

3D Vision and Machine Perception

Prof. Kyungdon Joo

3D Vision & Robotics Lab.

AI Graduate School (AIGS) & Computer Science and Engineering (CSE)

Contents

- Image Downsampling
- Image Upsampling (=interpolation)

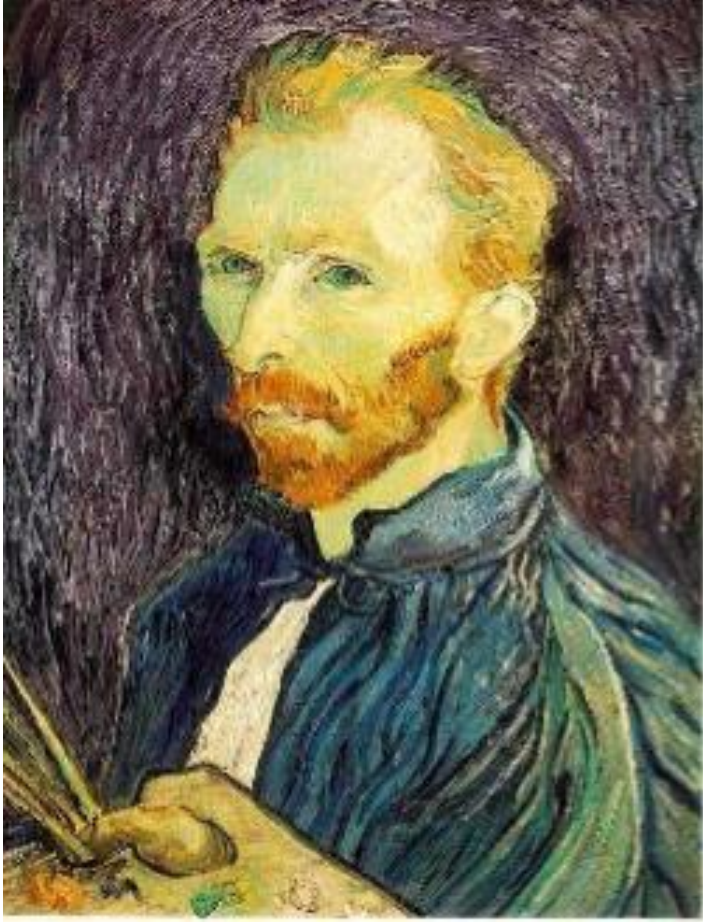
Image Downsampling

A portrait of Vincent van Gogh's 'Self-Portrait with Bandaged Ear'. The painting shows a man with a thick, curly beard and hair, rendered in shades of yellow and orange. He has a pale, greenish-yellow complexion and piercing blue eyes. His right ear is bandaged, and he is looking slightly to the left. The background is a dark, textured, brownish-purple. The brushstrokes are visible and expressive.

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

Naïve image subsampling

- Throw away half the rows and columns



$1/2$

delete even rows
delete even columns



$1/4$

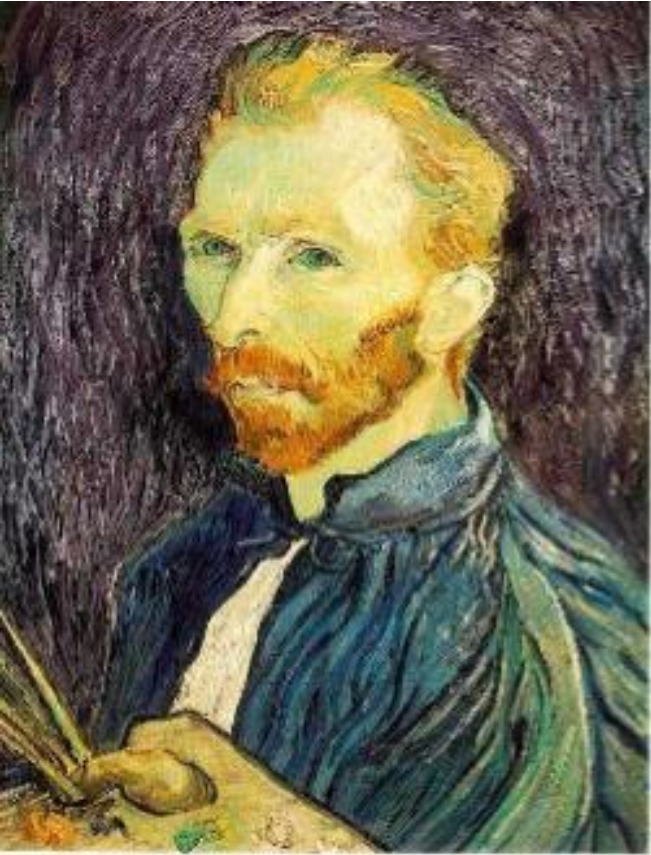
delete even rows
delete even columns



$1/8$

- What is the problem with this approach?

Naïve image subsampling



$1/2$



$1/4$ ($2\times$ zoom)

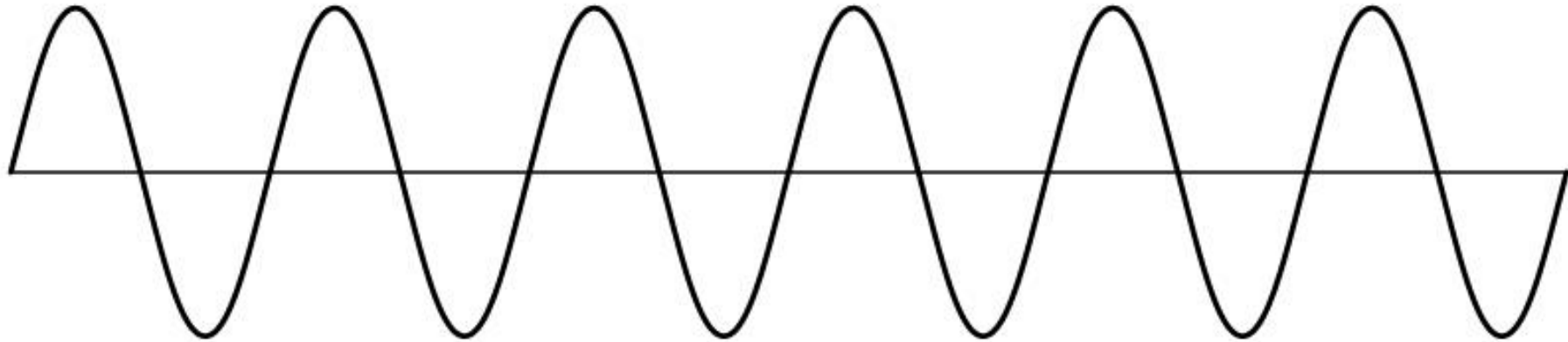


$1/8$ ($4\times$ zoom)

