Image Pyramids

Introduction

1. We have already seen translation invariant linear filters
2. We would also like to see scale invariant processing
3. How can we do that?
4. Can we use template matching?

Figure 23.1: Objects in images appear at arbitrary locations and with arbitrary image sizes.

Image Pyramids

1. Image pyramids (i.e., multiresolution representations for images) are a useful data structure for analyzing and manipulating images over a range of spatial scales.

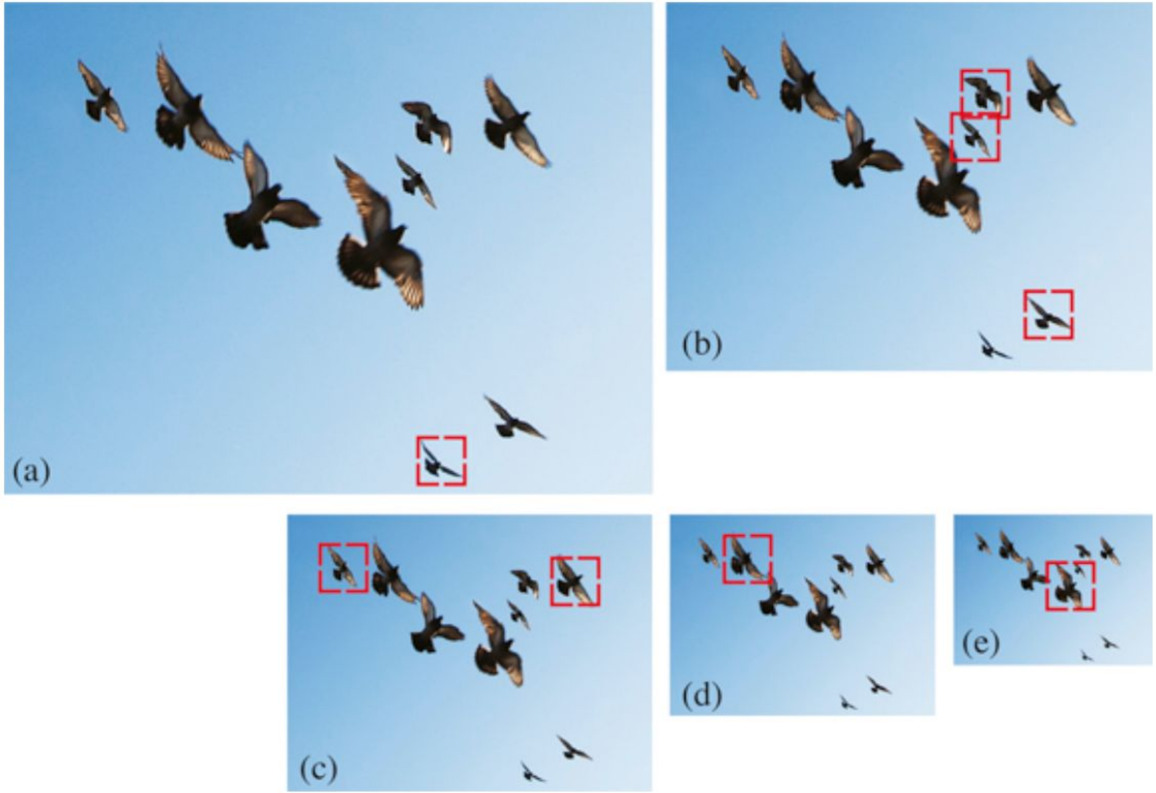

Figure 23.2: Multiscale image pyramid. Each image is 25 percent smaller than the previous one. The red box indicates the size of a template used for detecting birds. As the size of the template is fixed; it will only be able to detect the birds that tightly fit inside the box. By running the same template across many levels in this pyramid, different instances of birds are detected at different scales.

Gaussian Pyramid

1. convolving the image with a low-pass filter
2. subsampling by a factor of 2 the result.
3. Each level is obtained by filtering the previous level with the fourth binomial filter with a stride of 2 (on each dimension)

