DD2423 Image Processing and Computer Vision

DISCRETE FOURIER TRANSFORM

Mårten Björkman

Computer Vision and Active Perception

School of Computer Science and Communication

November 1, 2012

Fourier transform

$$\mathcal{F}(\mathsf{f}(x)) = \int_{x \in \mathbb{R}^n} \mathsf{f}(x) e^{-i\omega^T x} dx = \hat{\mathsf{f}}(\omega)$$

$$\mathcal{F}^{-1}(\hat{\mathsf{f}}(\omega)) = \frac{1}{(2\pi)^n} \int_{\omega \in \mathbb{R}^n} \hat{\mathsf{f}}(\omega) e^{i\omega^T x} d\omega$$

Terminology:

Frequency spectrum: $\hat{\mathsf{f}}(\omega) = \mid \hat{\mathsf{f}}(\omega) \mid e^{i\phi(\omega)}$

Fourier spectrum: $\mid \hat{\mathsf{f}}(\omega) \mid$

Power spectrum: $\mid \hat{\mathsf{f}}(\omega) \mid^2$

Why are we interested in a decomposition of the signal into harmonic components?

Sinusoids and cosinusoids are eigenfunctions of convolution

$$e^{i\omega t} \rightarrow \boxed{\qquad\qquad} \rightarrow A(\omega)e^{i\omega t}$$

$$e^{i\omega t} = \cos \omega t + i \sin \omega t$$

Thus we can understand what the system (e.g filter) does to the different components (frequencies) of the signal (image)

Theorem 2

Convolution in the spatial domain is same as multiplication in the Fourier (frequency) domain

$$\mathcal{F}(\mathrm{h} * \mathrm{f}) = \mathcal{F}(\mathrm{h})\mathcal{F}(\mathrm{f})$$

$$\mathrm{f} \rightarrow \boxed{*h} \rightarrow \mathrm{g{=}h{*}f} \qquad \hat{\mathrm{f}} \rightarrow \boxed{\hat{\mathrm{h}}} \rightarrow \hat{\mathrm{g}} = \hat{\mathrm{h}}\,\hat{\mathrm{f}}$$

Usage:

- For analysis and understanding of convolution operators.
- Some filters may be easily represented in the Fourier domain.
- Implementation: when size of the filter is too large it is more effective to use multiplication in the Fourier domain.

Convolution

$$\mathcal{F}(h{*}f) = \mathcal{F}(h)\mathcal{F}(f)$$

Proof:

$$\mathcal{F}(h{*}f)(\omega) = \int_{x\in\mathbb{R}^n}(\int_{\eta\in\mathbb{R}^n} h(x-\eta)f(\eta)d\eta)e^{-i\omega^T x}dx \quad \{\text{rewrite}\}$$

$$= \int_{\eta\in\mathbb{R}^n}(\int_{x\in\mathbb{R}^n} h(x-\eta)e^{-i\omega^T(x-\eta)}dx)f(\eta)e^{-i\omega^T\eta}d\eta \quad \{\text{with} \quad (x-\eta)=\zeta\}$$

$$= \int_{\eta\in\mathbb{R}^n}(\int_{\zeta\in\mathbb{R}^n} h(\zeta)e^{-i\omega^T\zeta}d\zeta)f(\eta)e^{-i\omega^T\eta}d\eta \quad \{\text{separate}\}$$

$$= (\int_{z\eta\in\mathbb{R}^n} h(\zeta)e^{-i\omega^T\zeta}d\zeta)(\int_{\eta\in\mathbb{R}^n} f(\eta)e^{-i\omega^T\eta}d\eta = \mathcal{F}(h)(\omega)\mathcal{F}(f)(\omega)$$

Spatial separability

- Given

$$\mathsf{h}(x,y) = \mathsf{h}_1(x)\mathsf{h}_2(y)$$

$$(\mathsf{h}*\mathsf{f})(x,y) = \int_{\eta\in\mathbb{R}^n}\int_{\zeta\in\mathbb{R}^n} \mathsf{h}(\eta,\zeta)\mathsf{f}(x-\eta,y-\zeta)\,d\eta\, d\zeta$$

$$= \int_{\eta\in\mathbb{R}^n} \mathsf{h}_1(\eta)\underbrace{\left(\int_{\zeta\in\mathbb{R}^n} \mathsf{h}_2(\zeta)\mathsf{f}(x-\eta,y-\zeta)\,d\zeta\right)}_{***} d\eta$$

  *** convolution of a column (fixed value of x) in y-direction

- If convolution mask h can be separated as above $\Rightarrow$ 2D convolution can be performed as a series of 1D convolutions.
- In the discrete case: If the mask is m^2 in size $\Rightarrow$ $2m$ operations / pixel instead of m^2.

In the Fourier domain

$$\hat{\mathsf{f}} = \int_{\omega_1 \in \mathbb{R}^n} \int_{\omega_2 \in \mathbb{R}^n} \mathsf{f}(x,y) e^{-i(\omega_1 x_1 + \omega_2 x_x)} dx_1 dx_2$$
$$= \int_{\omega_1 \in \mathbb{R}^n} e^{-i\omega_1 x_1} (\int_{\omega_2 \in \mathbb{R}^n} \mathsf{f}(x,y) e^{-i\omega_2 x_2} dx_2) dx_1$$

- A Fourier transform in 2D can always be performed as a series of two 1D Fourier transforms.

Applications

Filering techniques typically modify frequency characteristics:

- Enhance edges (increase high frequency)
- Remove noise (decrease high frequency)
- Smooth (decrease high frequency, increase low frequency)

Observation: Convolution is a spatial operation on an image.

Images can be converted to their frequency component prior to filtering to facilitate direct manipulation of image frequency characteristics.

Spatial versus frequency domain

The Fourier Transform converts spatial image data into a frequency representation. Both representations contain equivalent information.

Spatial Domain

+ Intuitive Representation

- Filtering with large kernels may result in long processing times.

+ Kernels applied directly to spatial data.

Frequency Domain

- Non-intuitive representation

+ Filtering with large kernels can be performed very quickly

- Image and Kernel must first be converted to frequency domain, modified, then reconverted. Engineering filters often easier.

