CAP 5415 Computer Vision
Fall 2005

Dr. Alper Yilmaz

Univ. of Central Florida

www.cs.ucf.edu/courses/cap5415/fall2005

Office: CSB 250

Histogram Segmentation Code

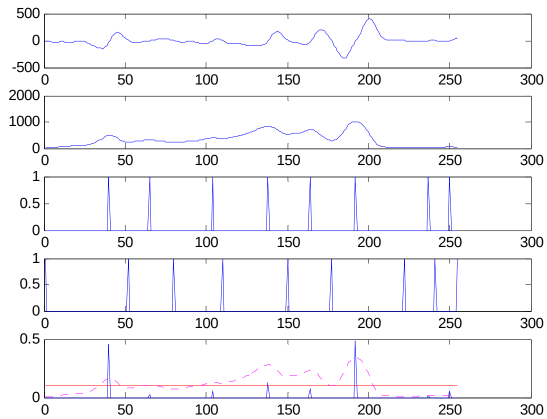
```
%smooth
for i=1:5
    hst=conv2(hst,gauss,'same');
end
%compute derivative
dr1=conv2(hst,dr,'same');
%find peaks and valleys
pw =dr1(1:254).*dr1(2:255);
peaks_valleys(find(pw<0))=1;

control=find(peaks_valleys==1);
for i=1:size(control,2)
    if (dr1(control(i)-2)>0)
        valleys(control(i))=1;
    end
    if (dr1(control(i)-2)<0)
        peaks(control(i))=1;
    end
end

peaklocs=find(peaks==1);
valleylocs=find(valleys==1);
```

```
current_valley=1;
for i=1:size(peaklocs,2)
    Va = hst(valleylocs(current_valley));
    Vb = hst(valleylocs(current_valley+1));
    P  = hst(peaklocs(i));
    W  = valleylocs(current_valley+1)-
     valleylocs(current_valley);
    N  = 0 ;
    for
     j=valleylocs(current_valley):valleylocs
     (current_valley+1)
        N = N + hst(j);
    end
    val1 = 1-((Va+Vb)/(2*P));
    val2 = 1-(N/(W*P));
    if (val1>0 && val1>0)
        peakiness(peaklocs(i)) = val1*val2;
    end;
    current_valley = current_valley +1;
end
```

Plots



Derivative

Histogram

Peaks

Valleys

Peakiness

Alper Yilmaz, Fall 2005 UCF

Recap
Region Properties

- Area
- Centroid
- Moments
- Perimeter
- Compactness
- Orientation

$$A = \sum_{x=0}^{m} \sum_{y=0}^{n} B(x,y)$$

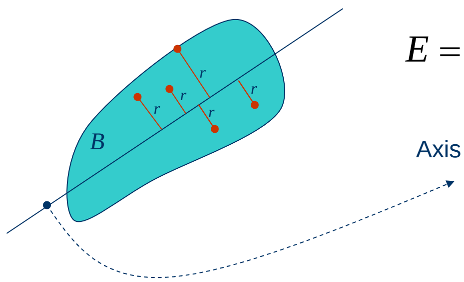
$$\bar{x} = \frac{\sum_{x=0}^{m} \sum_{y=0}^{n} xB(x,y)}{A} \qquad \bar{y} = \frac{\sum_{x=0}^{m} \sum_{y=0}^{n} yB(x,y)}{A}$$

$$\mu_{pq} = \int\int (x-\bar{x})^p (y-\bar{y})^q B(x,y) \; d(x-\bar{x}) d(y-\bar{y})$$

