

KECE471 Computer Vision

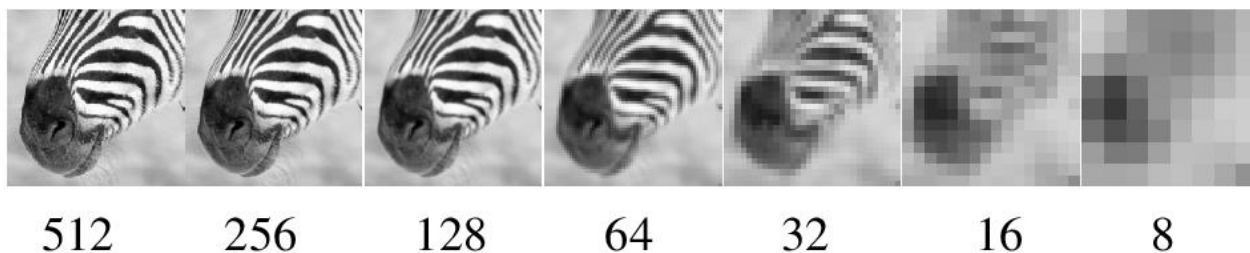
Pyramidal Image Representation

Chang-Su Kim

Sections 7.7 and 9.2, Computer Vision by Forsyth and Ponce

Note: Most contents were extracted from the lecture notes of Prof. Kyoung Mu Lee.

Image Pyramid: Example



- A curve corresponds to
- a hair on the nose in the biggest image
 - a stripe in the medium size image
 - nose itself in the smallest image

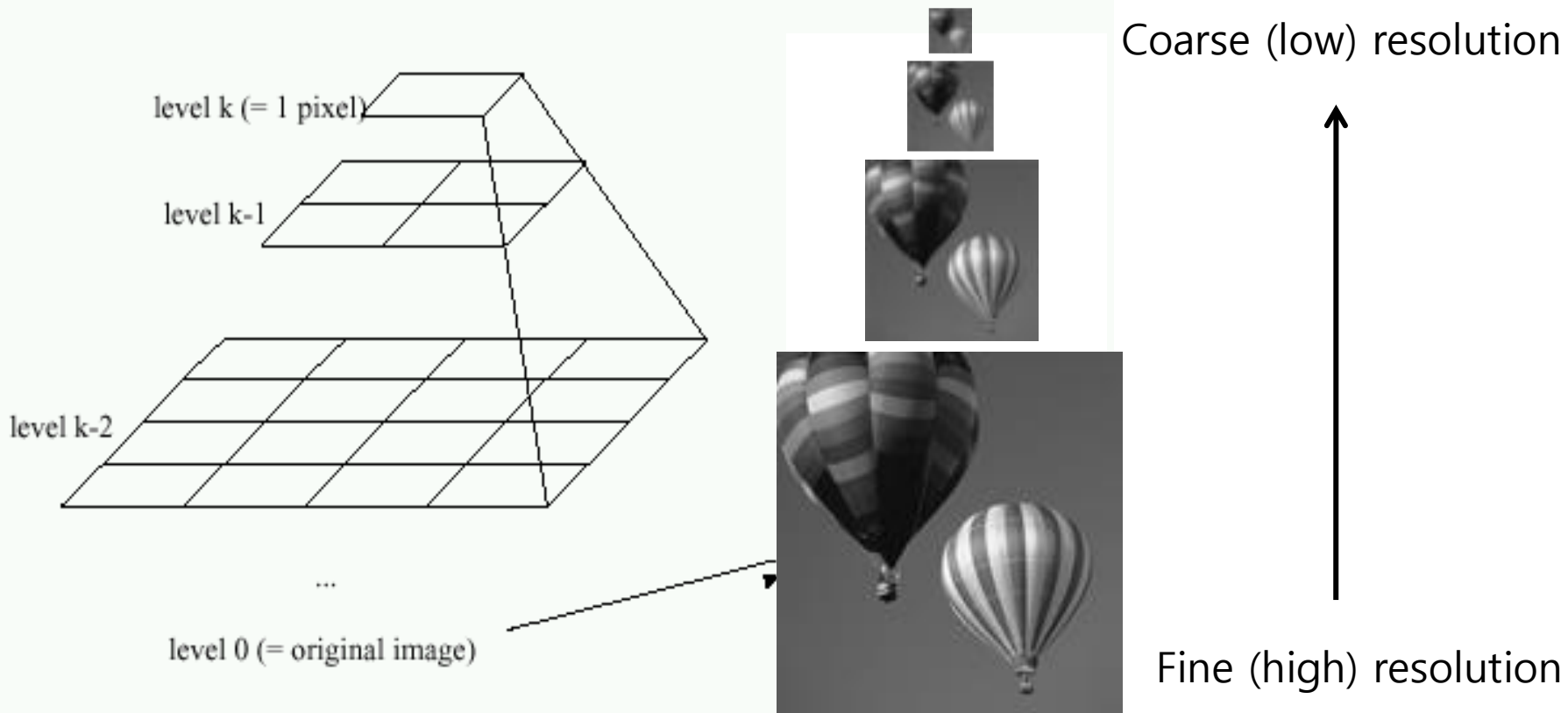


Pyramidal Representation

- Pyramidal representation is a kind of scaled representation
- Both large and small scaled information are interesting
 - Big bars and small bars
 - Stripes and hairs

Image Pyramid (it is not an Egyptian tomb)

Idea: Represent $N \times N$ image as a “pyramid” of $1 \times 1, 2 \times 2, 4 \times 4, \dots, 2^k \times 2^k$ images (assuming $N = 2^k$)



Aliasing

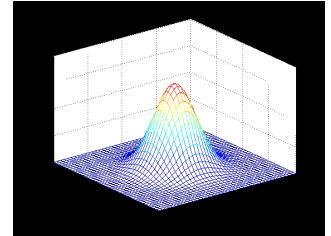
- Lowpass filtering is required before downsampling to avoid aliasing
- Anti-aliasing filtering
 - A Gaussian filter is often used



