Lecture 8 Early Vision

ECEN5283 Computer Vision

Dr. Guoliang Fan

Oklahoma State University

Goals

To introduce the basic concept of "early vision" in both biological visual processing and computer vision.

To study the basic concepts about 2D linear filtering with some examples.

Parallel Pathways in Visual Processing

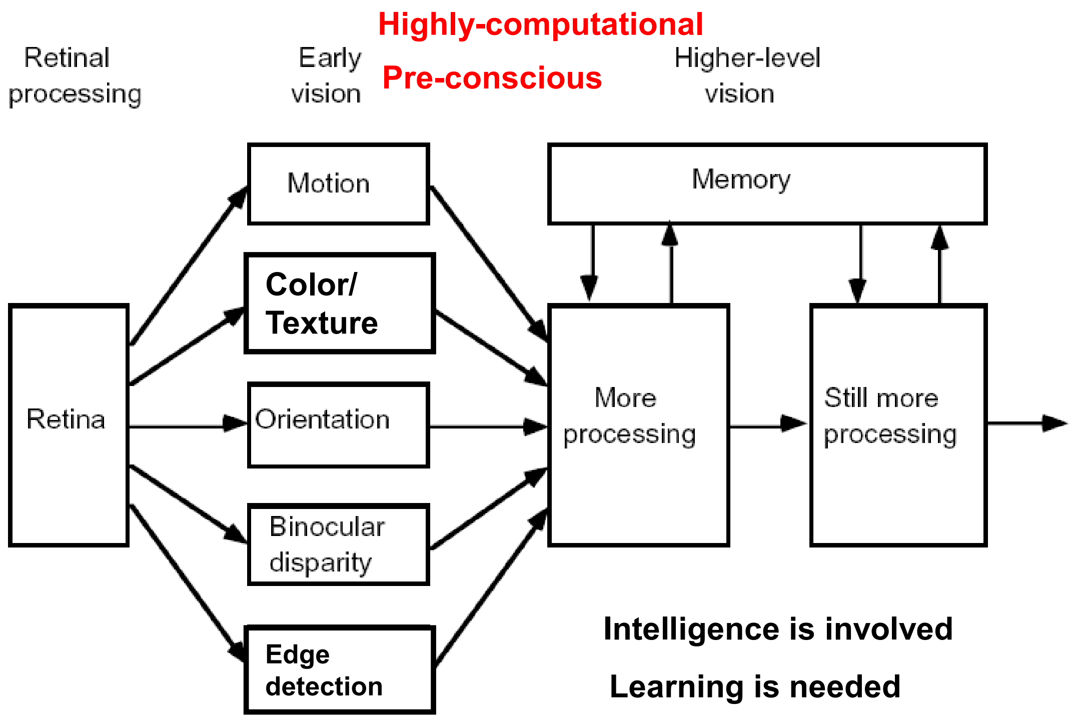


Fig.1.1
A generic diagram for visual processing. In this approach, early vision consists of a set of parallel pathways, each analyzing some particular aspect of the visual stimulus.

The Plenoptic Function and the Elements of Early Vision
web.mit.edu/persci/people/adelson/pub_pdfs/elements91.pdf

Early Vision for Computer Vision

In biological vision systems like our own, this means "pre-conscious" vision before any thinking takes place.

- *Raw, high-information, intrinsic* data are extracted from the light patterns incident on the retina.
- These processes tend to be *highly computational*, involving filtering, coding, feature detection, etc., as opposed to *symbolic* or *cognitive operation*.

For computer vision systems, similar "low-level" tasks have been defined, e.g., as edge detection, texture analysis, image denoising.

In practice, an early vision process may attempt to emulate biological vision, or involve other specific tasks, such as linear filtering.

Linear Filtering

Objective: Transform the image intensities (1) to enhance or extract certain desirable image features or (2) to suppress undesirable image attributes, such as noise and outliers

Implementations

- *2D Linear Filtering*
  - *Shift-Invariant & Linearity*
  - Separable 2D convolution
  - *Low-pass filtering:* 2D average vs Gaussian smoothing
- Frequency-domain image representation
  - Discrete Fourier Transform (DFT)
  - Magnitude and phase spectra
  - Linear phase filtering

Shift-Invariant Linear System

Most imaging or filtering systems have, to a good approximation, three significant properties:

