Image Processing: Transforms (part 1)

4C8: Digital Media Processing

Ussher Assistant Professor François Pitié

2021/2022

Department of Electronic & Electrical Engineering , Trinity College Dublin

adapted from original material written by Prof. Anil Kokaram.

Transforms Overview

- 2D Z-transform
- 2D Fourier Transform - Continuous and Discrete (2D DFT)
- Sampling Theorem
- 2D Filters and stability
- Some applications and the need for non-linear filters

Additional reference for this handout can be found in the book *Two-Dimensional Signal and Image Processing by Jae. S. Lim*

The 2D Z-Transform

The 2D Z-transform

Recall 1D Z Transform of signal x_n

$$X(z) = \sum_{n=-\infty}^{\infty} x_n z^{-n} \tag{1}$$

This is just a polynomial in z (a complex number). It is used to solve difference equations and also helps with stability of IIR filters.

The Z-transform of a sequence $g[h,k]$ is denoted $G(z_1, z_2)$:

$$G(z_1, z_2) = \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} g[h,k] z_1^{-h} z_2^{-k} \tag{2}$$

The z_1 and z_2 components act on the sequence $g[h,k]$ along vertical and horizontal directions.

The Z transform in 2D is a function of 2 complex numbers z_1, z_2. It exists in a 4D space, so it is very difficult to visualise.

Given a finite length sequence $g[h,k] =$ `[1 2;3 4]`. The Z-transform can be written down directly (Top left is [0,0]):

$$G(z_1, z_2) = 1 + 3z_1^{-1} + 2z_2^{-1} + 4z_1^{-1}z_2^{-1} \tag{3}$$

A non-causal signal $g[h,k] =$ `[1 2 3;4 5 6;7 8 9]`. The Z-transform can be written down as (Middle is [0,0]):

$$\begin{aligned} G(z_1, z_2) = & z_1 z_2 + 2z_1 + 3z_1 z_2^{-1} \\ & + 4z_2 + 5 + 6z_2^{-1} \\ & + 7z_1^{-1}z_2 + 8z_1^{-1} + 9z_1^{-1}z_2^{-1} \end{aligned} \tag{4}$$

Given signal $X(z_1, z_2)$ input into a system with transfer function $H(z_1, z_2)$ output $Y(z_1, z_2)$ is $X(z_1, z_2)H(z_1, z_2)$. [Just like 1D]

$H(z_1, z_2)$ is the Z-transform of the impulse response. [Just like 1D]

FIR filters

FIR = Finite Impulse Response

$H(z_1, z_2) = 4 - z_1^{-1} - z_1 - z_2^{-1} - z_2$ is an FIR filter.

We can verify this by deriving the difference equation.

$$Y(z_1, z_2) = X(z_1, z_2)H(z_1, z_2)$$

$$\begin{aligned} Y(z_1, z_2) &= X(z_1, z_2)(4 - z_1^{-1} - z_1 - z_2^{-1} - z_2) \\ &= 4X(z_1, z_2) - z_1^{-1}X(z_1, z_2) - z_1X(z_1, z_2) \\ &\qquad - z_2^{-1}X(z_1, z_2) - z_2X(z_1, z_2) \end{aligned}$$

$$y[h, k] = 4x[h, k] - x[h-1, k] - x[h+1, k] - x[h, k-1] - x[h, k+1]$$

Here $y[h, k]$ only depends on input $x[h, k]$, so when input stops, output stops in finite number of samples; hence FIR filter.

OR we could work out the impulse response $H(z_1, z_2) \rightleftarrows p[h, k]$.

