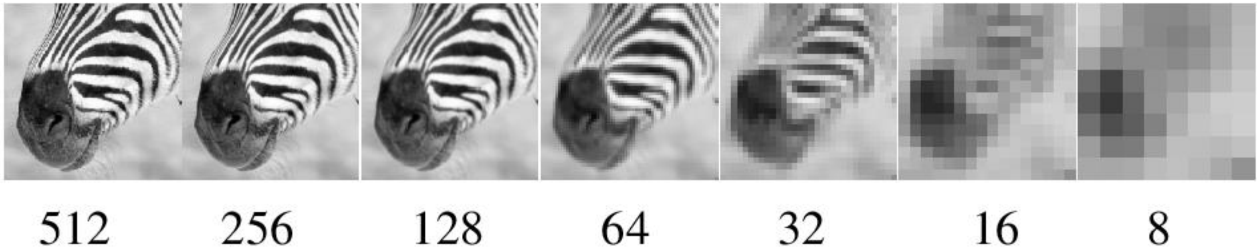
KECE471 Computer Vision

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

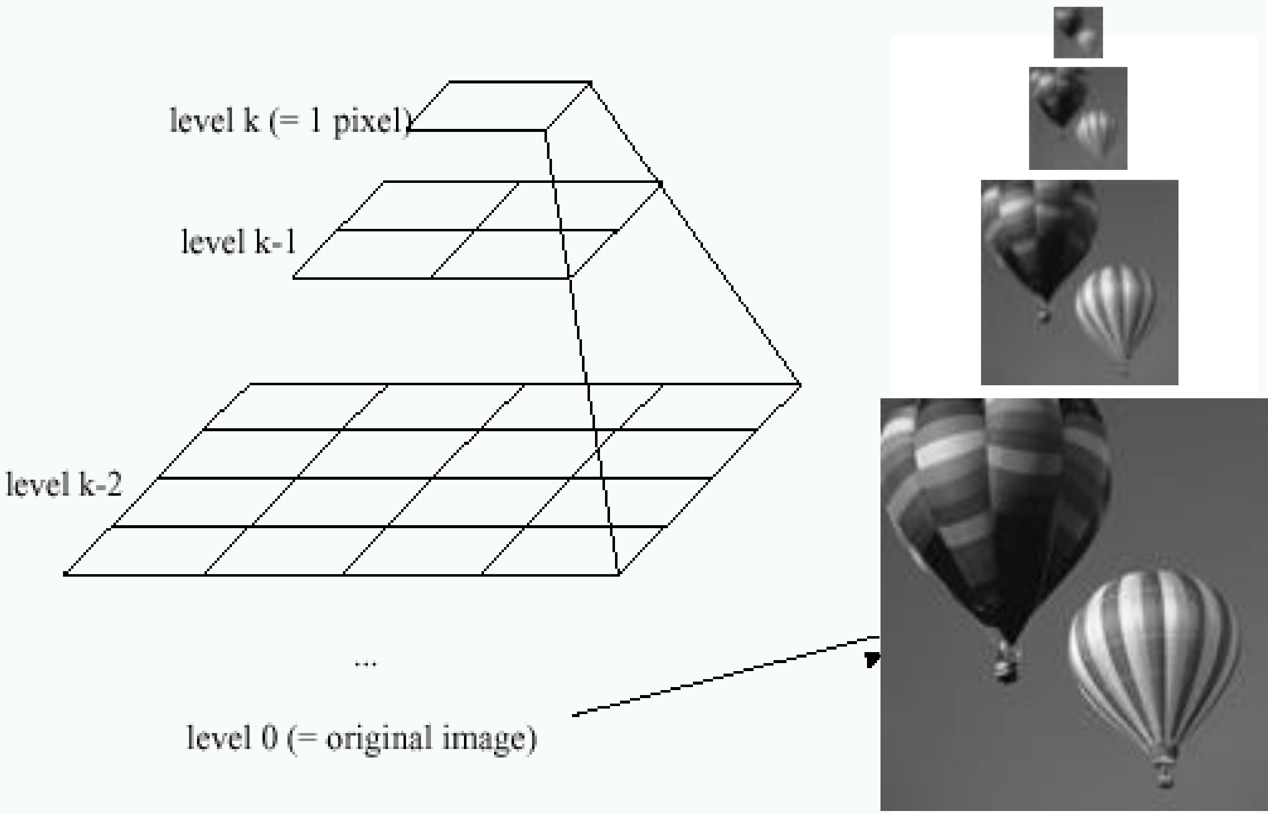

Pyramidal Representation

- 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 NxN image as a "pyramid" of $1\text{x}1$, $2\text{x}2$, $4\text{x}4$,..., $2^k\text{x}2^k$ images (assuming $N=2^k$)