Discrete Fourier Transform in 2D

$$\hat{\mathsf{f}}(u,v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathsf{f}(m,n) e^{-2\pi i \left(\frac{mu}{M} + \frac{nv}{N}\right)} \tag{1}$$

$$\mathsf{f}(m,n) = \frac{1}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \hat{\mathsf{f}}(u,v) e^{+2\pi i \left(\frac{mu}{M} + \frac{nv}{N}\right)} \tag{2}$$

Terminology:

- Fourier spectrum: $|F(u,v)| = \sqrt{R^2(u,v) + I^2(u,v)}$
- Phase angle: $\phi(u,v) = \tan^{-1} \frac{I(u,v)}{R(u,v)}$
- Power spectrum: $P(u,v) = |F(u,v)|^2 = R^2(u,v) + I^2(u,v)$
- The magnitude is simply the peak value, and the phase determines where the origin is, or where the sinusoid starts.

Relation continuous/discrete Fourier transform

- Continuous

$$\hat{\mathsf{f}}(\omega) = \int_{x \in \mathbb{R}^n} \mathsf{f}(x) e^{-i\omega^T x} dx$$

- Discrete

$$\hat{\mathsf{f}}(u) = \frac{1}{\sqrt{M}^n} \sum_{x \in I^n} \mathsf{f}(x) e^{-\frac{2\pi i u^T x}{M}}$$

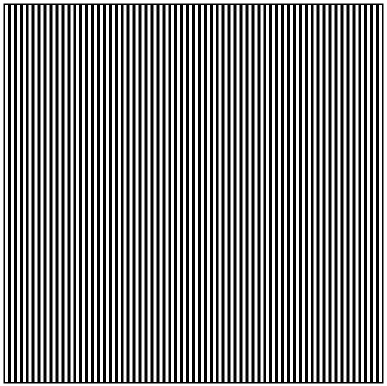
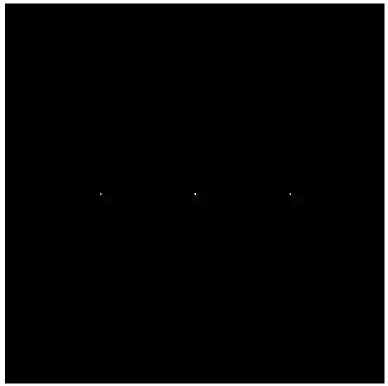
- Frequency variables are related (in 1D) by

$$\omega = \frac{2\pi u}{M}$$

- Note: u assumes values $0 ... M-1 \Rightarrow \omega \in [0, 2\pi)$.
- By periodic extension, we can map this integral to $[-\pi, \pi)$.

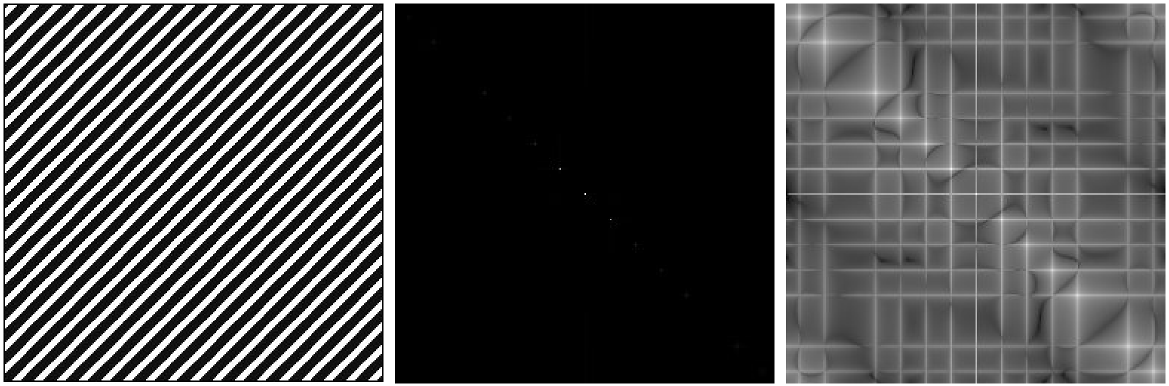
More insight (2 pixel wide stripes)

- The maximum frequency which can be represented in the spatial domain are one pixel wide stripes (period=2): $\omega_{max} = 2\pi\frac{1}{2} = \pi$
- So, 2 pixel wide stripes (period=4) give $\omega = 2\pi\frac{1}{4} = \frac{1}{2}\omega_{max}$
- Plotting magnitude of the Fourier transform: two points are halfway between the center and the edge of the image, i.e. the represented frequency is half of the maximum. One point in the middle shows the DC-value (image mean).

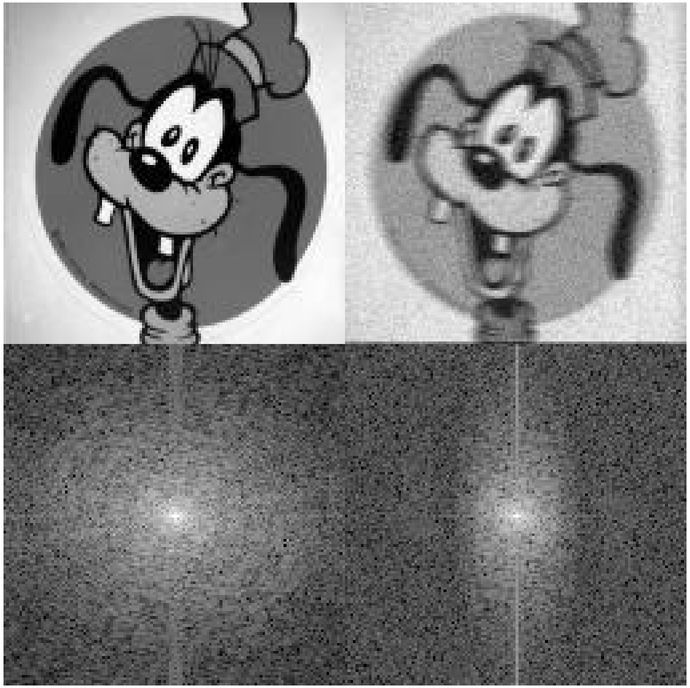
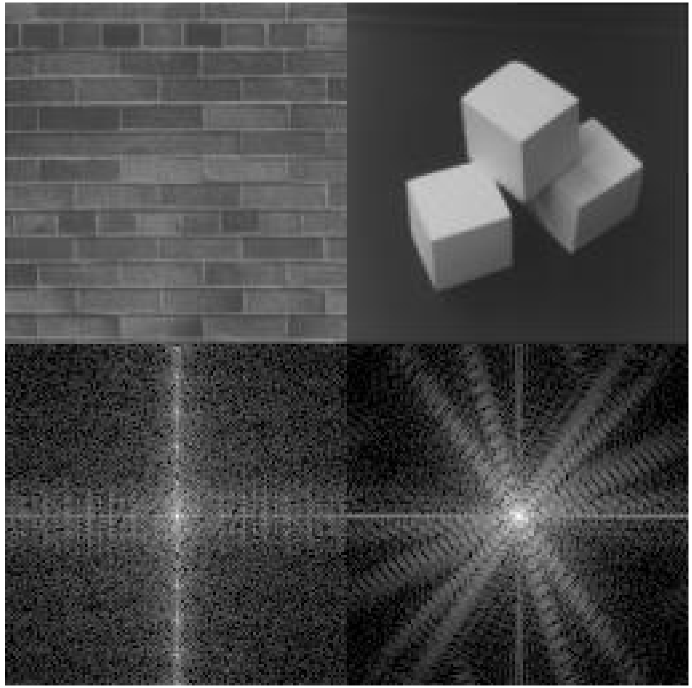