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

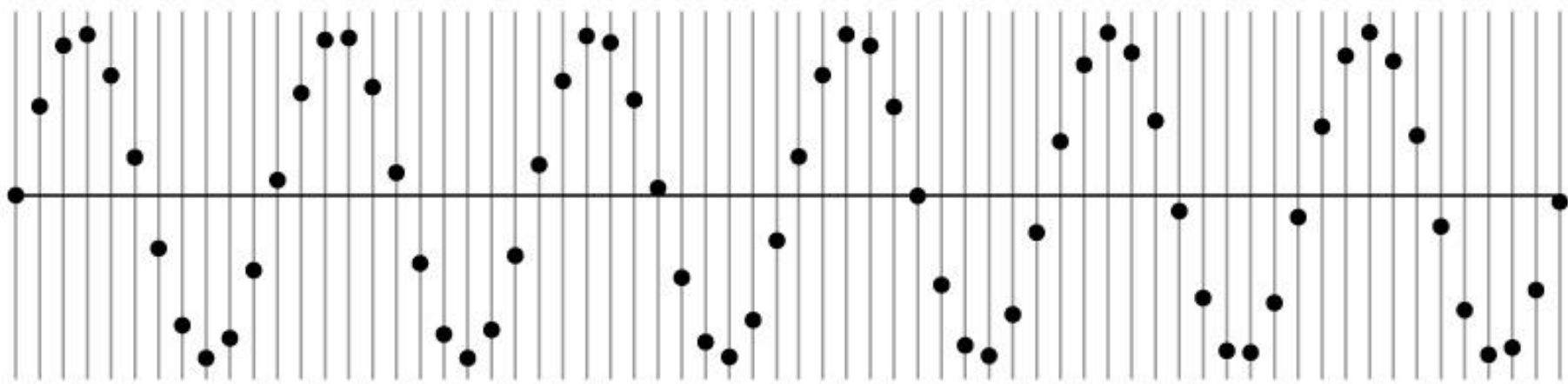
Sampling

- Very simple example: a sine wave
- How would you discretize this signal?



Sampling

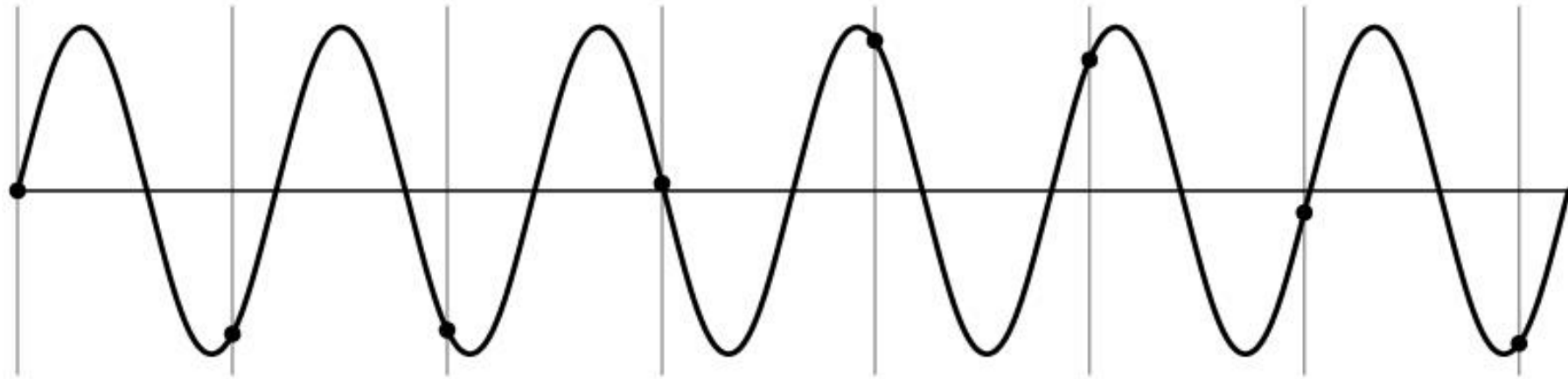
- Very simple example: a sine wave
- How many samples should I take?
- Can I take as many samples as I want?



- Can I take as few samples as I want?

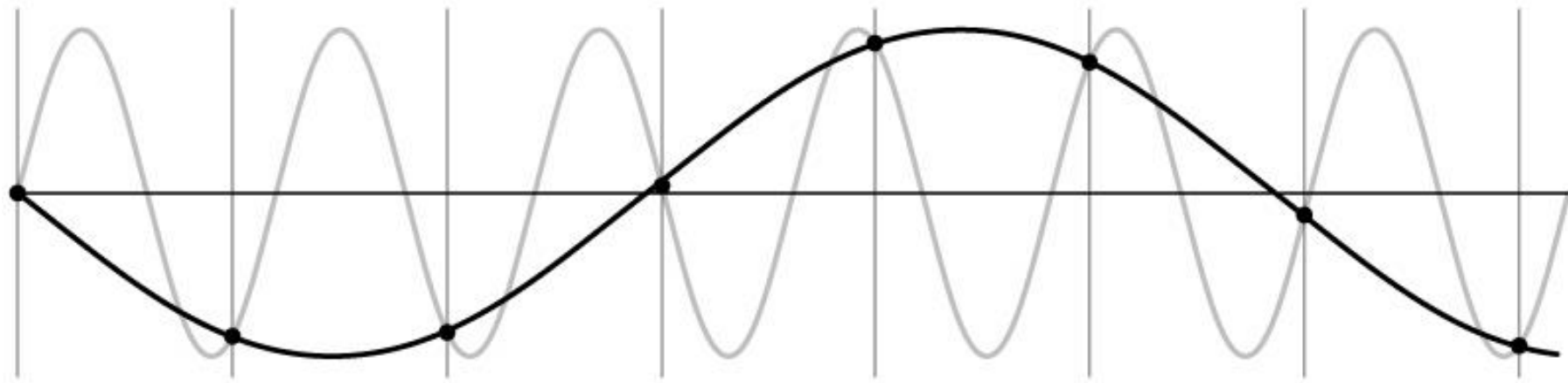
Undersampling

- Very simple example: a sine wave
- Unsurprising effect: information is lost.



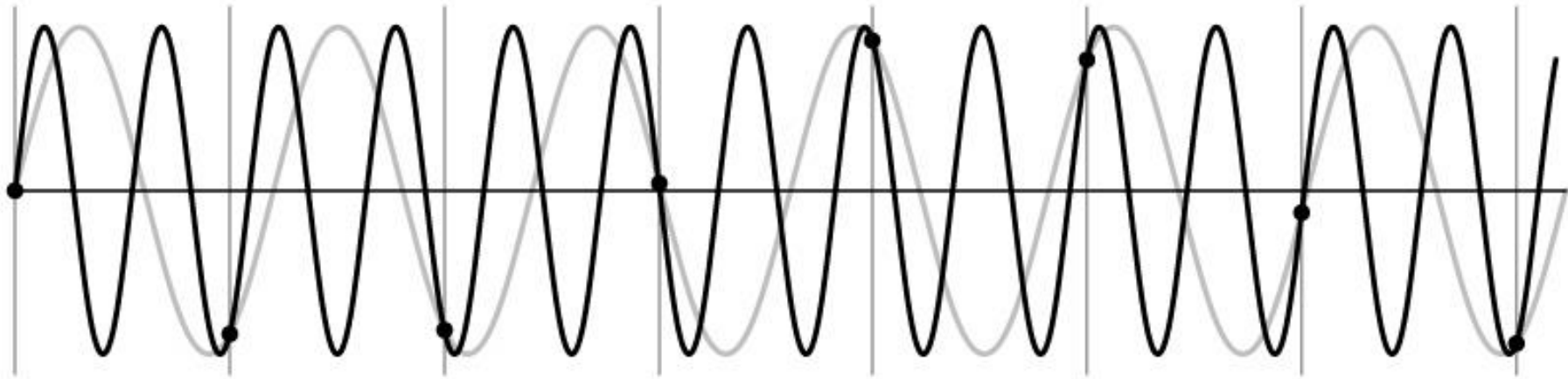
Undersampling

- Very simple example: a sine wave
- Unsurprising effect: information is lost.
- Surprising effect: can confuse the signal with one of lower frequency.