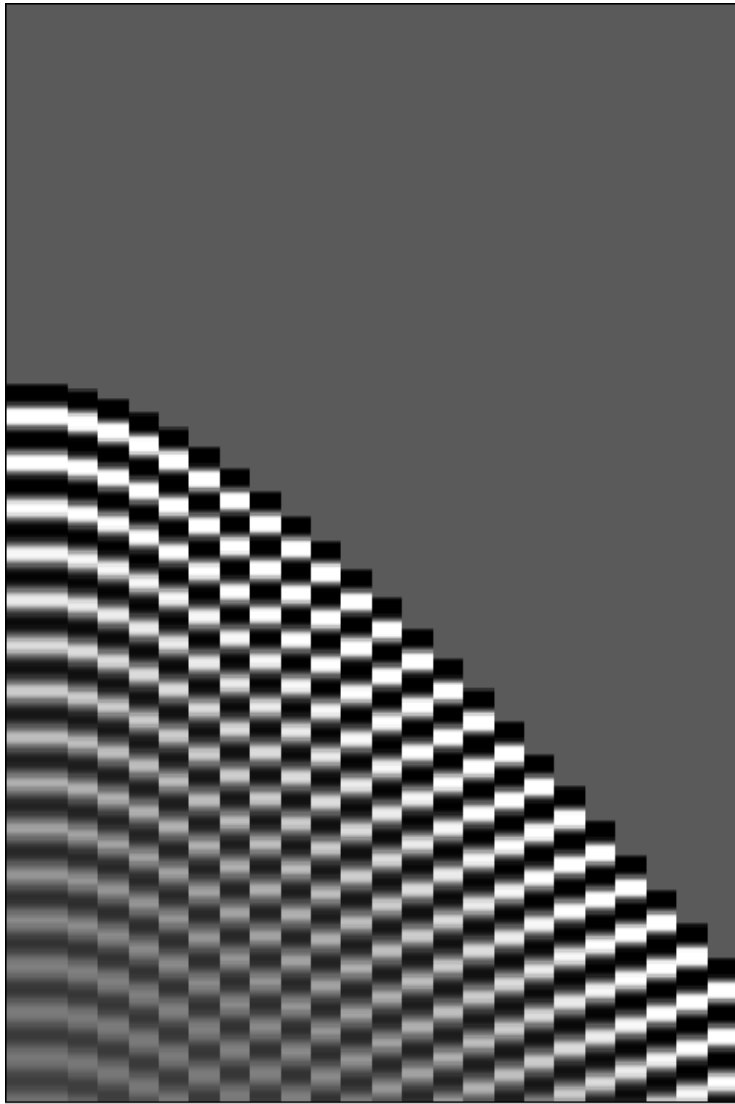
Without anti-aliasing filtering

Aliasing

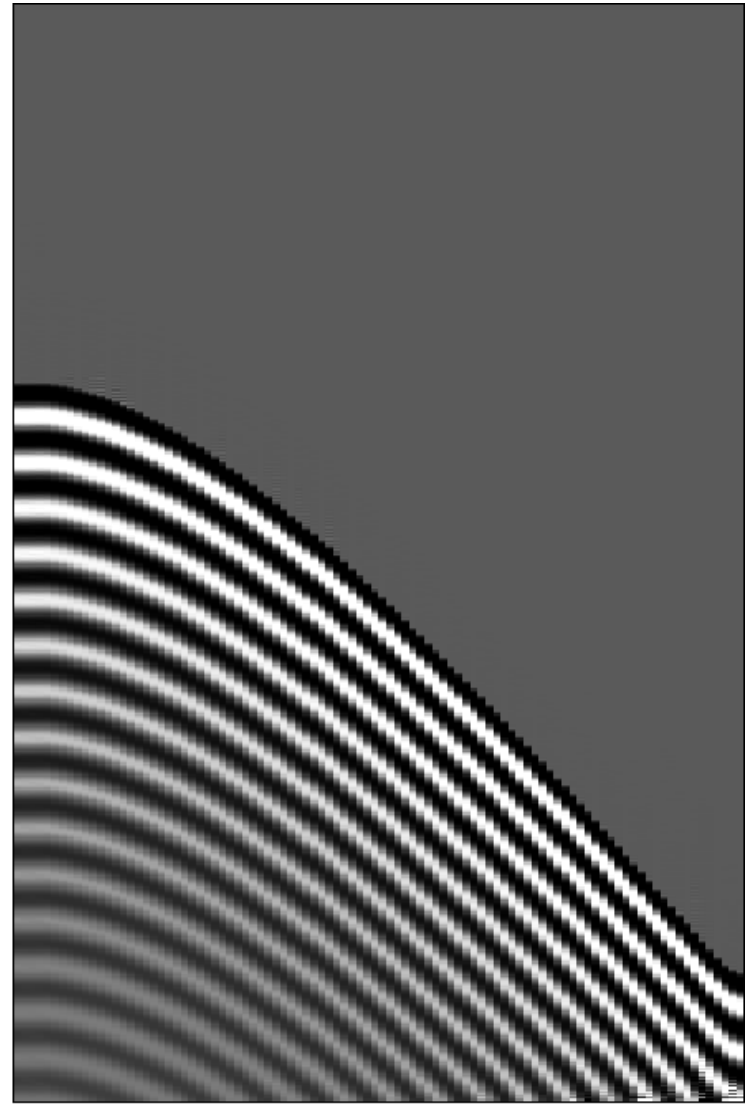
- Lowpass filtering is required before subsampling to avoid aliasing
- Anti-aliasing filtering
 - A Gaussian filter is often used