More insight

- If you take the logarithm of the Fourier transform, you see many minor frequencies. The reason is that since an image can only be represented by square pixels, the diagonals cannot be represented without discretization noise.



Example images and Fourier transforms

Exercise

Given a simple 4-pixel "image" $I(c) = [3, 2, 2, 1]$, what is its Fourier Transform $F(v)$?

Exercise

Given a simple 4-pixel "image" $I(c) = [3, 2, 2, 1]$, what is its Fourier Transform $F(v)$?

$$F(v) = \frac{1}{N}\sum_{c=0}^{N-1} I(c)e^{-j2\pi\frac{vc}{N}}$$

$$F(0) = \frac{1}{4}\sum_{c=0}^{3} I(c)e^{-j2\pi vc/4} = \frac{1}{4}\sum_{c=0}^{3} I(c)e^{0} = \frac{1}{4}[I(0)+I(1)+I(2)+I(3)]1 = \frac{1}{4}[3+2+2+2+1] = 2$$

$$F(1) = \frac{1}{4}\sum_{c=0}^{3} I(c)e^{-j2\pi(1)c/4} = \frac{1}{4}\left[3e^{0} + 2e^{-j\pi/2} + 2e^{-j\pi} + 1e^{-j\pi 3/2}\right] = \frac{1}{4}[3+2(-j)+2(-1)+1(j)] = \frac{1}{4}[1-j]$$

$$F(2) = \frac{1}{4}\sum_{c=0}^{3} I(c)e^{-j2\pi(2)c/4} = \frac{1}{4}\left[3e^{0} + 2e^{-j\pi} + 2e^{-j2\pi} + 1e^{-j3\pi}\right] = \frac{1}{4}[3+(-2)+2+(-1)] = \frac{1}{2}$$

$$F(3) = \frac{1}{4}\sum_{c=0}^{3} I(c)e^{-j2\pi(3)c/4} = \frac{1}{4}\left[3e^{0} + 2e^{-j\pi 3/2} + 2e^{-j3\pi} + 1e^{-j\pi 9/2}\right] = \frac{1}{4}[3+2j+2(-1)+1(-j)] = \frac{1}{4}[1+j]$$

Property I - Separability

$$\hat{\mathsf{f}}(u,v) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathsf{f}(m,n) e^{-2\pi i (\frac{mu}{M} + \frac{nv}{N})} \tag{3}$$

$$\hat{\mathsf{f}}(u,v) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} (\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \mathsf{f}(m,n) e^{-2\pi i \frac{nv}{n}}) e^{-2\pi i \frac{mu}{M}}$$

$$\mathsf{f}(m,n) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} (\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \hat{\mathsf{f}}(u,v) e^{2\pi i \frac{nv}{n}}) e^{2\pi i \frac{mu}{M}}$$

2D DFT can be implemented as a series of 1D DFTs along each column, followed by 1D DFTs along each row.

Property II - Linearity

$$\mathcal{F}\left[a\,\mathsf{f}_1(m,n)\ +\ b\,\mathsf{f}_2(m,n))\right] = a\ \hat{f}_1(u,v)\ +\ b\ \hat{f}_2(u,v)$$

$$a\,\mathsf{f}_1(m,n)\ +\ b\,\mathsf{f}_2(m,n)) = \mathcal{F}^{-1}\left[a\ \hat{f}_1(u,v)\ +\ b\ \hat{f}_2(u,v)\right]$$

You can add two functions (images) or rescale a function, either before or after computing the Fourier transform. It leads to the same result.

Property III - Modulation

$$\mathcal{F}\left[f(m,n)e^{2\pi i(\frac{mu_0}{M}+\frac{nv_0}{N})}\right] = \hat{f}(u-u_0, v-v_0)$$

If the original function is multiplied with the above exponential and transformed, it will result in a shift of the origin of the frequency plane to point (u_0, v_0).