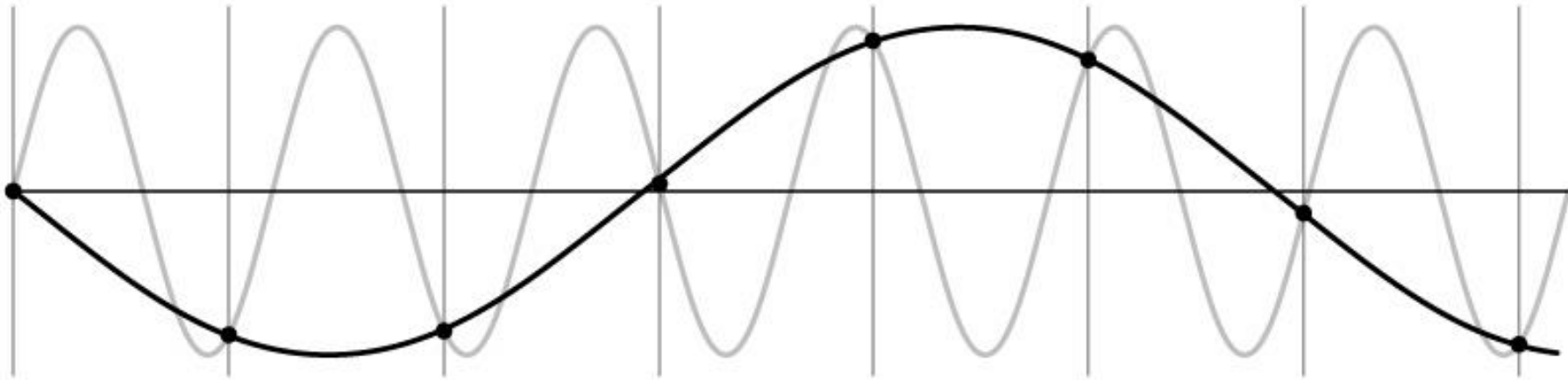
Undersampling

- 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.



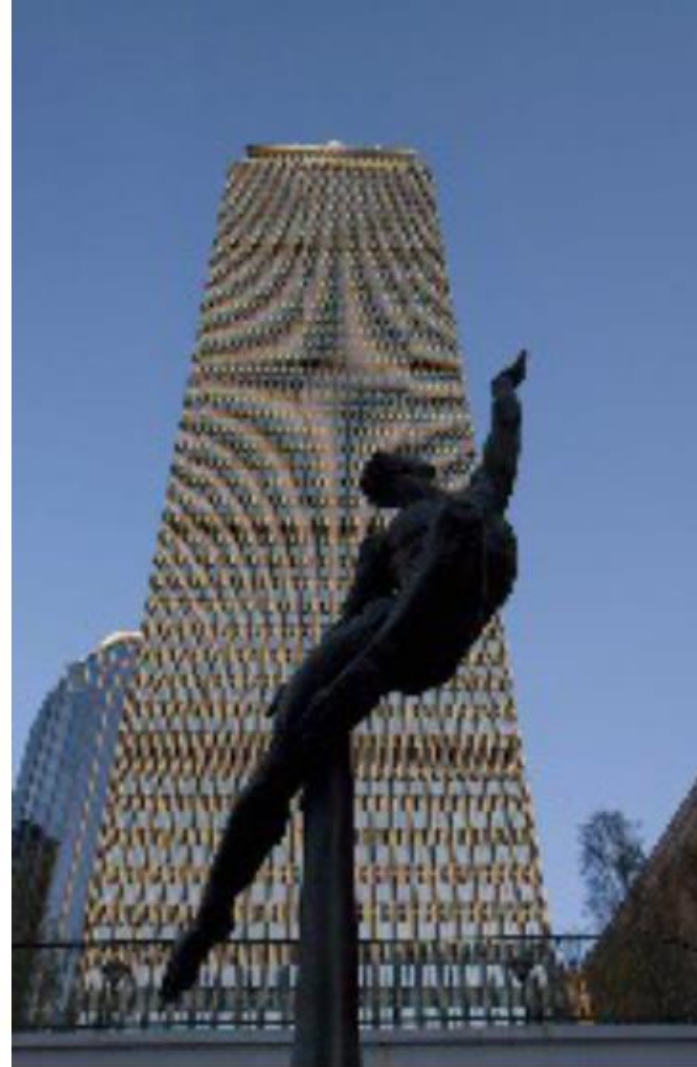
Aliasing

- Fancy term for: Undersampling can disguise a signal as one of a lower frequency



Aliasing in photographs

- This is also known as “moire”



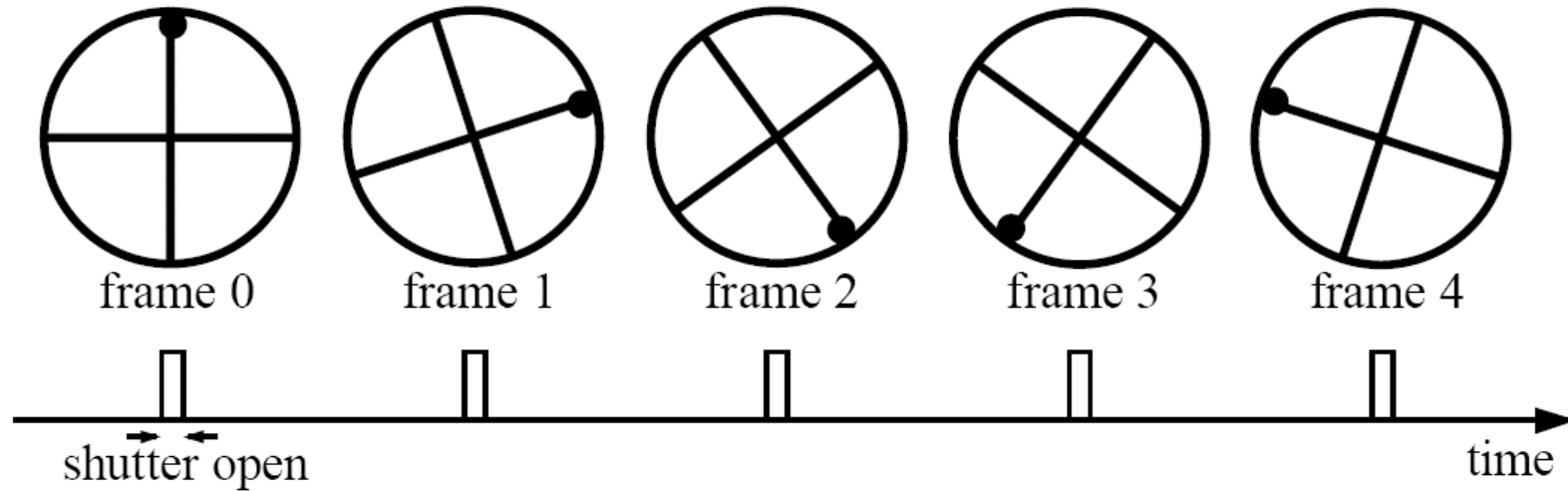
Source: F. Durand

Temporal aliasing

Imagine a spoked wheel moving to the right (rotating clockwise).

