EECS150 - Digital Design
Lecture 14 – FIFO 2 and SIFT

Oct. 15, 2013

Prof. Ronald Fearing
Electrical Engineering and Computer Sciences
University of California, Berkeley

(slides courtesy of Prof. John Wawrzynek)

http://www-inst.eecs.berkeley.edu/~cs150

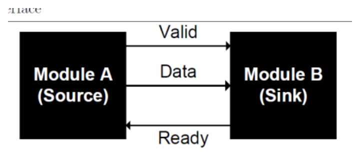
Recap and Outline

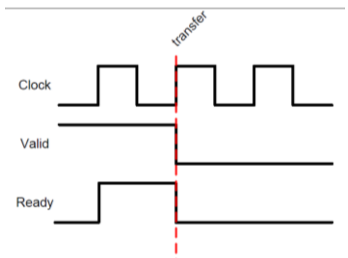
- MicroBlaze connections to feature detector (FSL) and frame buffer (Processor Local Bus)

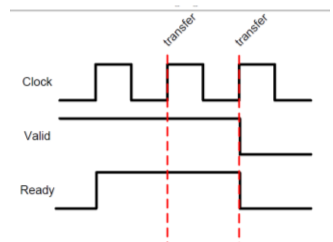
Outline for Today

- Ready/Valid Handshaking
- more FIFO details
- SIFT Algorithm

Valid/Ready Handshake



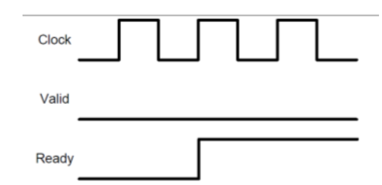


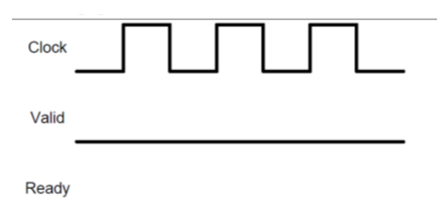


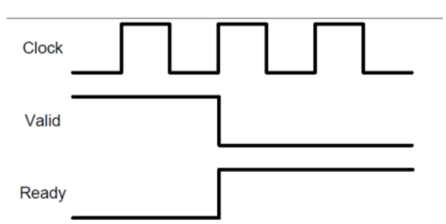
~cs150/fa13/resources/ReadyValidInterface.pdf

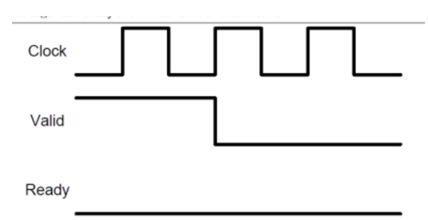
3

Valid/Ready Examples







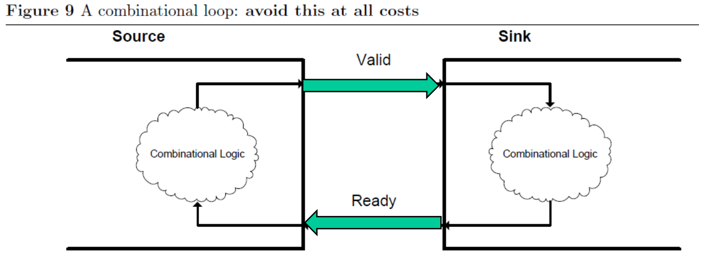


4

Valid/Ready with FIFO

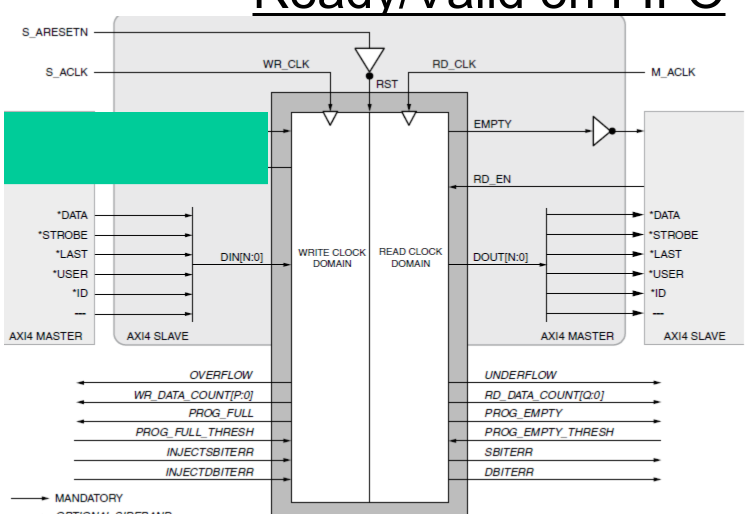
Figure 8 A FIFO functioning as both a source and sink

Figure 9 A combinational loop: avoid this at all costs



5

Ready/Valid on FIFO



not avail for virtex5!

