Vision and Perception

Lecture 2: Image pyramids

Sapienza
Università di Roma

References

Basic reading:

- Szeliski textbook, Sections 3.4, 3.5

Additional reading:

The original Laplacian pyramid paper

- Burt and Adelson, “The Laplacian Pyramid as a Compact Image Code,” IEEE ToC 1983.

Overview of today's lecture

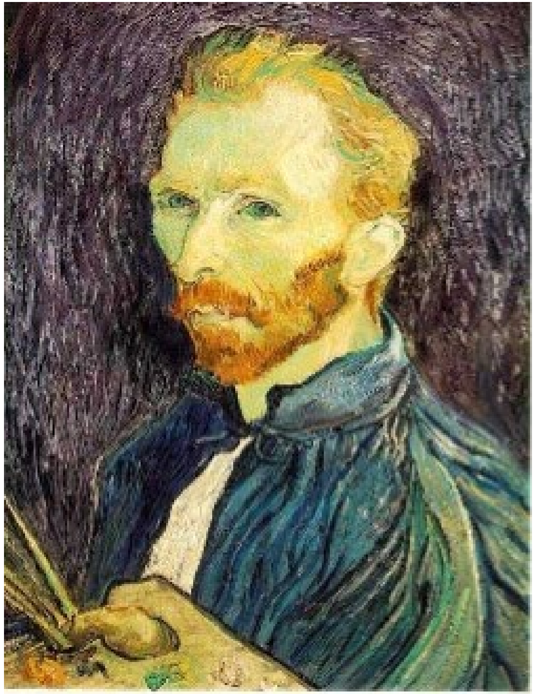
- Image downsampling
- Aliasing
- Gaussian image pyramid
- Laplacian image pyramid

Image downsampling


This image is too big to fit on the screen.
How would you reduce it to half its size?

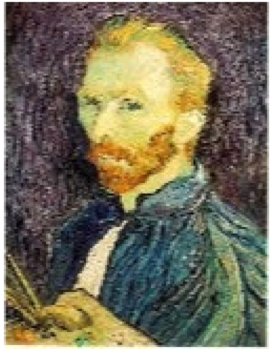
Naïve image downsampling

Throw away half the rows and columns



1/2

delete even rows
delete even columns



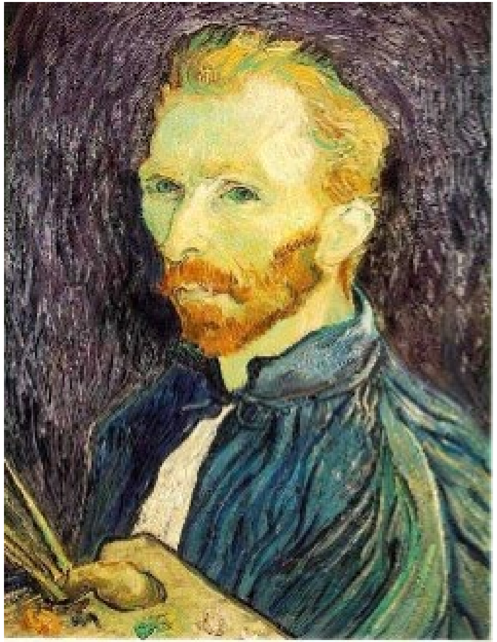
1/4

delete even rows
delete even columns

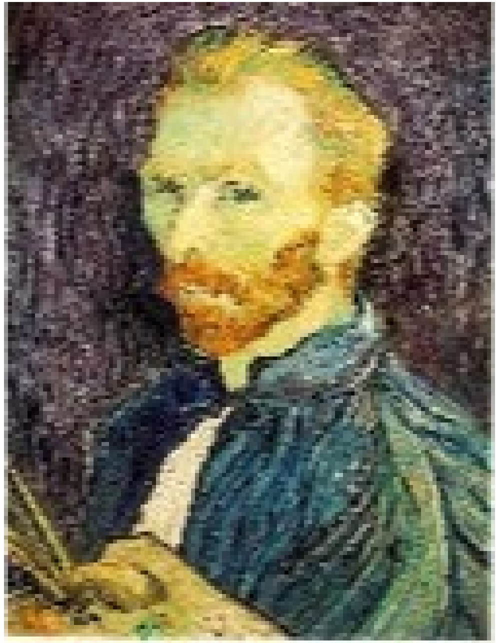
1/8

What is the problem with this approach?

Naïve image downsampling



1/2

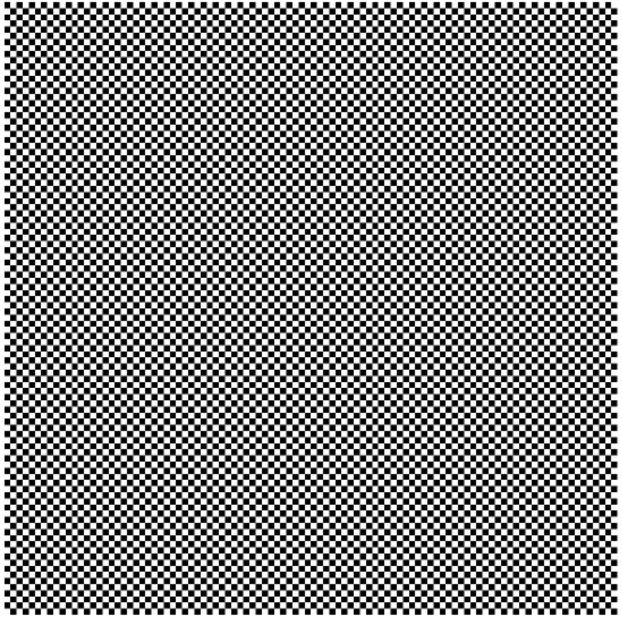


1/4 (2x zoom)

1/8 (4x zoom)

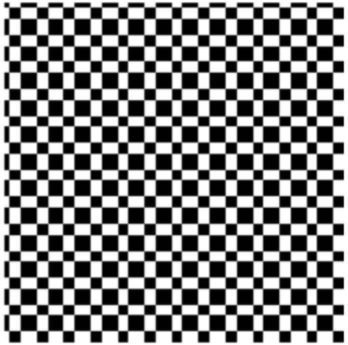
What is the 1/8 image so pixelated (and do you know what this effect is called)?

Spatial undersampling

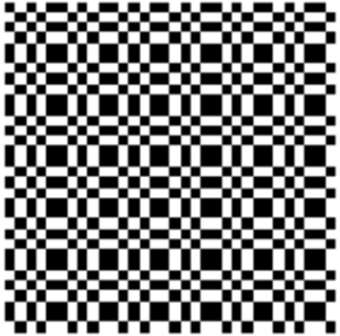


Checkerboard with 10 x 10 pixel squares

Downsampled images



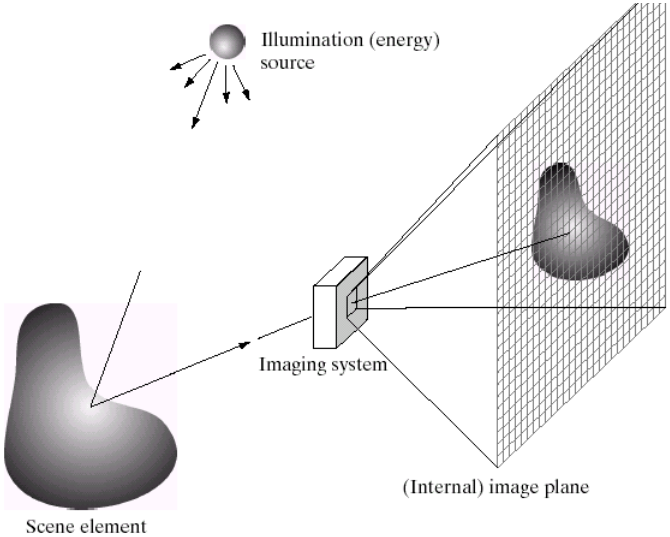
1/10



1/16

Aliasing

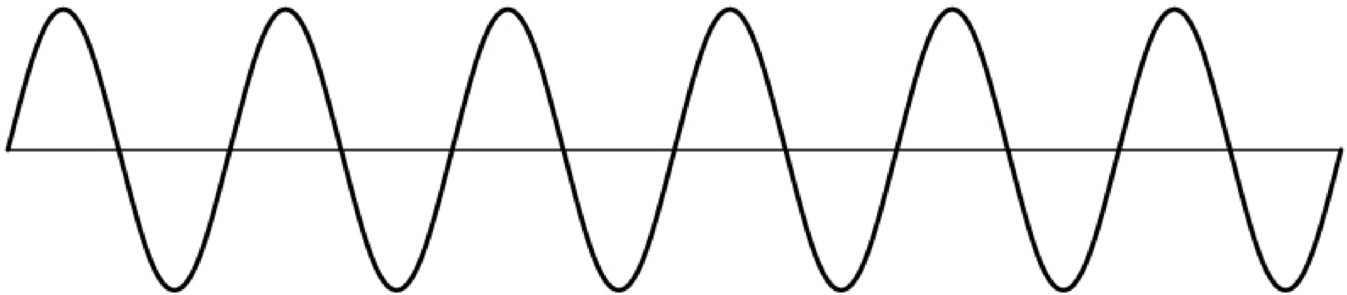
Reminder



Images are a *discrete*, or *sampled*, representation of a *continuous* world

Sampling

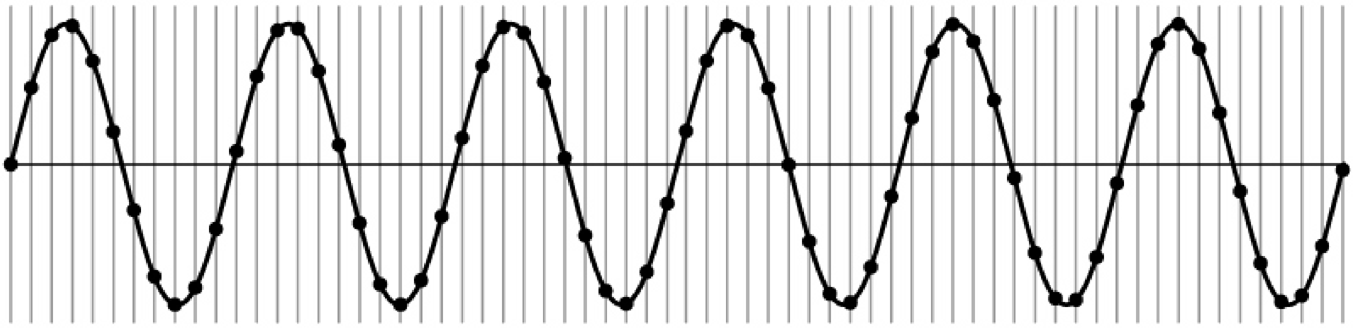
Very simple example: a sine wave



How would you discretize this signal?

