Spatial-Domain Convolution Filters

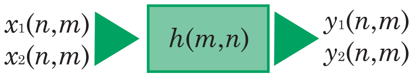
Consider a **linear space-invariant (LSI)** system as shown:

The two separate inputs to the LSI system, $x_1(m)$ and $x_2(m)$, and their corresponding outputs are given as

$$x_1(m) \to y_1(m) \quad \text{and} \quad x_2(m) \to y_2(m)$$

Thus, an LSI system has the following properties:

- **Superposition property**
  A **linear system** follows **linear superposition**. Consider the following LSI system with two inputs $x_1(m,n)$ and $x_2(m,n)$, and their corresponding outputs $y_1(m,n)$ and $y_2(m,n)$.



  The linear superposition is given as

$$x_1(m,n) + x_2(m,n) \to y_1(m,n) + y_2(m,n)$$

- **Space invariance property**
  Consider that if the input $x(m)$ to a linear system is shifted by M, then the corresponding output is also shifted by the same amount of space, as follows:

$$x(m-M) \to y(m-M).$$

  Futhermore, $h(m,n;m',n') \cong Tr[\delta(m-m',n-n')] = h(m-m',n-n';0,0)$, where Tr is the transform due to the linear system, as shown in the above figure. Hence, $h(m,n;m',n') = h(m-m',n-n')$; the system is defined as LSI or as **linear time invariant (LTI)**.

- **Impulse response property**
  A linear space-invariant system is completely specified by its **impulse response**. Since any input function can be decomposed into a sum of time-delayed weighted impulses, the output of a linear system can be calculated by superposing the sum of the impulse responses. For impulse at the origin, the output is $h(m,n;0,0) \cong Tr[\delta(m-0,n-0)]$.

Convolution

As a consequence of **LSI** properties, the output of a linear shift-invariant system can be calculated by a **convolution** integral since the superposition sum simplifies to a convolution sum due to the shift-invariant property. The output of an LSI system is given as

$$y(m) = \int_{-\infty}^{\infty} f(m,z)x(z)dz$$

Following the **shift invariance** of LSI, convolution is obtained, given as

$$y(m) = \int_{-\infty}^{\infty} f(m-z)x(z)dz$$

Convolution describes the processing of an image within a **moving window**. Processing within the window always happens on the original pixels, *not* on the previously calculated values. The result of the calculation is the output value at the center pixel of the moving window. The following steps are taken to obtain the convolution:

1. Flip the window in x and y.
2. Shift the window.
3. Multiply the window weights by the corresponding image pixels.
4. Add the weighted pixels and write to the output pixel.
5. Repeat steps 2–4 until finished.

A problem with the moving window occurs when it runs out of pixels near the image border. Several 'trick' solutions for the border region exist:

- Repeat the nearest valid output pixel.
- Reflect the input pixels outside the border and calculate the convolution.
- Reduce the window size.
- Set the border pixels to zero or mean image.
- Wrap the window around to the opposite side of the image (the same effect produced by filters implemented inthe Fourier domain), i.e., the circular boundary condition.

Convolution and Correlation in the Fourier Domain

Correlation is defined in the **spatial domain**, and the inverse Fourier transform must be performed to return to the spatial domain. In the case of **autocorrelation**,

$$F\{f(x) \otimes g(x)\} = F(u)G^*(u)$$

reduces to

$$F\{f(x) \otimes f(x)\} = |F(u)|^2$$

The inverse of this **intensity** spectrum generates the autocorrelation output:

$$f(x) \otimes f(x) = F^{-1}|F(u)|^2$$

The process of multiplying a **Fourier transform** by F^* and then taking the inverse transform is called **matched filtering**. This Fourier transform property of correlation forms the basis of performing matched filtering in a computer using FFTs.