$$C = \frac{Perimeter^2}{4\pi A}$$

Alper Yilmaz, Fall 2005 UCF

Recap
Region Orientation



$$E = \iint r^2 B(x,y)\,dxdy$$

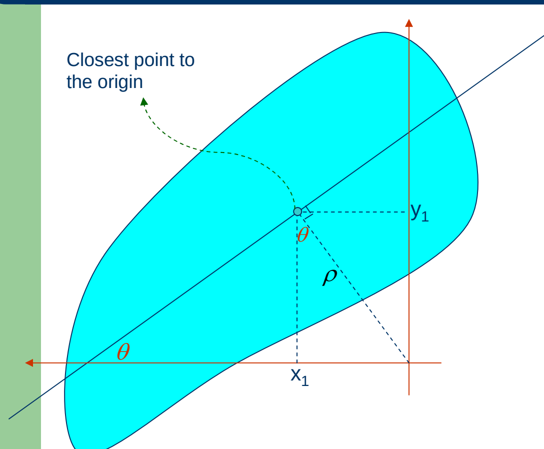
Axis of 2nd moment

$$m_{20}^2 = \sum_{x=0}^{m}\sum_{y=0}^{n} x^2 B(x,y)$$

$$m_{02}^2 = \sum_{x=0}^{m}\sum_{y=0}^{n} y^2 B(x,y)$$

$$m_{11}^2 = \sum_{x=0}^{m}\sum_{y=0}^{n} xy B(x,y)$$

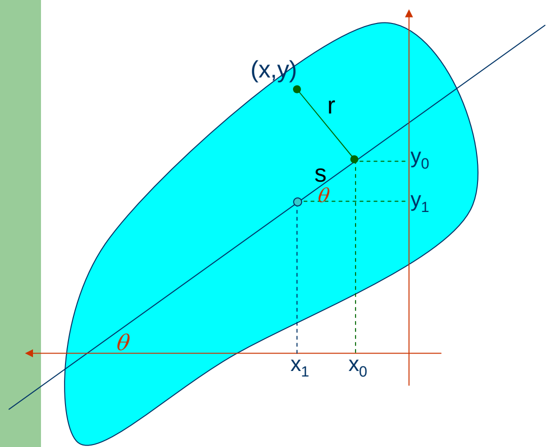
Recap
Region Orientation



$$x_1 = -\rho \sin\theta$$

$$y_1 = \rho \cos\theta$$

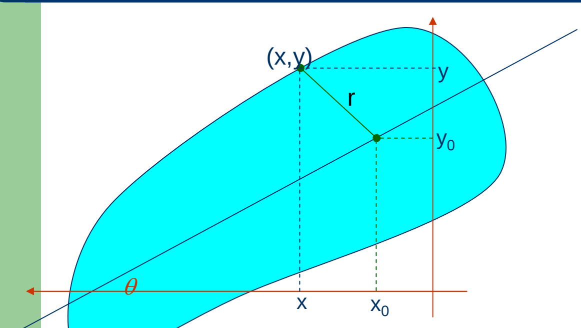
Recap
Region Orientation



$$x_0 = x_1 + s\cos\theta$$
$$y_0 = y_1 + s\sin\theta$$

$$x_0 = (-\rho\sin\theta) + s\cos\theta$$
$$y_0 = (\rho\cos\theta) + s\sin\theta$$

Recap
Region Orientation



$$r^2 = (x - x_0)^2 + (y - y_0)^2$$

$$x_0 = (-\rho\sin\theta) + s\cos\theta$$
$$y_0 = (\rho\cos\theta) + s\sin\theta$$

Substitute x_0 and y_0

$$r^2 = (x + \rho\sin\theta - s\cos\theta)^2 + (y - \rho\cos\theta - s\sin\theta)^2$$

$$r^2 = x^2 + y^2 + \rho^2 + s^2 - 2s(x\cos\theta + y\sin\theta) + 2\rho(x\sin\theta - y\cos\theta)$$

Recap
Region Orientation

$$r^2 = x^2 + y^2 + \rho^2 + s^2 - 2s(x\cos\theta + y\sin\theta) + 2\rho(x\sin\theta - y\cos\theta)$$

$$\frac{\partial}{\partial s} r^2 = 2s - 2(x\cos\theta + y\sin\theta)$$

$$\boxed{s = x\cos\theta + y\sin\theta}$$

Substitute s back to r

Recap
Region Orientation

$$r^2 = x^2 + y^2 + \rho^2 + s^2 - 2s(x\cos\theta + y\sin\theta) + 2\rho(x\sin\theta - y\cos\theta)$$

$$r^2 = x^2 + y^2 + \rho^2 - (x\cos\theta + y\sin\theta)^2 + 2\rho(x\sin\theta - y\cos\theta)$$
$$r^2 = x^2 + y^2 + \rho^2 - x^2\cos^2\theta - 2xy\sin\theta\cos\theta - y^2\sin^2\theta + 2\rho(x\sin\theta - y\cos\theta)$$
$$r^2 = x^2(1 - \cos^2\theta) + y^2(1 - \sin^2\theta) - 2xy\sin\theta\cos\theta + 2\rho(x\sin\theta - y\cos\theta) + \rho^2$$
$$r^2 = x^2\sin^2\theta + y^2\cos^2\theta - 2xy\sin\theta\cos\theta + 2\rho(x\sin\theta + y\cos\theta) + \rho^2$$
$$r^2 = (x\sin\theta - y\cos\theta)^2 + 2\rho(x\sin\theta - y\cos\theta) + \rho^2$$
$$r^2 = (x\sin\theta - y\cos\theta + \rho)^2$$

Recap
Region Orientation

$$a = x\sin\theta - y\cos\theta$$

$$\frac{\partial}{\partial\rho}E = \frac{\partial}{\partial\rho}\left(\iint (x\sin\theta - y\cos\theta + \rho)^2 B(x,y)dxdy\right)$$

$$\frac{\partial}{\partial\rho}r^2 = \frac{\partial}{\partial\rho}\iint \left(a^2 + 2a\rho + \rho^2\right)B(x,y)dxdy$$

$$\iint 2aB(x,y)dxdy + \iint 2\rho B(x,y)dxdy = 0$$

$$\sin\theta\iint xB(x,y)dxdy - \cos\theta\iint yB(x,y)dxdy + \rho A = 0$$

$$\rho = -\sin\theta\frac{\iint xB(x,y)dxdy}{A} + \cos\theta\frac{\iint yB(x,y)dxdy}{A}$$