Coarse (low) resolution

Fine (high) resolution

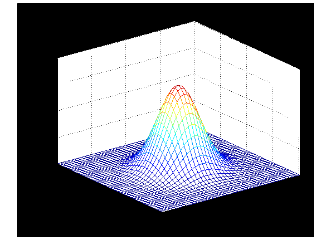
Aliasing

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

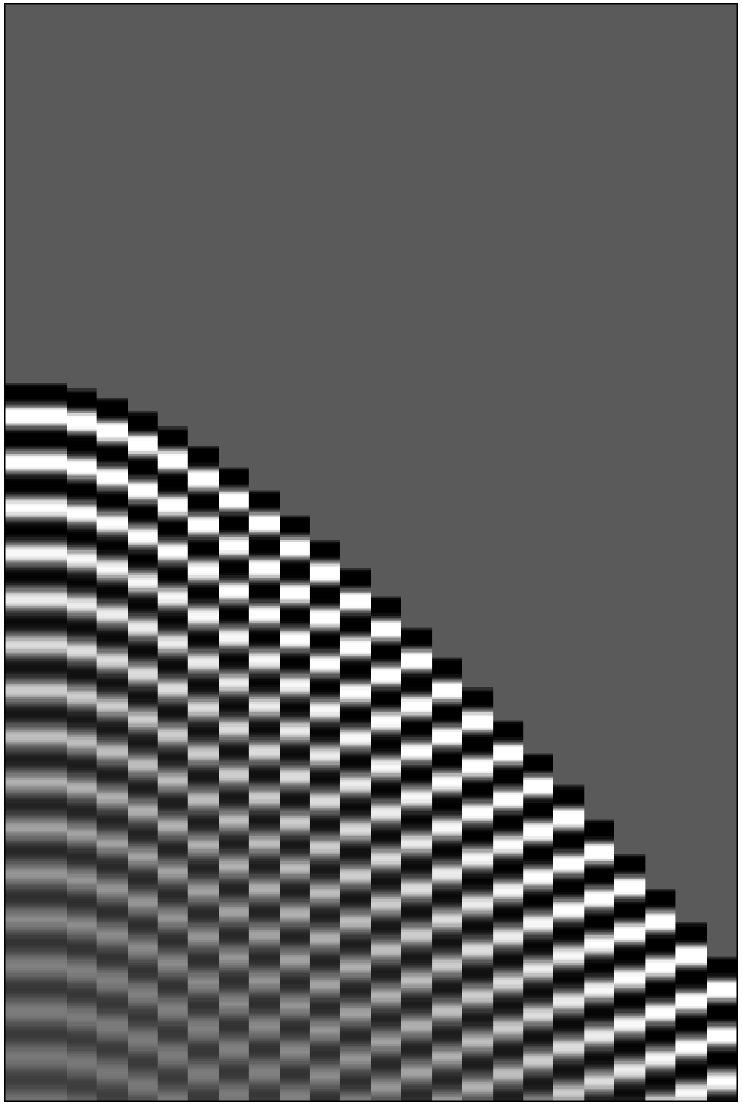
Without anti-aliasing filtering

Aliasing

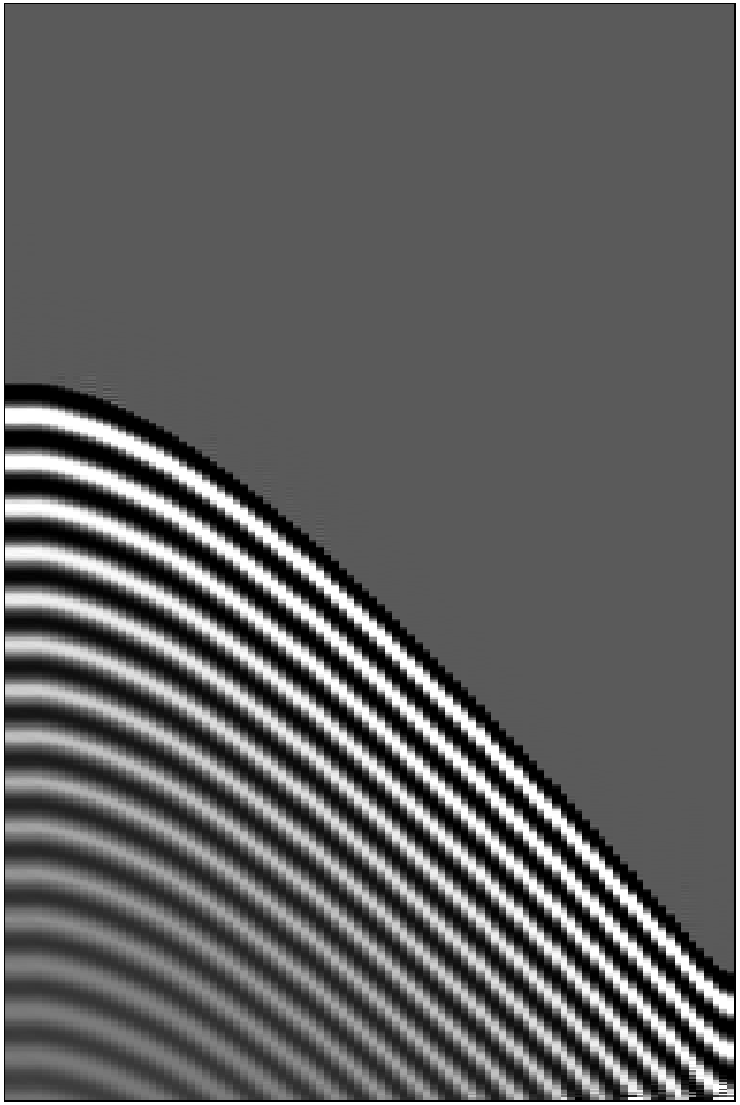
- Lowpass filtering is required before subsampling to avoid aliasing
- Anti-aliasing filtering
  - A Gaussian filter is often used