Mark wheel with dot so we can see what's happening.

If camera shutter is only open for a fraction of a frame time (frame time = $1/30$ sec. for video, $1/24$ sec. for film):



Without dot, wheel appears to be rotating slowly backwards!
(counterclockwise)



Anti-aliasing

- How would you deal with aliasing?

Anti-aliasing

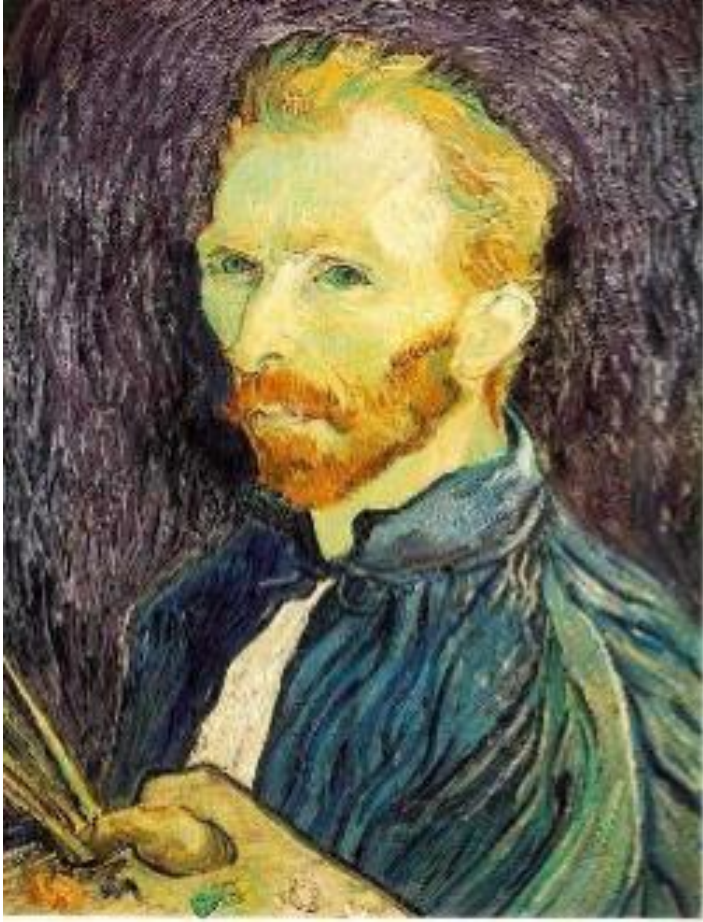
- How would you deal with aliasing?
- Approach 1: Oversample the signal

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?

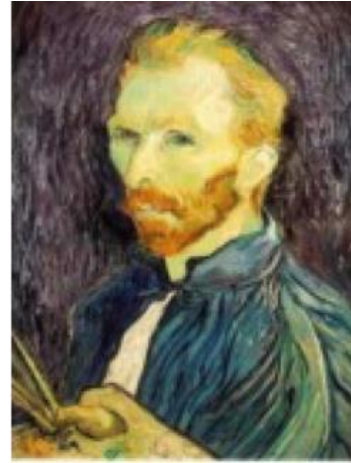
Better image downsampling

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



$1/2$

Gaussian filter
delete even rows
delete even columns



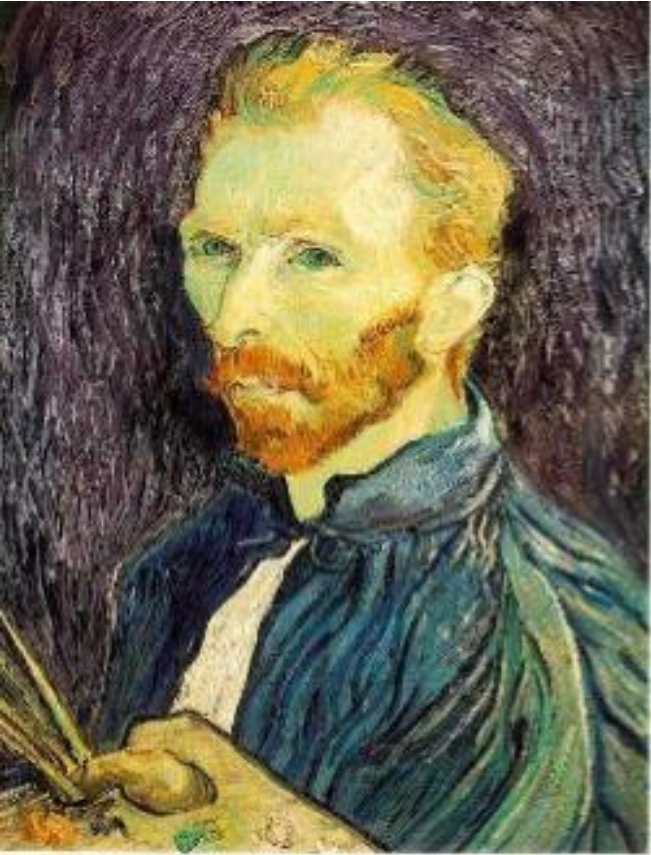
$1/4$

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



$1/2$

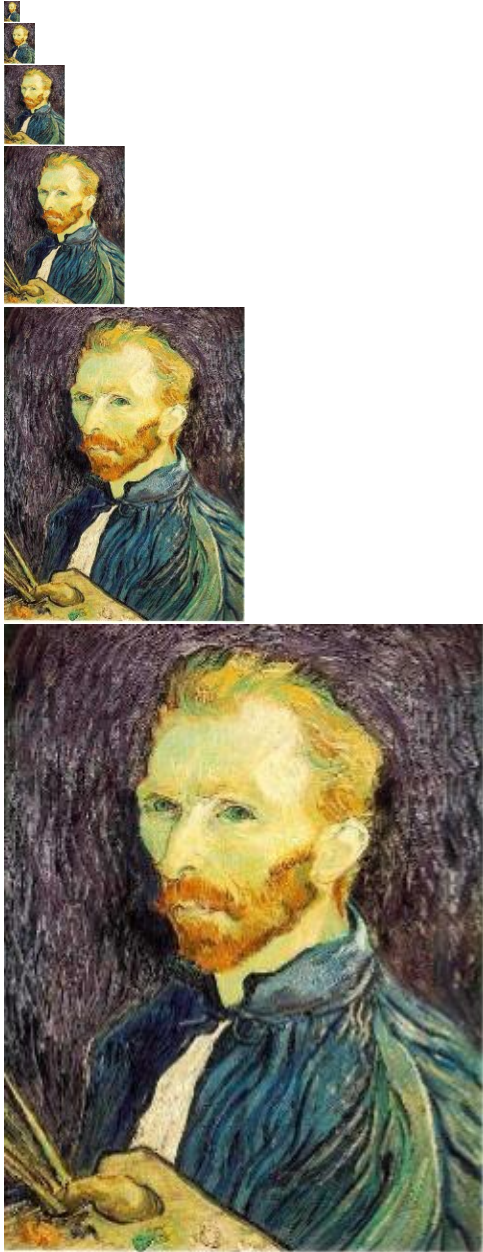


$1/4$ (2x zoom)