$$\boxed{\rho = -\bar{x}\sin\theta + \bar{y}\cos\theta}$$

Alper Yilmaz, Fall 2005 UCF

Recap
Region Orientation

Substitute ρ back to r

$$E = \iint \left((x-\bar{x})\sin\theta - (y-\bar{y})\cos\theta\right)^2 B(x,y)dxdy$$

$$x' = x - \bar{x}, \quad y' = y - \bar{y}$$

$$E = \iint (x'\sin\theta - y'\cos\theta)^2 B(x,y)dxdy$$

$$E = \sin^2\theta\underbrace{\iint x'^2 B(x,y)dxdy}_{a} - 2\sin\theta\cos\theta\underbrace{\iint x'y' B(x,y)dxdy}_{b} + \cos^2\theta\underbrace{\iint y'^2 B(x,y)dxdy}_{c}$$

$$E = a\sin^2\theta - 2b\sin\theta\cos\theta + c\cos^2\theta$$

take derivative wrt θ and equating it to 0 → $\boxed{\tan 2\theta = \frac{b}{a-c}}$

Alper Yilmaz, Fall 2005 UCF

Homework from last lecture

- Show that following holds $E_1 = E_2$
  where

$$E_1 = a\sin^2\theta - b\sin\theta\cos\theta + c\cos^2\theta$$

$$E_2 = \frac{1}{2}(a+c) - \frac{1}{2}(a-c)\cos 2\theta - \frac{1}{2}b\sin 2\theta$$

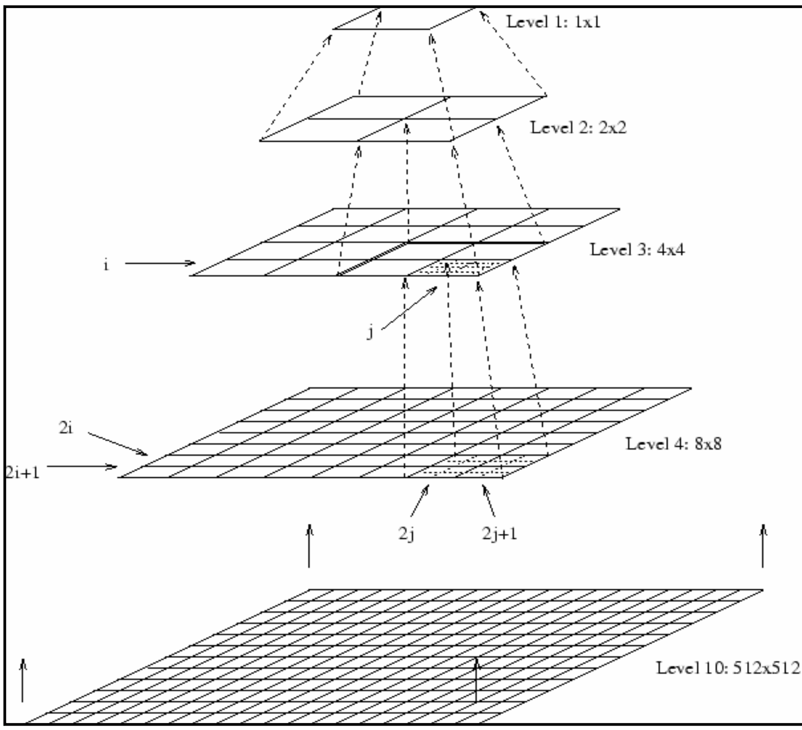
Pyramid Representation

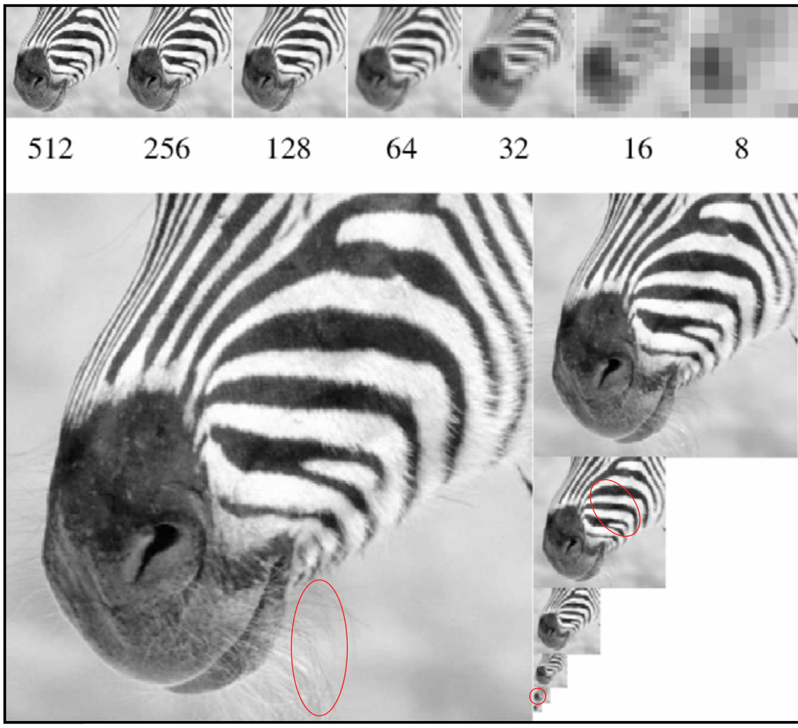
Gaussian
Laplacian

Gaussian Pyramid

- Very useful for representing images
- Image Pyramid is built by using multiple copies of image at different *scales*.
- Each level in the pyramid is ¼ of the size of previous level
- The highest level is of the highest resolution
- The lowest level is of the lowest resolution

