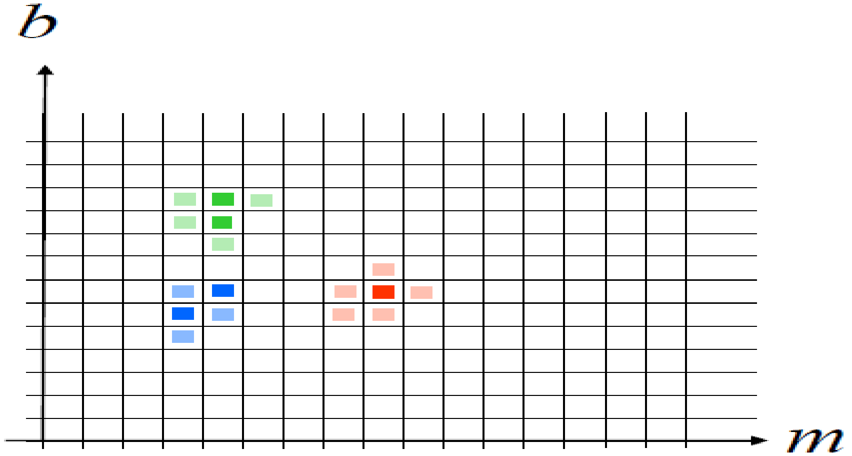


Images source :Allan Jepson [8]

Hough transform : Global features [1][13]



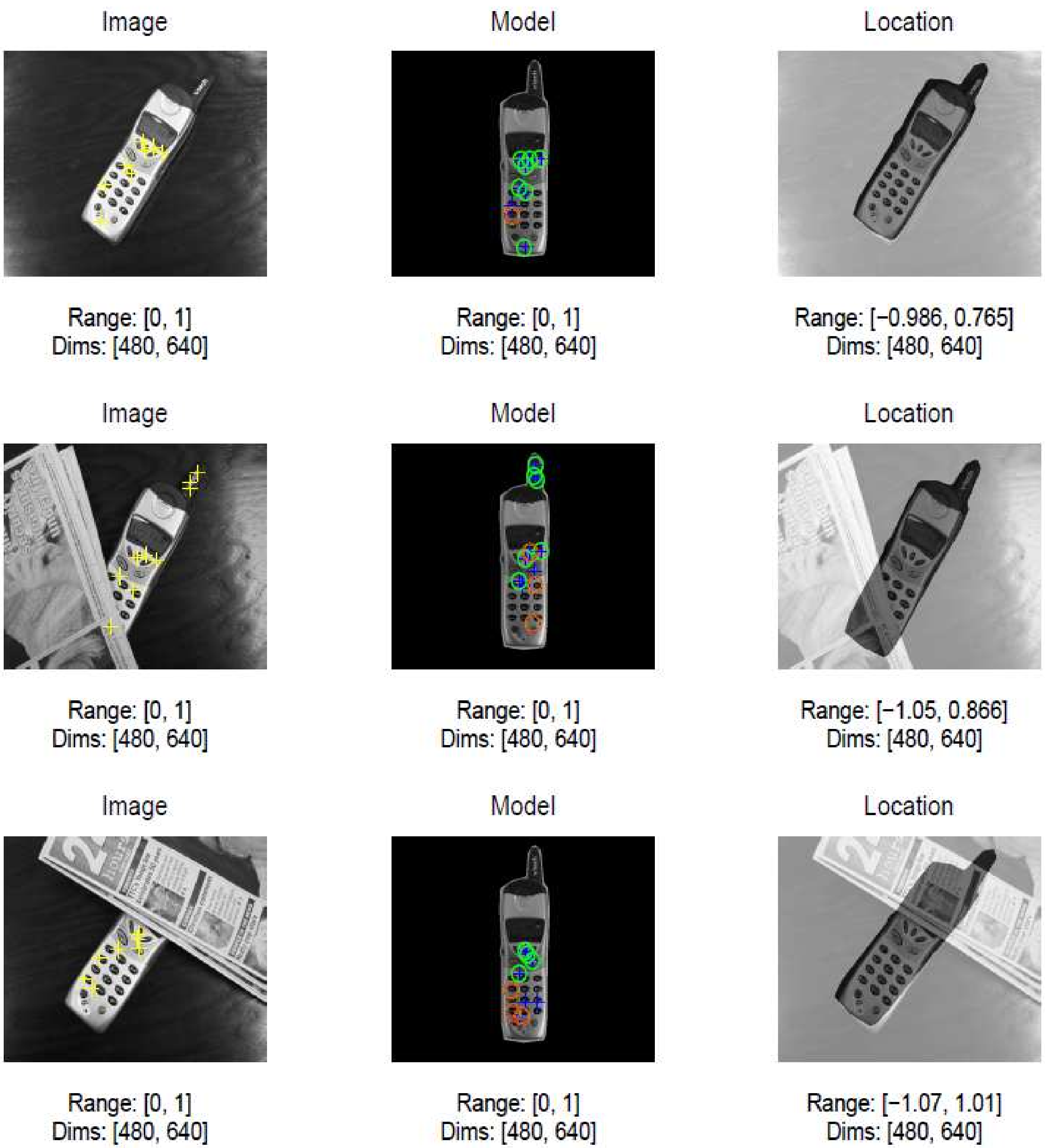




Images source :[13]