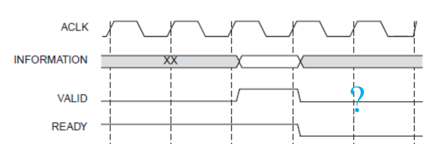


Figure 2-4: **AXI4 FIFO Timing Diagram**

6

Clock Domain Crossing FIFO

WB valid ready

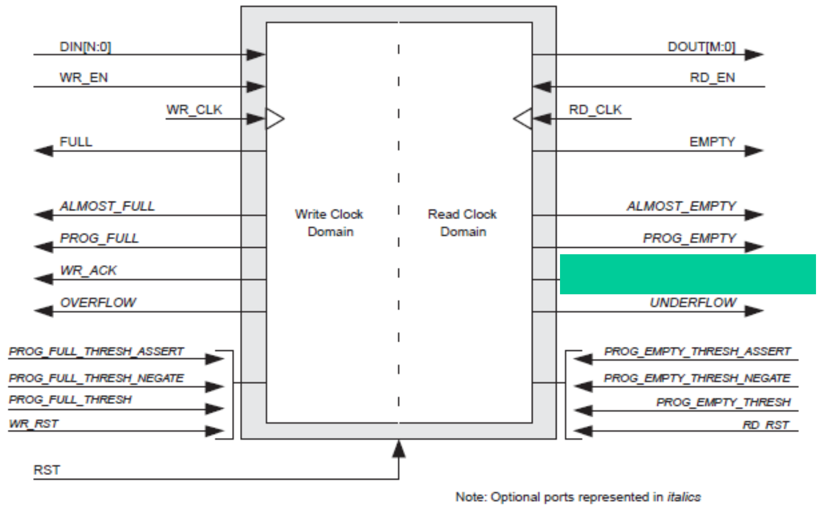


Figure 5-1: **FIFO with Independent Clocks: Write and Read Clock Domains**

7

Internal of FIFO

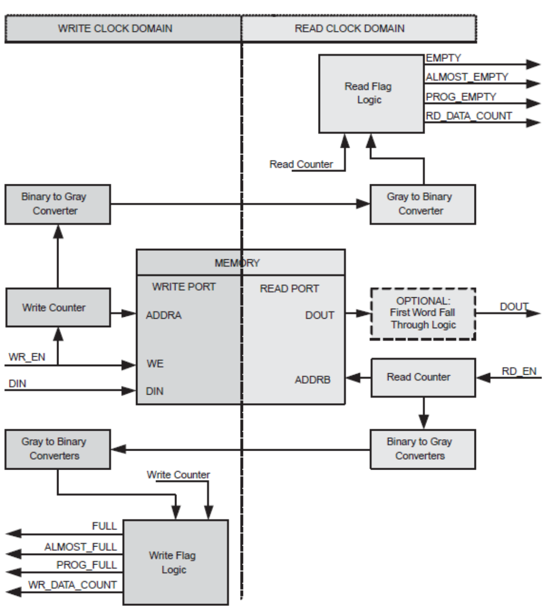
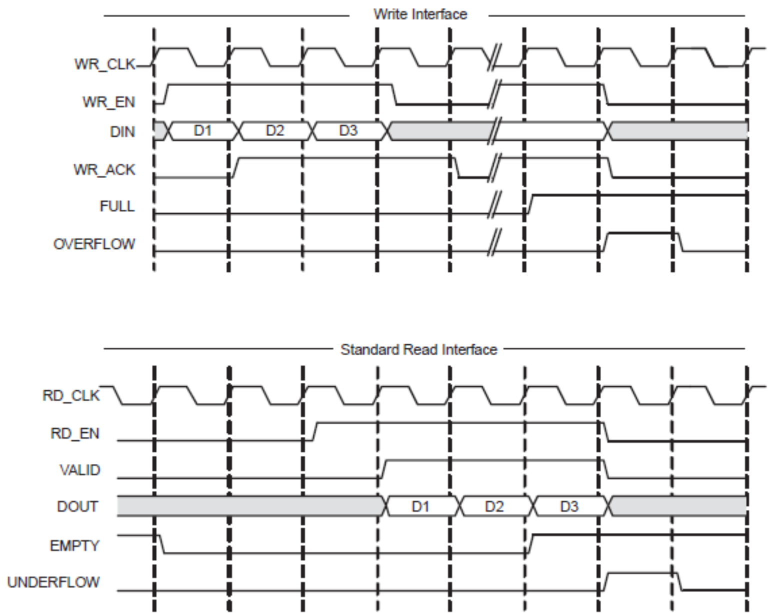