Alper Yilmaz, Fall 2005 UCF

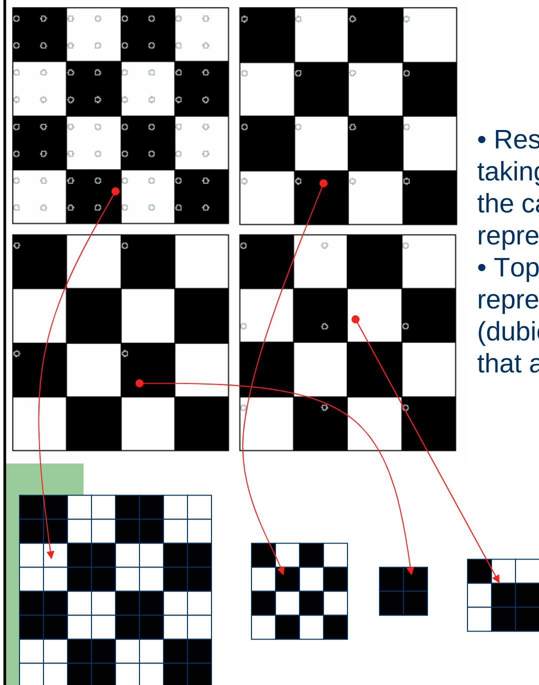




A bar in the big images is a hair on the zebra's nose; in smaller images, a stripe; in the smallest, the animal's nose

Aliasing Problem

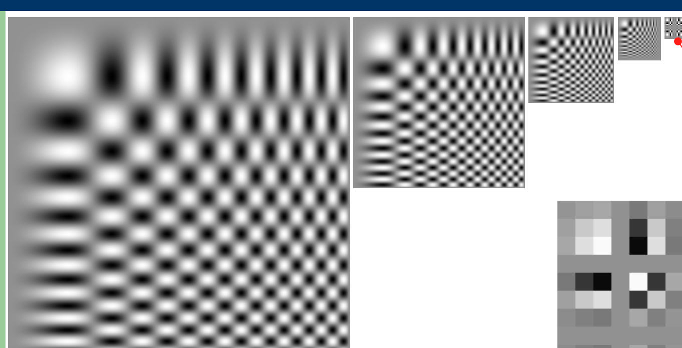
- Can't shrink an image by taking every second pixel
- If we do, characteristic errors appear
  - In the next few slides
  - Typically, small phenomena look bigger; fast phenomena can look slower
  - Common phenomenon
    - Wagon wheels rolling the wrong way in movies
    - Checkerboards misrepresented in ray tracing
    - Striped shirts look funny on color television



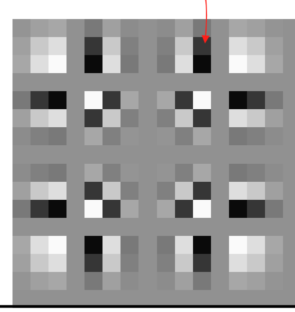
- Resample the checkerboard by taking one sample at each circle. In the case of the top left board, new representation is reasonable.
- Top right also yields a reasonable representation. Bottom left is all black (dubious) and bottom right has checks that are too big.

Alper Yilmaz, Fall 2005 UCF

Pyramid With Every Other Pixel

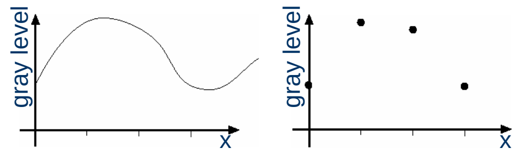


Constructing a pyramid by taking every second pixel leads to layers that badly misrepresent the top layer

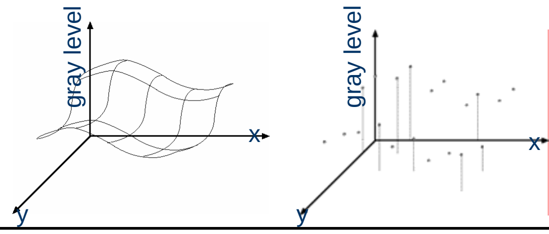


Sampling

- 1D



- 2D



z, Fall 2005 UCF

Smoothing

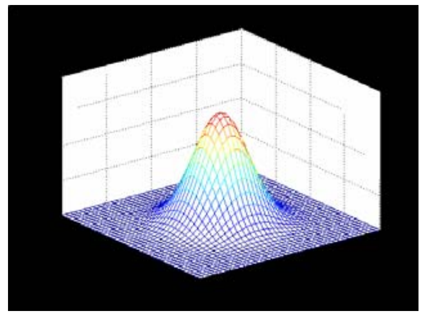
- High frequencies (edges) lead to trouble with sampling.
- Solution: suppress edges before sampling
- Common solution: use a Gaussian
  - Convolve image with Gaussian filter

