Digital Convolution — E186 Handout

The convolution of two continuous signals is defined as

$$y(t) = h(t) * x(t) \triangleq \int_{-\infty}^{\infty} x(\tau)h(t-\tau)d\tau = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau$$

i.e., convolution operation is commutative. Also it is associative:

$$h * (g * x) = (h * g) * x$$

As a typical example, $y(t)$ is the output of a system characterized by its impulse response function $h(t)$ with input $x(t)$.

Convolution in discrete form is

$$y(n) = \sum_{m=-\infty}^{\infty} x(n-m)\ h(m) = \sum_{m=-\infty}^{\infty} h(n-m)\ x(m)$$

If $h(m)$ is finite, i.e.,

$$h(m) = \begin{cases} h(m) & |m| \leq k \\ 0 & |m| > k \end{cases}$$

the convolution becomes

$$y(n) = \sum_{m=-k}^{k} x(n-m)\ h(m)$$

In time domain, all realistic systems are causal

$$y(n) = 0 \qquad \text{if } n < 0$$

However, in image processing, we often consider convolution in spatial domain where causality does not apply.

If $h(m)$ is symmetric (almost always true in image processing), i.e.,

$$h(-m) = h(m)$$

the convolution becomes

$$y(n) = \sum_{m=-k}^{k} x(n+m)\, h(m)$$

If $x(m)$ is also finite (always true in reality), i.e.,

$$x(m) = \begin{cases} x(m) & 0 \leq m < N \\ 0 & otherwise \end{cases}$$

for $x(n+m)$ to be in the valid non-zero range, its index $(n+m)$ has to satisfy:

$$0 \leq (n+m) \leq N-1$$

correspondingly for $y(n)$ to be non-zero, its index (n) has to satisfy:

$$-m \leq n \leq N-m-1$$

When $m = k$, the lower bound $n = -k$ is reached, and when $m = -k$, the upper bound $n = N+k-1$ is reached. In other words, there are $N+2k$ valid elements in the output:

$$y(n), \qquad (-k \leq n \leq N+k-1)$$

This convolution can be best understood graphically (where the index of $y(n)$ is rearranged).

In image processing, all the discussions above for one-dimensional convolution are generalized into two dimensions, and h is called a convolution kernel.