With anti-aliasing filtering



Aliasing

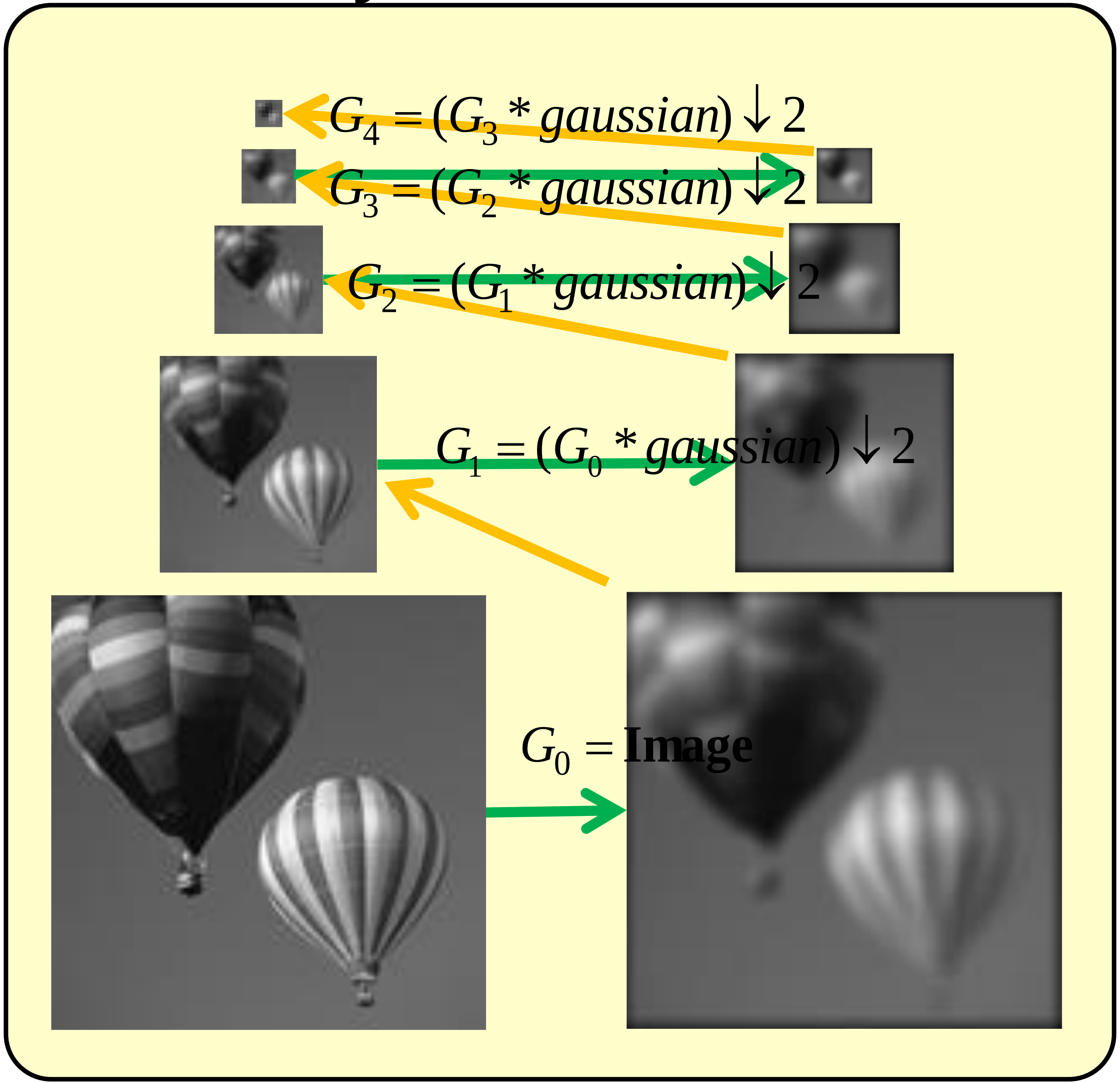


Anti-aliased

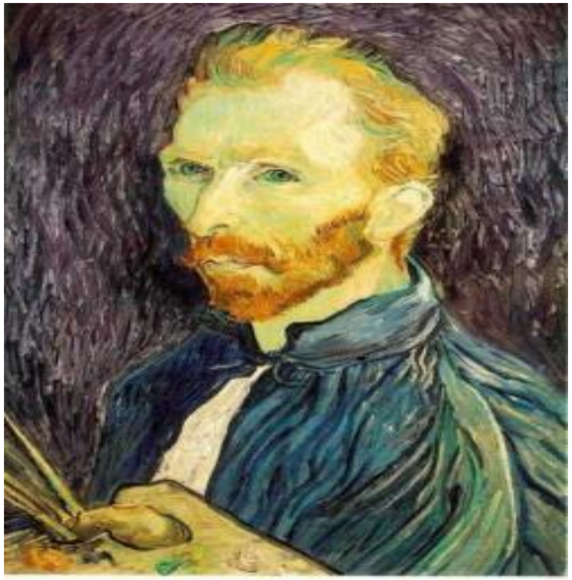
Gaussian Pyramid

Gaussian filtering

Downsampling



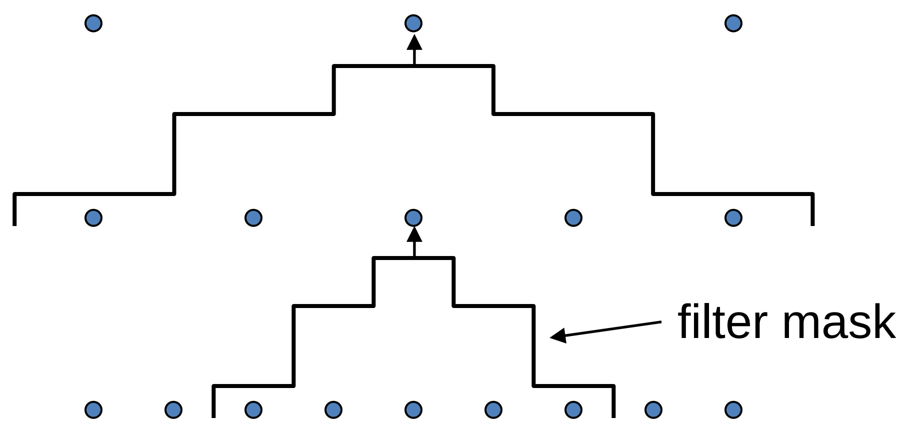
Construction of a Gaussian Pyramid



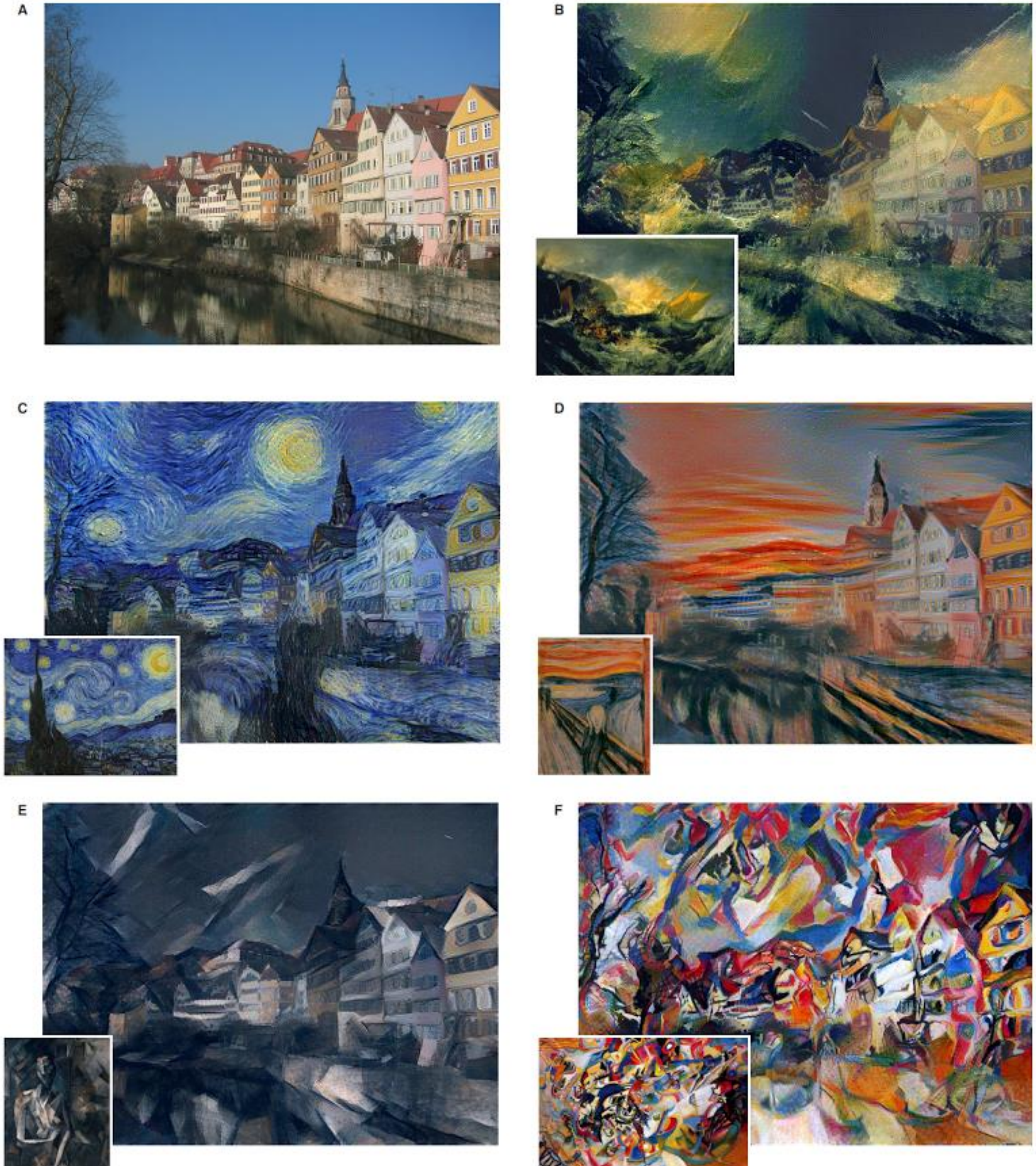
Gaussian 1/2

G 1/4

G 1/8