With anti-aliasing filtering



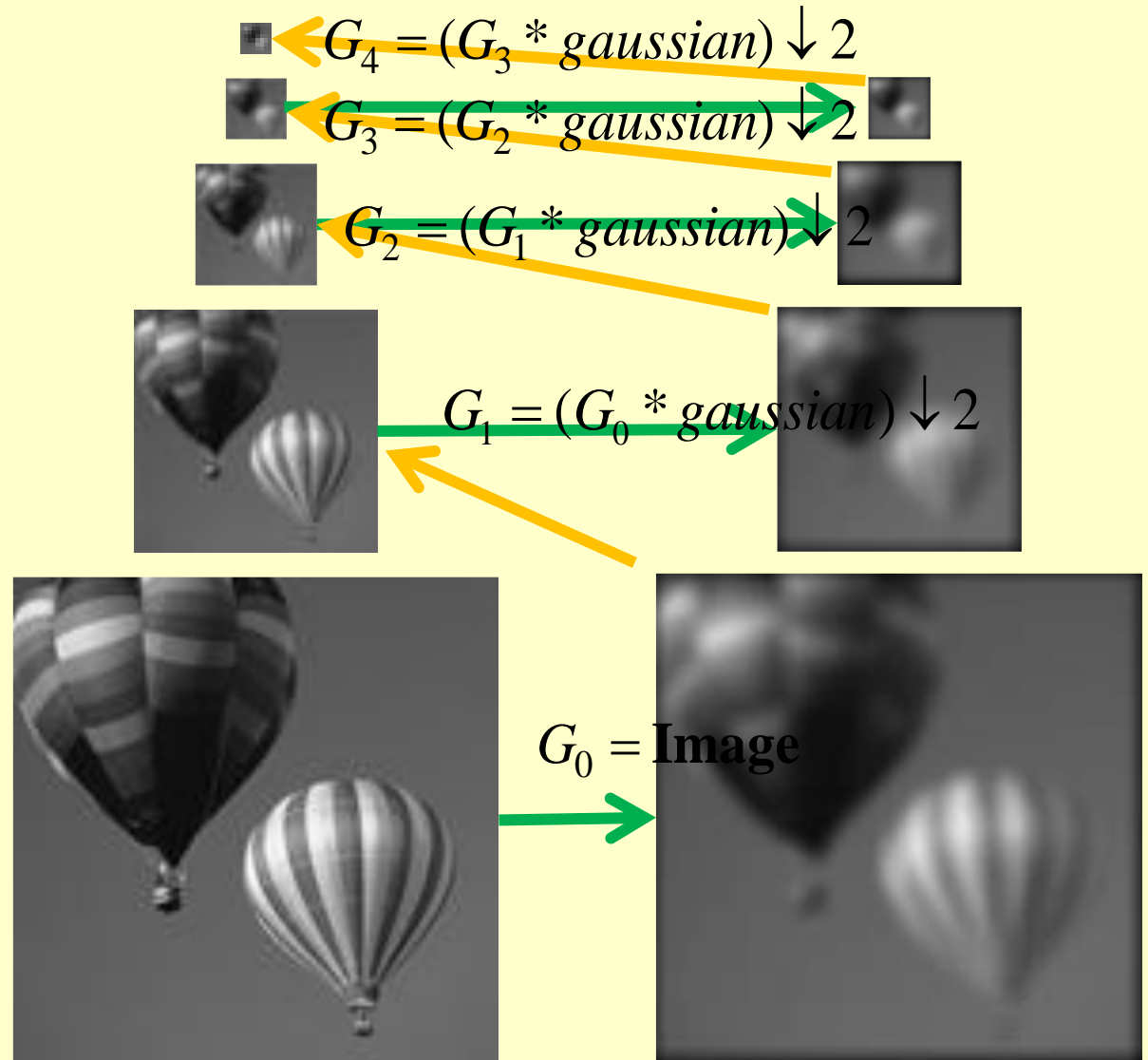
Aliasing



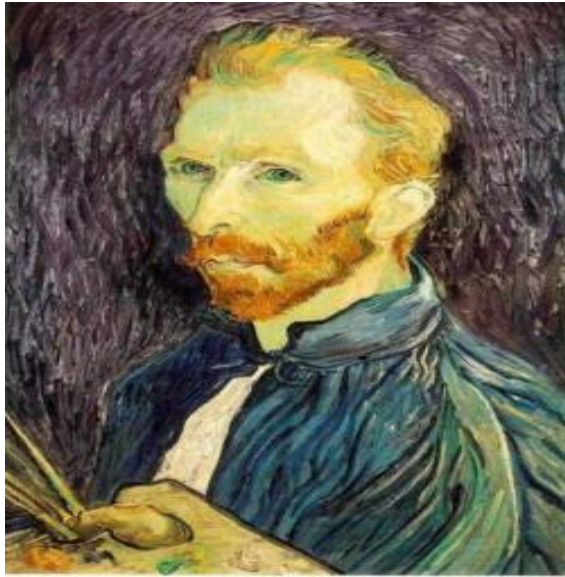
Anti-aliased

Gaussian Pyramid

-  Gaussian filtering
-  Downsampling



Construction of a Gaussian Pyramid



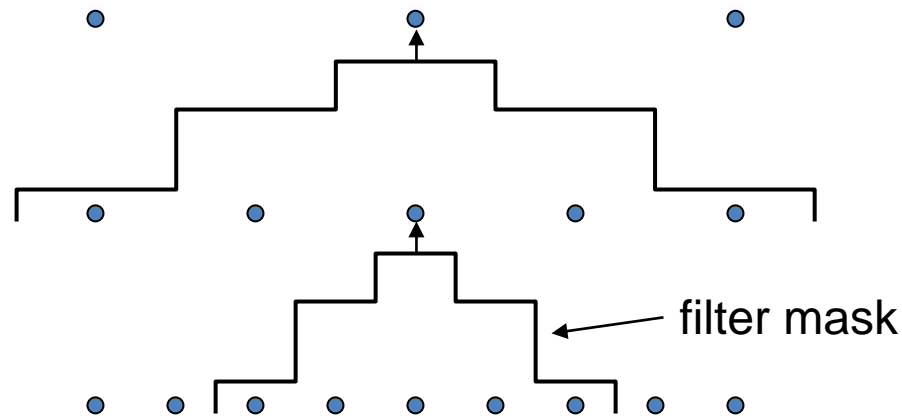
Gaussian 1/2



G 1/4

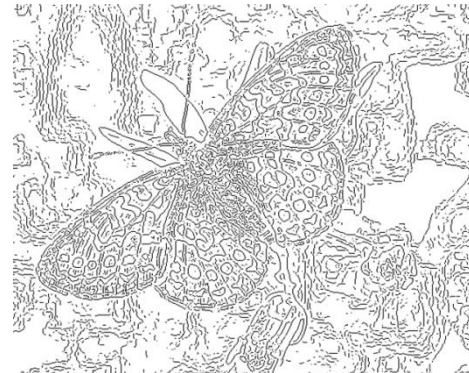


G 1/8

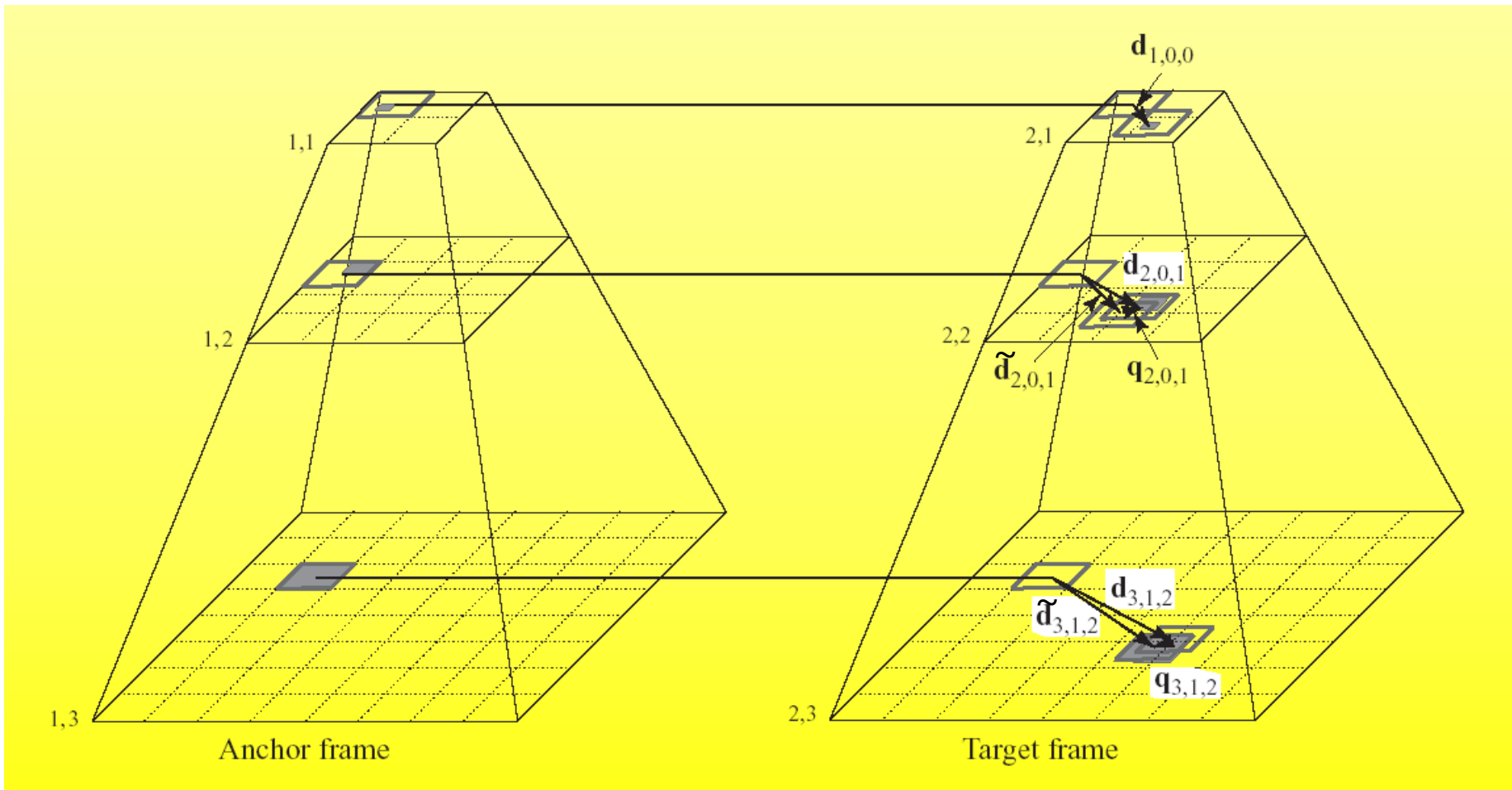


Applications of Gaussian Pyramids

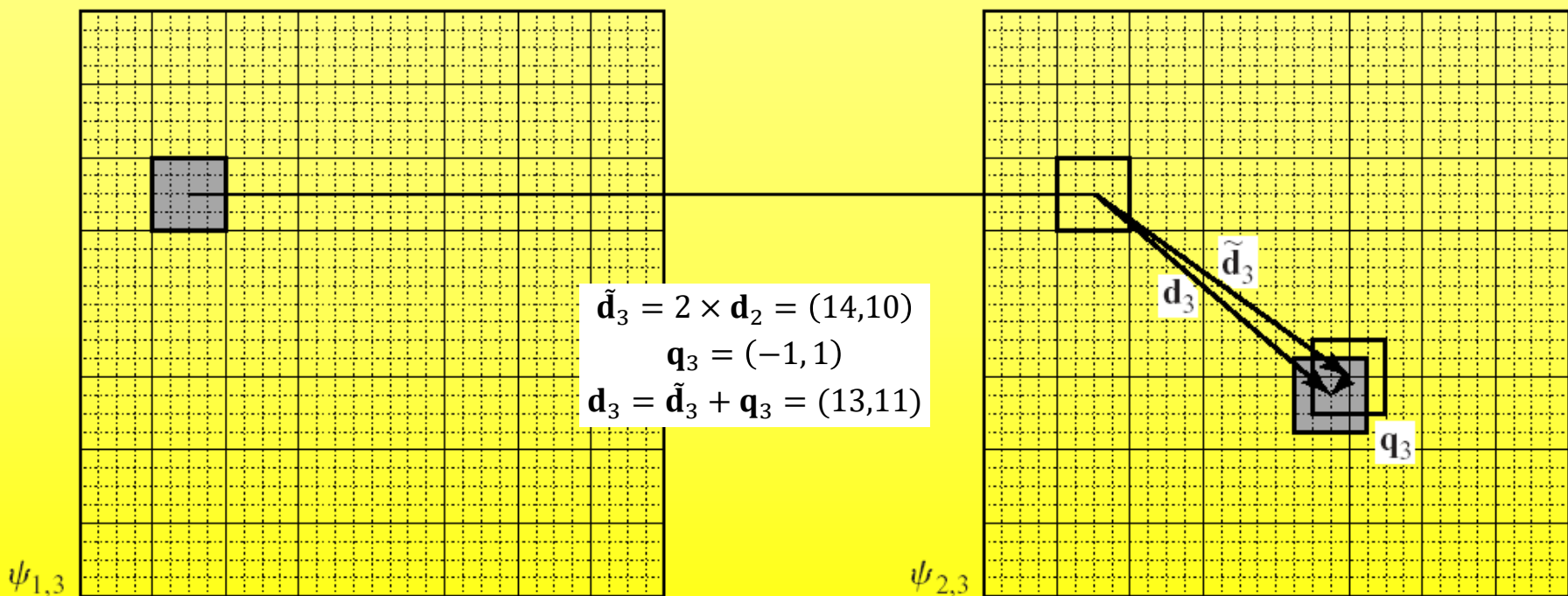