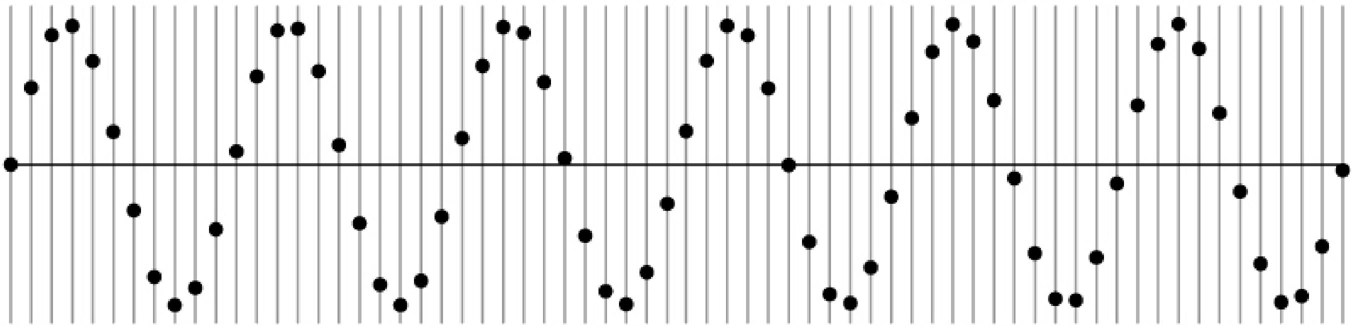
Sampling

Very simple example: a sine wave



Sampling

Very simple example: a sine wave

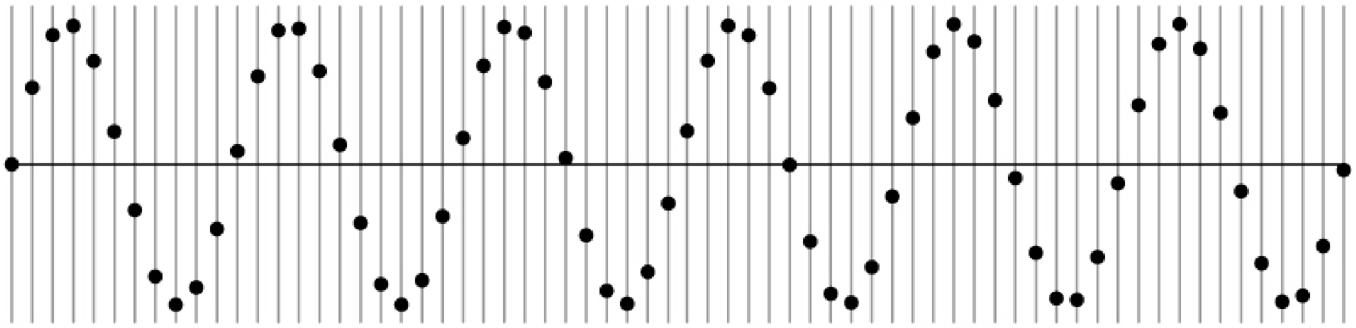


How many samples should I take?

Can I take as *many* samples as I want?

Sampling

Very simple example: a sine wave

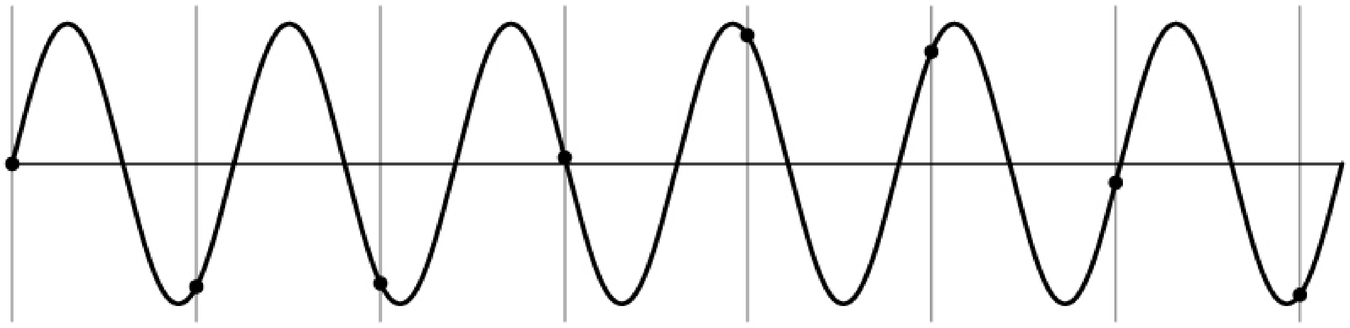


How many samples should I take?

Can I take as *few* samples as I want?

Undersampling

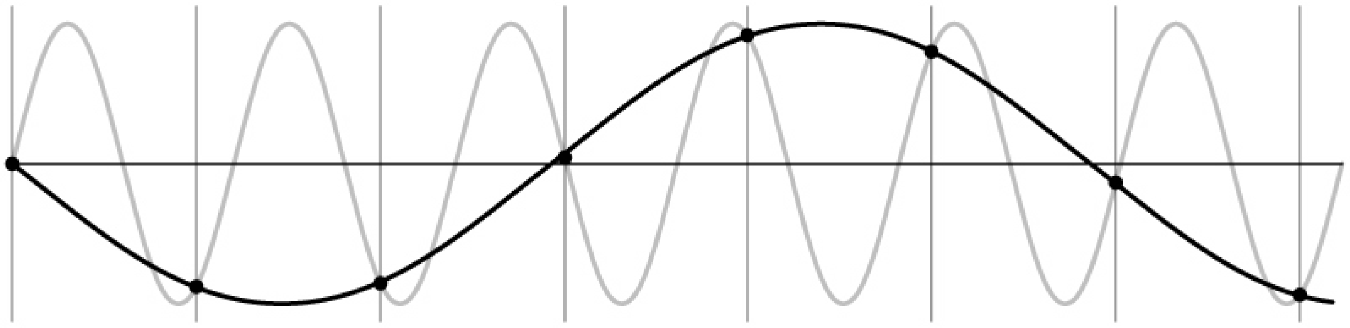
Very simple example: a sine wave



Unsurprising effect: information is lost.

Undersampling

Very simple example: a sine wave

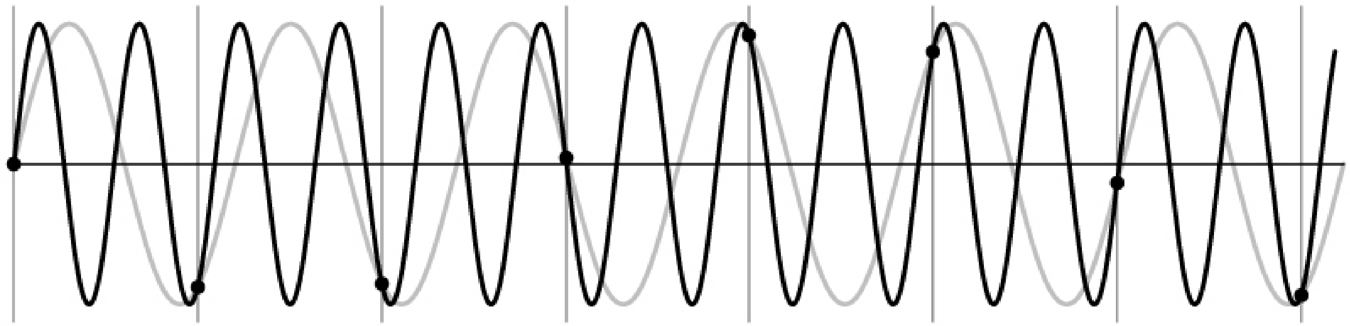


Unsurprising effect: information is lost.

Surprising effect: can confuse the signal with one of *lower* frequency.

Undersampling

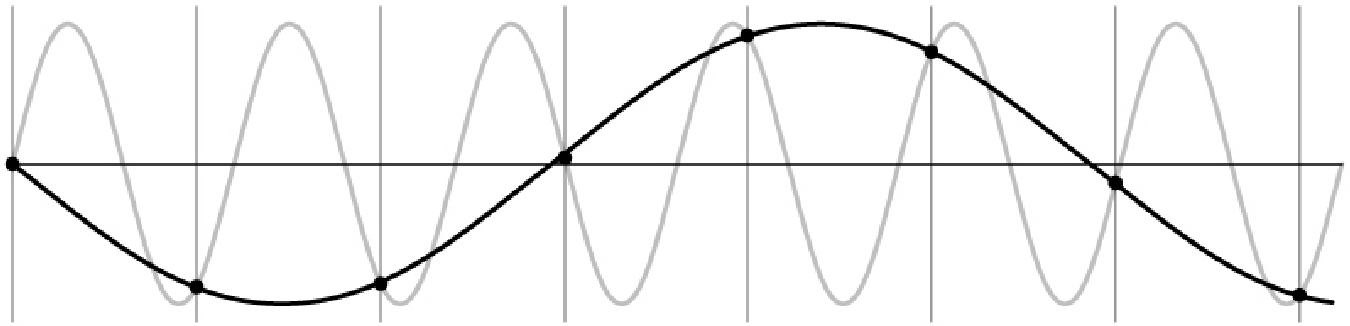
Very simple example: a sine wave



Unsurprising effect: information is lost.
Surprising effect: can confuse the signal with one of *lower* frequency.
Note: we could always confuse the signal with one of *higher* frequency.

Aliasing

Term for: *Undersampling can disguise a signal as one of a lower frequency*

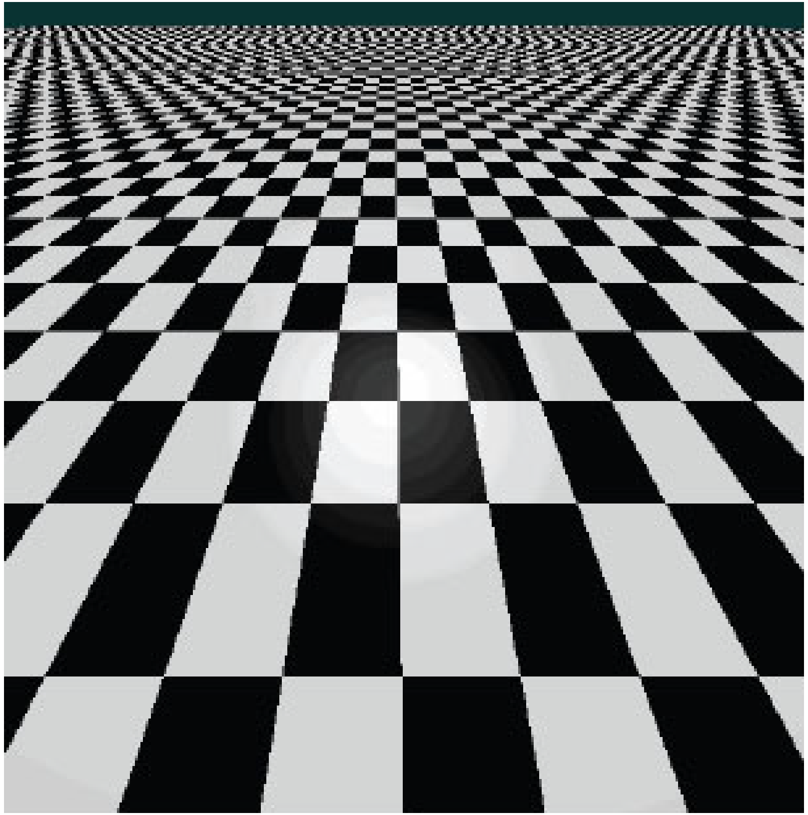


Unsurprising effect: information is lost.