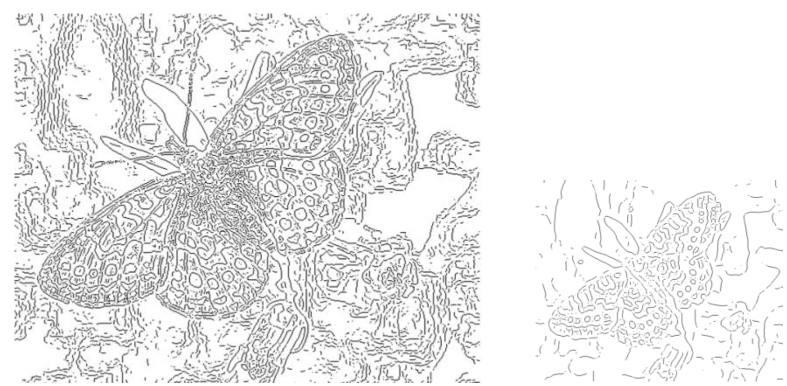


http://web.mit.edu/persci/people/adelson/pub_pdfs/pyramid83.pdf

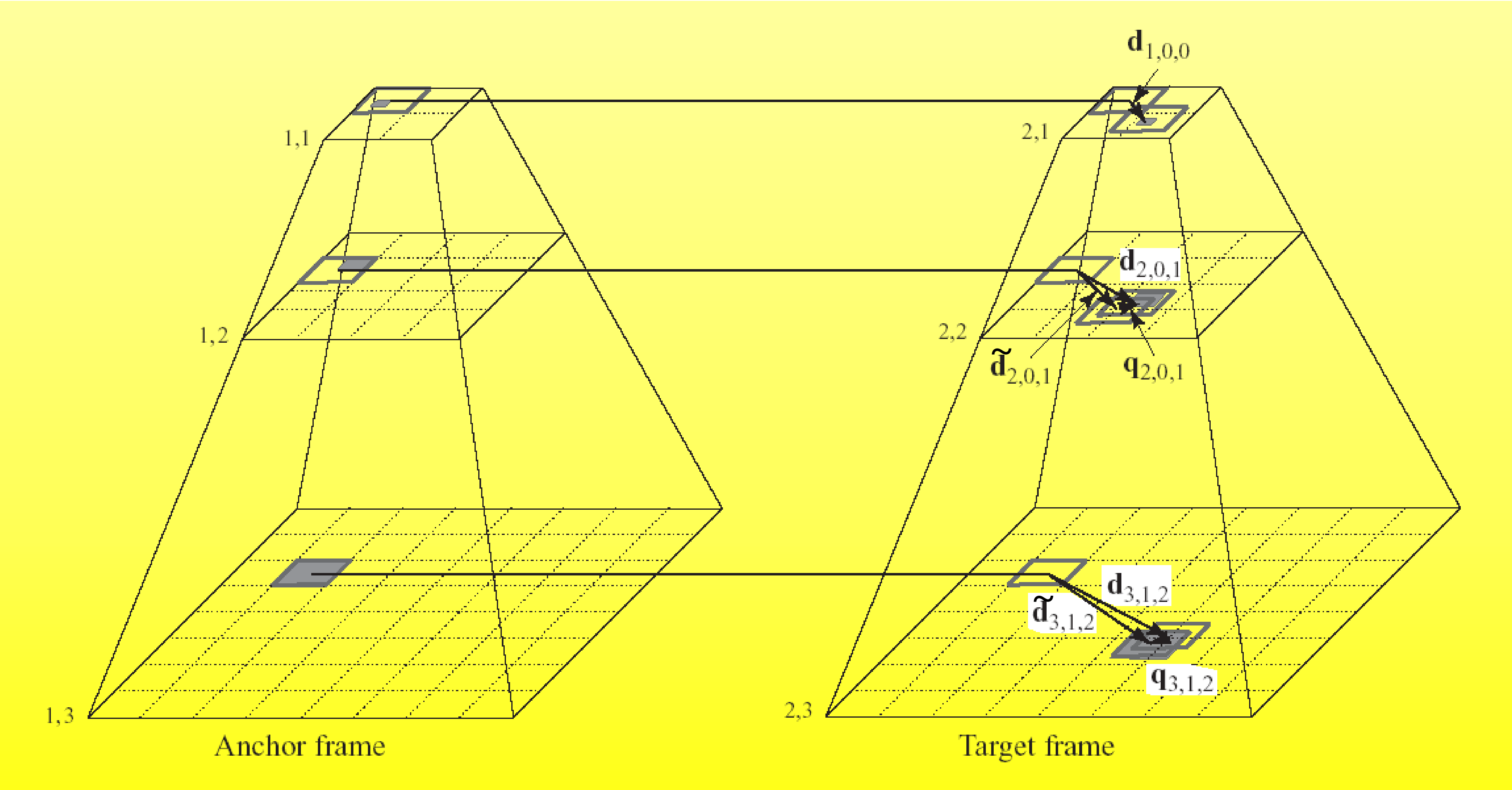

A
B
C
D
E
F

Applications of Gaussian Pyramids

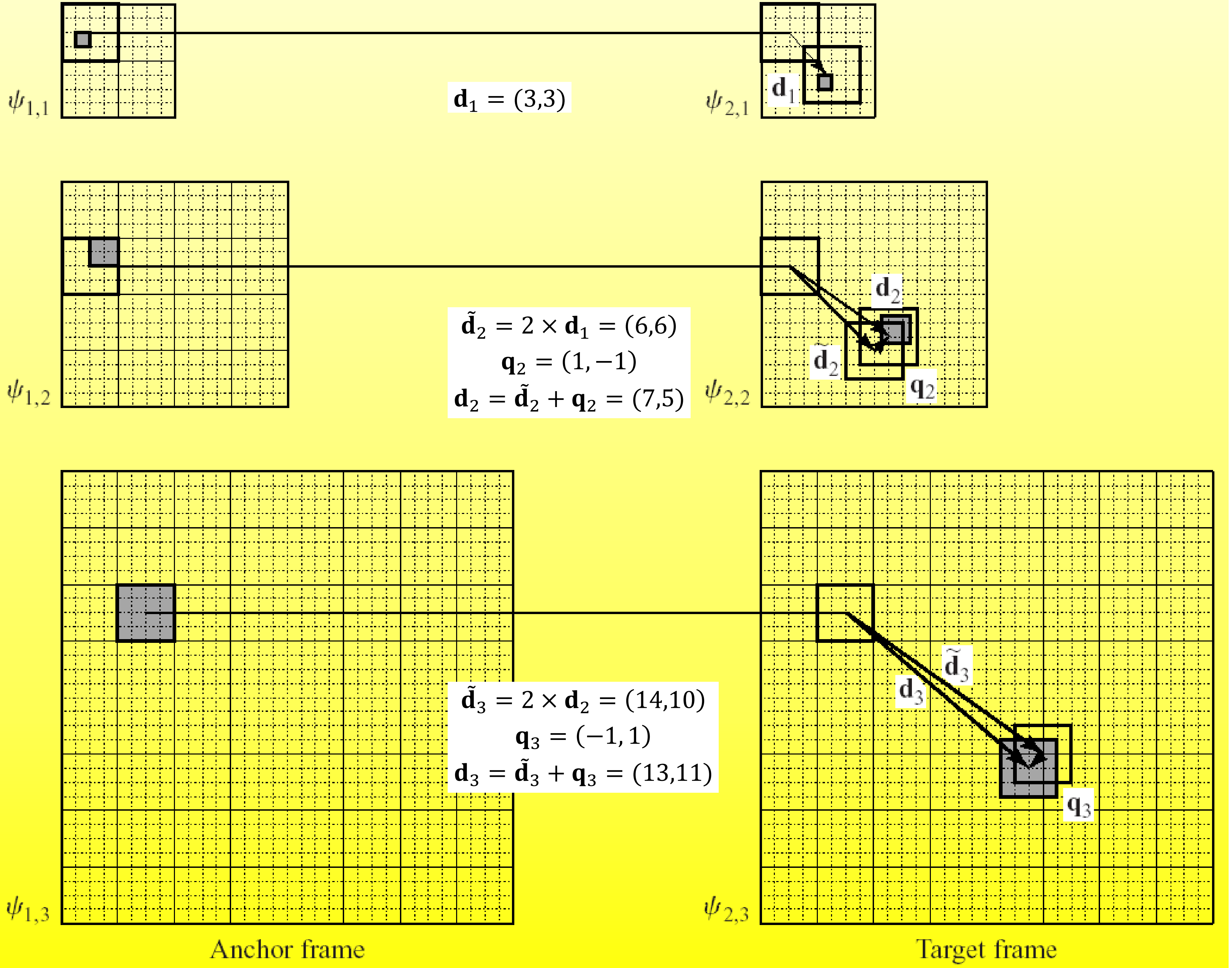
- Search for correspondence
  - look at coarse scales, then refine with finer scales
- Edge tracking
  - a "good" edge at a fine scale has parents at a coarser scale