- Superposition (additivity):

$$R(f+g) = R(f) + R(g)$$

- Scaling (homogeneity):

$$R(kf) = kR(f)$$

Linearity

- Shift Invariance (SI) or time-invariance (TI):

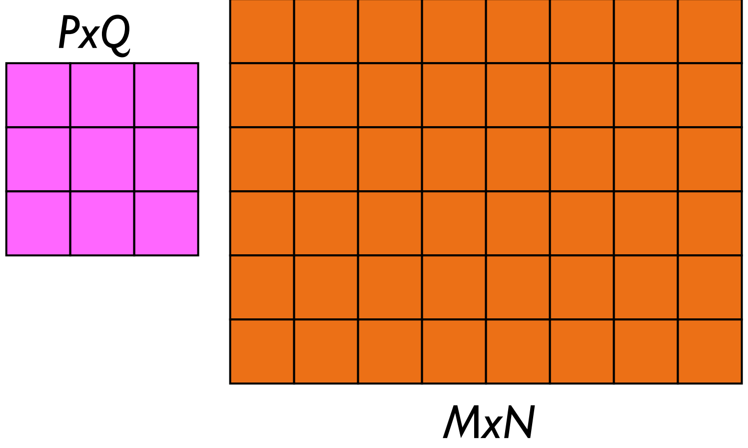
$$x(t) \rightarrow y(t) \Rightarrow x(t-t_0) \rightarrow y(t-t_0)$$

2-D Discrete Convolution

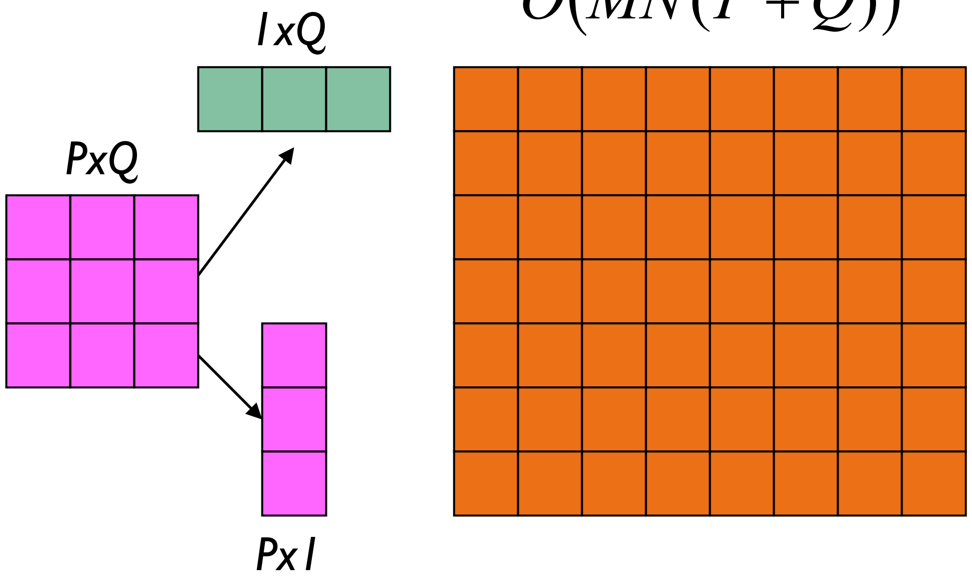
Definition $$r[i,j]=\sum_{u,v} h[i-u,j-v]f[u,v]$$

- Separable 2D convolution $h[i,j]=f[i]h[j]$

$O(MNPQ)$



$O(MN(P+Q))$



Low-pass Filtering

Original Image

Noisy Image

Denoised Image

2-D Discrete Convolution Example (1)

Example 1: Local average smoothing (Separable?)

$$r[i,j] = \frac{1}{(2k+1)^2} \sum_{u=i-k}^{i+k} \sum_{v=j-k}^{j+k} f[u,v]$$

Usually, we construct a (2k+1)x(2k+1) kernel for a smoothing filter.

$$\begin{pmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{pmatrix} \rightarrow \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} * \begin{pmatrix} 1/3 & 1/3 & 1/3 \end{pmatrix}$$