Transform of a transform: Taking the transform of a transform reproduces the original function with its axes reversed:

$$F\{F(u,v)\} = f(-x,-y)$$

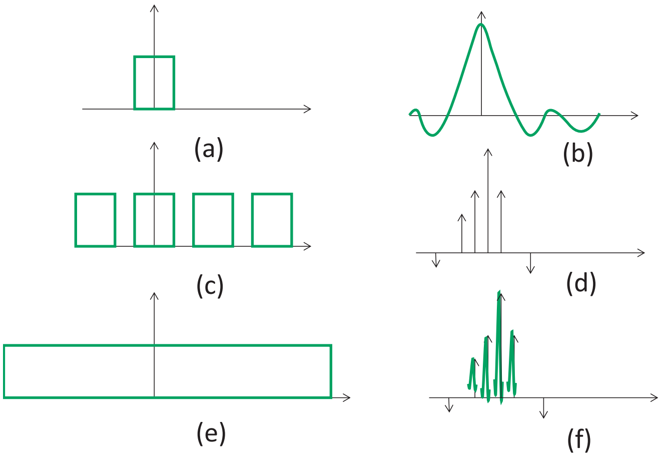
Optically, a lens performs a Fourier transform on its input at the **Fourier plane**. Interestingly, it actually performs a forward transform. This property explains why an image is flipped at the output when two lenses are used to image an input. It is because two lenses perform two forward transforms.

Convolution: The most famous property of the Fourier transform is the transform of convolution property, which states that the Fourier transform of a convolution of two functions is obtained by simply multiplying the individual Fourier transforms:

$$F\{f(x) * g(x)\} = F(u)G(u)$$

Spectrum of a Finite Periodic Signal

Single pulse: Assume a rectangular pulse as shown in (a). Its spectrum is a **sinc function** [sinc(x)/x], as shown in (b).



Infinite periodic series: If a **periodic signal**, as shown in (c), is created by repeating the pulse, this case is equivalent to convolving with a series of impulse functions (remember that in sampling, it was a multiplication). The frequency domain will be modified by a multiplication with the spectrum of this impulse series.

> **Spectrum**: The spectrum of the periodic function will be a series of impulses that have an envelope like the sinc function, as shown in (d). As a result, the spectrum of this periodic function is discrete.

Truncated periodic series: Assume that instead of an infinite series, the periodic function is truncated at a width A, as shown in (e). This is equivalent to multiplying by a huge rectangular function.

> **Spectrum**: The spectrum will be modified by convolution with the spectrum of this huge rectangular function, which is a narrow sinc function of width $1/A$. Imagine erecting a narrow sinc function at the location of the impulse function in the Fourier domain, as shown in (f).

Spectrum of a Finite Periodic Signal (cont.)

Mathematically, the single rectangular pulse is given as

$$\mathrm{rect}(x/a) \leftrightarrow \mathrm{sinc}(au)$$

The periodic **rect function** is expressed as a **convolution** with a series of impulse functions:

$$\mathrm{rect}(x/a) \otimes \sum \delta(x - nb) \leftrightarrow \mathrm{sinc}(au) \cdot \sum \delta(u - n/b)$$

The Fourier transform is equivalent to a series of impulses the magnitude of which is enveloped by a sinc function, sinc(au):

$$\left[\mathrm{rect}(x/a) \otimes \sum \delta(x - nb)\right] \mathrm{rect}(x/A) \leftrightarrow$$
$$\left[\mathrm{sinc}(au) \cdot \sum \delta(u - n/b)\right] \otimes \mathrm{sinc}(Au)$$

Limiting the number of the rect series by a huge rect with width A produces a very narrow sinc function sinc(Au) convolving the series of delta functions that are under another sinc envelope, sinc(au). Thus, the delta functions acquire a certain width due to this narrow sinc convolving with it.

Assume that the rectangular series is replaced by a **triangular function**:

$$\mathrm{tri}(x/a) \leftrightarrow \mathrm{sinc}^2(au)$$

This will thus change the envelope of the **spectrum** from sinc to sinc^2. The spectrum shown in (d) on the previous page will be modified by the envelope. However, a finite width on the series will have the same effect of erecting a sinc at the bottom of the impulse functions. Mathematically, the change of the function sinc(Au) to sinc^2(Au) in the above equation will be the spectrum for the finite triangular series.

When such a pulse train is passed through a communication channel with finite bandwidth (or a lens with finite aperture), the spectrum will be truncated. The resulting rectangular pulse train will appear smoothed at the edge as the high-frequency content is minimized by the bandwidth. This effect explains why imaging with a finite-aperture lens causes degradation in the fine details of an image because the aperture acts as a low-pass filter.