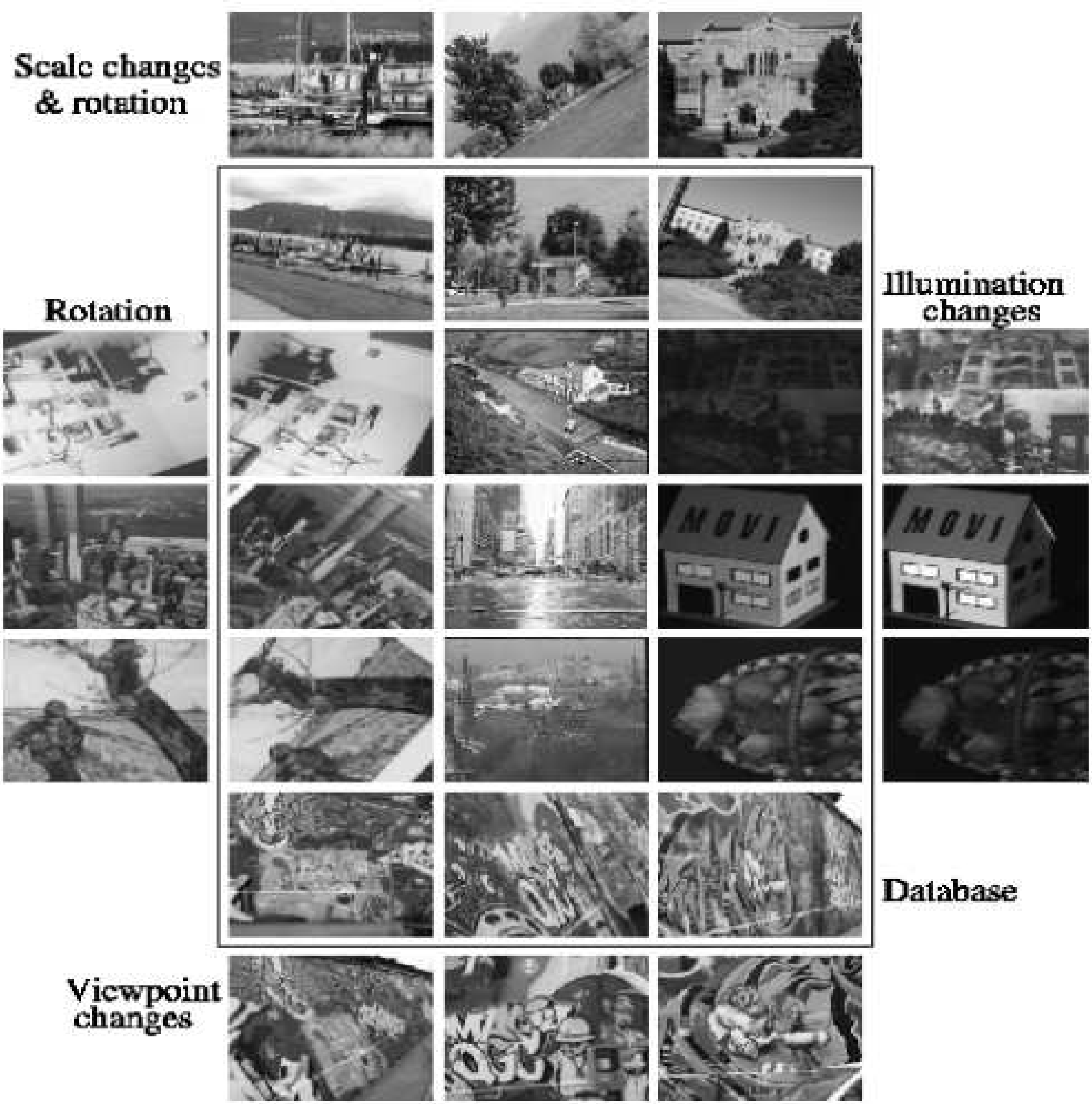
SIFT example : Find the cellphone ?