Surprising effect: can confuse the signal with one of *lower* frequency.

Note: we could always confuse the signal with one of *higher* frequency.

Aliasing in textures



Aliasing in photographs

This is also known as “moire”

Anti-aliasing

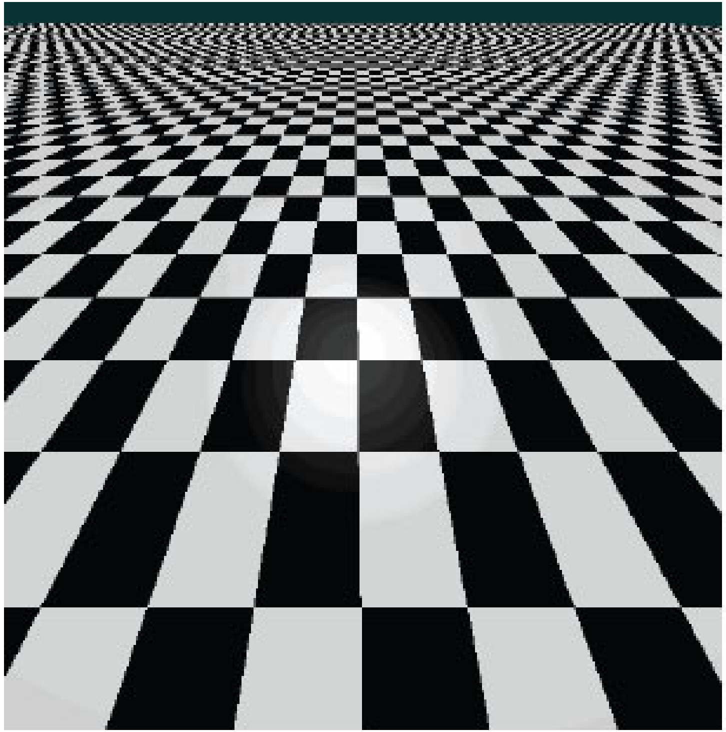
How would you deal with aliasing?

Anti-aliasing

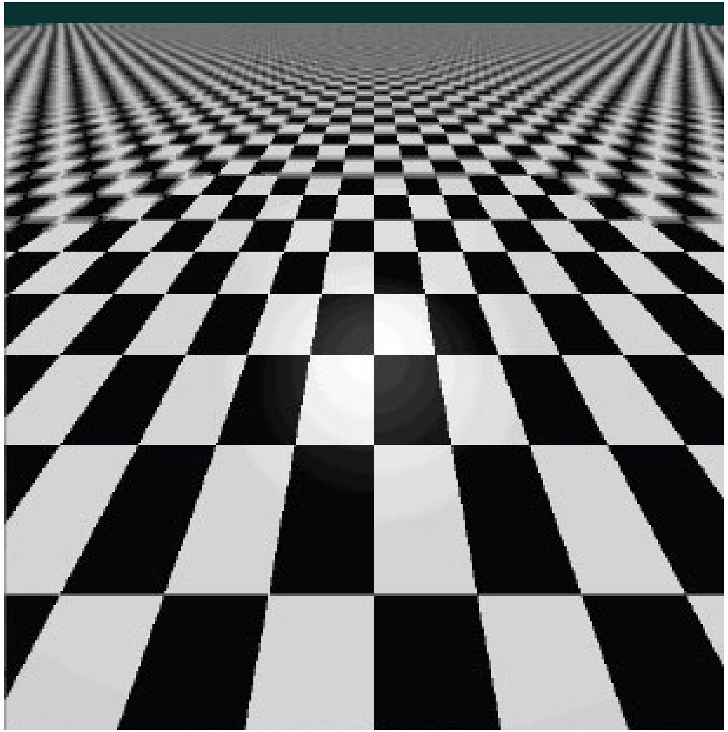
How would you deal with aliasing?

Approach 1: Oversample the signal

Anti-aliasing in textures



aliasing artifacts



anti-aliasing by oversampling

Anti-aliasing

How would you deal with aliasing?

Approach 1: Oversample the signal

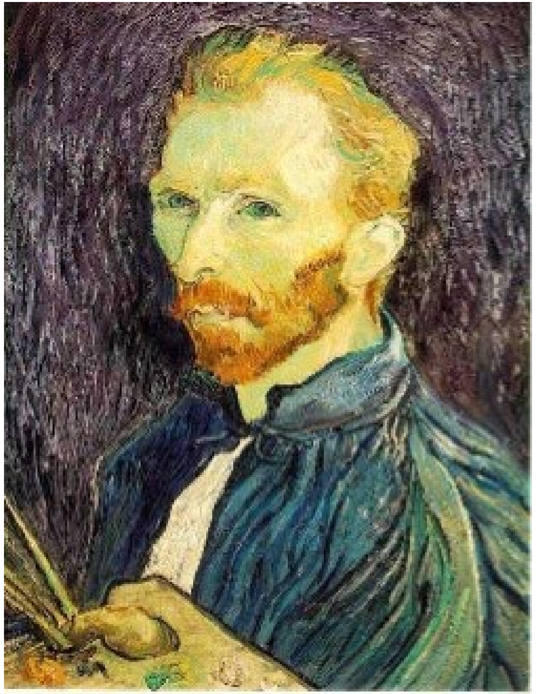
Approach 2: Smooth the signal

- Remove some of the detail effects that cause aliasing.
- Lose information, but better than aliasing artifacts.

How would you smooth a signal?

Better image downsampling

Apply a smoothing filter first, then throw away half the rows and columns



1/2

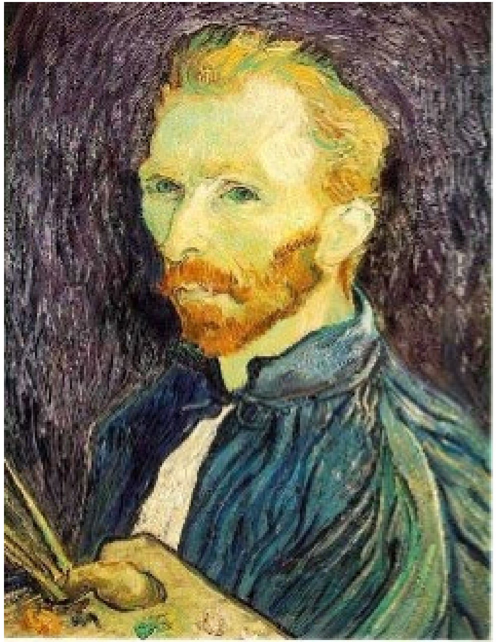
Gaussian filter
delete even rows
delete even columns

1/4

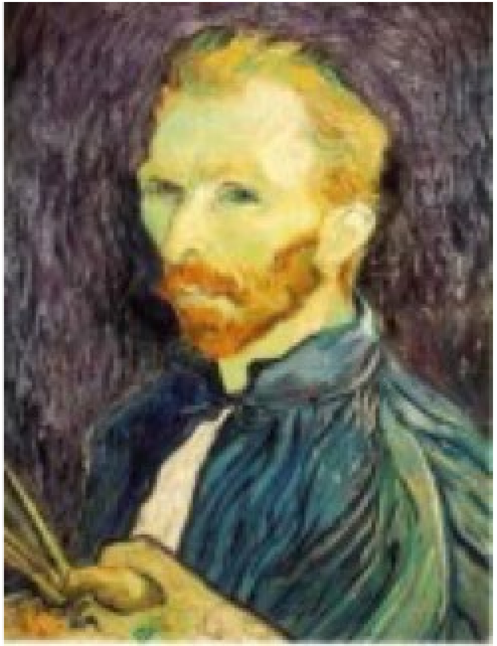
Gaussian filter
delete even rows
delete even columns

1/8

Better image downsampling



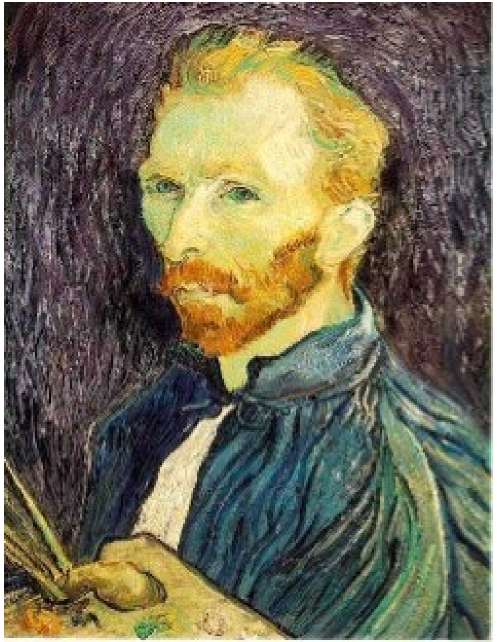
1/2



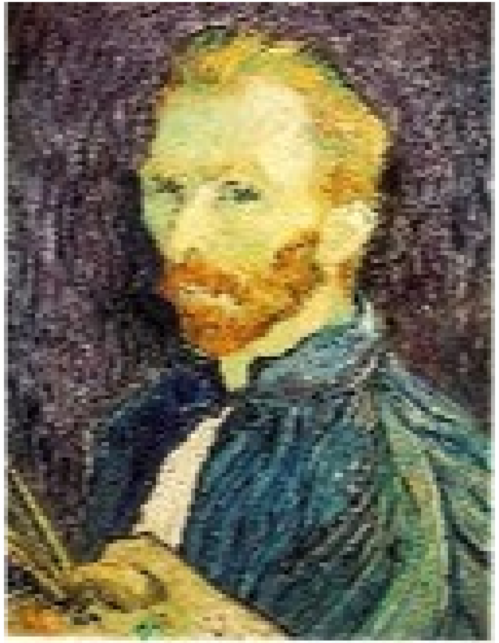
1/4 (2x zoom)

1/8 (4x zoom)

Naïve image downsampling



1/2



1/4 (2x zoom)

1/8 (4x zoom)

Anti-aliasing

Question 1: How much smoothing do I need to do to avoid aliasing?

Question 2: How many samples do I need to take to avoid aliasing?