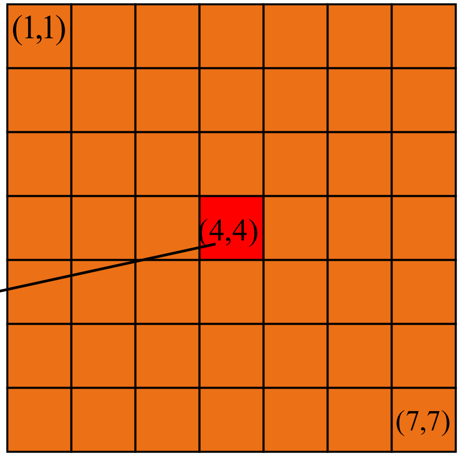
2-D Discrete Convolution Example (2)

Example 2: Gaussian smoothing (Separable?)

$$g(x,y)=\frac{1}{2\pi\delta^2}\exp\left(-\frac{x^2+y^2}{2\delta^2}\right)\to f(x)h(y)=\frac{\exp\left(-\frac{x^2}{2\delta^2}\right)}{\sqrt{2\pi}\delta}\frac{\exp\left(-\frac{y^2}{2\delta^2}\right)}{\sqrt{2\pi}\delta}$$

Usually, we construct a (2k+1)x(2k+1) kernel for a smoothing filter.

$$h[i,j]=\frac{1}{2\pi\delta^2}\exp\left(-\frac{(i-k-1)^2+(j-k-1)^2}{2\delta^2}\right)$$



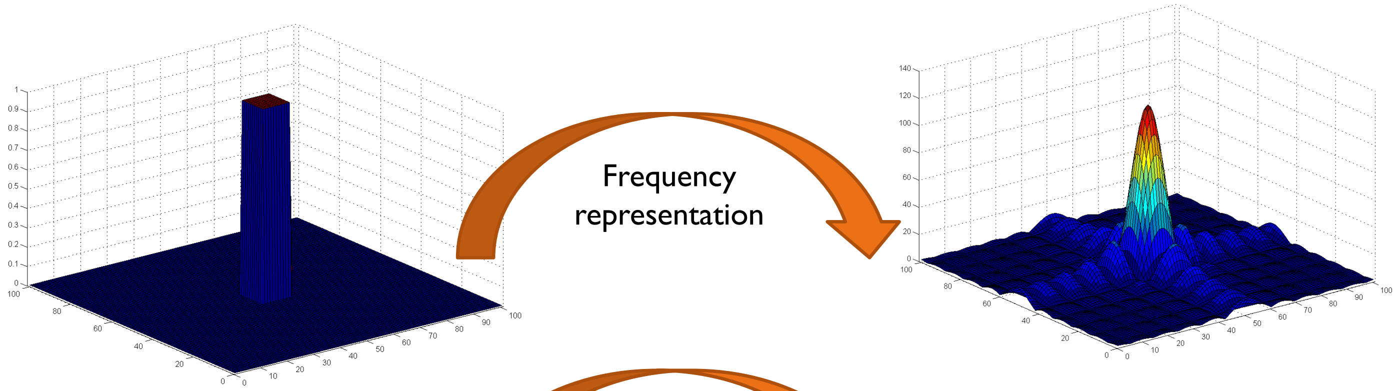
$k=3$

$(k+1,k+1)$

Two Smoothing Operators

Average smoothing:

$$f(i,j)=\frac{1}{(2k+1)^2}(i,j=1,\ldots,(2k+1))$$



Gaussian smoothing:

$$g(x,y)=\frac{1}{2\pi\delta^2}\exp\left(-\frac{x^2+y^2}{2\delta^2}\right)$$

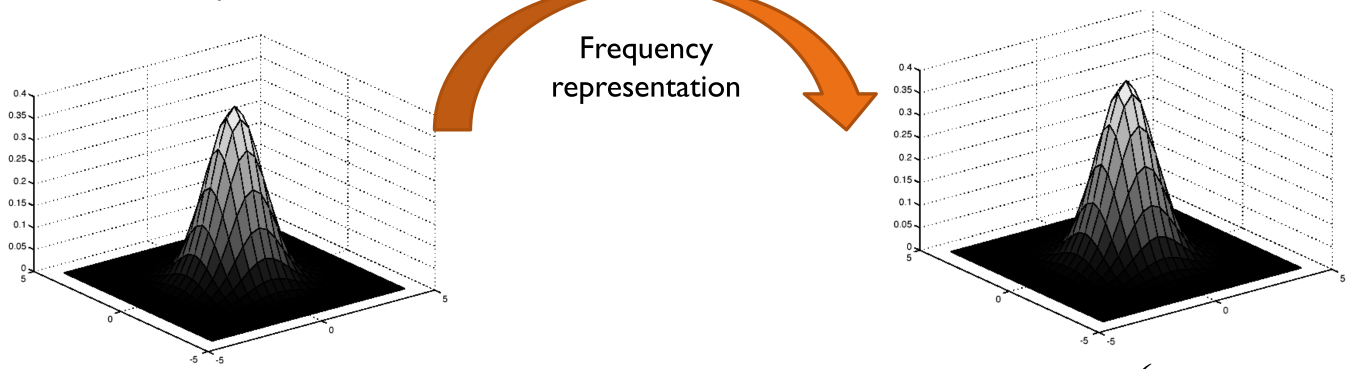


Figure 8.2. The symmetric Gaussian kernel in 2D. This view shows a kernel scaled so that its sum is equal to one; this scaling is quite often omitted. The kernel shown has $\sigma = 1$. Convolution with this kernel forms a weighted average which stresses the point at the center of the convolution window, and incorporates little contribution from those at the boundary. Notice how the Gaussian is qualitatively similar to our description of the point spread function of image blur; it is circularly symmetric, has strongest response in the center, and dies away near the boundaries.

$$G(u,v)=\frac{1}{2\pi}\exp\left(-\frac{(u^2+v^2)}{2(1/\delta)^2}\right)$$

Average and Gaussian Smoothing

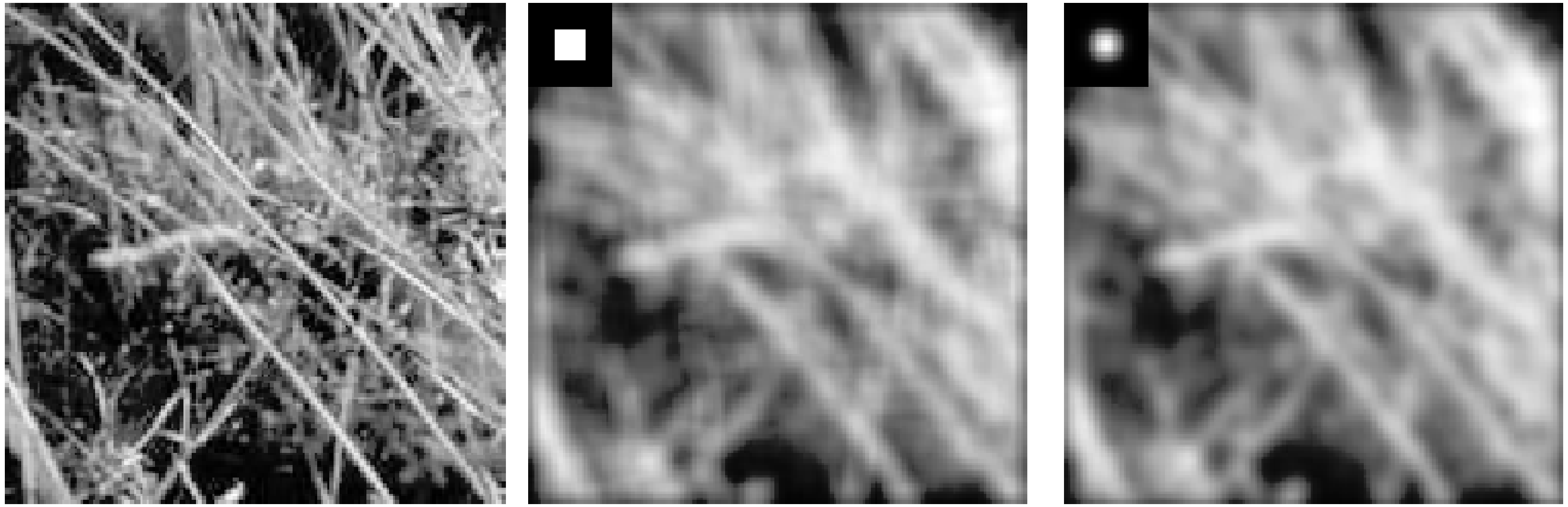


Figure 8.1. Although a uniform local average may seem to give a good blurring model, it generates effects that are not usually seen in defocussing a lens. The images above compare the effects of a uniform local average with weighted average. The image at the top shows a view of grass. On the left in the second row, the result of blurring this image using a uniform local model and on the right, the result of blurring this image using a set of Gaussian weights. The degree of blurring in each case is about the same, but the uniform average produces a set of narrow vertical and horizontal bars — an effect often known as **ringing**.