For $(u_0, v_0) = (M/2, N/2)$

$$e^{2\pi i(\frac{mu_0}{M}+\frac{nv_0}{N})} = e^{\pi i(m+n)} = (-1)^{m+n}$$

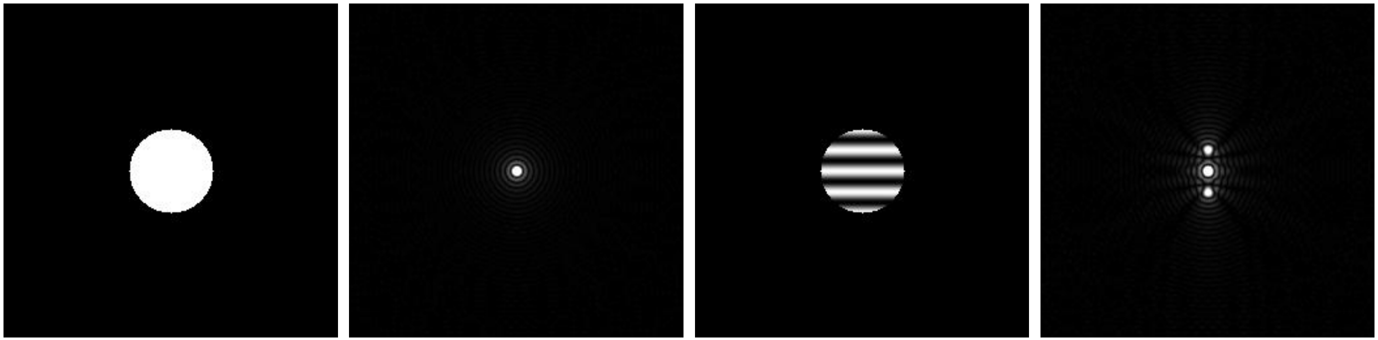
and

$$f(m,n)(-1)^{m+n} \Longleftrightarrow \hat{f}(u-M/2, v-N/2)$$

Conclusion: the origin of the Fourier transform can be moved to the center of the frequency square by multiplying the original function by $(-1)^{m+n}$.

Property III - Modulation/Frequency translation

From left: Original image, magnitude of the Fourier spectrum, original multiplied by $1+2\cos\omega y$ at a relative frequency of 16, magnitude of the Fourier spectrum.

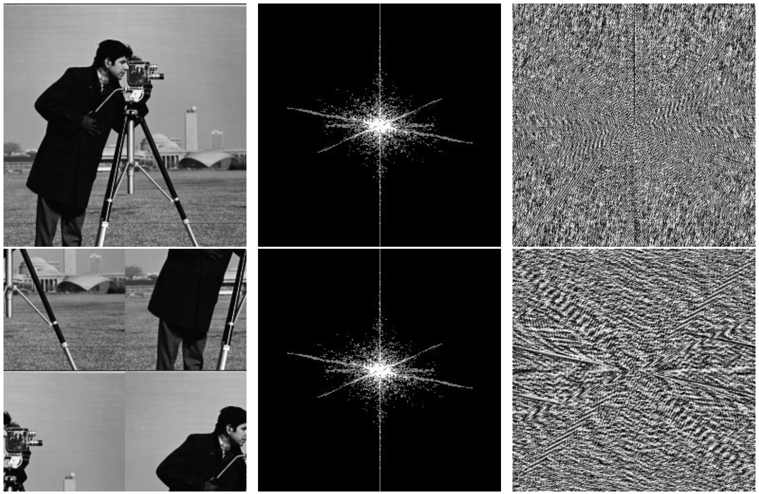


Note: $1+2\cos\omega y = 1+e^{i\omega y}+e^{-i\omega y}$.

Property IV - Translation

$$\mathcal{F}\left[\mathrm{f}(m-m_0,n-n_0)\right] = \hat{f}(u,v)e^{-2\pi i(\frac{m_0 u}{M}+\frac{n_0 v}{N})}$$

Conclusion: if the image is moved, the resulting Fourier spectrum undergoes a phase shift, but magnitude of the spectrum remains the same.

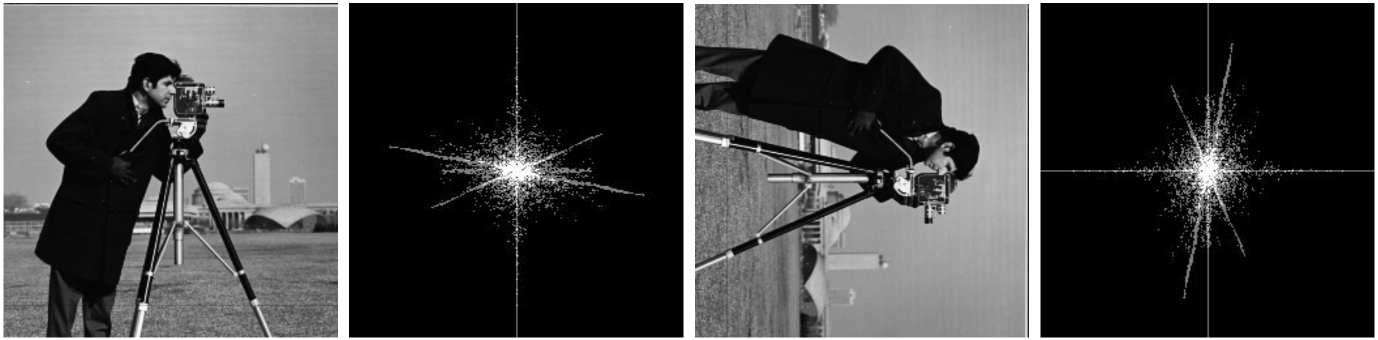


The shift in $f(x,y)$ does not affect the magnitude of its Fourier transform.

$$|\hat{f}(u,v)e^{-2\pi i(\frac{m_0 u}{M}+\frac{n_0 v}{N})}| = |\hat{f}(u,v)|$$

Property V - Rotation

Rotation of the original image rotates $\hat{f}$ by the same angle.



Exercise: Introduce polar coordinates and perform direct substitution.

Property VI - Scaling

$$A = \begin{pmatrix} S_1 & & \\ & \ddots & \\ & & S_n \end{pmatrix} \quad \text{(diagonal)}$$

$$\mathrm{g}(x) = \mathrm{f}\,(S_1 x_1, \ldots, S_n x_n)$$

$$\hat{\mathrm{g}}(\omega) = \frac{1}{|S_1 \ldots S_n|}\hat{\mathrm{f}}\,(\frac{\omega_1}{S_1}, \ldots, \frac{\omega_n}{S_n})$$

Conclusion: compression in spatial domain is same as expansion in Fourier domain (and vise versa).