Figure 5-2: **Functional Implementation of a FIFO with Independent Clock Domains**

WB: gray code
000,001,101,100,
110,111,011,010,
000

8

Write and Read with Indep Clocks



9

Read vs Read w/FWFT

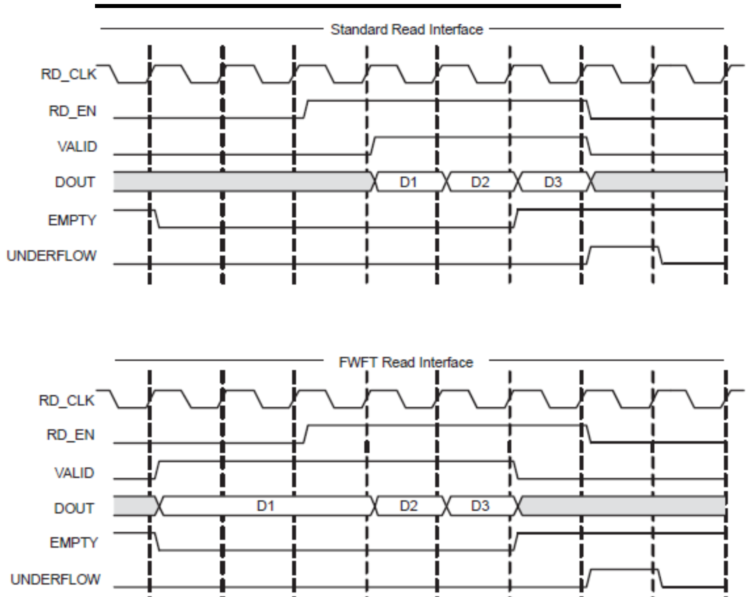


Figure 5-10: **Handshaking Signals for a FIFO with Independent Clocks**

10

The SIFT
(Scale Invariant Feature Transform)
Detector and Descriptor

- developed by David Lowe
- University of British Columbia
- US patent

Lowe, David G. (2004). Distinctive image features from scale-invariant key points. *International Journal of Computer Vision* 60(2): 91-110.

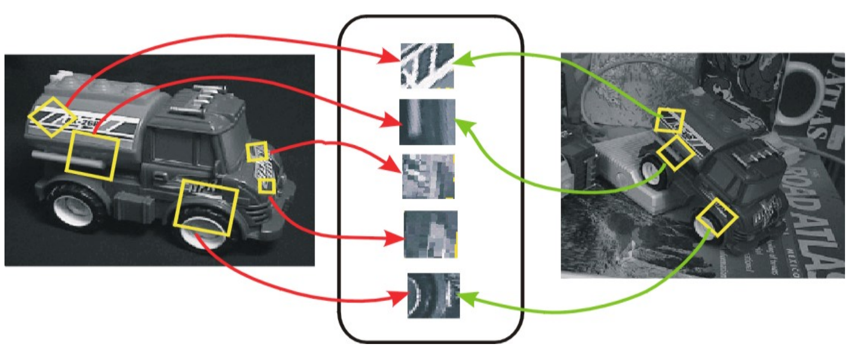
courses.cs.washington.edu/courses/cse576/11sp/.../SIFT_white2011.ppt

http://demo.ipol.im/demo/82/wait?key=ECE94E2AEE6F0D1CCD5265DB4E69D224&show=antmy_detect&action=cust_sift_matching

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Idea of SIFT

- Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



SIFT Features

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Claimed Advantages of SIFT

- **Locality:** features are local, so robust to occlusion and clutter (no prior segmentation)
- **Distinctiveness:** individual features can be matched to a large database of objects
- **Quantity:** many features can be generated for even small objects
- **Efficiency:** close to real-time performance
- **Extensibility:** can easily be extended to wide range of differing feature types, with each adding robustness

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington 13

Overall Procedure at a High Level

1. Scale-space extrema detection

   Search over multiple scales and image locations. HW

   SW

2. Keypoint localization

   Fit a model to determine location and scale.
   Select keypoints based on a measure of stability.

3. Orientation assignment

   Compute best orientation(s) for each keypoint region.