- Template matching
  - e.g. Detecting faces



Hierarchical Block Matching

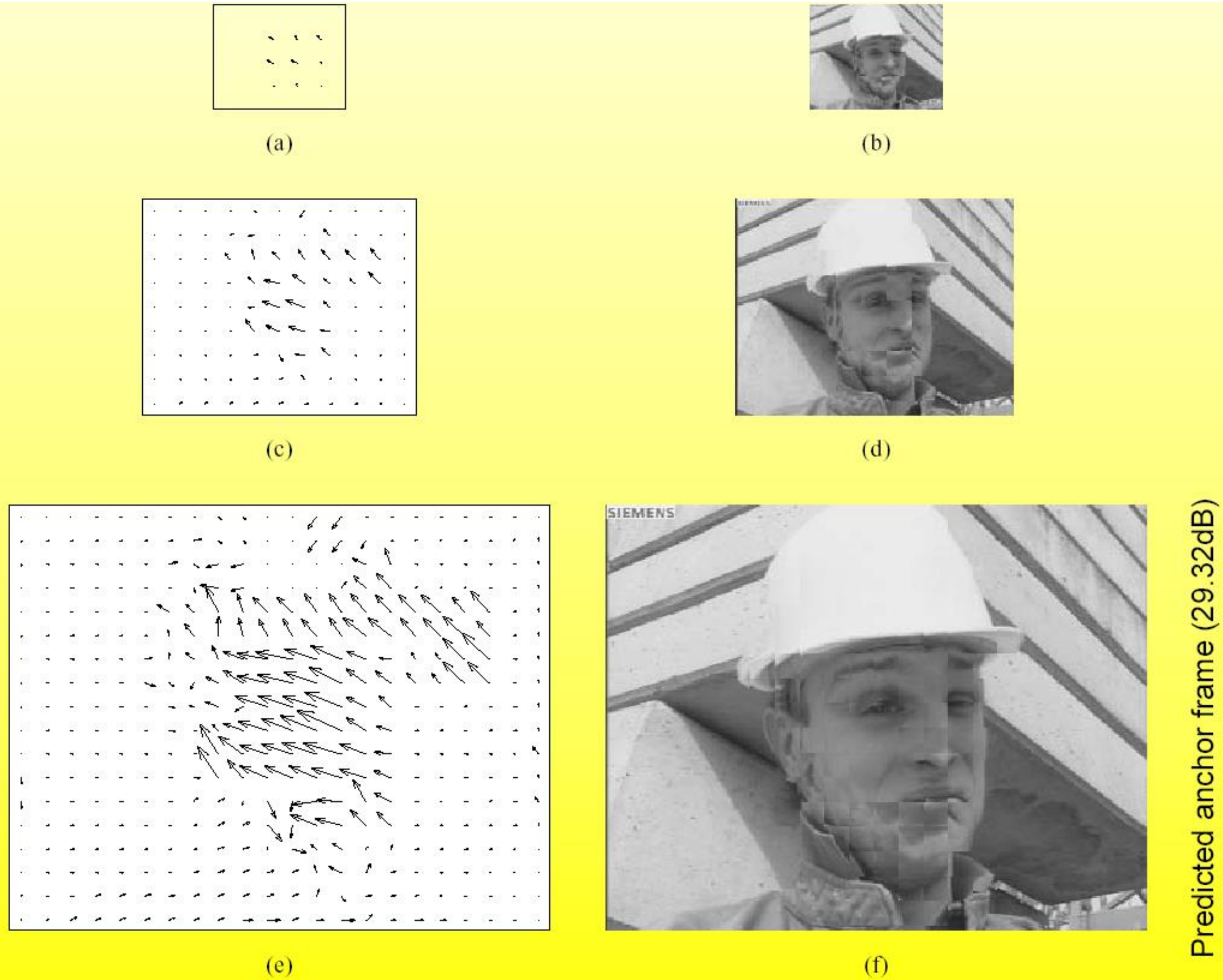


- Lower resolution motion vector is used to predict higher resolution motion vector (e.g. $d_{2,0,1}$ is used to predict $d_{3,1,2}$)
  - Reduction of computational complexity
  - More reliable motion vector estimation