Alper Yilmaz, Fall 2005 UCF

Gaussian Filter

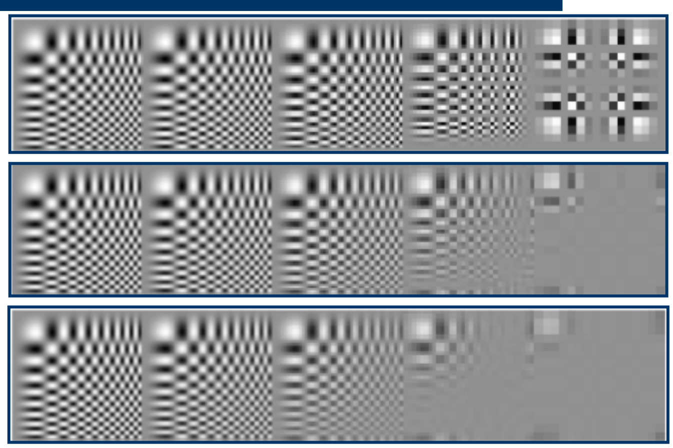
$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right)$$

- gaussian*gaussian=another gaussian
- Symmetric filter
- gaussians are low pass filters
  - removes noise



Alper Yilmaz, Fall 2005 UCF

Scaled images



Alper Yilmaz, Fall 2005 UCF

Applications of Scaled Images

- Search for correspondence
  - look at coarse scales, then refine with finer scales
- Edge tracking
  - "Good" edge at a fine scale has parents at a coarser scale
- Control of detail and computational cost in matching
  - Finding stripes
  - Very important in texture representation

Alper Yilmaz, Fall 2005 UCF

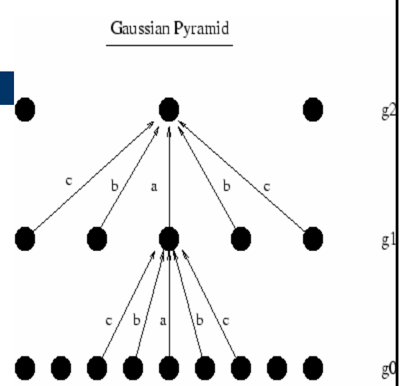
Gaussian Pyramid

- Let w be Gaussian filter

$$g_l(i) = \sum_{m=-2}^{2} \hat{w}(m) g_{l+1}(2i+m)$$

$$g_l(2) = \hat{w}(-2)g_{l-1}(4-2) + \hat{w}(-1)g_{l-1}\hat{w}(4-1) + \hat{w}(0)g_{l-1}(4) + \hat{w}(1)g_{l-1}(4+1) + \hat{w}(2)g_{l-1}(4+2)$$

$$g_l(2) = \hat{w}(-2)g_{l-1}(2) + \hat{w}(-1)g_{l-1}\hat{w}(3) + \hat{w}(0)g_{l-1}(4) + \hat{w}(1)g_{l-1}(5) + \hat{w}(2)g_{l-1}(6)$$



g_0 IMAGE

$g_l =$ REDUCE$[g_{l-1}]$

Convolution Kernel *w*

$$[w(-2), w(-1), w(0), w(1), w(2)]$$

- Symmetric

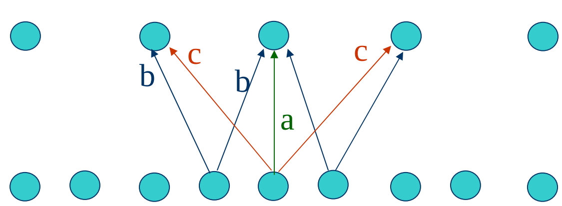
$$w(i) = w(-i) \Rightarrow [c, b, a, b, c]$$

- Sum must be 1

$$a + 2b + 2c = 1$$

Reduce 1D Convolution Kernel *w*

- All nodes at a given level must contribute the same total weight to the nodes at the next higher level



$$a + 2c = 2b$$

Convolution Kernel (5x1)