4. Keypoint description

   Use local image gradients at selected scale and rotation to describe each keypoint region.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington 14

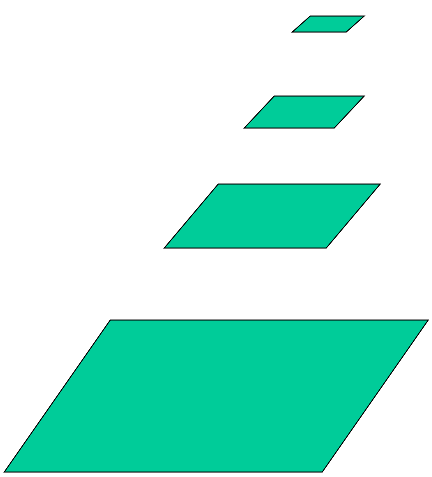
1. Scale-space extrema detection

- Goal: Identify locations and scales that can be repeatably assigned under different views of the same scene or object.
- Method: search for stable features across multiple scales using a continuous function of scale.
- Prior work has shown that under a variety of assumptions, the best function is a Gaussian function.
- The scale space of an image is a function $L(x,y,\sigma)$ that is produced from the convolution of a Gaussian kernel (at different scales) with the input image.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

15

Aside: Image Pyramids



And so on.

3rd level is derived from the 2nd level according to the same function

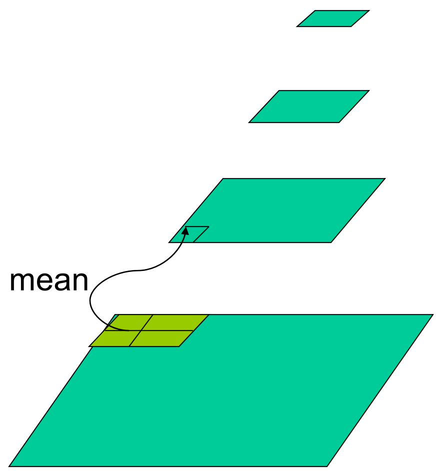
2nd level is derived from the original image according to some function

Bottom level is the original image.

16

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Aside: Mean Pyramid



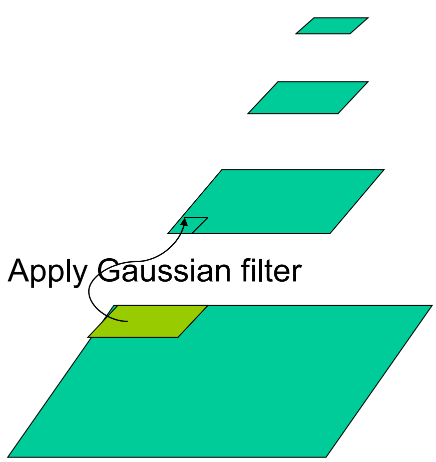
And so on.

At 3rd level, each pixel is the mean of 4 pixels in the 2nd level.

At 2nd level, each pixel is the mean of 4 pixels in the original image.

Bottom level is the original image.

Aside: Gaussian Pyramid
At each level, image is smoothed and reduced in size.

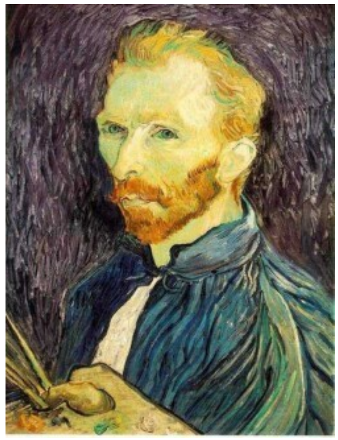


And so on.

At 2nd level, each pixel is the result of applying a Gaussian mask to the first level and then subsampling to reduce the size.

Bottom level is the original image.

Example: Subsampling with Gaussian pre-filtering



Gaussian 1/2

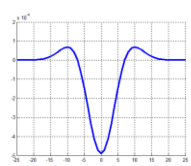
G 1/4

G 1/8

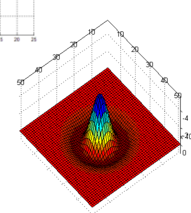
Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

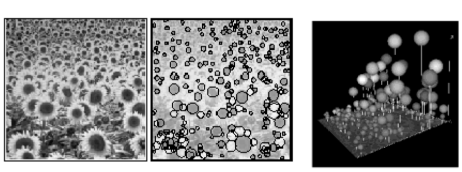
19

Lowe's Scale-space Interest Points



- **Laplacian of Gaussian** kernel
  - Scale normalised (x by scale^2)
  - Proposed by Lindeberg
- Scale-space detection
  - Find local maxima across scale/space
  - A good "blob" detector