Let's say example input is $x[h,k] = [1\ 2; 3\ 4]$.
So $X(z_1, z_2) = 1 + 2z_2^{-1} + 3z_1^{-1} + 4z_1^{-1}z_2^{-1}$ (eq. (3)) and then

$$\begin{aligned}
Y(z_1, z_2) =& X(z_1, z_2)H(z_1, z_2) \\
=& \left(4 - z_1^{-1} - z_1 - z_2^{-1} - z_2\right) \\
& \times \left(1 + 2z_2^{-1} + 3z_1^{-1} + 4z_1^{-1}z_2^{-1}\right) \\
=& - z_1 - 3z_1 z_2^{-1} - z_2 - 1 + 3z_2^{-1} - 2z_2^{-2} \\
& - 3z_1^{-1}z_2 + 7z_1^{-1} + 11z_1^{-1}z_2^{-1} - 4z_1^{-1}z_2^{-2} - 3z_1^{-2} - 4z_1^{-2}z_2^{-1}
\end{aligned}$$

$$y[h,k] = Z^{-1}\Big(Y(z_1, z_2)\Big) \equiv \begin{bmatrix} 0 & -1 & -3 & 0 \\ -1 & \color{red}{-1} & 3 & -2 \\ -3 & 7 & 11 & -4 \\ 0 & -3 & -4 & 0 \end{bmatrix}$$

The red value represents the value of the signal at $[h,k] = [0,0]$.

IIR filters

IIR = Infinite Impulse Response

$$H(z_1, z_2) = \frac{H_1(z_1, z_2)}{H_2(z_1, z_2)} \tag{5}$$

For example: $H(z_1, z_2) = \dfrac{1}{4 - z_1^{-1} - z_1 - z_2^{-1} - z_2}$

Denoting the input and output as $x[h, k]$ and $y[h, k]$:

$$Y(z_1, z_2) = X(\cdot)H(\cdot) = \frac{X(z_1, z_2)}{(4 - z_1^{-1} - z_1 - z_2^{-1} - z_2)}$$

$$Y(z_1, z_2)(4 - z_1^{-1} - z_1 - z_2^{-1} - z_2) = X(z_1, z_2)$$

$$y[h, k] = \frac{1}{4}\left(x[h, k] + y[h+1, k] + y[h-1, k] + y[h, k+1] + y[h, k-1]\right)$$

The output depends on previous outputs *and* future outputs. Hence impulse response does not stop in finite number of samples, hence IIR filter.

Stability of filters

LIM pages 102–123

Stability of LSI systems requires that the impulse response $p[h,k]$ of the system is finitely summable i.e.

$$\sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |p[h,k]| < \infty \tag{6}$$

In 1D this criterion is exposed by examining the poles of the system and their position relative to the unit circle.

Stability of non-separable 2D IIR filters is difficult to analyse. There is no polynomial factorisation theorem for n-D, and we have a 2-D Z-Xform! No notion of poles because $H_2(z_1, z_2) = 0$ maps out a SURFACE in the complex 4D z-space.

We shall not bother with non-separable 2D IIR filter stability for this reason.

Separable filters and stability

We already dealt with filters that had a separable impulse response. Their system transfer functions are then

$$H(z_1, z_2) = H_1(z_1)H_2(z_2) \tag{7}$$

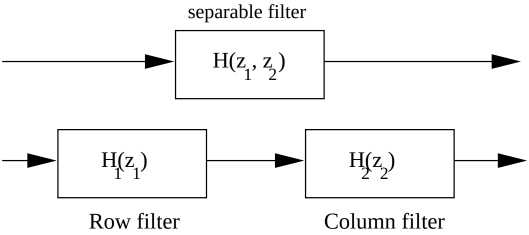
These types of systems are implemented by processing along rows (or columns) and then processing along columns (or rows). Revisit that low pass filter again.

Given $H_1(z_1) = z_1^{-1} + 2 + z_1$ and $H_2(z_2) = z_2^{-1} + 2 + z_2$. What is the impulse response of the system?

$$\begin{aligned} H(z_1, z_2) &= H_1(z_1)H_2(z_2) \\ &= (z_1^{-1} + 2 + z_1)(z_2^{-1} + 2 + z_2) \\ &= z_1^{-1}z_2^{-1} + 2z_1^{-1} + z_1^{-1}z_2 + 2z_2^{-1} + 4 + 2z_2 + z_1z_2^{-1} + 2z_1 + z_1z_2 \end{aligned}$$

$$\Rightarrow \quad Z^{-1}\left(H(z_1, z_2)\right) \equiv \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Separable filters and stability



The stability analysis of separable filters is therefore conducted in the same manner as 1-D filters.