Property VII - Periodicity

The DFT and its inverse are periodic with period N, for an N x N image.
This means:

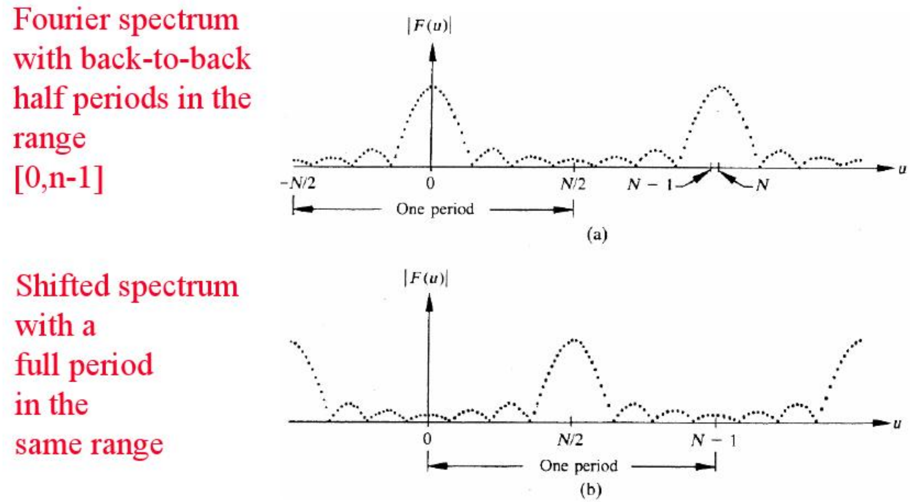
$$\hat{f}(u,v) = \hat{f}(u+N,v) = \hat{f}(u,v+N) = \hat{f}(u+N,v+N)$$

This property defines the implied symmetry in the Fourier spectrum as well as that $\hat{f}$ repeats itself infinitely. However, only one period is enough to reconstruct the original function f.

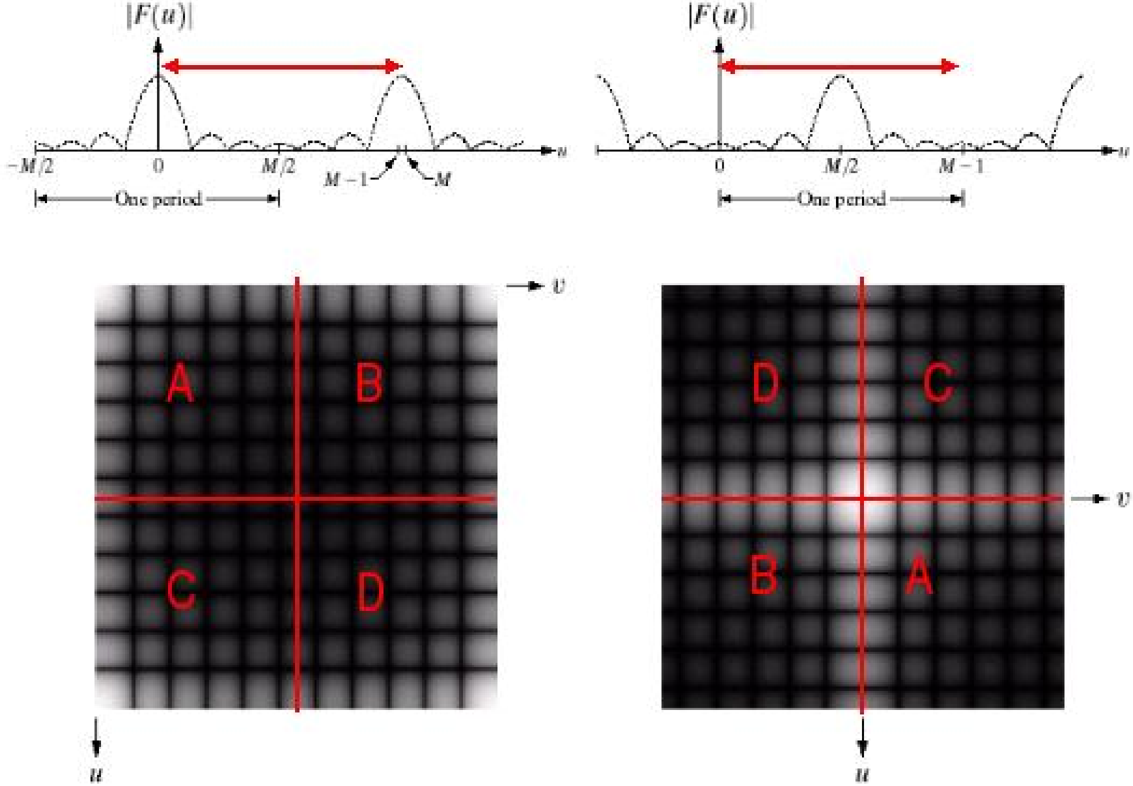
Property VIII - Conjugate Symmetry

The Fourier transform satisfies $\hat{f}(u,v) = \hat{f}^*(-u,-v)$ and $|\hat{f}(u,v)| = |\hat{f}(-u,-v)|$.

With periodicity and above, we have that $\hat{f}$ has period N and is (conjugate) symmetric around the origin.



Thus we don't need $2N^2$ (N^2 real and N^2 imaginary) values to represent a $N \times N$ image in Fourier domain, but N^2 if we exploit the symmetry.

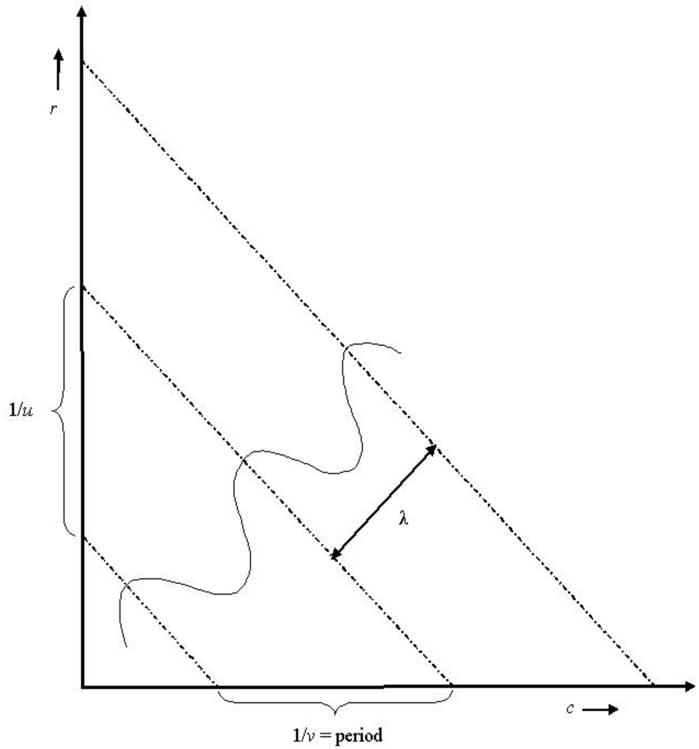


a b
c d

FIGURE 4.34
(a) Fourier spectrum showing back-to-back half periods in the interval $[0, M-1]$.
(b) Shifted spectrum showing a full period in the same interval.
(c) Fourier spectrum of an image, showing the same back-to-back properties as (a), but in two dimensions.
(d) Centered Fourier spectrum.