$$w(0) = a$$

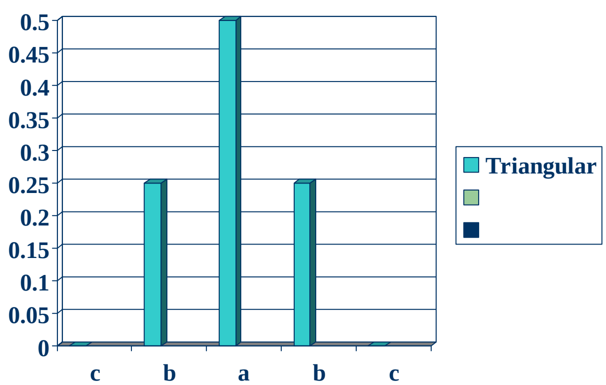
$$w(-1) = w(1) = \frac{1}{4}$$

$$w(-2) = w(2) = \frac{1}{4} - \frac{a}{2}$$

a=0.4 GAUSSIAN
a=0.5 TRIANGULAR

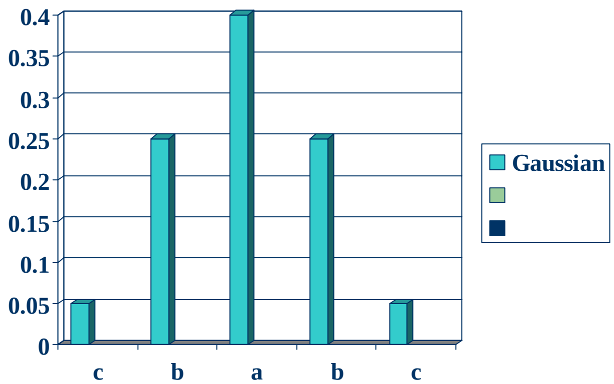
Alper Yilmaz, Fall 2005 UCF

Triangular (5x1)



Alper Yilmaz, Fall 2005 UCF

Approximate Gaussian (5x1)



Alper Yilmaz, Fall 2005 UCF

Two Dimensions?

- Gaussian is separable

$$\hat{I}(x, y) = I(x, y) * G(x, y)$$

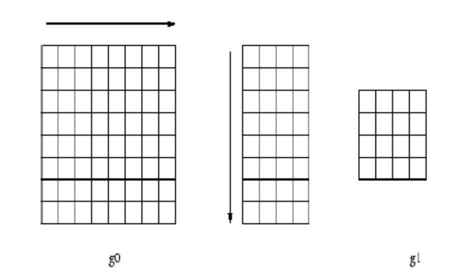
$$\hat{I}(x, y) = I(x, y) * G(x) * G(y)$$

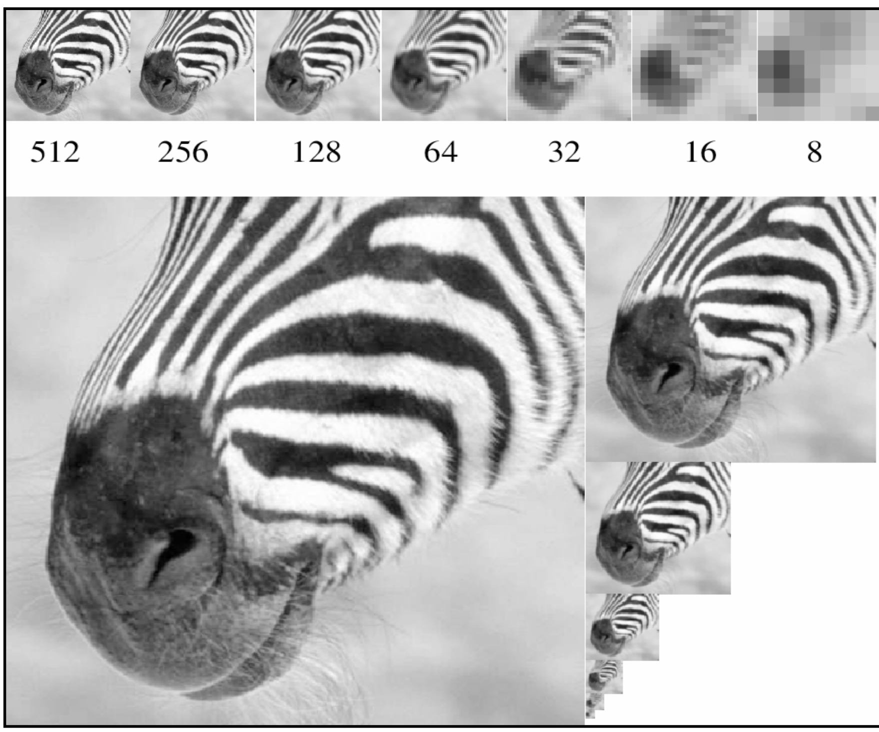
$G(x) = G^T(y)$ transpose

Alper Yilmaz, Fall 2005 UCF

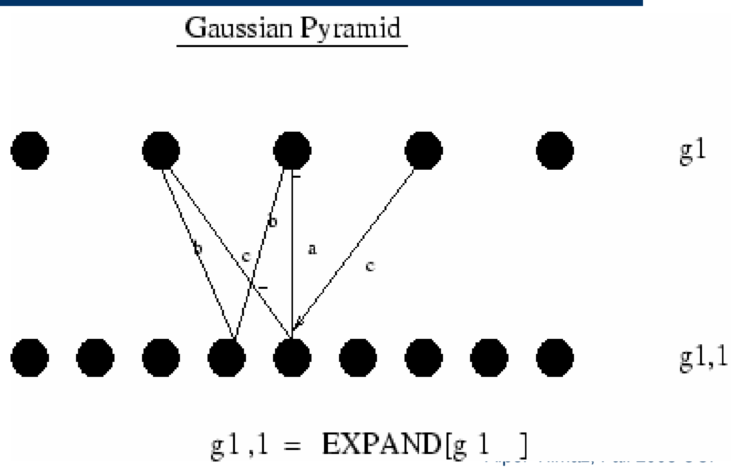
Gaussian Pyramid Algorithm

- Apply 1D mask to alternate pixels along each row of image.
- Apply 1D mask to alternate pixels along each column of resulting image from previous step.