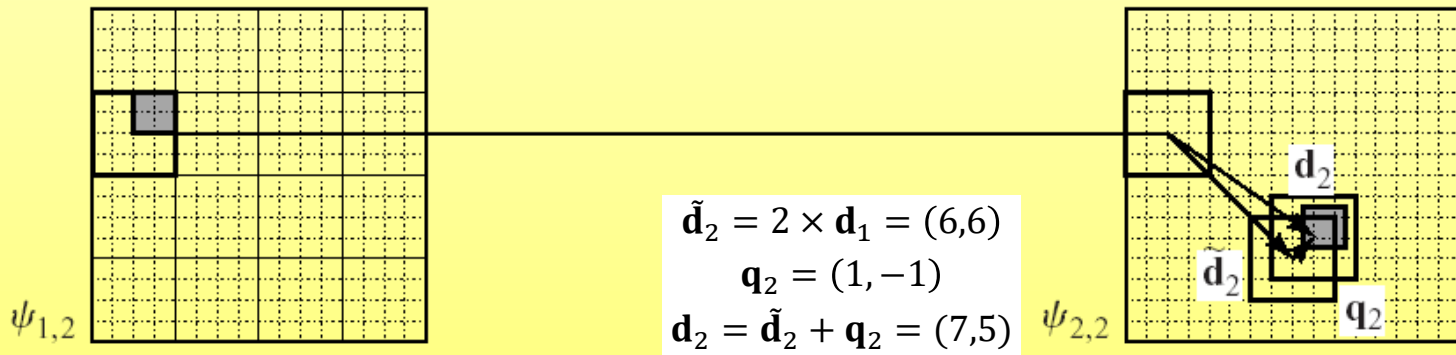
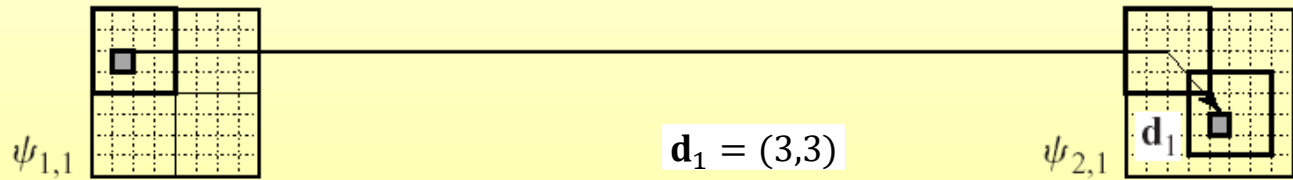
- Search for correspondence
 - look at coarse scales, then refine with finer scales
- Edge tracking
 - a “good” edge at a fine scale has parents at a coarser scale
- Template matching
 - e.g. Detecting faces



Hierarchical Block Matching



- Lower resolution motion vector is used to predict higher resolution motion vector (e.g. $d_{2,0,1}$ is used to predict $d_{3,1,2}$)