We just have to make sure that the row and column filters are separately STABLE, and then the whole system is stable. Can use usual pole/zero analysis in these cases.

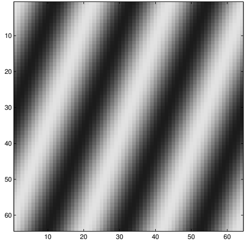
Fourier Transform

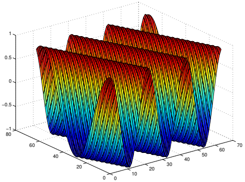
2D Fourier Analysis (Continuous 2-D signals)

The idea is to represent a signal as a sum of pure sinusoids of different amplitudes and frequencies.

In 1D the sinusoids are defined by frequency and amplitude.

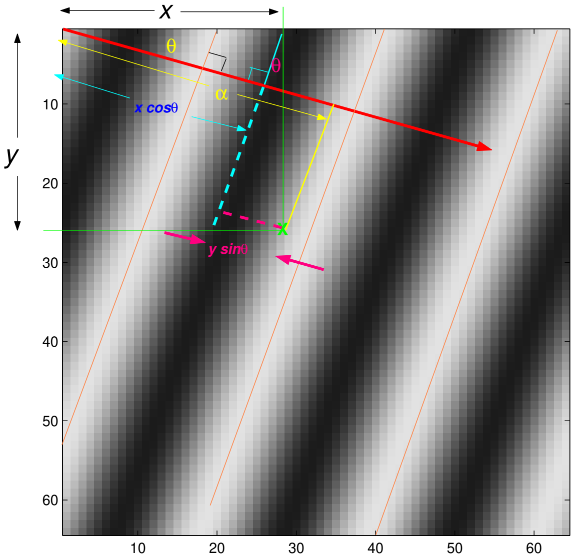
In 2D these sinusoids have a *direction* as well
e.g. $f(x, y) = a\cos(\omega_1 x + \omega_2 y + \phi)$





$a = 1.0$ $\omega_1 = 0.29$ rad/pel $\omega_2 = 0.11$ rad/pel Wave is directed at 20 degrees off horizontal, frequency is 0.05 cycles per pel (0.31 rad/pel) *in that direction* and phase lag $\phi = 0$.

How do ω_1, ω_2 relate to direction?



How do ω_1, ω_2 relate to direction?

In the direction θ: $f(\alpha) = a\cos(\omega_0\alpha + \phi)$, where ϕ is just some phase lag. Given any point (x, y), $\alpha = x\cos\theta + y\sin\theta$, thus:

$$f(x, y) = a\cos(\omega_0[x\cos\theta + y\sin\theta] + \phi)$$

Compare this with

$$f(x, y) = a\cos(\omega_1 x + \omega_2 y + \phi)$$

Therefore $\omega_1 = \omega_0\cos\theta$ and $\omega_2 = \omega_0\sin\theta$ Here $a = 1.0$; $\theta = -20°$; $\phi = 0°$; $\omega_0 = 0.05$ radians per pel.