$1/8$ (4x zoom)

Gaussian image pyramid



- The name of this sequence of subsampled images

Algorithm

repeat:

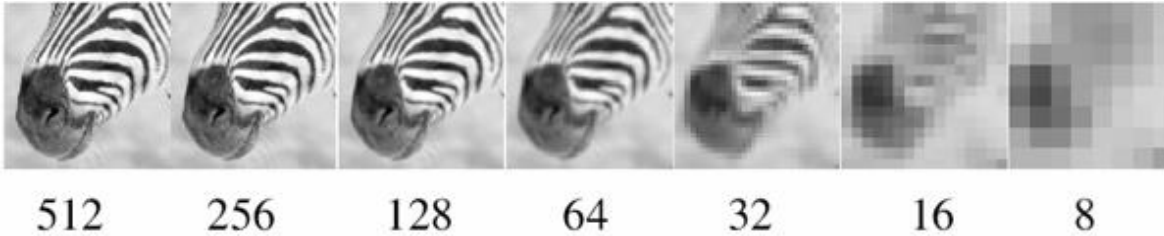
 filter

 subsample

until min resolution reached

Some properties of the Gaussian pyramid

- What happens to the details of the image?
 - They get smoothed out as we move to higher levels.

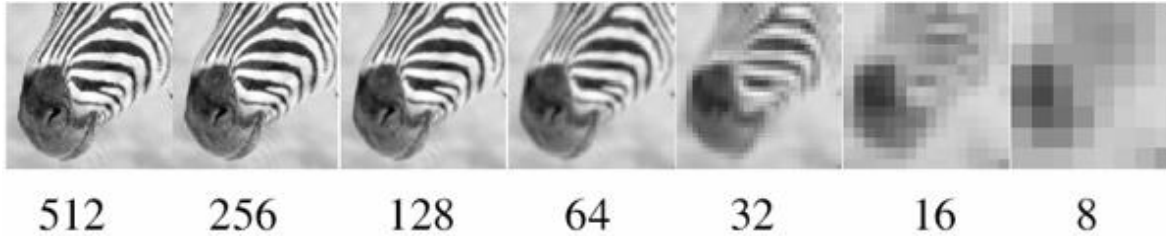


- What is preserved at the higher levels?



Some properties of the Gaussian pyramid

- What happens to the details of the image?
 - They get smoothed out as we move to higher levels.

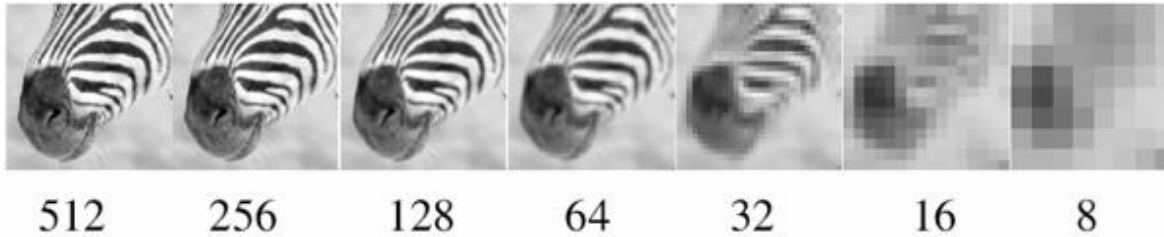


- What is preserved at the higher levels?
 - Mostly large uniform regions in the original image.
- When is it useful?



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.

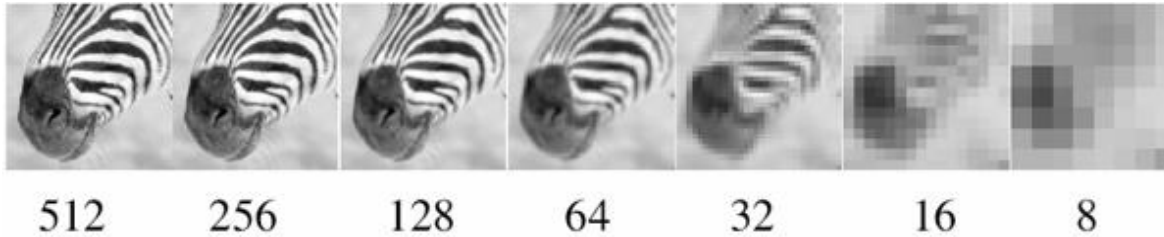
- When is it useful?

- 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.

- When is it useful?

- How would you reconstruct the original image from the image at the upper level?
 - That's not possible.



Blurring is lossy

- What does the residual look like?



level 0

-



level 1 (before downsampling)

=

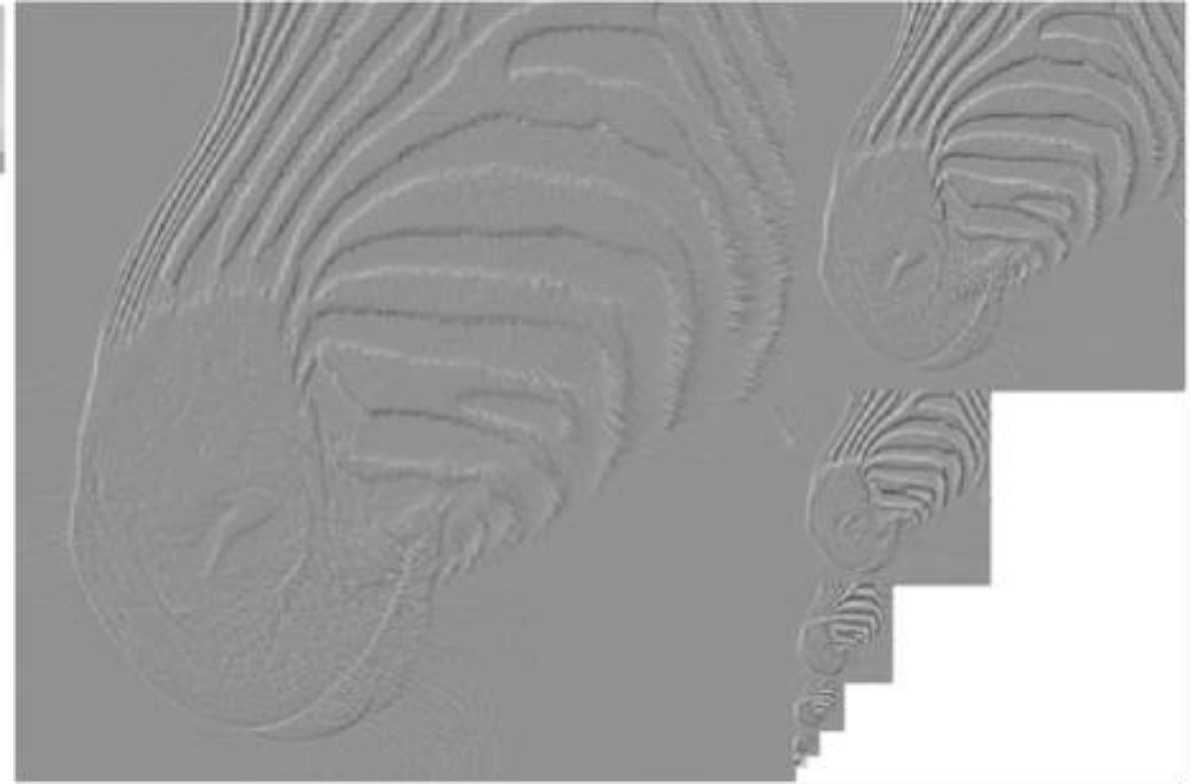
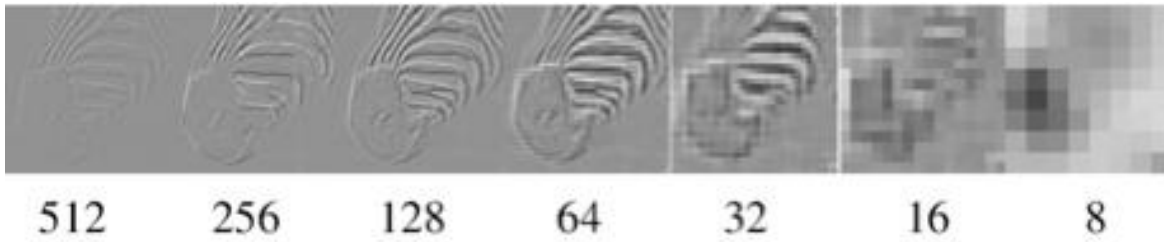