[8]



Images source : Allan jepson [8]

SIFT , Features database [13]

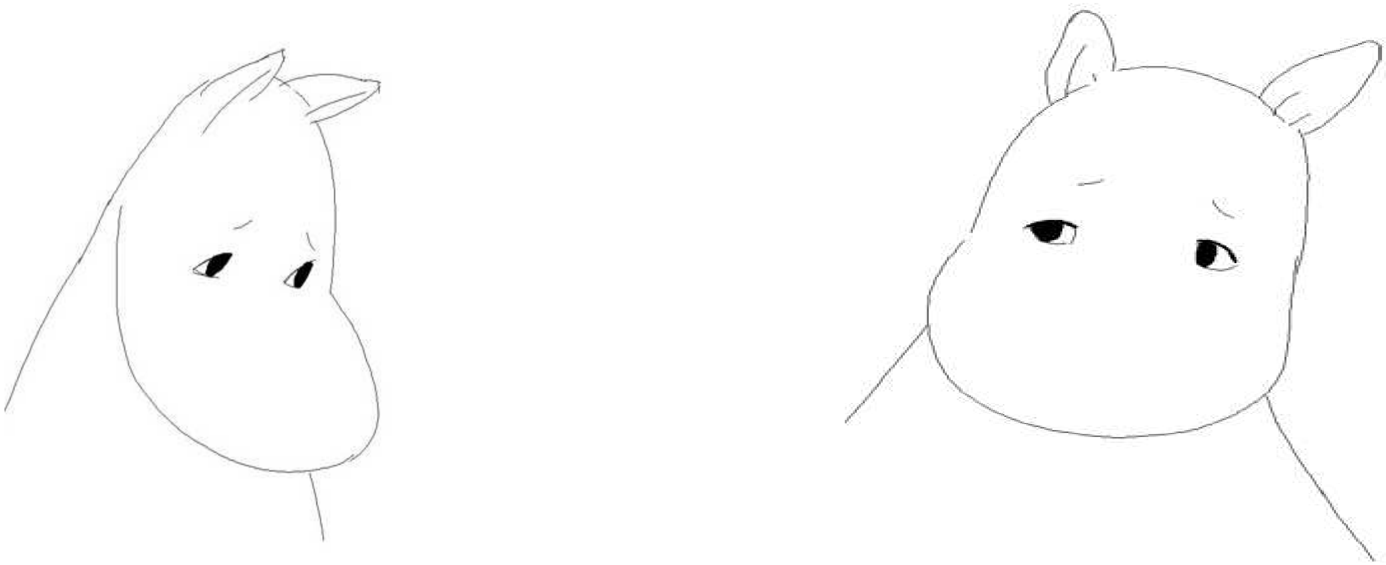


Images source :[13]

SIFT limitation [6]

Large illumination change

Non rigid deformations



SIFT Application [5]

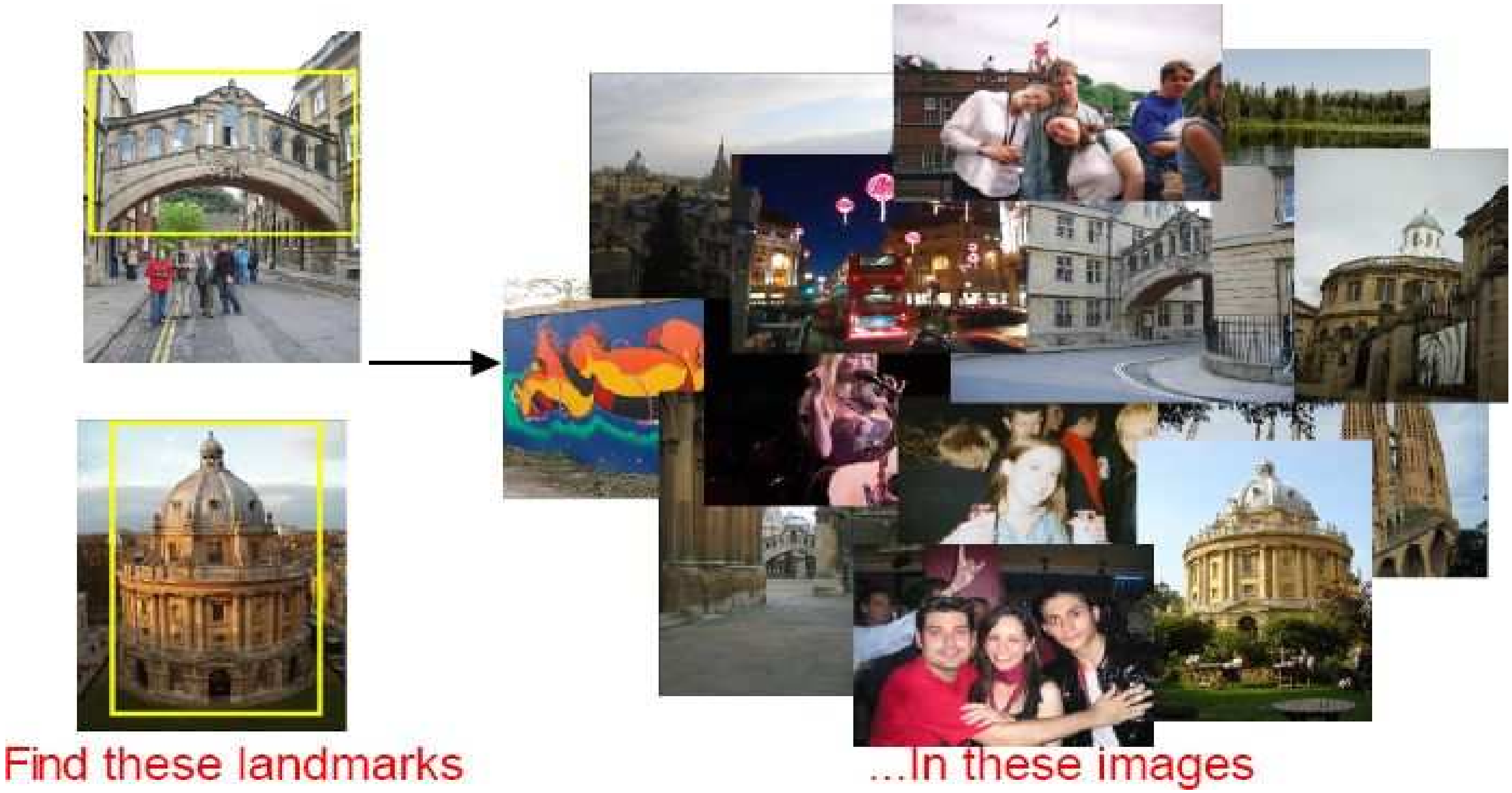
LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