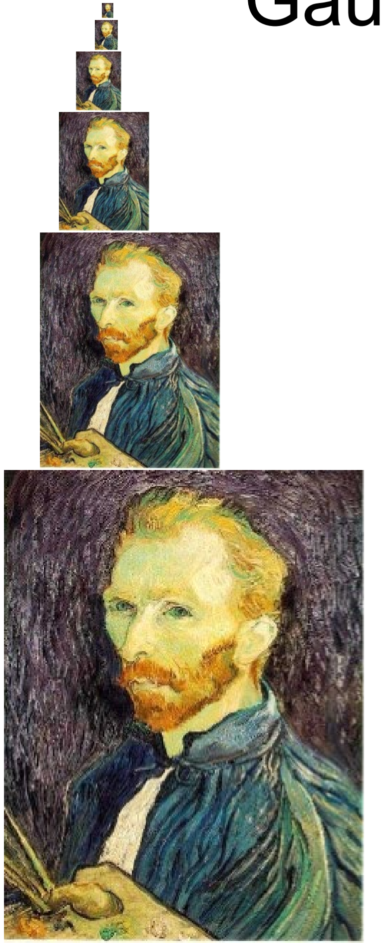
Answer to both: Enough to reach the Nyquist limit 2*fmax.

- The reason all of this matters for images is that through application of the Fourier Series, any signal of finite length can be represented as a sum of sinusoids.
- This means that even if a picture has no discernable wave pattern, it can still be represented as a sequence of sinusoids of different frequencies.
- The highest frequency that can be represented in the image is half the Nyquist rate (sampling frequency).

We'll see what this means soon.

Gaussian image pyramid

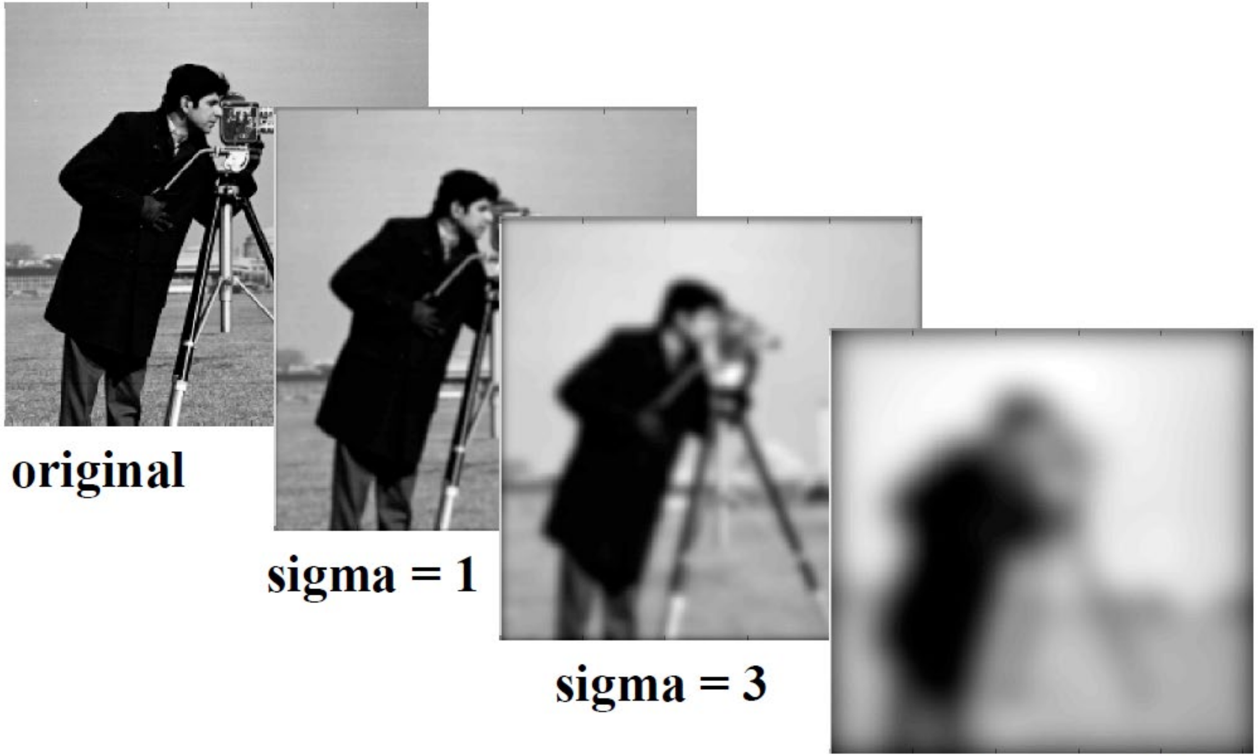
Gaussian image pyramid



The name of this sequence of subsampled images

Idea for Today:

Form a Multi-Resolution Representation



original

sigma = 1

sigma = 3

sigma = 10

Pyramid representation

Because a large amount of smoothing limits the frequency of features in the image, we do not need to keep all the pixels around!

Strategy: progressively reduce the number of pixels as we smooth more and more. Leads to a “pyramid” representation if we subsample at each level.

Synthesis: Smooth image with a Gaussian and downsample. Repeat.

- Gaussian is used because it is self-reproducing (enables incremental smoothing).

Constructing a Gaussian pyramid

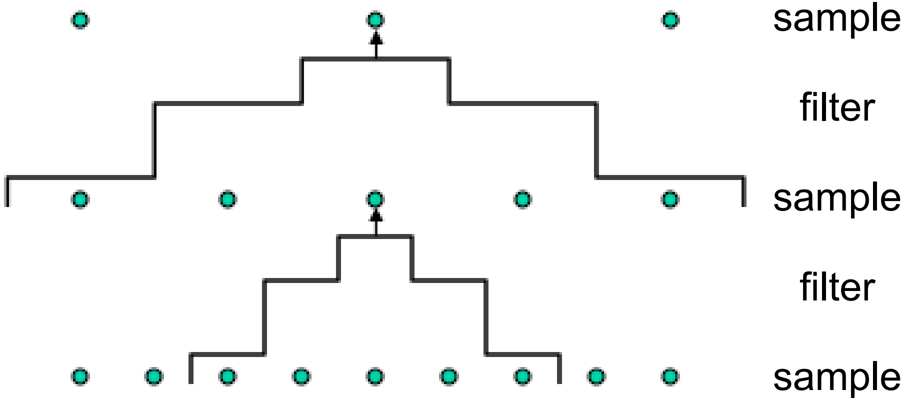
Algorithm

repeat:

filter

subsample

until min resolution reached



Question: How much bigger than the original image is the whole pyramid?

Constructing a Gaussian pyramid

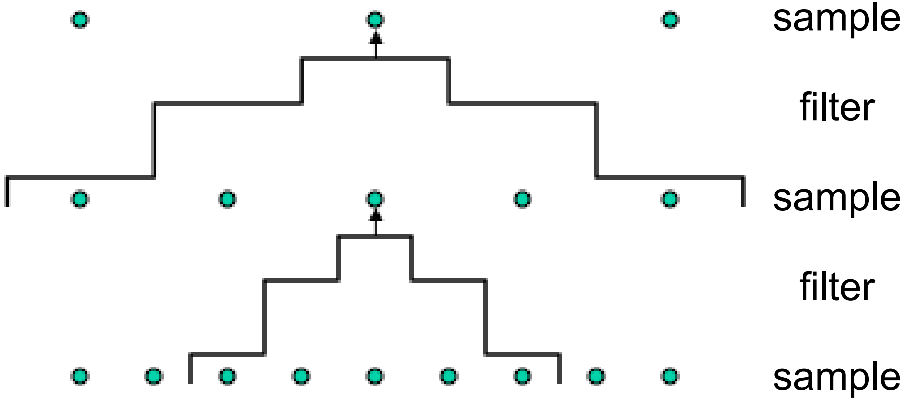
Algorithm

repeat:

filter

subsample

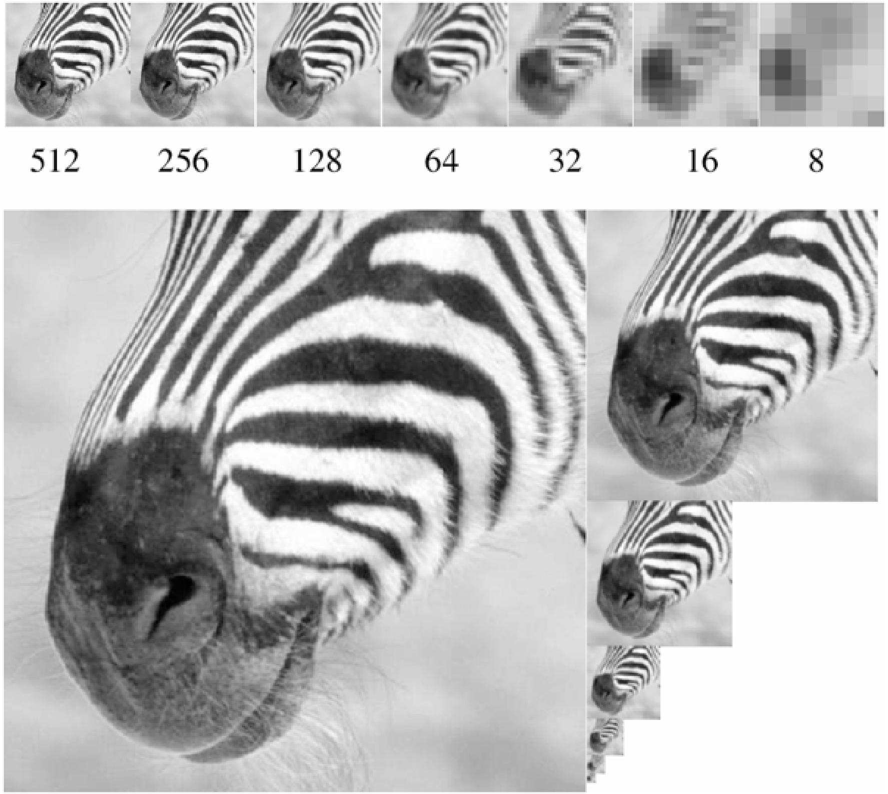
until min resolution reached



Question: How much bigger than the original image is the whole pyramid?

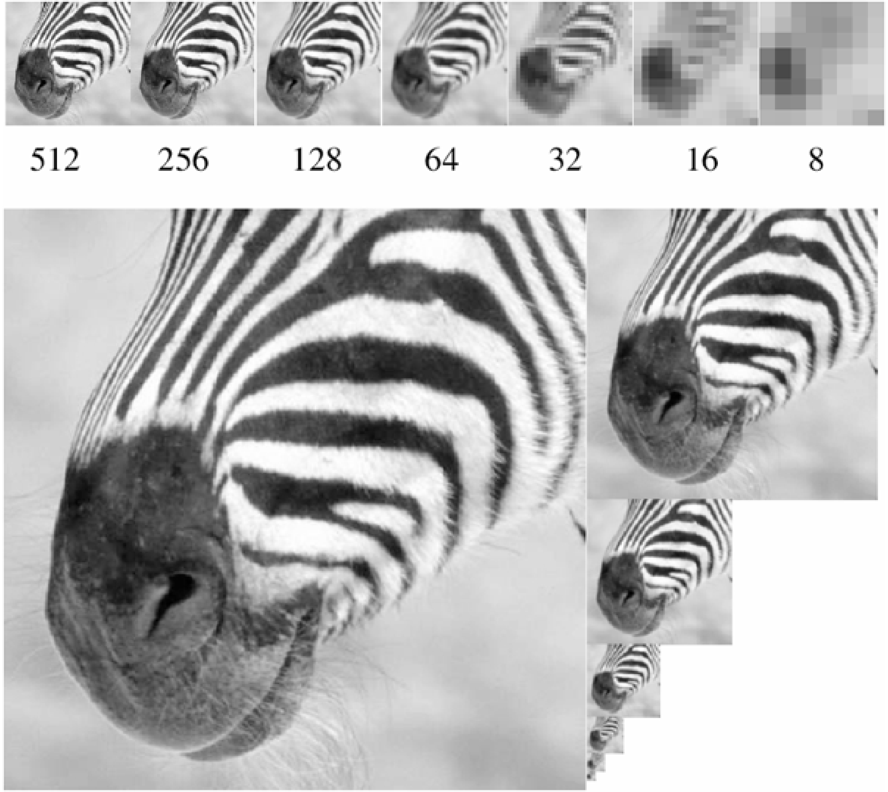
Answer: Just 4/3 times the size of the original image!