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

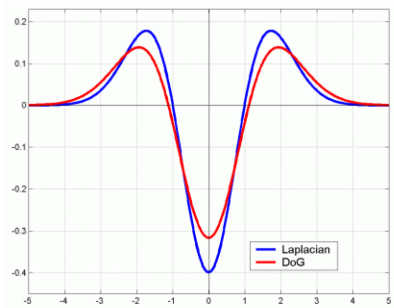
$$\nabla^2 G(x, y, \sigma) = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2}$$

[T. Lindeberg IJCV 1998]

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

20

Lowe's Scale-space Interest Points: Difference of Gaussians



(Difference of Gaussians)

- Gaussian is an ad hoc solution of heat diffusion equation

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

- Hence

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

- k is not necessarily very small in practice, e.g. $2^{1/3}$

$D(x,y,\sigma)=$ $DoG = G(x, y, k\sigma) - G(x, y, \sigma)$

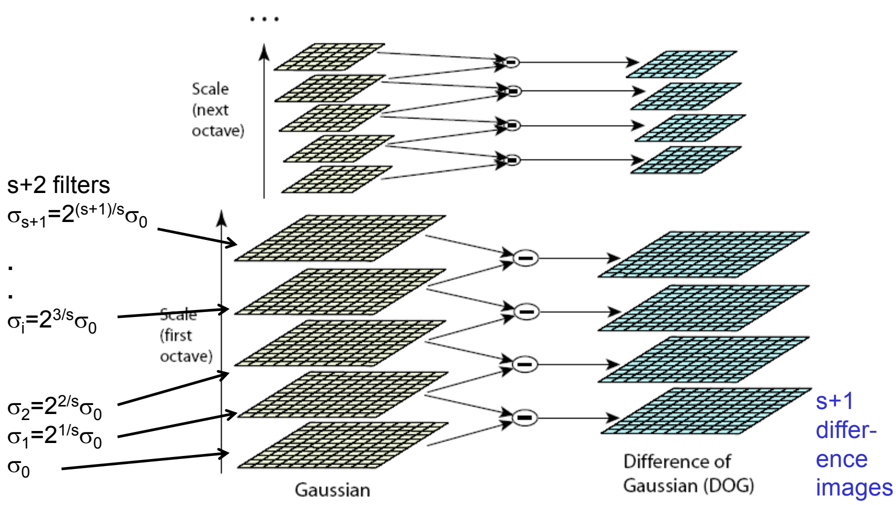
Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Lowe's Pyramid Scheme

- Scale space is separated into octaves:
  - Octave 1 uses scale σ
  - Octave 2 uses scale 2σ
  - etc.
- In each octave, the initial image is repeatedly convolved with Gaussians to produce a set of scale space images.
- Adjacent Gaussians are subtracted to produce the DoG
- After each octave, the Gaussian image is down-sampled by a factor of 2 to produce an image ¼ the size to start the next level.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Lowe's Pyramid Scheme



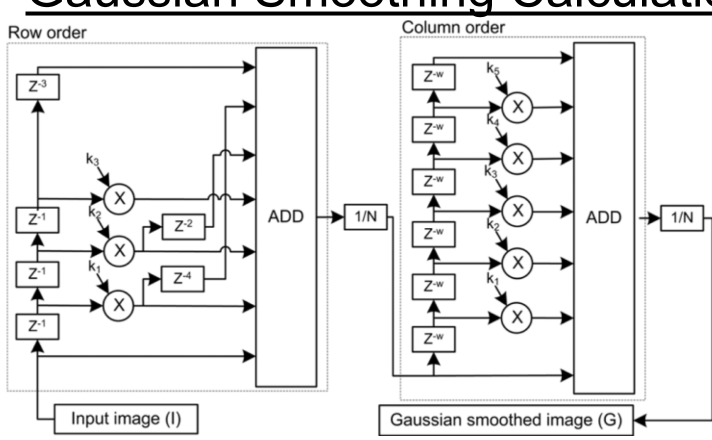
s+3 images including original   Lowe, Fig.1

The parameter **s** determines the number of images per octave.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