Expand 1D



Expand 1D

$$g_{l,n}(i) = \sum_{p=-2}^{2} \hat{w}(p) g_{l,n-1}\left(\frac{i-p}{2}\right)$$

$$g_{l,n}(4) = \hat{w}(-2) g_{l,n-1}\left(\frac{4-2}{2}\right) + \hat{w}(-1) g_{l,n-1}\left(\frac{4-1}{2}\right) + \hat{w}(0) g_{l,n-1}\left(\frac{4}{2}\right) + \hat{w}(1) g_{l,n-1}\left(\frac{4+1}{2}\right) + \hat{w}(2) g_{l,n-1}\left(\frac{4+2}{2}\right)$$

$$g_{l,n}(4) = \hat{w}(-2) g_{l,n-1}(1) + \hat{w}(0) g_{l,n-1}(2) + \hat{w}(2) g_{l,n-1}(3)$$

Expand 1D

$$g_{l,n}(i) = \sum_{m=-2}^{2} \hat{w}(p) g_{l,n-1}(\frac{i-p}{2})$$

$$g_{l,n}(3) = \hat{w}(-2) g_{l,n-1}(\frac{3-2}{2}) + \hat{w}(-1) g_{l,n-1}(\frac{3-1}{2}) + \hat{w}(0) g_{l,n-1}(\frac{3}{2}) + \hat{w}(1) g_{l,n-1}(\frac{3+1}{2}) + \hat{w}(2) g_{l,n-1}(\frac{3+2}{2})$$

$$g_{l,n}(3) = \hat{w}(-1) g_{l,n-1}(1) + \hat{w}(1) g_{l,n-1}(2)$$

The Laplacian Pyramid

- Similar to edge detected images
- Most pixels are zero
- Can be used in image compression

Constructing Laplacian Pyramid

- Compute Gaussian pyramid

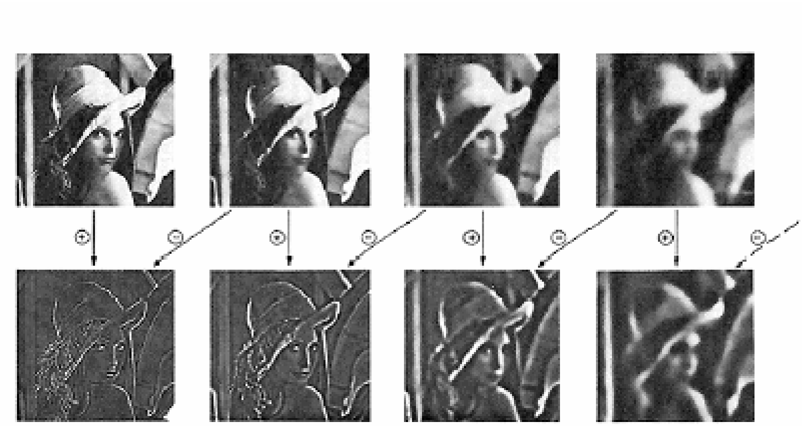
$$g_k, g_{k-1}, g_{k-2}, \ldots g_2, g_1$$

- Compute Laplacian pyramid as follows:

$$
\begin{aligned}
L_k &= g_k - EXPAND(g_{k-1}) \\
L_{k-1} &= g_{k-1} - EXPAND(g_{k-2}) \\
L_{k-2} &= g_{k-2} - EXPAND(g_{k-3}) \\
&\vdots \\
L_1 &= g_1
\end{aligned}
$$

Alper Yilmaz, Fall 2005 UCF

Laplacian Pyramid



Reconstructing Image

$$g_1 = L_1$$
$$g_2 = EXPAND(g_1) + L_2$$
$$g_3 = EXPAND(g_2) + L_3$$
$$\vdots$$
$$g_k = EXPAND(g_{k-1}) + L_k$$