The 2D Fourier Transform

Recall 1D Fourier Transform of signal $x(t)$

$$X(\omega) = \int x(t)e^{-j\omega t}dt \tag{8}$$

The 2D Fourier Transform is similar. except with TWO frequency axes for horizontal and vertical frequency.

$$\boxed{\begin{aligned} F(\omega_1, \omega_2) &= \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} f(x,y)e^{-j(\omega_1 x + \omega_2 y)}dxdy \\ f(x,y) &= \frac{1}{4\pi^2}\int_{\omega_2=-\infty}^{\infty}\int_{\omega_1=-\infty}^{\infty} F(\omega_1,\omega_2)e^{j(\omega_1 x+\omega_2 y)}d\omega_1 d\omega_2 \end{aligned}}$$

Convolution: $f_1(x,y) \circledast f_2(x,y) \rightleftarrows F_1(\omega_1,\omega_2)F_2(\omega_1,\omega_2)$

Parseval: $\int_y\int_x |f(x,y)|^2 dxdy = \frac{1}{4\pi^2}\int_{\omega_2}\int_{\omega_1}|F(\omega_1,\omega_2)|^2 d\omega_1 d\omega_2$

Fourier Transform Identities

1) The 2D Fourier Transform is separable:

$$F(\omega_1, \omega_2) = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{\infty} f(x,y) e^{-j(\omega_1 x + \omega_2 y)} dx dy$$

$$= \int_{y=-\infty}^{\infty} \left[\int_{x=-\infty}^{\infty} f(x,y) e^{-j\omega_1 x} dx \right] e^{-j\omega_2 y} dy$$

Thus you can do a 1-D transform of rows first then do a 1-D transform of the result along columns, or vice-versa.

2) Fourier Transform of a convolution of 2 signals is the product of the individual Fourier Transforms:

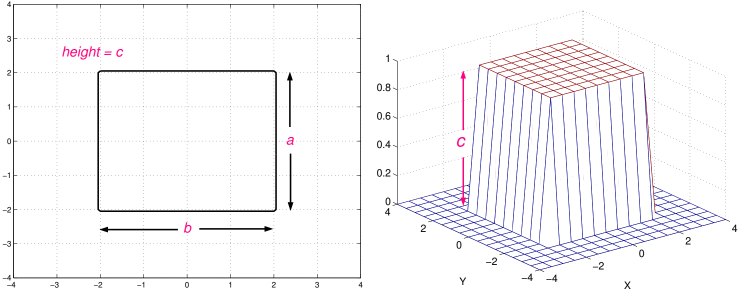
if $f(x,y) = f_x(x) \circledast f_y(y)$ then $F(\omega_1, \omega_2) = F_x(\omega_1) F_y(\omega_2)$

3) For REAL signals:

$$F(\omega_1, \omega_2) = F^*(-\omega_1, -\omega_2)$$

This means that the Fourier Transform (DFT) of an image will be centrally symmetric.

2D Fourier Xform example: $f(x, y) = \mathrm{rect}(ay, bx)$



Let's look at the Fourier Transform of a 2D rectangle pulse of height a, width b and value $c = 1$.

Note that this 2D rectangle pulse is actually the convolution of two 1D rectangle pulses:

$$f(x,y) = \mathrm{rect}(ay, bx) = \mathrm{rect}(ay) \circledast \mathrm{rect}(bx)$$

We know the 1D xform of a rectangular pulse. Say for $\mathrm{rect}(bx)$:

$$\int_{x=-b/2}^{b/2} e^{-j\omega_1 x} dx = \frac{1}{-j\omega_1}\left[e^{-j\omega_1 b/2} - e^{j\omega_1 b/2}\right] \tag{9}$$

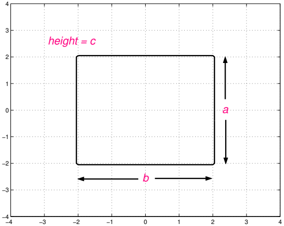
$$= \frac{\sin(b\omega_1/2)}{\omega_1/2} = b\mathrm{sinc}(\omega_1 b/2) \tag{10}$$

Thus

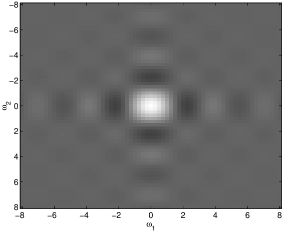
$$F(\omega_1, \omega_2) = a \times b \times \mathrm{sinc}(\omega_1 b/2) \times \mathrm{sinc}(\omega_2 a/2)$$