$\psi_{1,1}$
$\mathbf{d}_1 = (3,3)$
$\psi_{2,1}$
$\mathbf{d}_1$
$\psi_{1,2}$
$\tilde{\mathbf{d}}_2 = 2 \times \mathbf{d}_1 = (6,6)$
$\mathbf{q}_2 = (1,-1)$
$\mathbf{d}_2 = \tilde{\mathbf{d}}_2 + \mathbf{q}_2 = (7,5)$
$\psi_{2,2}$
$\mathbf{d}_2$
$\tilde{\mathbf{d}}_2$
$\mathbf{q}_2$
$\psi_{1,3}$
$\tilde{\mathbf{d}}_3 = 2 \times \mathbf{d}_2 = (14,10)$
$\mathbf{q}_3 = (-1,1)$
$\mathbf{d}_3 = \tilde{\mathbf{d}}_3 + \mathbf{q}_3 = (13,11)$
$\psi_{2,3}$
$\tilde{\mathbf{d}}_3$
$\mathbf{d}_3$
$\mathbf{q}_3$
Anchor frame
Target frame

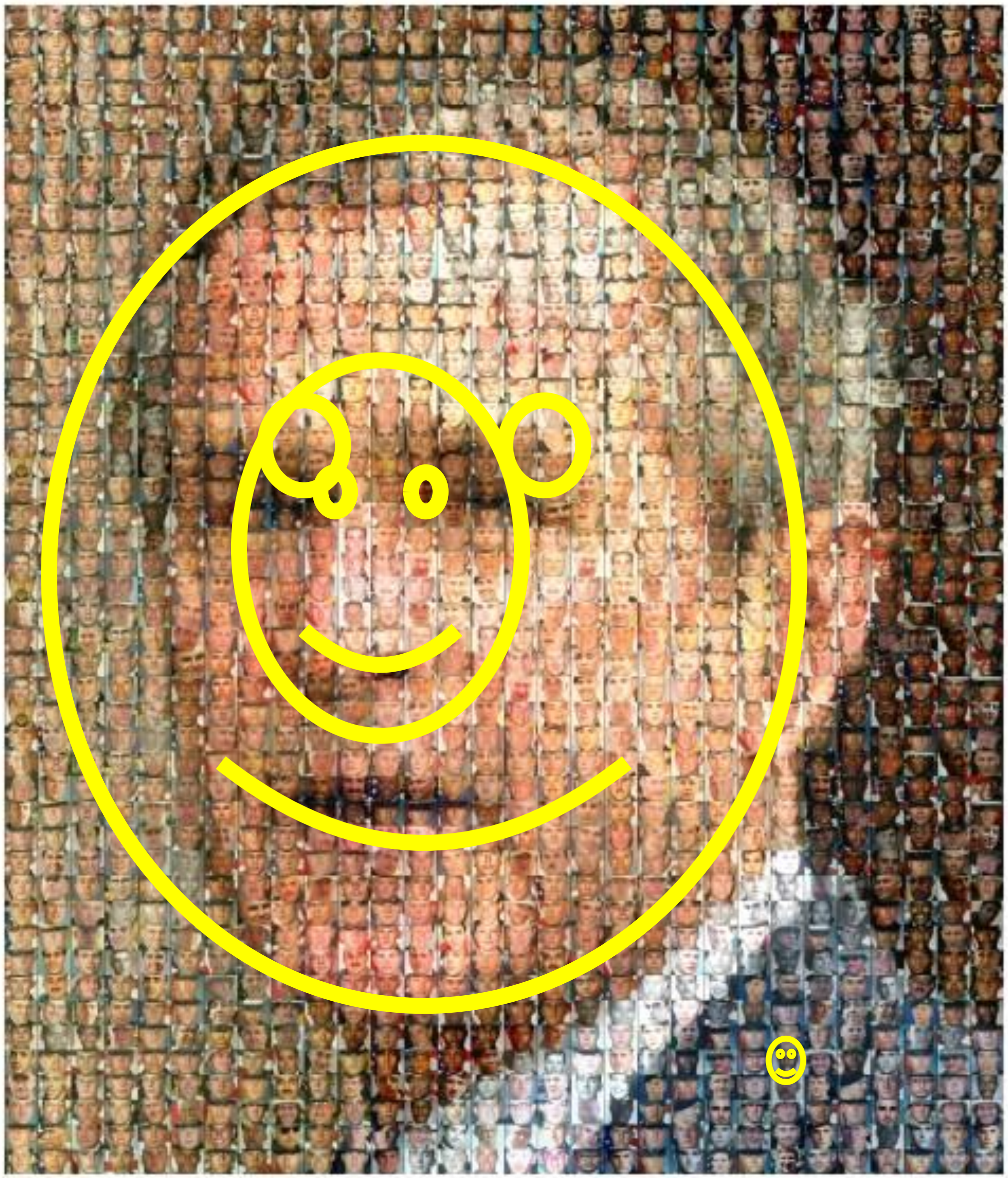
Non-Hierarchical Block Matching Algorithm