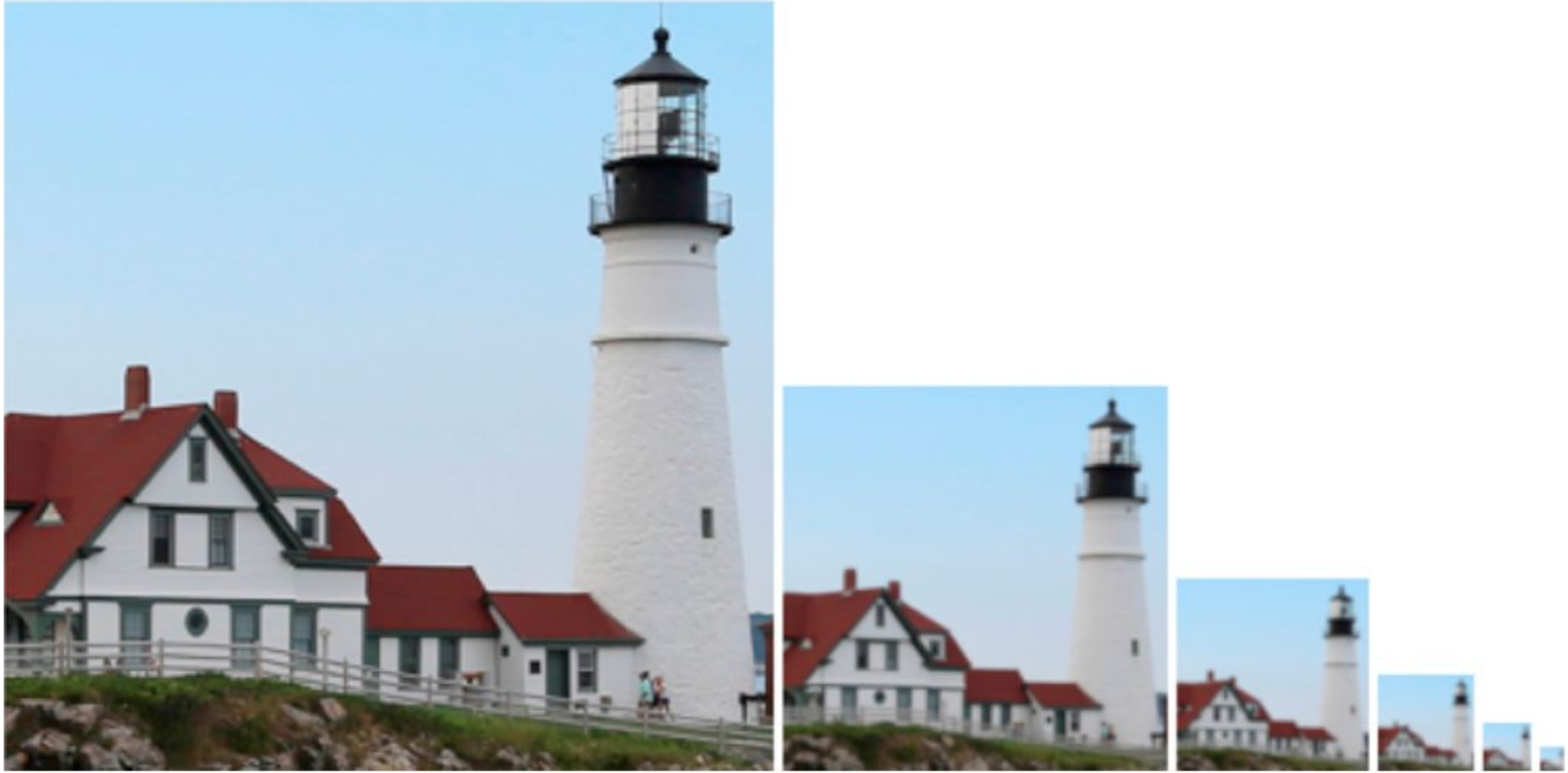

Figure 23.3: Gaussian pyramid with six levels. The first level, $\mathbf{g}_0$, is the input image. The Gaussian pyramid is built for each color channel independently.

Laplacian Pyramid

1. Gives us representations at different frequency bands
2. “The Laplacian pyramid is simple: it represents, at each level, what is present in a Gaussian pyramid image of one level, but not present at the level below it.
3. We calculate that by expanding the lower-resolution Gaussian pyramid image to the same pixel resolution as the neighboring higher-resolution Gaussian pyramid image, **then subtracting the two**.
4. This calculation is made in a recursive, telescoping fashion

Figure 23.4: Laplacian pyramid, including the tiny low-pass residual as the last image. The Laplacian pyramid is built for each color channel independently.

Laplacian Pyramid

1. Laplacian pyramid is an invertible transform, but only if we keep the low-pass residual
2. The Laplacian pyramid can be considered as a special type of deep neural net with it's encoder decoder component.