Some properties of the Gaussian pyramid



What happens to the details of the image?

Some properties of the Gaussian pyramid

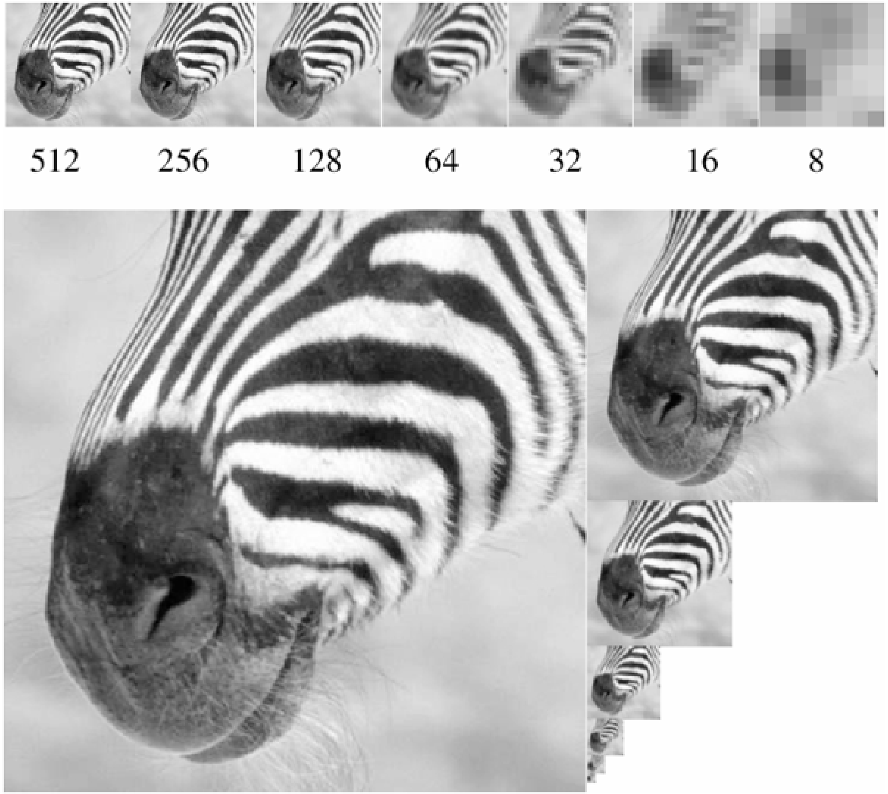


What happens to the details of the image?

- They get smoothed out as we move to higher levels.

What is preserved at the higher levels?

Some properties of the Gaussian pyramid



What happens to the details of the image?

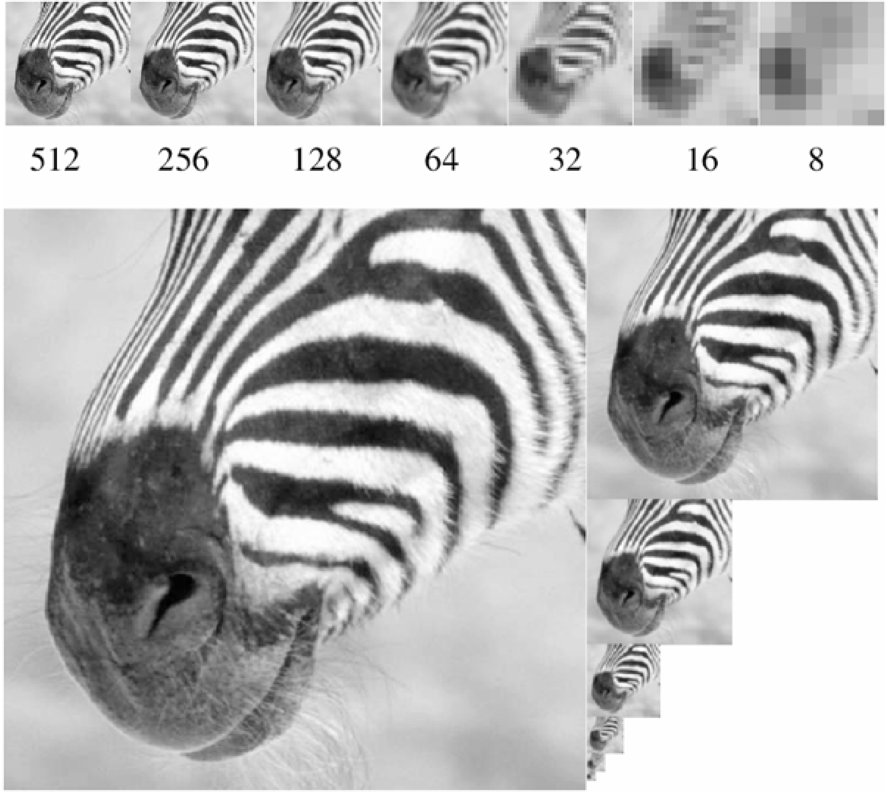
- They get smoothed out as we move to higher levels.

What is preserved at the higher levels?

- Mostly large uniform regions in the original image.

How would you reconstruct the original image from the image at the upper level?

Some properties of the Gaussian pyramid



What happens to the details of the image?

- They get smoothed out as we move to higher levels.

What is preserved at the higher levels?

- Mostly large uniform regions in the original image.

How would you reconstruct the original image from the image at the upper level?

- That’s not possible.

Blurring is lossy

\-

=

level 0

level 1 (before downsampling)

residual

What does the residual look like?

Blurring is lossy

-

=

level 0

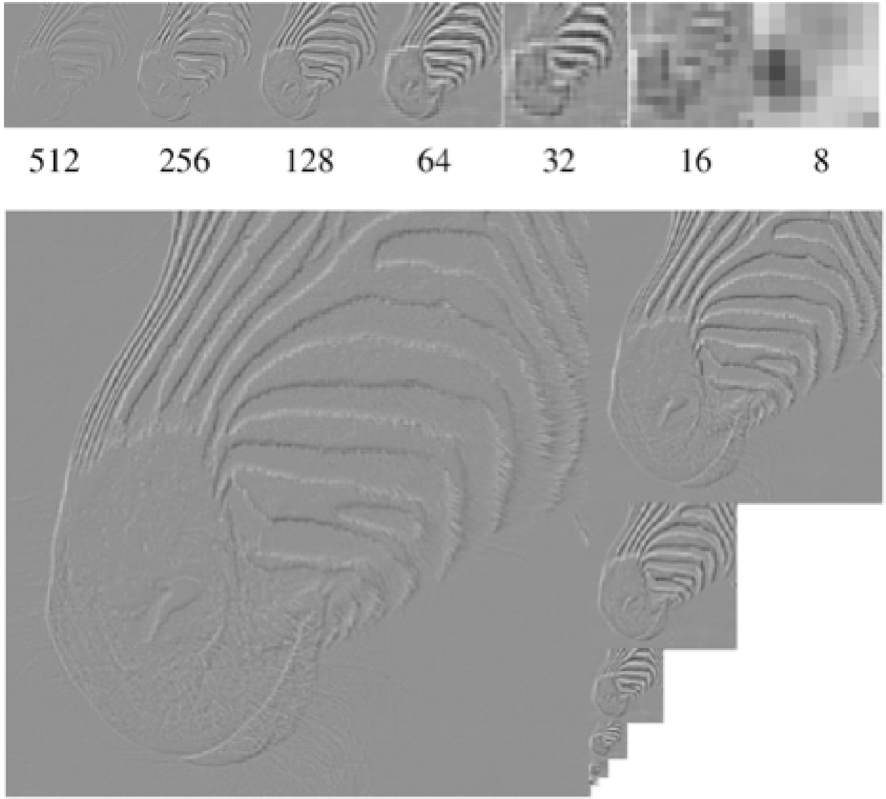
level 1 (before downsampling)

residual

Can we make a pyramid that is lossless?

Laplacian image pyramid

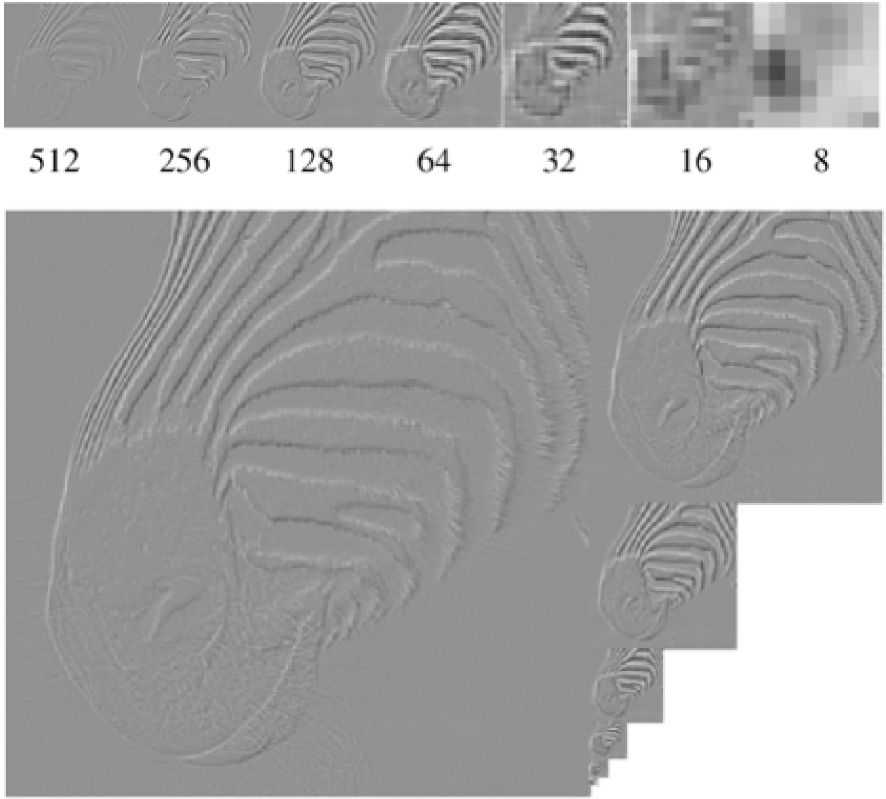
Laplacian image pyramid



At each level, retain the residuals instead of the blurred images themselves.