Wavelength in two dimensions

The wavelength of the sinusoid is: $\lambda = \frac{1}{\sqrt{u^2+v^2}}$, where (u,v) are the frequencies along (r,c) and the periods are $1/u$ and $1/v$.



Theorem 3 and some definitions

Multiplication in the spatial domain is same as convolution in the Fourier domain.

$$\mathcal{F}(\mathsf{hf}) = \mathcal{F}(\mathsf{h}) * \mathcal{F}(\mathsf{f})$$

Exercise: Prove it!

Transfer functions

- A linear, shift invariant system (such as a filter) is completely specified by it's response to an impulse, which is called the impulse response.
- The transfer function H is the Fourier transform of the impulse response.
- Using the convolution theorem to describe the effects of the system:

$$g(x,y) = h(x,y) * f(x,y)$$

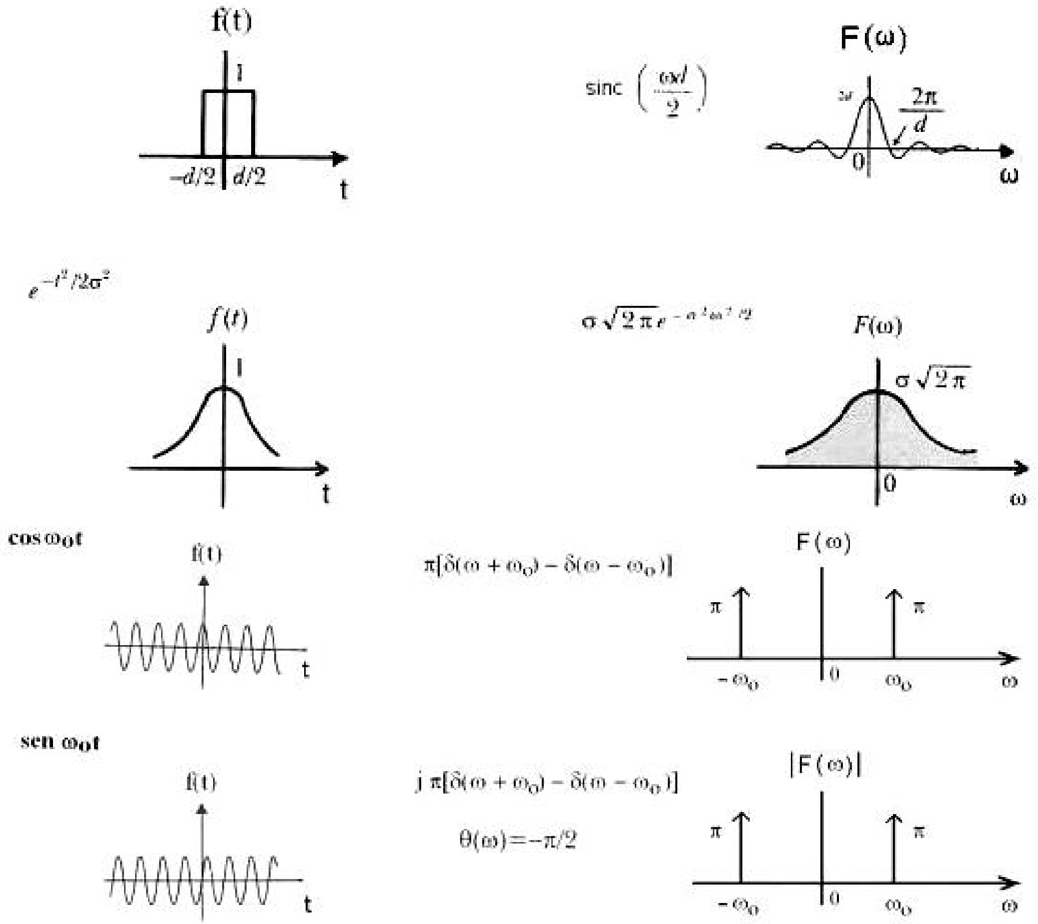
$$G(u,v) = H(u,v) \cdot F(u,v)$$

- Convolution with an impulse function

$$h(x,y) * \delta(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x_0,y_0)\delta(x-x_0,y-y_0)dx_0dy_0 = h(x,y)$$

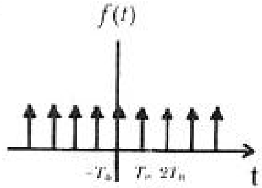
results in a “copy” of $h(x,y)$ to the location of the impulse

Transfer function examples

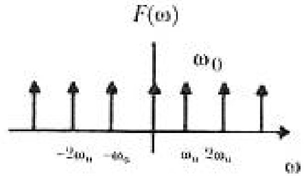


Transfer function examples

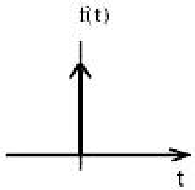
$\delta_{T_0}(t)$



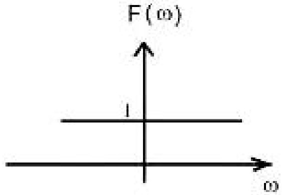
$\omega_0 \delta_{\omega_0}(\omega)$



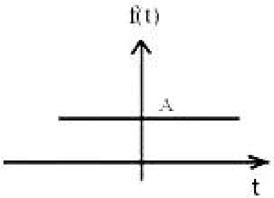
δ(t)