23

Gaussian Smoothing Calculation

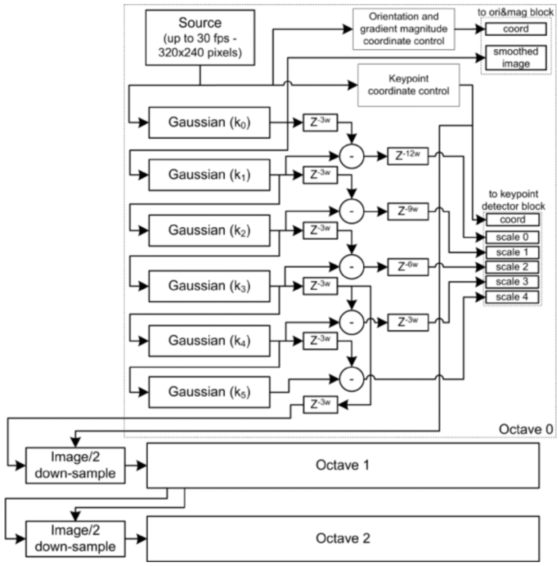


separable assumption WB

Example: ``A Parallel Hardware Architecture for Scale and Rotation Invariant Feature Detection," Bonato, et al., IEEE TRANS. ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 18, NO. 12, DECEMBER 2008

24

Gaussian Smoothing



Example: ``A Parallel Hardware Architecture for Scale and Rotation Invariant Feature Detection,'' Bonato, et al., IEEE Trans. on Circuits and Systems for Video Tech., vol. 18, no. 12, Dec. 2008.

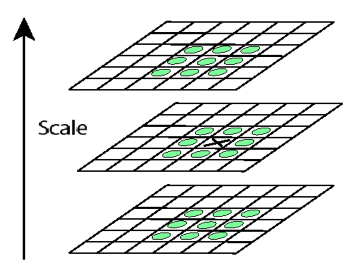
2. Key point localization

s+2 difference images.
top and bottom ignored.
s planes searched.

- Lowe, Fig.2

- Detect maxima and minima of difference-of-Gaussian in scale space

- Each point is compared to its 8 neighbors in the current image and 9 neighbors each in the scales above and below



For each max or min found, output is the **location** and the **scale**.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Keypoint Detection

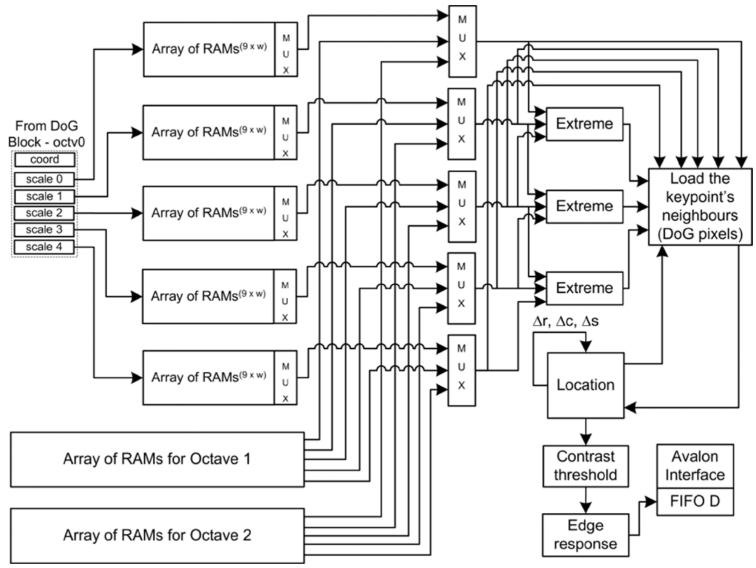


Fig. 8. : ``A Parallel Hardware Architecture for Scale and Rotation Invariant Feature Detection,'' Bonato, et al., IEEE Trans. on Circuits and Systems for Video Tech., vol. 18, no. 12, Dec. 2008.

27

Overall Procedure at a High Level

1. Scale-space extrema detection

   Search over multiple scales and image locations. HW

---

2. Keypoint localization SW

   Fit a model to determine location and scale.
   Select keypoints based on a measure of stability.

3. Orientation assignment

   Compute best orientation(s) for each keypoint region.

4. Keypoint description

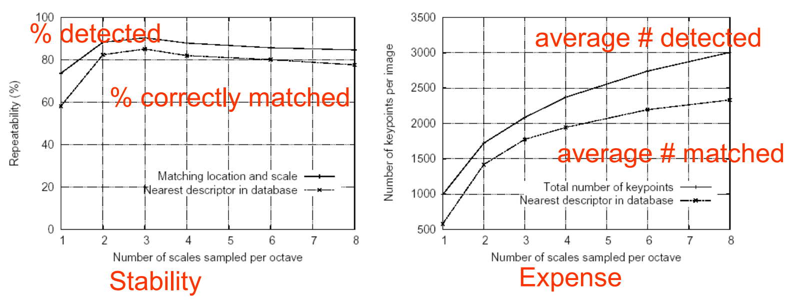
   Use local image gradients at selected scale and rotation to describe each keypoint region.