residual

- Can we make a pyramid that is lossless?

Laplacian image pyramid

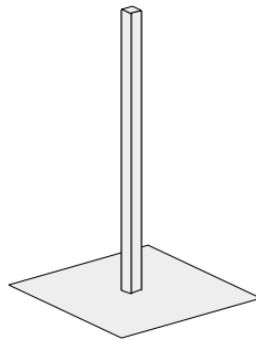
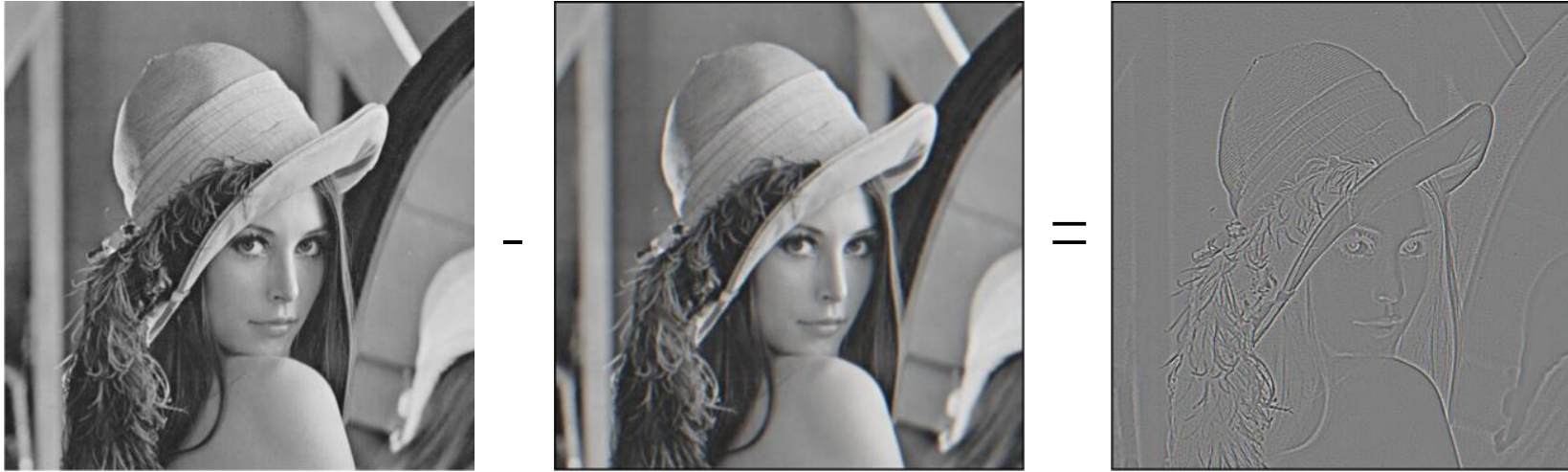
- At each level, retain the residuals instead of the blurred images themselves.
- It can reconstruct the original image.



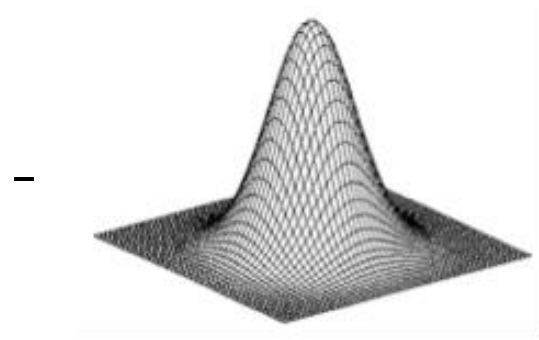
D.I.Y.

Why is it called a Laplacian pyramid?

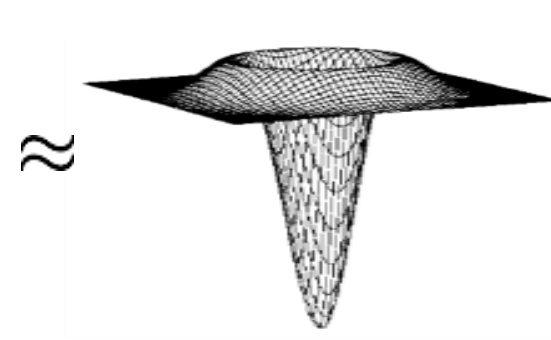
- Difference of Gaussians approximates the Laplacian



unit



Gaussian



Laplacian

What are image pyramids used for?

image compression



multi-scale
texture mapping

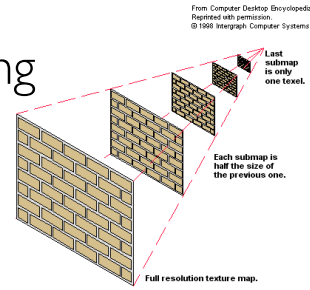
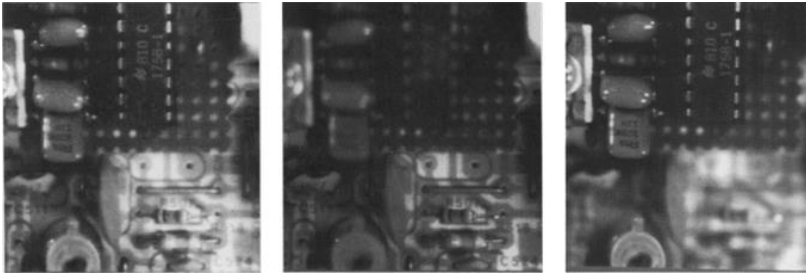