 - Reduction of computational complexity
 - More reliable motion vector estimation



Anchor frame

Target frame

Non-Hierarchical Block Matching Algorithm

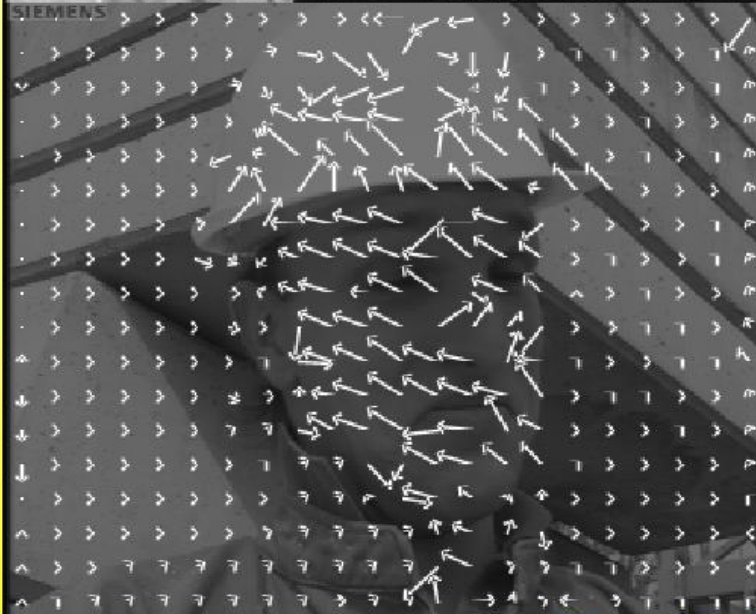
target frame



anchor frame



Motion field

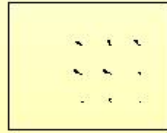


Predicted anchor frame (29.86dB)



Example: Half-pel EBMA

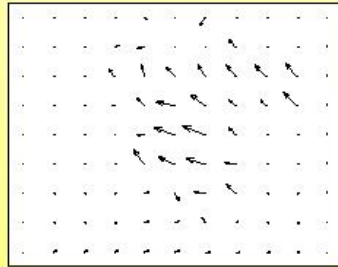
Hierarchical Block Matching Algorithm



(a)



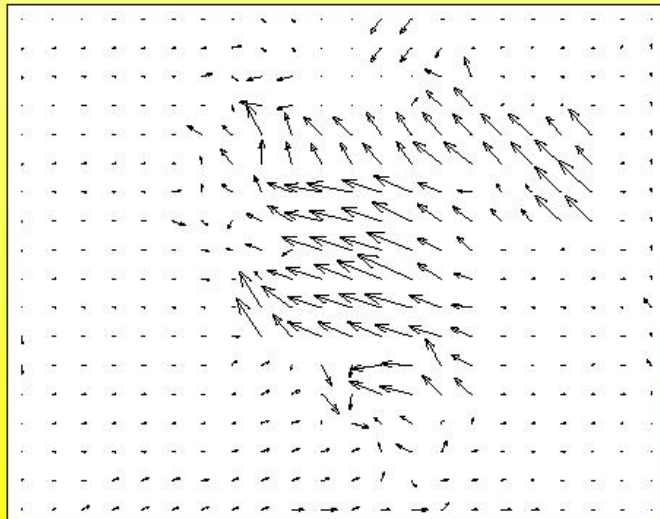
(b)