2D Fourier Xform example: $f(x,y) = \mathrm{rect}(a,b,c)$

$f(x,y)$

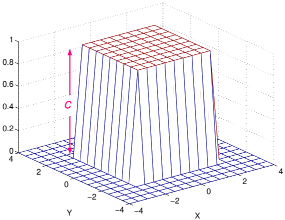


$\rightleftarrows$

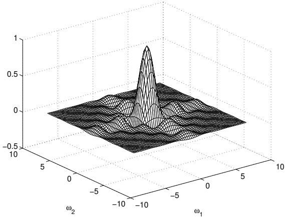


$F(\omega_1,\omega_2)$

$f(x,y)$



$\rightleftarrows$



$F(\omega_1,\omega_2)$

More Fourier Identities (proofs ≈ similar to 1D)

Given $$f(x, y) \rightleftarrows F(\omega_1, \omega_2)$$

Shift Theorem: $$f(x - d_x, y - d_y) \rightleftarrows F(\omega_1, \omega_2) e^{-j(\omega_1 d_x + \omega_2 d_y)}$$

Coordinate Xformation: $$f(\mathbf{A}\mathbf{x}) \rightleftarrows \frac{1}{||\mathbf{A}||} F(\left(\mathbf{A}^{-1}\right)^T \Omega)$$

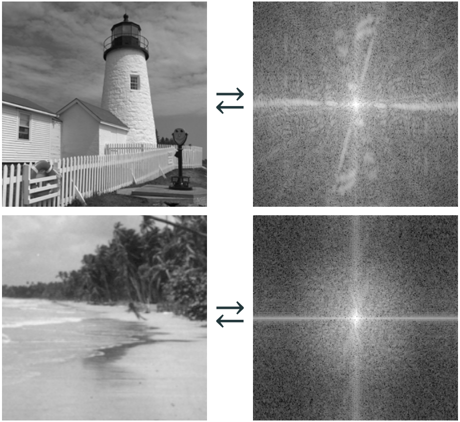
where $\mathbf{x} = [x \quad y]^T$ and $\mathbf{A}$ is a 2×2 matrix representing a general 2D transformation e.g.

$$\text{Zoom: } \mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \qquad (\mathbf{A}^{-1})^T = \begin{bmatrix} 1/a & 0 \\ 0 & 1/a \end{bmatrix}$$

$$\text{Rotation: } \mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad (\mathbf{A}^{-1})^T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Therefore, magnifying the image by a factor a (zooming), causes the spectrum to *shrink* in area by that factor, and rotating an image by θ causes the spectrum to rotate by the same angle.

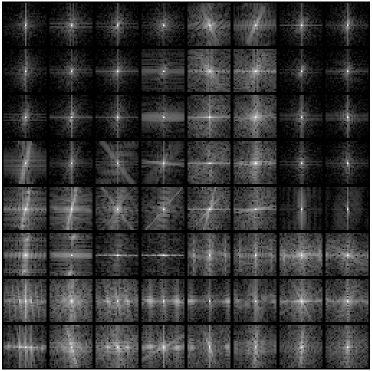
Fourier Xform of images (Log Power Spectra dB - 2D DFT)



Most natural images have an exponential drop off in spectral power with increasing frequency. Taking the Fourier transform of a signal causes a compaction of the energy into low-frequency regions.

Fourier Xform of images (Log Power Spectra dB - 2D DFT)

$\rightleftarrows$



The picture has been split into 64 32×32 blocks.

(remember data windowing is an issue: we used 2D Hamming here!)

Fourier Xform of images (Log Power Spectra dB - 2D DFT)

- Images are not statistically homogeneous over large areas
- Edges in the image correspond to highly directional spectra
- All the blocks have a strong DC component (i.e. average intensity is non-zero)
- This energy compaction is the key to the use of transforms for image compression