Hierarchical Block Matching Algorithm



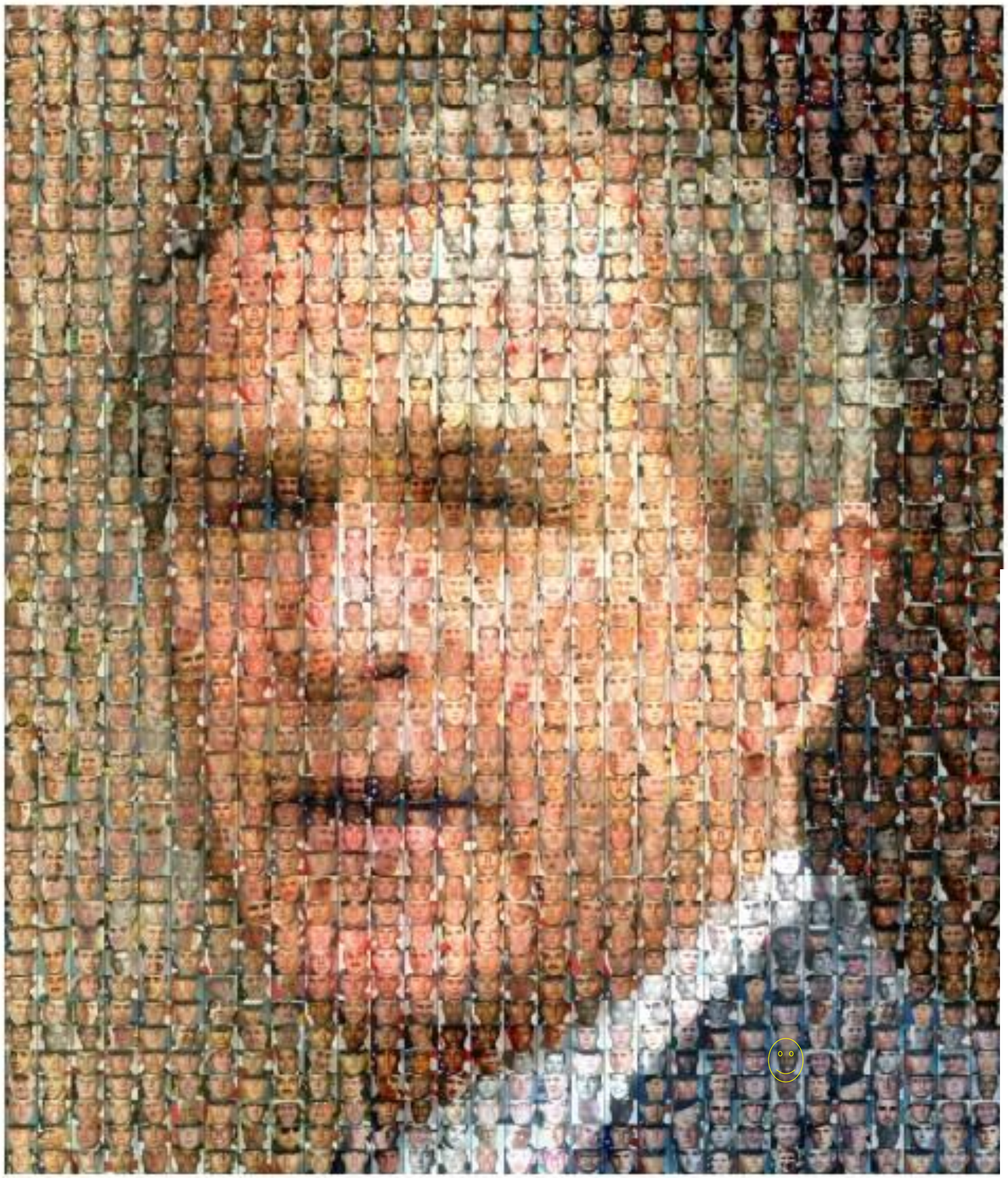
Example: Three-level HBMA

Template Matching

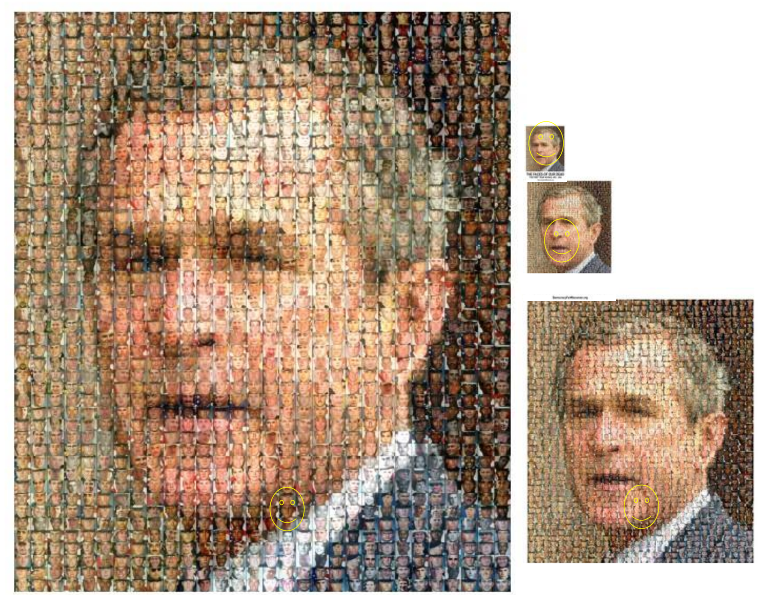


- Strategy 1
  - Use templates of different sizes
  - For large templates, matching is costly

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

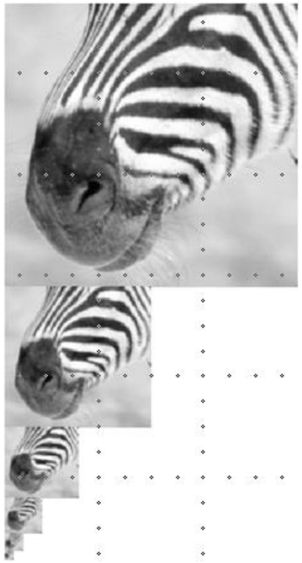
Laplacian Pyramid

- Gaussian Pyramid
  - G_0
  - G_1=D(G_0)
  - G_2=D(G_1)
  - G_3=D(G_2)
- D
  - Gaussian filtering
  - then Downsampling

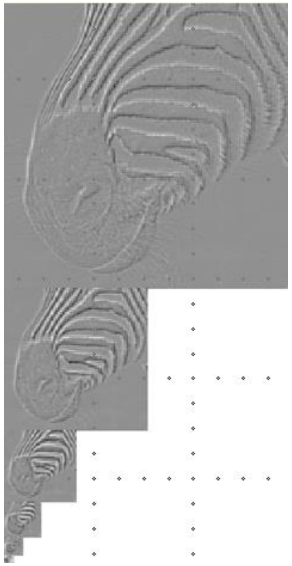
- Laplacian Pyramid
  - L_0=G_0-U(G_1)
  - L_1=G_1-U(G_2)
  - L_2=G_2-U(G_3)
  - L_3=G_3
- U
  - Upsampling

Laplacian Pyramid

- Gaussian Pyramid
  - G_0
  - $G_1=D(G_0)$
  - $G_2=D(G_1)$
  - $G_3=D(G_2)$



- Laplacian Pyramid
  - $L_0=G_0-U(G_1)$
  - $L_1=G_1-U(G_2)$
  - $L_2=G_2-U(G_3)$
  - $L_3=G_3$