(c)



(d)



(e)

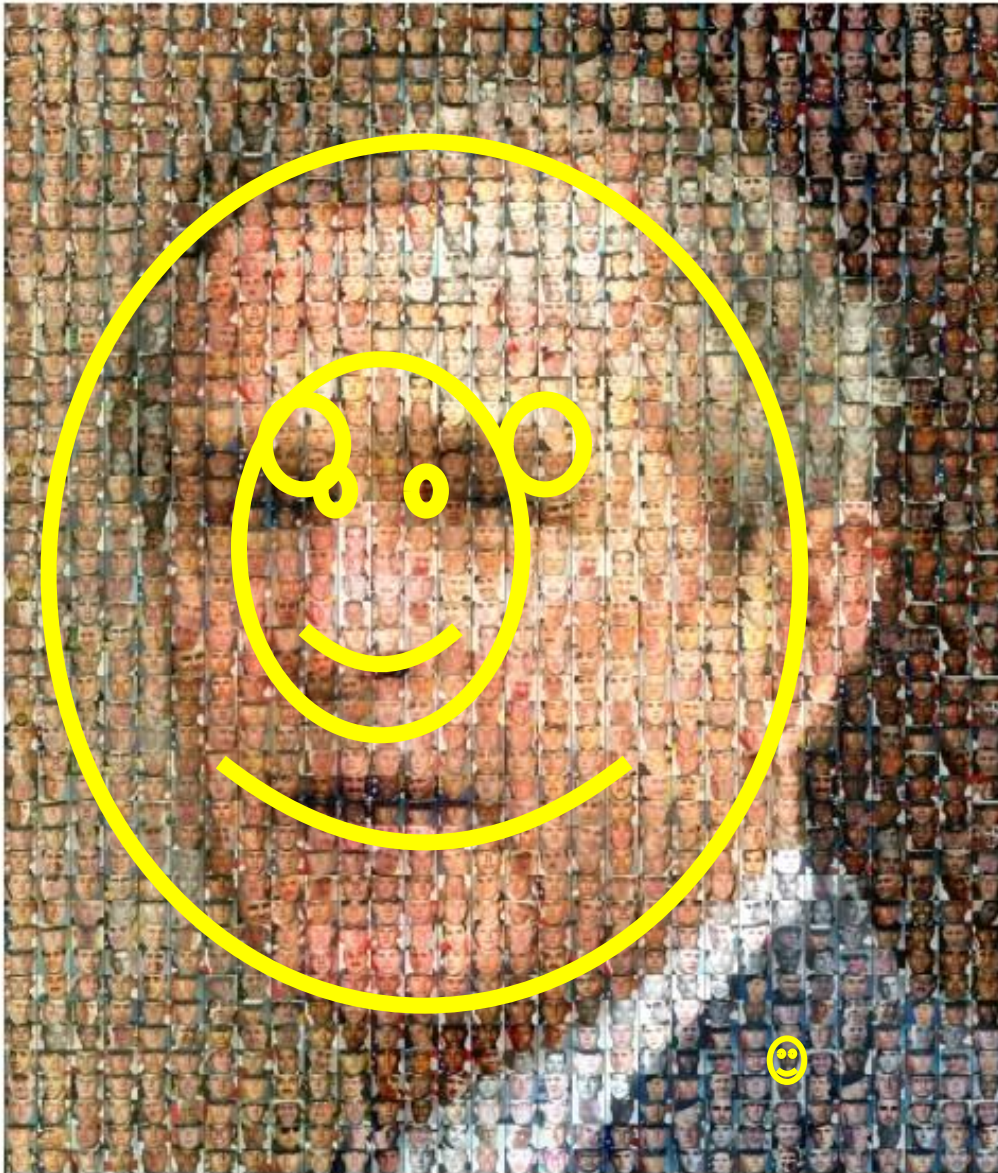


(f)

Predicted anchor frame (29.32dB)