Can we reconstruct the original image using the pyramid?

Laplacian image pyramid



At each level, retain the residuals instead of the blurred images themselves.

Can we reconstruct the original image using the pyramid?

- Yes we can!

What do we need to store to be able to reconstruct the original image?

Let’s start by looking at just one level

level 0

=

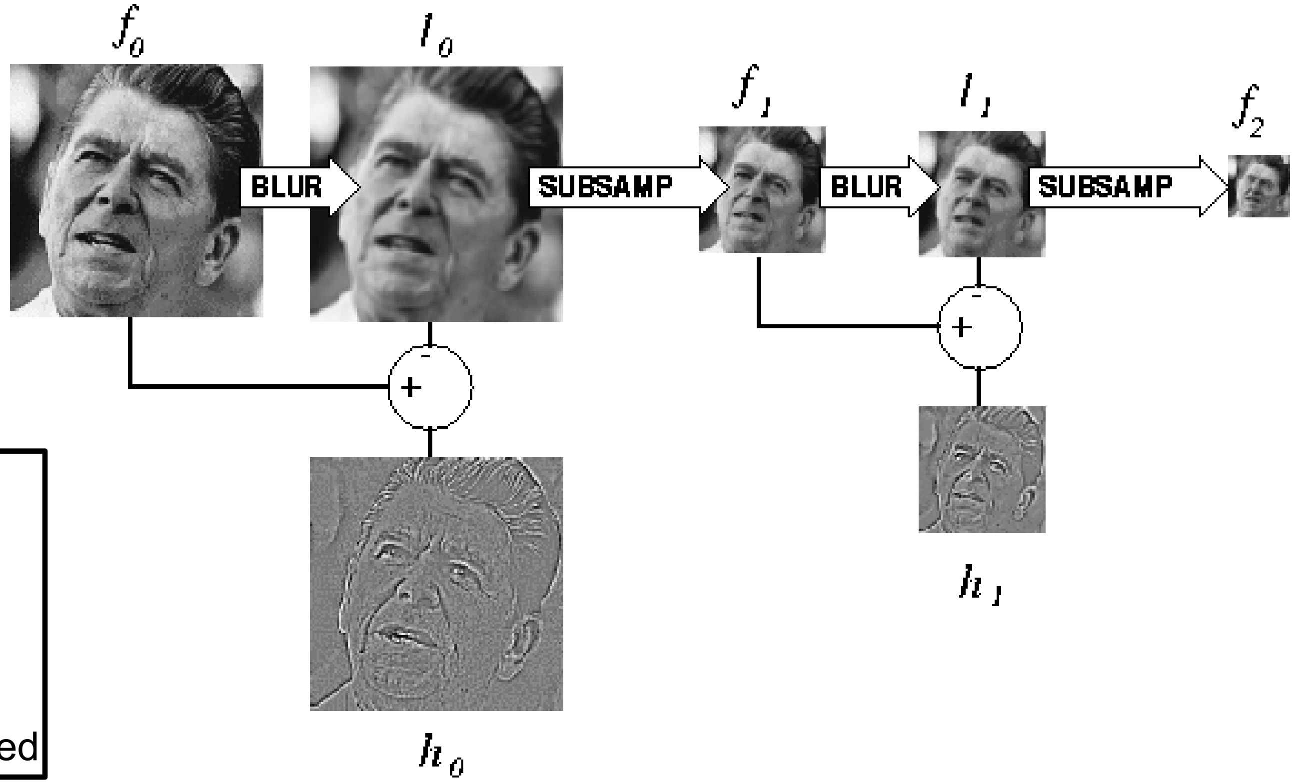
level 1 (upsampled)

+

residual

Does this mean we need to store both residuals and the blurred copies of the original?

Constructing a Laplacian pyramid



Algorithm

repeat:

filter

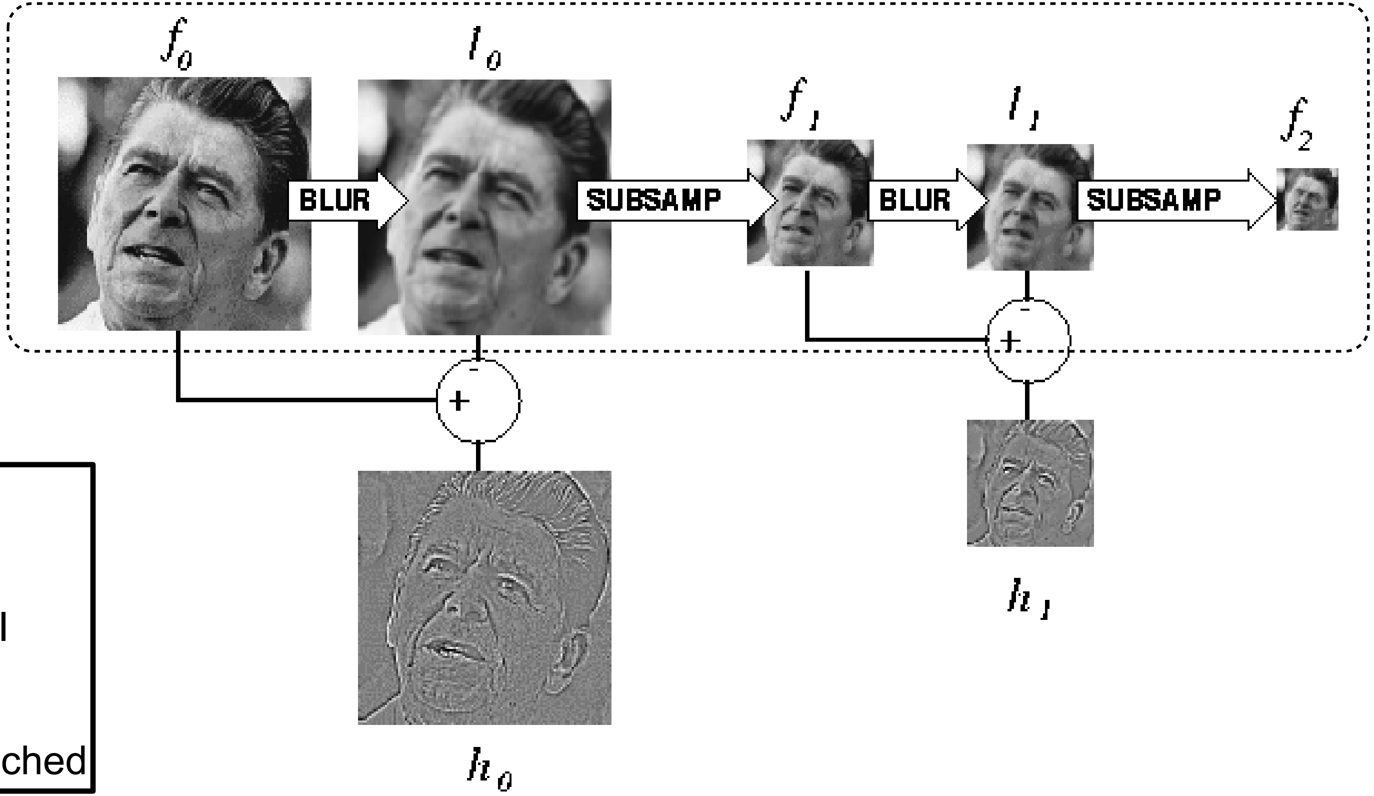
compute residual

subsample

until min resolution reached

Constructing a Laplacian pyramid

What is this part?



Algorithm

repeat:

filter

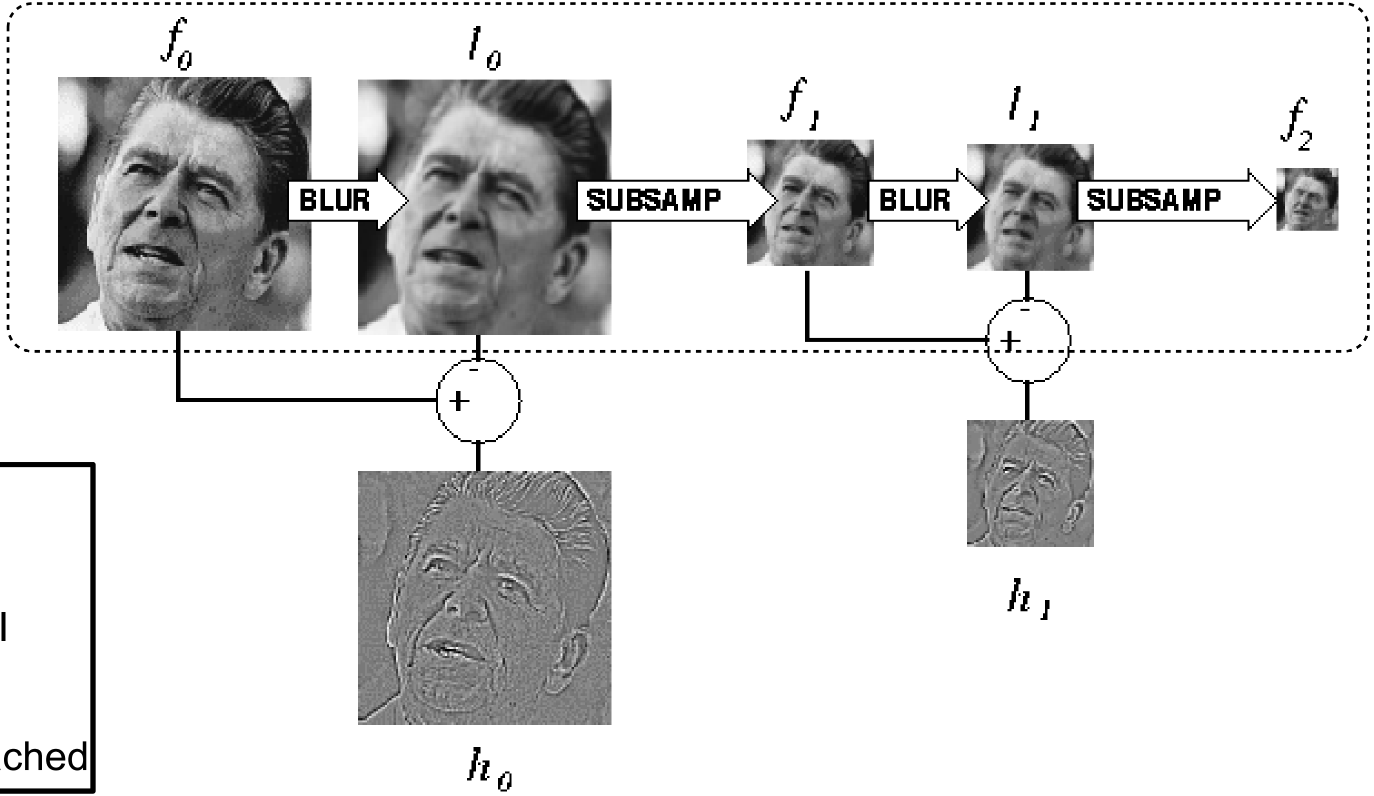
compute residual

subsample

until min resolution reached

Constructing a Laplacian pyramid

It's a Gaussian pyramid.