28

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

got here

Scale-space extrema detection: experimental results over 32 images that were synthetically transformed and noise added.



- Sampling in scale for efficiency
  - How many scales should be used per octave? S=?
    - More scales evaluated, more keypoints found
    - S < 3, stable keypoints increased too
    - S > 3, stable keypoints decreased
    - S = 3, maximum stable keypoints found

Lowe, Fig.3

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

Keypoint localization

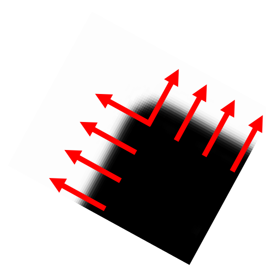
- Once a keypoint candidate is found, perform a detailed fit to nearby data to determine
  - location, scale, and ratio of principal curvatures
- In initial work keypoints were found at location and scale of a central sample point.
- In newer work, they fit a 3D quadratic function to improve interpolation accuracy.
- The Hessian matrix was used to eliminate edge responses.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

31

Corners as distinctive interest points

Since M is symmetric, we have $$M = X \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} X^T$$

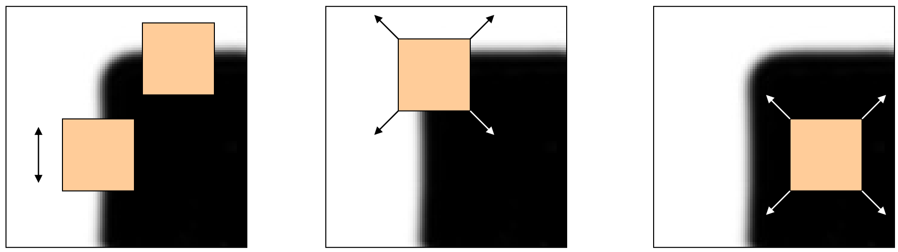


$$Mx_i = \lambda_i x_i$$

The *eigenvalues* of M reveal the amount of intensity change in the two principal orthogonal gradient directions in the window.

slide credit:
CS 143, Brown Univ
James Hays, 2011

Corners as distinctive interest points



"edge":
$\lambda_1 >> \lambda_2$
$\lambda_2 >> \lambda_1$

"corner":
λ_1 and λ_2 are large,
$\lambda_1 \sim \lambda_2$;

"flat" region
λ_1 and λ_2 are small;

One way to score the cornerness:

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

slide credit:
CS 143, Brown Univ
James Hays, 2011

Eliminating the Edge Response

see paper for details if interested

- Reject flats:
  - $|D(\hat{\mathbf{x}})|$ < 0.03
- Reject edges:

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

Let α be the eigenvalue with larger magnitude and β the smaller.

$$\mathrm{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta,$$
$$\mathrm{Det}(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta.$$

Let r = α/β.
So α = rβ

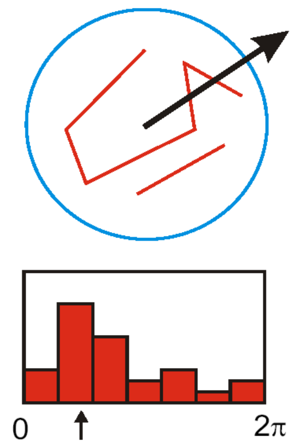
$$\frac{\mathrm{Tr}(\mathbf{H})^2}{\mathrm{Det}(\mathbf{H})} = \frac{(\alpha+\beta)^2}{\alpha\beta} = \frac{(r\beta+\beta)^2}{r\beta^2} = \frac{(r+1)^2}{r},$$

$(r+1)^2/r$ is at a min when the 2 eigenvalues are equal.

- r < 10

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington 34

3. Orientation assignment



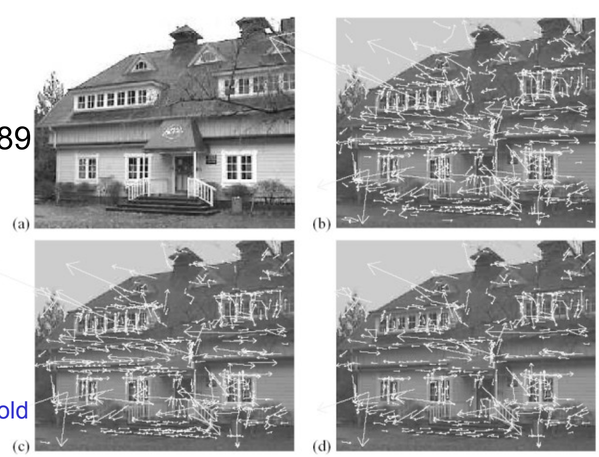
- Create histogram of local gradient directions at selected scale
- Assign canonical orientation at peak of smoothed histogram
- Each key specifies stable 2D coordinates (x, y, scale,orientation)