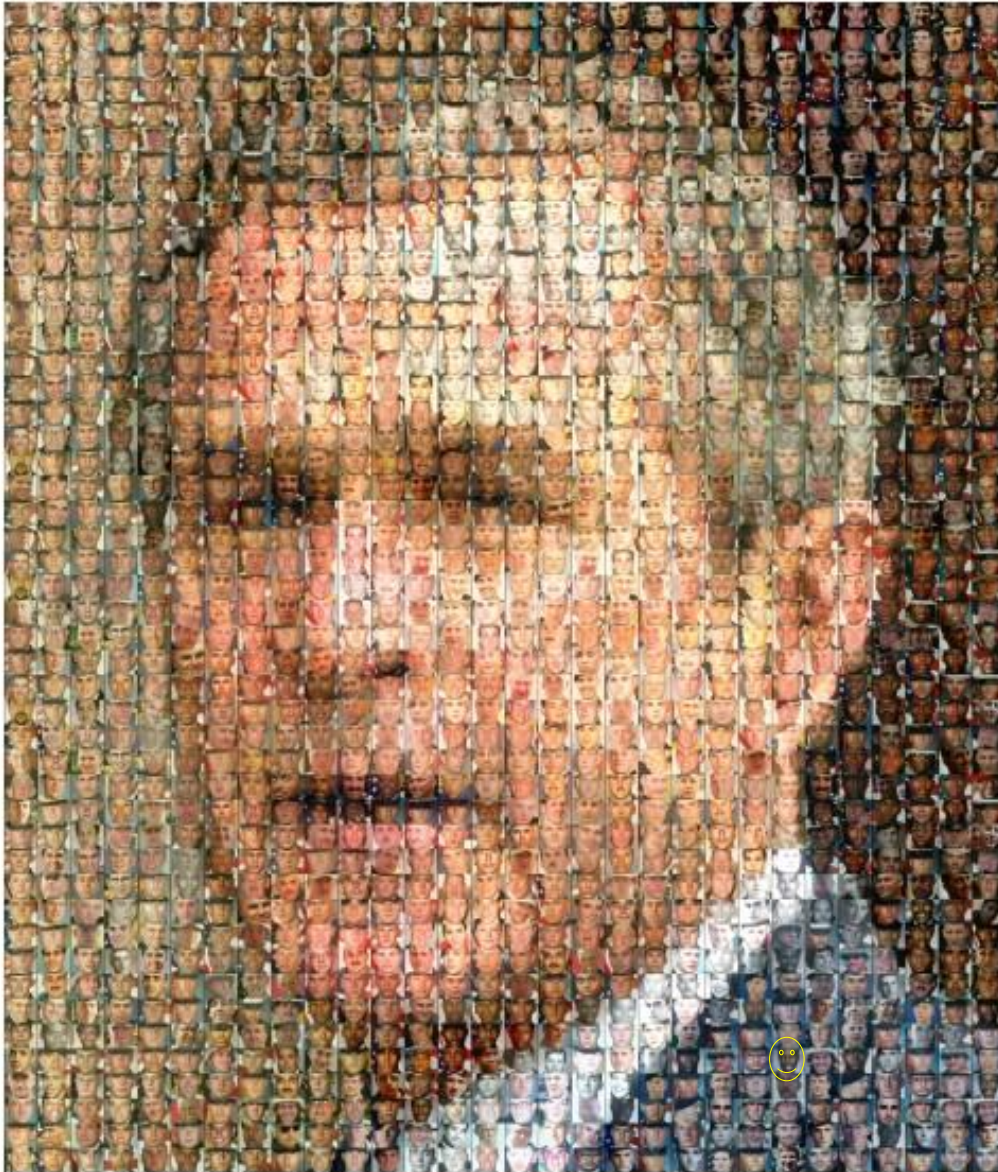
Example: Three-level HBMA

Template Matching

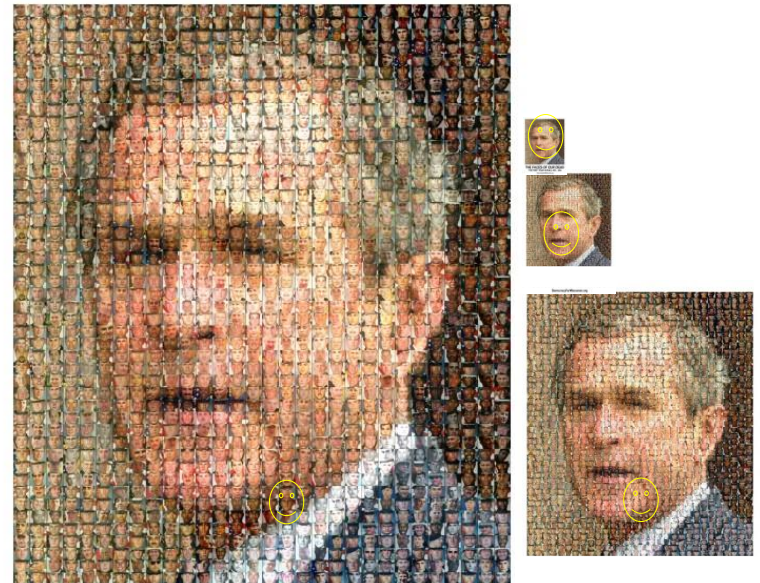


- Strategy 1
 - Use templates of different sizes
 - For large templates, matching is costly

Template Matching



- Strategy 2
 - Apply a fixed-size template to the Gaussian pyramid



Laplacian Pyramid

- It removes redundancies in Gaussian Pyramid
- Similar to edge images
- Most pixels are zero
- It can be used in point detection and image compression