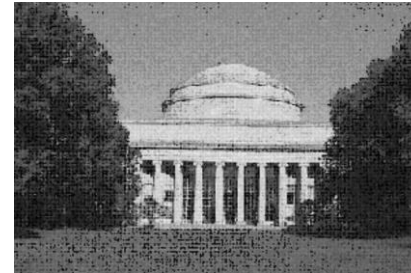
image blending



focal stack compositing



denoising



multi-scale detection



multi-scale registration

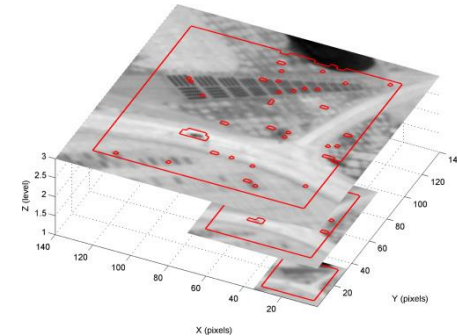


Image interpolation

Original image:  × 10



Nearest-neighbor interpolation



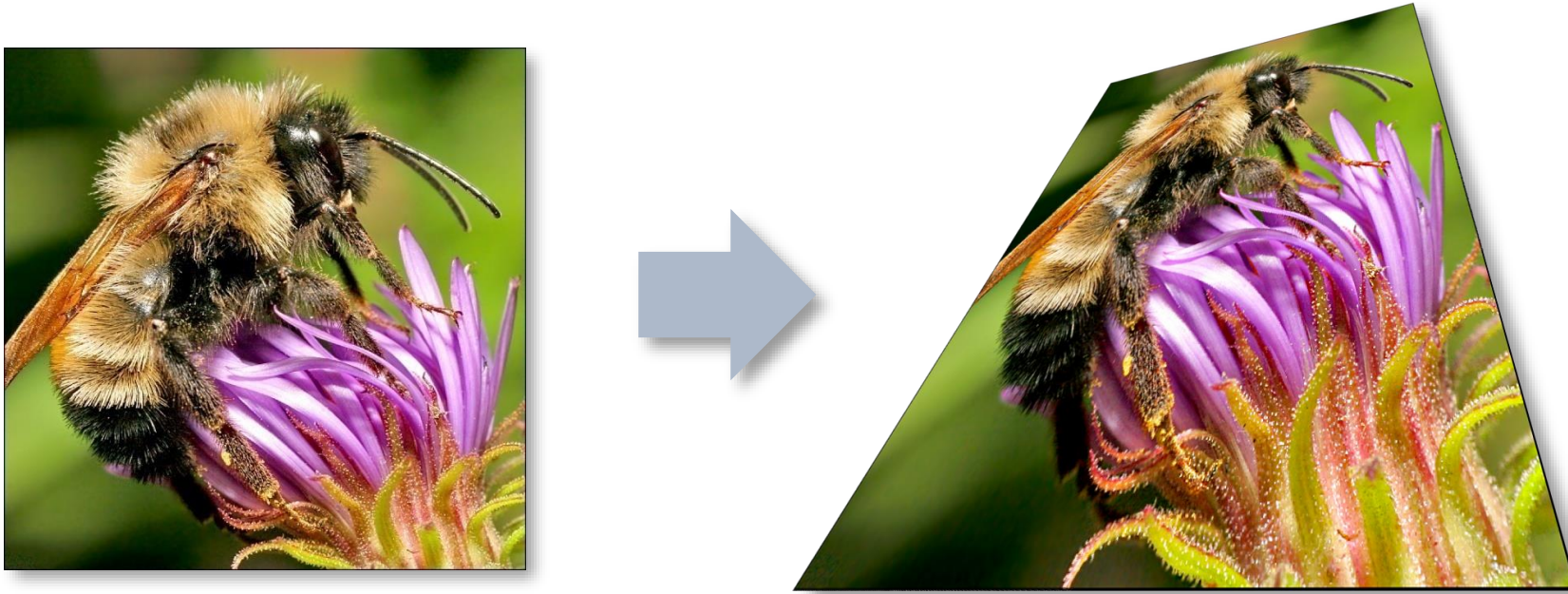
Bilinear interpolation



Bicubic interpolation

Image interpolation

Also used for *resampling*



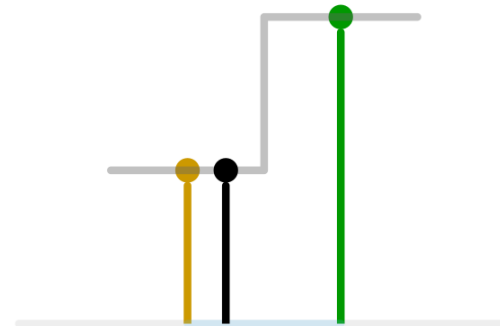
Interpolation

E.g., Bilinear interpolation

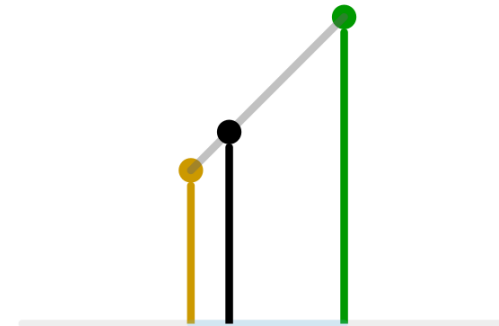
1. Linear interpolation along x-axis

$$f(x, y_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}),$$

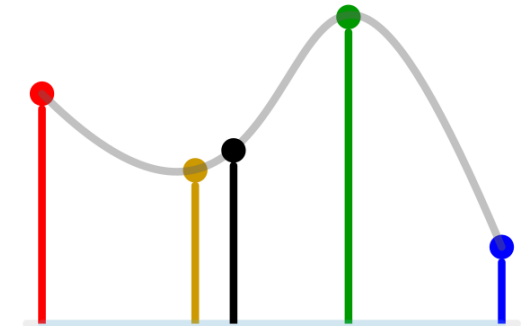
$$f(x, y_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}).$$