- Object recognition using SIFT features
- Robot localization and mapping
- Panorama stitching
- 3D scene modeling, recognition and tracking
- Others,(human action recognition)

SIFT Demo, Visual. Geometry Group Oxford University [7]

LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

http://www.robots.ox.ac.uk/~vgg/research/oxbuildings/index.html



Images source :Visual. Geometry Group[7]

Conclusion [6]

- In wide use both in academia and industry
- Many available implementations:
  - Binaries available at Lowe's website
  - C/C++ open source by A. Vedaldi (UCLA)
  - C# library by S. Nowozin (Tu-Berlin)
- Protected by a patent

Slide credit : Ofir Pele [6]

References

LAPPEENRANTA UNIVERSITY OF TECHNOLOGY

[1] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, 60, 2 (2004), pp. 91-110, retreived 3-11-2011 from : http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf

[2] David G. Lowe, "Object recognition from local scale-invariant features," International Conference on Computer Vision, Corfu, Greece (September 1999), pp. 1150-1157 retreived 3-11-2011 from : http://www.cs.ubc.ca/~lowe/papers/iccv99.pdf

[3] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors. In PAMI 27(10):1615-1630 . retreived 3-11-2011 from : http://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_pami2004.pdf

[4] Jason Clemons, University of Michigan ,course EECS598, lecture presentation 10_1 retreived from : http://www.eecs.umich.edu/~silvio/teaching/EECS598/lectures/lecture10_1.pdf

[5] http://en.wikipedia.org/wiki/Scale-invariant_feature_transform

[6] Ofir Pele,The Rachel and Selim Benin School of Computer Science and Engineering , computer vision 2006 course slides retrieved 3-11-2011 from : http://www.cse.huji.ac.il/course/2006/compvis/lectures/SIFT.ppt

[7] Visual. Geometry Group,. University of Oxford , Web-scale particular object search : http://www.robots.ox.ac.uk/~vgg/research/oxbuildings/index.html

[8] Allan Jepson,Sift tutorial , University of Toronto,CS department, course csc2503 : http://www.cs.toronto.edu/~jepson/csc2503/tutSIFT04.pdf

[9] http://en.wikipedia.org/wiki/Image_matching

[10] http://en.wikipedia.org/wiki/Feature_detection_%28computer_vision%29

[11] http:// en.wikipedia.org/wiki/Blob_detection

[12] Ilan Shimshoni, University of Haifa, Slide presentation about SIFT from 048972 Course in Multiple View Geometry : http://mis.hevra.haifa.ac.il/~ishimshoni/MVG/SIFT_KA.ppt

[13] Presentation slides : Cornell University, CS department , CS664 course Lecture 21: SIFT, object recognition, dynamic programming : http://www.cs.cornell.edu/courses/cs664/2005fa/Lectures/lecture21.pdf

[14] Ballard J.S. Blair, MIT-WHOI joint program, difference of Gaussian paper retrieved 2-11-2011 from : http://www.ballardblair.com/projects/Difference_of_Gaussian_paper.pdf

[15] http://en.wikipedia.org/wiki/Scale-space