Algorithm

repeat:

filter

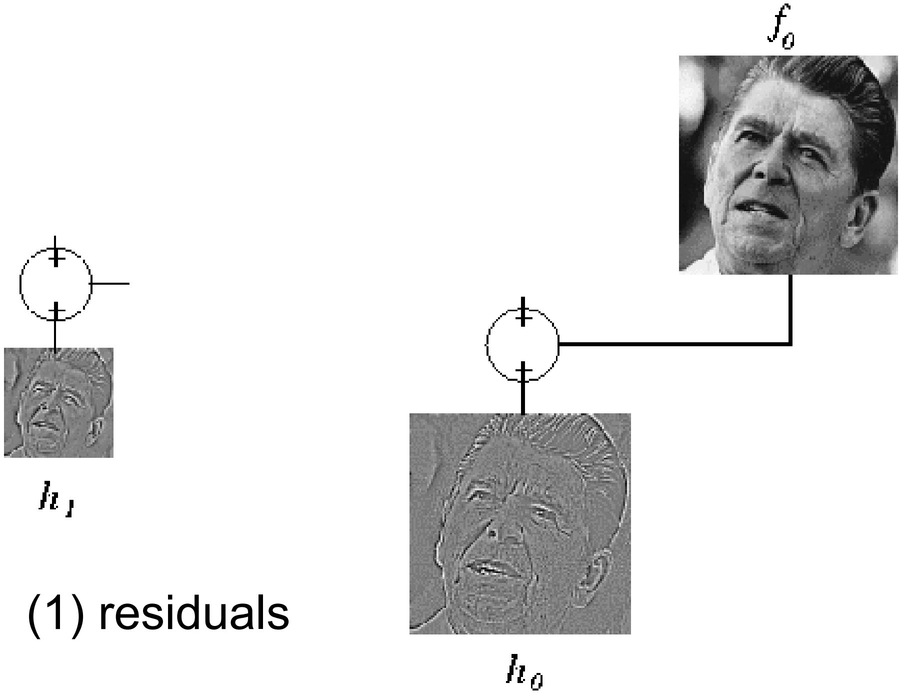
compute residual

subsample

until min resolution reached

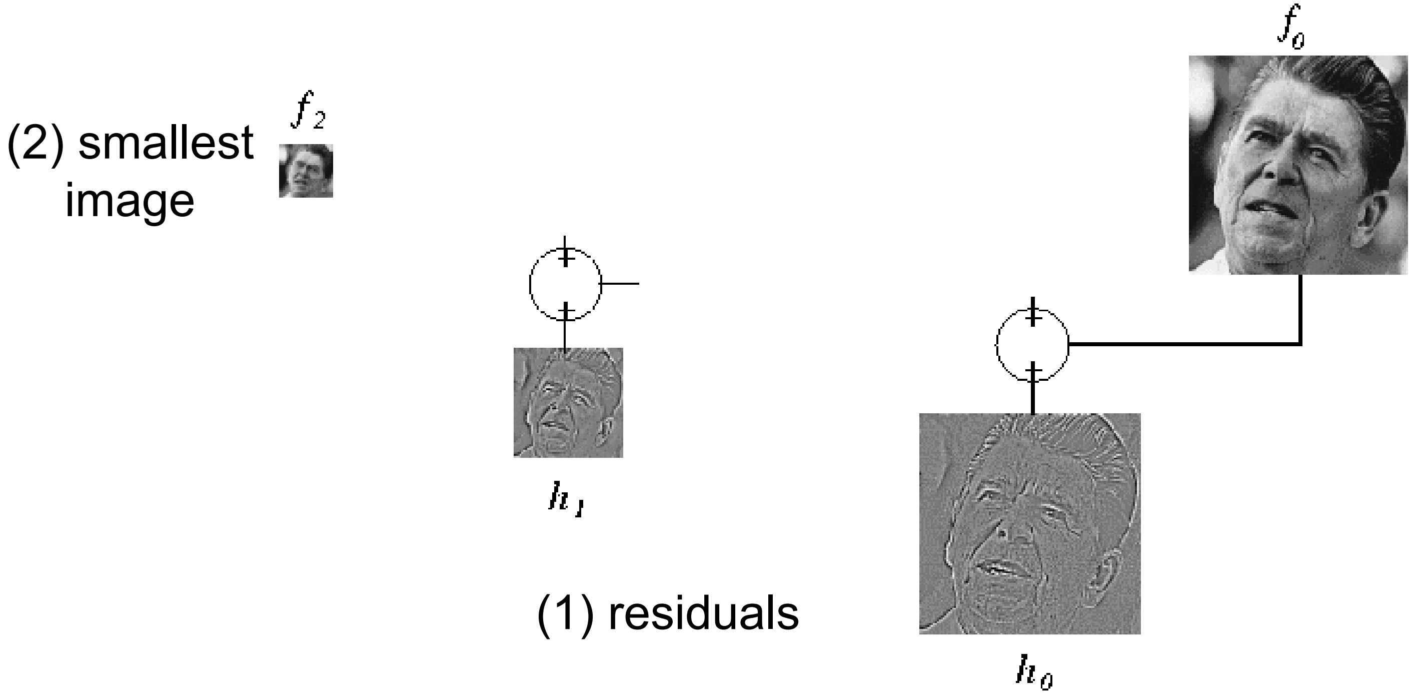
What do we need to construct the original image?

What do we need to construct the original image?

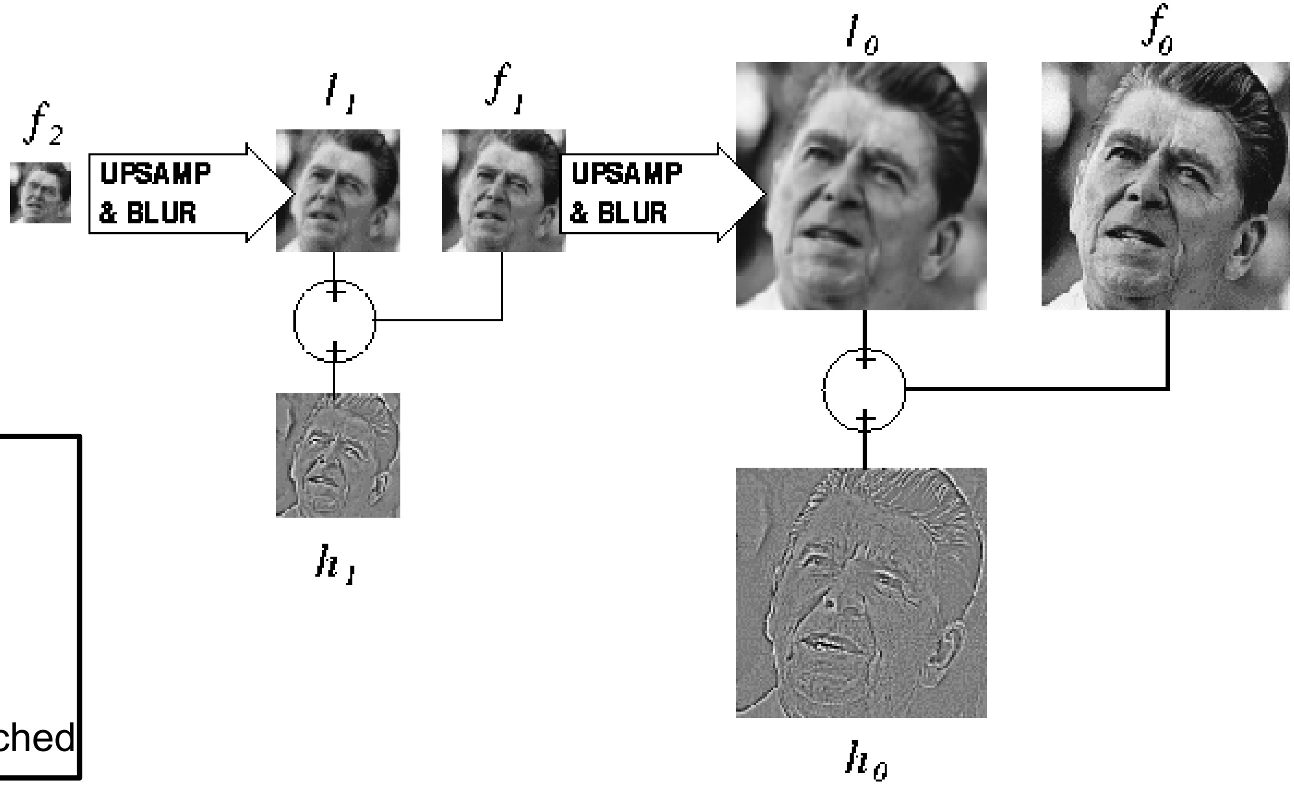


(1) residuals

What do we need to construct the original image?



Reconstructing the original image



Algorithm

repeat:

upsample

sum with residual

until orig resolution reached

Gaussian vs Laplacian Pyramid

Shown in opposite order for space.

Why is it called a Laplacian pyramid?

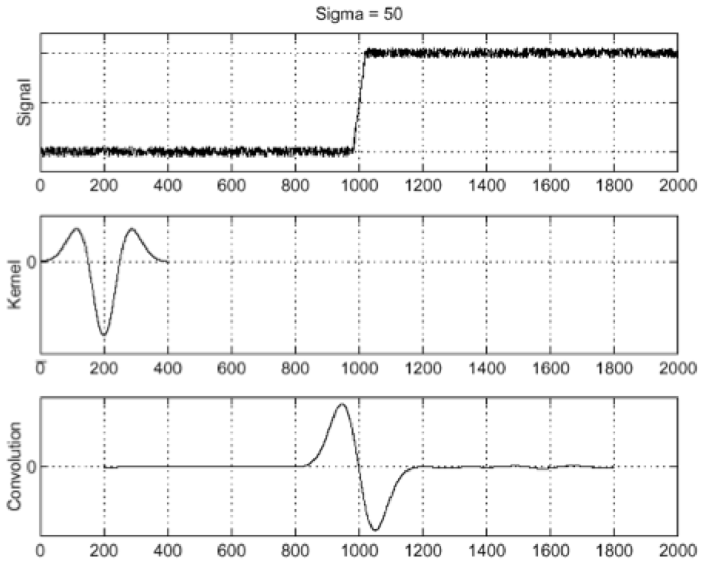
Reminder: Laplacian of Gaussian (LoG) filter