1D nearest-neighbour



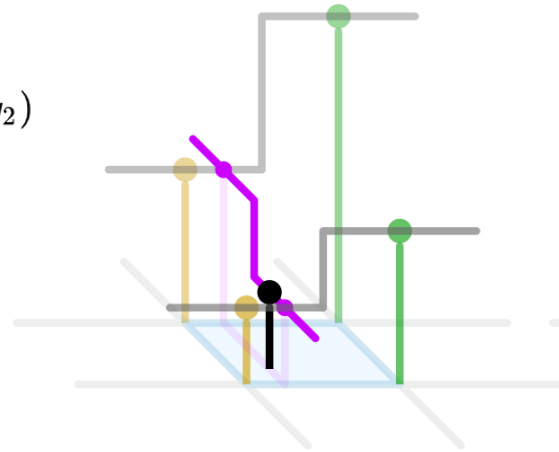
Linear



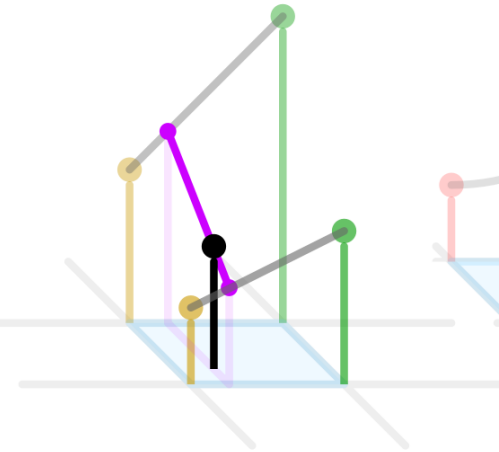
Cubic

2. Linear interpolation along y-axis

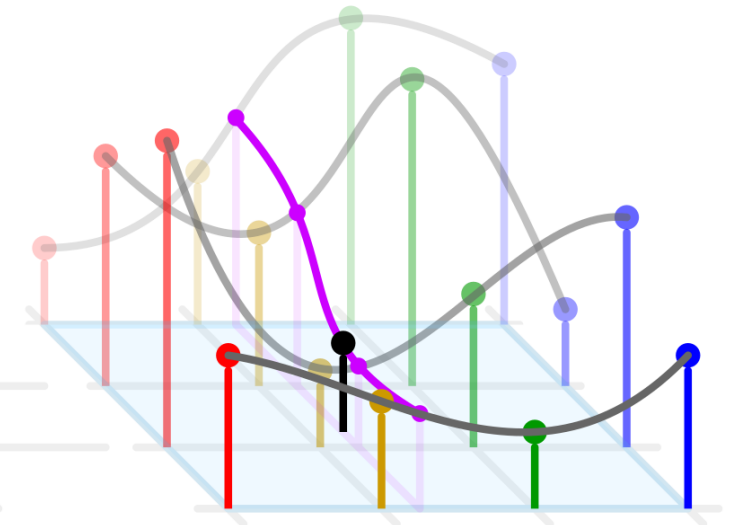
$$f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} f(x, y_1) + \frac{y - y_1}{y_2 - y_1} f(x, y_2)$$



2D nearest-neighbour



Bilinear



Bicubic

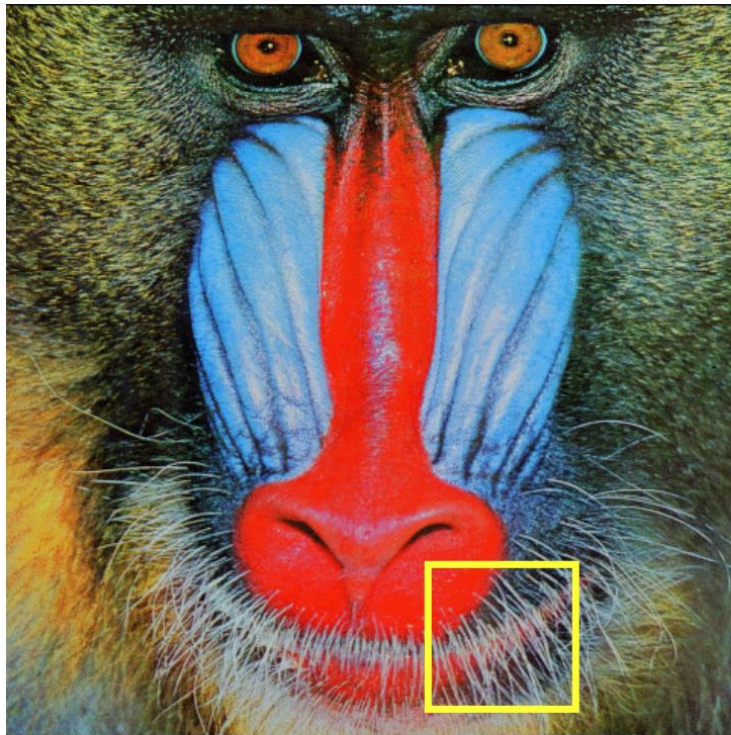
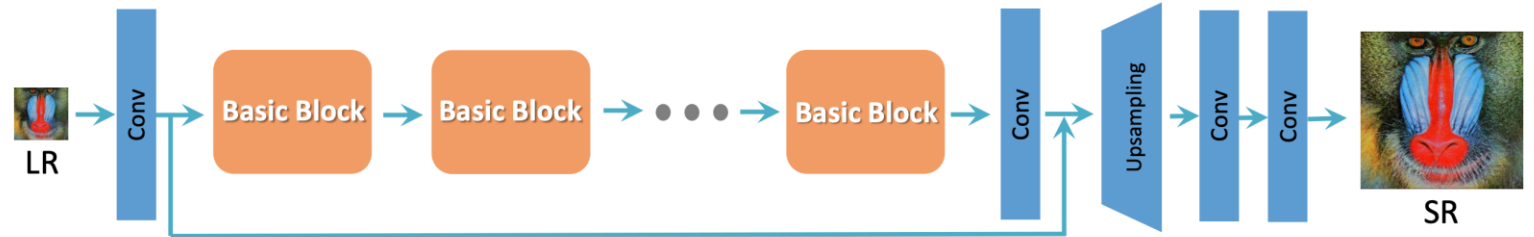
Advanced interpolation

- Depixelating Pixel Art

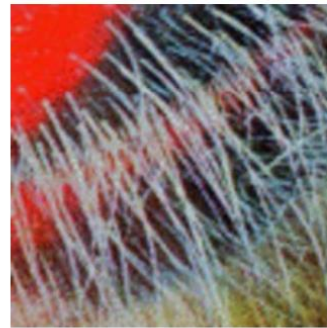


Advanced interpolation

- Super-resolution



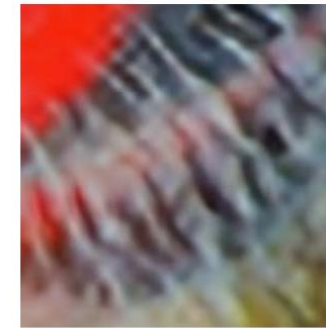
baboon from Set14
(PSNR / Perceptual Index)



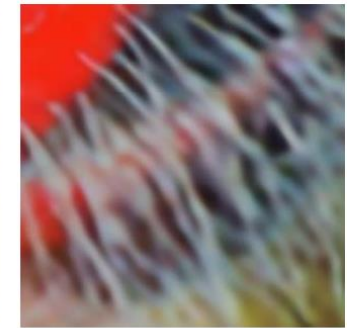
HR
(∞ / 3.59)



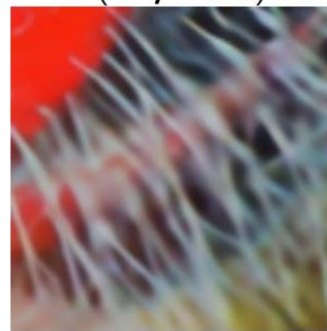
Bicubic
(22.44 / 6.70)



SRCNN
(22.73 / 5.73)



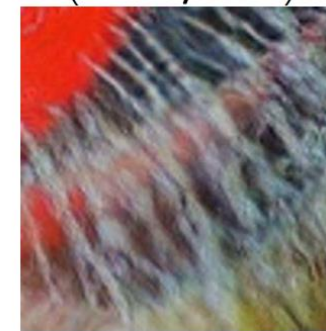
EDSR
(23.04 / 4.89)



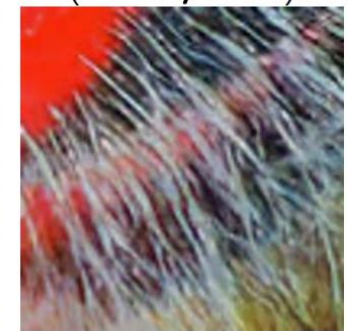
RCAN
(23.12 / 4.20)



EnhanceNet
(20.87 / 2.68)



SRGAN
(21.15 / 2.62)



ESRGAN(ours)
(20.35 / 1.98)

ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks, ECCV workshop 2018