Laplacian Pyramid

- Gaussian Pyramid

- G_0

- $G_1 = D(G_0)$

- $G_2 = D(G_1)$

- $G_3 = D(G_2)$

- D

- Gaussian filtering

- then

- Downsampling

- Laplacian Pyramid

- $L_0 = G_0 - U(G_1)$

- $L_1 = G_1 - U(G_2)$

- $L_2 = G_2 - U(G_3)$

- $L_3 = G_3$

- U

- Upsampling

Laplacian Pyramid

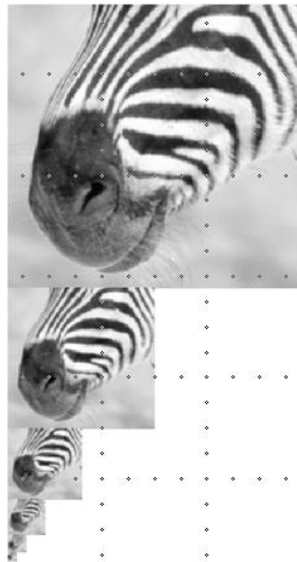
- Gaussian Pyramid

- G_0

- $G_1 = D(G_0)$

- $G_2 = D(G_1)$

- $G_3 = D(G_2)$



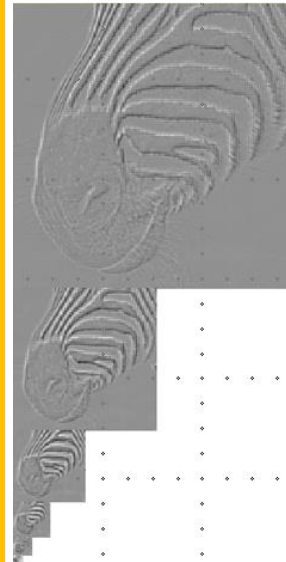
- Laplacian Pyramid

- $L_0 = G_0 - U(G_1)$

- $L_1 = G_1 - U(G_2)$

- $L_2 = G_2 - U(G_3)$

- $L_3 = G_3$



Laplacian Pyramid

- Analysis

- $L_0 = G_0 - U(G_1)$

- $L_1 = G_1 - U(G_2)$

- $L_2 = G_2 - U(G_3)$

- $L_3 = G_3$

- Synthesis

- $G_0 = L_0 + U(G_1)$

- $G_1 = L_1 + U(G_2)$

- $G_2 = L_2 + U(G_3)$

- $G_3 = L_3$

Gaussian Pyramid



512

256

128

64

32

16

8



Laplacian Pyramid



512

256

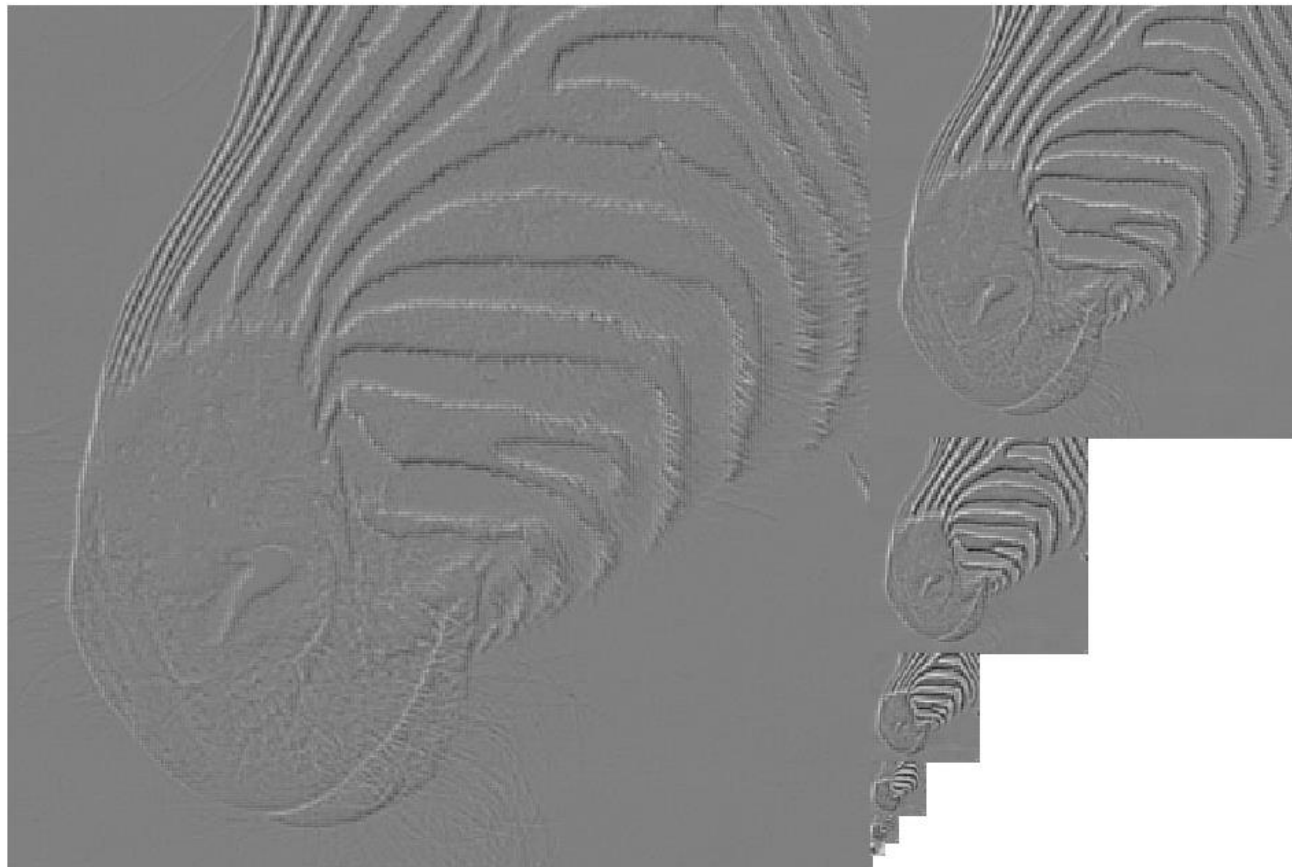
128

64

32

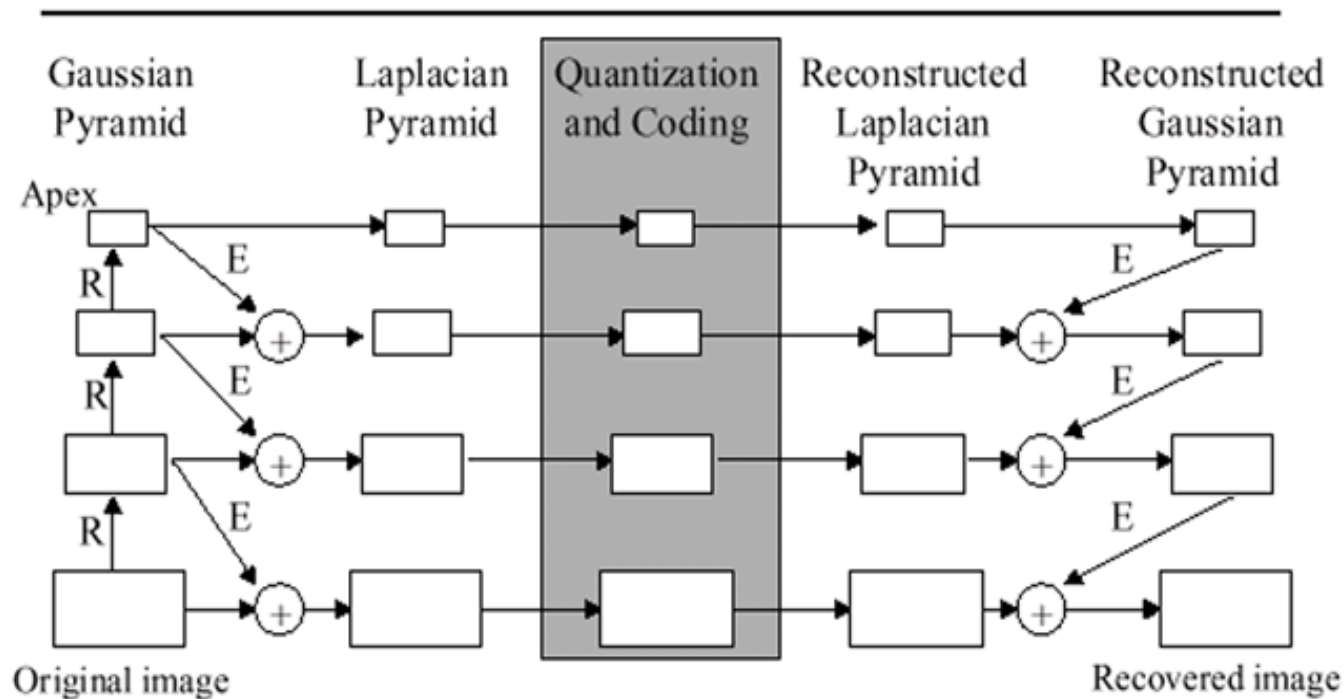
16

8



Laplacian Pyramid for Compression

The Laplacian Pyramid



Supplementary Video

Geometric Order Learning for Rank Estimation

Anonymous NeurIPS 2022 submission