If 2 major orientations, use both.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

35

Keypoint localization with orientation



233x189

832
initial keypoints

729
keypoints after gradient threshold

536
keypoints after ratio threshold

Lowe, Fig.5

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

36

4. Keypoint Descriptors

- At this point, each keypoint has
  - location
  - scale
  - orientation
- Next is to compute a descriptor for the local image region about each keypoint that is
  - highly distinctive
  - as invariant as possible to variations such as changes in viewpoint and illumination

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

37

Normalization

- Rotate the window to standard orientation

- Scale the window size based on the scale at which the point was found.

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

38

Lowe's Keypoint Descriptor (shown with 2 X 2 descriptors over 8 X 8)

gradient magnitude and orientation at each point weighted by a Gaussian

orientation histograms: sum of gradient magnitude at each direction

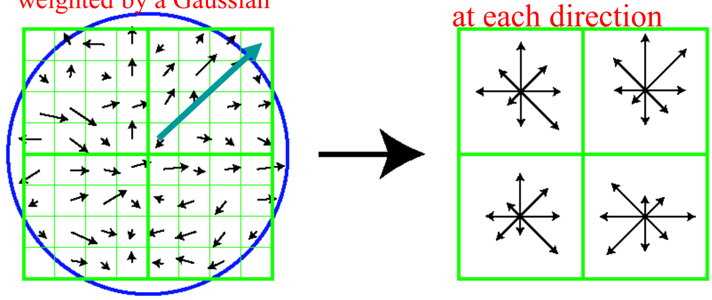


Image gradients

Keypoint descriptor

In experiments, 4x4 arrays of 8 bin histogram is used, a total of 128 features for one keypoint

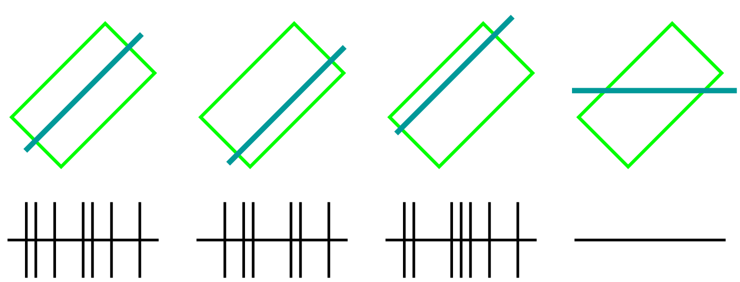
WB

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

39

Biological Motivation

- Mimic complex cells in primary visual cortex
- Hubel & Wiesel found that cells are sensitive to *orientation* of edges, but insensitive to their *position*
- This justifies spatial pooling of edge responses



["Eye, Brain and Vision" – Hubel and Wiesel 1988]

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

40

Lowe's Keypoint Descriptor

- use the normalized region about the keypoint
- compute gradient magnitude and orientation at each point in the region
- weight them by a Gaussian window overlaid on the circle
- create an orientation histogram over the 4 X 4 subregions of the window
- 4 X 4 descriptors over 16 X 16 sample array were used in practice. 4 X 4 times 8 directions gives a vector of **128 values**.

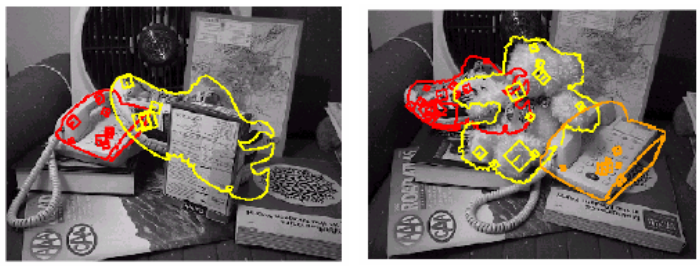
Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

41

Using SIFT for Matching "Objects"

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

42



Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

43

Uses for SIFT

- Feature points are used also for:
  - Image alignment (homography, fundamental matrix)
  - 3D reconstruction (e.g. Photo Tourism)
  - Motion tracking
  - Object recognition
  - Indexing and database retrieval
  - Robot navigation
  - … many others

[Photo Tourism: Snavely et al. SIGGRAPH 2006]

Slides courtesy of Prof. Linda Shapiro, Dept. of CSE, U. Washington

44

Conclusions

•Ready/Valid Hand Shaking
•SIFT details