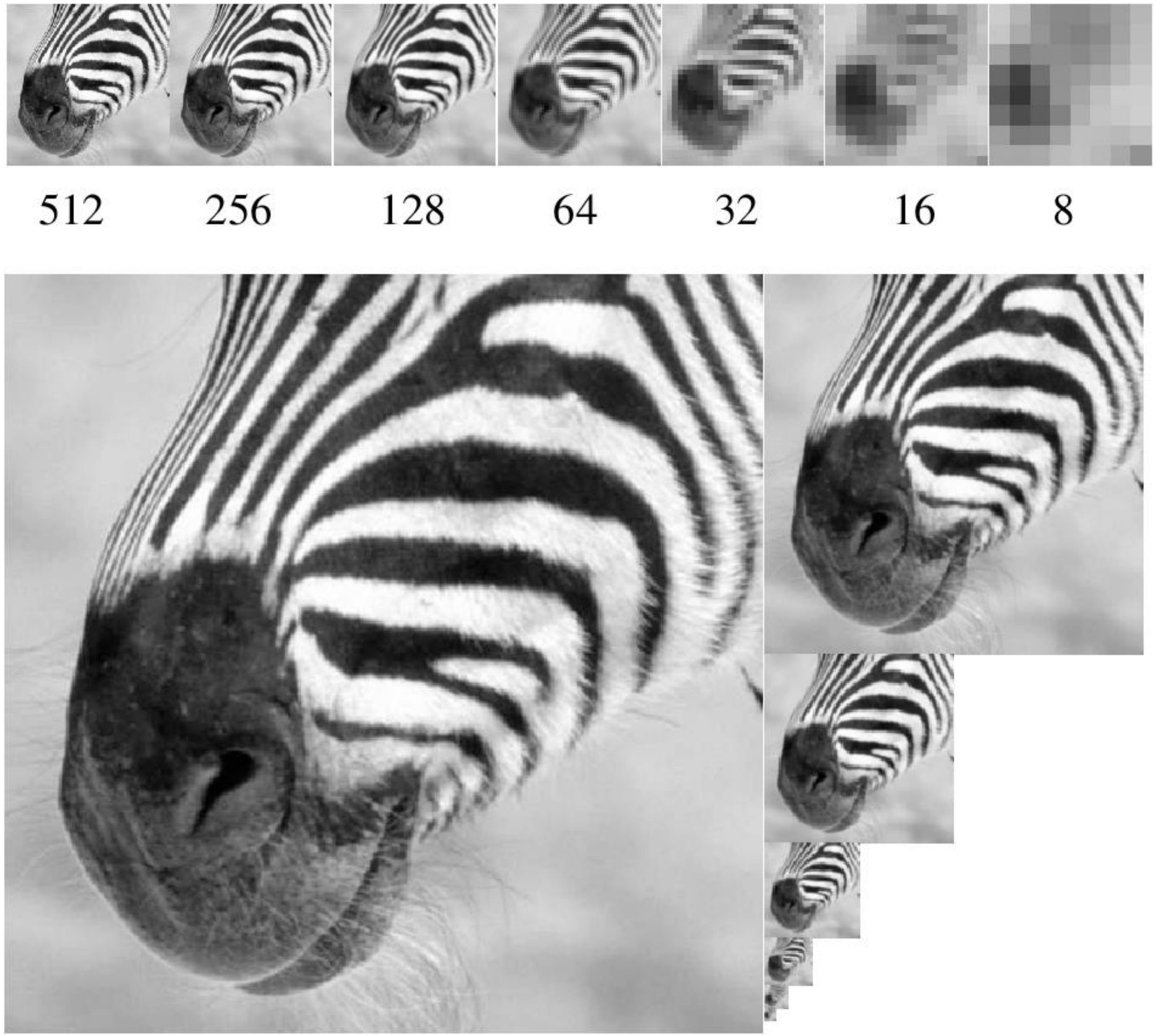


Laplacian Pyramid

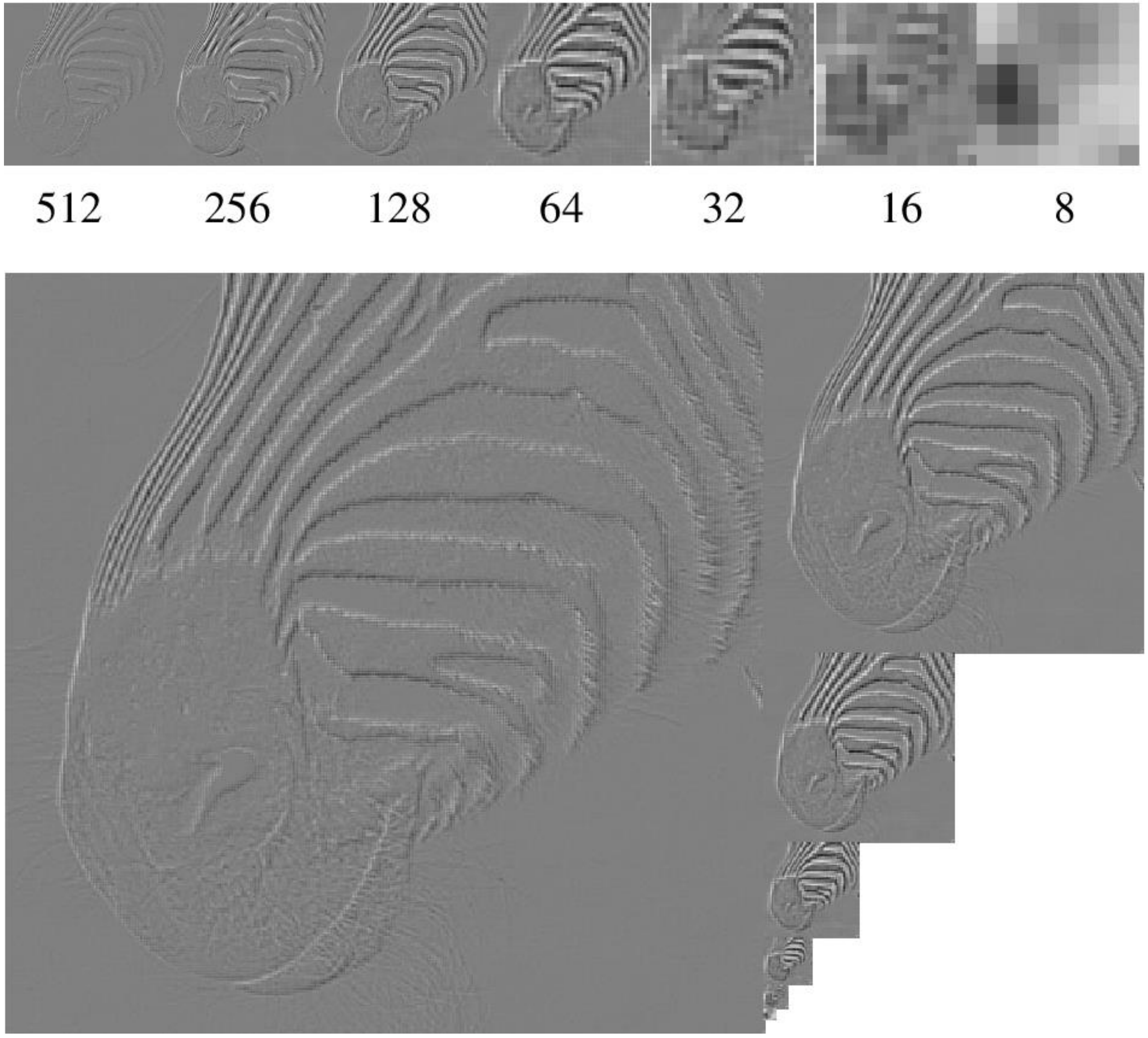
- Analysis
  - $L_0=G_0-U(G_1)$
  - $L_1=G_1-U(G_2)$
  - $L_2=G_2-U(G_3)$
  - $L_3=G_3$

- 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



Laplacian Pyramid



Laplacian Pyramid for Compression