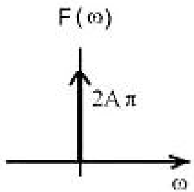
1



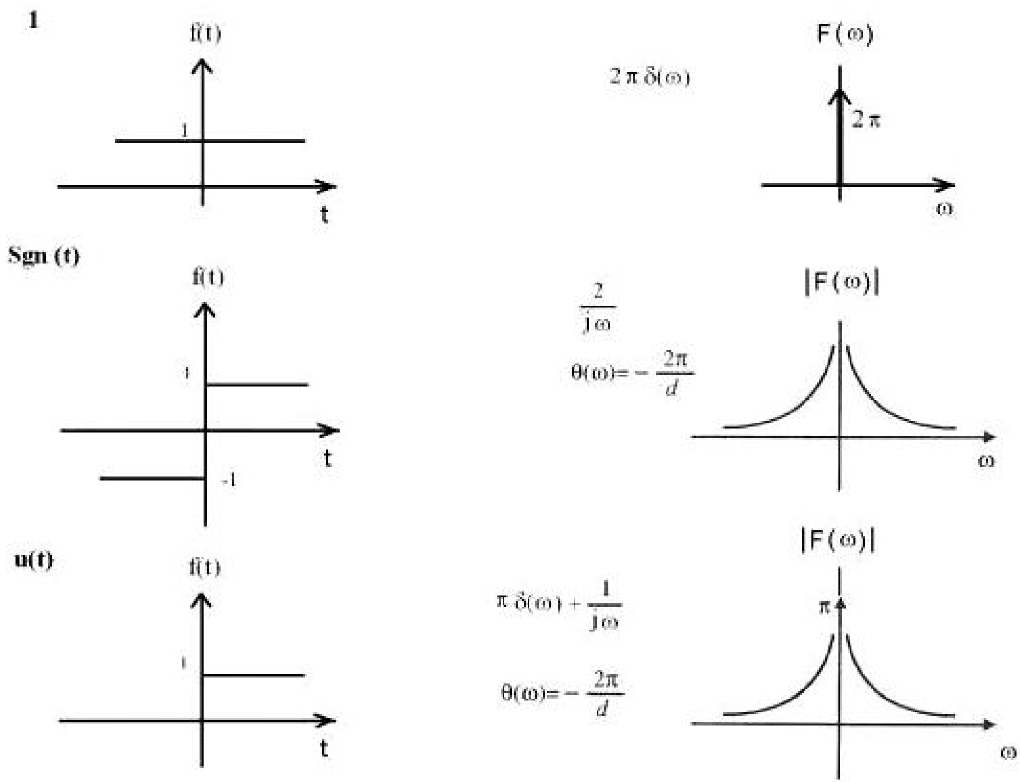
A



2Aπδ(ω)

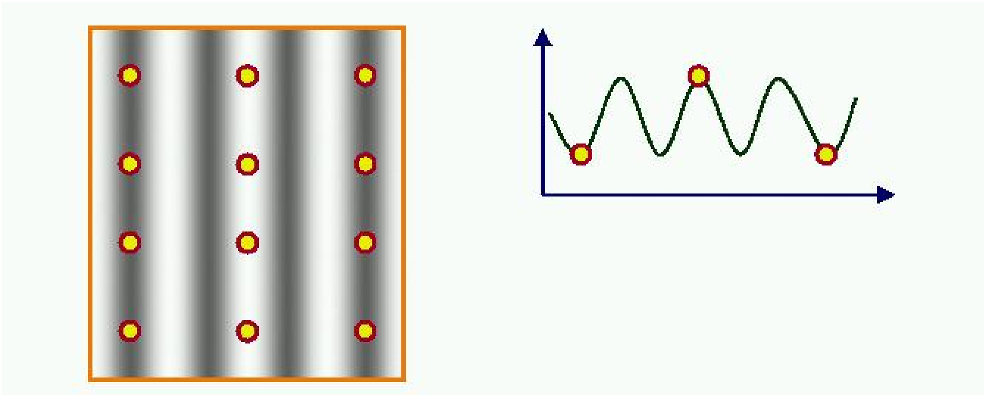


Transfer function examples



Back to sampling

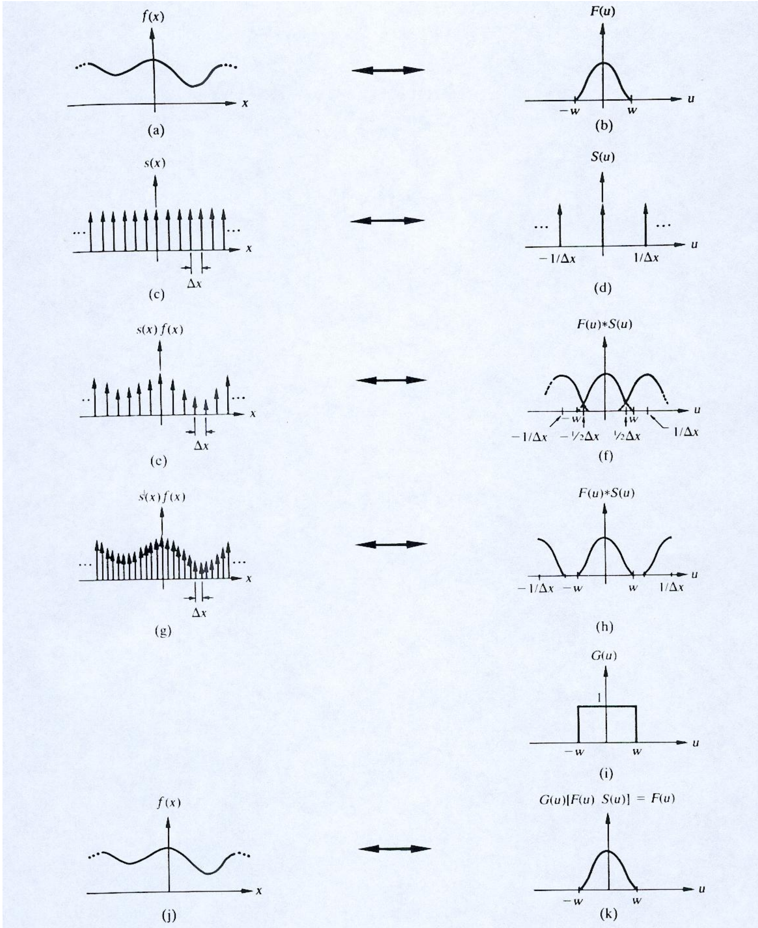
- How many samples are required to describe the given signal without loss of information?
- What signal can be reconstructed given the current sampling rate?