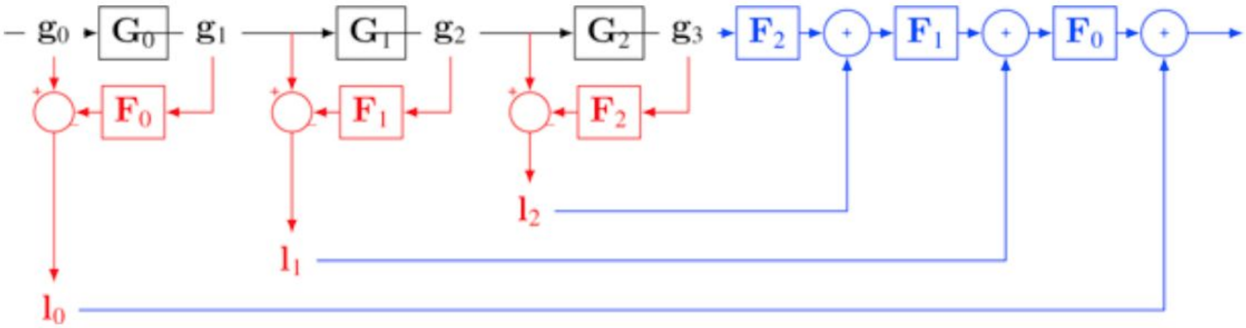

Figure 23.5: The diagram shows the Gaussian pyramid (black), the Laplacian pyramid (red) and the Laplacian inversion for a three-level Laplacian pyramid (blue).

Image Blending

1. We can blend two images using mask

Figure 23.6: Two images to be blended and the blending mask.

Image Blending

1. Using the Laplacian pyramid, we can transition from one image to the next over many different spatial scales to make a gradual transition between the two images.
2. We first build the Laplacian pyramid for the two input images
3. Second step is to build the Gaussian pyramid of the mask
4. In the third step, we combine the three pyramids to compute the Laplacian pyramid of the blended image

$$\ell_{\text{out}} = \ell^A * \mathbf{m} + \ell^B * (1 - \mathbf{m})$$

5. Finally, the fourth step consists in collapsing (i.e., decoding) the resulting pyramid to produce the blended image

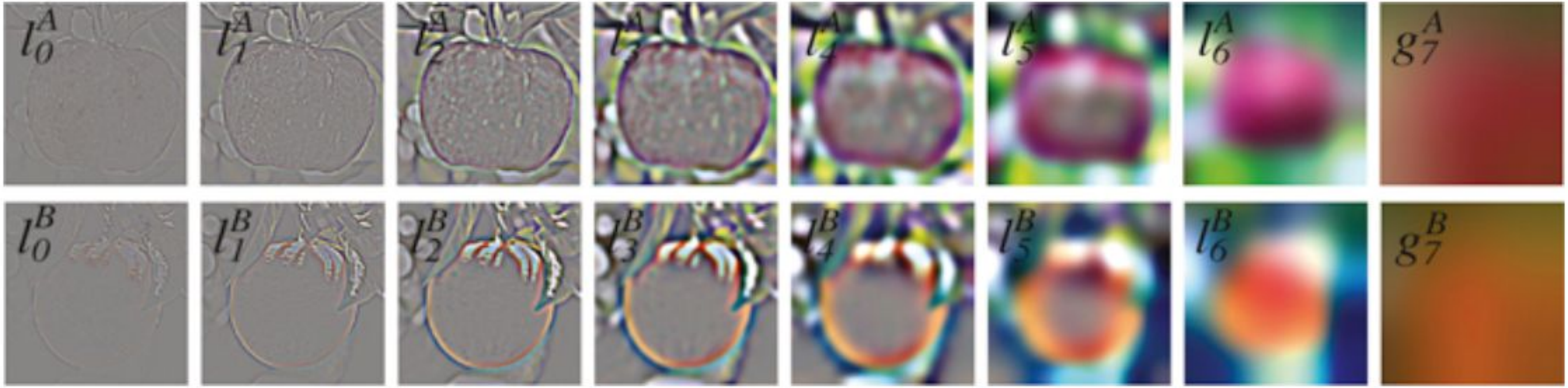

Figure 23.8: Laplacian pyramids (seven levels and Gaussian residual) for both input images.

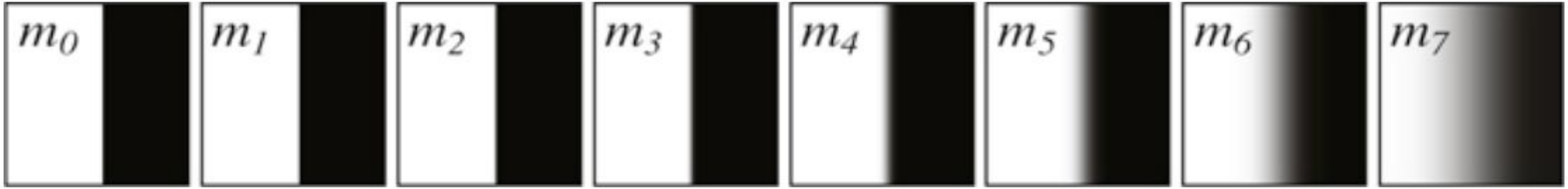

Figure 23.9: Gaussian pyramid of the mask.

References

1. Foundations of Computer Vision - Chapter 23