As with derivative, we can combine Laplace filtering with Gaussian filtering

input

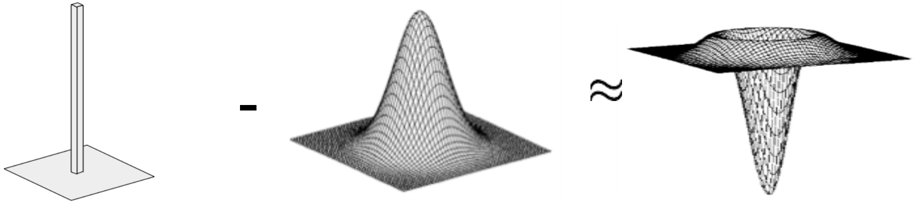
Laplacian of Gaussian

output

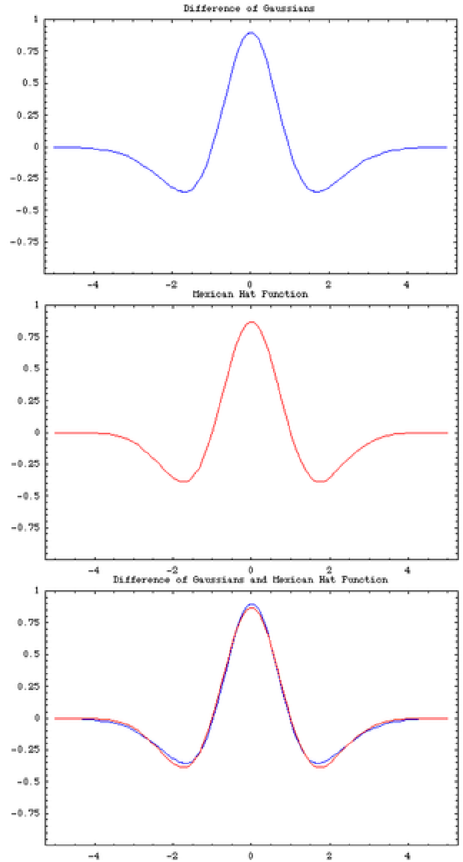


“zero crossings” at edges

Why is it called a Laplacian pyramid?



unit Gaussian Laplacian



Difference of Gaussians approximates the Laplacian

What are image pyramids used for?

image compression

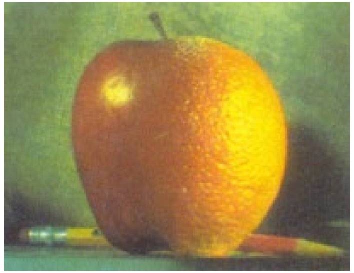
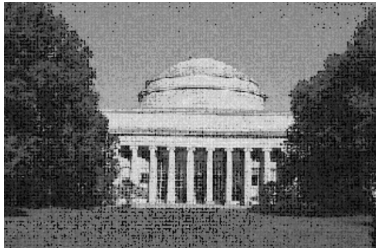


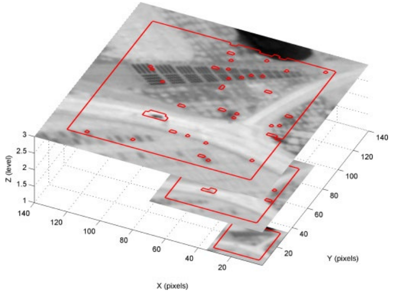
image blending



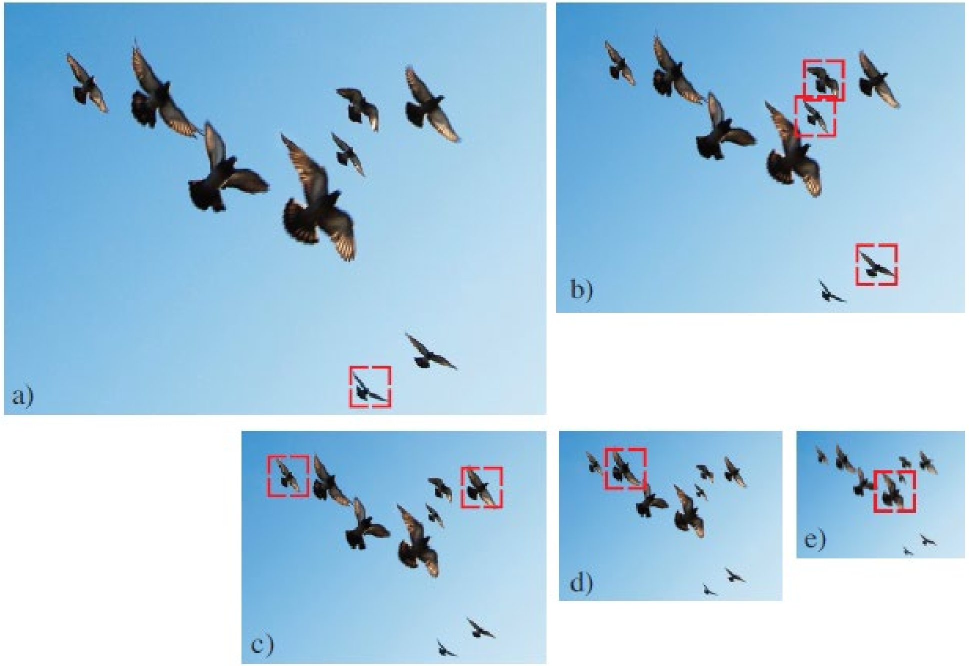
denoising

multi-scale detection

multi-scale registration

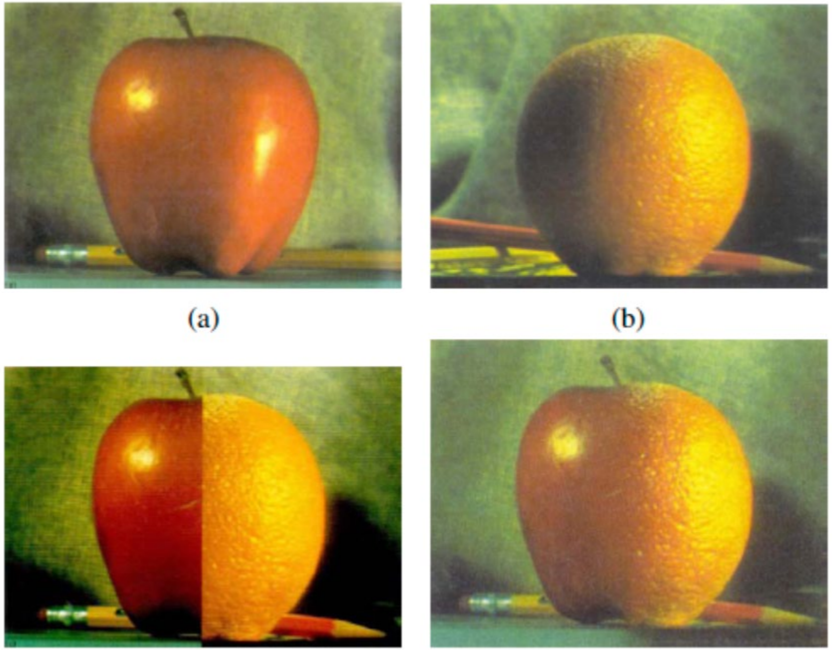


Multi-scale image analysis

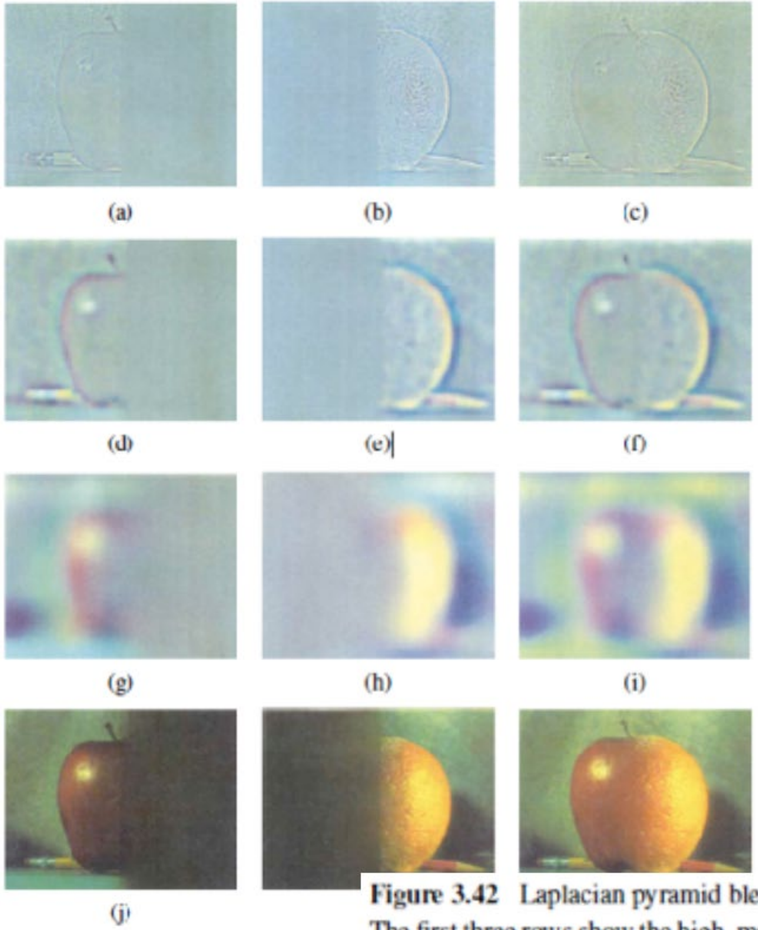


- Multiscale image pyramid.
  Each image is 25% smaller than the previous one.
  - The red box indicates the size of a template used for detecting flying birds.
  - As the size of the template is fixed, it will only be able to detect the birds that tightly t inside the box. Birds that are smaller or larger will not be detected within a single scale.
  - By running the same template across many levels in this pyramid, different birds instances are detected at different scales.

Image blending



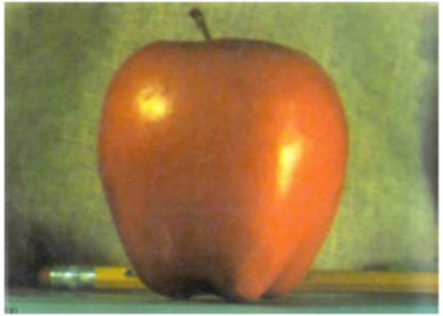
(a) (b)



Szeliski, Computer Vision, 2010

Figure 3.42 Laplacian pyramid blending details (Burt and Adelson 1983b) © 1983 ACM. The first three rows show the high, medium, and low frequency parts of the Laplacian pyramid (taken from levels 0, 2, and 4). The left and middle columns show the original apple and orange images weighted by the smooth interpolation functions, while the right column shows the averaged contributions.

Image blending

- Build Laplacian pyramid for both images: LA, LB
- Build Gaussian pyramid for mask: G
- Build a combined Laplacian pyramid: L(j) = G(j) LA(j) + (1-G(j)) LB(j)
- Collapse L to obtain the blended image

Aude Oliva, Antonio Torralba, and Philippe G. Schyns. 2006. Hybrid images. ACM Trans. Graph. 25, 3 (July 2006), 527-532.

Other types of pyramids

Steerable pyramid: At each level keep multiple versions, one for each direction.

Wavelets: Huge area in image processing

Acknowledgement: some slides and material from Bernt Schiele, Mario Fritz, Michael Black, Bill Freeman, Fei-Fei, Justin Johnson, Serena Yeung, R. Szelisky, Fabio Galasso, Ioannis Gkioulekas