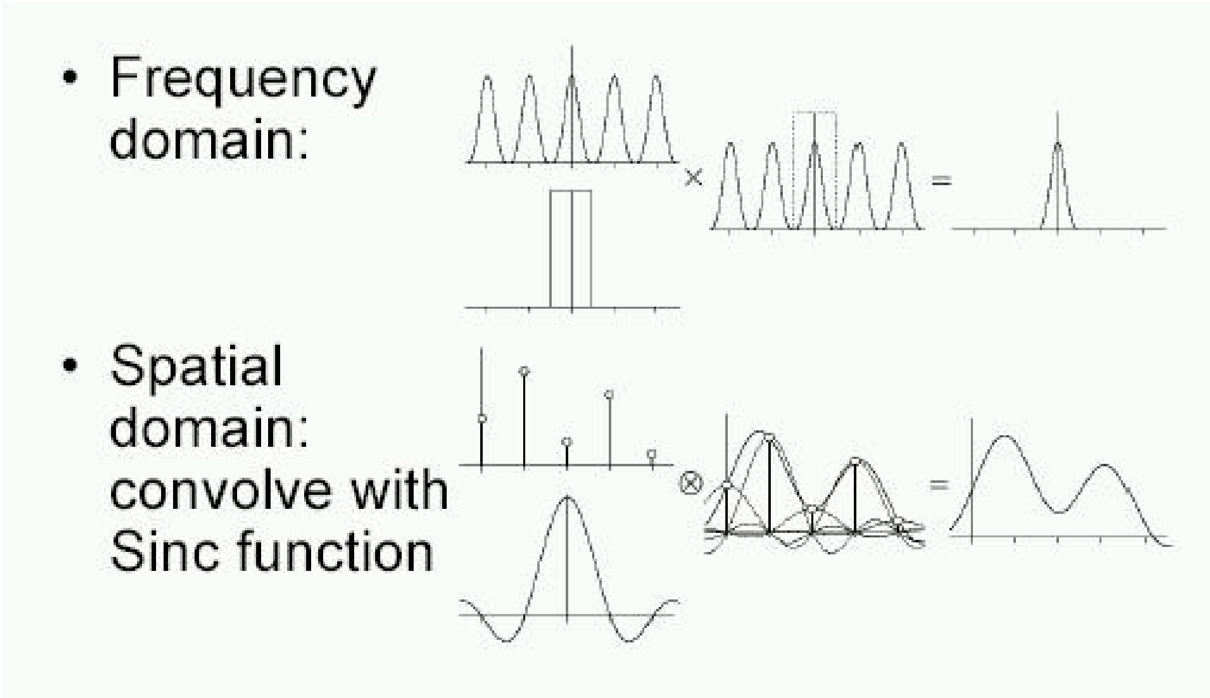
Sampling Theorem

- A signal is band limited if its highest frequency is bounded. This frequency is called the bandwidth.
- The sine/cosine component of the highest frequency determines the highest "frequency content" of the signal.
- If the signal is sampled at a rate equal or greater to than twice its highest frequency, the original signal can be completely recovered from it samples (Shannon).
- The minimum sampling rate for band limited function is called Nyquist rate.

Reconstruction



Reconstruction



Aliasing

- If the signal is undersampled, aliasing occurs.
- Prevent aliasing by:
  - increasing sampling rate, or
  - decreasing highest frequency before sampling.

Reconstruction Kernel

- For perfect reconstruction, we need to convolve with a sinc function.
  - It is the Fourier transform of the box function.
  - It has infinite support.
- May be approximated by a Gaussian, cubic or even “tent” function.

Parseval's equality

The total energy contained in an image summed across all (x, y) is equal to the total energy of its Fourier Transform summed across all frequencies.

$$\int_{x\in\mathbb{R}^n} \mathsf{f}(x)\mathsf{g}(x)dx = \frac{1}{(2\pi)^n}\int_{\omega\in\mathbb{R}^n} \hat{\mathsf{f}}(\omega)\hat{\mathsf{g}}^*(\omega)d\omega$$

Proof:

$$\int_{\omega\in\mathbb{R}^n} \hat{\mathsf{f}}(\omega)\hat{\mathsf{g}}^*(\omega)d\omega = \int_{\omega\in\mathbb{R}^n} \hat{\mathsf{f}}(\omega)\left(\int_{x\in\mathbb{R}^n} \mathsf{g}(x)e^{i\omega^T x}dx\right)d\omega$$

$$= \int_{x\in\mathbb{R}^n} \mathsf{g}(x)\underbrace{\left(\int_{\omega\in\mathbb{R}^n} \hat{\mathsf{f}}(\omega)e^{i\omega^T x}d\omega\right)}_{(2\pi)^n \mathsf{f}(x)} dx = (2\pi)^n \int_{x\in\mathbb{R}^n} \mathsf{f}(x)\mathsf{g}(x)dx$$

If $\mathsf{f} = \mathsf{g}$ then

$$\int_{x\in\mathbb{R}^n} \mathsf{f}(x)^2 dx = \frac{1}{(2\pi)^n}\int_{\omega\in\mathbb{R}^n} |\,\hat{\mathsf{f}}\,|^2\, d\omega.$$

Summary of good questions

- What is the Fourier transform of a convolution? Why is it important?
- What does separability of filters mean?
- How do you interpret a point in the Fourier domain in the spatial domain?
- How do you apply a discrete Fourier transform?
- What happens to the Fourier transform, if you translate an image?
- What happens to the Fourier transform, if you rotate an image?
- In what sense is the Fourier transform symmetric?
- What does the Sampling Theorem mean in practice?
- What can you do to get rid of aliasing in the sampling process?

Readings

- Gonzalez and Woods: Chapter 4
- Introduction to Lab 2