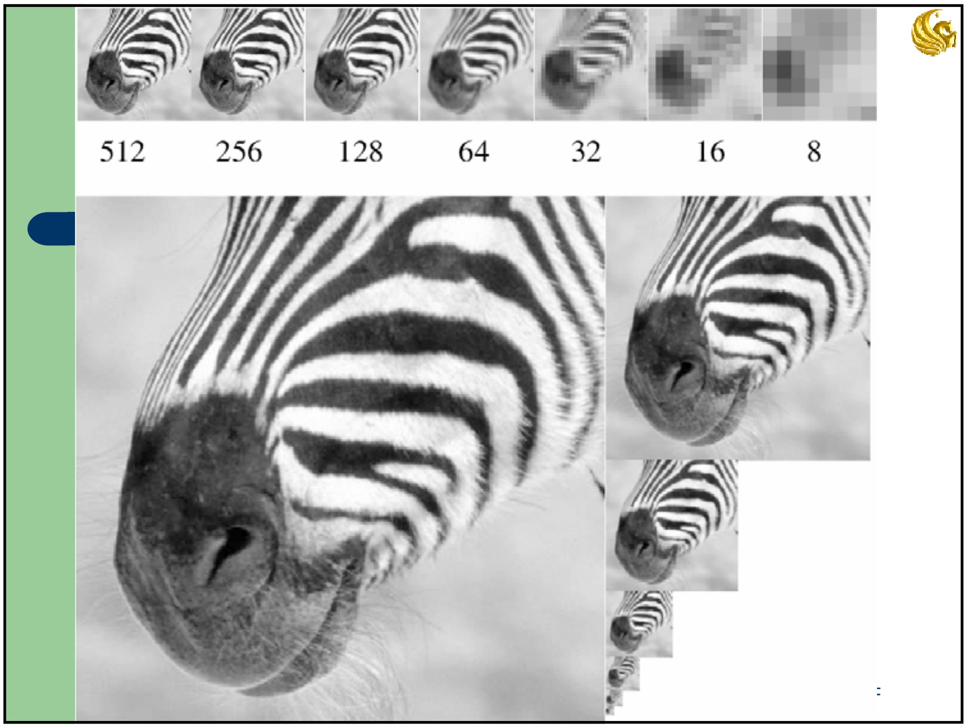
The Laplacian Pyramid

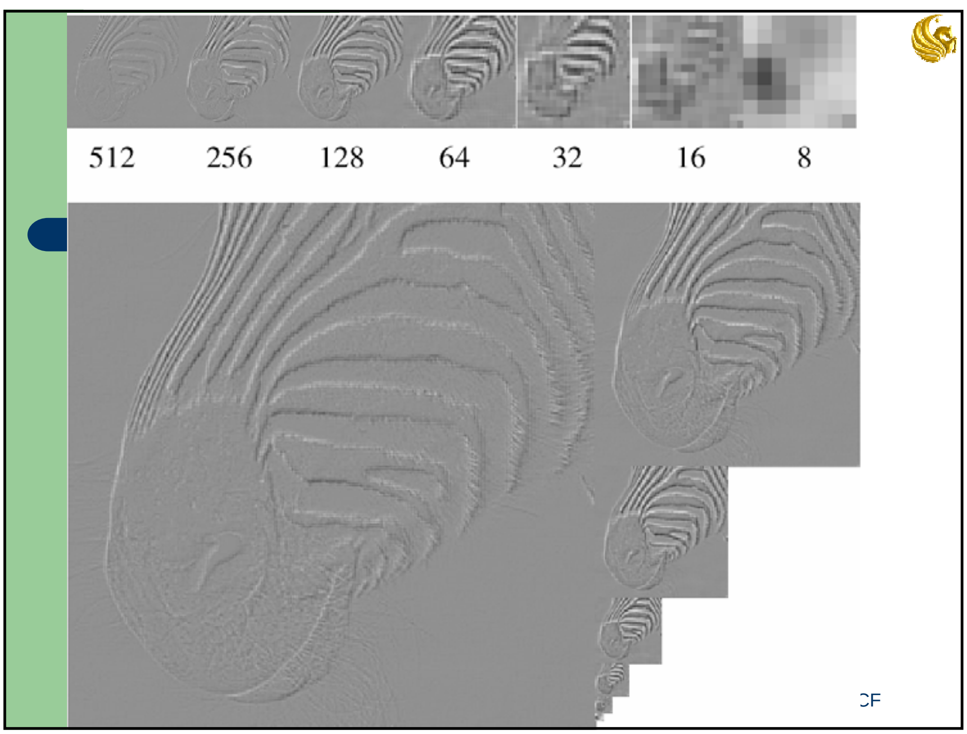
- Synthesis (Coding)
  - Compute Gaussian pyramid
  - Compute Laplacian pyramid

$$L_1 = g_1 - EXPAND[g_2]$$
$$L_2 = g_2 - EXPAND[g_3]$$
$$L_3 = g_3 - EXPAND[g_4]$$
$$L_4 = g_4$$

- Analysis (Decoding)
  - Compute Gaussian pyramid from Laplacian pyramid
  - g_1 is reconstructed image

$$g_4 = L_4$$
$$g_3 = EXPAND[g_4] + L_3$$
$$g_2 = EXPAND[g_3] + L_2$$
$$g_1 = EXPAND[g_2] + L_1$$


512
256
128
64
32
16
8


512
256
128
64
32
16
8

Programming Assignment

- Write the Gaussian Pyramid algorithm
  - The algorithm should be capable of providing any number of resolutions
  - Report should include scaled images of the Lenna image.
  - Due 12 October, 2005

Alper Yilmaz, Fall 2005 UCF